BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

DESIGNING ENERGY-EFFICIENT SUB-THRESHOLD LOGIC CIRCUITS USING EQUALIZATION AND NON-VOLATILE MEMORY CIRCUITS USING MEMRISTORS

by

MAHMOUD ZANGENEH

B.S., Amirkabir University of Technology (Tehran Polytechnic), 2007
M.S., University of Tehran, 2010

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2015

© 2015 by
Mahmoud Zangeneh
All rights reserved

Approved by

First Reader

Ajay Joshi, PhD
Assistant Professor of Electrical and Computer Engineering

Second Reader

Allyn Hubbard, PhD
Professor of Biomedical Engineering
Professor of Electrical and Computer Engineering

Third Reader

Bennett Goldberg, PhD
Professor of Physics
Professor of Biomedical Engineering

Fourth Reader

M. Selim Ünlü, PhD
Professor of Biomedical Engineering
Professor of Electrical and Computer Engineering

Acknowledgments

I would like to express my deepest appreciation to the committee members for reviewing this dissertation and providing me precious feedback. Then I want to thank my adviser, Prof. Ajay Joshi, who led my research and worked with me for many interesting research topics. Also it is my pleasure to work with Zafar Takhirov in our group, who provided many experimental results and valuable suggestions. I want to thank members in our group Schuyler Eldridge, Boyou Zhou, and Chao Chen for making our research in the lab enjoyable.

Also I wish to thank Prof. Selim Ünlü and Prof. Bennett Goldberg for their suggestions in my research. They helped me overcome many technical obstacles when we were writing papers together. I also want to thank several members in their groups, including Dr. Ronen Adato, Dr. Berkin Cilingiroglu, Aydan Uyar, and Dr. Abdulkadir Yurt for all their contributions when we were working on research projects together.

In addition, a thank you to thank Prof. Ronald Knepper, Dr. David Freedman, Prof. Michelle Sander and Prof. Min-Chang Lee. I worked as a teaching fellow for their classes at Boston University. I enjoyed when helping undergraduate students with their homework and lab projects, and I also obtained a lot of useful knowledge for my later research.

Finally, I want to thank my father, Majid Zangeneh, my mother, Zarin Aghai Rashti and my brother, Masoud Zangeneh for their continuous support and encouragement. They helped me focus on the research and work hard for achieving my education and career goals. I hope to provide my parents better lives after graduation and spend more time with them.

DESIGNING ENERGY-EFFICIENT SUB-THRESHOLD LOGIC CIRCUITS USING EQUALIZATION AND NON-VOLATILE MEMORY CIRCUITS USING MEMRISTORS

MAHMOUD ZANGENEH

Boston University, College of Engineering, 2015

Major Professor: Ajay Joshi, PhD, Assistant Professor of Electrical and Computer Engineering

ABSTRACT

The very large scale integration (VLSI) community has utilized aggressive complementary metal-oxide semiconductor (CMOS) technology scaling to meet the ever-increasing performance requirements of computing systems. However, as we enter the nanoscale regime, the prevalent process variation effects degrade the CMOS device reliability. Hence, it is increasingly essential to explore emerging technologies which are compatible with the conventional CMOS process for designing highly-dense memory/logic circuits. Memristor technology is being explored as a potential candidate in designing non-volatile memory arrays and logic circuits with high density, low latency and small energy consumption. In this thesis, we present the detailed functionality of multi-bit 1-Transistor 1-memRistor (1T1R) cell-based memory arrays. We present the performance and energy models for an individual 1T1R memory cell and the memory array as a whole. We have considered TiO_2- and HfO_x-based memristors, and for these technologies there is a sub-10% difference between energy and performance

computed using our models and HSPICE simulations. Using a performance-driven design approach, the energy-optimized TiO_2-based RRAM array consumes the least write energy (4.06 pJ/bit) and read energy (188 fJ/bit) when storing 3 bits/cell for 100 nsec write and 1 nsec read access times. Similarly, HfO_x-based RRAM array consumes the least write energy (365 fJ/bit) and read energy (173 fJ/bit) when storing 3 bits/cell for 1 nsec write and 200 nsec read access times.

On the logic side, we investigate the use of equalization techniques to improve the energy efficiency of digital sequential logic circuits in sub-threshold regime. We first propose the use of a variable threshold feedback equalizer circuit with combinational logic blocks to mitigate the timing errors in digital logic designed in sub-threshold regime. This mitigation of timing errors can be leveraged to reduce the dominant leakage energy by scaling supply voltage or decreasing the propagation delay. At the fixed supply voltage, we can decrease the propagation delay of the critical path in a combinational logic block using equalizer circuits and, correspondingly decrease the leakage energy consumption. For a 8-bit carry lookahead adder designed in UMC 130 nm process, the operating frequency can be increased by 22.87% (on average), while reducing the leakage energy by 22.6% (on average) in the sub-threshold regime. Overall, the feedback equalization technique provides up to 35.4% lower energy-delay product compared to the conventional non-equalized logic. We also propose a tunable adaptive feedback equalizer circuit that can be used with sequential digital logic to mitigate the process variation effects and reduce the dominant leakage energy component in sub-threshold digital logic circuits. For a 64-bit adder designed in 130 nm our proposed approach can reduce the normalized delay variation of the critical path delay from 16.1% to 11.4% while reducing the energy-delay product by 25.83% at minimum energy supply voltage. In addition, we present detailed energy-performance models of the adaptive feedback equalizer circuit. This work serves as

a foundation for the design of robust, energy-efficient digital logic circuits in sub-threshold regime.

Contents

1 Introduction **1**

1.1 Background and Motivation 1

1.2 Contribution and Organization 6

2 Memristor Technology and Modeling **10**

2.1 Introduction 10

2.2 Memristor Device Technology 10

2.3 Summary 18

3 Design of Multi-bit RRAM Array **19**

3.1 Introduction 19

3.2 Related Work 19

3.3 RRAM Cell Design 21

3.4 RRAM Array Architecture 23

3.5 Performance Models 25

3.6 Energy Models 36

3.7 Memory Technology Comparison 43

3.8 PVT Variation Analysis of n-bit RRAM Cell 45

3.9 Summary 53

4 Sub-threshold Logic Design using Feedback Equalization **55**

4.1 Introduction 55

4.2 Related Work 56

4.3 Equalized Flip flop versus Conventional Flip flop 58
4.4 Experimental Results 63
4.4.1 Performance improvement at the fixed supply voltage 64
4.4.2 Leakage reduction at the fixed operating frequency 67
4.4.3 Mitigating process variations 68
4.5 Effect of Technology Scaling 69
4.6 Summary 71

5 Tunable Sub-threshold Logic Circuits using Adaptive Feedback Equalization 72

5.1 Introduction 72
5.2 Adaptive Equalized Flip flop versus Conventional Flip flop 72
5.3 Modeling of Feedback Equalizer Circuits 81
5.4 Evaluation 90
5.4.1 Improvement of Energy Efficiency 91
5.4.2 Maintaining Robustness Using Post-Fabrication Tuning 96
5.4.3 Mitigating Voltage/Temperature Variations 99
5.4.4 Effect of Technology Scaling 101
5.4.5 Comparison with other Sub-threshold Design Techniques . . . 103
5.4.6 Memristor-based Feedback Equalization Technique 103
5.5 Summary 105

6 Conclusion and Future Work 106

6.1 Conclusion 106
6.2 Future Work 109
6.2.1 Equalized Flip Flop with Bypass 109
6.2.2 Equalization Techniques for Near-threshold Voltage Computing Applications 110

6.2.3 Exploring the Architectural Impact of Using Non-volatile Memristor-based On-chip Cache 111

References **114**

Curriculum Vitae **122**

List of Tables

1.1 Comparison between current emerging nonvolatile memory technologies. 4

2.1 Parameters of TiO_2-based (Strukov et al., 2008) and HfO_x-based (Ielmini, 2011), (Sheu et al., 2009) memristors used for modeling and simulations. 13

3.1 Comparison between the reference voltages determined using analytical model (AM) and HSPICE simulation (HS) for a read$_{sub}$ access time of $T_R(TiO_2) = 1, 2$ $nsec$ in the 2-bit/cell 1T1R RRAM. $V_{LL}(TiO_2) = 0.48$ V is chosen to reach to at least 25 mV difference between the two adjacent reference voltages. The average error is 5.7% for TiO_2. . . . 27

3.2 Comparison between the reference voltages determined using analytical model (AM) and HSPICE simulation (HS) for a read$_{sub}$ access time of $T_R(HfO_x) = 200, 400$ $nsec$ in the 2-bit/cell 1T1R RRAM. $V_{LL}(HfO_x) = 0.7$ V is chosen to reach to at least 25 mV difference between the two adjacent reference voltages. The average error is 0.151% for HfO_x. 27

3.3 Comparison between analytical model (AM) and HSPICE simulations (HS) for energy dissipated in the cell while reading 2-bit RRAM cell with a read access time of $T_R(TiO_2) = 1nsec$ and $T_R(HfO_x) = 200nsec$. The average error is 8.44% and 0.038% for TiO_2 and HfO_x respectively. 37

3.4 Transition times of different components in the multi-bit RRAM array. 40

3.5 $3\sigma/\mu$ of the 3-bit TiO_2-based and HfO_x-based 1T1R cell specification variations due to LER, OTF and RDD. Here WT = Write time, WE = Write energy, RE = Read energy and RD = Read destructiveness. 49
3.6 $3\sigma/\mu$ of the 3-bit TiO_2-based and HfO_x-based 1T1R cell specifications due to voltage variations for ($3\sigma_{V_{ref}} = 6\%$) and ($3\sigma_{V_{ref}} = 10\%$). . . . 50
3.7 $3\sigma/\mu$ of the 3-bit TiO_2-based and HfO_x-based 1T1R cell specifications due to temperature variations (ΔT). 50

4.1 Comparison between the characteristics of the equalized flip flop (E-flip flop) with the conventional non-equalized master-slave flip flop (NE-flip flop) at different supply voltages operating in sub-threshold regime. Feedback equalization technique reduces the propagation delay of the 8-bit carry-lookahead adder CMOS logic whereas the setup time and t_{c-q} delay of the conventional flip flop is smaller than the equalized flip flip. 60
4.2 Comparison between the minimum energy point and the corresponding operating frequency of the equalized logic (E-logic) vs. non-equalized (NE-logic) design of various logic blocks. 66
4.3 Energy savings in scaled-down equalized logic compared to baseline non-equalized and equalized logic at the minimum energy supply voltage at zero word error rate operation for 8-bit carry lookahead adder. 69

5.1 Comparison between the timing characteristics of the original non-equalized design, the equalized design with 1 feedback path ON and the buffer-inserted non-equalized design. 78
5.2 Comparison between the minimum energy point and the corresponding operating frequency of the NE-logic vs. E-logic design of various logic blocks. 95

5.3 Energy savings in scaled-down E-logic compared to baseline NE-logic and E-logic at the minimum energy supply voltage with zero word error rate operation. 95

5.4 Comparison between the total delay, total energy and delay variation of the digital logic (64-bit adder) at minimum energy supply voltage when the conventional upsizing method (Kwong et al., 2009) has been used together with adaptive feedback equalizer circuit in sub-threshold regime. 98

5.5 Comparison between the normalized delay variation and energy-delay product (EDP) of the equalized logic (E-logic) vs. non-equalized (NE-logic) and buffer-inserted non-equalized design of various logic blocks assuming $\sigma_{V_T} = 10$ mV. 99

List of Figures

1·1 Saturation of clock frequency with CMOS scaling process (ISSCC Reports 2013). 2
1·2 Trend of total power consumption with CMOS scaling process (ISSCC Reports 2013). 3
1·3 Storage capacity of nonvolatile memory technologies (ISSCC Reports 2013). 5

2·1 Physical structure of (a) TiO_2-based memristor between 2 Pt contacts consisting of a highly conductive doped region and a highly resistive undoped region, where L = thickness of the memristor and W = thickness of the conductive region, and (b) HfO_x-based memristor showing conductive filament growth/narrowing process where ϕ_{min} and ϕ_{max} are the minimum and maximum filament diameters, respectively. . . . 12
2·2 Equivalent resistance of memristor devices. 13
2·3 Dynamic window function of the memristor state showing the nonlinear behavior of the memristor for different control parameter p. The current sign function prevents the state from getting stuck at the two boundaries. 15
2·4 Rate of diameter change for HfO_x-based memristors in filament growth model (Ielmini, 2011) for set (V>0) and reset (V<0) operations as a function of voltage across the memristor. 16

3·1 1-transistor 1-memristor (1T1R) RRAM cell. 22

3·2 n-bit/cell RRAM array architecture. 24
3·3 Equivalent circuit of 1T1R cell for read$_{sub}$ (left) and write$_{sub}$/refresh$_{sub}$ (right) operation. 25
3·4 Bitline voltage of a 2-bit/cell TiO_2-based RRAM for different bitline voltage development times. 28
3·5 Bitline voltage of a 2-bit/cell HfO_x-based RRAM for different bitline voltage development times. 28
3·6 Read time of a multi-bit RRAM cell for different number of bits per cell. 29
3·7 Contour plots of the number of consecutive non-destructive read cycles in multi-bit TiO_2-based RRAM for different n and V_{LL} values ($x = 0.9$). 31
3·8 Contour plots of the number of consecutive non-destructive read cycles in multi-bit HfO_x-based RRAM for different n and V_{LL} values ($x = 0.9$). 32
3·9 Comparison between analytical model (AM) and HSPICE simulations (HS) for bitline voltage and energy dissipation in different TiO_2-based and HfO_x-based 2-bit RRAM *write/refresh* operation. The V_{LL} voltage is 1.5 V for all transitions. For bitline voltage, the average error is 9.81% for TiO_2-based cell and 5.19% for HfO_x-based cell, while for energy dissipation the average error is 8.71% for TiO_2-based cell and 5.25% for HfO_x-based cell. 33
3·10 Contour plots for set time (nsec) in the 2 bits/cell TiO_2-based RRAM. 34
3·11 Contour plots for set time (nsec) in the 2 bits/cell HfO_x-based RRAM. 35
3·12 Contour plots for average read energy (pJ) in multi-bit TiO_2 RRAMs. We maintain at least 25 mV difference between adjacent reference voltages for reliable read operation. 38

3·13 Contour plots for average read energy (pJ) in multi-bit HfO_x RRAMs. We maintain at least 25 mV difference between adjacent reference voltages for reliable read operation. 38
3·14 Energy dissipated in different components of the multi-bit TiO_2-based RRAM array in read operation for uniform (left) and non-uniform (right) state assignments (T_R=1nsec). 41
3·15 Energy dissipated in different components of the multi-bit HfO_x-based RRAM array in read operation considering uniform (left) and non-uniform (right) state distributions (T_R=200nsec). 41
3·16 Energy dissipated in different components of the TiO_2-based RRAM array in write operation (T_W=100nsec). 43
3·17 Energy dissipated in different components of the HfO_x-based RRAM array in write operation (T_W=1nsec). 43
3·18 Comparison of read time/energy between different memory technologies. 44
3·19 Comparison of write time/energy between different memory technologies. 45
3·20 Uniform state distribution of the multi-bit TiO_2-based memristor caused by OTF. The memristor state distribution for each number of bits/cell is such that maximum process noise margin would be achieved. 45
3·21 Non-uniform state distribution of the multi-bit TiO_2-based memristor caused by OTF. The memristor state distribution for each number of bits/cell is such that maximum read noise margin would be achieved. 46
3·22 Uniform state distribution of the multi-bit HfO_x-based memristor caused by LER. The memristor state distribution for each number of bits/cell is such that maximum process noise margin would be achieved. 48

3·23 Non-uniform state distribution of the multi-bit HfO_x-based memristor caused by LER. The memristor state distribution for each number of bits/cell is such that maximum read noise margin would be achieved. 48
3·24 Diameter change of HfO_x-based memristors as a function of temperature for different applied voltages in filament growth model. Diameter shows higher variation with temperature at lower loadline voltages. . 51
3·25 Effective thermal resistance of a 3-bit TiO_2-based RRAM as a function of memristor state. 52

4·1 Feedback equalizer (designed using a variable threshold inverter (Sridhara et al., 2008)) can be combined with a traditional master-slave flip flop to design an equalized flip flop. 59
4·2 DC response of the variable threshold circuit in sub-threshold regime. The switching threshold of the inverter is modified based on the previous sampled output data. 61
4·3 Comparison between the timing waveforms of the input node of the conventional flip flop (A), output node of the conventional flip flop (B), input node of the equalized flip flop (C), output node of the equalized flip flop (D), output node of the variable threshold inverter (E). Feedback circuit makes sharper transitions in the waveforms of the logic output node helping the equalized flip flop sample the correct data. . 62
4·4 Operating frequency of the 8-bit carry lookahead adder for zero word error rate as function of different sub-threshold supply voltages. The equalized logic (E-logic) can run 22.87% (on average) faster than the non-equalized logic (NE-logic). 63

4·5 Comparison between the total consumed energy as well as the dynamic/leakage components of the 8-bit carry lookahead adder for different supply voltages. At the minimum energy supply voltage, the equalized logic is burning 18.4% less total energy compared to the non-equalized version. 64

4·6 Comparison between the energy consumed by the equalized (E-logic) vs. non-equalized (NE-logic) 8-bit carry lookahead adder for different supply voltages with fixed performance ($f = 1.28\ MHz$) at zero word error rate. The non-equalized logic design consumes minimum energy at 300 mV. The equalized flip flop enables 30 mV supply voltage scaling leading to 16.72% lower total consumed energy. The equalized flip flop cannot operate at $V_{DD} < 270mV$ due to the occurrence of timing errors. 67

4·7 Energy-delay product of the scaled-down equalized 8-bit carry lookahead adder for zero word error rate operation. We can achieve reliable operation even when the transistors in the equalized logic design are scaled down to as small as $75\% \times W_{baseline}$. 68

4·8 Energy-delay product of a 8-bit carry lookahead adder designed using equalized logic (E-logic) vs. non-equalized logic (NE-logic) at zero word error rate at different technology nodes. The equalized logic approach reduces the energy-delay product of the sub-threshold logic by up to 26.46% across all technology nodes in the minimum energy supply voltage. 70

5·1 Adaptive feedback equalizer circuit with multiple feedback paths (designed using a variable threshold inverter (Sridhara et al., 2008)) can be combined with a traditional master-slave flip flop to design an adaptive equalized flip flop. 73
5·2 Circuit diagram of classic master-slave positive edge-triggered flip flop (Rabaey et al., 2003). 74
5·3 DC response of the adaptive feedback equalizer circuit with 2 different feedback paths in sub-threshold regime. The switching threshold of the inverter is modified based on the previous sampled output data. . 75
5·4 Block diagram of the original non-equalized design (a), equalized design with 1 feedback path ON (b) and buffer-inserted non-equalized design (c). 77
5·5 Comparison between the timing waveforms of the non-equalized logic design and the equalized logic design of a 64-bit adder. Here, the waveforms include the clock signal (A), input node of the conventional flip flop (B), output node of the conventional flip flop (C), input node of the equalized flip flop (D), output node of the equalized flip flop (E). Feedback circuit enables sharper transitions in the waveforms of the combinational logic output node helping the equalized flip flop sample the correct data. Here the feedback path 2 is OFF. 79
5·6 Maximum feedback strength in adaptive equalized flip flop. The switching threshold of the adaptive equalized flip flop should be larger than the maximum amplitude of the glitch. 81

5·7 Comparison between analytical model (AM) and HSPICE simulations (HS) for equivalent channel resistance of MOSFET devices operating in sub-threshold regime. The average error between the derived model and HSPICE simulation results is 6.96% in the entire sub-threshold regime. 82
5·8 Contour plots for the Δt_{s-t} (ns) of the adaptive equalized flip flop. Control path strength and feedforward path strength values are normalized to minimum-sized transistor sizes. 83
5·9 Contour plots for the Δt_{c-q} (ns) of the adaptive equalized flip flop. Control path strength and feedforward path strength values are normalized to minimum-sized transistor sizes. 84
5·10 Contour plots for the t_{PD-equ} (ns) of the critical path in the equalized logic (64-bit adder). Control path strength and feedforward path strength values are normalized to minimum-sized transistor sizes. . . 85
5·11 Comparison between analytical model (AM) contour plots for the total delay (ns) of the critical path in an equalized 64-bit adder with HSPICE simulations (HS). 88
5·12 Comparison between analytical model (AM) contour plots for the total energy (fJ/operation) of the equalized 64-bit adder with HSPICE simulations (HS). 89
5·13 Operating frequency of the 64-bit adder for zero word error rate as function of different sub-threshold supply voltages. The equalized logic (E-logic) can run 18.91% (on average) faster than the non-equalized logic (NE-logic). 91

5·14 Comparison between the total consumed energy as well as the dynamic/leakage components of the 64-bit adder for different supply voltages. Operating at the respective minimum energy supply voltage, the equalized logic is burning 10.85% less total energy compared to the non-equalized logic. 92
5·15 Block diagram of the 32-bit Array Multiplier. 93
5·16 Block diagram of the 3-tap 16-bit finite impulse response (FIR) filter. 94
5·17 Energy-delay product of the scaled-down equalized 64-bit adder for zero word error rate operation. We can achieve reliable operation even when the transistors in the equalized logic design are scaled down to as small as 75%$\times W_{baseline}$. 96
5·18 Delay distribution of the critical path in the 64-bit adder designed in UMC 130 nm process. The 3 $\times\sigma/\mu$ of the non-equalized logic (NE-logic), the equalized logic (E-logic) with 2 different feedback strengths and the buffer-inserted NE-logic are 16.1%, 11.4%, 7.14% and 15% for σ_{V_T} = 10 mV at the minimum energy supply voltage, respectively. Here, E-logic designs are operating at 300 mV. 97
5·19 Delay distribution of the critical path in the 64-bit adder designed in UMC 130 nm process considering supply voltage variation. 100
5·20 Delay distribution of the critical path in the 64-bit adder designed in UMC 130 nm process considering temperature variation. 101
5·21 Energy-delay product of a 64-bit adder designed using equalized logic (E-logic) vs. non-equalized logic (NE-logic) at zero word error rate at different technology nodes. The equalized logic approach reduces the energy-delay product of the sub-threshold logic by up to 23.6% across all technology nodes in the minimum energy supply voltage. 102

5·22 Adaptive memristor-based feedback equalizer circuit 104

6·1 Bypass flip flop design 109

6·2 Energy-delay trade-off in combinational logic. Traditional operation region is around minimum-delay point (MDP). Ultra low-energy region is around minimum-energy point (MEP) (Markovic et al., 2010). 111

6·3 Schematic of the 8T2R Memristor-based Nonvolatile (Rnv8T) SRAM cell (Chiu et al., 2012). 112

Chapter 1

Introduction

1.1 Background and Motivation

The design of fast and low-power memory and logic circuits is a critical part of designing very large scale integration (VLSI) chips that are extensively used today in applications ranging from biomedical implants to handheld devices to laptops/desktops to large data centers. Starting in the 1970's, the VLSI community was able to use complementary metal-oxide semiconductor (CMOS) technology scaling predicted by Moore's law (Moore, 1965) to sustain the historic improvement in performance and power of VLSI systems. Nevertheless, as CMOS technology is pushed to its atomic limits, the ability of hardware engineers to achieve power and performance improvements with every new technology generation becomes increasingly difficult (Kuhn, 2012; Borkar et al., 2004; Chuang et al., 2007) (Figure 1·1, Figure 1·2). In addition, the performance of most computer systems is increasingly limited by the capacity, access latency and energy consumption of on-chip memory blocks in today's computing systems. In particular, the size of the die limits the storage capacity of the memory block. Furthermore, most modern mobile applications require the use of nonvolatile memory to avoid losing data when the power supply is switched off to suppress the dynamic and static power consumption of the digital VLSI chips. Therefore, considering the degradation in CMOS device reliability, the limited available area, the energy

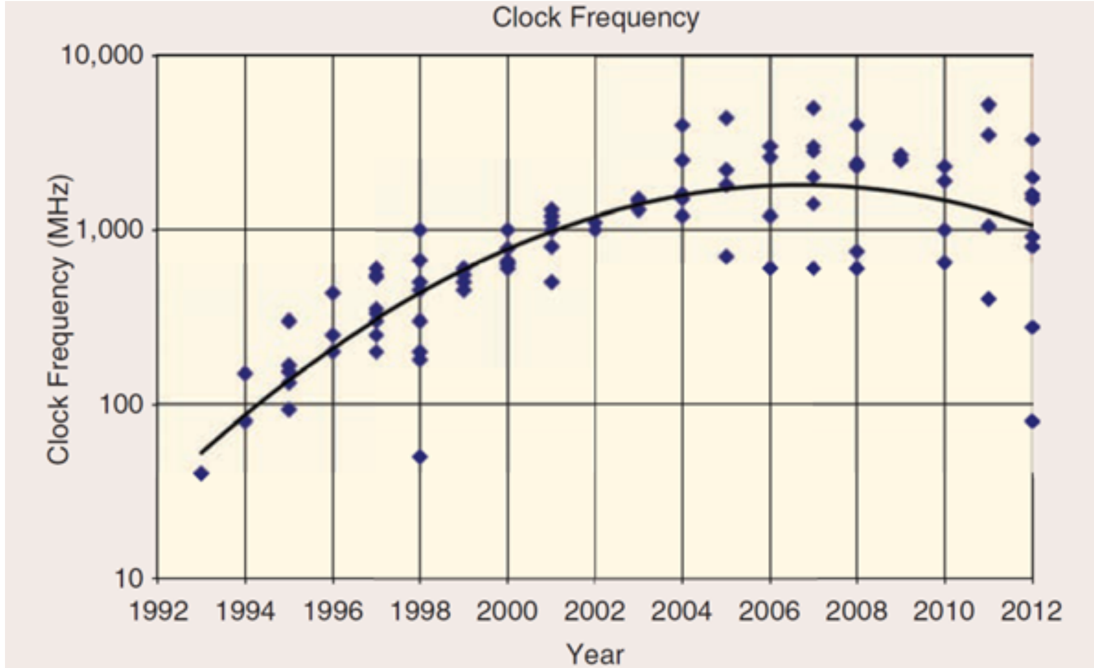

Figure 1·1: Saturation of clock frequency with CMOS scaling process (ISSCC Reports 2013).

and access latency for the design of on-chip memory blocks, it is necessary to explore emerging technologies and alternate circuit-level solutions which are compatible with the conventional CMOS process for designing highly-dense memory arrays. This dissertation addresses the use of recently-explored memristor technology to design dense nonvolatile on-chip memory arrays for today's computer architectures. 1-Transistor 1-memRistor (1T1R) cell based resistive random access memory (RRAM) arrays have low access latency, low access energy and large density (that can allow us to fit the entire working set of an application on the processor chip (Chiu et al., 2012)).

Table 1.1 shows a head-to-head comparison of the various nonvolatile emerging technologies. Each technology has its pros and cons, which has made it difficult to identify a successor to CMOS technology. Among these technologies, PCRAM requires large energy for its resistive switching behavior (Chung et al., 2011), FeRAM suffers from signal degradation in scaling process (Qazi et al., 2011) and MRAM has high en-

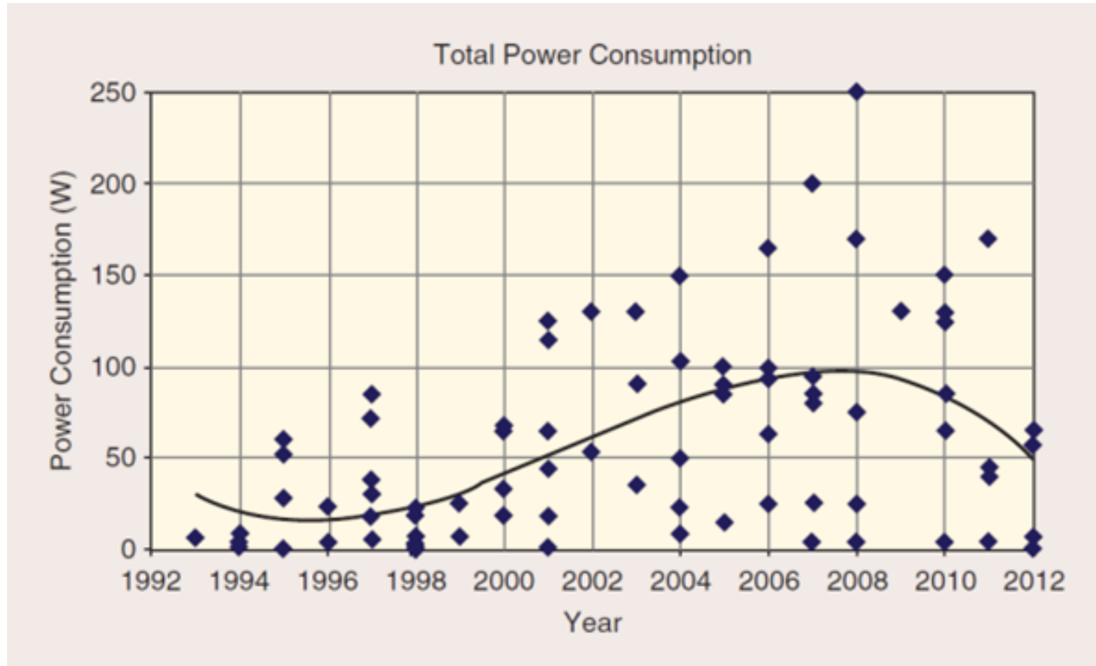

Figure 1·2: Trend of total power consumption with CMOS scaling process (ISSCC Reports 2013).

durance but it scales poorly and consumes large power due to large write currents (Nebashi et al., 2009). Among these emerging memory technologies, RRAM has been demonstrated to have high density capability due to multi-level cells (MLC) and cross-point array structures (Chen et al., 2009a). RRAM technology (memristor) has simple structure, high resistance ratio, fast-switching operation and device scalability beyond 10 nm technology node (Chiu et al., 2012). Figure 1·3 compares the storage capacity of the emerging RRAM technology with other modern nonvolatile memory technologies. The storage capacity of RRAM technology is approaching the storage capacity of Flash technology. HP Labs has already announced plans to commercialize memristor-based RAM and predicted that RRAM could eventually replace traditional memory technologies (Strukov et al., 2008). Therefore, the two-terminal memristor devices have been well-accepted as storage elements and are considered viable replacements to conventional CMOS-based memory designs. The memristor

Memory type	PCRAM (Chung et al., 2011)	MRAM (Nebashi et al., 2009)	FeRAM (Qazi et al., 2011)
Cell	1T1R	1T1R	1T1C
R/W time (ns)	76/20e3	12	200/134
Energy ($\frac{nJ}{bit}$)	15.3	0.9/1.3	9.77
Endurance	10^7	10^{16}	10^{13}
Retention	>10 yrs	>10 yrs	>10 yrs
Density ($\frac{Mb}{mm^2}$)	15.7	0.35	0.93
Tech. (nm)	58	90	130

Table 1.1: Comparison between current emerging nonvolatile memory technologies.

can be considered as a variable resistor which can be programmed by changing the voltage drop across the memristor or changing the current injected into the memristor. Here, programming amounts to changing the value of the memristance which leads to two different states for the memristor. These two states can correspond to storage of logic 0 and logic 1 in the memristor.

The memristor technology can also be used in designing low-energy logic circuits. Although historically higher performance has been the main motivation behind the CMOS scaling process, speed is not the ultimate goal for all modern applications of integrated circuits (ICs). Instead, a wide class of applications are emerging for which power and more importantly energy is the main problem. Ultra-low power sub-threshold circuits are becoming prominent in emerging embedded applications including wireless sensor networks and medical instruments where low energy operation is the main constraint instead of performance.

In sub-threshold circuits, scaling supply voltage into the sub-threshold region significantly reduces the dynamic energy consumed by digital circuits (Kwong et al., 2009). Scaling the supply voltage also lowers down the leakage current due to reduction in the drain induced barrier lowering (DIBL) effect resulting in considerable lower leakage power. However, as the supply voltage is scaled below the threshold voltage of

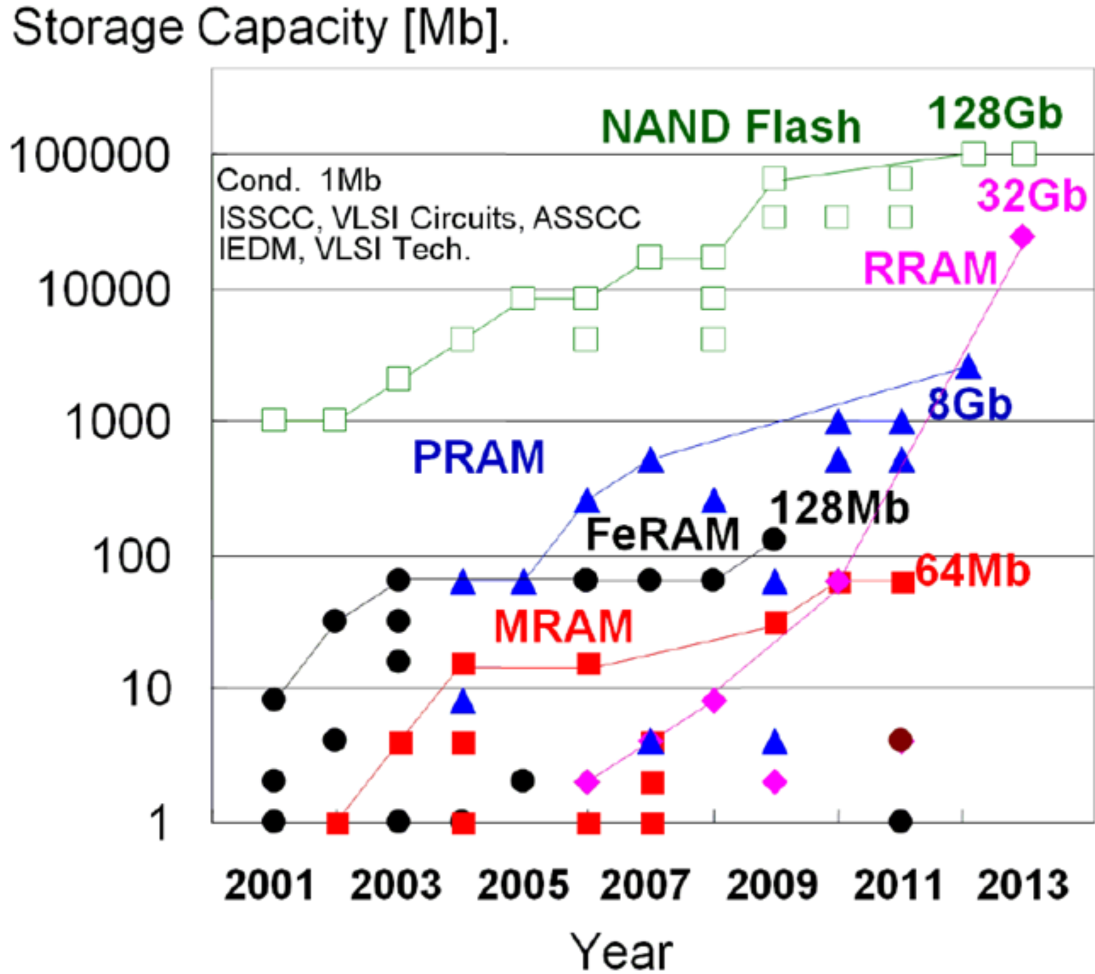

Figure 1·3: Storage capacity of nonvolatile memory technologies (ISSCC Reports 2013).

the transistors, the propagation delay of the logic gates increases enormously leading to the rise in leakage energy of the active devices operating in sub-threshold regime as the leakage power integrates over a longer period of time. As we scale the supply voltage, the two opposite trends in the leakage and the dynamic energy components lead to a minimum energy supply voltage and it has been shown in (Wang and Chandrakasan, 2005) that the minimum energy supply voltage of digital circuits occurs below the threshold voltage of the transistors.

Sub-threshold digital circuits, however, suffer from the degraded I_{ON}/I_{OFF} ratios

resulting in a failure in providing rail-to-rail output swings when restricted by aggressive timing constraints. Moreover, circuits working in weak inversion region suffer from process variations that directly affect the threshold voltage, which in turn has a significant impact on the drive current due to the exponential relationship between the drive current and the threshold voltage (V_T) of the transistors in sub-threshold regime. These degraded I_{on}/I_{off} ratios and process-related variations thus make sub-threshold circuits highly susceptible to timing errors which can further lead to complete system failures. Since the standard deviation of V_T varies inversely with the square root of the channel area (Pelgrom et al., 1989), one common approach to overcome the process variation is to upsize the transistors (Kwong et al., 2009). Similarly, increasing the logic path depth to leverage the statistical averaging of the delay across gates has been proposed in (Verma et al., 2008) to overcome process variations. A joint approach of choosing transistor sizes and logic depths that mitigate the impact of process variations has been proposed in (Zhai et al., 2005). Similarly, (Choi et al., 2004) proposes using gates of different drive strengths to overcome process variations. These approaches, however, increase the transistor parasitics, which in turn increases the energy consumption. Body-biasing approaches have also been proposed to mitigate the impact of variations in (Jayakumar and Khatri, 2005) and (Liu et al., 2011). It however necessitates extra complex on-chip circuitry to generate the required voltage for the substrate terminal of the CMOS devices to reduce the dominant leakage energy of the sub-threshold logic. Therefore, alternate circuit-level approaches are required to alleviate the timing errors while minimizing the energy consumption of the circuit.

1.2 Contribution and Organization

In Chapter 2, we summarize the recent efforts in design and modeling of memristor devices and then explain the detailed functionality of our target TiO_2 and HfO_x-based memristors. We also develop a new state function for HfO_x-based memristor devices and present it in this chapter. A new reliable SPICE netlist for HfO_x memristors is proposed based on the change in the conductive filament diameter.

In Chapter 3, we provide a detailed discussion of the functionality of an n-bit 1T1R RRAM cell followed by a description of the architecture of a memory array designed using this RRAM cell as the building block. We discuss the implementation of memory cells and arrays using both TiO_2 and HfO_x-based memristors. We also discuss our performance and energy models for the n-bit 1T1R memory arrays designed using TiO_2- and HfO_x-based memristors. We validate our performance and energy models against HSPICE simulations, and the difference is less than 10% for both n-bit TiO_2- and HfO_x-based 1T1R cells. Using energy and performance constraints, we determine the optimum number of bits/cell in the multi-bit RRAM array to be 3. The total write and read energy of the 3 bits/cell TiO_2-based RRAM array is 4.06 pJ/bit and 188 fJ/bit for 100 nsec and 1 nsec write and read access times while the optimized 3 bits/cell HfO_x-based RRAM array consume 365 fJ/bit and 173 fJ/bit for 1 nsec and 200 nsec write and read access times, respectively. We explore the trade-off between the read energy consumption and the robustness against process variations for uniform and non-uniform memristor state assignments in the multi-bit RRAM array. Using the proposed models, we analyze the effects of process, voltage and temperature variations on performance and energy consumption and the reliability of n-bit 1T1R memory cells. Our analysis show that multi-bit TiO_2 RRAM is more sensitive to Oxide Thickness Fluctuation (OTF) while HfO_x RRAM is more sensitive

to Line Edge Roughness (LER) and is more susceptible to voltage and temperature variations.

In Chapter 4, we explore the design of feedback equalizer circuits for digital logic circuits. The key idea here is to explore the use of communications-inspired techniques in the design of robust energy-efficient digital logic circuits. Feedback equalization for above-threshold regime has previously been proposed by (Takhirov et al., 2012) and we explore it for sub-threshold circuits. Using a feedback equalizer circuit that adjusts the switching thresholds of the gates (just before the flip flops) based on the prior sampled outputs, we can reduce the propagation delay of the critical path in the combinational logic block to make the sub-threshold system more robust to timing errors and at the same time reduce the dominant leakage energy of the entire design. We implement a non-equalized and an equalized design of an 8-bit carry lookahead adder in UMC 130 nm process using static complementary CMOS logic. In the equalized design, we could reduce the propagation delay of the critical path of the sub-threshold logic and correspondingly lower the dominant leakage energy, leading to 35.4% decrease in energy-delay product of the conventional non-equalized design at minimum energy supply voltage. Using the feedback equalizer circuit, we obtain 16.72% reduction in energy through voltage scaling while maintaining an operating frequency of 1.28 MHz. We show that the equalized sub-threshold 8-bit carry lookahead adder requires lower upsizing to tolerate process variation effects leading to 20.72% lower total energy.

