Compact Modeling and Simulation for

Digital Circuit Aging

by

Jyothi Bhaskarr Velamala

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2012 by the
Graduate Supervisory Committee:

Yu Cao, Chair
Lawrence Clark
Chaitali Chakrabarti
Hongbin Yu

ARIZONA STATE UNIVERSITY

December 2012

ABSTRACT

Negative bias temperature instability (NBTI) is a leading aging mechanism in modern digital and analog circuits. Recent NBTI data exhibits an excessive amount of randomness and fast recovery, which are difficult to be handled by conventional power-law model (t^n). Such discrepancies further pose the challenge on long-term reliability prediction under statistical variations and Dynamic Voltage Scaling (DVS) in real circuit operation. To overcome these barriers, the modeling effort in this work (1) practically explains the aging statistics due to randomness in number of traps with log(t) model, accurately predicting the mean and variance shift; (2) proposes cycle-to-cycle model (from the first-principle of trapping) to handle aging under multiple supply voltages, predicting the non-monotonic behavior under DVS (3) presents a long-term model to estimate a tight upper bound of dynamic aging over multiple cycles, and (4) comprehensively validates the new set of aging models with 65nm statistical silicon data. Compared to previous models, the new set of aging models capture the aging variability and the essential role of the recovery phase under DVS, reducing unnecessary guard-banding during the design stage.

With CMOS technology scaling, design for reliability has become an important step in the design cycle, and increased the need for efficient and accurate aging simulation methods during the design stage. NBTI induced delay shifts in logic paths are asymmetric in nature, as opposed to averaging effect due to recovery assumed in traditional aging analysis. Timing violations due to aging, in particular, are very sensitive to the standby operation regime of a digital circuit.

In this report, by identifying the critical moments in circuit operation and considering the asymmetric aging effects, timing violations under NBTI effect are correctly predicted. The unique contributions of the simulation flow include: (1) accurate modeling of aging induced delay shift due to threshold voltage (V_{th}) shift using only the delay dependence on supply voltage from cell library; (2) simulation flow for asymmetric aging analysis is proposed and conducted at critical points in circuit operation; (3) setup and hold timing violations due to NBTI aging in logic and clock buffer are investigated in sequential circuits and (4) proposed framework is tested in VLSI applications such DDR memory circuits. This methodology is comprehensively demonstrated with ISCAS89 benchmark circuits using a 45nm Nangate standard cell library characterized using predictive technology models. Our proposed design margin assessment provides design insights and enables resilient techniques for mitigating digital circuit aging.

DEDICATION

To my parents.

ACKNOWLEDGMENTS

This thesis arose in part out of years of research that has been done since I came to NIMO group. By that time, I have worked with a great number of people whose contributions in assorted ways to the research and the making of the thesis deserved special mention. It is a pleasure to convey my gratitude to them all in my humble acknowledgment.

In the first place I would like to record my gratitude to my advisor and mentor Dr. Yu Cao for his supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences throughout the work. Above all and the most needed, he provided me unflinching encouragement and support in various ways. I am grateful to my committee members Dr. Lawrence Clark, Dr. Chaitali Chakrabarti, and Dr. Hongbin Yu for their time and efforts in reviewing this work.

I would also like to convey thanks to Dr. Takashi Sato from Kyoto University, and Dr. Vijay Reddy from Texas Instruments for providing us the silicon data, and Dr. Gilson Wirth from UFRGS for his valuable insights in this work. I thank them for their constructive discussions and suggestions on this research work.

My thanks to Varsha Balakrishnan, Ketul Sutaria, Rui Zheng and Venkatesa Sarma, for their valuable contributions to this research project. I would also like to thank all the other NIMOs: Chi-Chao Wang, Saurabh Sinha, Yun Ye, Min Chen, Wei Zhao, Naveen Suda, Cheng Xu and Zihan Xu. You were a pleasure to work with.

I am indebted to my family for their unconditional love and support. Finally, I would like to thank all my friends who were also important in the successful realization of this thesis.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

CHAPTER

1 INTRODUCTION 1

1.1 Introduction 1

1.2 Limiting Factors of Circuit Lifetime 3

1.3 BTI and its Variability with Technology Scaling 6

1.4 Impact of NBTI at Device and Circuit Level 10

1.5 Current Research Efforts on NBTI 13

1.6 Thesis Contributions 17

1.7 Thesis Outline and Organization 20

2 RELIABILITY PHYSICS 21

2.1 Introduction 21

2.2 Reaction-Diffusion Mechanism 22

2.3 NBTI Degradation Model 24

2.4 RD Theory: Limitations 27

2.5 Trapping/Detrapping Theory 28

3 COMPACT AGING MODELS 30

3.1 Motivation 30

3.2 Model for Static Stress 33

3.3 Aging Prediction in DVS Operation 37

CHAPTER Page

3.4 Longterm Prediction 41

3.5 Silicon Validation 43

3.5.1 Test Chip and Measurement Plan 43

3.5.2 Aging Statistics 45

3.5.3 Model Validation 46

3.6 Circuit Aging Prediction under Multiple V_{dd} 54

3.6.1 Aging Prediction under Variability 54

3.6.2 Multiple V_{dd} in DVS 56

4 FAILURE DIAGNOSIS OF ASYMMETRIC AGING 58

4.1 Circuit Aging 58

4.2 Previous Works 59

4.3 Asymmetric Aging 61

4.4 Gate Level Aging Models 64

4.5 Asymmetric Aging Analysis 69

4.5.1 Simulation Method 69

4.5.2 Timing Violations in Sequential Circuits 71

4.6 Impact on VLSI Design 77

5 STATISTICAL AGING PREDICTION 79

5.1 Statistical Circuit Aging 79

5.2 Timing Violations under NBTI 81

5.3 Supple Voltage Variations 85

CHAPTER Page

6 SUMMARY AND FUTURE WORK 87

6.1 Thesis Conclusions 87

6.2 Future Work 89

References 90

APPENDIX

A SCRIPT OF AGING ANALYSIS IN SYRA TOOL 98

LIST OF TABLES

Table Page

1. Summary of NBTI Degradation Models 26

2. Summary of TD Aging Models 42

3. Demonstration of Setup Violations 73

4. Setup Violations in ISCAS89 Circuits 76

5. Fresh and Aged Circuit Analysis 81

6. Setup Violation Comparison between Models 84

LIST OF FIGURES

Figure Page

1.1 Variability Effects due to Scaling 1

1.2 Reliability Effects due to Scaling 2

1.3 Lifetime Limitations from NBTI and HCI 3

1.4 Shift in RO Frequency with Scaling 4

1.5 6T SRAM Cell Sizing with Area 6

1.6 Vmin Dependence on Size of Cache Array 7

1.7 Process Variations with Scaling 8

1.8 Aging Variability with transistor area and Scaling 9

1.9 V_{th} and Mobility Change due to NBTI 10

1.10 V_{th} shift due to Various Gate Bias 11

1.11 Shift in frequency in 11 state RO 11

1.12 RelXpert Framework 15

1.13 Reliability Analysis: Bottom-to-Top Approach 18

2.1 Reaction-Diffusion Mechanism (a) 1D diffusion in NBTI (b) 2D Hot Carrier Trapping in HCI 22

2.2 Recovery of V_{th} shift as a function of Temperature 27

2.3 Discrete V_{th} shifts observed during Pure Recovery 28

2.4 Trapping/De-trapping Mechanism due to Capture and Emission in MOS Device 29

3.1 Variation in (a) Time Exponent from RD Model (b) Drastic Variation in Lifetime 31

Figure Page

3.2 Non-Monotonic V_{th} Shift under DVS 37

3.3 Vth shift during Stress and Recovery 38

3.4 Shift in RO Frequency with Scaling 43

3.5 Microphotograph of Fabricated Test-Chip 44

3.6 Variation of Model Parameters from 20ks 45

3.7 Log(t) Model Prediction of Mean and Variance 46

3.8 Model Validation confirming Logarithmic Dependence 47

3.9 V_{th} shift in DVS: Model Validation 48

3.10 Validation of T-D Model predicting Dynamic Behavior under Voltage Tuning 49

3.11 Log(t) Model Sensitivity to Stress Duration 50

3.12 Aging Prediction under Stress Voltages and Duty Cycles 51

3.13 Long-term Model predicting the Tight Upper Bound from Cycle-to-Cycle Model 52

3.14 Degradation Behavior under Wide Range of Duty Cycle 53

3.15 Aging Prediction in RO due to Device Variations in Aging 54

3.16 Distribution of RNM in a 6T SRAM Cell under both RD and TD Models 55

3.17 Aging Prediction in DVS under Voltage Tuning 56

3.18 Aging Prediction in DVS with Frequency 57

4.1 Asymmetric Aging in Inverter Chain 61

Figure Page

4.2 Critical Points for Aged Timing Analysis 62

4.3 Validation of Delay Model in INV Gate 65

4.4 Validation of Delay Model in NAND2 Gate 66

4.5 Two Switching Cases in NOR2 gate 66

4.6 Validation of Delay Models in Two Switching Cases in NOR2 gate 67

4.7 Statistical Timing Analysis Framework under BTI 69

4.8 Inverter Chain between Two Sequential Elements 72

4.9 Logic and Clock Paths in ISCAS89 Circuits 73

4.10 Distribution of Path Delays, Setup Violations and Timing Violations Increase with Stress Time 74

4.11 Delay Shift in Half Cycle Path of DDR Memory Controller 77

5.1 DVS Operation Pattern in Intel65nm Processor and Predicted Threshold Voltage shift 79

5.2 Distribution of (a) path delays (b) Violations under TD and RD Models.. 83

5.3 Prediction of Delay Shift in a Ring Oscillator under Sequence of Two Voltages 85

Chapter 1

INTRODUCTION

1.1 Introduction

Moore in 1965 predicted that the number of transistors that can be placed on an integrated circuit will be approximately doubled every two years [17]. Although it was initially an observation or forecast, the prediction has turned more into a 'self-fulfilling prophecy' than a law. Moore's law has turned into the driving force for technological advancement in the semiconductor industry. The increased transistor count has directly led to improved capabilities in the digital devices such as processing speed, power, memory capacity, etc. At 45nm, present microprocessors pack close to 800 million transistors and have processing speeds of up to 4GHz.

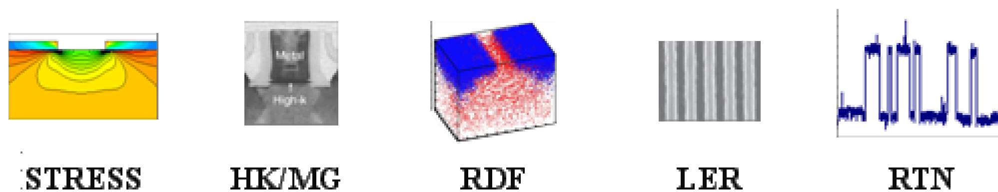


Figure 1.1. Variability effects observed due to technology scaling.

The evolution of electronics will continually be driven by scaling of CMOS technology and the increase in total number of devices per chip. Arising out of the advancement of technology are certain problems. The challenges of cost effectively designing and manufacturing electronic systems are increased in the presence of multiple variability and reliability issues.