In Chapter 5, we propose using an adaptive feedback equalizer circuit in the design of tunable sub-threshold digital logic circuits. This adaptive feedback equalizer circuit can reduce energy consumption and improve performance of the sub-threshold digital logic circuits. At the same time, the tunability of this feedback equalizer circuit enables post-fabrication tuning of the digital logic block to overcome worse than

expected process variations as well as lower energy and improve performance. We implement a non-equalized and an equalized design of a 64-bit adder in UMC 130 nm process using static complementary CMOS logic. Using the equalized design, the normalized variation of the total critical path delay can be reduced from 16.1% (non-equalized) to 11.4% (equalized) while reducing the energy-delay product by 25.83% at minimum energy supply voltage. Moreover, we show that in case of worse than expected process variation, the tuning capability of the equalizer circuit can be used post fabrication to reduce the normalized variation ($3\sigma/\mu$) of the critical path delay with minimal increase in energy. We also present detailed delay and energy models of the equalized digital logic circuit operating in the sub-threshold regime.

Chapter 2

Memristor Technology and Modeling

2.1 Introduction

Memristor is a two-terminal nanodevice that has been recently analyzed for its potential applications in memory design and logic design of both traditional and neuromorphic computing systems. It is a relatively well-explored device in terms of modeling, design methodology and its physical switching mechanism between two or more stable states. This chapter summarizes the current efforts on design and modeling of memristor devices and then explains the detailed functionality of our target TiO_2 and HfO_x-based memristors. A new state function that we developed for HfO_x-based memristor devices is also presented in this chapter.

2.2 Memristor Device Technology

Memristors provide a functional relationship between the charge and flux which was first postulated in (Chua, 1971). Several oxide-based memristor devices have been proposed as storage elements in the design of RRAM arrays. HfO_x and TaO_x have been widely used as switching elements in RRAM cells (Chen et al., 2009b), (Chen et al., 2009a), and (Lee et al., 2011). Although several fabricated RRAM prototypes based on different switching materials have been reported in the literature, only a few reliable device models have been proposed for large-scale circuit-level simula-

tions (Ielmini, 2011), (Bersuker et al., 2011), and (Lu et al., 2011). A numerical model of filament growth based on thermally activated ion migration, which accounts for the resistance switching characteristics is proposed in (Ielmini, 2011). This model (primarily developed for HfO_x-based 1T1R cell) matches the measurement results for different metal oxide RRAM configurations (HfO_x/ZrO_x, NiO). The authors in (Guan et al., 2012a) analyze the variation of switching parameters in RRAM devices using a trap-assisted-tunneling (TAT) current solver considering the stochastic generation and recombination of oxygen vacancies. The compact model for the proposed RRAM switching behavior in (Guan et al., 2012a) is introduced in (Guan et al., 2012b), while the measurement results of the HfO_x-based prototypes verify this model in (Yu et al., 2012).

There are multiple efforts in place to develop accurate analytical and SPICE models for the two-terminal memristor elements (Pickett et al., 2009), (Zangeneh and Joshi, 2012), (Ielmini, 2011). An analytical TiO_2 memristor model and the corresponding SPICE code that express both the static transport tunneling gap width and the dynamic behavior of the memristor state based on the measurement results are proposed in (Pickett et al., 2009) and (Abdalla and Pickett, 2011), respectively. The authors in (Kvatinsky et al., 2013) developed a simplified yet accurate analytical model for the TiO_2 tunnel barrier phenomena analyzed in (Pickett et al., 2009) with improved run times. In (Biolek et al., 2009), the authors developed a mathematical model for the prototype of memristor previously reported in (Strukov et al., 2008) with dependent voltage and current sources as well as an auxiliary capacitor which functions as integrator to calculate the state of the memristor. The authors in (Rak and Cserey, 2010) presented a schematic diagram of the memristor SPICE macromodel based on a simplified window function for the rate of change of state. A magnetic flux controlled SPICE model for memristors is proposed in (Batas and Fiedler, 2011)

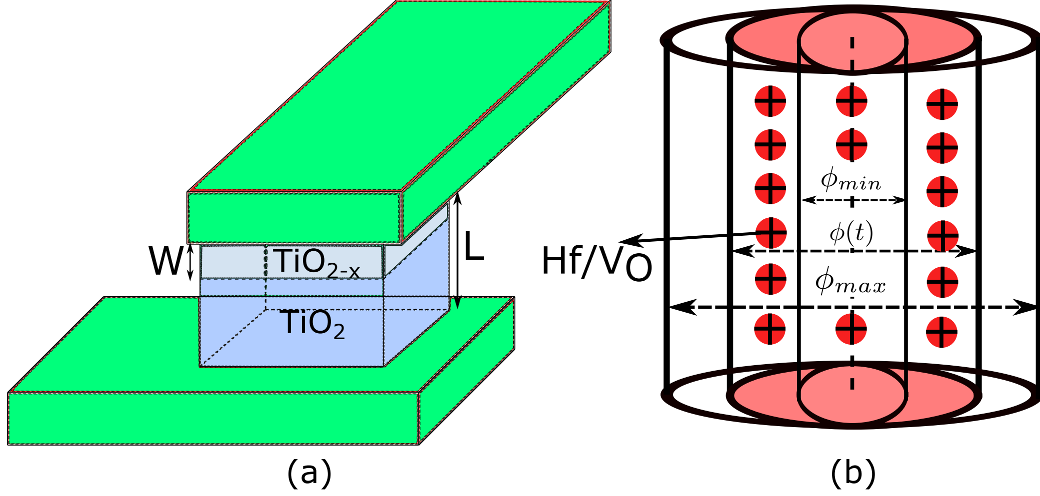

Figure 2·1: Physical structure of (a) TiO_2-based memristor between 2 Pt contacts consisting of a highly conductive doped region and a highly resistive undoped region, where L = thickness of the memristor and W = thickness of the conductive region, and (b) HfO_x-based memristor showing conductive filament growth/narrowing process where ϕ_{min} and ϕ_{max} are the minimum and maximum filament diameters, respectively.

based on an exponential relationship for memristor I-V characteristics. In this work we focus on titanium dioxide (TiO_2)- and hafnium oxide (HfO_x)-based memristor implementations.

The TiO_2-based memristor was first fabricated by HP (Strukov et al., 2008). The fabricated prototype had a highly resistive thin layer of TiO_2 and a second conductive deoxygenized TiO_{2-x} layer (see Figure 2·1a). The change in the oxygen vacancies due to a voltage applied across the memristor modulated the dimension of the conductive region in the memristor. This resulted in a high resistance state and a low resistance state corresponding to the resistive and conductive region of operation, respectively. The effective 'memristance' of the memristor device can be calculated using Equation (2.1) (proposed in (Strukov et al., 2008)).

$$M(t) = R_{ON}x(t) + R_{OFF}(1 - x(t)). \tag{2.1}$$

Parameter	TiO_2	HfO_x
$R_{ON}(\Omega)$	100	3K
$R_{OFF}(\Omega)$	16K	10M
$L(nm)$	10	20
$E_{A0}(eV)$	-	1.2
$A(ms^{-1})$	-	1
α	-	0.3
$\rho(\mu\Omega cm)$	-	400
$k_{th}(Wm^{-1}K^{-1})$	-	20

Table 2.1: Parameters of TiO_2-based (Strukov et al., 2008) and HfO_x-based (Ielmini, 2011), (Sheu et al., 2009) memristors used for modeling and simulations.

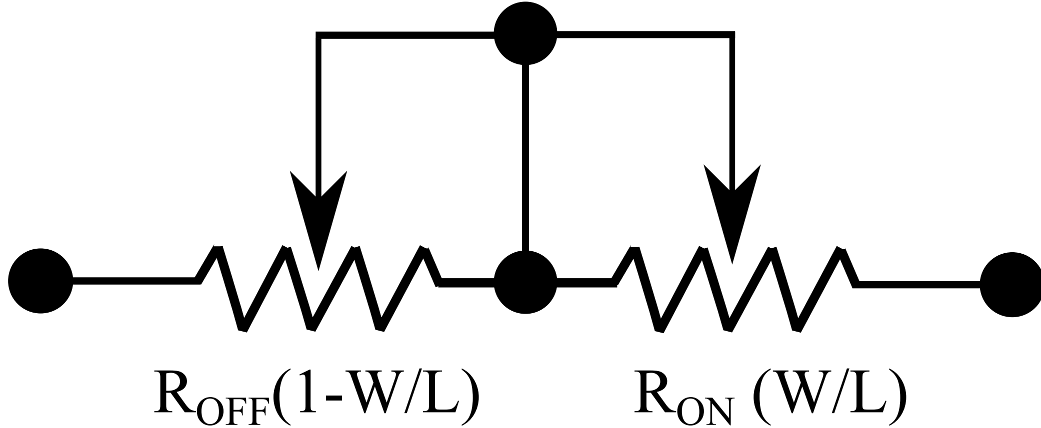

Figure 2·2: Equivalent resistance of memristor devices.

Here, R_{ON} and R_{OFF} are the minimum and maximum memristances, respectively, and $x(t)$ is the state of the memristor (Eshraghian et al., 2011) (see Figure 2·2). This state of the memristor can be calculated as $w(t)/L$, where $w(t)$ is the thickness of the conductive doped region as a function of time, and L is the memristor thickness.

The rate of change of the memristor state follows the ionic drift model which is a function of the memristor physical parameters and the current through the memristor. As the current itself varies with time, the change of memristor state exhibits nonlinear behavior. This nonlinear behavior can be expressed using a window function shown in Equation (2.2) (Eshraghian et al., 2011).

$$\frac{dx}{dt} = \frac{\mu_v R_{ON}}{L^2} i(t) F(x(t), p) \tag{2.2}$$

In Equation (2.2), $\mu_v \approx 3 \times 10^{-8} m^2/s/V$ (Witrisal, 2009) is the average dopant mobility, $F(x(t), p)$ is the window function, where the parameter p controls the memristor nonlinearity. Increasing p yields a flat window function for larger memristor states. Window functions that consider the linear ionic drift, and the nonlinear behavior that appears at the boundaries of the memristor state, have been proposed in (Benderli and Wey, 2009) and (Joglekar and Wolf, 2009). However, both these window functions get stuck at the memristor state boundaries. We use the window function proposed in (Biolek et al., 2009) for developing the performance and energy models of the TiO_2-based RRAM cell. This function models the nonlinear behavior of the rate of change of state without getting stuck at the boundaries and is given in Equation (2.3).

$$F(x(t), p) = 1 - (x - sgn(-i(t))^{2p} \tag{2.3}$$

Here, $i(t)$ is the current through the memristor, sgn is a sign function that prevents the state of the cell from getting stuck at the borders and p is the control parameter. Figure 2·3 shows a plot of the window function for different p values.

In case of the HfO_x-based memristor, the set/reset (changing memristor resistance to R_{ON}/R_{OFF}) process is performed by increasing/decreasing the diameter of the conductive filament (CF) using positively charged oxygen vacancies (V_O) or Hf ions migration in a thermally activated hopping process in the filament growth model (Ielmini, 2011). Applying a voltage across the HfO_x-based memristor forces the positive ions to move along the direction of the electric field while increasing the maximum tem-

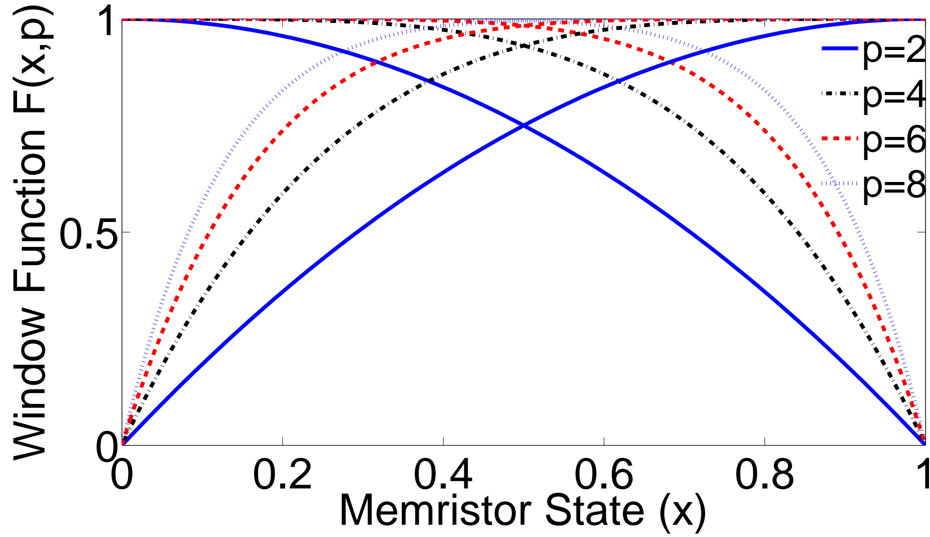

Figure 2·3: Dynamic window function of the memristor state showing the nonlinear behavior of the memristor for different control parameter p. The current sign function prevents the state from getting stuck at the two boundaries.

perature along the CF and changing the effective cross section diameter of the CF (see Figure 2·1b). This rate of change of diameter was derived in (Ielmini, 2011) and is given by

$$\frac{d\phi}{dt} = Ae^{-\frac{E_{A0}-\alpha qV}{kT_0(1+\frac{V^2}{8T_0\rho k_{th}})}} \tag{2.4}$$

where, ϕ is the CF diameter, A is a pre-exponential constant, E_{A0} is the energy barrier for ion hopping, α is the barrier lowering coefficient, q is the elementary charge, V is the applied voltage across the memristor, k is the the Boltzmann constant, T_0 is the room temperature, ρ is the electrical resistivity and k_{th} is the thermal conductivity. A similar expression with a negative rate of change is used for modeling the reset process in HfO_x-based memristors. As voltage is applied across the HfO_x-based memristor, its cross section area changes and the instantaneous resistance of the CF changes

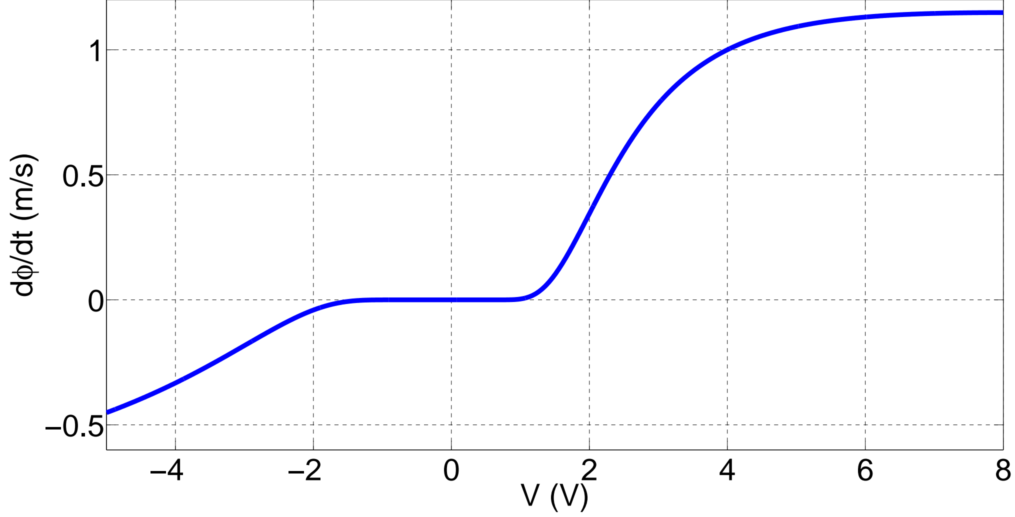

Figure 2·4: Rate of diameter change for HfO_x-based memristors in filament growth model (Ielmini, 2011) for set (V>0) and reset (V<0) operations as a function of voltage across the memristor.

according to $R(t) = 4\rho L/\pi\phi(t)^2$. The rate of change of the diameter for HfO_x-based memristors in filament growth model for set and reset operations is shown in Figure 2·4. The nominal parameter values of the memristor used for generating this plot are listed in Table 2.1. To minimize the destruction of the stored data during read operation, we maintain the voltage across the memristor to be greater than -1.7 V. Similarly, during write operation we maintain the applied voltage between 1 V to 4 V to minimize the set operation time.

To find the instantaneous memristance of the HfO_x RRAM, we define a new state function for HfO_x memristors as in Equation (2.5).

$$x(t) = C\left(1 - \frac{\phi_{min}^2}{\phi(t)^2}\right) \tag{2.5}$$

where the coefficient C is

$$C = \frac{\phi_{max}^2}{\phi_{max}^2 - \phi_{min}^2} = (1 - 1/\beta) \tag{2.6}$$

Here, ϕ_{max} and ϕ_{min} are the maximum and minimum CF diameters corresponding to R_{ON} and R_{OFF}, and $\beta = R_{OFF}/R_{ON}$. This state function can be plugged into equation (2.1) to calculate the effective memristance. Considering the rate of change of the CF diameter in (2.4) and the state function in (2.5), we define the rate of change of the HfO_x-based memristor state in Equation (2.7).

$$\frac{dx}{dt} = \frac{2C\sqrt{(1-x/C)^3}}{\phi_{min}}\frac{d\phi}{dt}. \tag{2.7}$$

The corresponding HSPICE netlist that we developed for HfO_x-based memristors is:

```
.SUBCKT memristorHfOx PLUS MINUS phi
.PARAM phimin='sqrt(4*ro*L/(3.14*Roff))'
.PARAM phimax='sqrt(4*ro*L/(3.14*Ron))'
.PARAM C='phimax*phimax/(phimax*phimax-
phimin*phimin)'
Csv phi 0 1
.IC V(phi) 0.3
Emem PLUS AUX VOL='I(Emem)*(V(phi)*Ron+
(1-V(phi))*Roff)'
Rtest AUX MINUS 1
Gsv 0 phi
CUR='C*phimin*phimin*POW(sqrt(phimin*phimin
/(1-(phimax*phimax-phimin*phimin)*V(phi)/
(phimax*phimax))),-3)*2*A*exp(-1*(EA0-alpha*
```

```
q*V(PLUS,MINUS))/(k*T0*(1+POW(V(PLUS,MINUS)
,2)/(8*T0*ro*kth)))) * sgn(I(Emem)) * sgn((1-V(phi)+
sgn(sgn(-I(Emem))+1))) * sgn((sgn(V(phi))+
sgn(I(Emem))+1))'
.ENDS memristorHfOx
```

The rate of change of the HfO_x-based memristor state is modeled as a voltage-controlled current source, and the combination of *sgn* functions guarantees the reliable set/reset operations, and the normalized memristor state does not get stuck when approaching 1 or 0.

2.3 Summary

In this chapter, we first described the behavioral functionality of the TiO_2 memristor based on the ionic drift model. We then proposed the state function for the HfO_x-based memristors. A new reliable SPICE netlist for HfO_x memristors was proposed based on the change in the conductive filament diameter.

Chapter 3

Design of Multi-bit RRAM Array

3.1 Introduction

In Chapter 2 we provided a detailed description of the memristor technology. In this chapter we provide a detailed discussion of the functionality of an n-bit 1T1R RRAM cell followed by a description of the architecture of a memory array designed using this RRAM cell as the building block. We discuss the implementation of memory cells and arrays using both TiO_2 and HfO_x-based memristors. We also discuss our performance and energy models for the n-bit 1T1R memory arrays designed using TiO_2- and HfO_x-based memristors.

3.2 Related Work

Several memory circuit/architecture topologies have been proposed in the literature based on the memristive structures. The authors in (Jo et al., 2009) used a Si-based memristive system to fabricate high-density crossbar arrays with high yield and OFF/ON ratio. A memristor-based TiO_2 memory cell is introduced in (Ho et al., 2011) and its functionality is evaluated using system-level simulations. An energy-efficient dual-element TiO_2-based memory structure is proposed in (Niu et al., 2010a), in which each memory cell contains two memristors that store the complementary states. Similarly, a 2-bit storage memristive cell is proposed in (Manem and

Rose, 2011). Both these multi-bit memory cells have large area. Content addressable memory (CAM) designed using TiO_2 memristors has been introduced in (Eshraghian et al., 2011). A memristor-based Look Up Table (LUT) design has been introduced in (Chen et al., 2012) to replace the SRAM-based FPGA design while achieving higher density. In (Fei et al., 2012), the functionality, performance and power of several CMOS/memristor based circuits with memory applications have been verified using a simulator based on a Modified Nodal Analysis. An analysis of the peripheral circuitry of the crossbar array architecture is presented in (Xu et al., 2011). A nonvolatile 8T2R SRAM cell that uses two HfO_x-based 1T1R cells along with the conventional 6T SRAM structure is introduced in (Chiu et al., 2012) for low power mobile applications. A bridge-like neural synaptic circuit with 5 TiO_2-based memristors which is capable of performing sign/weight setting and synaptic multiplication operations is introduced in (Kim et al., 2012b). A memristor emulator composed of the basic circuit-level elements is designed in (Kim et al., 2012a). The authors in (Liauw et al., 2012) presented a 3D-FPGA with stacked RRAM technology achieving lower energy-delay product (EDP) and smaller area compared to the conventional 2D-FPGA design. In (Xue et al., 2012), the authors proposed adaptive write and read circuits for RRAM arrays to enhance yield and β ratio while eliminating large power consumption rising from the resistance fluctuations.

Memristors are highly vulnerable to process variation and several authors have analyzed its impact on the functionality of the memristive structures. Line-Edge Roughness (LERs) caused by uncertainties in the process of lithography and etching (Jiang et al., 2009), Oxide Thickness Fluctuations (OTFs) caused during sputtering or atomic layer deposition, and Random Discrete Doping (RDDs), which leads to randomness in resistivity of the conductive as well as the resistive region of the memristor, are generally the main causes of process variations. The authors in (Niu et al., 2010b)

have analyzed the effect of cross section area and oxide thickness variations on the memristor resistance. The authors in (Hu et al., 2011a) have analyzed the effect of LER and OTF on the state $x(t)$, the rate of change of state $dx(t)/dt$ and power dissipation variations of TiO_2-based memristor. Using an Error Correcting Code (ECC) design that is commonly used in conventional DRAM memory, the authors in (Niu et al., 2012) propose the detection and mitigation of errors rising from process variations in both MOS-based and crossbar memristive RRAM cells. The authors in (Sheu et al., 2011) have used a Parallel-Series Reference-Cell (PSRC) scheme to decrease the reference current fluctuations in 1T1R RRAM structure. Moreover, using a Process-Temperature-Aware Dynamic BL-bias (PTADB) circuit, they lower the read disturbance caused by bitline voltage variations.

We present the detailed energy and performance models of multi-level 1T1R RRAM cells that use TiO_2- and HfO_x-based memristors. For the HfO_x-based array design, we use the filament growth model in (Ielmini, 2011) that has been validated against measurement results. We determine the optimum number of bits per RRAM cell that consumes the least energy while being constrained by cell performance. We apply the Monte-Carlo methodology in (Hu et al., 2011a) to model the effects of LER, OTF and RDD on the functionality of multi-bit HfO_x as well as TiO_2 RRAM cells.

3.3 RRAM Cell Design

The circuit of the 1T1R RRAM cell is similar to a DRAM cell and consists of an access transistor and a memristor as storage element (see Figure 3·1). Similar to DRAM, the access transistor is enabled for both read and write operations. As the memristor device shows considerable nonlinearity when approaching the states of 0 ($R_m = R_{OFF}$) and 1 ($R_m = R_{ON}$), it increases the required set/reset operation times

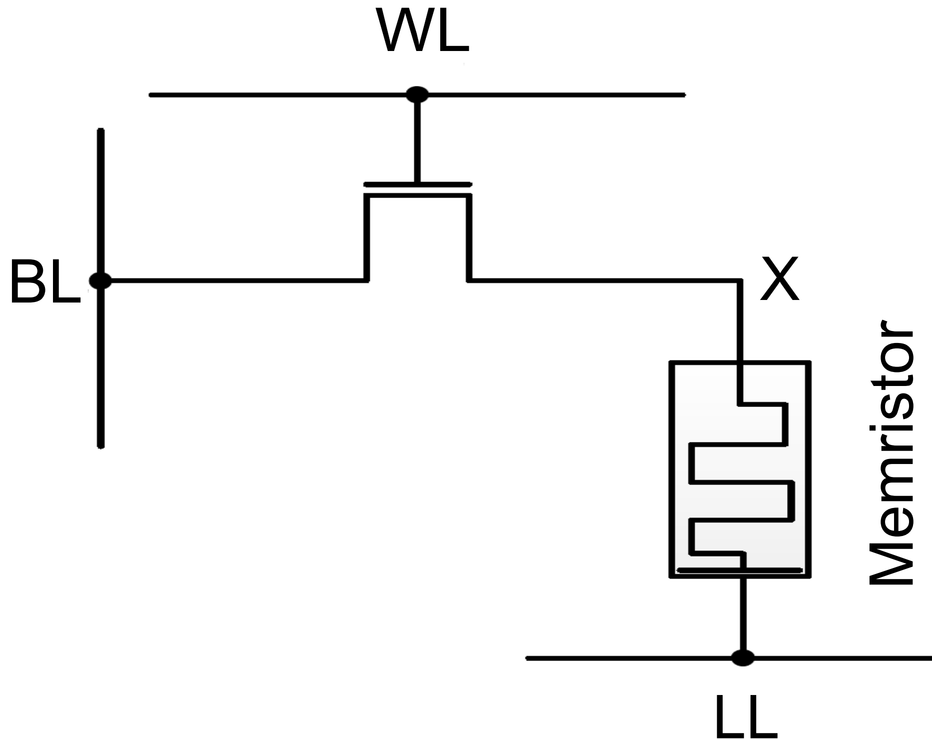

Figure 3·1: 1-transistor 1-memristor (1T1R) RRAM cell.

at the two boundaries. We therefore ignore the states smaller than 0.1 and larger than 0.9 for faster set/reset i.e. write operations. The n bits of a cell are stored in the 2^n distinct sub-ranges in the range of 0.1 to 0.9. For an n-bit cell design, the state assignment can be done such that maximum noise margin would be achieved. For example, for a 2-bit RRAM cell, a memristor state below 0.3 corresponds to 00, a memristor state between 0.3 and 0.5 corresponds to 01, a memristor state between 0.5 and 0.7 corresponds to 11 and a memristor state above 0.7 corresponds to 10. We use Gray coding to increase the robustness and minimize the probability of getting two bits in error in the read operation. We refer to this assignment as uniform state assignment. A non-uniform state assignment could also be used for the n-bit cell. A comparison of the two assignments is presented in Section 3.8.

To perform the read operation, the loadline is driven to charge the bitline through the memristor and access transistor. The read operation of the n-bit RRAM cell may be destructive and could require periodic refreshing of the cell data. For threshold-based memristor technologies recent measurement results have shown that if the drive voltage is less than a threshold, the state does not change for fast read operations (see Figure 2·4). The TiO_2 RRAM - based on the ionic drift model - is not a threshold-based technology (Kvatinsky et al., 2013) and shows more destructiveness during read cycles. A detailed analysis of the read destructiveness in multi-bit RRAM cells is proposed in Section 3.5.

The write operation always consists of two sub-operations – read followed by write as we need to know the data currently stored in the cell to determine the exact voltage that needs to be applied across the memristor to write new data. To perform the write operation, a positive or negative voltage is applied across the memristor for transitions to higher or lower states, respectively. The current flowing through the memristor changes the size of conductive region (in ionic drift model) or changes the diameter of the conductive filament (in filament growth model), thus increasing or decreasing the 'memristance'. In the rest of the thesis, we refer to the memory read and write operations as read_{top} and write_{top}, and the sub-operations as read_{sub}, refresh_{sub} and write_{sub}. Thus $\text{read}_{top} = \text{read}_{sub} + \text{refresh}_{sub}$, while $\text{write}_{top} = \text{read}_{sub} + \text{write}_{sub}$.

3.4 RRAM Array Architecture

The overall architecture of a memory array built using 1T1R RRAM cells is similar to the conventional DRAM array i.e. a wordline is used to select a row of cells, and a bitline is shared by the cells in a column for reading/writing (see Figure 3·2). In

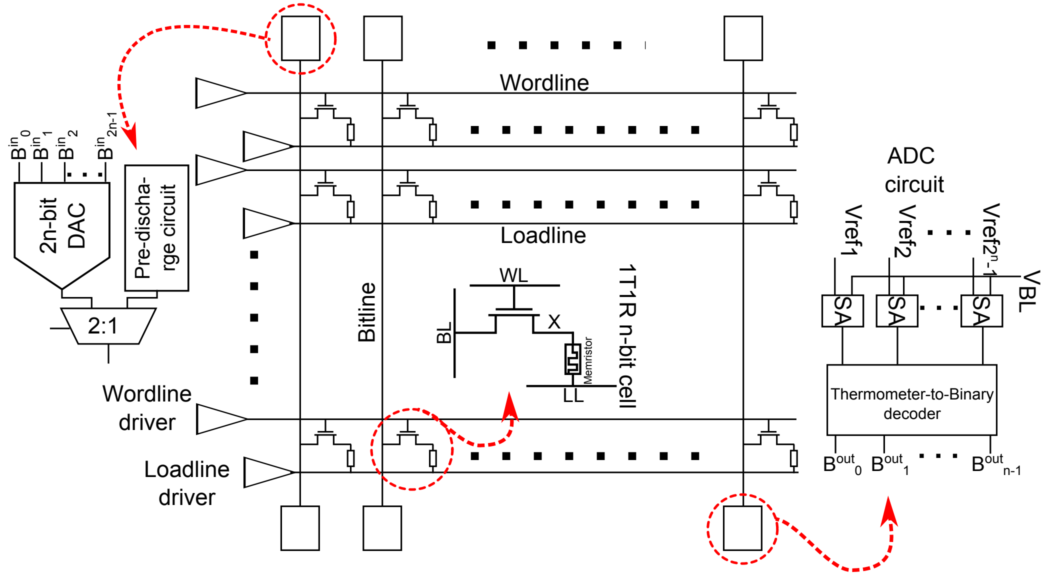

Figure 3·2: n-bit/cell RRAM array architecture.

an RRAM array architecture, to perform the read$_{sub}$ operation, we first discharge the bitline (BL) to 0 V, and then enable the wordline (WL) and loadline (LL) for a fixed predefined time. For the n-bit/cell array, when the WL and LL are enabled, the bitline charges to one of the 2^n distinct voltages corresponding to the 2^n distinct data values (i.e. the memristor state) stored in the cell. For instance in a 2-bit/cell array, there will be 4 distinct data values. An analog-to-digital converter (ADC) can be used to retrieve the n bits in each cell during the read operation. Each n-bit ADC consists of $2^n - 1$ differential sense-amplifiers, each having the V_{BL} as one input and a unique reference voltage (V_{refi}) as the other input. For example a 2-bit/cell array needs 3 differential sense amplifiers. The $2^n - 1$ sense amplifiers are shared by all the cells in the column. The sense amplifiers could be shared between columns to relax the area constraints on sense amplifier design. The rail-to-rail outputs of the sense amplifiers are fed to thermometer-to-binary code decoders that determine the exact data stored in the n-bit 1T1R cell and is given by bit B_0^{out} to B_{n-1}^{out}. We use the multiplexer-based decoder introduced in (Sail and Vesterbacka, 2004) which has

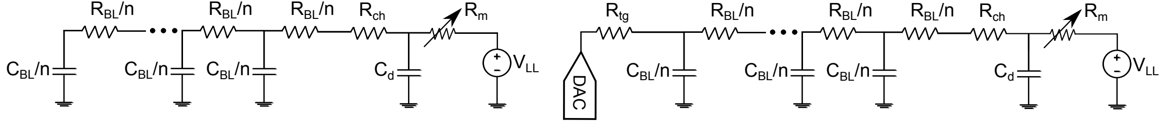

Figure 3·3: Equivalent circuit of 1T1R cell for read$_{sub}$ (left) and write$_{sub}$/refresh$_{sub}$ (right) operation.

a short critical path and consumes low power.

To perform the write$_{sub}$ operation, one of the $2^{2n} - 2^n$ different voltages (corresponding to the $2^n(2^n - 1)$ possible transitions for the n-bit RRAM cell) need to be applied across the memristor. For example, a 2-bit/cell array needs 12 voltages corresponding to 12 different transitions. The refresh$_{sub}$ operation would be similar to the write$_{sub}$ operation and the applied voltage will depend on the mechanism used for refresh operation. A 2n-bit multiplexer-based digital-to-analog converter (DAC) can be used to generate the voltages to be applied across the memristor for write$_{sub}$/refresh$_{sub}$ operation. During write$_{sub}$/refresh$_{sub}$ operation, the outputs B_0^{out} and B_{n-1}^{out} are connected to the B_0^{in} and B_{n-1}^{in} inputs (corresponding to the current stored bits) and the data to be written into the cell is connected to the B_n^{in} and B_{2n-1}^{in} inputs of the 2n-bit DAC. This ensures the DAC generates the correct voltage to be applied to the bitline for writing the data. For the 2-bit/cell array, we need a 4-bit DAC that generates 12 different set/reset voltages and an ADC with 3 sense amplifiers.

3.5 Performance Models

As discussed in Section 3.3, the $read_{top}$ and $write_{top}$ operation of the n-bit 1T1R cell consists of read$_{sub}$ + refresh$_{sub}$ and read$_{sub}$ + write$_{sub}$ operations, respectively. The equivalent circuit model for the 1T1R RRAM cell during read$_{sub}$ operation is shown in Figure 3·3. Here, R_m is the equivalent time-variant resistance of the memristor and R_{ch} is the access transistor channel resistance while operating in the triode region. The

transmission gate which is part of the pre-discharging path of the bitline capacitor is not included here as that transmission gate is switched OFF as soon as BL is discharged resulting in very high equivalent resistance for the transmission gate. C_{BL} and C_d are the bitline capacitor and access transistor junction capacitor, respectively. Also R_{BL} is the total resistance of the bitline. The bitline voltage at the end of read$_{sub}$ operation (i.e. after time T_R) will be

$$V_{BL} = V_{LL}(1 - e^{\frac{-T_R}{(R_m(t)+R_{ch}+0.5R_{BL})C_{BL}}}). \tag{3.1}$$

Here the time constant of the junction capacitor (C_d) is much smaller than that of the bitline capacitor (C_{BL}), and hence $C_{BL} + C_d$ has been approximated to be equal to C_{BL}. Also, the term $0.5R_{BL}C_{BL}$ is the intrinsic time constant of the bitline modeled as a distributed RC-line. We assume the bitline, wordline and loadline to be 1 mm long, each with total capacitance of 200 fF and total resistance of 6.5 $K\Omega$ corresponding to copper metal line with 50 nm × 50 nm cross-section area. In addition, we assume the distributed RC-line model with 80 segments for all of the interconnects in the RRAM array architecture. For a n-bit RRAM cell, equation (3.1) can be used to define the $2^n - 1$ reference voltages to be input to the sense amplifiers that are used to differentiate between the different stored values while performing read$_{sub}$ operation. For example, for a 2-bit RRAM cell, we can use equation (3.1) to determine the three different reference voltages to differentiate between the four different stored values. The bitline voltage depends on the data stored in the memristor, i.e. the memristor state. For $V_{ref1} > V_{BL}$, $V_{ref1} < V_{BL} < V_{ref2}$, $V_{ref2} < V_{BL} < V_{ref3}$ and $V_{ref3} < V_{BL}$ the stored data is 00, 01, 11 and 10, respectively. In Table 3.1 and Table 3.2, we compare the reference voltages calculated using the analytical model shown in Equation (3.1) and using HSPICE simulation using 22 nm PTM technology (Ptm,). The parameters of TiO_2 and HfO_x-based memristors that are

Reference Voltages	AM TiO_2 1 nsec	HS TiO_2 1 nsec	AM TiO_2 2 nsec	HS TiO_2 2 nsec
V_{ref1}	137.5 mV	120.5 mV	162 mV	158.17 mV
V_{ref2}	168 mV	153.5 mV	190.4 mV	188.69 mV
V_{ref3}	215.5 mV	202.5 mV	228.9 mV	231.37 mV

Table 3.1: Comparison between the reference voltages determined using analytical model (AM) and HSPICE simulation (HS) for a read$_{sub}$ access time of $T_R(TiO_2) = 1, 2\ nsec$ in the 2-bit/cell 1T1R RRAM. $V_{LL}(TiO_2) = 0.48\ V$ is chosen to reach to at least 25 mV difference between the two adjacent reference voltages. The average error is 5.7% for TiO_2.

Reference Voltages	AM HfO_x 200 nsec	HS HfO_x 200 nsec	AM HfO_x 400 nsec	HS HfO_x 400 nsec
V_{ref1}	94.5 mV	94.79 mV	100.85 mV	100.84 mV
V_{ref2}	130.5 mV	130.9 mV	135.25 mV	135.23 mV
V_{ref3}	214 mV	214.6 mV	204.8 mV	204.81 mV

Table 3.2: Comparison between the reference voltages determined using analytical model (AM) and HSPICE simulation (HS) for a read$_{sub}$ access time of $T_R(HfO_x) = 200, 400\ nsec$ in the 2-bit/cell 1T1R RRAM. $V_{LL}(HfO_x) = 0.7\ V$ is chosen to reach to at least 25 mV difference between the two adjacent reference voltages. The average error is 0.151% for HfO_x.

used in modeling and HSPICE simulations are summarized in Table 2.1. Here the read time of 1, 2 ns (for TiO_2) and 200, 400 ns (for HfO_x) is chosen based on the nominal β value for the two types of memristors (see Table 2.1). HfO_x has larger β and R_{OFF} values compared to TiO_2, and therefore it needs higher read time for reliable read operation. If we ignore the destructiveness (changing the memristance) during read$_{sub}$ in the analytical model for simplicity, the resulting average error is 5.7% for TiO_2 and 0.151% for HfO_x.

To ensure a reliable read operation, there should be sufficient difference in the four different voltages developed on the bitline corresponding to the 4 different data that can be stored in the 2-bit cell. For very large bitline voltage development times, the

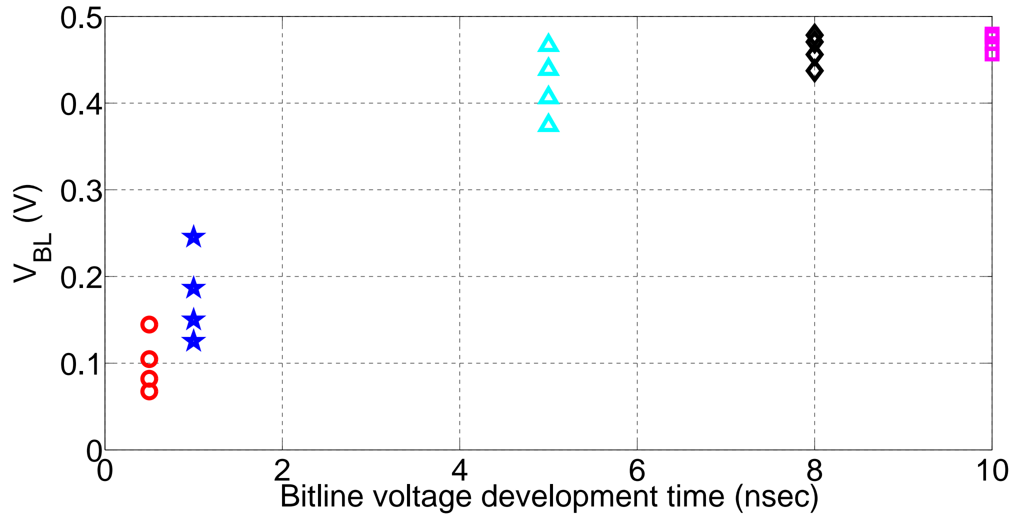

Figure 3·4: Bitline voltage of a 2-bit/cell TiO_2-based RRAM for different bitline voltage development times.

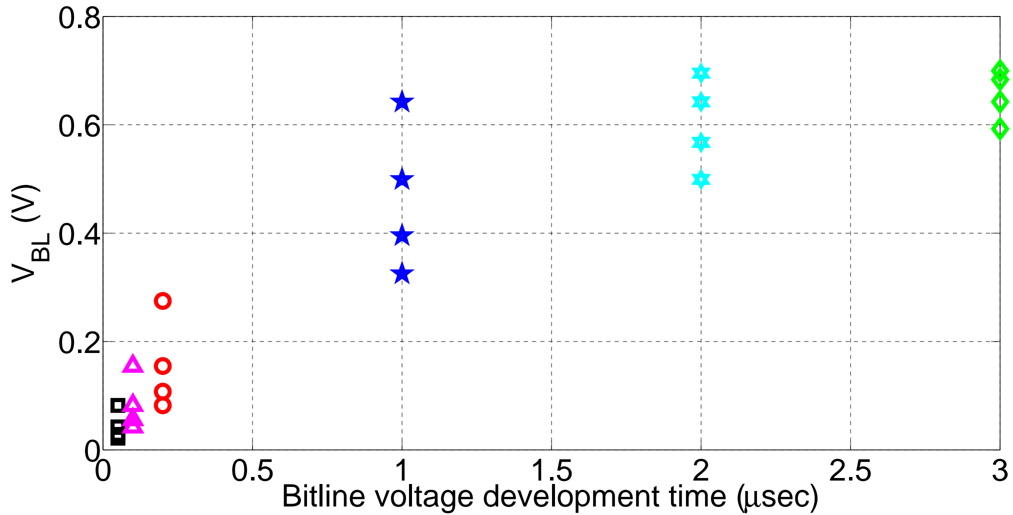

Figure 3·5: Bitline voltage of a 2-bit/cell HfO_x-based RRAM for different bitline voltage development times.

bitline can get completely charged to the load line voltage (V_{LL}). At the same time, for very small bitline voltage development times, the difference in the bitline voltages may not be large enough for the sense-amplifiers to correctly determine the data stored in the cell. The bitline voltage of TiO_2- and HfO_x-based 2-bit/cell RRAM cells for various bitline voltage development times during read operation are shown in

Figures 3·4 and 3·5, respectively. For our 2-bit/cell RRAM array example, we design our sense amplifier such that it needs at least 12.5 mV differential inputs. Hence, we need at least 25 mV difference between the adjacent bitline voltages corresponding to the 4 different data that can be stored in the 2-bit cell. The V_{ref} inputs to the three sense amplifiers are chosen based on bitline voltages (corresponding to the four different data that can be stored in the cell) while ensuring the 12.5 mV differential input. So for the TiO_2- and HfO_x-based 2-bit/cell RRAM cells we choose bitline development time of 1 nsec and 200 nsec, respectively. In the TiO_2-based cell, for the 1 nsec read access time, the four different bitline voltages are 125 mV, 150 mV, 186 mV and 245 mV. The corresponding V_{ref1}, V_{ref2} and V_{ref3} values are 137.5 mV, 168 mV, and 215.5 mV, respectively. Similarly, in the HfO_x-based cell, for the 200 nsec read access time, the four different bitline voltages are 82 mV, 107 mV, 154 mV and 274 mV. The corresponding V_{ref1}, V_{ref2} and V_{ref3} values are 94.5 mV, 130.5 mV, and 214 mV, respectively. The read times as a function of number of bits/cell (n) is illustrated in Figure 3·6. These read times have been chosen using the same approach as described above for the 2 bits/cell RRAM cell. As the value of n increases, we need larger read times to ensure the reliable read operation.

As discussed in section 3.3, the read$_{sub}$ operation of the 1T1R cell can be destructive. The read destructiveness of TiO_2-based memristors is larger compared to HfO_x-based memristors for the same loadline voltage (V_{LL}). The TiO_2-based memristor therefore needs to be refreshed more frequently than HfO_x-based memristor. Considering the rate of change of state for TiO_2 RRAM in Equation (2.2), the number of consecutive read operations that will not destruct the stored data in multi-bit TiO_2-based 1T1R RRAM cell, i.e. the refresh threshold can be written as (Zangeneh and Joshi, 2014a)

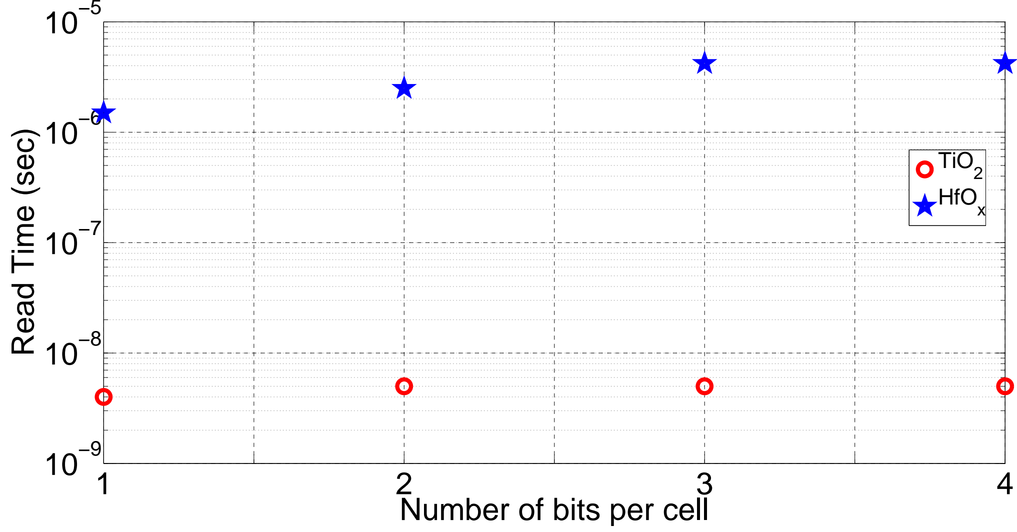

Figure 3·6: Read time of a multi-bit RRAM cell for different number of bits per cell.

$$t_{ref-TiO_2} \approx \frac{(x_{max} - x_{min})(R_m(x) + R_{ch})}{2^n \gamma T_R V_{LL}(1 - (x-1)^{2p})}. \tag{3.2}$$

Here, $R_m(x)$ is the resistance of the memristor for each state, n is the number of bits/cell, T_R is the read access time, x_{max} and x_{min} are the maximum and minimum normalized memristor states (0.9 and 0.1 in this work), respectively and $\gamma = \frac{\mu_v R_{ON}}{L^2}$. Large V_{LL}, n and R_{ON} values (smaller β) necessitate more frequent refresh operation in the multi-bit RRAM cell. The contour plots of the number of consecutive non-destructive read operations in multi-bit TiO_2 RRAM is shown in Figure 3·7 for different n (number of bits/cell) and V_{LL} values for a memristor with initial state of $x = 0.9$. In case of the highly destructive multi-bit TiO_2 memristor, we explored two different refresh schemes: A refresh operation can be performed after each read cycle to compensate for destructiveness (Niu et al., 2010b). In this refresh scheme, we apply a -V_{LL} for the same duration as $read_{sub}$. This doubles the read energy and lowers the performance of the RRAM array. A second refresh approach is to use a

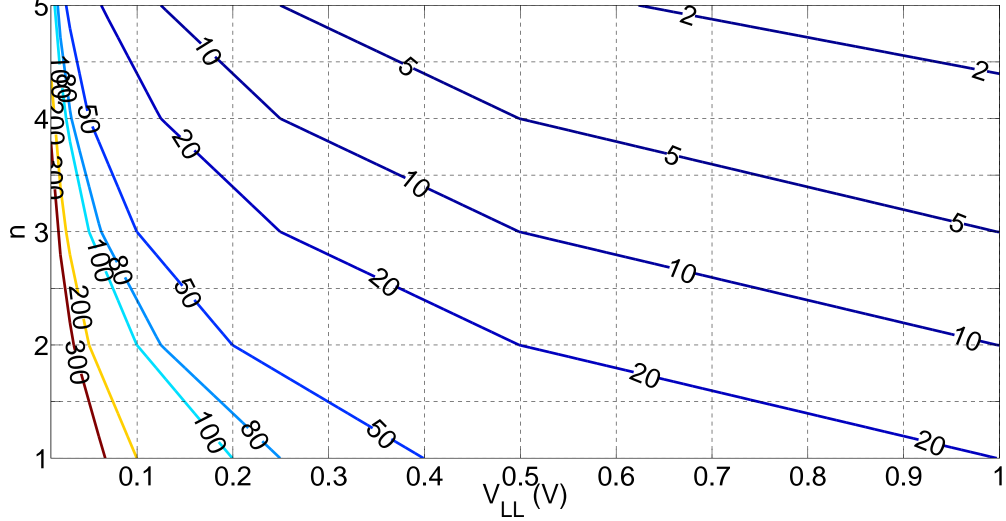

Figure 3·7: Contour plots of the number of consecutive non-destructive read cycles in multi-bit TiO_2-based RRAM for different n and V_{LL} values ($x = 0.9$).

counter to track the current state of the memristor as well as the number of consecutive read operations. A refresh operation is done once the number of consecutive read operations on the multi-bit TiO_2 RRAM cell exceeds the threshold. For instance, in a 3-bit/cell TiO_2-based RRAM array with $V_{LL} = 0.1\ V$, 50 consecutive read cycles will result in loss of data (see Figure 3·7), so a 6-bit counter will be required to track the magnitude of destructiveness and perform refresh operation. Although the counter-based refresh approach seems more beneficial in multi-bit TiO_2 RRAM compared to the read followed by refresh scheme, our analysis shows that the energy and area overhead of the counter-based approach makes it infeasible.

Considering the rate of change of state for HfO_x RRAM in Equation (2.7), the number of consecutive non-destructive read operations in multi-bit HfO_x-based 1T1R RRAM cell will be

$$t_{ref-HfO_x} = \frac{\phi_{min}(x_{max} - x_{min})}{2^{n+1} T_R C \sqrt{(1 - x/C)^3} \frac{d\phi}{dt}}. \tag{3.3}$$

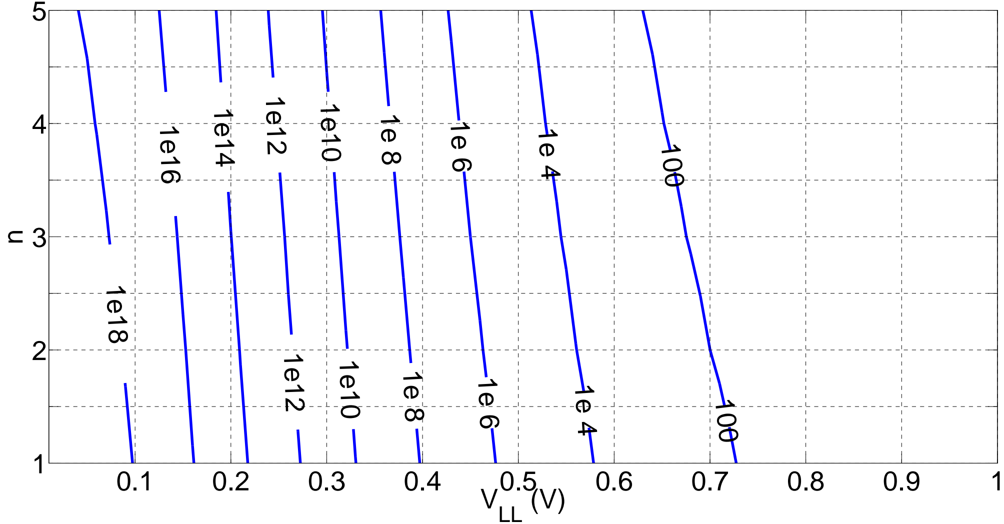

Figure 3·8: Contour plots of the number of consecutive non-destructive read cycles in multi-bit HfO_x-based RRAM for different n and V_{LL} values ($x = 0.9$).

The corresponding contour plots of the number of consecutive non-destructive read operations for different n and V_{LL} values for a memristor with initial state of $x = 0.9$ is shown in Figure 3·8. The threshold-based conductive filament growth mechanism in HfO_x memristor makes it more resilient to read destructiveness compared to ion-drift mechanism based TiO_2 memristors. As can be seen in Figure 3·8, for small read voltage values a large number of consecutive read operations are required to destruct the current state in multi-bit HfO_x RRAM technology. The refresh threshold proposed in Equation (3.3) and shown in Figure 3·8 exceeds the maximum allowed number of accesses (endurance) in the HfO_x-based RRAMs reported in (Sheu et al., 2009) (see Table 1.1) and (Chiu et al., 2012) which practically makes HfO_x a non-destructive memristor technology at small read voltages. In case large voltages are used for $read_{sub}$ operation, then we might observe destructiveness of memristor state. To combat this, we propose to use a counter that tracks the current state of the memristor as well as the number of consecutive read operations. A refresh operation is done once the number of read operations exceeds the threshold given by Equation

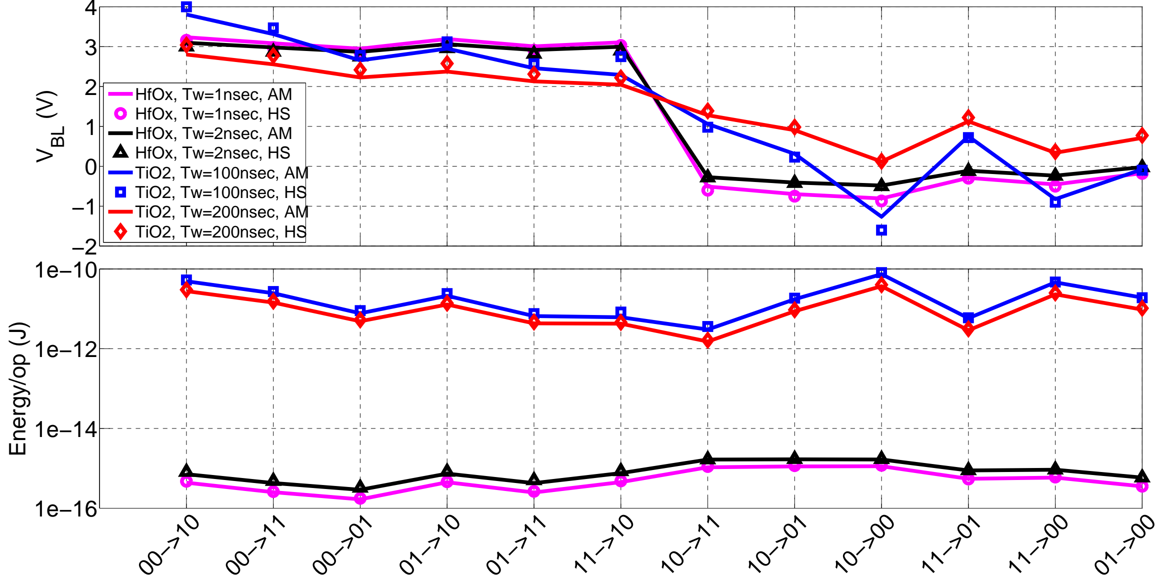

Figure 3·9: Comparison between analytical model (AM) and HSPICE simulations (HS) for bitline voltage and energy dissipation in different TiO_2-based and HfO_x-based 2-bit RRAM $write/refresh$ operation. The V_{LL} voltage is 1.5 V for all transitions. For bitline voltage, the average error is 9.81% for TiO_2-based cell and 5.19% for HfO_x-based cell, while for energy dissipation the average error is 8.71% for TiO_2-based cell and 5.25% for HfO_x-based cell.

(3.3).

The equivalent circuit model for the refresh$_{sub}$/write$_{sub}$ operation of a 1T1R RRAM cell is shown in Figure 3·3. For the TiO_2-based memristor the refresh$_{sub}$/write$_{sub}$ operation model uses the window function proposed in (Biolek et al., 2009). The switching time of the bitline capacitor and the junction capacitor are orders of magnitude lower than the switching time of the memristor. Hence, we do not consider these two capacitors in our analytical models. Given the threshold voltage (V_{th}) drop across the access transistor (i.e. R_{ch}), the expression for memristor current during refresh$_{sub}$/write$_{sub}$ operation is

$$i_w(t) = \frac{V_{BL} - V_{th} - V_{LL}}{R_m(t)}. \tag{3.4}$$

Using the window function in Equation (2.3) and the rate of change of state in Equation (2.2), the refresh$_{sub}$/write$_{sub}$ time can be approximated as

$$T_W = \frac{R_{OFF} Q_i}{(V_{BL} - V_{LL} - V_{th})\gamma} \tag{3.5}$$

where $\gamma = \frac{\mu_v R_{ON}}{L^2}$. Here $Q_i = \int_{x_i}^{x_{i+1}} \frac{1-x}{1-x^4}\, dx$ is the nonlinear delay integral for transitions to higher memristor states where x_i is the state of memristor and $Q_i = \int_{x_{i+1}}^{x_i} \frac{1-x}{1-(x-1)^4}\, dx$ is the nonlinear delay integral for transitions to lower memristor states (note that here Q_i could be negative leading to a negative voltage across the memristor for transitions to lower states). Here the resistance of the memristor is approximated as $R_m(t) \approx R_{OFF}(1 - x(t))$ for simplicity. The integrals are determined from the window function we considered previously to model the nonlinearity of the memristor at the boundaries in equation (2.3) with $p = 2$. For the n-bit RRAM cell, the limits of the nonlinear delay integral Q_i will change based on 2^n different states. As an example, for the 2-bit cell we compared the required bitline voltages for 12 possible write$_{sub}$ transitions for 100 ns and 200 ns time period in TiO_2-based 1T1R memory cells in Figure 3·9. The V_{LL} voltage is maintained at 1.5 V for all transitions. The average error between the analytical model and the HSPICE simulation results for a 2-bit TiO_2-based 1T1R memory cell is 9.81%.

For the HfO_x-based memristor using the rate of change of state in (2.7), the set/reset time of the 1T1R RRAM cell can be modeled as

$$T_W = \frac{\phi_{min}}{2C}\left(\frac{d\phi}{dt}\right)^{-1} U_i \tag{3.6}$$

where $U_i = \int_{x_i}^{x_{i+1}} \frac{dx}{\sqrt{(1-x/C)^3}}$ is the nonlinear delay integral for HfO_x-based memristors

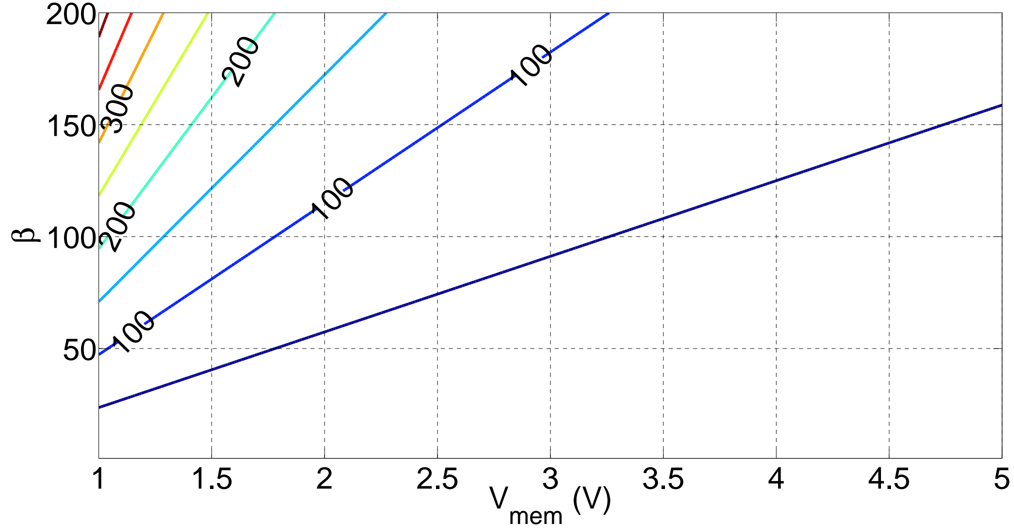

Figure 3·10: Contour plots for set time (nsec) in the 2 bits/cell TiO_2-based RRAM.

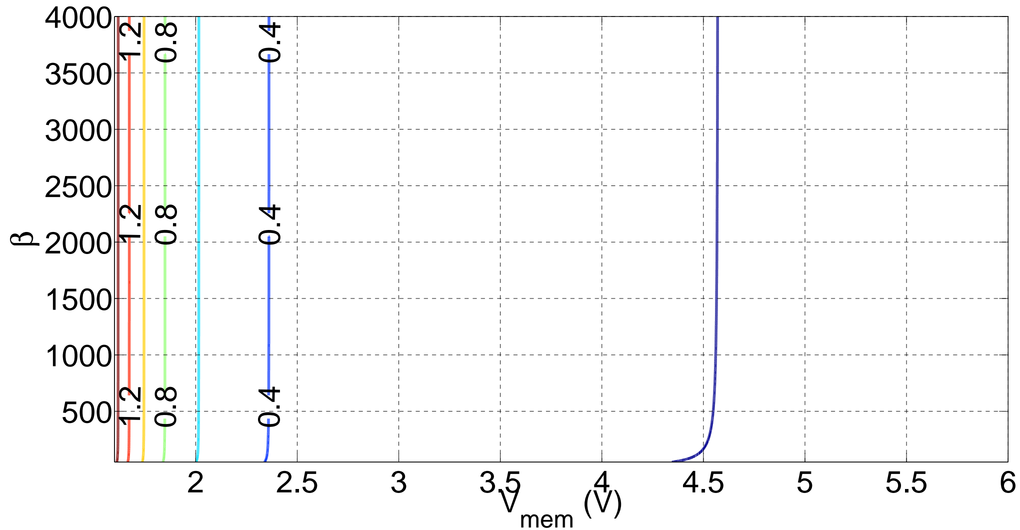

Figure 3·11: Contour plots for set time (nsec) in the 2 bits/cell HfO_x-based RRAM.

for transitions to higher states and $U_i = \int_{x_i+1}^{x_i} \frac{dx}{\sqrt{(1-x/C)^3}}$ is the nonlinear delay integral for HfO_x-based memristors for transitions to lower states. For the n-bit RRAM cell, the limits of the nonlinear delay integral U_i will change based on 2^n different states. Similar to the TiO_2-based memristor, there is a threshold voltage drop across the access transistor for set operation. The HfO_x cell write access time in (3.6) does not

include the 0%-90% distributed RC-line transition time for bitline ($R_{BL}C_{BL}$) which will later be included in the whole RRAM array design specification. Comparing results from the analytical model and the HSPICE simulation for 1 ns and 2 ns time period for a 2-bit HfO_x-based 1T1R memory cell in Figure 3·9, the average error is 5.19%. The modeling error for HfO_x-based cell is different from the TiO_2-based cell due to the different electrical parameters for each type of cell (see Table 2.1).

The contour plots for the set time constraints of 2 bits/cell TiO_2-based and HfO_x-based RRAM is shown in Figure 3·10 and Figure 3·11. Write speed is limited by the voltage applied across the memristor (V_{mem}). The write operation of HfO_x-based memristor is faster compared to TiO_2-based due to the faster rate of change of state of HfO_x memristors.

3.6 Energy Models

In this section, we present the models for energy consumption during read$_{sub}$ and write$_{sub}$/ refresh$_{sub}$ operation. It should be noted that the energy consumed in the wordline, bitline and loadline depends on the aspect ratio of the memory array. Once the array structure is finalized the energy can be determined based on bitline capacitance (C_{BL}), loadline capacitance (C_{LL}) and wordline capacitance (C_{WL}). The energy dissipated in the cell during read$_{sub}$ operation (for both TiO_2 and HfO_x) can be expressed as

$$E_R = \int_0^{T_R} V_{LL} i_R(t)\, dt \tag{3.7}$$

where $i_R(t)$ is the memristor current during the read$_{sub}$ operation. Using the RC circuit model in Figure 3·3, the energy dissipated in the n-bit RRAM cell at the end

Cell Data	AM TiO_2	HS TiO_2	AM HfO_x	HS HfO_x
00	12.03 fJ	12.84 fJ	11.51 fJ	11.50 fJ
01	14.40 fJ	15.58 fJ	15.03 fJ	15.02 fJ
11	17.91 fJ	19.68 fJ	21.65 fJ	21.65 fJ
10	23.57 fJ	26.48 fJ	38.47 fJ	38.47 fJ

Table 3.3: Comparison between analytical model (AM) and HSPICE simulations (HS) for energy dissipated in the cell while reading 2-bit RRAM cell with a read access time of $T_R(TiO_2) = 1nsec$ and $T_R(HfO_x) = 200nsec$. The average error is 8.44% and 0.038% for TiO_2 and HfO_x respectively.

of read$_{sub}$ operation will be

$$E_R = C_{BL}V_{LL}^2(1 - e^{\frac{-T_R}{(R_m(t)+R_{ch}+0.5R_{BL})C_{BL}}}). \tag{3.8}$$

Table 3.3 compares the energy dissipation calculated from the analytical model and determined using HSPICE simulation during read$_{sub}$ operation of a 2-bit TiO_2-based RRAM cell having a latency of 1 $nsec$ as well as a 2-bit HfO_x-based RRAM cell having a latency of 200 $nsec$ for different stored data values. The average error is 8.44% and 0.038% for TiO_2 and HfO_x respectively.

The read energy contour plots for different number of bits/cell for both TiO_2- and HfO_x-based RRAMs are illustrated in Figure 3·12 and Figure 3·13. For each value of bits/cell and each read timing constraint, we find the V_{LL} value that gives at least 25 mV difference between two adjacent reference voltages of the sense amplifiers for reliable read operation. The difference between the reference voltages of the sense amplifiers is determined by the offset voltage of the input transistors in the voltage sense amplifiers and could be further reduced by increasing area at the expense of power (Schinkel et al., 2007). Higher number of bits/cell requires larger drive voltages to increase read noise margin and therefore consumes more energy during

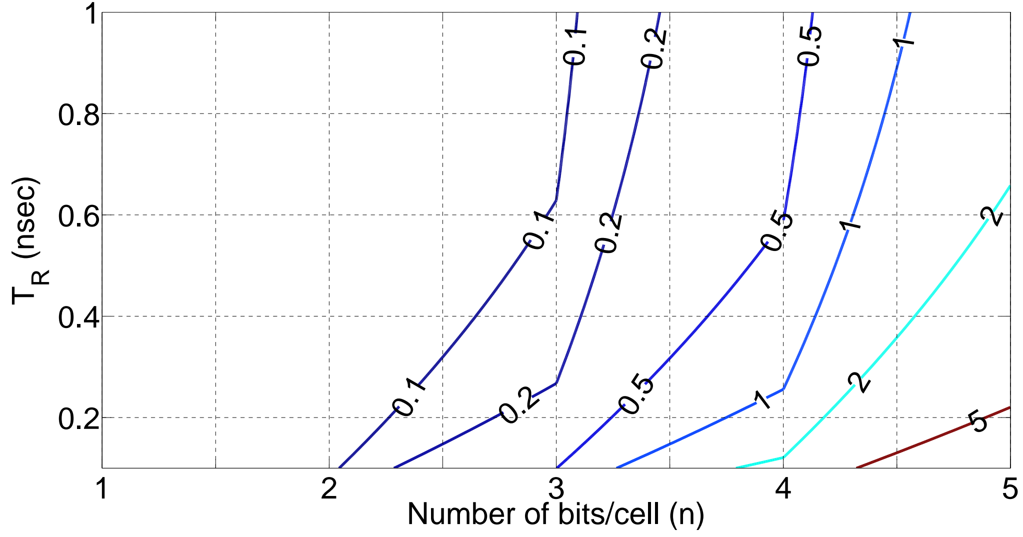

Figure 3·12: Contour plots for average read energy (pJ) in multi-bit TiO_2 RRAMs. We maintain at least 25 mV difference between adjacent reference voltages for reliable read operation.

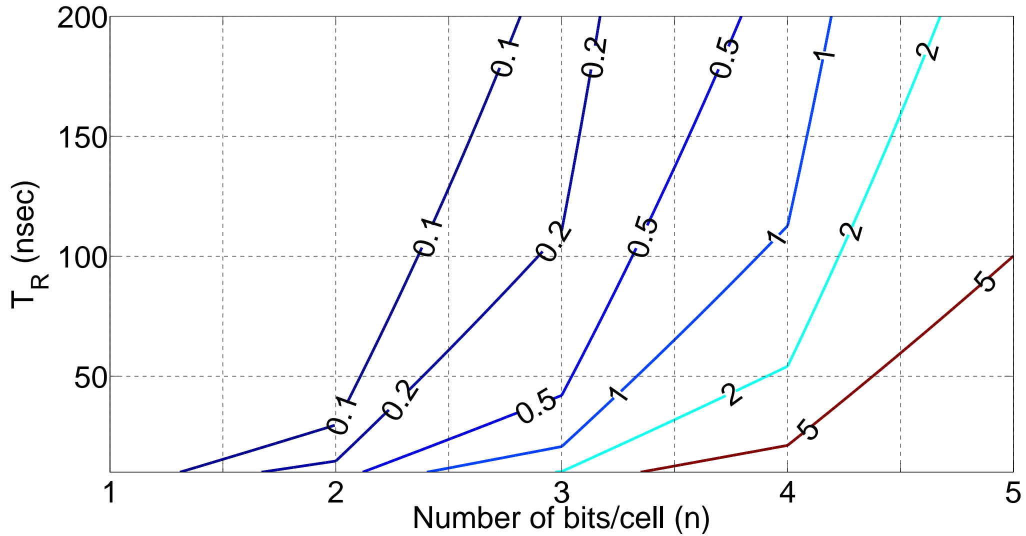

Figure 3·13: Contour plots for average read energy (pJ) in multi-bit HfO_x RRAMs. We maintain at least 25 mV difference between adjacent reference voltages for reliable read operation.

read operation. Larger read times require lower drive voltages and dissipate lower amount of energy.

The instantaneous current of the memristor while performing refresh$_{sub}$/write$_{sub}$ op-

eration in the TiO_2-based cell is determined by Equation (3.4). Considering the V_{th} voltage drop across the access transistor, the energy dissipated in the cell during refresh$_{sub}$/write$_{sub}$ operation can be calculated as

$$E_W = \int_0^{T_W} (V_{BL} - V_{th} - V_{LL}) i_W(t)\, dt = \frac{(V_{BL} - V_{th} - V_{LL}) B_i}{\gamma} \tag{3.9}$$

where $\int i_W(t)dt = P_i \ / \ \gamma$ and $B_i = \int_{x_i}^{x_{i+1}} \frac{dx}{1-x^4}$ is the nonlinear energy integral for transitions to higher memristor states and $P_i = \int_{x_{i+1}}^{x_i} \frac{dx}{1-(x-1)^4}$ is the nonlinear energy integral for transitions to lower memristor states. The dissipated energy in the diffusion capacitor of the access transistor is ignored since it's much smaller than the overall cell energy. For the n-bit RRAM cell, the limits of the nonlinear energy integral P_i will change based on 2^n different states.

Figure 3·9 compares the energy dissipated in a 2-bit 1T1R cell for write$_{sub}$ in 12 possible transitions calculated using the analytical model and the HSPICE simulation for TiO_2-based configurations with transition time of $T_W = 100nsec$ and $200nsec$. The average error is 8.71%.

The write$_{sub}$/refresh$_{sub}$ energy in the HfO_x-based memristor is modeled as

$$E_W = \int_0^{T_W} V^2 / R(t)\, dt \tag{3.10}$$

where V is the voltage across the memristor. Here using $R(t) = (1 - x(t)/C) R_{OFF}$, the closed form expression for write$_{sub}$/refresh$_{sub}$ energy in n-bit 1T1R HfO_x-based cell is

$$E_W = \frac{V^2 \phi_{min} (\frac{d\phi}{dt})^{-1}}{2CR_{OFF}} S_i \tag{3.11}$$

Component	Transition Time
Wordline	1.3 nsec
Loadline	1.3 nsec
ADC	1 nsec
Mux-based DAC	1 nsec

Table 3.4: Transition times of different components in the multi-bit RRAM array.

where $S_i = \int_{x_i}^{x_{i+1}} \frac{dx}{\sqrt{(1-x/C)^5}}$ is the nonlinear energy integral for HfO_x-based memristors. Since there is a threshold voltage drop across the access transistor, the write voltage (V) in (3.11) is chosen as one threshold voltage below the difference between V_{BL} and V_{LL} voltages. In the n-bit RRAM cell, the limits of the nonlinear energy integral S_i will change based on 2^n different states. The average error between the dissipated energy of a 2-bit HfO_x RRAM cell model and the simulation results is 5.25% (see Figure 3·9). We do not consider the effect of subthreshold leakage in our energy analysis since all transistors are working in strong-inversion region of operation.

Using the energy models, we compare the different energy components of the 1T1R RRAM array for different number of bits/cell. The transition times of different components (other than the cell) in the RRAM array have been assumed constant for different number of bits and are summarized in Table 3.4. The energy consumption in different components of the RRAM array during read operation for TiO_2-based RRAMs is illustrated in Figure 3·14. Cell energy increases during read operation for higher number of bits. This is due to higher loadline voltages required for providing sufficient read noise margin for higher number of bits/cell. Since the read process of multi-bit TiO_2 RRAM is destructive (see Figure 3·7), we consider the energy of read followed by a refresh operation in Figure 3·14. The total wordline energy is constant across all cells. The number of sense amplifiers increases with number of bits/cell ($2^n - 1$ sense amplifiers for n-bit RRAM cell), and so the energy/bit of the

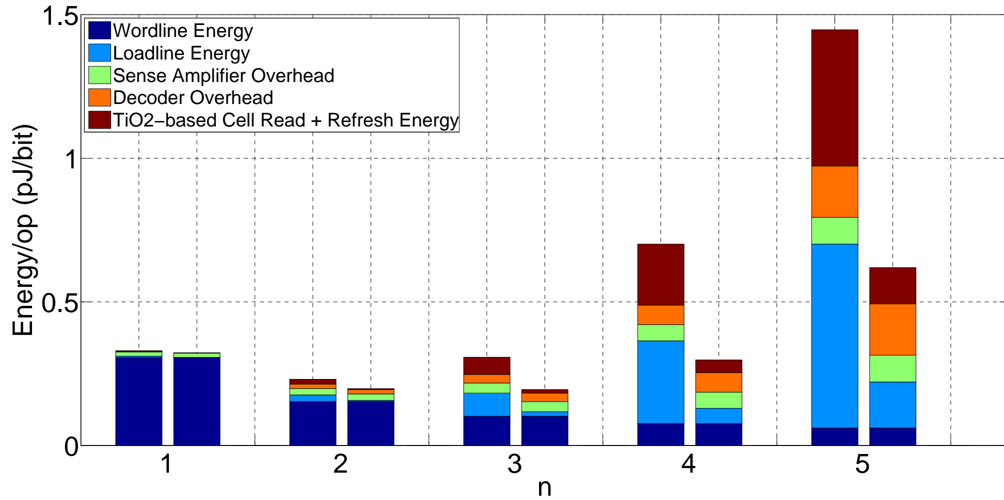

Figure 3·14: Energy dissipated in different components of the multi-bit TiO_2-based RRAM array in read operation for uniform (left) and non-uniform (right) state assignments (T_R=1nsec).

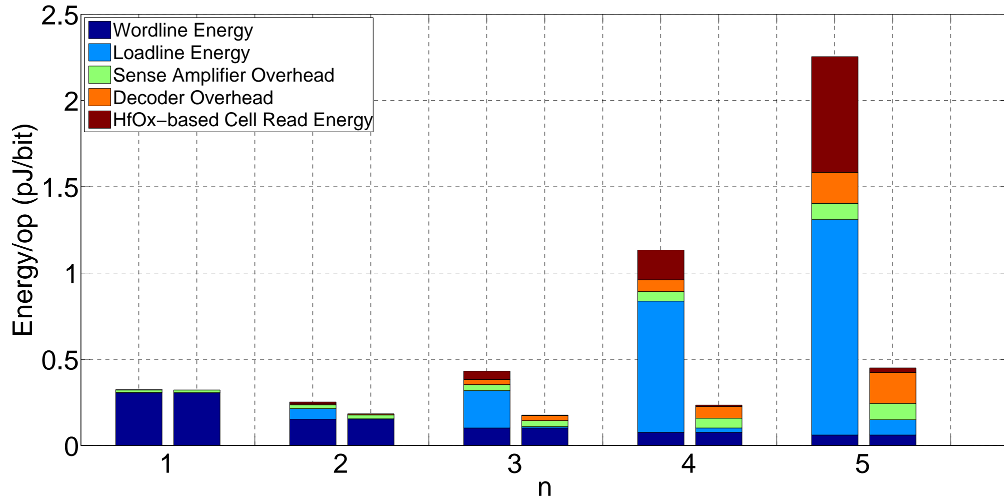

Figure 3·15: Energy dissipated in different components of the multi-bit HfO_x-based RRAM array in read operation considering uniform (left) and non-uniform (right) state distributions (T_R=200nsec).

sense amplifiers increases. The same trend is observed for the decoder energy as the number of multiplexers increases with number of bits/cell.

To increase the read reliability of multi-bit RRAM array, we assume there should be at

least 25 mV difference between two adjacent reference voltages. One way to reach this voltage difference is to use uniform state assignment and increase the V_{LL} voltage. In the uniform state assignment scheme, there is a fixed distance between two adjacent states. Another way of reaching the 25 mV difference between two adjacent reference voltages is by lowering V_{LL} voltages, and choosing the appropriate memristor states such that the read reliability would be maximized. This approach is called non-uniform state assignment where the 0.1 to 0.9 range for the state of a memristor is not uniformly shared between the 2^n different data that can be stored in the cell. Comparing uniform and non-uniform state assignment strategies, the non-uniform state assignment consumes lower energy due to lower V_{LL} values. The minimum total read energy/operation is consumed at n=2 for uniform state assignment and n=3 for non-uniform state assignment. Considering the same throughput constraint (# bits/cell n=3) for both cases, non-uniform state assignment consumes 32.1% less energy than uniform state assignment.

Using the same approach, we show the energy consumption in different components of the RRAM array during read operation for HfO_x-based RRAMs using uniform and non-uniform state assignments in Figure 3·15. The refresh energy of the multi-bit HfO_x memristor is amortized across the different components of the array. Compared to TiO_2 and considering the same throughput constraint (n=3) the HfO_x RRAM array using non-uniform state assignment has 59.07% lower total read energy consumption.

The energy consumed in the various components of the RRAM array during write operation for both TiO_2- and HfO_x-based RRAMs are shown in Figure 3·16 and Figure 3·17. Since the size of mux-based DAC increases with number of bits per RRAM cell, the energy consumption of the DAC increases accordingly. The total wordline and loadline energy is constant across all cells. We determine the cell energy

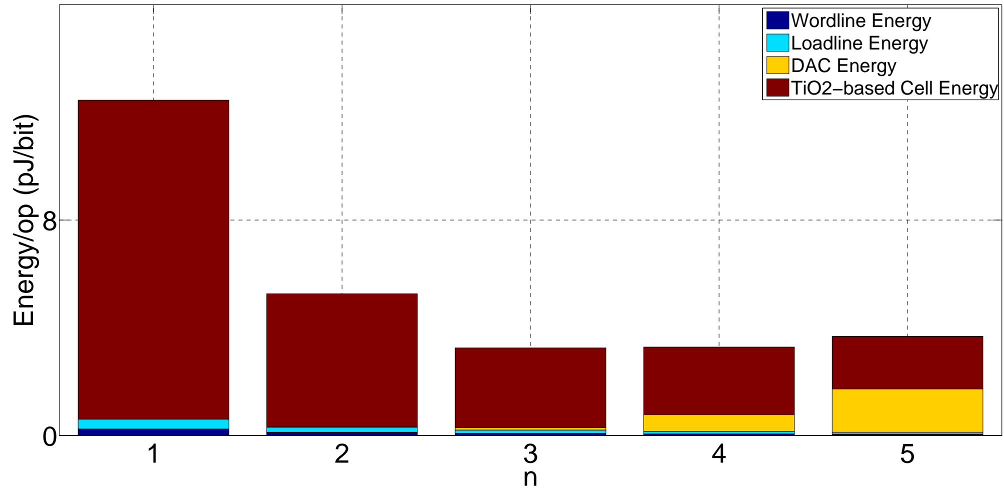

Figure 3·16: Energy dissipated in different components of the TiO_2-based RRAM array in write operation (T_W=100nsec).

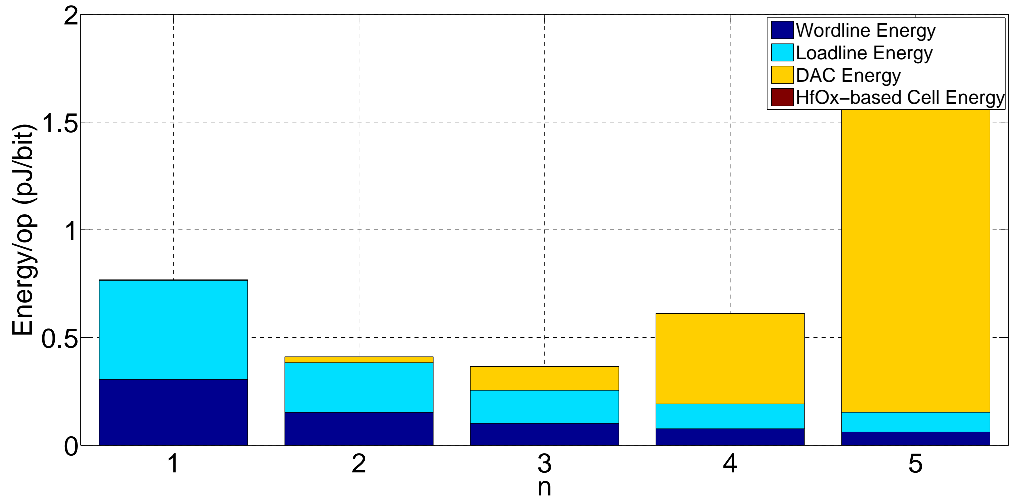

Figure 3·17: Energy dissipated in different components of the HfO_x-based RRAM array in write operation (T_W=1nsec).

by using the average energy value of all possible transitions for the n-bit cell. The TiO_2 cell energy dominates the energy dissipated in all the array components due to large set/reset time and lower resistance values for TiO_2 RRAM, while the HfO_x cell energy is much smaller than the energy in the remaining array components. The minimum total write energy/operation is consumed at n=3 for both cases.