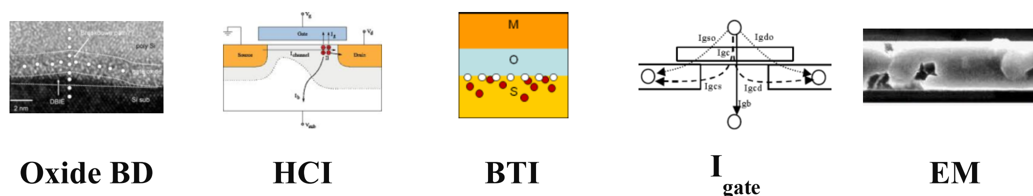


Figure 1.2. Reliability effects seen with increase in operation time.

The variability effects observed with technology scaling are layout dependent stress, high-K metal gate effects (HK/MG), random dopant fluctuations (RDF), line edge roughness (LER), and random telegraph noise (RTN), shown in Figure 1.2. These effects are exhibited after the fabrication of transistors and are termed as the fresh effects impacting the device and circuit performance. Further, aging effects are observed with the operation time, manifesting as the change in device parameters during the course of operation. Circuit performance degrades over time, which is called circuit aging. Such effects are Negative Bias Temperature Instability (NBTI), Channel Hot Carrier (CHC), Time Dependent Dielectric Breakdown (TDDB) etc. These mechanisms have known to affect the transistors since the 1970s but have become more pronounced in the nano-scale regime due to processing and scaling changes introduced to improve device and circuit performance [18]. The result of these aging mechanisms is the degradation in the circuit characteristics such as drain current which ultimately may lead to logic failure and reduce lifetime of a device.

1.2 Limiting Factors of Lifetime

As the reliability concerns become more severe with continuous scaling, it is critical to understand, simulate and mitigate their impact during the design stage. Among these aging effects, NBTI is the primary reliability mechanism which limits the circuit life time, as it increases threshold voltage at the device level. Figure 1.3 shows the plot of critical voltages as a function of the gate oxide thickness. In the high V_{dd} region, hot carrier induced drive current degradation limits the device lifetime. In the low V_{dd} region, however, V_{th} change due to BTI for the PMOS becomes to limit it. and lifetime is limited by NBTI below 180nm technology node.

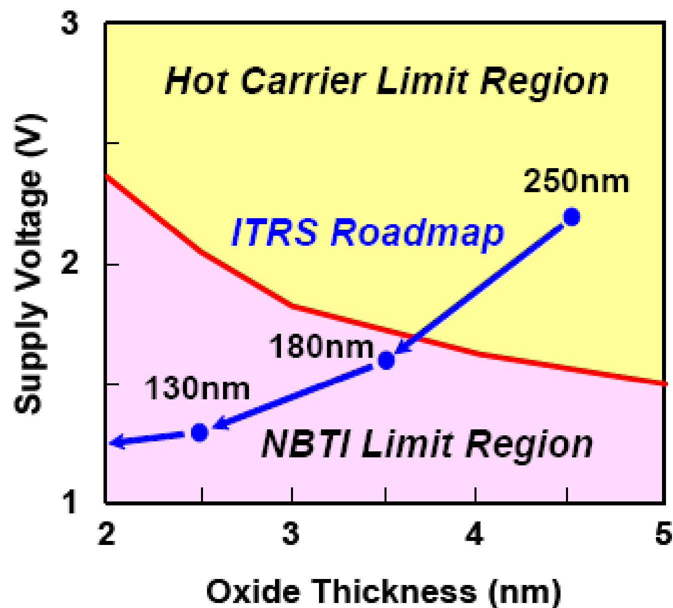


Figure 1.3: The transition of lifetime limitations from NBTI and HCI mechanisms, with gate oxide thickness.

Though the physical explanation of NBTI has been known for decades, it gained attention recently due to aggressive gate oxide scaling and less aggressive

operating supply voltage scaling have exacerbated NBTI degradation. According to ITRS [18], the oxide thickness will reduce less than 1nm and the operating supply voltage for future sub-32nm technology node. Such thin oxide and high voltage results in increase of vertical electric field, which leads to more severe NBTI degradation. The introduction of the high-k dielectrics from 45nm technology can further expedite the NBTI degradation [14]. Since the high-k gate stack actually consists of two layers, namely a high-k gate stack layer and an interfacial SiO_2 layer [15], the overlaying high-k film induces more defects into the SiO_2 interfacial layer which also increases.

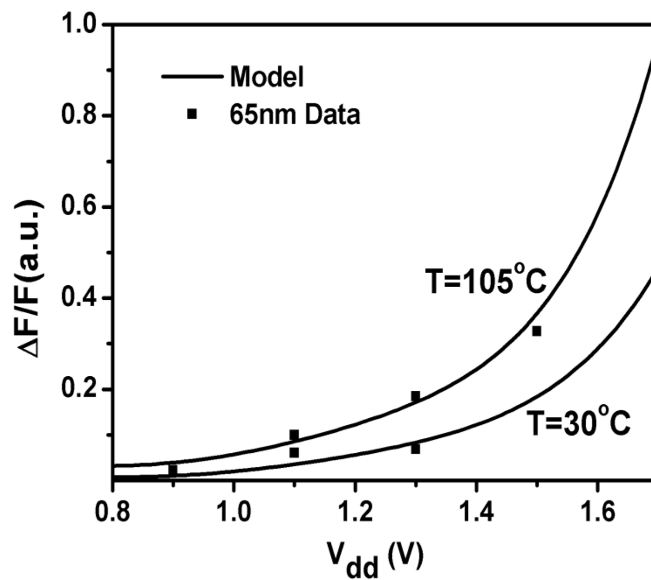


Figure 1.4: Shift in frequency of a RO with technology scaling.

NBTI is prominent in PMOS devices and occurs due to the generation of the interface traps at the Si-SiO_2 interface when a negative gate bias is applied to the PMOS. It manifests itself as an increase in the V_{th} of the PMOS transistor [2,

6, 16–18]. Removal of the stress can anneal some of the interface traps resulting in partial recovery [17], i.e., reduced V_{th} degradation. Over ten years' stress, it shifts V_{th} up to 50mV, translating to more than 20% degradation in circuit speed. In extreme cases, it may even cause functional failures to occur over time [6, 19]. CHC is prominent in NMOS and causes the generation of the interface traps at the Si-SiO_2 interface near the drain end when the gate of NMOS switches. It also result in V_{th} degradation and this degradation cannot be recovered. Due to the V_{th} degradation, these reliability effects result in poor drive current, lower noise margin and shorter device and circuit lifetime [20, 21].

Furthermore, with scaling the shift in RO frequency due to NBTI is studied in Figure 1.4. From the figure, lower V_{dd} helps reduce the degradation rate. On the other side, if V_{dd} is too low, then circuit performance sensitivity to V_{th} shift is elevated, which eventually cancels the benefit due to scaling.

1.3 BTI and its Variability with Technology Scaling

Transistor degradation due to PMOS negative bias temperature instability (NBTI) has become a major reliability mechanism that affects product lifetime for both maximum operating frequency (F_{max}) and minimum operating voltage (V_{min}) in advanced logic technologies. [5], [21] F_{max} degradation due to PMOS NBTI can be handled by either building in enough design margins so that the degradation does not become an issue, by implementing a guard band or both. NBTI induced transistor V_{th} mismatch can further increase memory cell imbalances and degrade V_{min} characteristics over time. In addition to built-in design margin, error correction techniques are often leveraged [65]. Other circuit specific fail modes due to PMOS NBTI degradation can range from degraded analog performance, reduced yield, to even clock slowing down resulting in min-delay fails.

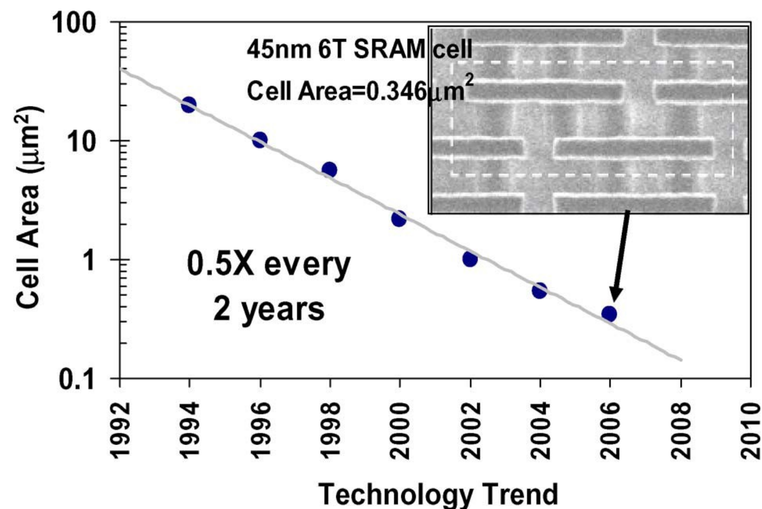


Figure 1.5: 6T SRAM cell sizing with technology scaling showing 2X cell area scaling for every two years [67-68].

The SRAM devices are among the smallest in the technology and the SRAM static noise margin is highly sensitive to the device V_{th} mismatch. Fig. 1.5 shows the 6T SRAM cell area technology trend showing a 2X cell area scaling every two years. As mentioned, the individual transistor footprint in the cache cell is much smaller when compared to the overall cell size, and that layout regularity is used to help minimize process variations. For the recent technology generations, "wide" design topology (as shown in the inset of Fig. 1.5) was used, which helped improve variation by aligning poly in a single direction [66].

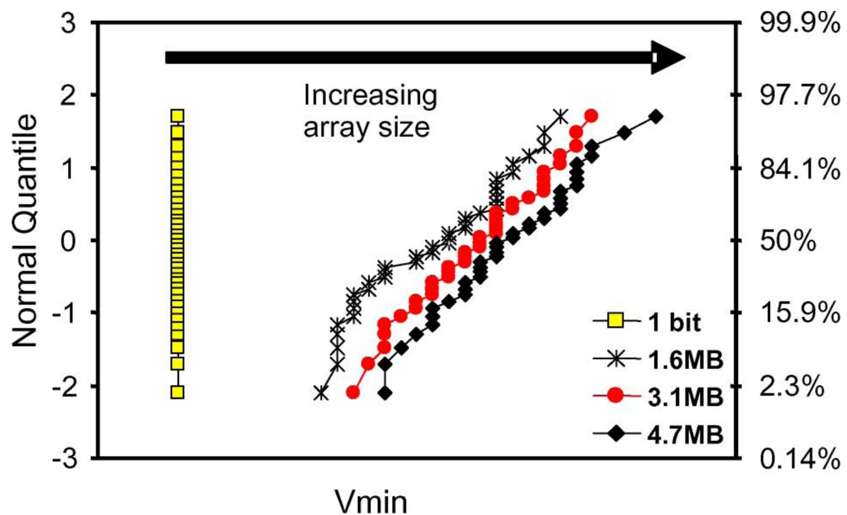


Figure 1.6. Vmin dependence on cache array size of 6T SRAM [67].