3.7 Memory Technology Comparison

In this section, we compare the performance and energy of the designed multi-bit array with other types of memory technologies. Figures 3·18 and 3·19 show a comparison of the read and write time vs. energy of different state-of-the-art memory technologies with the designed multi-bit TiO_2-/HfO_x-based RRAM array. For TiO_2 and HfO_x we consider the minimum energy points corresponding to different number of bits/cell in uniform state assignment scheme. The optimized multi-bit RRAM array designed in this work has lower energy consumption compared to other emerging nonvolatile memory technologies such as MRAM, FeRAM and PCRAM based on recently-published measurement data. The write access time of HfO_x RRAM is small compared to other types of nonvolatile memory technologies but the TiO_2 RRAM write time is large. On the other hand, the read access time of HfO_x RRAM is large compared to other types of nonvolatile memory technologies but the TiO_2 RRAM read time is small. The lower energy and access time of the optimized multi-bit RRAM array makes it a promising replacement for CMOS-based nonvolatile memory technologies.

3.8 PVT Variation Analysis of n-bit RRAM Cell

As was mentioned in section 2.1, OTF and LER cause variations in memristor geometry (Niu et al., 2010b), (Hu et al., 2011b), (Hu et al., 2011a) and RDD causes randomness in resistivity which directly impacts the performance and energy dissipation of RRAM cells. In this section, we apply the Monte-Carlo methodology (Hu et al., 2011a) to our models for both TiO_2-based and HfO_x-based memristors to analyze the influence of OTF, LER and RDD on the performance and energy of the n-bit 1T1R RRAM cell. For our analysis, we exclude the variations in the energy and

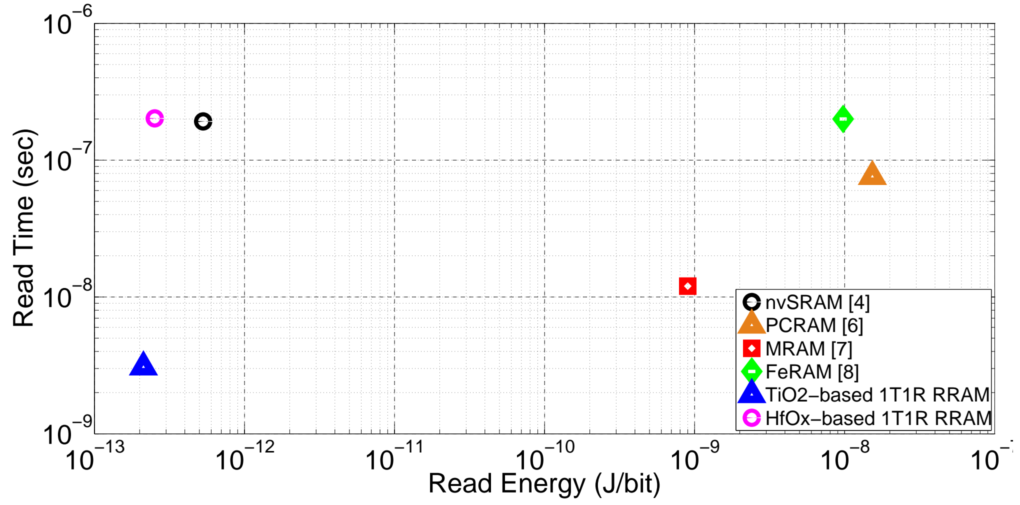

Figure 3·18: Comparison of read time/energy between different memory technologies.

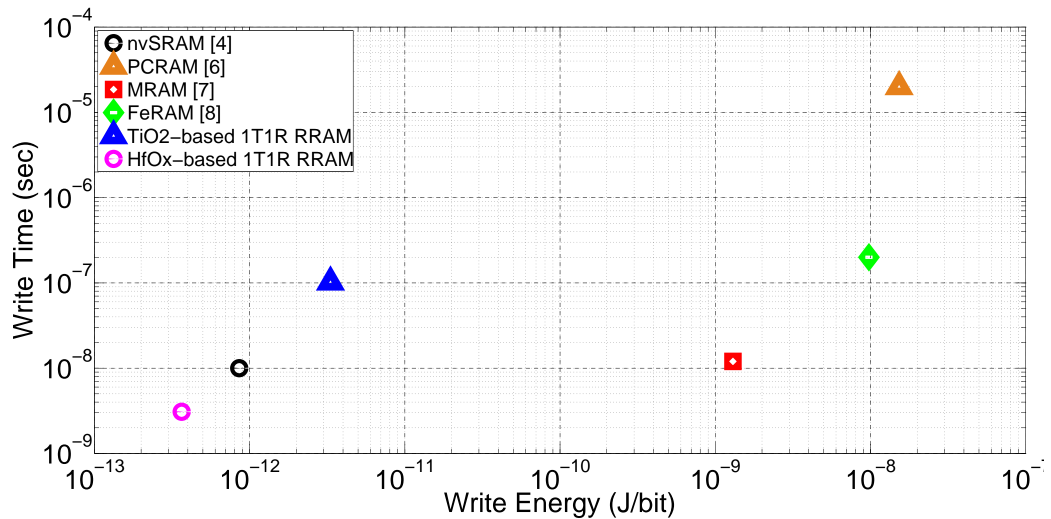

Figure 3·19: Comparison of write time/energy between different memory technologies.

performance of the CMOS devices due to PVT variations to isolate and quantify the true impact of PVT variations on the memristors device functionality and the cell as a whole.

The LER of the memristor has been modeled as a combination of the low and high frequency domain disturbances in (Hu et al., 2011a), and (Wang et al., 2009), and is

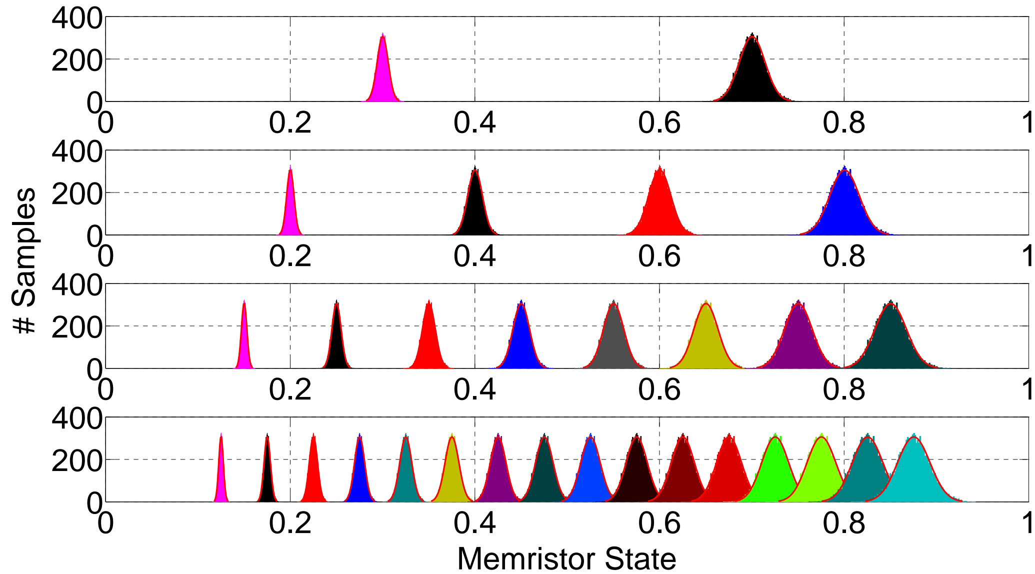

Figure 3·20: Uniform state distribution of the multi-bit TiO_2-based memristor caused by OTF. The memristor state distribution for each number of bits/cell is such that maximum process noise margin would be achieved.

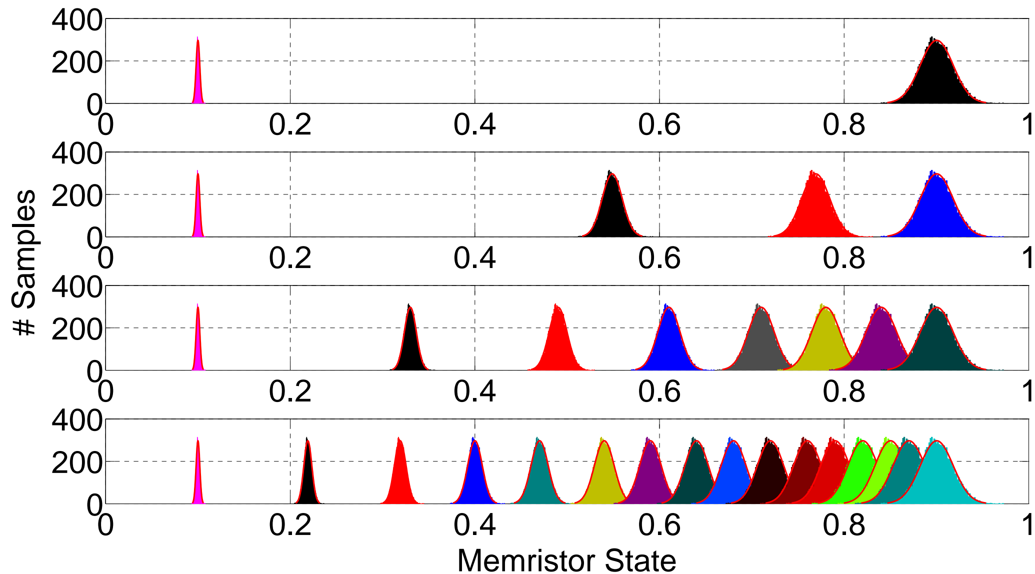

Figure 3·21: Non-uniform state distribution of the multi-bit TiO_2-based memristor caused by OTF. The memristor state distribution for each number of bits/cell is such that maximum read noise margin would be achieved.

given by

$$LER = L_{LF}.sin(f_{max}.r) + L_{HF}.z \tag{3.12}$$

where the sinusoid function with the amplitude of L_{LF} describes the low frequency domain variations. Here $f_{max} = 1.8\ MHz$ is the mean of the low frequency range with a uniform distribution represented as $r \in U(-1, 1)$. L_{HF} accounts for the high frequency variations and z is considered to have a normal distribution function as $N(0, 1)$. The effect of OTF is usually modeled as a Gaussian distribution with a $\sigma = 2\%$ deviation from the nominal memristor thickness (Hu et al., 2011a), (Niu et al., 2010b). Also RDD has been modeled as having a Gaussian distribution with $\sigma = 2\%$ (Hu et al., 2011b) in the resistivity term in both ionic drift and filament growth models for TiO_2 and HfO_x-based RRAMs.

Considering the nominal parameters in Table 2.1, we explore the effect of OTF, LER, and RDD on the states of both TiO_2-based and HfO_x-based RRAMs. The state definition for ionic drift-based TiO_2 RRAM model is only a function of the ratio of the doped region to memristor thickness. The movement of dopants along the memristor thickness defines memristance (see Figure 2·1a). Therefore, the state assignment will only be affected by OTF. In other words, LER and RDD will not change the state assignment of TiO_2-based RRAMs according to ionic drift memristor model. The impact of OTF on TiO_2-based RRAM with uniform and non-uniform state assignments for different number of stored bits ($1 \leq n \leq 4$) for 10000 samples are illustrated in Figures 3·20 and 3·21. The multi-bit TiO_2-based 1T1R RRAM cell is resilient to OTF-based process variations up to n=3 for uniform state assignment and up to n=2 for non-uniform state assignment, where no overlap is observed between adjacent states.

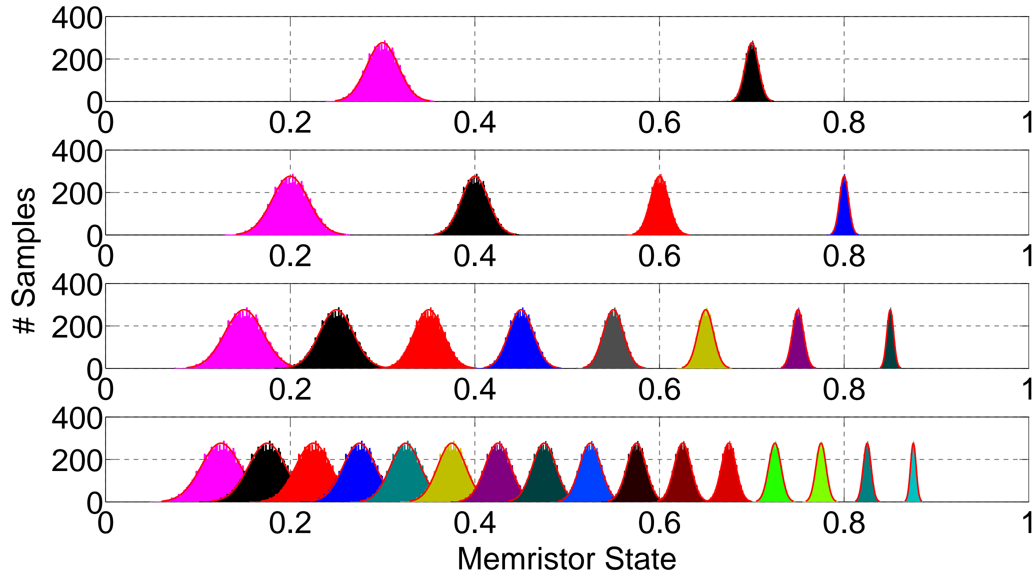

Figure 3·22: Uniform state distribution of the multi-bit HfO_x-based memristor caused by LER. The memristor state distribution for each number of bits/cell is such that maximum process noise margin would be achieved.

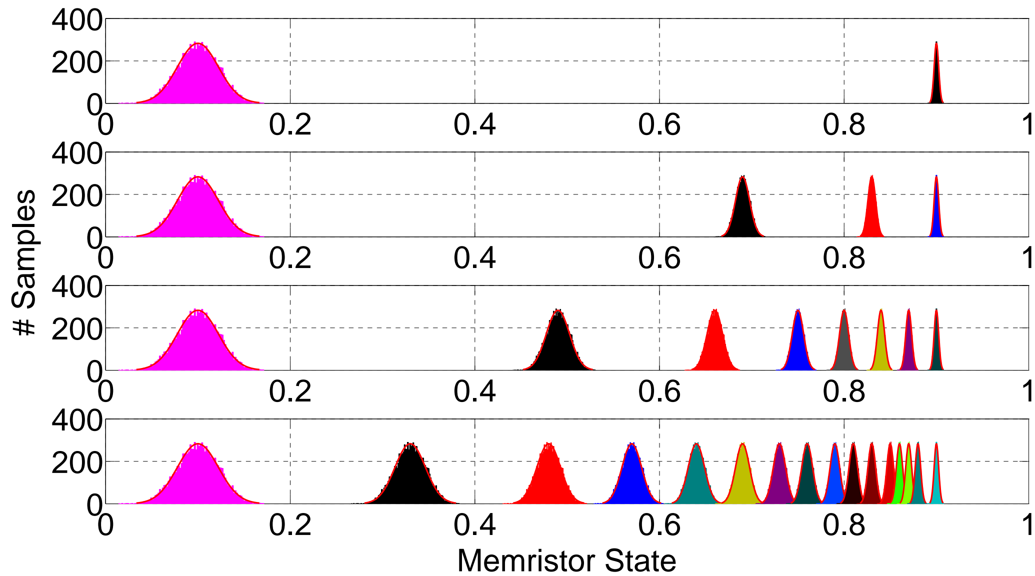

Figure 3·23: Non-uniform state distribution of the multi-bit HfO_x-based memristor caused by LER. The memristor state distribution for each number of bits/cell is such that maximum read noise margin would be achieved.

Parameter	LER TiO_2	OTF TiO_2	RDD TiO_2	LER HfO_x	OTF HfO_x	RDD HfO_x
$x(t)$	0%	6.01%	0%	12.58%	0%	0%
WT	0%	17.79%	8.52%	7.36%	0%	0.95%
WE	7.28%	12.03%	5.99%	21.96%	5.93%	5.03%
RE	3.78%	3.06%	3.01%	6.32%	5.34%	5.31%
RD	5.47%	7.50%	6.54%	3.65%	0%	63.25%

Table 3.5: $3\sigma/\mu$ of the 3-bit TiO_2-based and HfO_x-based 1T1R cell specification variations due to LER, OTF and RDD. Here WT = Write time, WE = Write energy, RE = Read energy and RD = Read destructiveness.

The state definition for filament growth-based HfO_x RRAM model is only a function of filament diameter. Therefore, the state assignment will only be affected by LER. OTF and RDD will not change the state assignment of HfO_x-based RRAMs. The uniform and non-uniform state distributions of the HfO_x-based RRAM for different number of stored bits ($1 \leq n \leq 4$) are illustrated in Figure 3·22 and 3·23. The multi-bit HfO_x-based 1T1R RRAM cell is resilient to LER-based process variations up to n=3 where no overlap is observed between adjacent states.

Table 3.5 summarizes the effect of LER, OTF and RDD on the state assignment, write time, write energy, read energy and read destructiveness of the 3-bit TiO_2-based and HfO_x-based 1T1R cells. As discussed earlier, the TiO_2 memristor state is only affected by OTF whereas the HfO_x memristor state is only affected by LER. The impact of LER, OTF and RDD is quantified as $(3\sigma/\mu)\times 100\%$ value of each parameter. OTF has higher impact on the TiO_2 specifications compared to LER. Also OTF has the highest impact on the write time variations for the multi-bit TiO_2 memristor since the TiO_2 set/reset time is a quadratic function of memristor thickness based on (3.5). Similarly, the effect of OTF on the write energy and read destructiveness of the TiO_2 RRAM is higher than LER. The variation in read destructiveness changes the refresh threshold which affects the reliability of read operation. OTF and LER

Parameter	$3\sigma = 6\%$ TiO_2	$3\sigma = 10\%$ TiO_2	$3\sigma = 6\%$ HfO_x	$3\sigma = 10\%$ HfO_x
WT	5.92%	8.68%	8.10%	13.75%
WE	5.99%	9.01%	6.84%	11.55%
RE	12.08%	17.94%	11.92%	17.93%
RD	5.91%	9%	146.5%	229.98%

Table 3.6: $3\sigma/\mu$ of the 3-bit TiO_2-based and HfO_x-based 1T1R cell specifications due to voltage variations for ($3\sigma_{V_{ref}} = 6\%$) and ($3\sigma_{V_{ref}} = 10\%$).

Parameter	$\Delta T = 10$ TiO_2	$\Delta T = 30$ TiO_2	$\Delta T = 10$ HfO_x	$\Delta T = 30$ HfO_x
WT	4.47%	13.59%	7.56%	23.31%
WE	4.52%	13.75%	7.88%	23.40%
RE	0.07%	0.23%	24.79%	97.7%
RD @ $V_{LL} = 0.4V$	6.48%	19.22%	129.53%	427.75%

Table 3.7: $3\sigma/\mu$ of the 3-bit TiO_2-based and HfO_x-based 1T1R cell specifications due to temperature variations (ΔT).

have similar impact on read energy since it is mostly dominated by bitline variations according to (3.8). It should be noted that OTF has a minimal impact on the write time and read destructiveness of the HfO_x-based 1T1R cell as these two parameters are independent of the oxide thickness (see equations (2.7) and (3.6)). LER has the highest impact on the write energy variations of the multi-bit HfO_x memristor due to its sensitivity to filament diameter fluctuations based on (3.11). The rate of change of diameter in the filament growth model has higher sensitivity to RDD at lower voltages. In other words, high set/reset voltages limit the effect of RDD in write time variations of HfO_x-based RRAMs. However, read destructiveness significantly changes with RDD since the applied read voltages are considerably low compared to write (set/reset) voltages which deteriorates the read reliability of the HfO_x-based RRAMs.

The power supply noise in VLSI chips causes variations in the supply voltage applied to the various transistors in a circuit, which in turn causes variations in performance

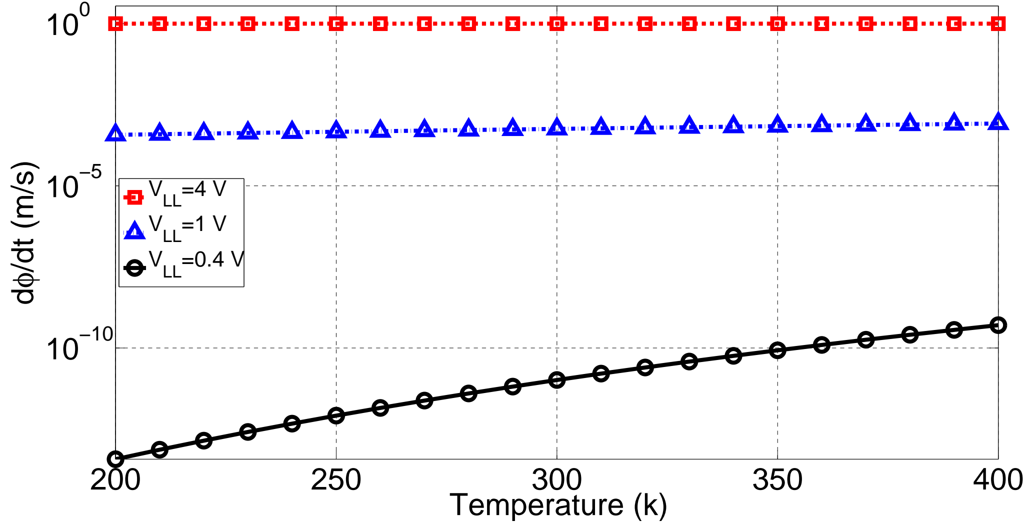

Figure 3·24: Diameter change of HfO_x-based memristors as a function of temperature for different applied voltages in filament growth model. Diameter shows higher variation with temperature at lower loadline voltages.

and energy dissipation. Table 3.6 summarizes the impact of voltage variations on write time, write energy, read energy and read destructiveness of a 3-bit RRAM cell. Without loss of generality, we explore two cases where each voltage reference has been assumed to have a Gaussian distribution with $3\sigma = 6\%$ and $3\sigma = 10\%$ of the nominal value. We calculate the write time and energy variations considering 56 possible transitions for the 3-bit 1T1R RRAM cell. The write time and write energy of the HfO_x RRAM has more variations compared to TiO_2 since these two parameters are exponential functions of applied voltage in HfO_x RRAM according to (3.6) and (3.11). Comparing the rate of state change in equations (2.2) and (2.7), the destructiveness of the HfO_x-based memristor state is considerably more sensitive to voltage fluctuations. This will significantly affect the refresh threshold in Equation (3.3) (see Table 3.6). The read energy has similar amount of variations due to voltage fluctuations for both materials according to (3.8).

We also analyzed the impact of temperature variations on performance and energy

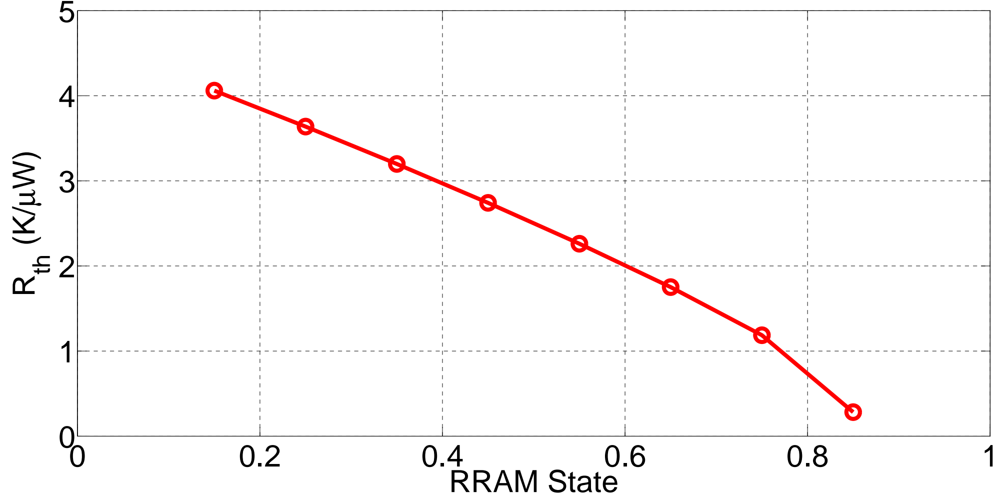

Figure 3·25: Effective thermal resistance of a 3-bit TiO_2-based RRAM as a function of memristor state.

metrics of both TiO_2-based and HfO_x-based memristors in the 3-bit RRAM cell. The temperature dependency of the ionic drift model has been modeled in (Strukov and Williams, 2011) where thermal resistance of the filament, defined as the ratio between the maximum temperature increase in the filament and the dissipated electrical power (Russo et al., 2009), for the state 1 (R_{ON}) and state 0 (R_{OFF}) in the TiO_2 filament are derived as:

$$R_{th}(R_{ON}) = L/(8k_M A_{CF}) \tag{3.13}$$

$$R_{th}(R_{OFF}) \approx (2ArcSinh[L/(\sqrt{A_{CF}})] - 1.5)/(4k_I L). \tag{3.14}$$

Here, $k_M = 30W/mK$ and $k_I = 3W/mK$ (Strukov and Williams, 2011) are the thermal conductances of the metal and insulator corresponding to titanium oxide thin films with oxygen vacancies conductive channels and A_{CF} is the filament area. The change in resistance of the RRAM based on ionic drift model follows $\Delta R_{R_{OFF},R_{ON}} \propto \Delta T/(R_{th}I^2)$ where I is the RRAM current.

Table 3.7 summarizes the impact of temperature variations on write time, write energy, read energy and destructiveness of both TiO_2-based and HfO_x-based memristors in the 3-bit RRAM cell. We explore two cases with nominal ambient temperature and variations of $\Delta T = 10K$ and $\Delta T = 30K$. Temperature variations have a larger impact on the read destructiveness of the HfO_x-based memristor. The rate of change of diameter in HfO_x-based RRAMs due to temperature variations increases at lower applied voltages based on filament growth model in (2.4) (see Figure 3·24). The variations in write time and write energy of HfO_x RRAM is higher than TiO_2 due to the exponential temperature term in these metrics for HfO_x RRAM. The effect of temperature variation on the intermediate states of the multi-bit TiO_2 RRAM can be analyzed using the effective thermal resistance as $R_{th} = R_{th}(R_{ON})||R_{th}(R_{OFF})$ (Strukov and Williams, 2011) where the corresponding cross-section area for each state is plugged into the two thermal resistance expressions in (3.13) and (3.14). The effective thermal resistance of a 3-bit TiO_2-based RRAM is illustrated in Figure 3·25 for different memristor states. Temperature variations have minimal effect on the read energy fluctuations of TiO_2 RRAM since it is mostly affected by bitline resistance (according to (3.8)). This is however not the case for the HfO_x RRAM since its typical R_{OFF} value is orders of magnitude larger than the bitline resistance according to Table 2.1. This will dominate the effect of temperature variations in HfO_x RRAM read energy fluctuations with respect to bitline parasitic variations.

3.9 Summary

In this chapter, we presented the design and optimization of a n-bit 1T1R RRAM array designed using TiO_2- and HfO_x-based memristors. We first presented models for performance and energy of read and write operation in n-bit 1T1R RRAM cells designed using TiO_2- and HfO_x-based memristors. We validated our performance

and energy models against HSPICE simulations, and the difference is less than 10% for both n-bit TiO_2- and HfO_x-based 1T1R cells. Using energy and performance constraints, we determined the optimum number of bits/cell in the multi-bit RRAM array to be 3. The total write and read energy of the 3 bits/cell TiO_2-based RRAM array was 4.06 pJ/bit and 188 fJ/bit for 100 nsec and 1 nsec write and read access times while the optimized 3 bits/cell HfO_x-based RRAM array consumed 365 fJ/bit and 173 fJ/bit for 1 nsec and 200 nsec write and read access times, respectively. We explored the trade-off between the read energy consumption and the robustness against process variations for uniform and non-uniform memristor state assignments in the multi-bit RRAM array. Using the proposed models, we analyzed the effects of process, voltage and temperature variations on performance and energy consumption and the reliability of n-bit 1T1R memory cells. Our analysis showed that multi-bit TiO_2 RRAM is more sensitive to OTF while HfO_x RRAM is more sensitive to LER and is more susceptible to voltage and temperature variations.

Chapter 4

Sub-threshold Logic Design using Feedback Equalization

4.1 Introduction

The use of sub-threshold digital CMOS logic circuits is becoming increasingly popular in energy-constrained applications where high performance is not required. We propose using a novel feedback equalizer circuit to improve energy efficiency in sub-threshold digital logic circuits. The key idea here is to explore the use of techniques which are commonly used in communication theory in the design of robust energy-efficient digital logic circuits. Feedback equalization for above-threshold regime has previously been proposed by (Takhirov et al., 2012) and we will explore it for sub-threshold circuits. Using a feedback equalizer circuit that adjusts the switching thresholds of the gates (just before the flip flops) based on the prior sampled outputs, we can reduce the propagation delay of the critical path in the combinational logic block to make the sub-threshold system more robust to timing errors and at the same time reduce the dominant leakage energy of the entire design.

4.2 Related Work

Several techniques have been proposed to design robust ultra-low power sub-threshold circuits. As described earlier, transistor upsizing (Kwong et al., 2009) and increasing the logic path depth (Verma et al., 2008), (Zhai et al., 2005) can be used to overcome process variations. The use of gates of different drive strengths has also been proposed to overcome process variations (Choi et al., 2004). A detailed analysis on the timing variability and the metastability of the flip flops designed in sub-threshold region has been presented in (Lotze et al., 2008) and (Li et al., 2011), respectively. The authors in (Lotze and Manoli, 2012) have used the Schmitt Trigger structures in sub-threshold logic circuits to improve the I_{ON}/I_{OFF} ratio and effectively reduce the leakage from the gate output node. The authors in (Pu et al., 2010) proposed a design technique that uses a configurable V_T balancer to mitigate the V_T mismatch of transistors operating in sub-threshold regime. The authors in (Zhou et al., 2011) propose to boost the drain current of the transistors using minimum-sized devices with fingers to mitigate the inverse narrow width effect in sub-threshold domain. An analytical framework for sub-threshold logic gate sizing based on statistical variations has been proposed in (Liu et al., 2012) which provides narrower delay distributions compared to the state-of-the-art approaches. Body-biasing has also been proposed to mitigate the impact of variations (Jayakumar and Khatri, 2005). A controller that uses a sensor to first quantify the effect of process variations on sub-threshold circuits and then generates an appropriate supply voltage to overcome that effect has been proposed in (Mishra et al., 2009). In (De Vita and Iannaccone, 2007), the authors have used a current reference circuit to design a voltage regulator providing a supply voltage that makes the propagation delay of the sub-threshold digital circuits almost insensitive to temperature and process variations. Using differential dynamic logic in standby mode, the authors in (Liu and Rabaey, 2012) propose to suppress leak-

age in the sub-threshold circuits. Error detection and correction techniques have been widely used in resilient, energy-efficient above-threshold architectures (Tschanz et al., 2010), (Bowman and Tschanz, 2010), (Bull et al., 2011), (Chae and Mukhopadhyay, 2014), (Whatmough et al., 2013). The authors in (Tschanz et al., 2010) and (Bowman and Tschanz, 2010) have used a tunable replica circuit (with 3.5% leakage power overhead, 2.2% area overhead) and error-detection sequentials (with 5.1% leakage power overhead, 3.8% area overhead) to monitor critical path delays and mitigate dynamic variation guardbands for maximum throughput in above-threshold regime. Using an adaptive clock controller based on error statistics, the proposed processor architecture operates at maximum efficiency across a range of dynamic variations.

Equalization techniques have been proposed to design energy-efficient logic circuits operating in the above-threshold regime. The authors in (Takhirov et al., 2012) proposed to use the feedback equalizer circuit with Schmitt Trigger (FEST) to mitigate timing errors resulting from voltage scaling and in turn improve energy efficiency for above-threshold logic circuits. Using the FEST circuit, they lower down the critical supply voltage of a 4-bit Kogge-Stone adder as well as a 3-tap 4-bit finite impulse response (FIR) filter leading to 20% and 40% decrease in the total consumed energy, respectively. We use the equalization technique developed in (Takhirov et al., 2012) for designing logic circuits in sub-threshold regime.

We propose a circuit-level scheme that uses a communications-inspired feedback equalization technique in the critical path to mitigate the timing errors rising from aggressive voltage scaling in sub-threshold digital logic circuits. It should be noted that we are not designing sub-threshold communication circuits. We are proposing the design of sub-threshold logic circuits that leverage principles of communication theory. Several authors have already used feedback-based techniques to boost the weak low-voltage signals in global interconnections (Seo et al., 2007), (Singh et al.,

2008), (Schinkel et al., 2006), (Sridhara et al., 2008), (Kim and Seok, 2014). The authors in (Seo et al., 2007) and (Singh et al., 2008) proposed the self-timed regenerator (STR) technique to improve the speed and power for on-chip global interconnects leading to 14% delay improvement over the conventional repeater design in above-threshold regime. The authors in (Kim and Seok, 2014) proposed a reconfigurable interconnect design technique based on regenerators for ultra-dynamic-voltage-scaling (UDVS) systems to improve performance and energy efficiency across a large range of above-threshold supply voltages.

We propose using a feedback equalizer circuit in the design of sub-threshold digital logic circuits. This feedback equalizer circuit can reduce energy consumption and improve performance of the sub-threshold digital logic circuits. Using feedback equalizer circuits, we further scale down the operating voltage of the sub-threshold circuit to decrease the dynamic energy as well as the leakage energy in sub-threshold CMOS circuits.

4.3 Equalized Flip flop versus Conventional Flip flop

In this section, we first explain the use of the feedback equalizer circuit in the design of an equalized flip flop and then provide a detailed comparison of the equalized flip flop with a conventional flip flop in terms of area, setup time and performance. We propose the application of a feedback equalizer (designed using a variable threshold inverter (Sridhara et al., 2008) shown in Figure 4·1) along with the classic master-slave positive edge-triggered flip flop (Rabaey et al., 2003) to implement an equalized flip flop. The equalized flip flop dynamically modifies the switching threshold of the gate before the flip flop based on the previous sampled data. If the previous output of the gate is a zero, the equalized flip flop lowers down the switching threshold

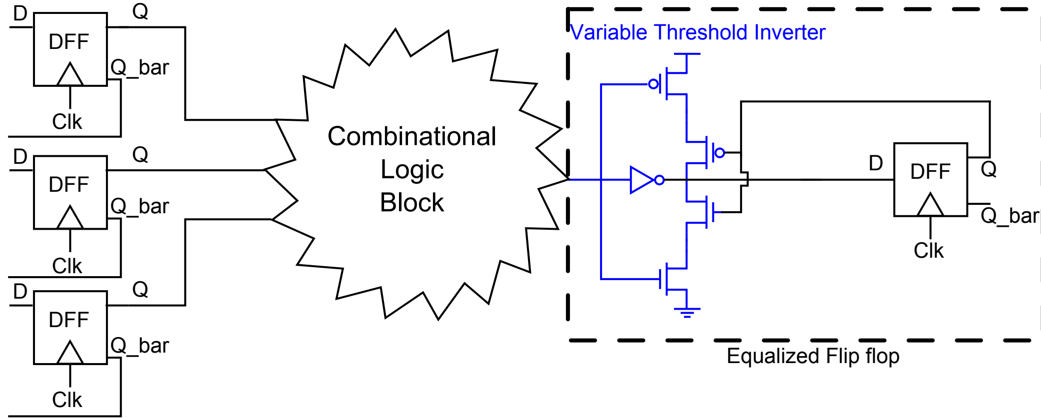

Figure 4·1: Feedback equalizer (designed using a variable threshold inverter (Sridhara et al., 2008)) can be combined with a traditional master-slave flip flop to design an equalized flip flop.

which speeds up the transition to one. Similarly if the previous output is one, the equalized flip flop increases the switching threshold which speeds up the transition to zero. In this configuration, the circuit adjusts the switching threshold and facilitates faster high-to-low and low-to-high transitions. The DC response of the feedback equalizer circuit in sub-threshold regime is shown in Figure 4·2. The switching of the variable threshold inverter is dynamically adjusted based on the previous sampled output data. Compared to the above-threshold regime, the reduced noise margin in weak inversion region does not allow for aggressively overscaling the supply voltage while using the variable threshold inverter. So we do not use the Schmidt Trigger circuit along with the feedback equalizer circuit as proposed in (Takhirov et al., 2012) for above-threshold operation. The equalized flip flop has 6 transistors more than the conventional master-slave positive edge-triggered flip flop (Rabaey et al., 2003). Compared to a classic master-slave flip flop with 22 transistors (7 inverters and 4 transmission gates (TG)), the area overhead of the equalized flip flop is around 27%. This area overhead gets amortized across the critical path of the sub-threshold logic.

The total power consumed by a digital circuit can be calculated using

$$P_T = P_{DYN} + P_L = C_{eff} V_{DD}^2 f + I_{leak} V_{DD} \tag{4.1}$$

In Equation (4.1), P_{DYN} and P_L are the dynamic and leakage power components of the digital circuit, respectively. C_{eff} is the average total capacitance of the entire circuit, V_{DD} is the supply voltage and f is the operating frequency of the circuit. I_{leak} is the leakage current and can be written as

$$I_{leak} = \mu_0 C_{ox} \frac{W}{L} (n-1) V_{th}^2 e^{\frac{\eta V_{DS} - V_T}{n V_{th}}} \tag{4.2}$$

In Equation (5.2), V_T is the transistor threshold voltage, V_{th} is the thermal voltage, n is the sub-threshold slope factor and η is the DIBL coefficient. There is an exponential relationship between the leakage current and the supply voltage (due to the DIBL effect and for $V_{DS} \approx V_{DD}$). Using the equalized flip flop, we can scale down the supply voltage while maintaining the zero word error rate at a given operating frequency and achieve lower dynamic power consumption (due to the quadratic relationship between the dynamic power and the supply voltage) as well as lower leakage power (due to smaller DIBL effect which exponentially decreases the leakage current). Similar to the area overhead, the dynamic power as well as the leakage overhead of the variable threshold inverter gets amortized across the entire sub-threshold combinational logic block.

Figure 5·5 illustrates the timing waveforms of the output carry bit of an 8-bit carry-lookahead adder implemented in UMC 130 nm process using static complementary CMOS logic. In the figure, we show the waveform for the input node of the non-

Supply voltage (mV)	Delay E-logic (nsec)	Delay NE-logic (nsec)	t_{c-q} E-flip flop (nsec)	t_{c-q} NE-flip flop (nsec)	Setup time E-flip flop (nsec)	Setup time NE-flip flop (nsec)
350	226	255	3.85	3.82	8.62	6.07
330	336	378	5.72	5.66	12.80	9.01
310	489	532	8.30	8.23	18.62	13.11
290	750	842	12.72	12.61	28.55	20.09
270	1064	1159	18.04	17.87	40.49	28.49
250	1661	1820	28.15	27.89	63.20	44.46

Table 4.1: Comparison between the characteristics of the equalized flip flop (E-flip flop) with the conventional non-equalized master-slave flip flop (NE-flip flop) at different supply voltages operating in sub-threshold regime. Feedback equalization technique reduces the propagation delay of the 8-bit carry-lookahead adder CMOS logic whereas the setup time and t_{c-q} delay of the conventional flip flop is smaller than the equalized flip flip.

equalized flip flop (NE-flip flop), the input node of the equalized flip flop (E-flip flop), the latched output for both cases and the output node of the variable threshold inverter. Compared to the signal at the input node of the non-equalized flip flop, the variable threshold circuit provides sharper transitions and decreases the propagation delay of the critical path of the sub-threshold logic. However, it should be noted that excessive positive feedback might lead to increased glitches at the input of the equalizer which increases the probability of occurrence of timing errors. Therefore, the transistors in variable threshold inverter need to be carefully sized to avoid the errors rising due to the glitches.