Fig. 1.6 shows an example of Vmin dependence on the SRAM cache array size. Both the magnitude of Vmin and the Vmin spread increase as cache array size is increased. This is largely due to device VT mismatch caused by the variability in the cells. Thus, understanding device variability that influences the circuit performance at both time0 and after device aging is critical. This is

becoming increasingly important as multicore desktop, mobile and server CPU products are embedded with increasing cache size. As technology scales, the operating electric field (E-field) in the MOSFET device is pushed to a higher value to meet higher performance needs. Higher E-field of operation may result in higher transistor degradation if nothing is done. The undesired increase in device degradation due to a higher operating field can be mitigated by intrinsic reliability improvement from the process side.

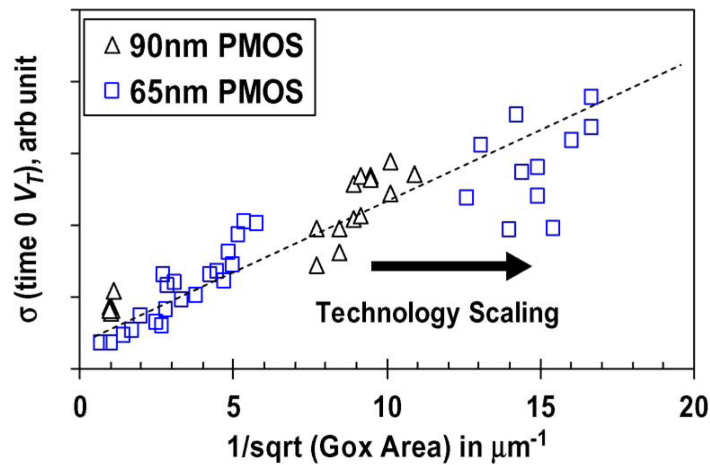


Figure 1.7. Process variations increase with technology scaling [67].

Figs. 1.7 and 1.8 show the V_{th} variability at time0 (before aging) and the variability after aging by NBTI stress, respectively. The increase in variability as a function of technology scaling is due to more random dopant fluctuations in recent technologies. Further, aging variability also increases due to randomness of available traps per device. Figs. 1.7 and 1.8 show that both variability in fresh

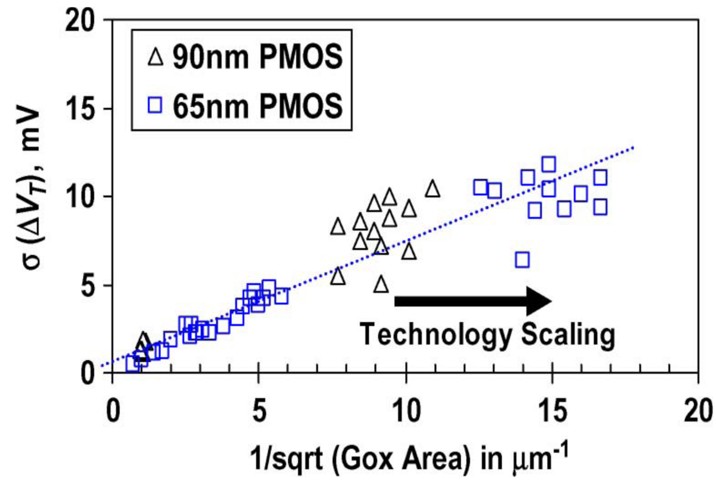


Figure 1.8: Increase in aging variability with technology scaling and function of transistor area [67].

device and aging variability increase inversely as a function of transistor area. This shows that along with variability, reliability is a critical issue in modern technologies and appropriate guard-banding margin needs to be fixed at the design stage.

1.4 Impact of NBTI at Device and Circuit Level

At the device level, the primary and major impact of NBTI is the increase in the absolute value of threshold voltage as shown in Figure 1.9. Mobility is also affected due to the larger Coulomb scattering and sub-threshold slope is increased due to aging.

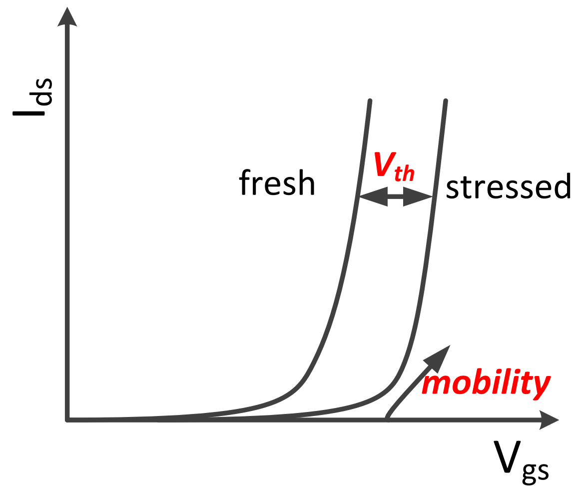


Figure 1.9: Impact on threshold voltage and mobility with aging due to NBTI.

The impact of aging on threshold voltage gets significant with scaling due to thinner oxide thickness. Also, if the amount of electrical stress applied at the gate is large, the degradation worsens leading to a greater increase in V_{th} as illustrated in Figure 1.10. The experiments were conducted on 65nm devices at 105°C where the drain currents are monitored and the V_{th}s are extracted from them. From the figure, it can be seen that time exponent value for V_{th} shift is 0.16 irrespective of the stress voltage applied.

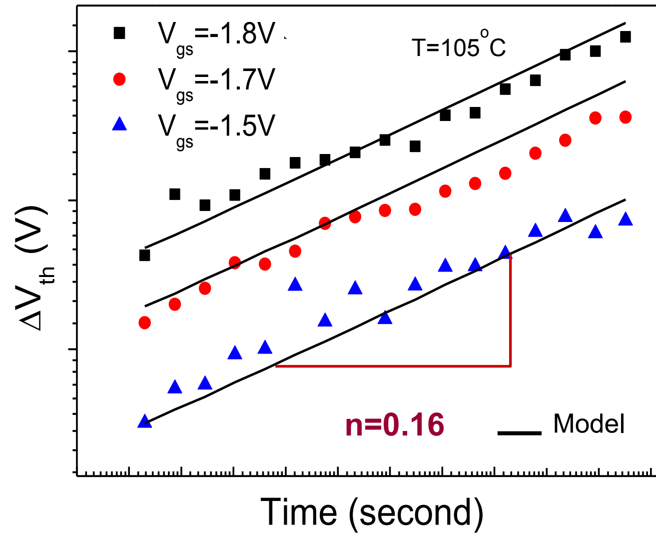


Figure 1.10: Shift in the threshold voltage due to NBTI under various gate stress voltages validated with the model.

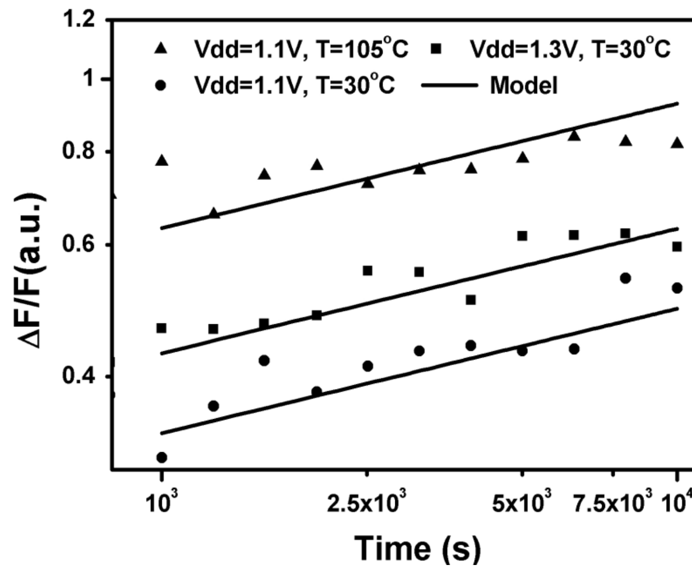


Figure 1.11: Shift in the frequency of 11 stage ring oscillator due to NBTI under various stress voltages and temperature.

At the circuit level, NBTI affects both analog and digital circuits. In digital design, aging impacts speed (major impact), power (both active and leakage),

noise margin, data stability and lifetime [26]. The shift in the threshold voltage decreases the on current, thereby reducing the frequency of operation in digital circuits. Figure 1.11 shows the change in frequency of 11 stage ring oscillator under different temperatures and stress voltages.

1.5 Current Research Efforts on NBTI

NBTI is a design phenomenon and requires the creation of design solutions to properly detect, protect, and solve its potential problems. Traditional research in this area has focused only on improving process technology, while VLSI designs rely on guard-banding to bypass the analysis and optimization of this time-dependent effect. Overall, there are four possible options to cope with this threat:

- Improve process technology: The aggressive scaling of CMOS technology inevitably reduces the reliability of each component, especially that of the thinner gate oxide. As NBTI emerges as the limiting factor of circuit life time in modern microprocessor and SoC design, its degradation rate is much faster as we introduced high-k gate dielectrics since the 45nm node. Further, NBTI is statistical in nature and improving the process technology is critical in order to reduce the aging variability. While traditional reliability research focuses on the improvement of the fabrication technology, it becomes much more difficult to continue that as we work with the extremely scaled transistor.
- Guard-bands: The idea is to slow down the clock frequency based on the worst degradation due to extreme temperature, voltage, percentage of time the PMOS transistors are on, etc. With worst-case guard-banding, all chips may suffer significant reduction in speed although most of them may not be stressed to worst levels in the field. Increase in standard deviations of aging further enhances worst-case guard-bands.

• Circuit Failure Prediction and Self-Correction: Circuit failure prediction [22] predicts the occurrence of a circuit failure before errors actually appear in system data and states. This is in contrast to traditional error detection where a failure is detected. As the reliability concerns become more severe with continuous scaling, it is critical after errors appear in system data and states. The idea is to collect information about various system parameters over time concurrently during normal system operation or during periodic on-line self-test, and analyze the collected data for failure prediction purposes. This approach is applicable for aging because of its gradual degradation characteristics.

• Simulate, predict, and mitigate its impact in the design stage: Some design techniques have been proposed to minimize NBTI effect, such as gate or transistor level sizing [21], input vector control [23], technology mapping and logic synthesis [24]. The implementation of these techniques relies on the worst case estimation of the circuit performance degradation during the design stage. Traditional worst case estimation does not consider the unique property of NBTI, which causes an excessive amount of over-margining. It is significant to develop new simulation method to accurately predict the circuit performance degradation.

To date, research work on NBTI has been active only within the communities of device and reliability physics [19-22, 45-50]. This is partially due to its complexity and emerging status, lack of design knowledge and CAD tools for managing the NBTI degradation [42-43]. Leading industrial companies develop their own models and tools to handle this effect. These tools, however, are usually proprietary and customized to a specific technology. Figure 1.12

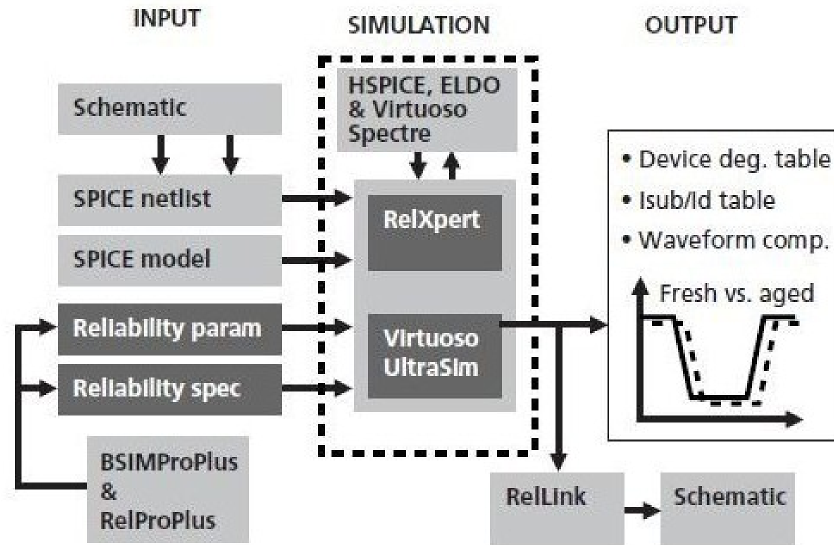


Figure 1.12: Framework of RelXpert, a commercial reliability tool.

shows the framework of RelXpert (extended version of *Berkeley reliability simulator*), a commercial reliability tool. In this framework, several device and reliability parameters are needed, and aging is predicted based on extrapolation and initial assumption of stress waveform. However, this tool is computationally expensive and consumes large portion of the memory usage. To overcome this problem, a generic simulation tool that can accurately predict the degradation would be extremely useful. A good circuit-aging simulator must have the traditional levels of high capacity, high speed and high accuracy. However, the simulation of NBTI effect is fundamentally difficult, since circuit degradation rate depends on both process and operation conditions such as V_{dd}, temperature (T), and input signal duty cycle (α) [30]. These parameters are not spatially or temporally uniform, but vary significantly from gate to gate and from time to time due to the uncertainty in circuit topologies and operations. While we can use high

temperature and high V_{dd} to guard-band the degradation, the strong dependence on α poses a particular challenge. Depending on the signal duty cycle and input patterns, over 75% of previous NBTI-induced degradation can be annealed by biasing the PMOS gate at supply voltage (V_{dd}) [1-5].

The simple static analysis may provide an extremely pessimistic estimate and consequently, result in over-margining. To estimate the degradation bound under various α's, a rudimentary approach resorts to exhaustive simulations. Yet such method is inhibitive in computation cost, especially for circuits with a large number of inputs. Under such condition, in order to improve circuit reliability and design predictability, it is critical to model, simulate, predict, and mitigate its impact in the design stage.

1.6 Thesis Contributions

NBTI effect was widely accepted to be caused by the Reaction-Diffusion (RD) mechanism, triggered by the breaking of Si-H bonds. However, RD model does not account for key recovery features and variation in the time exponent from silicon measurement. In this work, we propose compact aging models based on Trapping/De-trapping (TD) mechanism starting with the distribution of various trap properties. The model is statistical in nature, predicting the aging variability over time. Further, techniques such as Dynamic Voltage Scaling (DVS) are employed in today's design, in order to aggressively reduce power consumption. Such techniques further complicate the aging prediction as the stress and recovery phases are mixed, since operation under lower voltage results in recovery. To accurately handle aging prediction under these scenarios, this thesis proposes a cycle-to-cycle model, that predicts dynamic aging under DVS and a long-term model, that predicts the tight upper bound of multi cycle aging. The proposed model facilitates designers to avoid time consuming atomistic simulations to predict aging caused due trapping/de-trapping. The model accurately predicts the aging in low power operations such as dynamic voltage scaling, overestimated by the conventional t^n model.

The accurate aging models at the device level are critical to predict the circuit aging. The timing paths which meet the timing requirements in the fresh circuit may turn critical over time due to aging, leading to a timing violation. In this work, System level Reliability Analyzer (SyRA) tool is developed to estimate aging in circuit logic paths and estimate the timing violations. The proposed tool

is integrated with standard Static Timing Analysis (STA) flow and is capable of aging computation in circuits with thousands of gates with minimum design overhead. Further, the tool is used to estimate the guard banding overhead when RD and TD models are used and shows that RD based NBTI models significantly overestimates the aging.

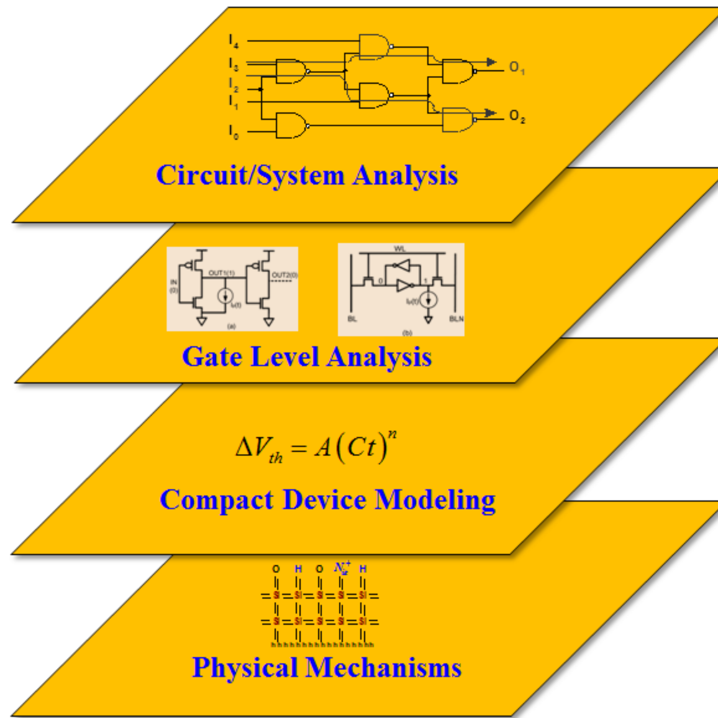


Figure 1.13: Bottom-to-top approach of proposed reliability analysis

This thesis presents a bottom-to-top approach of reliability analysis from device level to system level aging, as shown in Fig. 1.13. Physical mechanisms contributing for the aging are investigated at the device level. Device level compact aging models, that predict the V_{th} shift of the transistor under various operating conditions are proposed. Further, gate level models that use device

models to predict delay shift at the gate levels are presented. Further, the timing violations in sequential circuits due to aging are estimated. The entire approach transfers the microscopic understanding of trapping/de-trapping physics into system level reliability.

1.7 Thesis Outline and Organization

The organization of the report is as follows: chapter 2 presents the background of NBTI effect and presents the reliability physics of RD and TD mechanism. Trapping/De-trapping based compact aging models are proposed and validated with silicon data in chapter 3. Failure diagnosis due to asymmetric aging in sequential circuits is demonstrated using SyRA tool in chapter 4. Chapter 5 presents the statistical aging prediction at the circuit level under trapping/de-trapping and quantifies the TD model accuracy. Chapter 6 summarizes the key contributions of this thesis along with recommendations for future work.

Chapter 2

RELIABILITY PHYSICS

2.1 Introduction

NBTI (in PMOS devices) and CHC (in NMOS devices) are major reliability concerns in nanoscale regime. In fact, as the gate oxide thickness reaches 4nm, the threshold voltage (V_{th}) degradation of the PMOS transistor caused due to NBTI becomes the limiting factor of the circuit lifetime instead of the channel hot-carrier effect in the NMOS transistor. NBTI occurs when a high voltage is applied at the gate of a PMOS transistor at elevated temperature and is explained based on the reaction-diffusion (RD) mechanism. In NBTI, when a high negative electrical stress is applied at the gate of PMOS, interface charges are generated at the Si-SiO_2 interface [30]. NBTI manifests itself as an increase in V_{th}. Removal of stress anneals some of the interface traps resulting in partial recovery [29].

In this chapter, the reaction-diffusion (RD) model and its shortcomings are presented. Further, the physics of trapping/detrapping mechanism and its role in NBTI is presented. This chapter ends with the illustration of necessity of trapping/de-trapping based NBTI models.

2.2 Reaction-Diffusion Mechanism

Till date, the RD model is the only model that successfully explains the power-law dependence of shift in the threshold voltage due to both NBTI and CHC. Since both NBTI and CHC can be physically described as the generation of interface traps (charges) at the Si-SiO_2 interface, this common theoretical framework, RD model was proposed to explain both the effects. This model assumes that when a gate voltage is applied, it initiates a field dependent reaction at the semiconductor-oxide interface that generates interface traps by breaking the passivated Si-H bonds. Figure 2.1 shows the cross-section of a transistor to illustrate RD model.

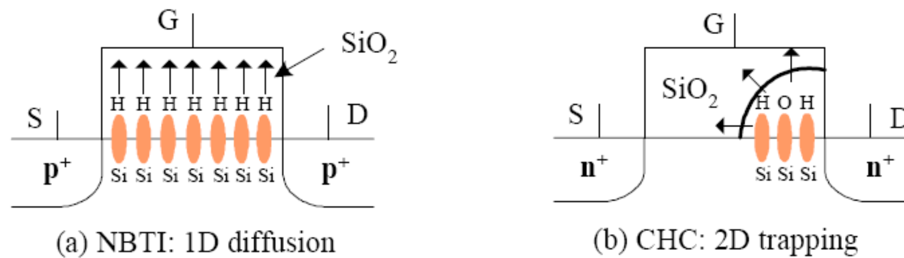


Figure 2.1: The reaction-diffusion mechanism (a) NBTI: 1D hydrogen species diffusion (b) CHC: 2D hot carriers trapping

There are two critical phases described in RD model

1. Reaction, in which some Si-H or Si-O bonds at the substrate/gate oxide interface are broken under electrical stress [33]. The species that trigger such reactions can be positive holes in NBTI and hot electrons in CHC [28]. Consequently, the interface charges are induced, causing increase of V_{th}. Given

the initial concentration of the Si-H bonds (N_o) and the concentration of inversion carriers (P), the generation rate of interface traps, N_{IT} is given by

$$\frac{dN_{IT}}{dt} = k_F (N_0 - N_{IT}) P - k_R N_H N_{IT} \tag{2.1}$$

where k_F and k_R are the reaction rates of forward and reverse reactions, N_0 is the hydrogen concentration. Akin to other reactions, the generation rate is an exponential function of electric field and temperature; it is also proportional to the density of reaction species, namely holes and electrons [31].

2. *Diffusion*, in which reaction generated interface charges diffuse away from the interface toward the gate, driven by the gradient of density. While NBTI happens uniformly in the channel, CHC impacts primarily the drain end, as shown in figure 2.1. The process influences the balance of reaction and is governed by

$$\frac{dN_H}{dt} = D_H \frac{d^2 N_H}{dx^2} \tag{2.2}$$

The solution of equations (2.1) and (2.2) exhibits a power-law dependence on the time [23]. The exact value of the power-law index indicates the type of diffusion species. However, the value of time exponent changes with operation regions which will be discussed in detail in future sections.