It has been shown in (Rabaey et al., 2003) that the setup time of the conventional master-slave positive edge-triggered flip flop is $t_{setup} = 3t_{inv} + t_{TG}$. Since the equalized flip flop uses an extra variable-threshold inverter at its output, the setup time of the equalized flip flop will be larger $t_{setup} \approx 4t_{inv} + t_{TG}$. The t_{c-q} delay of the conventional flip flop is $t_{c-q} = t_{inv} + t_{TG}$. Since the equalized flip flop has the variable threshold inverter as extra load at the output, the t_{c-q} delay of the equalized flip flop is $t_{c-q} = t_{inv} + \Delta t + t_{TG}$ which is slightly larger than the t_{c-q} delay of the conventional

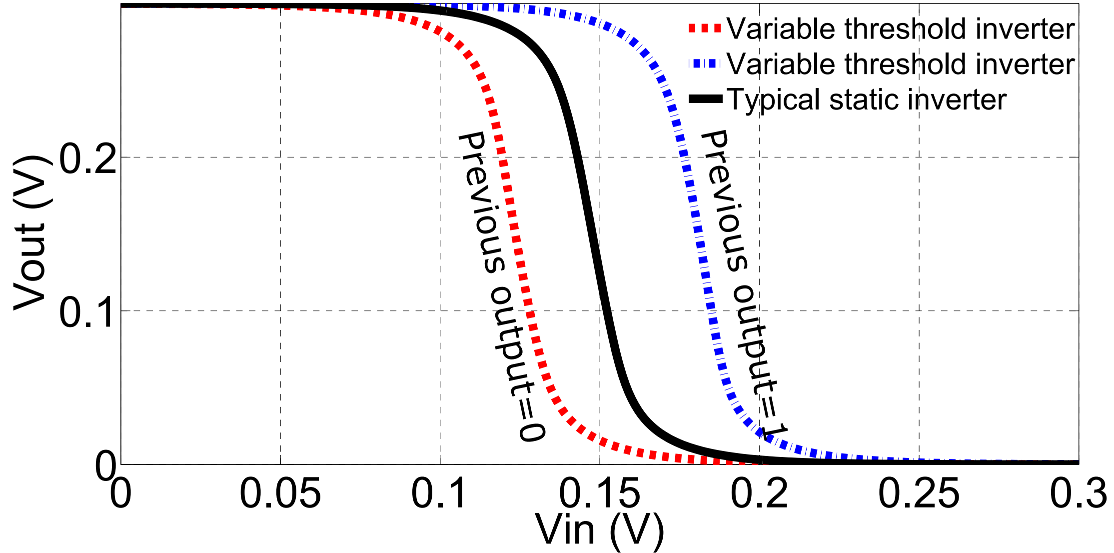

Figure 4·2: DC response of the variable threshold circuit in sub-threshold regime. The switching threshold of the inverter is modified based on the previous sampled output data.

flip flop. Here Δt is the increase in inverter delay due to the extra load. However, the feedback equalizer circuit can significantly lower down the propagation delay of the critical path by providing a faster charging (or discharging) path for the input capacitance of the flip flop. Table 4.1 compares the propagation delay, setup time and the t_{c-q} delay of the two 8-bit carry-lookahead adders designed with conventional flip flop and equalized flip flop in UMC 130 nm when operating with different supply voltages. The variable threshold inverter has been accurately sized to minimize the total delay of the critical path.

4.4 Experimental Results

In this section, we perform a detailed comparison, in terms of performance and energy consumption, of a sample 8-bit carry-lookahead adder designed in UMC 130 nm process using both equalized and non-equalized flip flops. We analyze the impact of

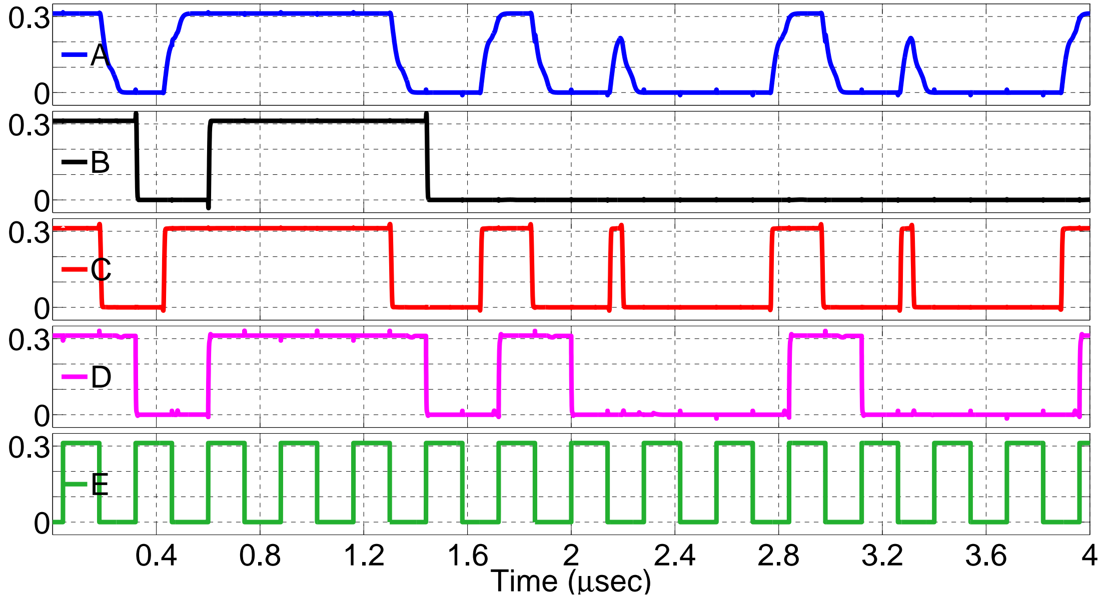

Figure 4·3: Comparison between the timing waveforms of the input node of the conventional flip flop (A), output node of the conventional flip flop (B), input node of the equalized flip flop (C), output node of the equalized flip flop (D), output node of the variable threshold inverter (E). Feedback circuit makes sharper transitions in the waveforms of the logic output node helping the equalized flip flop sample the correct data.

the proposed feedback equalization technique when the frequency of the sub-threshold logic is improved at a fixed supply voltage and also when the energy of the sub-threshold logic is reduced by scaling down the supply voltage at a fixed operating frequency. We also explore the use of the proposed feedback equalizer circuit to reduce the amount of transistor oversizing for mitigating the process variation effects.

4.4.1 Performance improvement at the fixed supply voltage

We first explore the case where the feedback equalizer circuit reduces the rise/fall time of the last gate and hence the critical path of the combinational logic block leading to a higher operating frequency. The variable threshold inverter can be used to reduce the propagation delay of the critical path at any operating supply volt-

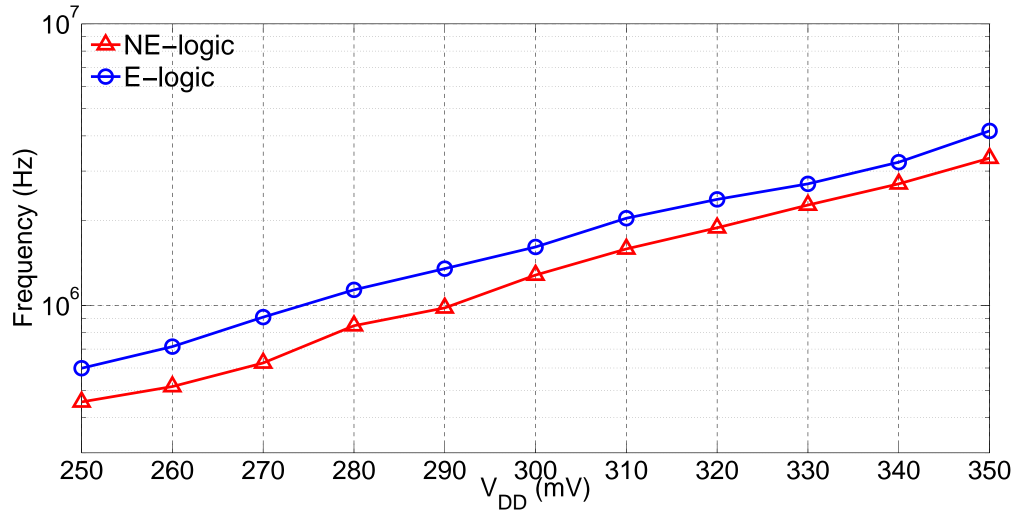

Figure 4·4: Operating frequency of the 8-bit carry lookahead adder for zero word error rate as function of different sub-threshold supply voltages. The equalized logic (E-logic) can run 22.87% (on average) faster than the non-equalized logic (NE-logic).

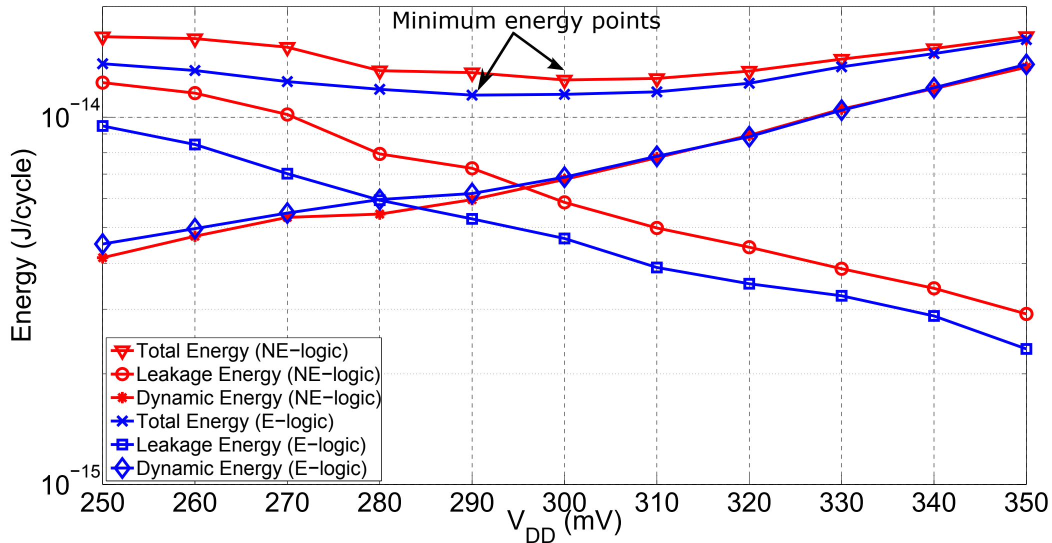

Figure 4·5: Comparison between the total consumed energy as well as the dynamic/leakage components of the 8-bit carry lookahead adder for different supply voltages. At the minimum energy supply voltage, the equalized logic is burning 18.4% less total energy compared to the non-equalized version.

age. Figure 5·13 shows the operating frequency of the 8-bit carry lookahead adder for different sub-threshold supply voltages at zero word error rate when using an

equalized and conventional flip flop. Here, we determined the optimum sizing for the feedback equalizer circuit that minimizes the propagation delay of the critical path and prevents glitches for zero error rate operation at each supply voltage data point. The sizing of the combinational logic block is the same for both the equalized and non-equalized circuit and is determined using the design methodology described in (Kwong et al., 2009) to address the degraded noise margin levels in sub-threshold regime. The operating frequency of the equalized logic is 22.87% (on average) higher than the non-equalized logic over the range of 250 mV to 350 mV. The amount of performance acceleration in aggressively scaled supply voltages is more promising compared to voltages close to the threshold as the variable threshold inverter is capable of significantly decreasing the large transition times of the logic designed in deep sub-threshold region. At 250 mV supply voltage, the equalized flip flop improves the operating frequency of the logic by 27.8% whereas the amount of performance improvement at 350 mV is 16.2%.

By reducing the propagation delay of the critical path, the feedback equalizer circuit is capable of reducing the dominant leakage energy of the digital logic in sub-threshold regime. Figure 5·14 illustrates a head-to-head comparison between the total energy, the dynamic energy and the leakage energy of the 8-bit carry lookahead adder for different supply voltages while using the equalized or conventional non-equalized flip flops. By adding the feedback equalizer to the conventional flip flop, the dynamic energy of the logic with the equalized flip flop is 3.47% (on average) larger than the logic designed with non-equalized conventional flip flop. This is negligible compared to the 22.6% reduction in the leakage component of the design. At the minimum energy supply voltage, the equalized logic consumes 18.4% less total energy compared to the non-equalized version. The feedback circuit drops the minimum energy supply voltage of the logic by 10 mV while maintaining the zero word error rate operation.

The leakage energy reduction mechanism of the feedback equalization technique in sub-threshold regime is due to the fact that the total delay along the critical path of the equalized logic decreases (the sub-threshold CMOS logic is running faster) leading to lower leakage energy according to (Kwong et al., 2009)

$$E_T = E_{DYN} + E_L = C_{eff}V_{DD}^2 + I_{leak}V_{DD}T_D \tag{4.3}$$

In Equation (5.1), E_T is the total dissipated energy, E_{DYN} and E_L are the dynamic and leakage components, respectively. $T_D = 1/f$ is the total delay along the critical path of a digital circuit.

Decreasing the dominant leakage energy component of the sub-threshold logic together with reducing the propagation delay of the critical path, the feedback equalization technique lowers the energy-delay product of the logic designed in weak inversion region. On average, the equalized 8-bit carry lookahead adder has 30.44% smaller energy-delay product value compared to the non-equalized logic over the range of 250 mV to 350 mV for zero word error rate operation. If we compare the energy-delay product at the respective minimum energy supply voltages, the equalized flip flop reduces the energy-delay product of the 8-bit carry lookahead adder by 35.4%. Table 5.5 compares the minimum energy point and the corresponding operating frequency of the equalized logic design (E-logic) vs. non-equalized logic design (NE-logic) of an 8-bit carry lookahead adder (CLA), 8-bit Array Multiplier and 3-tap 8-bit FIR filter, all designed in UMC 130 nm process. On an average, the equalization technique has 24.49% lower energy-delay product than the non-equalized logic design.

Logic block	NE-logic Energy (fJ/cycle)	E-logic Energy (fJ/cycle)	NE-logic Frequency (MHz)	E-logic Frequency (MHz)
8-bit CLA	12.63	10.3	1.28	1.62
8-bit Multiplier	16.27	15.24	1.22	1.49
8-bit FIR filter	100.32	94.84	0.64	0.71

Table 4.2: Comparison between the minimum energy point and the corresponding operating frequency of the equalized logic (E-logic) vs. non-equalized (NE-logic) design of various logic blocks.

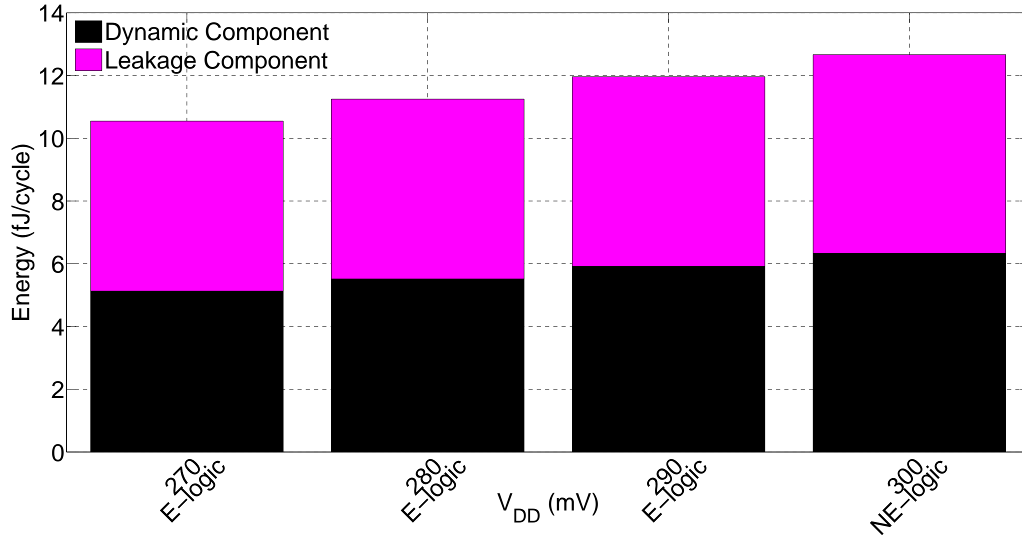

Figure 4·6: Comparison between the energy consumed by the equalized (E-logic) vs. non-equalized (NE-logic) 8-bit carry lookahead adder for different supply voltages with fixed performance ($f = 1.28\ MHz$) at zero word error rate. The non-equalized logic design consumes minimum energy at 300 mV. The equalized flip flop enables 30 mV supply voltage scaling leading to 16.72% lower total consumed energy. The equalized flip flop cannot operate at $V_{DD} < 270mV$ due to the occurrence of timing errors.

4.4.2 Leakage reduction at the fixed operating frequency

As described in Section 4.3, the equalized flip flop can be used to scale supply voltages (while maintaining the operating frequency) to lower down the dominant leakage energy by decreasing the leakage current of the sub-threshold logic. We designed

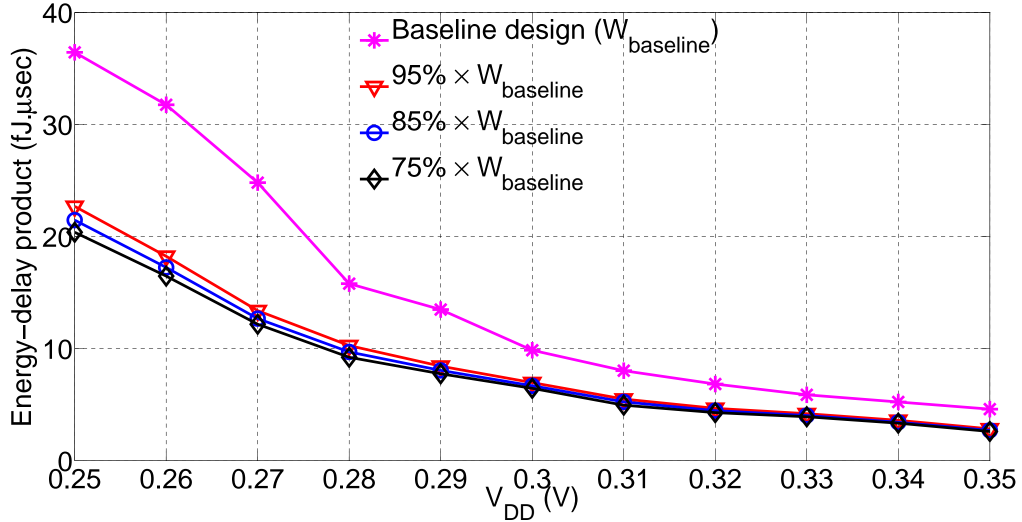

Figure 4·7: Energy-delay product of the scaled-down equalized 8-bit carry lookahead adder for zero word error rate operation. We can achieve reliable operation even when the transistors in the equalized logic design are scaled down to as small as $75\% \times W_{baseline}$.

the feedback equalizer circuit for each scaled supply voltage that ensured the reliable operation of the equalized design without any timing errors. Figure 4·6 illustrates the dynamic and leakage energy components of the 8-bit carry lookahead adder at the minimum energy supply voltage (of the non-equalized design) and below. The operating frequency of all design points with zero word error rate is $f = 1.28\ MHz$ (the frequency of the minimum energy supply voltage for the non-equalized design). Compared to the non-equalized design, the equalized design can operate at 30 mV lower supply voltage leading to 16.72% lower energy consumption. The equalized design cannot operate for $V_{DD} < 270mV$ due to the larger rise/fall times that lead to timing errors.

4.4.3 Mitigating process variations

Using the proposed feedback-based technique, the critical sizing approach used for designing the sub-threshold logic circuits in (Kwong et al., 2009) can be relaxed. The

Scaled-down equalized logic size	Total energy saving w.r.t non-equalized	Total energy saving w.r.t equalized
$95\%\times W_{baseline}$	12.87%	6.72%
$85\%\times W_{baseline}$	16.79%	10.92%
$75\%\times W_{baseline}$	20.72%	15.12%

Table 4.3: Energy savings in scaled-down equalized logic compared to baseline non-equalized and equalized logic at the minimum energy supply voltage at zero word error rate operation for 8-bit carry lookahead adder.

transistor sizing can be scaled down while ensuring the reliable operation using feedback equalizer circuit in presence of process variations. For the 8-bit carry-lookahead adder in UMC 130 nm process, the transistors sized using (Kwong et al., 2009) ($W_{baseline}$) can be scaled down to $75\%\times W_{baseline}$ while matching the operating frequency of the equalized design and non-equalized design. Figure 5·17 illustrates the energy-delay product of the scaled down equalized logic and baseline non-equalized logic for different sub-threshold supply voltages. At a given voltage, compared to the non-equalized design, the equalized design uses smaller transistors and has lower propagation delay resulting in a reduction of both dynamic and leakage energy. For a $3\sigma_{V_T} = 30mV$ variation in threshold voltage, the equalized design can reliably operate without the occurrence of any timing errors. Table 5.3 summarizes the amount of energy savings of the equalized logic with scaled down transistors compared to the baseline non-equalized and the equalized logic where the combinational logic has been sized according to the method proposed in (Kwong et al., 2009). Overall the feedback equalization along with transistor size scaling consumes up to 20.72% lower total energy compared to the conventional non-equalized design in sub-threshold regime.

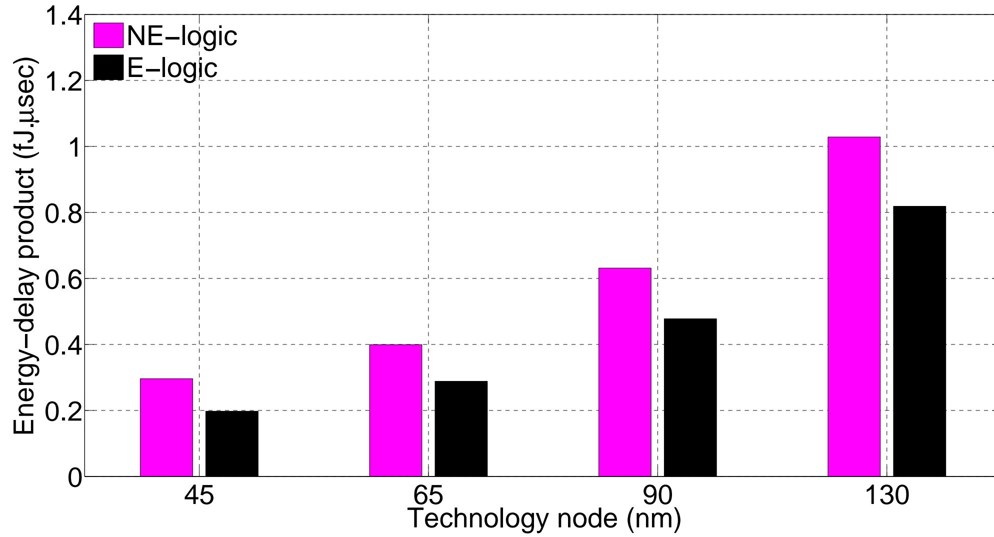

Figure 4·8: Energy-delay product of a 8-bit carry lookahead adder designed using equalized logic (E-logic) vs. non-equalized logic (NE-logic) at zero word error rate at different technology nodes. The equalized logic approach reduces the energy-delay product of the sub-threshold logic by up to 26.46% across all technology nodes in the minimum energy supply voltage.

4.5 Effect of Technology Scaling

In this section, we analyze the effect of technology scaling on the performance improvement and the energy reduction obtained using feedback equalization technique in sub-threshold regime. In scaled technology nodes, the contribution of leakage energy component dominates due to larger DIBL effect as well as smaller V_T values. Running the sub-threshold logic faster, the equalizer will more effectively reduce the leakage energy component and in turn decrease the energy-delay product in scaled technology nodes. Figure 5·21 illustrates the value of the energy-delay product of the 8-bit carry lookahead adder designed using PTM (Ptm,) for 4 different technology nodes and operating at zero word error rate at minimum energy supply voltage. Compared to the non-equalized logic design, the energy-delay product of the equalized logic design is 20.45%, 24.32%, 27.82% and 33.25% smaller at 130 nm, 90 nm,

65 nm and 45 nm technology nodes, respectively. On average, the equalized flip flop reduces the energy-delay product of the sub-threshold logic by up to 26.46% across all technology nodes at the minimum energy supply voltage.

4.6 Summary

In this chapter, we proposed the application of a variable threshold inverter-based feedback equalization circuit to reduce the dominant leakage energy of the digital CMOS logic operating in sub-threshold regime. Adjusting the switching thresholds based on the prior sampled outputs, the feedback equalization circuit enables a faster switching of the logic gate outputs and provides the opportunity to reduce the leakage current in weak inversion region. We implemented a non-equalized and an equalized design of an 8-bit carry lookahead adder in UMC 130 nm process using static complementary CMOS logic and managed to reduce the propagation delay of the critical path of the sub-threshold logic and correspondingly lower the dominant leakage energy, leading to 35.4% decrease in energy-delay product of the conventional non-equalized design at minimum energy supply voltage. Using the feedback equalizer circuit, we obtained 16.72% reduction in energy through voltage scaling while maintaining an operating frequency of 1.28 MHz. We showed that the equalized sub-threshold 8-bit carry lookahead adder requires lower upsizing to tolerate process variation effects leading to 20.72% lower total energy.

Chapter 5

Tunable Sub-threshold Logic Circuits using Adaptive Feedback Equalization

5.1 Introduction

The dominating process variation effects necessitates the application of adaptive circuits to mitigate timing errors in digital sub-threshold logic circuits. We propose using an adaptive feedback equalizer circuit in the design of tunable sub-threshold digital logic circuits. This adaptive feedback equalizer circuit can reduce energy consumption and improve performance of the sub-threshold digital logic circuits. At the same time, the tunability of this feedback equalizer circuit enables post-fabrication tuning of the digital logic block to overcome worse than expected process variations as well as lower energy and improve performance.

5.2 Adaptive Equalized Flip flop versus Conventional Flip flop

In this section, we first explain the use of the adaptive feedback equalizer circuit in the design of an adaptive equalized flip flop and then provide a detailed comparison of the equalized flip flop with a conventional flip flop in terms of area, setup time and performance. We propose the use of a variable threshold inverter (Sridhara et al.,

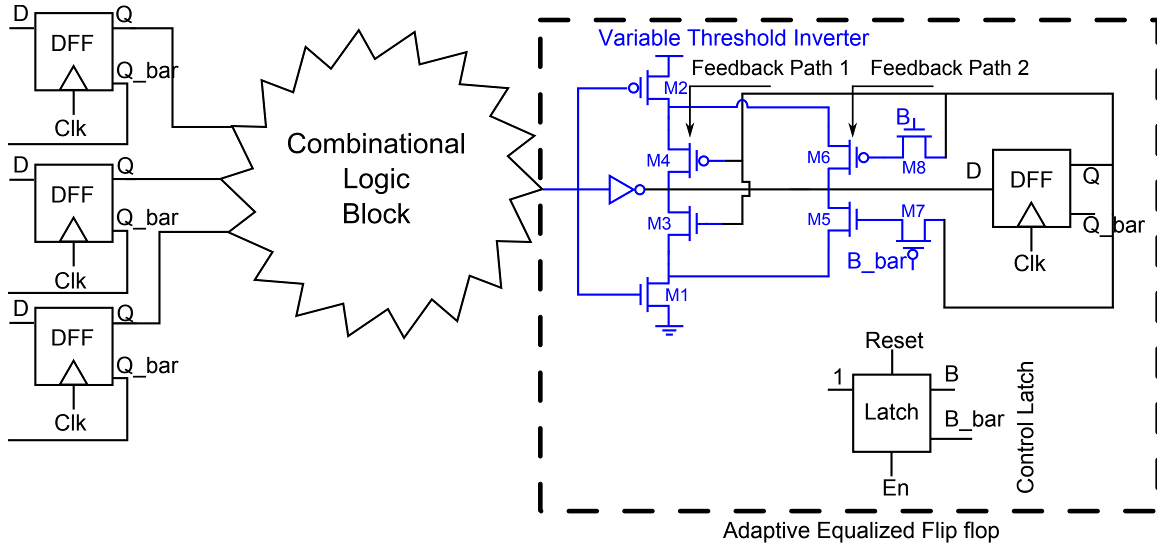

Figure 5·1: Adaptive feedback equalizer circuit with multiple feedback paths (designed using a variable threshold inverter (Sridhara et al., 2008)) can be combined with a traditional master-slave flip flop to design an adaptive equalized flip flop.

2008) (see Figure 5·1) as an adaptive feedback equalizer along with the classic master-slave positive edge-triggered flip flop (Rabaey et al., 2003) (see Figure 5·2) to design an adaptive equalized flip flop. This adaptive feedback equalizer circuit consists of 2 feedforward transistors (M1 and M2 in Figure 5·1) and 4 control transistors (M3 and M4 for feedback path 1 that is always ON and M5 and M6 for feedback path 2 that can be conditionally switched ON post-fabrication in Figure 5·1) that provide extra pull-up/pull-down paths in addition to the pull-up/pull-down path in the static inverter for the DFF input capacitance. The extra pull-up/pull-down paths are enabled whenever the output of the critical path in the combinational logic changes. The control transistors M5 and M6 are enabled/disabled through transistor switches (M7 and M8) that are controlled by an asynchronous control latch. The value of the static control latch is initially reset to 0 during chip bootup. After bootup, if required a square pulse is sent to the En terminal to set the output of the latch to 1 to switch ON M7 and M8 which enables feedback path 2.

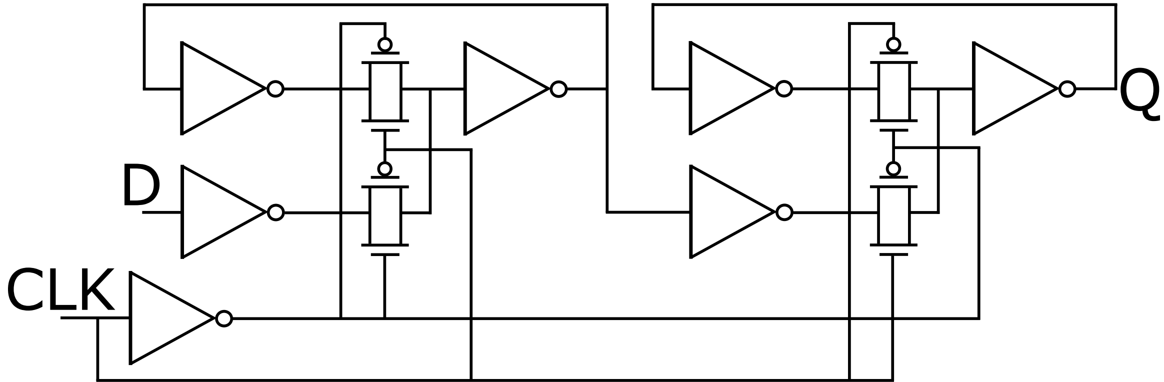

Figure 5·2: Circuit diagram of classic master-slave positive edge-triggered flip flop (Rabaey et al., 2003).

The adaptive equalized flip flop effectively modifies the switching threshold of the static inverter in the feedback equalizer based on the output of flip flop in the previous cycle. If the previous output of the flip flop is a zero, the switching threshold of the static inverter is lowered, which speeds up the transition of the flip flop input from zero to one. Similarly if the previous output is one, the switching threshold is increased which speeds up the transition to zero. Effectively, the circuit adjusts the switching threshold and facilitates faster high-to-low and low-to-high transitions of the flip flop input. Moreover, the smaller input capacitance of the feedback equalizer reduces the switching time of the last gate in the combinational logic block. Overall, this reduces the total delay of the sequential logic. The DC response of the adaptive feedback equalizer circuit with 2 different feedback paths in sub-threshold regime is shown in Figure 5·3.

The adaptive equalized flip flop has 8 more transistors than the conventional master-slave flip flop (Rabaey et al., 2003). Compared to a classic master-slave flip flop with 22 transistors (7 inverters and 4 transmission gates (TG)), the area overhead of the adaptive equalized flip flop is 36%. The area overhead of the control latch with 10 transistors (3 inverters and 2 transmission gates) is 45%. This area overhead gets

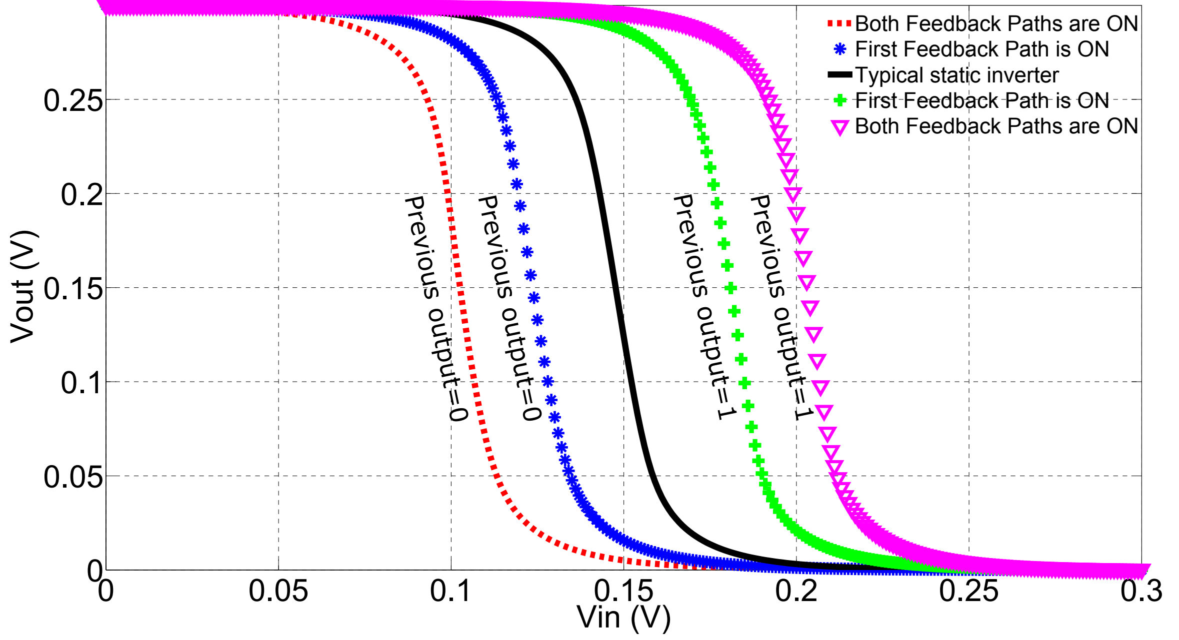

Figure 5·3: DC response of the adaptive feedback equalizer circuit with 2 different feedback paths in sub-threshold regime. The switching threshold of the inverter is modified based on the previous sampled output data.

amortized across the entire sequential logic block.

The total energy consumed by a digital circuit in the sub-threshold regime can be calculated using

$$E_T = E_{DYN} + E_L = C_{eff} V_{DD}^2 + I_{leak} V_{DD} T_D \tag{5.1}$$

In Equation (5.1), E_{DYN} and E_L are the dynamic and leakage energy components, respectively. C_{eff} is the total capacitance of the entire circuit, V_{DD} is the supply voltage and $T_D = 1/f$ is the total delay along the path of the digital logic block. Feedback equalization enables us to reduce the delay of the path in the digital logic block, which in turn reduces the leakage energy. In Equation (5.1), I_{leak} is the leakage current and can be written as

$$I_{leak} = \mu_0 C_{ox} \frac{W}{L} (n-1) V_{th}^2 e^{\frac{\eta V_{DS} - V_T}{n V_{th}}} \quad (5.2)$$

In Equation (5.2), V_T is the transistor threshold voltage, V_{th} is the thermal voltage, n is the sub-threshold slope factor and η is the DIBL coefficient. There is an exponential relationship between the leakage current and the supply voltage (due to the DIBL effect and because $V_{DS} \approx V_{DD}$). Using the equalized flip flop, we can scale down the supply voltage while maintaining the zero error rate at a given operating frequency and achieve lower dynamic energy consumption (due to the quadratic relationship between the dynamic energy and the supply voltage) as well as lower leakage energy (due to smaller DIBL effect which exponentially decreases the leakage current). Similar to the area overhead, the dynamic energy as well as the leakage energy overhead of the variable threshold inverter gets amortized across the entire sequential logic block.

The setup time of the conventional master-slave positive edge-triggered flip flop is $t_{s-t} = 3t_{inv} + t_{TG}$ (Rabaey et al., 2003). Since the adaptive equalized flip flop uses an extra variable-threshold inverter at its input, the setup time of the adaptive equalized flip flop will be larger $t_{s-t-equ} \approx 4t_{inv} + t_{TG}$ (Zangeneh and Joshi, 2014b). The clk-to-q delay of the conventional flip flop is $t_{c-q} = t_{inv} + t_{TG}$. Since the equalized flip flop has the variable threshold inverter as extra load at the output, the t_{c-q} delay of the equalized flip flop is $t_{c-q-equ} = t_{inv} + t_{TG} + \Delta t_{c-q}$ which is slightly larger than the t_{c-q} delay of the conventional flip flop. Here Δt_{c-q} is the increase in inverter delay due to the extra load of the adaptive feedback equalizer circuit. However, the adaptive feedback equalizer circuit can significantly lower down the propagation delay of the critical path because the small input capacitance of the feedback equalizer reduces the switching time of the last gate in the combinational logic. The hold time of the classic

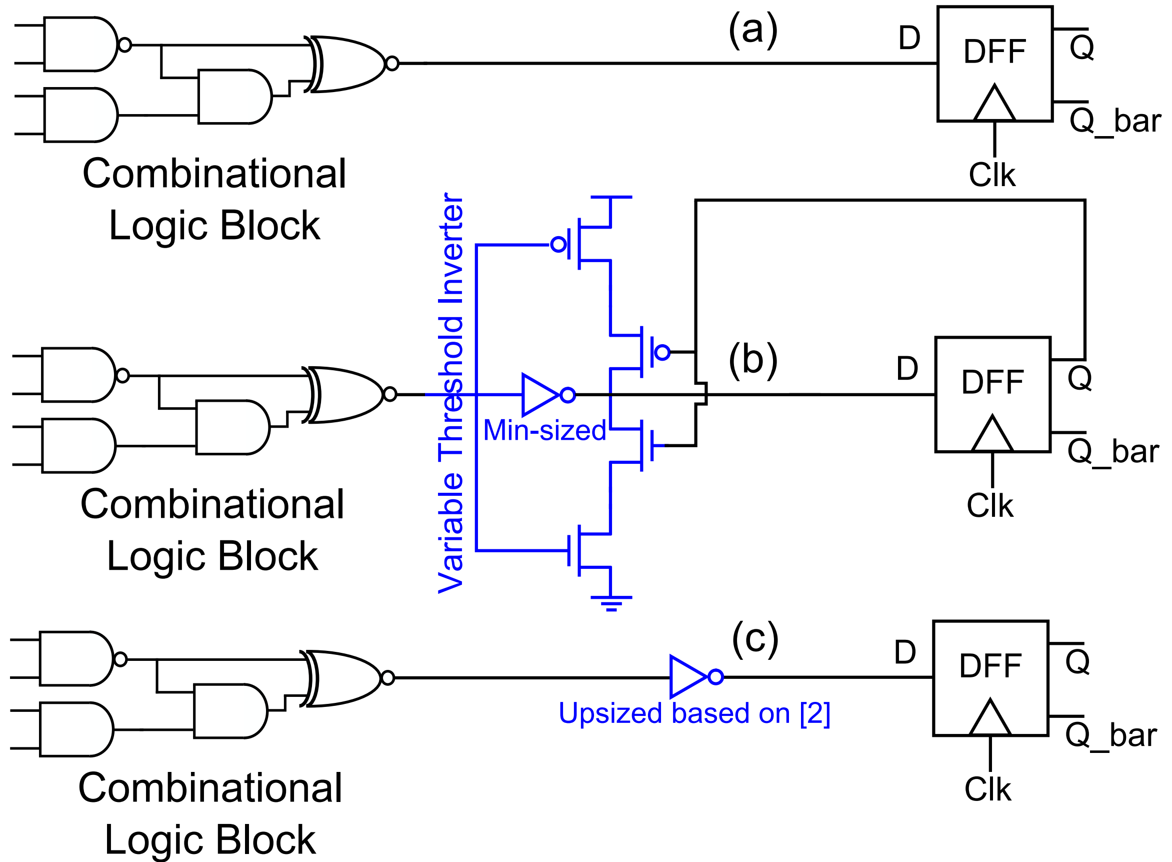

Figure 5·4: Block diagram of the original non-equalized design (a), equalized design with 1 feedback path ON (b) and buffer-inserted non-equalized design (c).

master-slave positive edge-triggered flip flop is zero (Rabaey et al., 2003). Therefore the adaptive feedback equalizer circuit does not impact the hold time violations.