2.3 NBTI Degradation Model

Negative bias temperature instability (NBTI) affects the drain current, V_{th}, etc., of the PMOS transistors and its impacts have been known since 1960's. However, the aggressive scaling in CMOS technology makes NBTI as one of the significant reliability concerns in nanoscale regime. Due to the difference in flat band voltage, the NMOS transistor has a negligible level of holes in the channel and thus, does not suffer from NBTI degradation.

For a PMOS transistor, there are two phases of NBTI depending upon the bias condition of the gate. During the phase 1 when $V_g=0$ (i.e., $V_{gs}=-V_{dd}$), interface traps are generated diffusing the hydrogen atoms broken from Si-H bonds towards the gate. This phase is referred as "stress" or "static NBTI". In phase 2, when $V_g=V_{dd}$ (i.e., $V_{gs}=0$), the PMOS device is under pure recovery as hydrogen atoms closer to the interface diffuse back to the interface and anneal the broken Si-H bonds. This phase is referred as "recovery" and has a significant impact on the estimation of NBTI during the dynamic switching in digital circuits.

2.3.1 Temporal Degradation due to Static NBTI

During the stress phase of NBTI, the forward reaction of equation (2.1) dominates. During the initial period of the stress phase, trap generation is slow [10]. Hence $\frac{dN_{IT}}{dt} \approx 0$ and $N_{IT} << N_0$. Thus, the equation (2.1) reduces to

$$N_H N_{IT} \approx \frac{k_F}{k_R} P N_0 \qquad (2.3)$$

With the continuation of forward reaction, H atom is produced and two H atoms combine to form H_2 molecule. The concentration of H_2 (N_{H2}) is related to the concentration of H (N_H) using

$$N_{H_2} = k_H N_H{}^2 \tag{2.4}$$

Driven by the gradient of the generated H_2 density, the H_2 current diffuses into the oxide and is governed by equation (2.2). After a time t, the diffusion front is at a distance of $\sqrt{D_{H_2} t}$ from the Si-SiO_2 interface. The total number of interface charges produced after time *t* is twice the number of H_2 molecules generated during that time since there are two hydrogen atoms in the hydrogen molecule. Thus

$$N_{IT} = 2 \int_0^{\sqrt{D_{H2} t}} N_{H_2}(x) dx \tag{2.5}$$

Solving equations (2.2) through (2.5), the number of interface traps generated as a function of time *t* is given by

$$N_{IT} = \frac{\sqrt{k_H} k_F N_0 P}{k_R} \left(D_{H_2} t \right)^{1/6} \tag{2.6}$$

where, the inversion hole density $P = C_{ox}(V_{gs} - V_{th})$

2.3.2 Temporal Degradation due to Dynamic NBTI

In Dynamic NBTI, stress phase is accompanied by recovery phase where the forward reaction term in equation (2.1) becomes zero, and no net generation of interface traps take place during this phase. Dynamic NBTI corresponds to the phase where the PMOS transistor undergoes alternate stress and recovery periods.

The number of interface traps generated during NBTI can be calculated using similar methodology and the amount of absolute shift in V_{th} as a function of interface charge is given by

$$\Delta V_{th} = \frac{qN_{IT}}{C_{ox}} \tag{2.7}$$

Table 1: Summary of NBTI degradation models

$\Delta V_{th}(t)$	Static NBTI	$(K_v{}^2 t)^n$
	Dynamic NBTI	$\left(\frac{n^2 K_v{}^2 \alpha C t_1 t}{\xi_1{}^2 t_{ox}{}^2 (1-\alpha)}\right)^n$
K_v	$\left(\frac{qt_{ox}}{\varepsilon_{ox}}\right)^3 K^2 C_{ox}(V_{gs} - V_{th})\sqrt{C}\exp\left(\frac{2E_{ox}}{E_0}\right)$	

The expressions for temporal degradation of threshold voltage due to NBTI in both static and dynamic cases are summarized in Table 1. C_{ox} is the oxide capacitance, E_{ox} is the oxide electric field and *n* is the time exponent, characteristic of NBTI. The theoretical value for power law index, *n* in both static and dynamic NBTI is 0.16. However, the variation of the time exponent from silicon data shows that deterministic modeling from RD approach cannot account for the aging variability.

2.4 RD Theory: Limitations

RD theory well explains the degradation behavior during the stress phase and predicts the aging under various bias voltages. However, the primary limitation of RD model is that it is deterministic in nature and cannot account for device to device aging variability. In addition, RD model predicts that the recovery upon stress removal is independent of the electric field and temperature during the recovery phase. However, Figure 2.2 shows that it is dependent on the temperature applied during the recovery [37], contradicting RD theory. Similarly, [37] show that recovery is also a function of the oxide electric field.

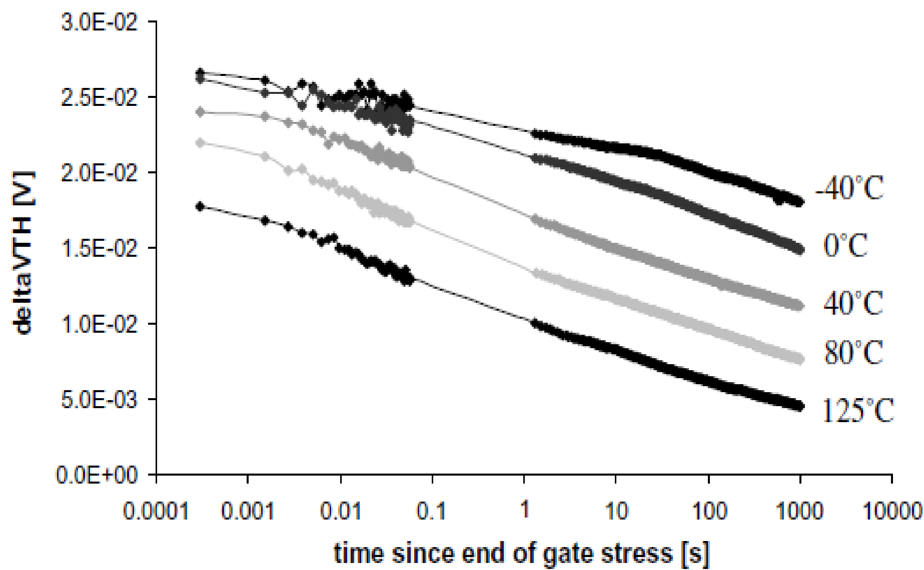


Figure 2.2: Recovery of V_{th} shift as a function of temperature [37].

These limitations question the physical background from RD theory and incline more towards trapping/de-trapping theory that explains these limitations.

2.5 Trapping/De-trapping Theory

Several works show the role of charge trapping/de-trapping mechanism in NBTI degradation as opposed to classical Reaction Diffusion (RD) theory [37]. Clear steps showing single trapping or de-trapping events have been reported through discrete V_{th} shifts [38]. Figure 2.3 shows the measured V_{th} of a device under pure recovery with discrete V_{th} shifts due to trapping/de-trapping events, confirming the necessity for TD based NBTI models for reliable aging prediction. According to trapping/de-trapping theory, traps located in the gate dielectric or at the silicon interface capture and re-emit some of the charge carriers responsible (Figure 2.4), for the current flowing between source and drain of a MOSFET [12]. When a trap captures a charged carrier, the drain current decreases because the number of carriers available in the channel changes and the charged trap state

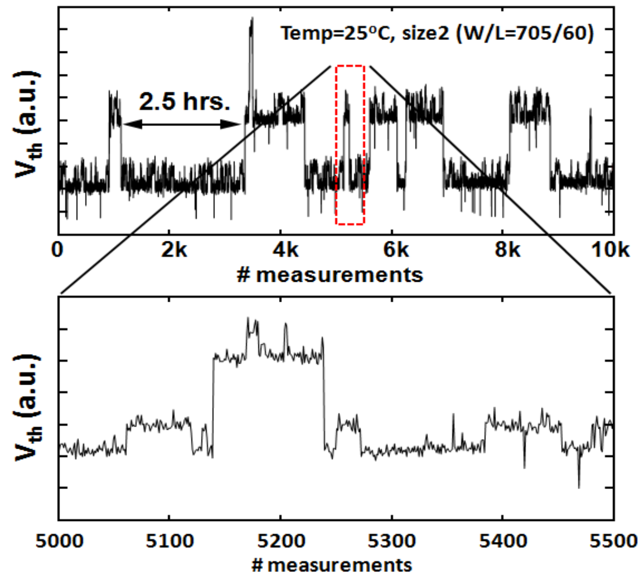


Figure 2.3: Discrete V_{th} shifts observed during pure recovery.

modulates the local V_{th} and acts as a scattering source and thus reducing the mobility. The occupation probability for the trap to be filled is independent of the stress time and the gradual shift in number of occupied traps results in the time evolution of V_{th}. Faster traps (i.e., shorter capture time constants) have a higher probability of getting filled compared to the slower traps. Further, the trap occupation probability increases with gate bias and temperature.

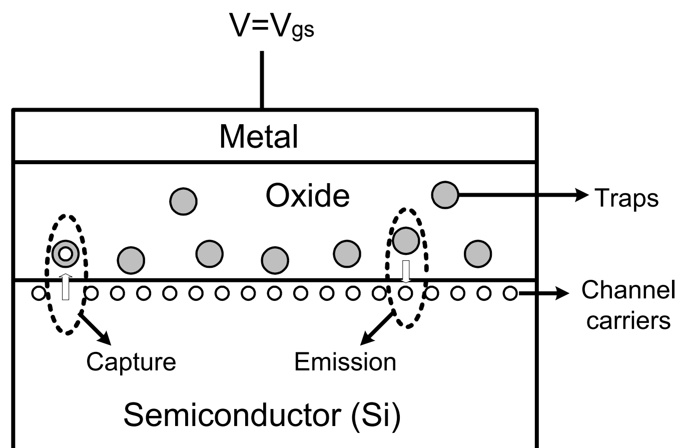


Figure. 2.4. Trapping and de-trapping mechanisms due to capture and emission in a MOS device.

The same model can be applied to the recovery process with the decrease in number of occupied traps. When the stress is removed, trap might emit the charge carrier and results in recovery. The probability of trapping depends on trap properties such as time constants, number of traps, location of traps and trap energies. Developing a compact aging model taking these trap properties into account is not trivial, however, will be significant for designers to predict aging and its variability.

Chapter 3

COMPACT AGING MODELS

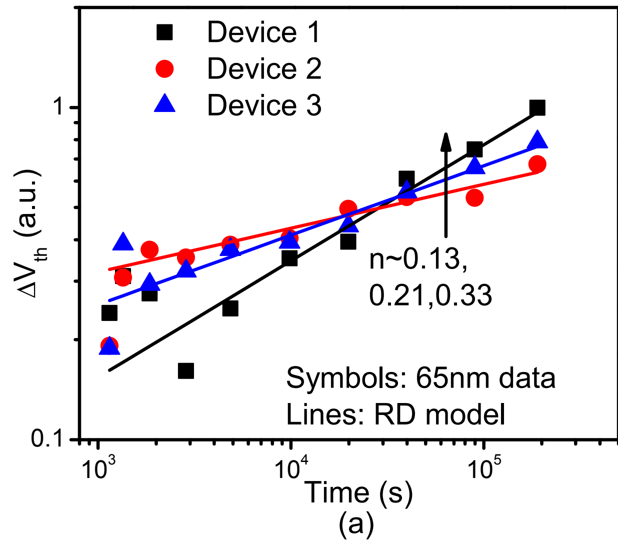
3.1 Motivation

As discussed in chapter 2, NBTI mechanism has been historically been explained by the classical Reaction Diffusion (RD) model, where holes initiate the breaking of Si-H bonds [3] through an electro-chemical reaction and result in the generation of interface states. This increase in the interface charge leads to increase in the absolute V_{th}, thus impacting the operation speeds of digital circuits [7]. According to RD theory, the threshold voltage shift (ΔV_{th}) follows a power law relationship with time: $\Delta V_{th} \propto t^n$ where the time slope, *n,* is an RD characteristic. The value of *n* is typically ~ 0.16 and is expected to be independent of process and operating parameters. However, Figure 3.1(a) illustrates that devices stressed under identical conditions and from the same process exhibit a wide range of time slopes, similar to that observed at the circuit [35]. Further, a small difference in *n* leads to a significant change in long-term prediction of the circuit lifetime, as shown in Figure 3.1(b), and this complicates the guard banding strategy.

To address this concern and improve the accuracy of reliability prediction, this work investigates the variability in aging using 65nm silicon data and proposes a robust prediction method based on the Trapping/De-trapping (TD) mechanism. The new model correctly predicts the mean and variance of V_{th} shift in the long term and demonstrates the following.

1. The temporal degradation follows a logarithmic behavior rather than the conventional power law expectation of the RD model.

2. The RD based t^n model significantly overestimates the aging and its variability.



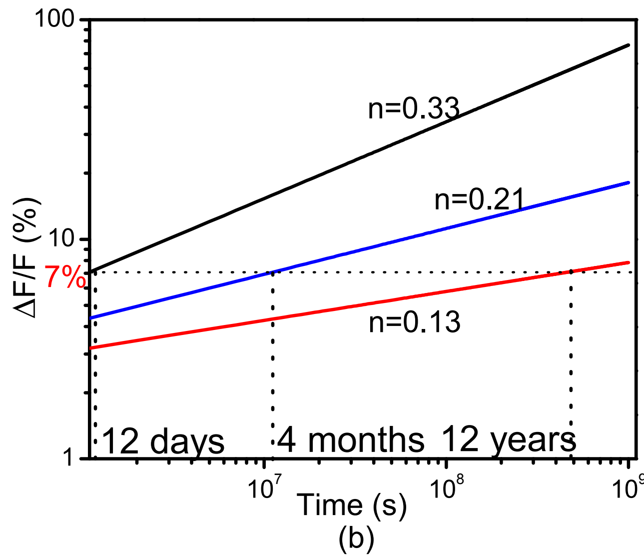


Figure 3.1. (a) Time exponent of the RD model extracted from 65nm silicon data (w/$V_{gstress}$=-1.8V) (b) A drastic variation of lifetime prediction for a ring oscillator (RO) assuming a 7% frequency shift at the end of the lifetime.

3. The log(t) model derived from trapping/de-trapping theory correctly predicts the aging variability due to the stochastic nature of number of available traps.

3.2 Model for Static Stress

The basic assumptions in the modeling effort based on T-D theory are the same as the ones used in modeling of low-frequency noise, since the charge trapping dynamics (capture and emission time statistics) that contribute to the degradation of device performance over time is similar to that causing low-frequency noise [59], [60]. The three main assumptions of the trap properties are

- The number of traps follows a Poisson distribution, common for a discrete process.
- Capture and emission time constants are uniformly distributed on logarithmic scale. This microscopic assumption is critical to derive the logarithmic time evolution at the macro scale.
- The distribution of trap energy is approximated as a U-shape, which is verified by silicon measurement and key to the voltage and temperature dependence on aging [61].

The location of traps is assumed to be close to the interface, which is appropriate for dielectric thickness<2nm.

Using the distributions of trap properties, atomistic Monte-Carlo simulations can be performed to estimate the threshold voltage shift. However, such simulations are time consuming and impractical to predict aging at the circuit level. Our focus is to use the assumptions of the trap properties and to derive a compact aging model that is statistical in nature. The model should facilitate the aging and its variability prediction during the design stage.

Based on the T-D theory, the V_{th} shift at a given stress time is the result of a number of traps ($n(t)$) occupied by the channel carriers [62]. The probability of a particular trap, initially empty (state 0) to be occupied (state 1) after an elapsed time t is given by $P_{01}(t)$. This occupation probability can be calculated by observing that

$$P_{01}(t+dt) = P_{01}(t)p_{11}(dt) + P_{00}(t)p_{01}(dt) \tag{3.1}$$

where $p_{01}(dt)=1/\tau_c$ and $p_{11}(dt)=1- p_{10}(dt)=1/\tau_e$. Solving the differential Eq. (2) from time t_0 to t,

$$\int_{P_{01}(t_0)}^{P_{01}(t+t_0)} \frac{dP_{01}}{\frac{1}{\tau_c} - \frac{1}{\tau_{eq}} P_{01}(t)} = \int_{t_0}^{t+t_0} dt \tag{3.2}$$

Integrating Eq. (3.2) leads to

$$P_{01}(t+t_0) = \frac{\tau_{eq}}{\tau_c}\left(1 - e^{-t/\tau_{eq}}\right) + P_{01}(t_0)e^{-t/\tau_{eq}} \tag{3.3}$$

where $1/\tau_{eq}=1/\tau_c+1/\tau_e$. τ_c, τ_e are random in nature, representing capture and emission time constants respectively, and dependent on bias point and temperature. The values are determined by [22]:

$$\tau_c = 10^p(1+e^{-q}) \tag{3.4}$$

$$\tau_e = 10^p(1+e^{+q}) \tag{3.5}$$

where $p \in [p_{min}, p_{max}]$, where p_{min} and p_{max} define the time constants for fastest and slowest traps respectively (p_{min}~1 and p_{max}>10). This assumption of the existence of defects with wide distribution of time constants is in line with recent NBTI data [19], [25]-[26]. Since p is assumed to be uniformly distributed, the

characteristic time constants are uniformly distributed on logarithmic scale. The parameter q is given by $(E_T-E_F)/kT$, where E_T is the trap energy and E_F is the Fermi energy level. The trap energy (relative to Fermi energy) is a function of applied electric field. Consequently, τ_c and τ_e are dependent on voltage and temperature.

The occupation probability of the trap at time t, assuming that it is under constant stress from time t_0=0 is obtained by substituting $P_{01}(0)$=1-$P_{01}(0)$=0 in Eq. (3.3):

$$P_{01}(t) = \frac{\tau_{eq}}{\tau_c}\left(1 - e^{-t/\tau_{eq}}\right) \tag{3.6}$$

By integrating the occupation probability and multiplying with the number of available traps, the average number of occupied traps, $n(t)$, is obtained:

$$n(t) = \left(\sum_{Ntr=0}^{\infty} \frac{N^{Ntr}e^{-N}}{N_{tr}!}\right)\int P_{01}(t, \tau_c, \tau_e) \tag{3.7}$$

where N is the Poisson parameter for the trap distribution. Substituting the logarithmic distribution of time constants, and the U shaped distribution of trap energies:

$$n(t) = \frac{N}{\ln 10(p_{\max} - p_{\min})}\int_0^{ET\max} \frac{g(E_T)dE_T}{1 + \exp\left(-\frac{E_T - E_F}{kT}\right)} \cdot \int_{10^{-p\min}t}^{10^{-p\max}t} \frac{e^{-u} - 1}{u}du \tag{3.8}$$

where $g(E_T)$ is the trap energy distribution, and p_{min} and p_{max} represent fast and slow traps, respectively. The trap energy, E_T changes as a function of electric field (E_{ox}). Assuming p_{min}~1 and p_{max}>10, and E_T~1/E_{ox},

$$n(t) = \frac{N}{\ln 10(p_{\max} - p_{\min})}\exp\left(\frac{\beta V_g}{T_{ox}kT}\right)\exp\left(\frac{-E0}{kT}\right)\left[A + B\log 10^{-p\max}t\right] \tag{3.9}$$

Eq. (3.9) describes the aging under a constant stress voltage and temperature. Similar to the R-D model [5], it is an exponential function of the stress voltage, temperature and T_{ox}. Furthermore, it has a statistical nature with N, an index for the number of traps per device. For the simplicity of model derivation and data analysis, Eq. (10) is compacted as:

$$\Delta V_{th}(t) = \phi\left[A + B\log(1 + Ct)\right] \tag{3.10}$$

Eq. (3.10) shows the logarithmic relation of degradation with stress time in contrary to the power law behavior. Such a time evolution has a far-reaching impact on the aging behavior.

3.3 Aging Prediction in DVS Operation

Today's circuits typically have a reduced activity factor (or duty cycle) through DVS, to reduce power consumption. Therefore, a significant portion of the operation is under pure recovery or under lower supply voltage. Traditional modeling methodology cannot be applied to the model NBTI degradation under such low-power circuit applications. This section focuses on developing compact aging model in DVS based on T-D theory.

3.3.1 Cycle-to-Cycle Model

Since the degradation is highly sensitive to the voltage (Eq. (3.9)), dynamic voltage scaling leads to different amounts of circuit aging. Figure 3.2 illustrates such two cases when the stress voltage changes. In case 1, stress

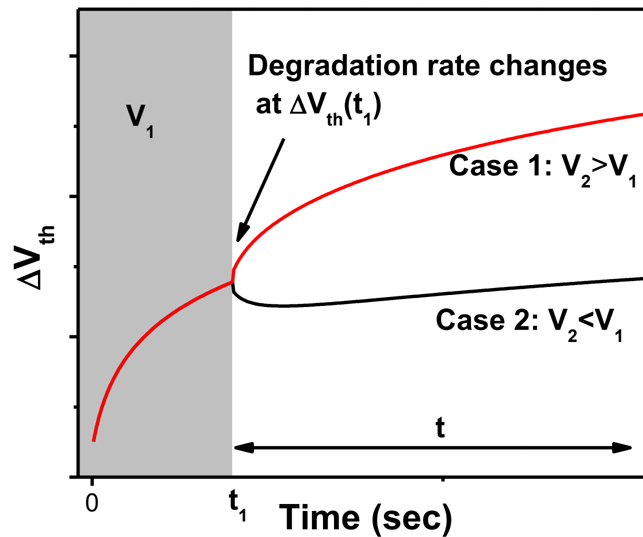


Figure 3.2. The V_{th} shift under DVS is non-monotonic; when the stress voltage is changed from V_1 to V_2, the degradation rate changes. If $V_2<V_1$, the device experiences initial recovery and degradation eventually catches up; If $V_2>V_1$, ΔV_{th} is accelerated.