We analyze the capability of the adaptive feedback equalizer circuit to reduce the transition time of the last gate in critical path of the sub-threshold logic and make a comparison with the original non-equalized design, and the buffer-inserted non-equalized design (see Figure 5·4). The classic buffer insertion technique (Figure 5·4(c)) will reduce the total delay along critical path of the sub-threshold logic. Like the gates in the combinational logic, the buffer used in Figure 5·4(c) is upsized to account for the process variation effects based on the design methodology proposed in (Kwong

Design methodology	Transition time (ns)	t_{c-q} (ns)	t_{s-t} (ns)
NE-logic	25.03	10.7	14.1
Buffer-inserted NE-logic	11.38	10.7	20.2
E-logic	2.9	11.5	21

Table 5.1: Comparison between the timing characteristics of the original non-equalized design, the equalized design with 1 feedback path ON and the buffer-inserted non-equalized design.

et al., 2009). Using a minimum-sized inverter instead of an upsized inverter would further lower down the delay but has lower reliability with respect to the dominant process variation effects in sub-threshold regime. So we propose to use a combination of minimum-sized inverter and feedback equalizer circuit along the critical path of the sub-threshold logic. Minimum-sized inverter reduces the total delay and the feedback equalizer mitigates the effect of process variation. Table 5.1 compares the timing characteristics of the original non-equalized logic (NE-logic) design, the buffer-inserted non-equalized logic and the equalized logic (E-logic) design with 1 feedback path ON. The adaptive feedback equalizer circuit reduces the propagation delay along the critical path of the digital sub-threshold logic while ensuring reliable operation compared to the non-equalized logic and the buffer-inserted design. Our analysis shows that the classic buffer insertion technique reduces the transition time of the last gate in critical path of the NE-logic by more than half and the proposed adaptive feedback equalizer circuit could further reduce the delay by 1/4. The setup time and the clk-to-q delay of the equalized flip flop is larger than that of the conventional flip flop, but the total delay of the E-logic is smaller than the total delay of the NE-logic.

Figure 5·5 illustrates the timing waveforms of the output carry bit of a 64-bit adder implemented in UMC 130 nm process using non-equalized logic (NE-logic) and equalized logic (E-logic). In the figure, we show the waveform of clock signal, the input

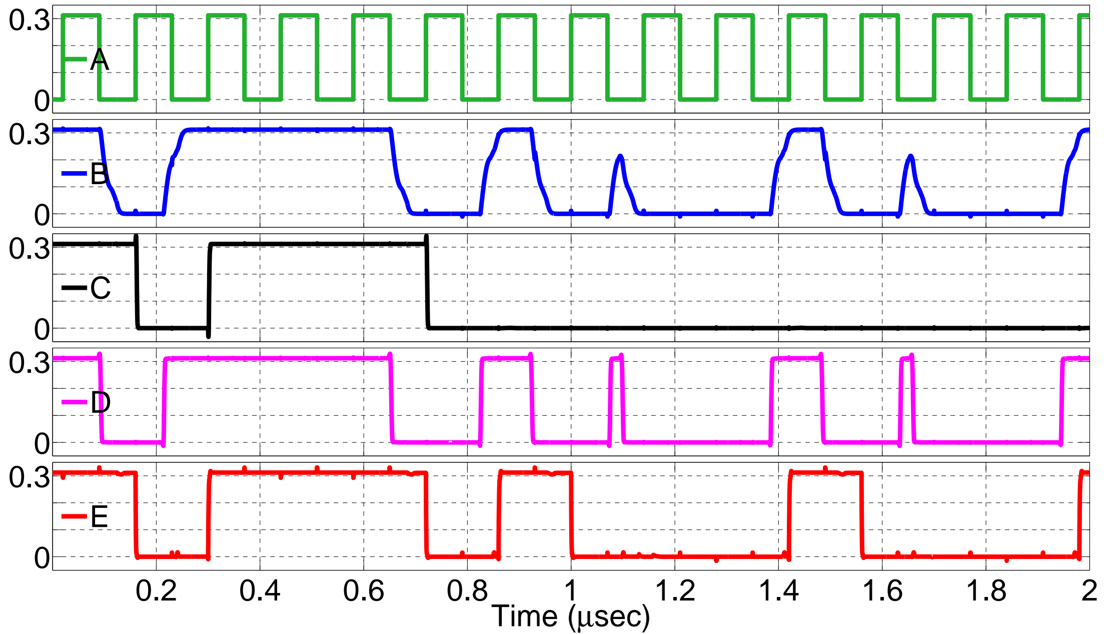

Figure 5·5: Comparison between the timing waveforms of the non-equalized logic design and the equalized logic design of a 64-bit adder. Here, the waveforms include the clock signal (A), input node of the conventional flip flop (B), output node of the conventional flip flop (C), input node of the equalized flip flop (D), output node of the equalized flip flop (E). Feedback circuit enables sharper transitions in the waveforms of the combinational logic output node helping the equalized flip flop sample the correct data. Here the feedback path 2 is OFF.

node of the non-equalized flip flop (NE-flip flop), the input node of the equalized flip flop (E-flip flop) and the flip flop output for both cases. Compared to the signal at the input node of the non-equalized flip flop, the variable threshold circuit enables sharper transitions and decreases the propagation delay of the critical path of the sub-threshold logic.

However, it should be noted that the equalized flip flop might sample the glitches due to the change in switching threshold. In order to avoid sampling of the glitch by the equalized flip flop, the positive edge of the clock signal should arrive after the occurrence of the glitch. Moreover, the switching threshold of the adaptive feedback equalizer circuit should still be larger than the amplitude of the glitch. This

would specify the maximum allowable feedback strength of the adaptive feedback equalization technique (maximum tolerable glitch amplitude shown in Figure 5·6). The sampling of a glitch leads to the marginal increase in the dynamic energy of the sequential logic block (0.72% increase in the 64-bit adder), but it has a negligible impact on the overall energy consumption as it is not the dominant energy component in the sub-threshold regime. The feedback equalizer circuit also reduces the pulse width of the glitch (by 41%). This decreases the required guardband in the clock period to avoid sampling the glitch (and hence we can reduce the clock period), which ultimately reduces the dominant leakage energy component of the sub-threshold logic block by 5.1% in the 64-bit adder at minimum-energy supply voltage.

To avoid the meta-stability problem in the equalized flip flop, both the setup time and hold time constraints should be satisfied. The setup time and the clk-to-q delay of the adaptive equalized flip flop is larger than that of the classic master-slave positive edge-triggered flip flop. However, the feedback equalizer circuit can lower down the propagation delay of the critical path since it significantly reduces the switching time of the last gate in the combinational logic. Therefore, if we match the clock period for both NE-logic and E-logic then the setup time condition is easily met. In fact, it should be noted that E-logic enables a reduction in the clock period.

The hold time constraint of the flip flop is as follows:

$$t_{hold} < t_{cdFF} + t_{cdlogic} \tag{5.3}$$

where t_{cdFF} is the minimum propagation delay of the flip flop and $t_{cdlogic}$ is the minimum propagation delay of logic. The hold time of the equalized flip flop is zero. So the hold time constraint is also fulfilled which insures the stability of feedback equalizer circuit in sub-threshold regime.

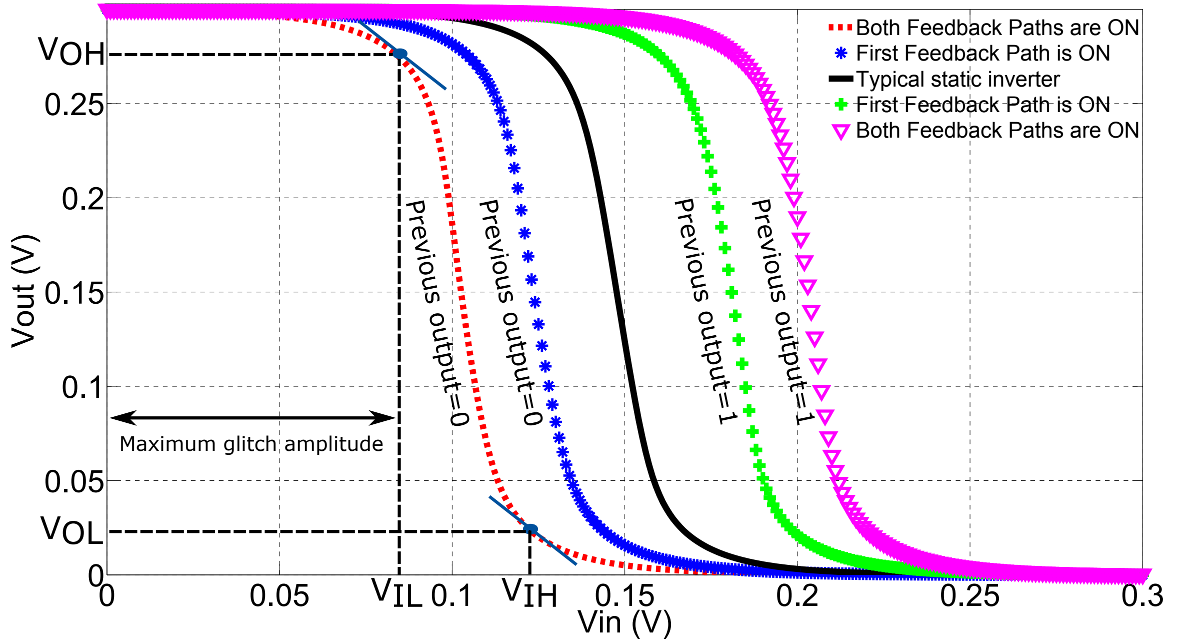

Figure 5·6: Maximum feedback strength in adaptive equalized flip flop. The switching threshold of the adaptive equalized flip flop should be larger than the maximum amplitude of the glitch.

5.3 Modeling of Feedback Equalizer Circuits

In this section, we present detailed analytical models for the performance and the energy of adaptive equalizer circuits operating in sub-threshold regime. Using these models, we determine the sizes for feedforward transistors and control transistors in the feedback equalizer circuit that minimize total delay and leakage energy for the equalized sub-threshold logic. Without loss of generality, we choose minimum-sized transistors for matching high-to-low and low-to-high propagation delay in the static inverter of the feedback equalizer circuit. As part of the effort, we first develop an analytical methodology to calculate the equivalent channel resistance of active MOSFET devices operating in sub-threshold regime. The proposed model is validated against HSPICE simulations using UMC 130 nm process.

The average channel resistance of MOSFET devices in sub-threshold regime can be

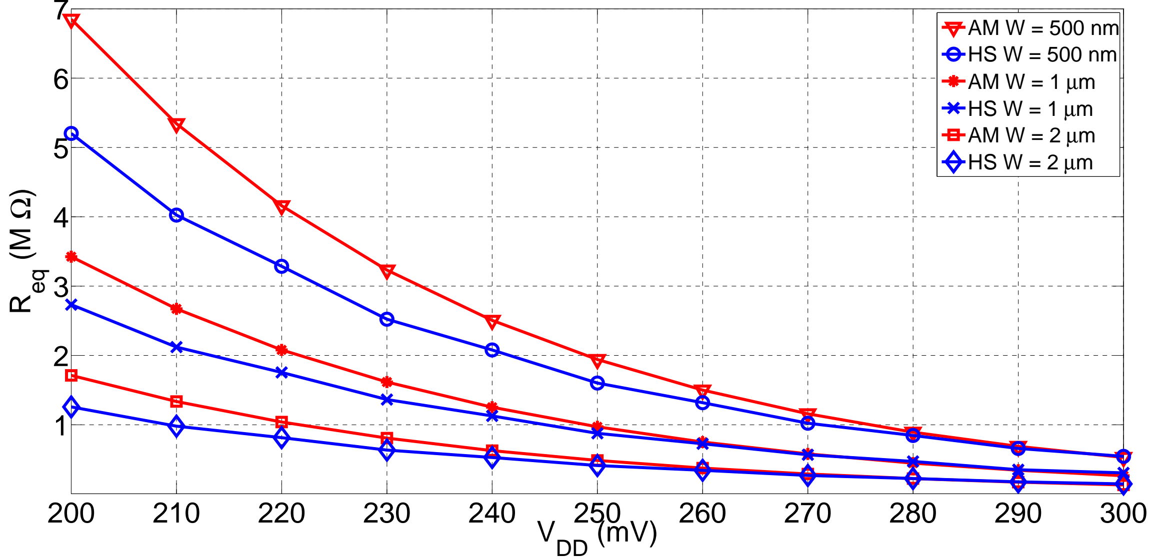

Figure 5·7: Comparison between analytical model (AM) and HSPICE simulations (HS) for equivalent channel resistance of MOSFET devices operating in sub-threshold regime. The average error between the derived model and HSPICE simulation results is 6.96% in the entire sub-threshold regime.

approximated as

$$R_{eq} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} R_{on}(t)\, dt = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{V_{DS}(t)}{I_D(t)}\, dt \tag{5.4}$$

where $R_{on}(t)$ is the finite switching resistance, $V_{DS}(t)$ is the drain to source voltage and $I_D(t)$ is the drain current. Assuming for the case of an NMOS discharging a load capacitor from V_{DD} to $V_{DD}/2$ (this is virtually the definition of the propagation delay), we can derive the value of the equivalent resistance using:

$$R_{eq} = \frac{1}{V_{DD}/2} \int_{V_{DD}/2}^{V_{DD}} \frac{v}{I_D}\, dv \tag{5.5}$$

Here, v is the auxiliary variable which accounts for the change in the V_{DS} voltage.

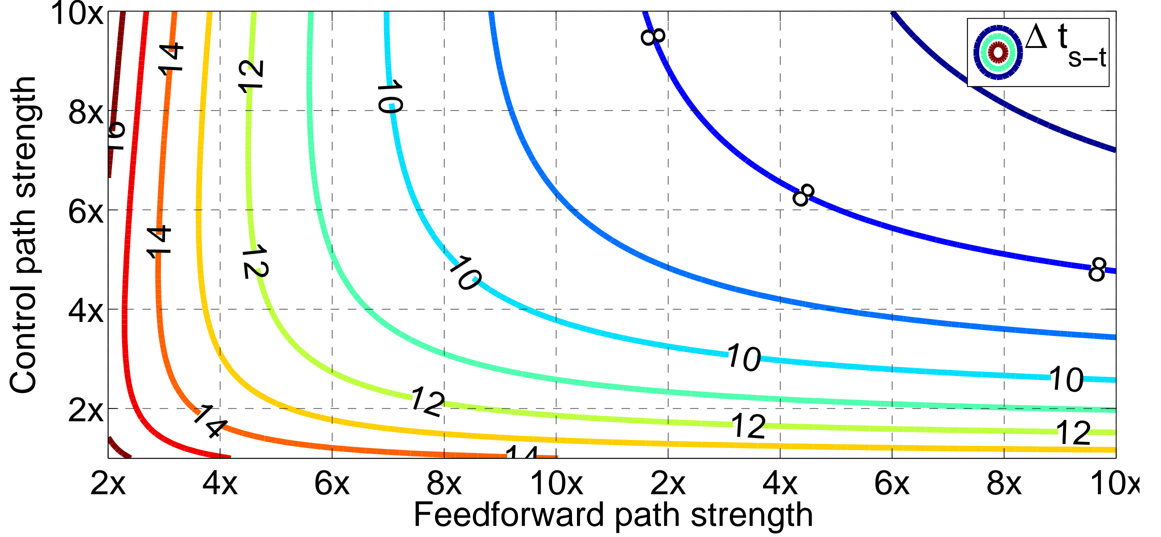

Figure 5·8: Contour plots for the Δt_{s-t} (ns) of the adaptive equalized flip flop. Control path strength and feedforward path strength values are normalized to minimum-sized transistor sizes.

The equivalent channel resistance in Equation (5.5) can be approximated as

$$R_{eq} \approx \frac{1}{I_0 \times V_{DD}/2} \int_{V_{DD}/2}^{V_{DD}} \frac{v}{1 - e^{-v/V_{th}}} \, dv \tag{5.6}$$

where the constant $I_0 = \mu_0 C_{ox} \frac{W}{L} (n-1) V_{th}^2 e^{\frac{V_{DD}-V_T}{nV_{th}}}$. The equivalent channel resistance in Equation (5.6) is valid for the case where the rise time of the input signal is smaller than the propagation delay of the logic gate in sub-threshold regime. Figure 5·7 compares the channel resistance of NMOS devices operating in sub-threshold regime calculated using Equation (5.6) with HSPICE simulations for 3 different channel widths using UMC 130 *nm* process. The average error between the derived model and HSPICE simulation is 6.96% in the entire sub-threshold regime.

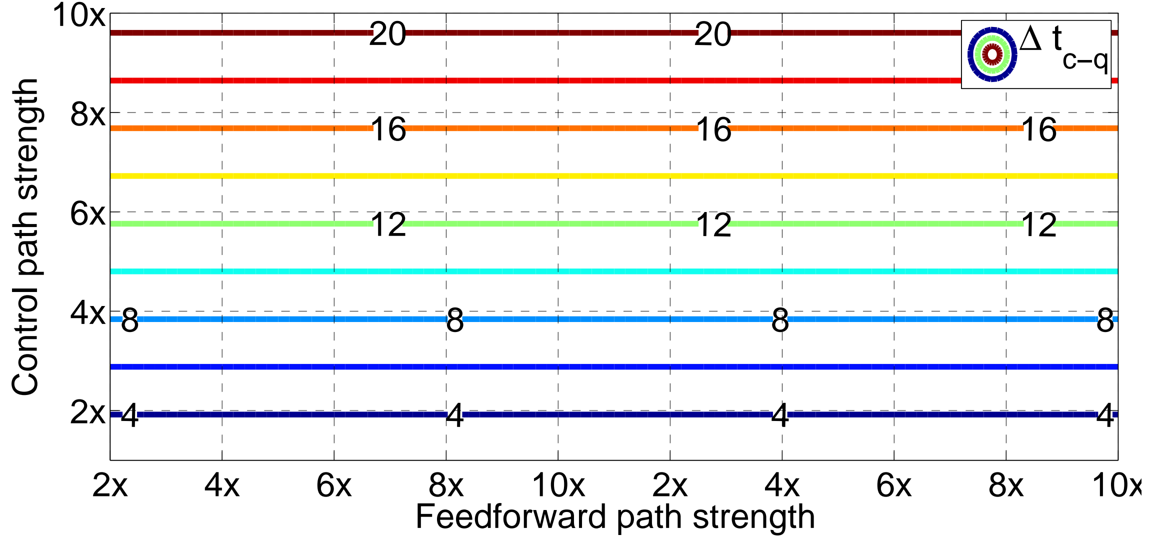

Figure 5·9: Contour plots for the Δt_{c-q} (ns) of the adaptive equalized flip flop. Control path strength and feedforward path strength values are normalized to minimum-sized transistor sizes.

The clock period constraint of a typical sequential digital logic block can be written as:

$$t_{clk} > t_{PD} + t_{s-t} + t_{c-q} \tag{5.7}$$

where t_{clk} is the clock period, t_{PD} is the propagation delay of logic, t_{s-t} is the setup time and t_{c-q} is the clk-to-q delay of the flip flop. In an equalized sequential logic block, the propagation delay of the equalized logic can be written as:

$$t_{PD-equ} = t'_{PD} + 0.69R_{out} \times C_{in-equ} \tag{5.8}$$

where t'_{PD} is the propagation delay of the combinational logic part excluding the final

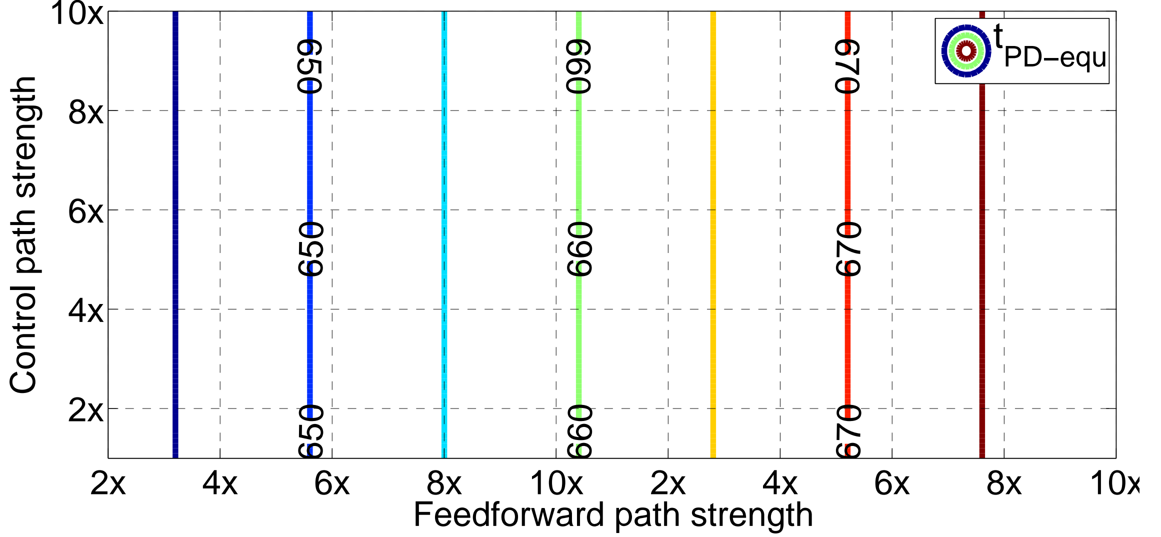

Figure 5·10: Contour plots for the t_{PD-equ} (ns) of the critical path in the equalized logic (64-bit adder). Control path strength and feedforward path strength values are normalized to minimum-sized transistor sizes.

gate, R_{out} is the output resistance of the final gate in the critical path of non-equalized logic, C_{in-equ} is the input capacitance of the feedback equalizer circuit and can be written as (see Figure 4·1):

$$C_{in-equ} = C_{stat-inv-g} + C_{M1-g} + C_{M2-g} \tag{5.9}$$

In Equation (5.9), $C_{stat-inv-g}$ is the input capacitance of the static inverter, C_{M1-g} and C_{M2-g} are the gate capacitance of feedforward transistors. The setup time of the equalized flip flop can be written as $t_{s-t-equ} = t_{s-t} + \Delta t_{s-t}$ where t_{s-t} is the setup time of the conventional non-equalized flip flop and Δt_{s-t} is due to the equalization overhead. Δt_{s-t} for a falling transition can be written as Equation (5.10):

$$\Delta t_{s-t} = 0.69[R_{M1} \times (C_{M1-d} + C_{M3,5-d}) + (R_{stat-inv}||(R_{M1} + R_{M3(5)})) \times C_T] \quad (5.10)$$

where $R_{stat-inv}$ and R_{M1} are the equivalent resistance of the typical static inverter and feedforward transistor, respectively. $R_{M3(5)}$ is the equivalent resistance of the control transistor for feedback path 1 or is the equivalent resistance of the control transistors for both feedback paths (if the second path is activated). C_{M1-d} is the drain junction capacitance of feedforward transistor, $C_{M3,5-d}$ is the junction capacitance of control transistor for feedback path 1 and 2, and $C_T = (C_{stat-inv-d} + C_{M3,4,5,6-d} + C_{in-FF})$ is the total capacitance at the output node of the variable threshold inverter. Here, $C_{stat-inv-d}$ is the drain junction capacitance of typical static inverter, $C_{M3,4,5,6-d}$ is the drain junction capacitance of control transistors for feedback path 1 and 2 and C_{in-FF} is the input capacitance of conventional non-equalized flip flop.

As it was mentioned in Section 4.3, the clk-to-q delay of the equalized flip flop is $t_{c-q-equ} = t_{inv} + \Delta t_{c-q} + t_{TG}$ where Δt_{c-q} is the increase in inverter delay due to the extra load of the adaptive feedback equalizer circuit. The Δt_{c-q} in the equalized flip flop can be written as Equation (5.11):

$$\Delta t_{c-q} = 0.69[R_{out-FF} \times (C_{M7,8-d} + C_{M3,4-g}) + (R_{out-FF} + R_{M7}) \times C_{M5-g} + R_{out-FF} \times C_{M6-g}] \quad (5.11)$$

Here, R_{out-FF} is the output resistance of non-equalized flip flop, R_{M7} and $C_{M7,8-d}$ are the equivalent resistance and drain/source capacitance of the M7 and M8 transistor

switches which enable/disable the control transistors. $C_{M3,4-g}$ and $C_{M5,6-g}$ are the gate capacitance of control transistors for feedback path 1 and 2, respectively. The total gate capacitance of the MOSFET in sub-threshold regime is size-dependent and can be written as (Sarpeshkar, 2010):

$$C_g = WC_{gso} + WC_{gdo} + WLC_{ox}(1 - 1/n) \tag{5.12}$$

where C_{gso} and C_{gdo} are the overlap capacitance per unit length at the source and drain, respectively and n is the sub-threshold slope factor. The total source or drain junction capacitance of the MOSFET in sub-threshold regime can be written as:

$$C_j = A.C_1 + (W + 2L_D).C_2 \tag{5.13}$$

where A represents the source or drain diffusion areas, C_1 represents the capacitance per unit area from the bottom of the source/drain diffusion region pointing into the bulk, C_2 is the capacitance per unit length of the sidewall regions, L_D is the length of the diffusion regions and $W + 2L_D$ represents the perimeter of the side wall.

To better understand the timing issues in equalized logic, the contour plots for the Δt_{s-t}, Δt_{c-q} of the adaptive equalized flip flop and t_{PD-equ} of the critical path in an equalized 64-bit adder designed in UMC 130 nm process are illustrated in Figure 5·8, 5·9 and 5·10, respectively. The contour plots are for different strengths of feedforward path and control path (normalized to minimum-sized transistor) of the feedback equalizer circuit. For this analysis we assume that only feedback path 1 is ON. From the delay models described in Equations (5.10) and (5.11), we can see that increasing the size of feedforward and control transistors (i.e. feedback strength) reduces the Δt_{s-t} overhead of the equalized flip flop. However the increase in the control path strength increases the Δt_{c-q} overhead (due to larger control transistors - M3, M4,

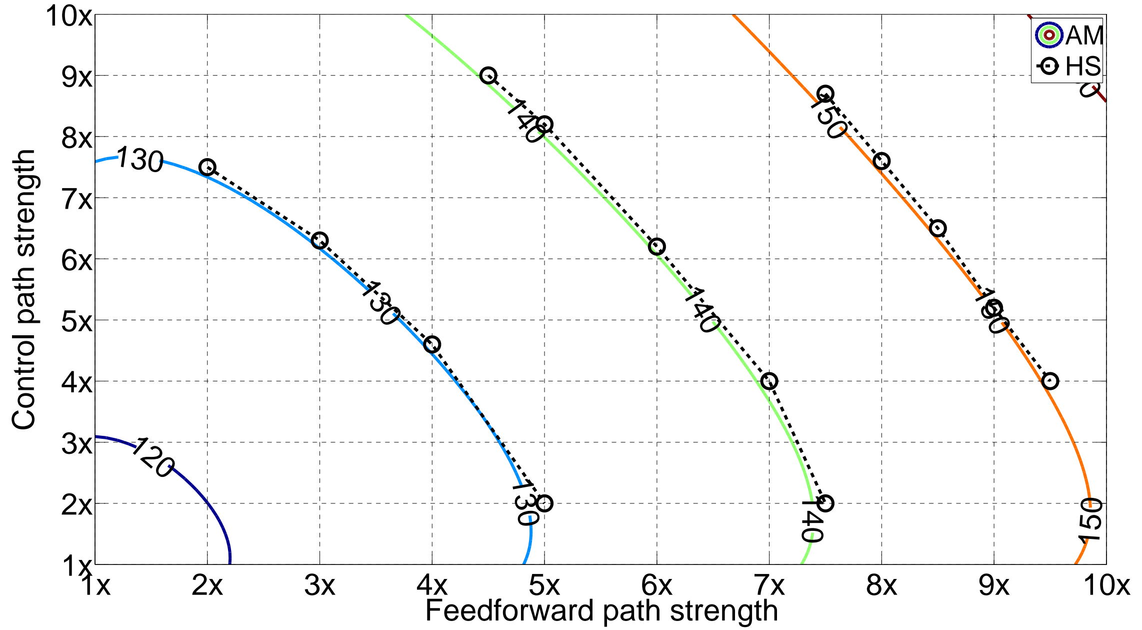

Figure 5·11: Comparison between analytical model (AM) contour plots for the total delay (ns) of the critical path in an equalized 64-bit adder with HSPICE simulations (HS).

M5 and M6) of the equalized flip flop. The change in the feedforward path strength does not have any impact on the clk-to-q delay (see Equation (5.11)). Similarly, the increase in the feedforward path strength increases the propagation delay of the logic (due to larger feedforward transistors - M1 and M2) and correspondingly increases the total delay of the critical path. The change in the control path strength does not have any impact on the critical path delay (see Equation (5.8)).

The contour plots for the total delay calculated from the analytical models of the different delay components for an equalized 64-bit adder designed in UMC 130 nm process are illustrated in Figure 5·11. The total delay is plotted for different normalized strengths of feedforward path (M1 and M2) and control path (M3, M4, M5 and M6) of the feedback equalizer circuit. Figure 5·11 also shows the total delay values from HSPICE simulations for various combinations of feedforward and control path strength. We can see that our models match well with HSPICE simulations. In

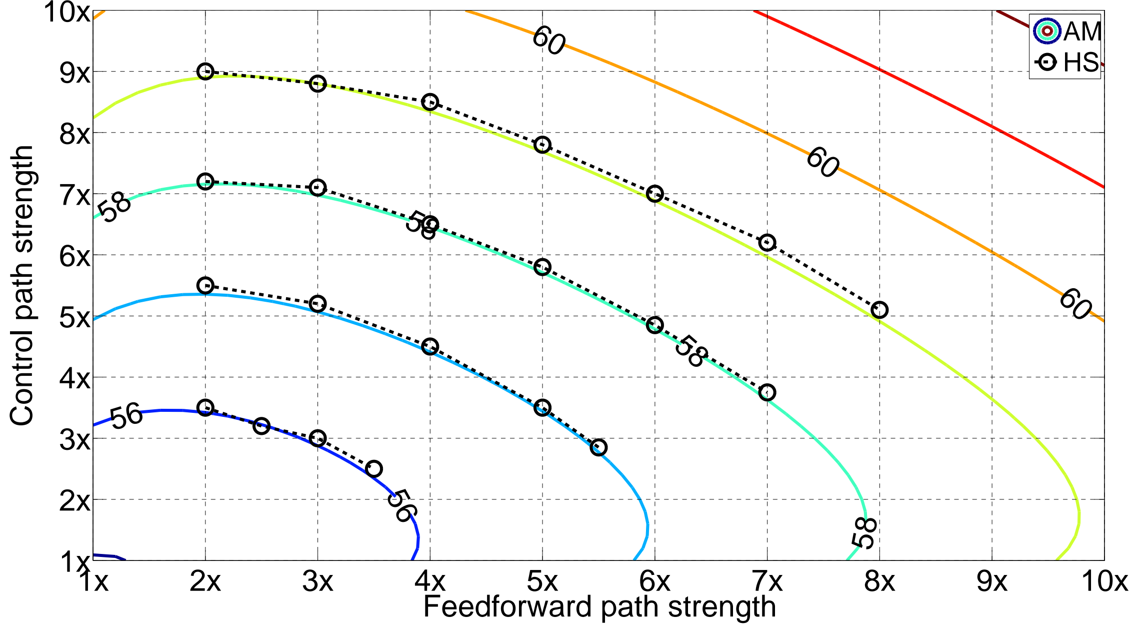

Figure 5·12: Comparison between analytical model (AM) contour plots for the total energy (fJ/operation) of the equalized 64-bit adder with HSPICE simulations (HS).

addition, Figure 5·11 shows that choosing the minimum possible size for the feedforward and control transistors will lead to the minimum latency for the equalized logic designed in sub-threshold regime.

The total energy consumed in the E-logic circuit can be calculated as

$$E'_T = E_T + E'_{leak} + E'_{dyn} \tag{5.14}$$

where E_T is the energy consumption of the E-logic circuit excluding the feedback equalizer circuit and can calculated using Equation (5.1). E'_{leak} is the leakage energy in the feedback equalizer circuit and can be calculated as $I'_{leak}V_{DD}T_{D-equ}$ where T_{D-equ} is the total latency of the equalized logic and can be written as $T_{D-equ} = t_{PD-equ} + t_{s-t-equ} + t_{c-q-equ}$ and I'_{leak} is the leakage current overhead of the adaptive feedback equalizer circuit and can be calculated as

$$I'_{leak} = \mu_0 C_{ox} \frac{\Sigma W_i}{L} (n-1) V_{th}^2 e^{\frac{\eta V_{DS} - V_T}{n V_{th}}} \tag{5.15}$$

Here, ΣW_i is sum of the widths for all of the transistors in the adaptive feedback equalizer circuit. The dynamic energy of the adaptive equalizer circuit (E'_{dyn}) can be calculated as $\Sigma C_{eff}(W_i)V_{DD}^2$ where $\Sigma C_{eff}(W_i)$ is the total parasitic capacitance due to all the transistors of the feedback equalizer circuit. A comparison between the analytical model contour plots for the total energy of the equalized 64-bit adder in UMC 130 *nm* process with HSPICE simulations is illustrated in Figure 5·12. The leakage energy component is directly proportional to the latency of the sub-threshold logic. Therefore using larger feedforward and control transistors increases the dominant leakage energy component of the digital logic in sub-threshold regime.

5.4 Evaluation

In this section, using a 64-bit adder designed in UMC 130 *nm* process as a sample circuit, we first explore the use of the feedback equalizer circuit to reduce energy consumption while maintaining reliable operation of the 64-bit adder. This is followed by the evaluation of the post-fabrication tunability property of the adaptive equalizer circuit to manage the occurrence of worse than expected process variations in the 64-bit adder circuit after fabrication. In addition we provide an evaluation of the use of feedback equalizer circuit in the 64-bit adder designed using aggressive technology nodes.

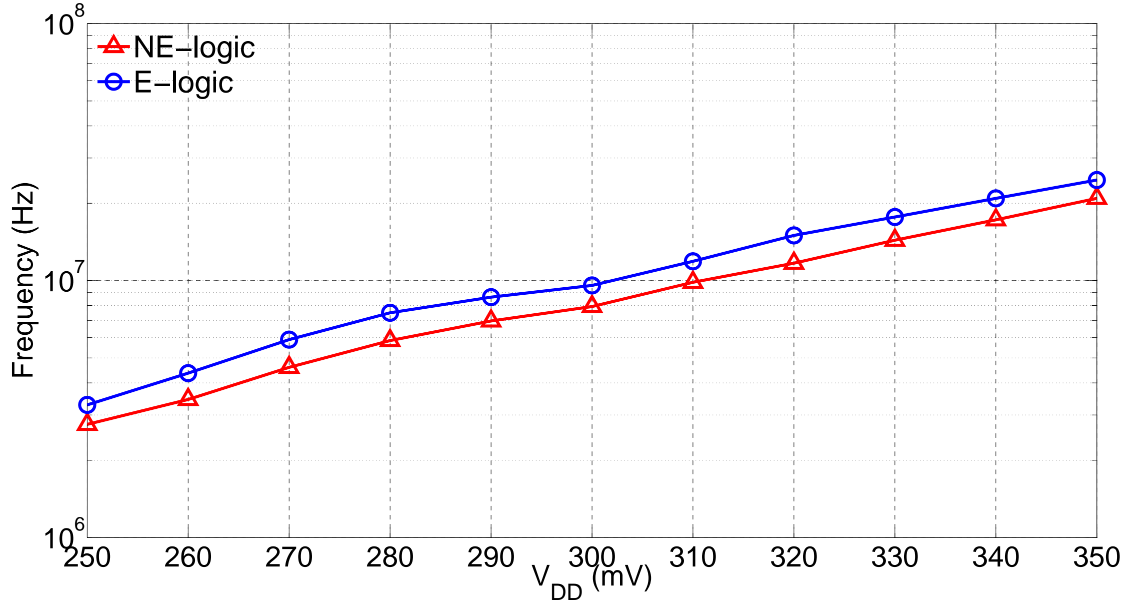

Figure 5·13: Operating frequency of the 64-bit adder for zero word error rate as function of different sub-threshold supply voltages. The equalized logic (E-logic) can run 18.91% (on average) faster than the non-equalized logic (NE-logic).

5.4.1 Improvement of Energy Efficiency

We first explore the case where the feedback equalizer circuit reduces the rise/fall time of the last gate and hence the delay of the critical path of the combinational logic block leading to a higher operating frequency without any change in supply voltage. In general, the variable threshold inverter can be used to reduce the propagation delay of the critical path at any operating supply voltage. Figure 5·13 shows the operating frequency of the 64-bit adder for different sub-threshold supply voltages at zero error rate for equalized logic (E-logic) and non-equalized logic (NE-logic) when only the first feedback path is ON. Here, we determined the optimum sizing for the feedback equalizer circuit that minimizes the propagation delay of the critical path and avoids sampling of glitches to achieve zero error rate operation at each supply voltage. The sizing of the combinational logic block is the same for both the E-logic and NE-logic

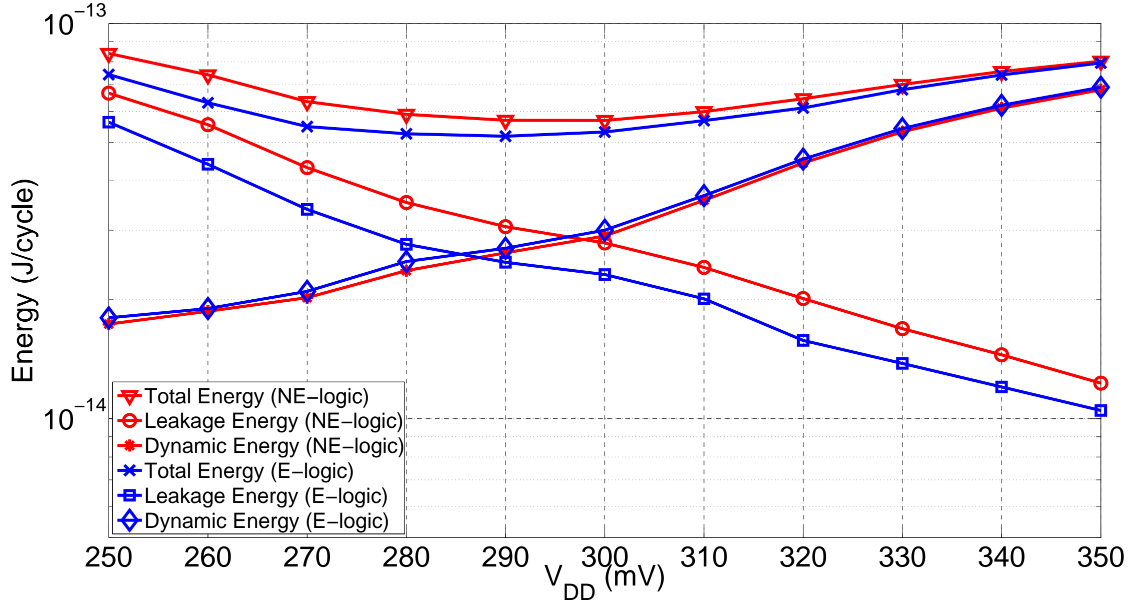

Figure 5·14: Comparison between the total consumed energy as well as the dynamic/leakage components of the 64-bit adder for different supply voltages. Operating at the respective minimum energy supply voltage, the equalized logic is burning 10.85% less total energy compared to the non-equalized logic.

and is determined using the design methodology described in (Kwong et al., 2009) (assuming $\sigma_{V_T} = 10$ mV) to address the degraded noise margin levels in sub-threshold regime. The operating frequency of the equalized logic is 18.91% (on average) higher than the non-equalized logic over the range of 250 mV to 350 mV.