voltage is changed from V_1 to a higher voltage V_2 and in case 2, V_2 is lower than V_1. To handle such a voltage transition, we replace t_0=t_1 to calculate the occupation probability at time t (time elapsed after t_1) using Eq. (3.3):

$$P_{01}(t+t_1)=\frac{\tau_{eq2}}{\tau_{c2}}\left(1-e^{-t/\tau_{eq2}}\right)+P_{01}(t_1)e^{-t/\tau_{eq2}}$$

(3.11)

where τ_{eq2}, τ_{c2} represent the time constants under voltage V_2. Using Eqs. (3.4) and (3.5), τ_{eq1}= τ_{eq2}, since τ_{eq} depends only on parameter p, which is independent of the voltage. Substituting this property and $P_{01}(t_1)$ from Eq. 7) in Eq. (3.11):

$$P_{01}(t+t_1)=\frac{\tau_{eq}}{\tau_{c1}}\left(1-e^{-t/\tau_{eq}}\right)-\frac{\tau_{eq}}{\tau_{c2}}\left(e^{-t/\tau_{eq}}-e^{-(t+t_1)/\tau_{eq}}\right) \tag{3.12}$$

Following similar steps from Eqs. (3.6) to (3.10), we arrive at a closed form solution for the degradation:

$$\Delta V_{th}(t) = \phi_2\left[A + B\log(1 + Ct)\right] + \phi_1.B\left[\log\left(\frac{1 + C(t + t1)}{1 + Ct}\right)\right] \tag{3.13}$$

where ϕ_1 corresponds to the voltage V_1 and ϕ_2 corresponds to V_2. Eq. (3.13) predicts the aging under multiply supply voltages used in DVS operation. When the voltage is changed to a lower voltage, traps emit some of the charge carriers, and the number of occupied traps reaches a new equilibrium. This behavior is shown in Figure 3.3, where the V_{th} shift initially recovers and the degradation eventually catches up. Such a non-monotonic behavior is correctly predicted from Eq. (3.13), which is a sum of two components. For $t \ll t_1$, the second component dominates and recovery is observed; and for $t \gg t_1$, the second component saturates to a constant value and the first component increases logarithmically with stress time. When the voltage is changed to a higher voltage, the degradation rate increases at the point of transition. Eq. (3.13) helps to develop aging

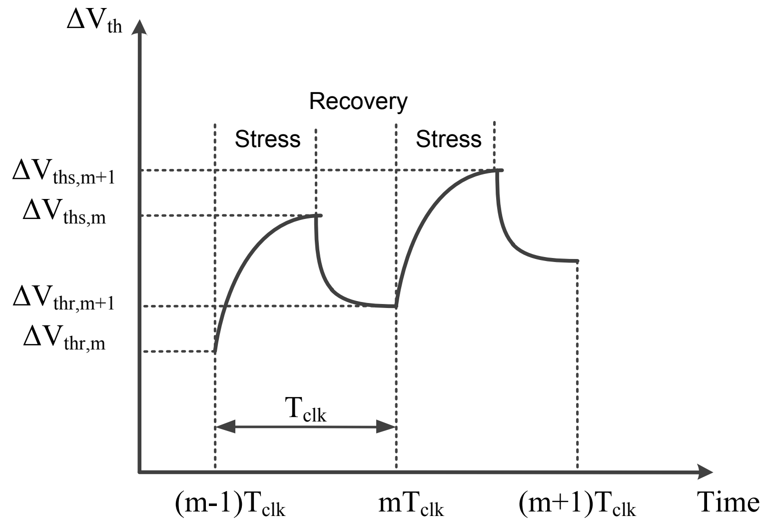


Figure 3.3. The V_{th} shift during the stress and recovery, representing the parameters used in multi cycle model.

prediction method under DVS for multiple cycles.

3.3.2. Multi Cycle Prediction

Aging prediction under voltage tuning for a single cycle is shown in the previous sub-section. Eq. (3.13) is used to obtain the aging under stress and recovery for $m=t/T_{clk}$ cycles. Eq. (3.13) can be re-written as:

$$\Delta V_{th}(t) = \phi_2\left[A + B\log(1+Ct)\right] + \Delta_{Vth}(t_1)\left[1 - \frac{k+\log(1+Ct)}{k+\log(1+C(t+t_1))}\right]$$

(3.14)

Figure 3.3 shows the change in degradation during the stress (higher voltage) and recovery (lower voltage) cycles. Assuming stress starts at the onset of *(m-1)*T_{clk}:

$$\Delta V_{ths,m} = \phi_1\left[A + B\log(1+C\alpha T_{Clk})\right] + \Delta_{Vthr,m-1}\beta_{1,m} \quad (3.15)$$

where

$$\beta_{1,m} = 1 - \frac{k+\log(1+\alpha T_{Clk})}{k+\log(1+C(mT_{Clk}+\alpha T_{Clk})} \quad (3.16)$$

where ϕ_1 corresponds to the voltage V_1, T_{Clk} is the clock period and α is the ratio of stress time under voltage $V_1$to the clock period. Similarly,

$$\Delta V_{thr,m} = \phi_2\left[A + B\log(1+C(1-\alpha)T_{Clk})\right] + \Delta_{Vths,m}\beta_{2,m} \quad (3.17)$$

$$\beta_{2,m} = 1 - \frac{k+\log(1+(1-\alpha)T_{Clk})}{k+\log(1+C(mT_{Clk}+(1-\alpha)T_{Clk})} \quad (3.18)$$

where ϕ_2 corresponds to the voltage V_2. Eqs. (3.15) and (3.17) facilitate to predict the aging under multiple cycles of DVS, thereby enabling the long-term prediction.

3.4. Long-term Prediction

A long term model to predict a tight upper bound based on multi-cycle prediction is necessary, in order to directly estimate aging at the end of a given operation time. Based on the multi cycle model in the previous subsection, $\Delta V_{ths,m}$ and $\Delta V_{ths,m+1}$ are connected by:

$$\Delta V_{ths,m+1} = \phi_1\left[A + B\log\left(1 + C\alpha T_{Clk}\right)\right] + \phi_2\left[A + B\log\left(1 + C(1-\alpha)T_{Clk}\right)\right]\beta_{1,m} + \Delta_{Vths,m}\left(1-\beta_{1,m}\right)\left(1-\beta_{2,m}\right) \tag{3.19}$$

Using Eq. (3.19) and repeatedly replacing the $\Delta V_{ths,m}$ by $\Delta V_{ths,m}$ for i=m,….,1, we arrive at:

$$\Delta V_{ths,m} = \phi_1\left[A + B\log\left(1 + C\alpha T_{Clk}\right)\right]\left(1 + \sum_{i=1}^{m}\prod_{j=m-i+1}^{m}\beta_{1,,j}.\beta_{2,,j}\right) + \phi_2\left[A + B\log\left(1 + C\left(1-\alpha\right)T_{Clk}\right)\right]\beta_{1,m}\left(1 + \sum_{i=1}^{m}\prod_{j=m-i+1}^{m}\beta_{1,,j-1}.\beta_{2,,j}\right)$$

(3.20)

Since obtaining a closed form solution for Eq. (3.20) is not straight-forward, we use the property $\beta_{1,m-1} < \beta_{1,m}$ and $\beta_{2,m-1} < \beta_{2,m}$,

$$\Delta V_{ths,m+1} \leq \phi_1\left[A + B\log\left(1 + C\alpha T_{Clk}\right)\right]\left(1 + \beta_{1,,m}.\beta_{2,,m} + \left(\beta_{1,,m}.\beta_{2,,m}\right)^2 + \ldots\right) + \phi_2\beta_{1,m}\left[A + B\log\left(1 + C\left(1-\alpha\right)T_{Clk}\right)\right]\left(1 + \beta_{1,,m}.\beta_{2,,m} + \left(\beta_{1,,m}.\beta_{2,,m}\right)^2 + \ldots\right) \tag{3.21}$$

Eq. (3.21) is a geometric series and the upper bound of degradation is obtained:

$$\Delta V_{ths,m} = \phi_1\left[A + B\log\left(1 + C\alpha T_{Clk}\right)\right]\frac{1}{1-\beta_{1,,m}.\beta_{2,,m}} + \phi_2\left[A + B\log\left(1 + C\left(1-\alpha\right)T_{Clk}\right)\right]\frac{\beta_{1,,m}}{1-\beta_{1,,m}.\beta_{2,,m}} \tag{3.22}$$

Eq. (3.22) is sensitive to the duty cycle, α (ratio of time under V_1 to time under V_2), time period (sum of operation times under V_1 and V_2 for a single cycle), and

the stress voltages. The new solution can accurately predict aging under DVS for wide range of duty cycles and voltage tuning.

Table 2: Summary of TD based aging models.

Constant Stress	$\Delta V_{th}(t) = \phi \cdot [A + B\log(1 + Ct)]$
Cycle-to-cycle	$\Delta V_{th}(t + t_1) = \Delta_1 + \Delta_2$ $\Delta_1 = \varphi\left(A + B\log(1 + Ct)\right),$ $\Delta_2 = \Delta V_{th}(t_1)\left(1 - \frac{k + \log(1 + Ct)}{k + \log(1 + C(t + t_1))}\right)$
Long-term	$\Delta V_{ths,m+1} = \varphi_1\left[A + B\log\left(1 + C\alpha.T_{clk}\right)\right]\frac{1}{1 - \beta_{1m}\beta_{2m}}$ $+\varphi_2\left[A + B\log\left(1 + C\left(1 - \alpha\right)T_{clk}\right)\right]\frac{\beta_{1m}}{1 - \beta_{1m}\beta_{2m}}$ $\beta_{1,m} = 1 - \frac{k + \log(1 + C.\alpha T_{clk})}{k + \log(1 + C(mT_{clk} + C.\alpha T_{clk})}$ $\beta_{2,m} = 1 - \frac{k + \log(1 + C\left(1 - \alpha\right)T_{clk})}{k + \log(1 + C(mT_{clk} + \left(1 - \alpha\right)T_{clk})}$

3.5. Silicon Validation

The new set of aging models proposed uses the first-principle probability equations to predict aging under constant stress and dynamic operations. To validate the models, statistical aging data is collected from a 65nm test chip. Various sequences of voltages are applied to probe both stress and recovery phases, and the models are validated across these voltages in this section.

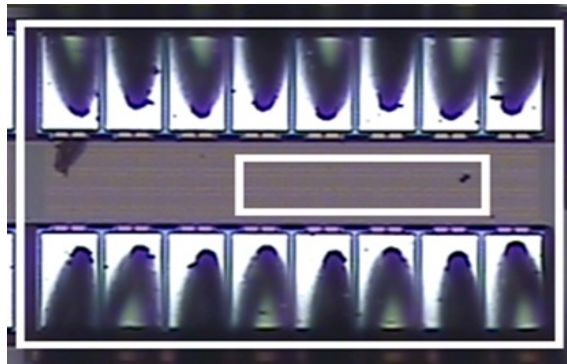


Figure 3.4. Microphotograph of the test chip with area 439x332um^2; 256 PMOS transistors with 4 different aspect ratios present in arrays.