By reducing the propagation delay of the critical path, the feedback equalizer circuit is capable of reducing the dominant leakage energy consumption of the digital logic in sub-threshold regime. Figure 5·14 illustrates a head-to-head comparison between the total energy, the dynamic energy and the leakage energy of the 64-bit adder for different supply voltages for the E-logic and NE-logic. By adding the feedback equalizer to the conventional flip flop, the dynamic energy of the E-logic is 2.69% (on average) larger than the NE-logic. This is negligible compared to the 18.5%

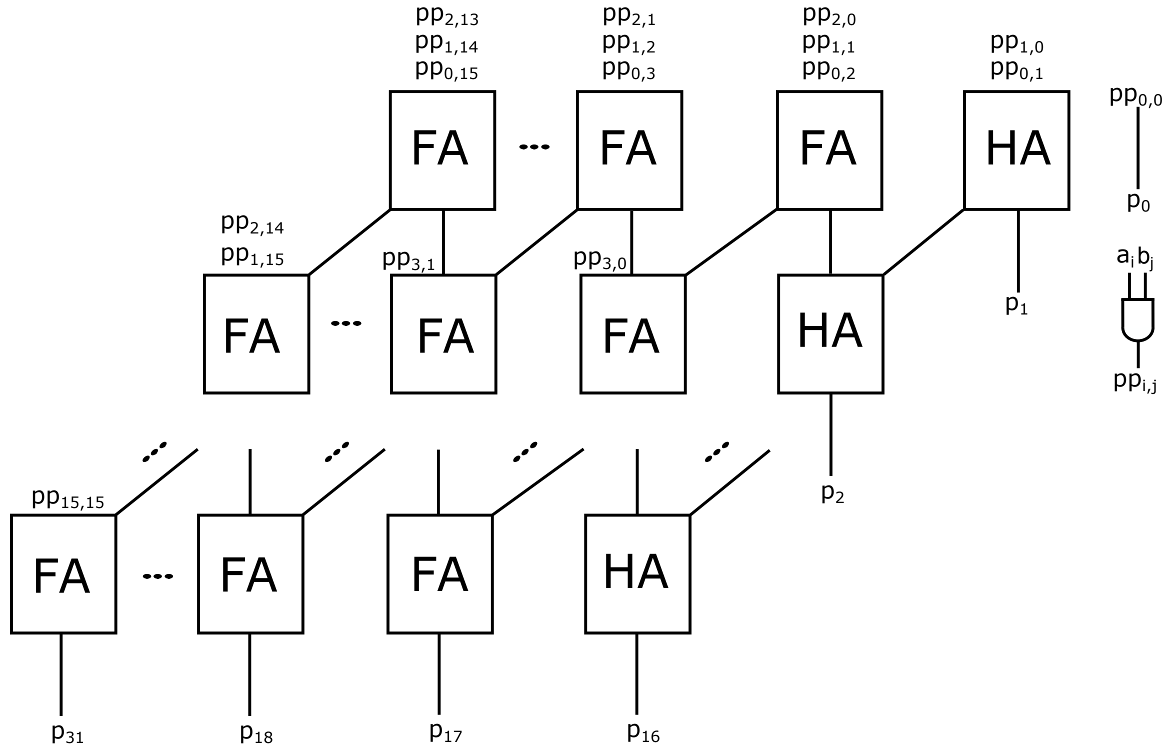

Figure 5·15: Block diagram of the 32-bit Array Multiplier.

reduction in the leakage energy (on average) of the design. The feedback circuit drops the minimum energy supply voltage of the E-logic by 10 mV while maintaining the zero error rate operation. If operated at the respective minimum energy supply voltage, the E-logic consumes 10.85% less total energy compared to the NE-logic and runs 8.04% faster. If both designs are operated at the minimum energy supply voltage of the NE-logic, the E-logic runs 19.1% faster and consumes close to 10% less energy.

By decreasing the dominant leakage energy component of the sub-threshold logic together with reducing the propagation delay of the critical path, the feedback equalization technique lowers the energy-delay product of the logic designed in weak inversion region. On average, the E-logic design of the 64-bit adder has 24.4% smaller energy-delay product value compared to the NE-logic design over the range of 250 mV to 350

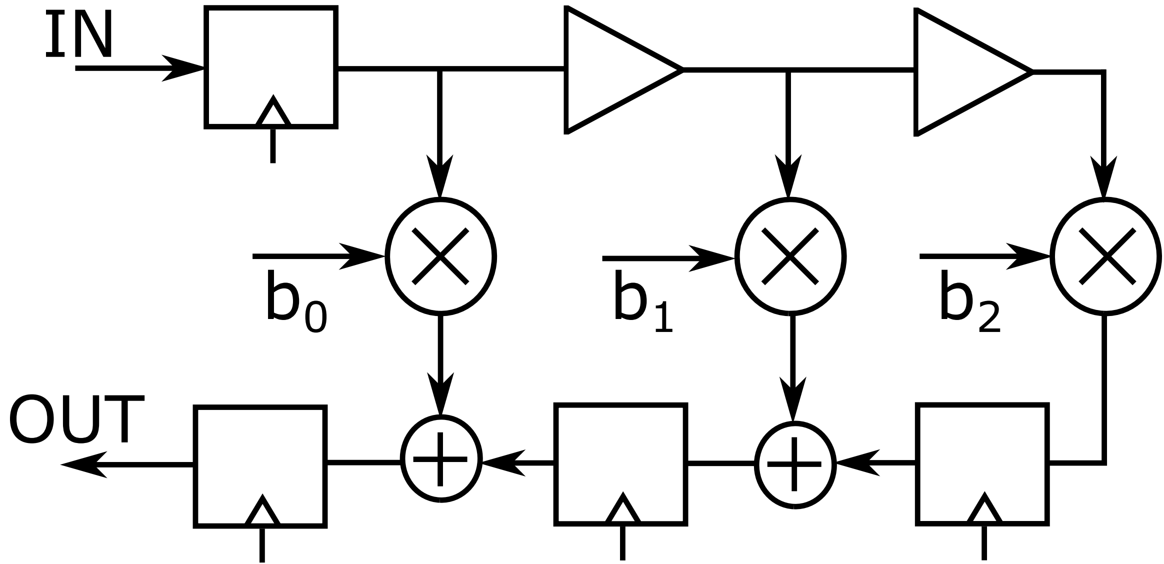

Figure 5·16: Block diagram of the 3-tap 16-bit finite impulse response (FIR) filter.

mV for zero word error rate operation. If we compare the energy-delay product at the respective minimum energy supply voltages, the use of equalized flip flop reduces the energy-delay product of the 64-bit adder by 25.83%. To further evaluate the viability of E-logic, we consider a 32-bit Array Multiplier and a 3-tap 16-bit Finite Impulse Response (FIR) filter. In general, our methodology will be applicable to other types of binary multipliers such as Wallace tree multiplier, Dadda multiplier, etc and other digital signal processing blocks with similar improvements. The block diagram of the 32-bit Array Multiplier and the 3-tap 16-bit FIR filter are illustrated in Figures 5·15 and 5·16, respectively. Table 5.2 compares the minimum energy point and the corresponding operating frequency of E-logic design vs. NE-logic design of a 64-bit Adder, 32-bit Array Multiplier and 3-tap 16-bit FIR filter all designed using Cadence Encounter in UMC 130 *nm* process. On an average, the E-logic design has 18.45% lower energy-delay product than the NE-logic design.

Using the proposed feedback-based technique, the critical sizing approach proposed in (Kwong et al., 2009) for designing the sub-threshold logic circuits can be relaxed

Logic block	NE-logic Energy (fJ/cycle)	E-logic Energy (fJ/cycle)	NE-logic Freq. (MHz)	E-logic Freq. (MHz)
64-bit Adder	57.1	50.9	7.69	9.52
32-bit Multiplier	319	298	3.18	3.44
16-bit FIR filter	503	470	2.78	3.01

Table 5.2: Comparison between the minimum energy point and the corresponding operating frequency of the NE-logic vs. E-logic design of various logic blocks.

Scaled-down E-logic size	Energy saving w.r.t NE-logic	Energy saving w.r.t E-logic
$95\%\times W_{baseline}$	9.63%	4.38%
$85\%\times W_{baseline}$	14.61%	9.75%
$75\%\times W_{baseline}$	19.39%	14.71%

Table 5.3: Energy savings in scaled-down E-logic compared to baseline NE-logic and E-logic at the minimum energy supply voltage with zero word error rate operation.

while ensuring the reliable operation in presence of process variations. Figure 5·17 compares the energy-delay product of the scaled down E-logic and NE-logic of the 64-bit adder in UMC 130 nm for different sub-threshold supply voltages and assuming a $3\sigma_{V_T} = 30mV$ systematic variability in threshold voltage. Here, the transistors sized using (Kwong et al., 2009) ($W_{baseline}$) for the NE-logic can be scaled down to $75\%\times W_{baseline}$ when using E-logic while ensuring reliable operation (no timing errors) at any given voltage. As a result the dynamic energy of E-logic decreases due to decrease in the transistor parasitic capacitances. For a given supply voltage all E-logic designs are operated at the same frequency. The E-logic with transistor sizing smaller than 75% of $W_{baseline}$ cannot operate at this frequency and has timing errors. Table 5.3 summarizes the amount of energy savings of the E-logic with scaled down transistors compared to the NE-logic and E-logic. Overall the feedback equalization along with transistor size scaling consumes up to 19.39% lower total energy compared to the NE-logic in sub-threshold regime.

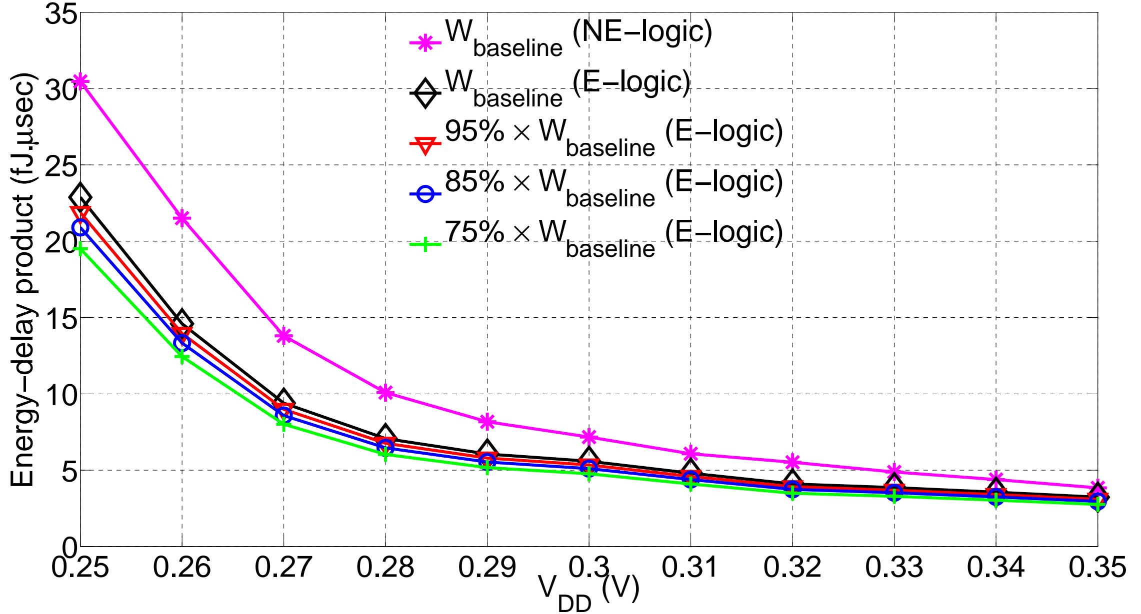

Figure 5·17: Energy-delay product of the scaled-down equalized 64-bit adder for zero word error rate operation. We can achieve reliable operation even when the transistors in the equalized logic design are scaled down to as small as $75\% \times W_{baseline}$.

5.4.2 Maintaining Robustness Using Post-Fabrication Tuning

In this section, we explore the use of the adaptive feedback equalizer circuit to mitigate worse than expected process variations. As described earlier, adaptive feedback equalizer circuit dynamically modifies the switching threshold of the inverter driving the flip flop and at the same time the smaller input capacitance of the feedback equalizer reduces the switching time of last gate in the combinational logic. This reduces the standard deviation σ of the total delay in the critical path. Figure 5·18 illustrates the distribution of total delay of the critical path in the 64-bit adder designed in UMC 130 nm process for different standard deviation values of threshold voltage. The delay distributions are shown for the NE-logic, for the buffer-inserted NE-logic, for the E-logic when only one feedback path is ON (1-FB) and when both feedback paths are ON (2-FB). The sizing of the combinational logic block is the

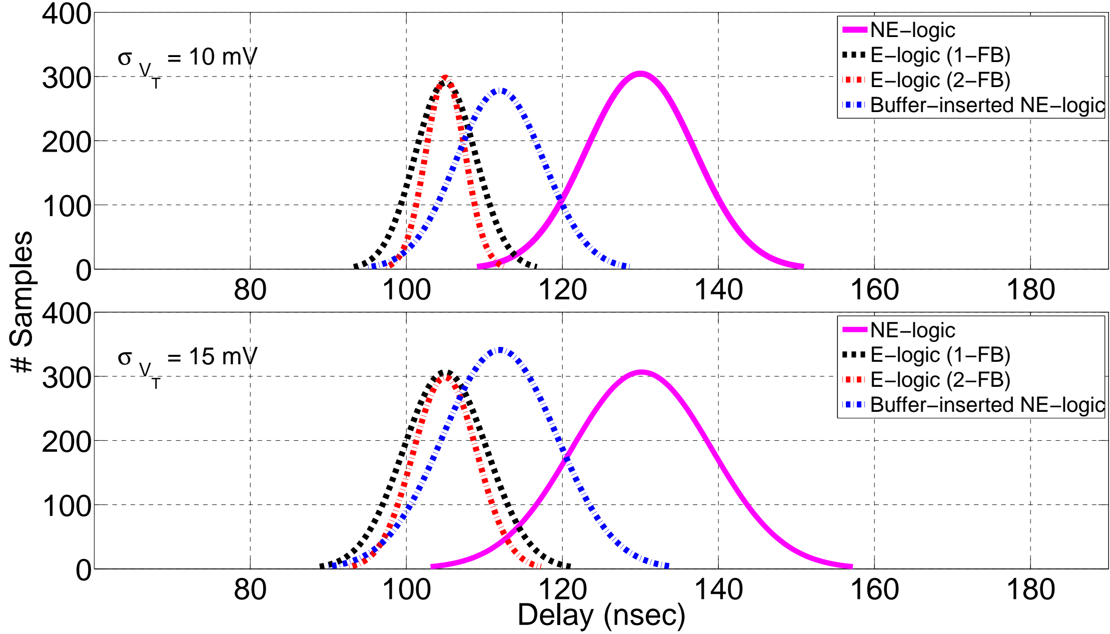

Figure 5·18: Delay distribution of the critical path in the 64-bit adder designed in UMC 130 nm process. The 3 $\times\sigma/\mu$ of the non-equalized logic (NE-logic), the equalized logic (E-logic) with 2 different feedback strengths and the buffer-inserted NE-logic are 16.1%, 11.4%, 7.14% and 15% for σ_{V_T} = 10 mV at the minimum energy supply voltage, respectively. Here, E-logic designs are operating at 300 mV.

same for both the E-logic and NE-logic and is determined using the design methodology described in (Kwong et al., 2009) and assuming σ_{V_T} = 10 mV. Considering $\Delta V_T = 3 \times \sigma_{V_T}$ = 30 mV variation in the threshold voltage of the transistors, the normalized delay variation (3 $\times\sigma/\mu$) of the NE-logic, E-logic (1-FB), E-logic (2-FB) and the buffer-inserted NE-logic are 16.1%, 11.4%, 7.14% and 15% respectively at the minimum energy supply voltage. Both the equalized designs have lower delay and lower total energy than the NE-logic designs. Between the two equalized designs, the E-logic (2-FB) design has lower normalized delay variation due to the extra pull-up/pull-down path in the feedback equalizer circuit. However, it has higher energy consumption due to more parasitics and higher dynamic/leakage energy components. Table 5.4 provides a head-to-head comparison of normalized delay variation, energy

Method	σ_{V_T} (mV)	$3\sigma/\mu$ (delay)	Energy (fJ/cycle)	Delay (ns)
NE-logic Upsized (Kwong et al., 2009)	10	16.1%	57.1	130
Buffer-inserted NE-logic	10	15%	53	113
E-logic Upsized + 1-FB	10	11.4%	50.9	105
E-logic Upsized + 2-FB	10	7.14%	52.4	105
NE-logic Upsized (Kwong et al., 2009)	15	20.8%	57.1	130
Buffer-inserted NE-logic	15	19.6%	53	113
E-logic Upsized + 1-FB	15	16%	50.9	105
E-logic Upsized + 2-FB	15	11.7%	52.4	105

Table 5.4: Comparison between the total delay, total energy and delay variation of the digital logic (64-bit adder) at minimum energy supply voltage when the conventional upsizing method (Kwong et al., 2009) has been used together with adaptive feedback equalizer circuit in sub-threshold regime.

and delay of NE-logic, buffer-inserted NE-logic, E-logic (1-FB) and E-logic (2-FB). In the E-logic design, the control latch consumes 2.43 nW on an average (Zangeneh and Joshi, 2015).

In our feedback equalizer circuit, we propose that the second feedback path is switched ON post fabrication if the σ_{V_T} variations are worse than expected. The second feedback path compensates for the increase in the variation in logic path delays due to worse than expected σ_{V_T} variations and reduces the normalized $3 \times\sigma/\mu$ of the total delay for the equalized logic. As an example, say we design a 64-bit adder using E-logic assuming a $\sigma_{V_T} = 10$ mV. With only one feedback path ON, the design has a $3 \times\sigma/\mu$ of 11.4% for the delay. If post-fabrication the σ_{V_T} is larger and is equal to 15 mV, then we can switch ON the second feedback path to achieve a $3 \times\sigma/\mu$ of close to 11.4% for the delay (see Figure 5·18 and Table 5.4). This will result in a 2.94% increase in energy. One could argue that we could design the 64-bit adder upfront to achieve a $3 \times\sigma/\mu$ for the delay that is smaller than 11.4% and that way even if σ_{V_T} is larger than expected, then we can still have a $3 \times\sigma/\mu$ closer to 11.4%. However, to do this we will need to use larger M3 and M4 transistors (see Figure 4·1) resulting in higher energy consumption in the baseline 1-FB E-logic design. So we propose that

	NE-logic (1-FB)	E-logic (2-FB)	E-logic	NE-logic (1-FB)	E-logic (2-FB)	E-logic
Logic block	$3\times\sigma/\mu$ (delay)	$3\times\sigma/\mu$ (delay)	$3\times\sigma/\mu$ (delay)	EDP (fJ.μs)	EDP (fJ.μs)	EDP (fJ.μs)
64-bit Adder	16.1%	11.4%	7.14%	7.42	5.34	5.5
32-bit Multiplier	12.2%	8.5%	6.1%	100.16	86.42	89
16-bit FIR filter	10.1%	6.7%	4.8%	180.57	156.04	160.6

Table 5.5: Comparison between the normalized delay variation and energy-delay product (EDP) of the equalized logic (E-logic) vs. non-equalized (NE-logic) and buffer-inserted non-equalized design of various logic blocks assuming $\sigma_{V_T} = 10$ mV.

the first feedback path should be designed to achieve a target $3\times\sigma/\mu$ specification for the delay for an expected σ_{V_T}. Our proposed feedback equalizer then provides the option of switching ON the second feedback path to achieve the target $3\times\sigma/\mu$ specification for the delay in case σ_{V_T} turns out to be worse than expected.

Table 5.5 compares the normalized delay variation and the energy-delay product (EDP) of the NE-logic design, buffer-inserted NE-logic, E-logic (1-FB) design and E-logic (2-FB) design of a 64-bit Adder, 32-bit Array Multiplier and 16-bit FIR filter all designed using Cadence Encounter in UMC 130 *nm* process. In each case, both the E-logic approaches have lower $3\times\sigma/\mu$ delay variation than the NE-logic. Between the two E-logic designs, E-logic (2-FB) provides more robustness (smaller $3\times\sigma/\mu$) but higher energy compared to E-logic (1-FB).

5.4.3 Mitigating Voltage/Temperature Variations

In this section we explore the use of the equalization technique to mitigate the effect of voltage and temperature variations on the performance of digital logic designed in sub-threshold regime. Figure 5·19 illustrates the distribution of total delay of the critical path in the 64-bit adder designed in UMC 130 nm process in case of supply voltage variations. The delay distributions are shown for the NE-logic, for the buffer-

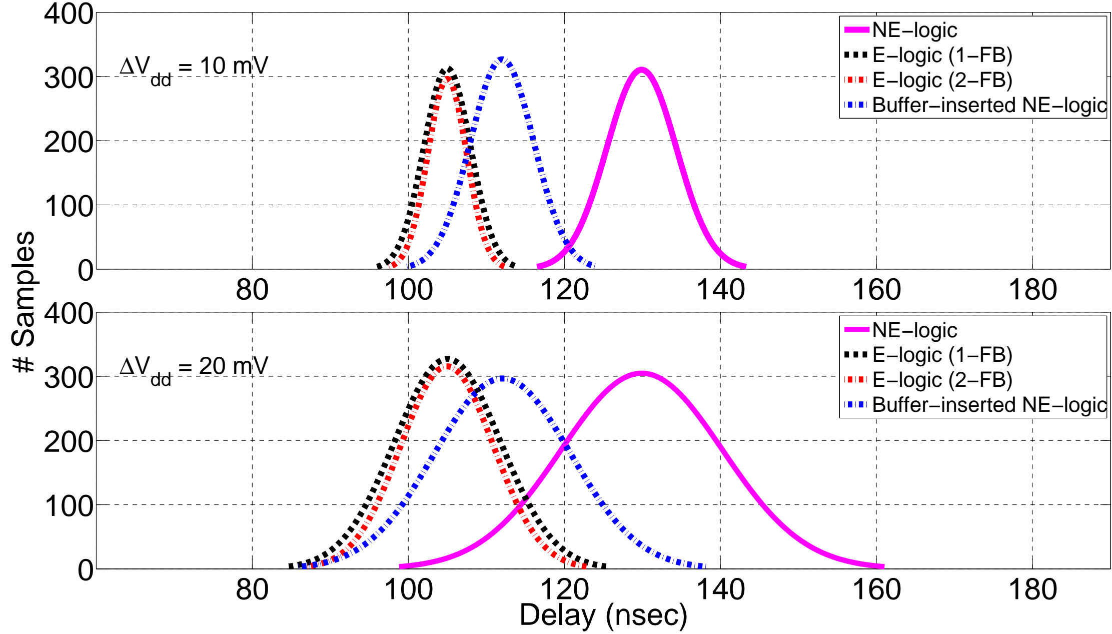

Figure 5·19: Delay distribution of the critical path in the 64-bit adder designed in UMC 130 *nm* process considering supply voltage variation.

inserted NE-logic, for the E-logic when only one feedback path is ON (1-FB) and when both feedback paths are ON (2-FB). Considering $\Delta V_{dd} = 10$ mV supply voltage variation, the feedback equalization technique reduces the worst-case delay of the sub-threshold logic by 20.44% compared to the original NE-logic (3.1% smaller $3\times\sigma/\mu$) and by 8.8% compared to the buffer-inserted NE-logic. Considering $\Delta V_{dd} = 20$ mV supply voltage variation, the feedback equalization technique reduces the worst-case delay of the sub-threshold logic by 22.23% compared to the original NE-logic (4.7% smaller $3\times\sigma/\mu$) and by 9.27% compared to the buffer-inserted NE-logic. Here, there is not much difference between the results from E-logic (1-FB) and E-logic (2-FB).

Figure 5·20 illustrates the distribution of total delay of the critical path in the 64-bit adder designed in UMC 130 nm process in case of temperature variations. The delay distributions are shown for the NE-logic, for the buffer-inserted NE-logic, for the E-logic when only one feedback path is ON (1-FB) and when both feedback paths are ON (2-FB). Considering $\Delta T = 10$ K temperature variation, the feedback equal-

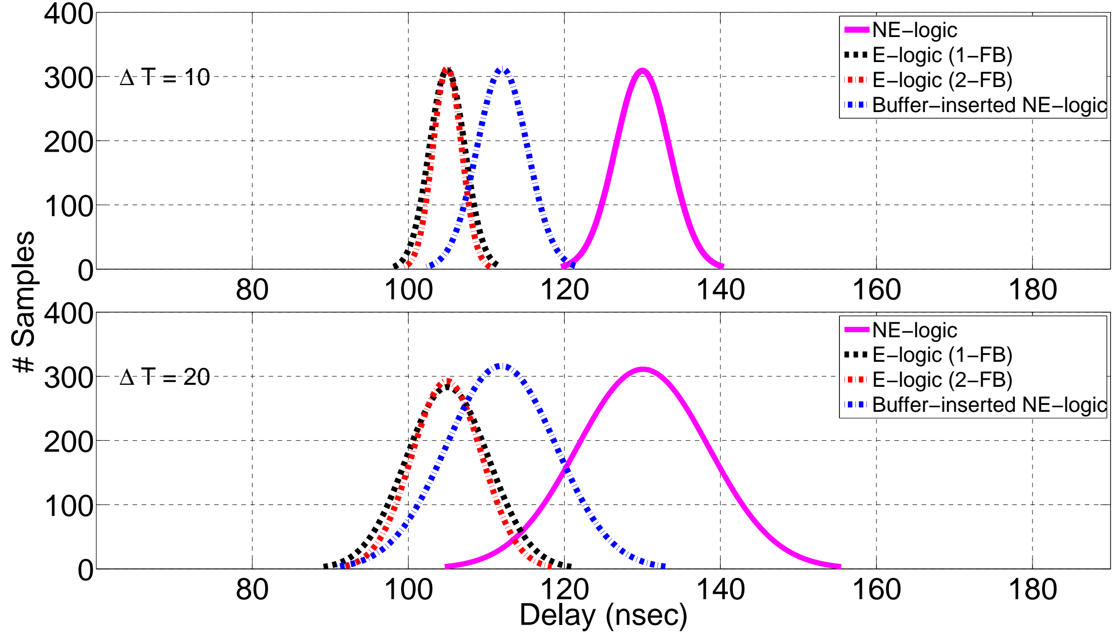

Figure 5·20: Delay distribution of the critical path in the 64-bit adder designed in UMC 130 nm process considering temperature variation.

ization technique reduces the worst-case delay of the sub-threshold logic by 21.27% compared to the original NE-logic (2.3% smaller $3\times\sigma/\mu$) and by 7.6% compared to the buffer-inserted NE-logic. Considering $\Delta T = 20$ K temperature variation, the feedback equalization technique reduces the worst-case delay of the sub-threshold logic by 22.42% compared to the original NE-logic (4.3% smaller $3\times\sigma/\mu$) and by 9.17% compared to the buffer-inserted NE-logic. Here, there is not much difference between the results from E-logic (1-FB) and E-logic (2-FB).

5.4.4 Effect of Technology Scaling

In this section, we analyze the effect of technology scaling on the performance improvement and the energy reduction that can be obtained using feedback equalization technique in sub-threshold regime. In scaled technology nodes, the contribution of leakage energy component increases due to larger DIBL effect as well as smaller V_T values. By running the sub-threshold logic faster, the feedback equalizer can reduce

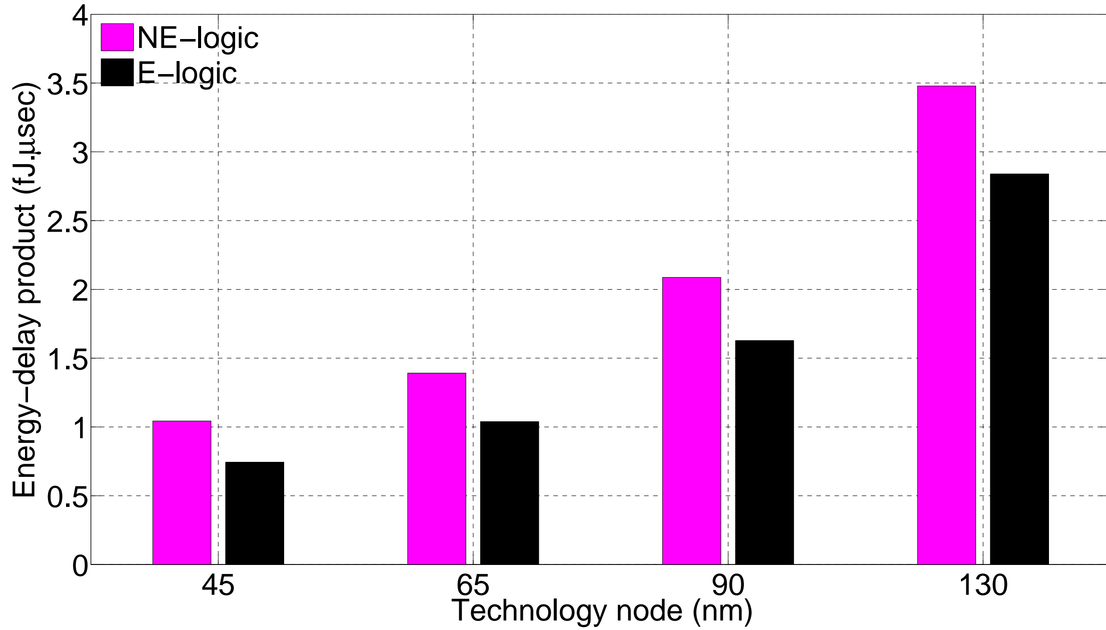

Figure 5·21: Energy-delay product of a 64-bit adder designed using equalized logic (E-logic) vs. non-equalized logic (NE-logic) at zero word error rate at different technology nodes. The equalized logic approach reduces the energy-delay product of the sub-threshold logic by up to 23.6% across all technology nodes in the minimum energy supply voltage.

the leakage energy component and in turn decrease the energy-delay product in scaled technology nodes. Figure 5·21 illustrates the value of the energy-delay product of the 64-bit adder designed using PTM (Ptm,) for 4 different technology nodes and operating at zero word error rate at minimum energy supply voltage. Here we assume the second feedback path is switched OFF. Compared to the NE-logic design, the energy-delay product of the E-logic design is 18.37%, 22.02%, 25.34% and 28.66% smaller in 130 nm, 90 nm, 65 nm and 45 nm technology nodes, respectively. On average, the equalized flip flop reduces the energy-delay product of the sub-threshold logic by 23.6% across all technology nodes when operating at their respective minimum energy supply voltages.

5.4.5 Comparison with other Sub-threshold Design Techniques

In this section, we compare different techniques proposed in (Kwong et al., 2009), (Zhai et al., 2005) and (Jayakumar and Khatri, 2005) with adaptive feedback equalizer circuit to mitigate process variations in digital sub-threshold logic circuits. Feedback equalization complements these existing techniques and can be used along with these techniques for sub-threshold circuit design. So we do not provide a head-to-head quantitative comparison for the proposed methodologies. We compare these techniques qualitatively in terms of circuit complexity, area and energy overhead. Increasing logic path depth (Zhai et al., 2005) requires additional logic gates in the critical path of the sub-threshold design to reduce the normalized (σ/μ) delay variation. This increases the parasitics and the dominant leakage energy of the design. Body-biasing (Jayakumar and Khatri, 2005) necessitates extra on-chip circuitry to generate the required voltage for the substrate terminal of the CMOS devices to reduce the dominant leakage energy of the sub-threshold logic (Tschanz et al., 2002). The upsizing design methodology proposed in (Kwong et al., 2009) increases the device parasitics which in turn increases the dynamic and leakage energy components of the entire digital sub-threshold logic block. The proposed adaptive feedback equalizer circuit has simple topology, negligible area and energy overhead and the capability to reduce the normalized delay variations post fabrication.

5.4.6 Memristor-based Feedback Equalization Technique

Chapter 4 shows the benefits of feedback equalization technique in reducing the dominant leakage energy component of digital logic circuits designed in sub-threshold regime. We explored the application of a memristor-based feedback equalization technique (shown in Figure 5·22) for adaptive tunable logic circuit design. The tun-

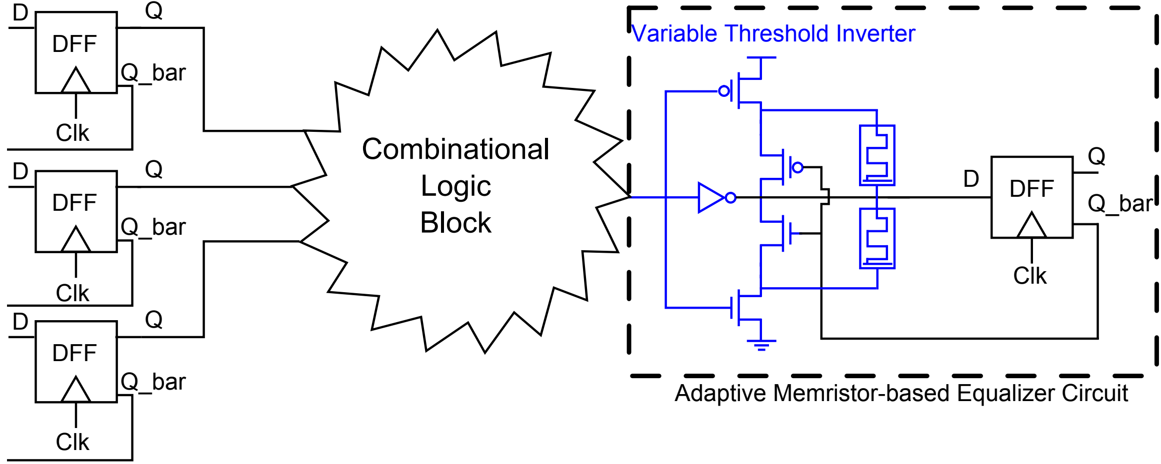

Figure 5·22: Adaptive memristor-based feedback equalizer circuit

ability of the feedback strength is possible by tuning the memristance value of the memristor in the feedback topology. We considered TiO_2 and HfO_x-based memristor technologies in the design of adaptive memristor-based feedback equalizer circuits. The use of TiO_2 or HfO_x-based RRAM devices in the feedback path would decrease the area/leakage overhead of the equalized flip flop and would provide the opportunity to dynamically adjust the feedback strengths and tune the digital logic block with respect to performance, energy and error rate constraints.

The resistance of transistor with $1\mu m$ width in 300 mV supply voltage is roughly $100K\Omega$ for UMC 130 nm technology. Considering the range of $R_{ON} = 100\Omega$ and $R_{OFF} = 16K\Omega$ for the TiO_2 memristors (which is outside the range of transistor resistance in sub-threshold regime), it is impossible to use the TiO_2 memristor technology as the tunable element of the feedback equalizer.

The minimum and maximum resistance value for the HfO_x RRAM technology is $R_{ON} = 3K\Omega$ and $R_{OFF} = 1M\Omega$, respectively. This makes the HfO_x-based memristor technology an appropriate alternative in the design of adaptive feedback equalizer circuits. As discussed in Chapter 3, there needs to be more than 3 V power supply

to program the HfO_x RRAM cell. This supply voltage should be provided off-chip for tuning the HfO_x-based adaptive feedback equalizer circuit.

We developed the required peripheral circuitry for programming the memristor-based feedback equalizer circuit. However, the large leakage energy overhead of the CMOS switches required for tuning the RRAM devices in the feedback path makes the memristor-based feedback equalizer circuit an inappropriate alternative for the design of sub-threshold digital logic systems.

5.5 Summary

In this chapter, we proposed the application of a tunable adaptive feedback equalizer circuit to reduce the normalized variation of total delay along the critical path and the dominant leakage energy of the digital CMOS logic operating in sub-threshold regime. Adjusting the switching thresholds of the gates before the flip flop based on the gate output in the previous cycle, the adaptive feedback equalizer circuit enables a faster switching of the gate outputs and provides the opportunity to reduce the leakage energy of digital logic in weak inversion region. We implemented a non-equalized and an equalized design of a 64-bit adder in UMC 130 nm process using static complementary CMOS logic. Using the equalized design the normalized variation of the total critical path delay can be reduced from 16.1% (non-equalized) to 11.4% (equalized) while reducing the energy-delay product by 25.83% at minimum energy supply voltage. Moreover, we showed that in case of worse than expected process variation, the tuning capability of the equalizer circuit can be used post fabrication to reduce the normalized variation ($3\sigma/\mu$) of the critical path delay with minimal increase in energy. We also presented detailed delay and energy models of the equalized digital logic circuit operating in the sub-threshold regime.

Chapter 6

Conclusion and Future Work

In this dissertation, we have proposed the use of memristor devices in the future non-volatile RRAM array architectures and propose the use of an adaptive feedback equalization technique to improve energy efficiency in ultra low-power subthreshold digital logic circuits. Memristor-based memory architectures have lower write and read energy dissipation compared to other emerging non-volatile memory technologies. The adaptive feedback equalizer circuit helps reduce the dominant process/temperature/voltage variation effects and improve energy efficiency in subthreshold regime.

6.1 Conclusion

In Chapter 2, we reviewed the current approaches on modeling the memristor devices and then explained the functionality of our target TiO_2 and HfO_x-based memristors. We also developed a new state function for HfO_x-based RRAM devices and presented it in this chapter. We then proposed a reliable SPICE netlist for HfO_x memristor technology based on the change in conductive filament diameter.

In Chapter 3, we first presented the detailed functionality of an n-bit 1T1R RRAM cell followed by the design and optimization of an n-bit 1T1R RRAM array architecture using this 1T1R RRAM cell as the building block. We discussed the detailed

implementation of memory cells and arrays using both TiO_2 and HfO_x-based RRAM technologies. We also presented models for performance and energy of read and write operation in n-bit 1T1R RRAM cells designed using TiO_2- and HfO_x-based RRAM devices. We validated our performance and energy models with sub-10% error against HSPICE simulations for both multi-bit TiO_2- and HfO_x-based 1T1R cells. Using energy and performance constraints, we determined the optimum number of bits/cell in the multi-bit RRAM array architecture to be 3. The total write and read energy of the 3 bits/cell TiO_2-based RRAM array was 4.06 pJ/bit and 188 fJ/bit for 100 nsec and 1 nsec write and read access times while the optimized 3 bits/cell HfO_x-based RRAM array consumed 365 fJ/bit and 173 fJ/bit for 1 nsec and 200 nsec write and read access times, respectively. We also presented a detailed analysis of the read destructiveness in multi-level TiO_2- and HfO_x-based RRAM devices. We investigated the trade-off between the robustness against process variations and the read energy consumption of multi-level RRAM array architectures for uniform and non-uniform memristor state assignments. Using the proposed models, we analyzed the destructive effects of process, voltage and temperature (PVT) variations on performance, energy consumption and the reliability of multi-level 1T1R memory cells. Our analysis indicated that multi-bit HfO_x RRAM is more sensitive to LER and is more susceptible to voltage and temperature variations, whereas TiO_2 memristor is more vulnerable to OTF.

In Chapter 4, we analyzed the use of a variable threshold inverter-based feedback equalization technique in the design of energy-efficient sequential logic circuits in sub-threshold regime. We applied the fundamentals of communications-inspired techniques in the design of robust digital logic circuits. Modifying the switching thresholds of the last gate of the critical path (just before the flip flops) based on the prior sampled outputs, the feedback equalization circuit enables a faster switching

of the logic gate outputs and provides the opportunity to reduce the leakage current in weak inversion region. We implemented a non-equalized and an equalized design of an 8-bit carry lookahead adder in UMC 130 nm process using static complementary CMOS logic and decreased the total propagation delay of the critical path of the sub-threshold logic and correspondingly reduced the dominant leakage energy, leading to 35.4% reduction in energy-delay product of the conventional non-equalized design at minimum energy supply voltage. Using the feedback equalizer circuit, we reduced the total energy by 16.72% through voltage scaling while maintaining an operating frequency of 1.28 MHz. Moreover, we demonstrated that the equalized 8-bit carry lookahead adder needs lower upsizing to tolerate the dominant process variation effects leading to 20.72% lower total energy in sub-threshold regime.

In Chapter 5, we proposed using an adaptive feedback equalization technique for tuning the sequential logic circuits designed in sub-threshold regime. This tunable feedback equalizer circuit could reduce energy consumption and improve performance of the sub-threshold digital logic circuits. At the same time, the tunability of this feedback equalizer circuit enabled post-fabrication tuning of the digital logic block to tolerate worse than expected process variations. We implemented a non-equalized and an equalized design of a 64-bit adder in UMC 130 nm process using static complementary CMOS logic. Using the equalized design the normalized variation of the total critical path delay could be further reduced from 16.1% (non-equalized) to 11.4% (equalized) while decreasing the energy-delay product by 25.83% at minimum energy supply voltage. In addition, in case of worse than expected process variation, we showed that the tuning capability of the equalizer circuit could be used post fabrication to reduce the normalized variation ($3\sigma/\mu$) of the critical path delay with negligible energy overhead. Furthermore, we analyzed the application of the proposed feedback equalizer circuit in reducing the temperature/voltage variation effects

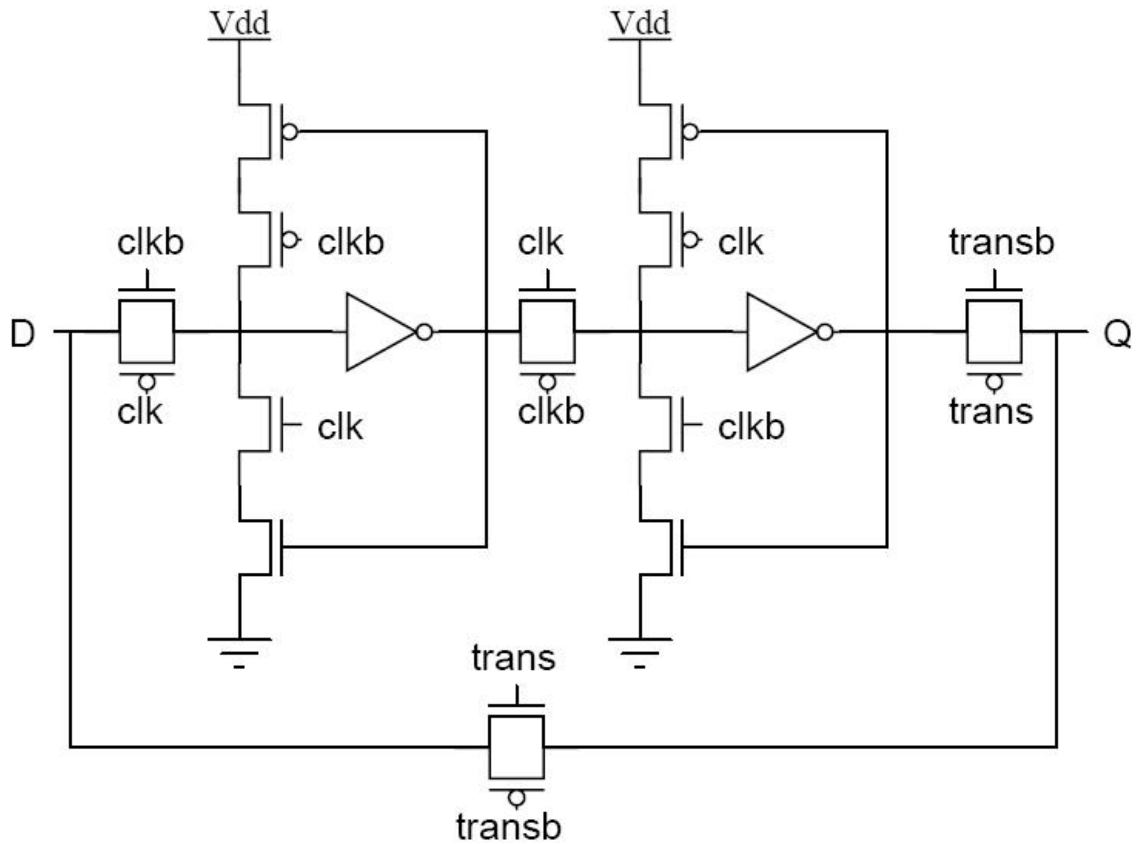

Figure 6·1: Bypass flip flop design

in sub-threshold regime. We also presented detailed delay and energy models of the equalized digital logic circuits operating in the sub-threshold regime. As part of the modeling approach, we developed an accurate analytical methodology to estimate the equivalent resistance of MOSFET devices operating in sub-threshold regime.

6.2 Future Work

6.2.1 Equalized Flip Flop with Bypass

Chapter 4 shows the benefits of feedback equalization technique in reducing the dominant leakage energy component of digital logic circuits designed in sub-threshold

regime. We will explore the design of a bypassed equalized flip flop which is similar to the equalized flip flop except for two transmission gates that are used to implement a bypass path in transparent mode. In a transparent flip flop, changes in D are immediately reflected in the output Q. This in turn effectively reduces the setup time and clk-to-q delay overhead of the classic master-slave positive edge-triggered flip flop. While using this design (shown in Figure 6·1) introduces an area penalty (due to the bypass path), it has less control overhead than the previously presented equalized flip flop, since only one signal (driving the additional transmission gates) needs to be toggled to bring the timing element into or out of transparent mode. Compared to the equalized flip flop discussed in Chapter 4, it compensates for the setup time overhead of the equalized flip flop and more effectively reduces the dominant leakage energy of the sub-threshold sequential logic.

6.2.2 Equalization Techniques for Near-threshold Voltage Computing Applications

Traditionally, the logic circuit community has always targeted the minimum-delay operational point (MDP). Nevertheless, with the emergence of sensor and biomedical applications that require ultra low-energy consumption, the VLSI circuits should be operated near the minimum-energy point (MEP). As the optimum energy-delay curve is quite flat around MEP (Markovic et al., 2010) (see Figure 6·2), a substantial amount of performance can be recovered by just backing off a bit (commonly called the near-threshold region). In Chapter 5, we demonstrated that feedback equalization technique could reduce the energy-delay product of the 64-bit adder by 25.83% at minimum energy supply voltage (MEP). Equalization techniques could be explored to improve the energy efficiency of the sequential logic in near-threshold regime (Kaul et al., 2012). Since a considerable amount of performance can be recovered by op-

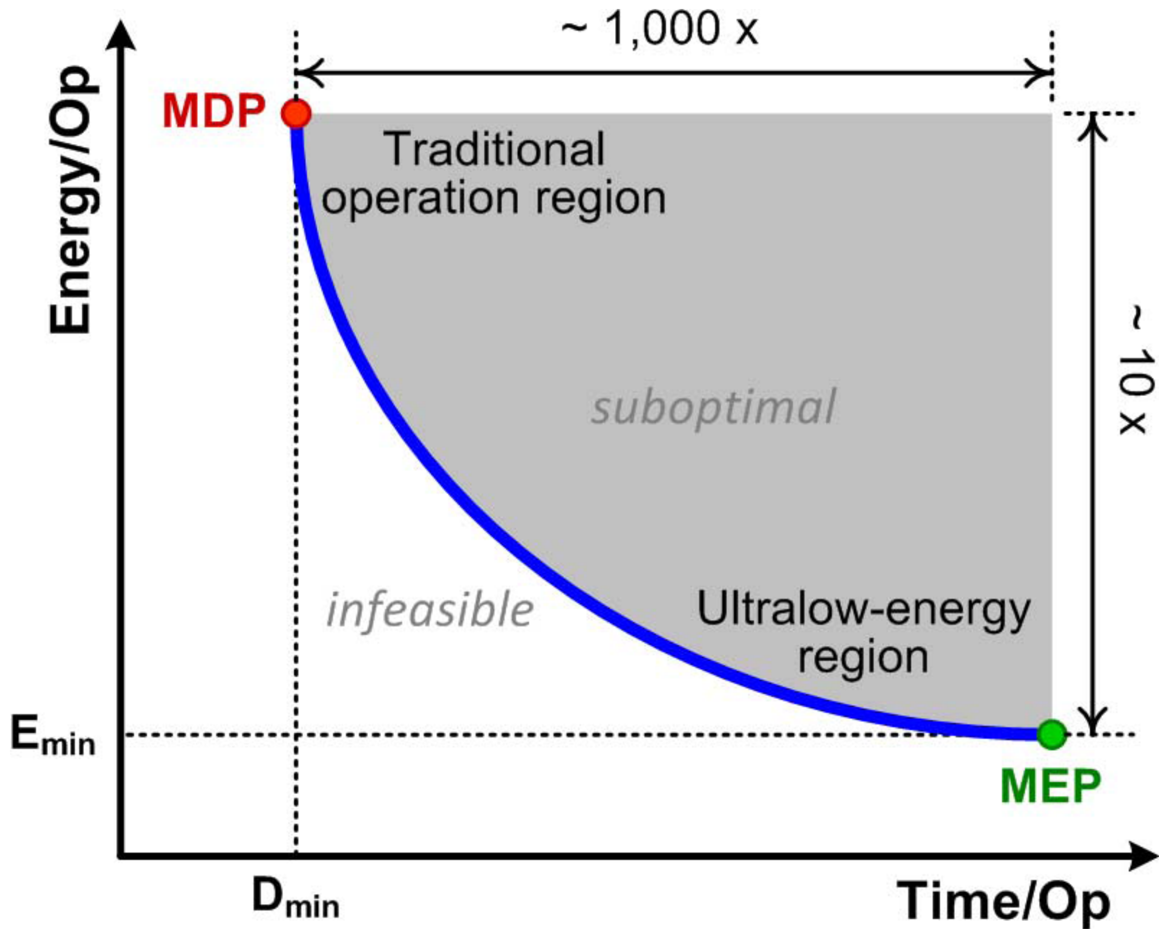

Figure 6·2: Energy-delay trade-off in combinational logic. Traditional operation region is around minimum-delay point (MDP). Ultra low-energy region is around minimum-energy point (MEP) (Markovic et al., 2010).

erating in near-threshold regime (Markovic et al., 2010), the targeted energy-delay product (EDP) values in moderate-inversion regime will be further lower than the values obtained in sub-threshold regime using equalization techniques.

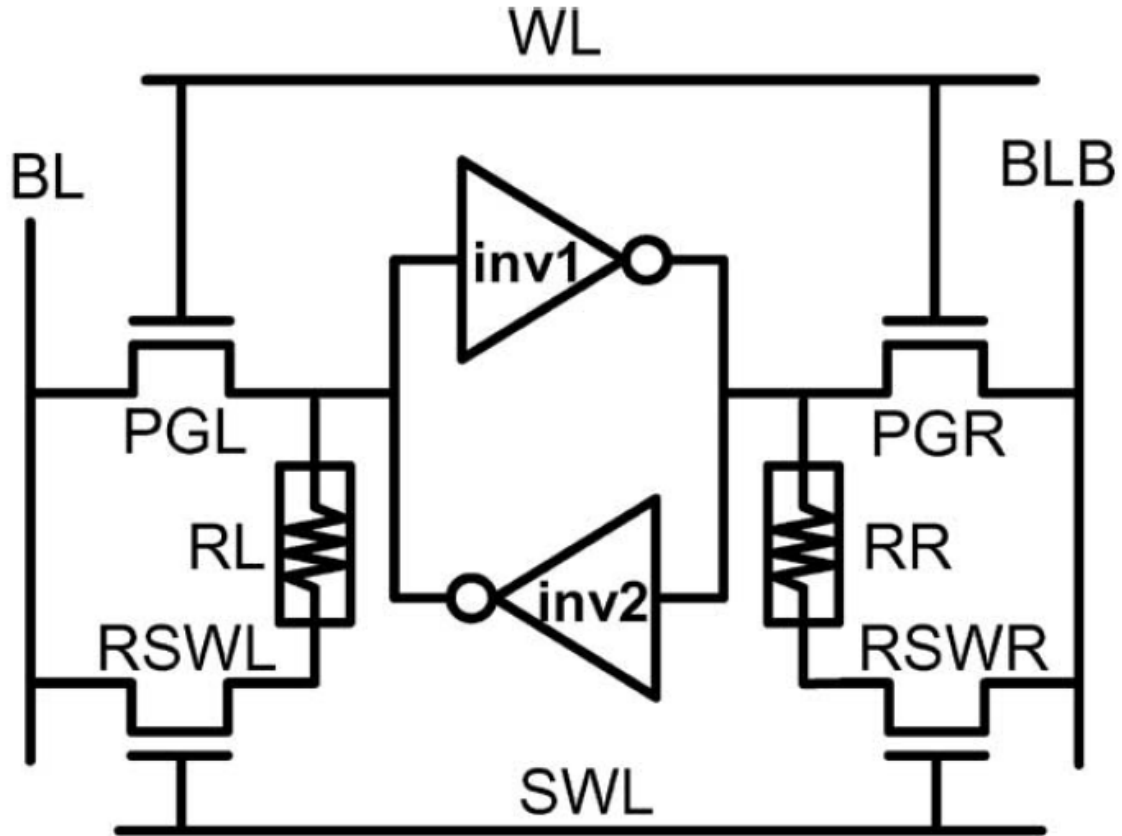

Figure 6·3: Schematic of the 8T2R Memristor-based Nonvolatile (Rnv8T) SRAM cell (Chiu et al., 2012).

6.2.3 Exploring the Architectural Impact of Using Non-volatile Memristor-based On-chip Cache

Historically, SRAMs have been used for on-chip memory as they provide high speed read/write operations. Dynamic voltage scaling (DVS) is a popular approach to suppress active-mode standby power and dynamic power by adjusting the operating voltage in on-chip cache design (Chiu et al., 2012). However, the conventional SRAM design is suffering from the dominant leakage energy component in standby mode. Moreover, it is impossible to preserve data in on-chip SRAM blocks when the power supply is turned off. Non-volatile memory (NVM) provides the opportunity to switch off the power supply to further suppress standby power and extend bat-

tery life of digital chips without a loss of data. The recently-proposed non-volatile SRAM (nvSRAM) (see Figure 6·3) integrates SRAM cells and memristor devices, forming a direct bit-to-bit connection in a 2D or vertical arrangement to achieve fast parallel data transfer and fast power-on/off speed suitable for biomedical or mobile applications (Chiu et al., 2012). This setup enables symmetric read/write access and fast parallel data transfer in systems. More importantly, it is able to operate in normal mode (using SRAM) or standby (using non-volatile RRAM) to achieve better performance and reduce energy consumption. We will explore other potential energy-efficient non-volatile on-chip memory structures and their applications in ultra low-power cache architecture design. In particular, we will explore the application of the nonvolatile 1T1R RRAM array architecture discussed in Chapter 3 in designing energy-efficient on-chip cache architectures. The TiO_2-based RRAM array discussed in Chapter 3 is not a threshold-based technology and is suitable for ultra low-power sub-threshold computing systems where rapid write operation is not required. The multi-level 1T1R TiO_2-based RRAM technology has considerably smaller area and lower read energy compared to the classic SRAM cell making it a promising alternative in the design of highly-dense cache storage architectures. In contrast, a threshold-based RRAM technology (HfO_x) will not be suitable in designing low-voltage cache memory blocks as it requires an out of range supply voltage to perform the read and write operations.

References

Predictive technology model (ptm).

Abdalla, H. and Pickett, M. D. (2011). Spice modeling of memristors. In *IEEE International Symposium on Circuits and Systems proceedings. ISCAS*, pages 1832–1835.

Batas, D. and Fiedler, H. (2011). A memristor spice implementation and a new approach for magnetic flux-controlled memristor modeling. *IEEE Transactions on Nanotechnology*, 10(2):250 –255.

Benderli, S. and Wey, T. A. (2009). On spice macromodelling of tio2 memristors. *Electronics Letters*, 45(7):377–379.

Bersuker, G., Gilmer, D. C., Veksler, D., Kirsch, P., Vandelli, L., and Padovani, A. (2011). Metal oxide resistive memory switching mechanism based on conductive filament properties. *Journal of Applied Physics*, 110(12).

Biolek, Z., Biolek, D., and Biolkova, V. (2009). Spice model of memristor with nonlinear dopant drift. *Radioengineering*, 18(2).

Borkar, S., Karnik, T., and De, V. (2004). Design and reliability challenges in nanometer technologies. In *Proceedings: ACM/IEEE Design Automation Conference*, page 75.

Bowman, K. A. and Tschanz, J. W. (2010). Resilient microprocessor design for improving performance and energy efficiency. In *Proceedings of the International Conference on Computer-Aided Design*, ICCAD '10, pages 85–88. IEEE Press.

Bull, D., Das, S., Shivashankar, K., Dasika, G., Flautner, K., and Blaauw, D. (2011). A power-efficient 32 bit arm processor using timing-error detection and correction for transient-error tolerance and adaptation to pvt variation. *IEEE Journal of Solid-State Circuits*, 46(1):18–31.

Chae, K. and Mukhopadhyay, S. (2014). A dynamic timing error prevention technique in pipelines with time borrowing and clock stretching. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 61(1):74–83.

Chen, Y.-C., Zhang, W., and Li, H. (2012). A look up table design with 3d bipolar rrams. In *Proceedings of the Asia and South Pacific Design Automation Conference*, pages 73 –78.

Chen, Y. S., Lee, H., Chen, P., Gu, P., Chen, C., Lin, W., Liu, W. H., Hsu, Y. Y., Sheu, S. S., Chiang, P.-C., Chen, W.-S., Chen, F., Lien, C., and Tsai, M.-J. (2009a). Highly scalable hafnium oxide memory with improvements of resistive distribution and read disturb immunity. In *2009 IEEE International Electron Devices Meeting (IEDM)*, pages 1 –4.

Chen, Y.-S., Wu, T.-Y., Tzeng, P.-J., Chen, P.-S., Lee, H.-Y., Lin, C.-H., Chen, P.-S., and Tsai, M.-J. (2009b). Forming-free hfo2 bipolar rram device with improved endurance and high speed operation. In *Proceedings of technical papers (International Symposium on VLSI Technology) VLSI-TSA*, pages 37 –38.

Chiu, P.-F., Chang, M.-F., Wu, C.-W., Chuang, C.-H., Sheu, S.-S., Chen, Y.-S., and Tsai, M.-J. (2012). Low store energy, low vddmin, 8t2r nonvolatile latch and sram with vertical-stacked resistive memory (memristor) devices for low power mobile applications. *IEEE Journal of Solid-State Circuits*, 47(6):1483 –1496.

Choi, S.-H., Paul, B., and Roy, K. (2004). Novel sizing algorithm for yield improvement under process variation in nanometer technology. In *Design Automation Conference, 2004. Proceedings. 41st*, pages 454–459.

Chua, L. (1971). Memristor-the missing circuit element. *IEEE Transactions on Circuit Theory*, 18(5):507 – 519.

Chuang, C.-T., Mukhopadhyay, S., Kim, J.-J., Kim, K., and Rao, R. (2007). High-performance sram in nanoscale cmos: Design challenges and techniques. In *Proceedings: International Workshop on Memory Technology, Design, and Testing (MTDT)*, pages 4 –12.

Chung, H., Jeong, B.-H., Min, B., Choi, Y., Cho, B.-H., Shin, J., Kim, J., Sunwoo, J., min Park, J., Wang, Q., Lee, Y.-J., Cha, S., Kwon, D., Kim, S., Kim, S., Rho, Y., Park, M.-H., Kim, J., Song, I., Jun, S., Lee, J., Kim, K., won Lim, K., ryul Chung, W., Choi, C., Cho, H., Shin, I., Jun, W., Hwang, S., Song, K.-W., Lee, K., whan Chang, S., Cho, W.-Y., Yoo, J.-H., and Jun, Y.-H. (2011). A 58nm 1.8v 1gb pram with 6.4mb/s program bw. In *2011 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 500 –502.

De Vita, G. and Iannaccone, G. (2007). A voltage regulator for subthreshold logic with low sensitivity to temperature and process variations. In *Proceedings IEEE International Solid-State Circuits Conference*, pages 530–620.

Eshraghian, K., Cho, K.-R., Kavehei, O., Kang, S.-K., Abbott, D., and Kang, S.-M. S. (2011). Memristor mos content addressable memory (mcam): Hybrid architecture for future high performance search engines. *IEEE Transactions on VLSI Systems*, 19(8):1407 –1417.

Fei, W., Yu, H., Zhang, W., and Yeo, K. S. (2012). Design exploration of hybrid cmos and memristor circuit by new modified nodal analysis. *IEEE Transactions*

on VLSI Systems, (6):1012 –1025.

Guan, X., Yu, S., and Wong, H.-S. (2012a). On the switching parameter variation of metal-oxide rram-2014;part i: Physical modeling and simulation methodology. *IEEE Transactions on Electron Devices*, 59(4):1172 –1182.

Guan, X., Yu, S., and Wong, H.-S. (2012b). On the variability of hfox rram: from numerical simulation to compact modeling. In *Proceedings Workshop on Compact Modeling*, pages 815–820.

Ho, Y., Huang, G., and Li, P. (2011). Dynamical properties and design analysis for nonvolatile memristor memories. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 58(4):724 –736.

Hu, M., Li, H., Chen, Y., Wang, X., and Pino, R. (2011a). Geometry variations analysis of tio2 thin-film and spintronic memristors. In *Proceedings of the Asia and South Pacific Design Automation Conference*, pages 25–30.

Hu, M., Li, H., and Pino, R. (2011b). Fast statistical model of tio2 thin-film memristor and design implication. In *Proceeding ICCAD – IEEE/ACM International Conference on Computer-Aided Design*, pages 345 –352.

Ielmini, D. (2011). Modeling the universal set/reset characteristics of bipolar rram by field- and temperature-driven filament growth. *IEEE Transactions on Electron Devices*, 58(12):4309 –4317.

Jayakumar, N. and Khatri, S. (2005). A variation-tolerant sub-threshold design approach. In *Design Automation Conference, 2005. Proceedings. 42nd*, pages 716–719.

Jiang, Z., Zhao, F., Jing, W., Prewett, P., and Jiang, K. (2009). Characterization of line edge roughness and line width roughness of nano-scale typical structures. In *Proceedings 4th Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS)*, pages 299 –303.

Jo, S. H., Kim, K., and Lu, W. (2009). High-density crossbar arrays based on a si memristive system. *Nano Letters*, 9(2):870–874.

Joglekar, Y. and Wolf, S. (2009). The elusive memristor: properties ofbasic electrical circuits. *European Journal of Physics*, 30(4):661–675.

Kaul, H., Anders, M., Hsu, S., Agarwal, A., Krishnamurthy, R., and Borkar, S. (2012). Near-threshold voltage (ntv) design-opportunities and challenges. In *Proceedings Design Automation Conference*, pages 1149–1154.

Kim, H., Sah, M., Yang, C., and Chua, L. (2012a). Memristor emulator for memristor circuit applications. *IEEE Transactions on Circuits and Systems I: Regular Papers*, page 1.

Kim, H., Sah, M., Yang, C., Roska, T., and Chua, L. (2012b). Neural synaptic weighting with a pulse-based memristor circuit. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 59(1):148 –158.

Kim, S. and Seok, M. (2014). Reconfigurable regenerator-based interconnect design for ultra-dynamic-voltage-scaling systems. In *Proceedings of the 2014 International Symposium on Low Power Electronics and Design*, ISLPED '14, pages 99–104, New York, NY, USA. ACM.

Kuhn, K. (2012). Considerations for ultimate cmos scaling. *IEEE Transactions on Electron Devices*, 59(7):1813 –1828.

Kvatinsky, S., Friedman, E., Kolodny, A., and Weiser, U. (2013). Team: Threshold adaptive memristor model. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(1):211 –221.

Kwong, J., Ramadass, Y., Verma, N., and Chandrakasan, A. (2009). A 65 nm sub-vt microcontroller with integrated sram and switched capacitor dc-dc converter. *IEEE Journal of Solid-State Circuits*, 44(1):115–126.

Lee, M., Lee, C. B., D., L., and R., L. S. (2011). A fast, high-endurance and scalable non-volatile memory device made from asymmetric ta2o5-x/tao2-x bilayer structures. *Nature Materials*, 10:625–630.

Li, D., Chuang, P.-J., Nairn, D., and Sachdev, M. (2011). Design and analysis of metastable-hardened flip-flops in sub-threshold region. In *2011 International Symposium on Low Power Electronics and Design (ISLPED)*, pages 157–162.

Liauw, Y. Y., Zhang, Z., Kim, W., Gamal, A., and Wong, S. (2012). Nonvolatile 3d-fpga with monolithically stacked rram-based configuration memory. In *Proceedings IEEE International Solid-State Circuits Conference (ISSCC)*, pages 406 –408.

Liu, B., Ashouei, M., Huisken, J., and De Gyvez, J. P. (2012). Standard cell sizing for subthreshold operation. In *Proceedings Design Automation Conference*, pages 962–967.

Liu, B., Pourshaghaghi, H., Londono, S., and de Gyvez, J. (2011). Process variation reduction for cmos logic operating at sub-threshold supply voltage. In *2011 14th Euromicro Conference on Digital System Design (DSD)*, pages 135–139.

Liu, T.-T. and Rabaey, J. (2012). A 0.25v 460nw asynchronous neural signal processor with inherent leakage suppression. In *2012 Symposium on VLSI Circuits (VLSIC)*, pages 158–159.

Lotze, N. and Manoli, Y. (2012). A 62 mv 130 nm cmos standard-cell-based design technique using schmitt-trigger logic. *IEEE Journal of Solid-State Circuits*, 47(1):47–60.

Lotze, N., Ortmanns, M., and Manoli, Y. (2008). Variability of flip-flop timing at sub-threshold voltages. In *Proceedings of International Symposium on Low Power Electronics and Design (ISLPED)*, pages 221–224.

Lu, W., Kim, K.-H., Chang, T., and Gaba, S. (2011). Two-terminal resistive switches (memristors) for memory and logic applications. In *Proceedings of the Asia and South Pacific Design Automation Conference*, pages 217 –223.

Manem, H. and Rose, G. S. (2011). A read-monitored write circuit for 1t1m multi-level memristor memories. In *Proceedings IEEE International Symposium on Circuits and Systems*, pages 2938 –2941.

Markovic, D. et al. (2010). Ultralow-power design in near-threshold region. *Proceedings of the IEEE*, 98(2):237–252.

Mishra, B., Al-Hashimi, B., and Zwolinski, M. (2009). Variation resilient adaptive controller for subthreshold circuits. In *Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09.*, pages 142–147.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, 38(8):114–117.

Nebashi, R., Sakimura, N., Honjo, H., Saito, S., Ito, Y., Miura, S., Kato, Y., Mori, K., Ozaki, Y., Kobayashi, Y., Ohshima, N., Kinoshita, K., Suzuki, T., Nagahara, K., Ishiwata, N., Suemitsu, K., Fukami, S., Hada, H., Sugibayashi, T., and Kasai, N. (2009). A 90nm 12ns 32mb 2t1mtj mram. In *Proceedings IEEE International Solid-State Circuits Conference (ISSCC)*, pages 462 –463,463a.

Niu, D., Chen, Y., and Xie, Y. (2010a). Low-power dual-element memristor based memory design. In *Proceedings International Symposium on Low Power Electronics and Design (ISLPED)*, pages 25–30.

Niu, D., Chen, Y., Xu, C., and Xie, Y. (2010b). Impact of process variations on emerging memristor. In *Proceedings Design Automation Conference*, pages 877–882.

Niu, D., Xiao, Y., and Xie, Y. (2012). Low power memristor-based reram design with error correcting code. In *Proceedings of the Asia and South Pacific Design Automation Conference*, pages 79 –84.

Pelgrom, M., Duinmaijer, A. C. J., and Welbers, A. (1989). Matching properties of mos transistors. *IEEE Journal of Solid-State Circuits*, 24(5):1433–1439.

Pickett, M. D., Strukov, D., Borghetti, J., Yang, J., Snider, G., Stewart, D., and Williams, S. (2009). Switching dynamics in titanium dioxide memristive devices. *Applied Physics*, 106:440–446.

Pu, Y., Pineda de Gyvez, J., Corporaal, H., and Ha, Y. (2010). An ultra-low-energy

multi-standard jpeg co-processor in 65 nm cmos with sub/near threshold supply voltage. *IEEE Journal of Solid-State Circuits*, 45(3):668–680.

Qazi, M., Clinton, M., Bartling, S., and Chandrakasan, A. (2011). A low-voltage 1mb feram in 0.13 um cmos featuring time-to-digital sensing for expanded operating margin in scaled cmos. In *Proceedings IEEE International Solid-State Circuits Conference*, pages 208 –210.

Rabaey, J., Chandrakasan, A., and Nikolić, B. (2003). *Digital Integrated Circuits*. Pearson Education.

Rak, A. and Cserey, G. (2010). Macromodeling of the memristor in spice. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pages 632–636.

Russo, U., Ielmini, D., Cagli, C., and Lacaita, A. (2009). Filament conduction and reset mechanism in nio-based resistive-switching memory (rram) devices. *IEEE Transactions on Electron Devices*, (2):186 –192.

Sail, E. and Vesterbacka, M. (2004). A multiplexer based decoder for flash analog-to-digital converters. In *Conference Proceedings. IEEE Region 10.; Institute of Electrical and Electronics Engineers. Hong Kong Section (TENCON)*, pages 250 – 253 Vol. 4.

Sarpeshkar, R. (2010). *Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems*. Cambridge books online. Cambridge University Press.

Schinkel, D., Mensink, E., Klumperink, E., van Tuijl, E., and Nauta, B. (2006). A 3-gb/s/ch transceiver for 10-mm uninterrupted rc-limited global on-chip interconnects. *IEEE Journal of Solid-State Circuits*, 41(1):297–306.

Schinkel, D., Mensink, E., Klumperink, E., Van Tuijl, E., and Nauta, B. (2007). A double-tail latch-type voltage sense amplifier with 18ps setup+hold time. In *Proceedings IEEE International Solid-State Circuits Conference*, pages 314 –605.

Seo, J., Singh, P., Sylvester, D., and Blaauw, D. (2007). Self-timed regenerators for high-speed and low-power interconnect. In *8th International Symposium on Quality Electronic Design, 2007. ISQED '07*, pages 621–626.

Sheu, S.-S., Chang, M.-F., Lin, K.-F., Wu, C.-W., Chen, Y.-S., Chiu, P.-F., Kuo, C.-C., Yang, Y.-S., Chiang, P.-C., Lin, W.-P., Lin, C.-H., Lee, H.-Y., Gu, P.-Y., Wang, S.-M., Chen, F., Su, K.-L., Lien, C.-H., Cheng, K.-H., Wu, H.-T., Ku, T.-K., Kao, M.-J., and Tsai, M.-J. (2011). A 4mb embedded slc resistive-ram macro with 7.2ns read-write random-access time and 160ns mlc-access capability. In *Proceedings IEEE International Solid-State Circuits Conference*, pages 200 –202.

Sheu, S.-S., Chiang, P.-C., Lin, W.-P., Lee, H.-Y., Chen, P.-S., Chen, Y.-S., Wu, T.-Y., Chen, F., Su, K.-L., Kao, M.-J., Cheng, K.-H., and Tsai, M.-J. (2009). A 5ns fast write multi-level non-volatile 1 k bits rram memory with advance write scheme. In *2009 Symposium on VLSI Circuits*, pages 82 –83.

Singh, P., Seo, J.-s., Blaauw, D., and Sylvester, D. (2008). Self-timed regenerators for high-speed and low-power on-chip global interconnect. *IEEE Transactions on Very Large Scale Integration Systems*, 16(6):673–677.

Sridhara, S., Balamurugan, G., and Shanbhag, N. (2008). Joint equalization and coding for on-chip bus communication. *IEEE Transactions on Very Large Scale Integration Systems*, 16(3):314–318.

Strukov, D. and Williams, R. (2011). Intrinsic constrains on thermally-assisted memristive switching. *Applied Physics A: Materials Science & Processing*, 102:851–855.

Strukov, D. B., Snider, G. S., Stewart, D. R., and Williams, R. S. (2008). The missing memristor found. *Nature*, 453(7191):80–83.

Takhirov, Z., Nazer, B., and Joshi, A. (2012). Error mitigation in digital logic using a feedback equalization with schmitt trigger (fest) circuit. In *2012 13th International Symposium on Quality Electronic Design (ISQED)*, pages 312–319.

Tschanz, J., Bowman, K., Lu, S.-L., Aseron, P., Khellah, M., Raychowdhury, A., Geuskens, B., Tokunaga, C., Wilkerson, C., Karnik, T., and De, V. (2010). A 45nm resilient and adaptive microprocessor core for dynamic variation tolerance. In *2010 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 282–283.

Tschanz, J., Kao, J., Narendra, S., Nair, R., Antoniadis, D., Chandrakasan, A., and De, V. (2002). Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage. In *Proceedings IEEE International Solid-State Circuits Conference*, volume 1, pages 422–478 vol.1.

Verma, N., Kwong, J., and Chandrakasan, A. (2008). Nanometer mosfet variation in minimum energy subthreshold circuits. *IEEE Transactions on Electron Devices*, 55(1):163–174.

Wang, A. and Chandrakasan, A. (2005). A 180-mv subthreshold fft processor using a minimum energy design methodology. *IEEE Journal of Solid-State Circuits*, 40(1):310–319.

Wang, X., Chen, Y., Xi, H., Li, H., and Dimitrov, D. (2009). Spintronic memristor through spin-torque-induced magnetization motion. *IEEE Electron Device Letters*, 30(3):294 –297.

Whatmough, P., Das, S., and Bull, D. (2013). A low-power 1ghz razor fir accelerator with time-borrow tracking pipeline and approximate error correction in 65nm cmos.

In *IEEE International Solid-State Circuits Conference*, pages 428–429.

Witrisal, K. (2009). Memristor-based stored-reference receiver - the uwb solution? *Electronics Letters*, 45(14):713 –714.

Xu, C., Dong, X., Jouppi, N., and Xie, Y. (2011). Design implications of memristor-based rram cross-point structures. In *Proceedings Design, Automation and Test in Europe*, pages 1 –6.

Xue, X. Y., Jian, W., Yang, J. G., Xiao, F. J., Chen, G., Xu, X. L., Xie, Y. F., Lin, Y., Huang, R., Zhou, Q. T., and Wu, J. G. (2012). A 0.13 um 8mb logic based cuxsiyo resistive memory with self-adaptive yield enhancement and operation power reduction. In *2012 Symposium on VLSI Circuits*, pages 42 –43.

Yu, S., Guan, X., and Wong, H.-S. (2012). On the switching parameter variation of metal oxide rram-;part ii: Model corroboration and device design strategy. *IEEE Transactions on Electron Devices*, 59(4):1183 –1188.

Zangeneh, M. and Joshi, A. (2012). Performance and energy models for memristor-based 1t1r rram cell. In *Proceedings of the Great Lakes Symposium on VLSI 2012*, pages 9–14.

Zangeneh, M. and Joshi, A. (2014a). Design and optimization of nonvolatile multibit 1t1r resistive ram. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 22(8):1815–1828.

Zangeneh, M. and Joshi, A. (2014b). Sub-threshold logic circuit design using feedback equalization. In *Proceedings Design, Automation and Test in Europe, 2014*, pages 1–6.

Zangeneh, M. and Joshi, A. (2015). Designing tunable sub-threshold logic circuits using adaptive feedback equalization. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*.

Zhai, B., Hanson, S., Blaauw, D., and Sylvester, D. (2005). Analysis and mitigation of variability in subthreshold design. In *Proceedings of the 2005 International Symposium on Low Power Electronics and Design, 2005.*, pages 20–25.

Zhou, J., Jayapal, S., Busze, B., Huang, L., and Stuyt, J. (2011). A 40 nm inverse-narrow-width-effect-aware sub-threshold standard cell library. In *2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC)*, pages 441–446.

CURRICULUM VITAE

Mahmoud Zangeneh
Email: zangeneh@bu.edu
Phone: +1(857)540-2061

EDUCATION

- **Doctor of Philosophy**, Electrical Engineering
  Boston University, Boston, MA, USA
  Thesis Topic: Designing Memory and Logic Circuits using Hybrid CMOS/Memristor Technology
  Advisor: Prof. Ajay Joshi
  September 2010 – May 2015

- **Master of Science**, Electrical Engineering
  University of Tehran, Tehran, Iran
  Thesis Topic: Analysis of Delay and Crosstalk Reduction Techniques in Global Interconnections in Modern VLSI Systems
  Advisor: Prof. Nasser Masoumi
  September 2007 – February 2009

- **Bachelor of Science**, Electrical Engineering
  Tehran Polytechnic, Tehran, Iran
  September 2002 – May 2007

TECHNICAL SKILLS

- Circuit Design Software Tools: Cadence SpectreRF, Hspice, Encounter RTL Compiler, Virtuoso Layout Editor, SimVision, LabView, Advanced Design System (ADS), Modelsim, Altium Designer, Synplify, Leonardo, L-Edit, S-Edit, Silvaco, ANSYS HFSS, Lumerical Device and FDTD Solutions

- Programming Languages: Verilog, VHDL, Visual C++, Visual Basic, Java, Python and Perl

- Mathematical Software and Languages: Matlab, Simulink, Maple

WORK EXPERIENCE

- **Research Assistant**, Boston University, Boston, MA, USA
  - Developed a novel feedback equalization circuit to enhance the operating frequency of standard digital CMOS logic while burning lower leakage energy and mitigating timing errors. The proposed feedback equalizer reduces the propagation delay time of the critical path based on the sampled data in the previous clock cycle. The functionality of the proposed feedback circuit has been verified using Cadence Encounter in an ASIC physical design flow tape-out (Verilog code, RTL design, Synthesis and Timing analysis, DRC/LVS-clean Layout design, Floorplanning, Placement & Routing to final GDS) using 0.13 μm UMC Design Kit Technology
  - Developed a CMOS-based adaptive equalizer circuit topology having various feedback strengths to dynamically tune the designed digital logic and meet the required performance constraints for an energy and error rate budget
  - Developed an optimization framework of multi-bit 1T1R RRAM arrays considering performance, energy, reliability and process variation constraints. The proposed technique provides the optimum number of bits/cell for nonvolatile RRAM arrays consisting of TiO_2- and HfO_x-based memristors
  - Proposed three types of Hardware Trojans based on the switching power, leakage power and critical path delay measurements. A Negative Bias Temperature Instability (NBTI) aging approach is used to create ultra low-leakage Hardware Trojans in the critical path of the AES block
  - Proposed the application of embedded nano-antennas in detecting the Hardware Trojans by comparing the optical pattern of the array of nano-antennas embedded in filler cells. The proposed technique provides a more effective method than the conventional electrical testing methodologies to identify the changes in the surrounding circuitry

  May 2011 – May 2015

- **Graduate Teaching Fellow**, Boston University, Boston, MA, USA
  - EC410 Introduction to Analog Electronics, Fall 2010, supervisor: Prof. Ronald Knepper
  - EC410 Introduction to Analog Electronics, Spring 2011, supervisor: Dr. David Freedman
  - EK307 Analysis and Design of Linear Circuits, Spring 2014, supervisor: Prof. Michelle Sander

– EK307 Analysis and Design of Linear Circuits, Spring 2015, supervisor: Prof. Min-Chang Lee

January 2015 – May 2015

PUBLICATION

1. B. Zhou, R. Adato, M. Zangeneh, J. Yang, A. Uyar, M. Selim Unlu, B. Goldberg and A. Joshi, "Detecting hardware trojans using backside optical imaging of embedded watermarks," *Proc. Design Automation Conference (DAC), 2015.*
2. M. Zangeneh, and A. Joshi, "Designing tunable sub-threshold logic circuits using adaptive feedback equalization," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, 2015.*
3. T. Cilingiroglu, M. Zangeneh, A. Uyar, W. Clem Karl, J. Konrad, A. Joshi, B. Goldberg and M. Selim Unlu, "Dictionary-based sparse representation for resolution improvement in Laser Voltage Imaging of CMOS Integrated Circuits," *Proc. Design, Automation and Test in Europe (DATE), 2015.*
4. M. Zangeneh, and A. Joshi, "Sub-threshold logic circuit design using feedback equalization," *Proc. Design, Automation and Test in Europe (DATE), 2014.*
5. M. Zangeneh, and A. Joshi, "Design and optimization of nonvolatile multibit 1T1R resistive RAM," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, vol. 22, no. 8, pp. 1815-1828, August 2014.*
6. M. Zangeneh, and A. Joshi, "Performance and energy models for memristor-based 1T1R RRAM cell," *Proc. ACM Great Lakes Symposium on VLSI, May 2012.*
7. M. Zangeneh and N. Masoumi, "MIRIM: Modified interleaved repeater insertion methodology to reduce delay uncertainty in global interconnections," *Journal of Electrical Systems and Signals, 2014.*
8. M. Zangeneh, H. Aghababa, and B. Forouzandeh, "Effects of device and peripheral parameters on transconductance of silicon nanowire transistors," *Journal of Nanoscience and Nanotechnology (JNN), Dec 2011.*
9. H. Aghababa, M. Zangeneh, A. Afzali-Kusha, and B. Forouzandeh, "Statistical delay modeling of read operation of SRAMs due to channel length variation," *Proc. International Symposium on Circuits and Systems (ISCAS), May 2010.*
10. M. Zangeneh, H. Aghababa, and B. Forouzandeh, "Design of a two-capacitor sample and hold circuit using a two-stage OTA with hybrid cascode compen-

sation,"*Proc. European Conference on Circuit Theory and Design (ECCTD), 2009.*

11. M. Zangeneh and N. Masoumi, "Throughput optimization for interleaved repeater-inserted interconnects in VLSI design,"*Proc. International NanoElectronics Conference, Jan 2010.*
12. M. Zangeneh and N. Masoumi, "An analytical delay reduction strategy for buffer-inserted global interconnects in VDSM technologies,"*Proc. European Conference on Circuit Theory and Design (ECCTD), Aug 2009.*
13. M. Zangeneh, H. Aghababa, and B. Forouzandeh, "Analysis of potential function in cylindrical nanowires,"*Proc. 18th (IEEE) International Conference of Semiconductors (Circuits and Systems), Oct 2008 .*
14. H. Hosseinzadegan, H. Aghababa, M. Zangeneh, A. Afzali-kusha, and B. Forouzandeh, "A compact current- voltage model for carbon nanotube field effect transistors,"*Proc. 18th (IEEE) International Conference of Semiconductors (Circuits and Systems), Oct 2008 .*
15. M. Zangeneh, and N. Masoumi, "Statistical delay metrics for binary RC Tree Interconnects in VDSM technology,"*Proc. 17th Iranian Conference on Electrical Engineering, May 2009 .*

AWARDS

- Top 1% among 20000 attendees in the Iranian National Entrance Exam for Graduate Studies in Electrical Engineering, May 2007.
- Top 0.1% among 500000 attendees in the Iranian National Entrance Exam for Undergraduate Studies, Aug 2002.
- Ranked 5th out of 100000 attendees to the Azad University Entrance Exam for Undergraduate Studies, Tehran, Iran, Sep 2002.

SCHOLARSHIPS

- Research Assistant, Boston University, 2011 – 2014
- Graduate Teaching Fellowship, Boston University, 2010 – 2011

LANGUAGES

- English
- Farsi
- French