3.5.1. Test Chip and Measurement Plan:

The measurement delay plays a crucial role in NBTI test since even a small measurement time leads to large recovery, resulting in inaccurate aging data. Hence, obtaining degradation data by removing stress from all devices leads to a large measurement error. One solution is to place multiple DUTs on a chip so that stress periods and threshold voltage measurements can be conducted in parallel. This approach is very expensive and needs a larger area. Contrary to

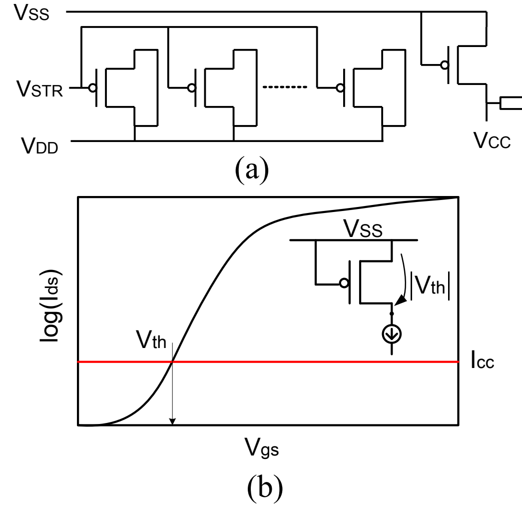


Figure 3.5. Measurement setup: (a) All devices are under stress except the device under measurement (b) V_{th} measurement using constant current method.

parallel measurement method, a parallelize stress period in a pipeline manner is implemented and V_{th} measurements for the DUTs are conducted in this work [35].

Figure 3.4 shows the microphotograph of test chip implemented in 65nm and 11 metal layer CMOS process. Total area of the test chip is 489x332um^2 and 128 PMOS devices of four different aspect ratios are implemented as DUTs. Aging measurements are conducted when all the devices are stressed at 1.8V and a temperature of 125°C for 200ks. A DUT in the transistor array is changed from the stress phase to the recovery phase using pass gates as switches. These measurements are required in order to analyze device to device statistical aging behavior under constant stress and dynamic operations. Further, measurements are conducted when all the devices are stressed under multiple V_{DD} to realize the aging in DVS operation. Figure 3.5 presents the test structure and measurement principle implemented in our work. Except for the device in measurement, all other devices are stressed and V_{th} is measured using constant current method with

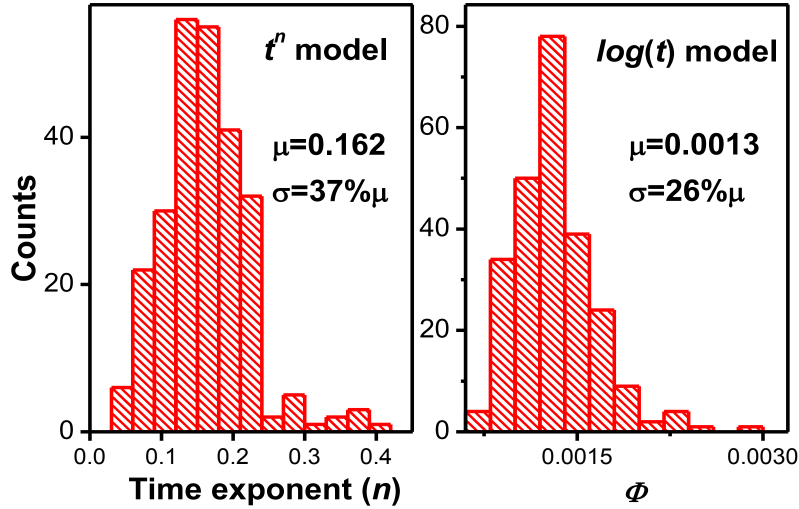


Figure 3.6. Variations of n and ϕ, as fitted from data under 20ks.

a resolution of 0.2mV. Measurement time for each DUT is less than 0.4ms, minimizing the recovery.

3.5.2. Aging Statistics

Statistical aging analysis is performed on the data collected from the test chip. Initial data is collected by stressing all the PMOS transistors at 1.8V for 200ks. From this stress data, model parameters for both power law and logarithmic models are extracted. Figure 3.6 presents the distribution of time exponent n, from Eq. (1) and ϕ from Eq. (11). Other model parameters do not suffer a high amount of variations. The time exponent has a mean value~0.16 (same as its characteristic value) and a standard deviation of 37% mean. Such a variation in n complicates the lifetime prediction, since even a small difference in n leads to a significant change in long term. On the other side, standard deviation in parameter ϕ is approximately 27% mean. The variation of ϕ is usually

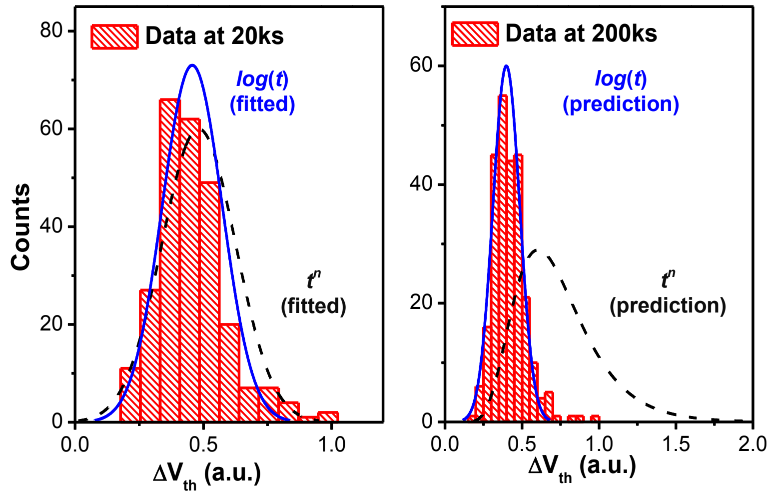


Figure 3.7. The *log*(*t*) model, fitted from data < 20ks, well predicts the long-term statistics, while the t^n model overestimates the aging.

attributed to the randomness in the number of traps [63], which has only linear impact on the V_{th} shift.

To further test the accuracy of model prediction, parameters for both models are initially extracted from stress data < 20ks. These model parameters are used to predict aging to 200ks by extrapolation. Figure 3.7 evaluates the prediction of both models with direct silicon data at 200ks. Although both models fit the data well under 20ks, t^n model overestimates the mean behavior by 40% and variance by 260% at 200ks. This behavior is due to the excessive sensitivity of power law model to the time exponent, n. The log(t) model correctly captures the mean and variance of the V_{th} shift, as it is linear to the parameter ϕ.

3.5.3. Model Validation under Constant Stress and DVS

As indicated in Eq. (3.9), the degradation follows a logarithmic dependence on stress time, and the aging rate is an exponential function of the

stress voltage. Figure 3.8 presents the validation of *log(t)* model under various constant stress voltages. From the figure, it is evident that the degradation has an exponential relation with stress voltage. Such an exponential dependence on voltage is similar to that predicted by the t^n model [3]. Figure 3.8 shows the linear dependence with time when the x-axis is plotted in log scale, confirming the logarithmic dependence on stress time. The time dependence of degradation is the major difference between R-D based t^n and T-D based *log(t)* compact models.

One of the major differences between the R-D and the T-D theory is the recovery behavior [3] [13]. By the R-D mechanism, the recovery is more gradual due to the slow diffusion process; the T-D mechanism expects fast and discrete recovery from individual trapped charges. In this work, various sequences of stress voltages are applied, in order to calibrate the model prediction by the logarithmic mode.

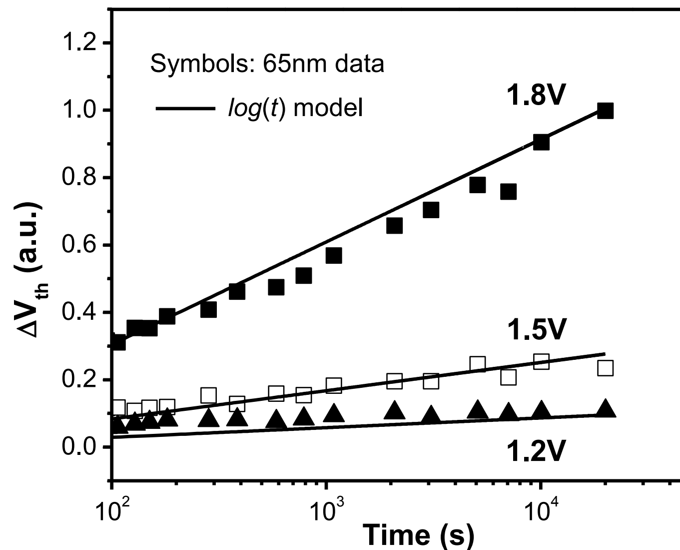


Figure 3.8. The T-D based compact model matches the logarithmic time dependence and exponential voltage dependence.

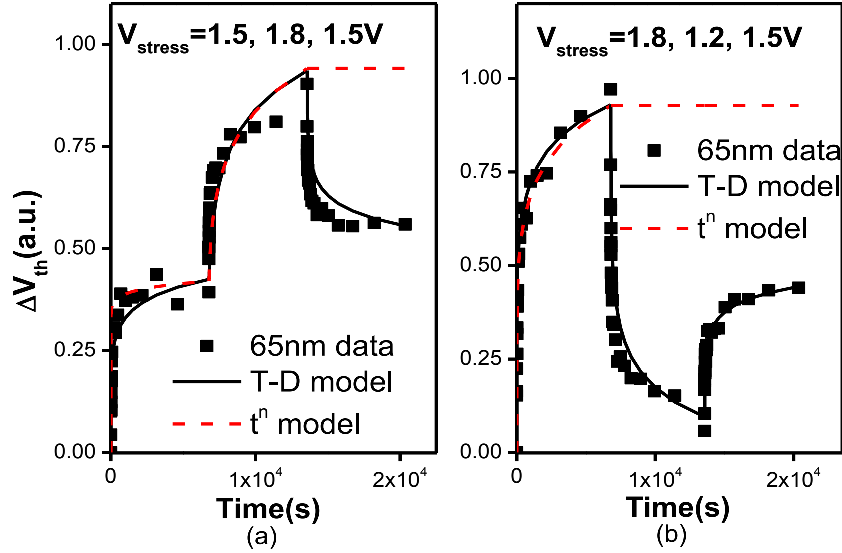


Figure 3.9. Threshold voltage shift under different V_{DD}; t^n model overestimates compared to T-D based *log(t)* model.

Figure 3.9 illustrates two cases where a PMOS device is stressed under different voltages sequentially. In case 1, the device is stressed at 1.5V, 1.8V, 1.5V for 5000s during each period. When the stress voltage is changed from 1.5V to 1.8V, the degradation rate is increased. The number of occupied traps increases and V_{th} shift increases exponentially as shown in Figure 3.9a. This behavior is captured by both t^n and *log(t)* models. When the voltage is changed back to 1.5V from 1.8V, the degradation undergoes recovery due to traps emitting the excess charge carriers (Figure 3.9a). This recovery behavior is captured by trapping/de-trapping model and t^n model predicts a very low degradation at a lower stress voltage using the boundary condition [13]. In case 2, the device is stressed at 1.8V, 1.2V, 1.5V for 5000s during each period. The recovery is furthermore significant when operated under a much lower voltage of 1.2V as presented in

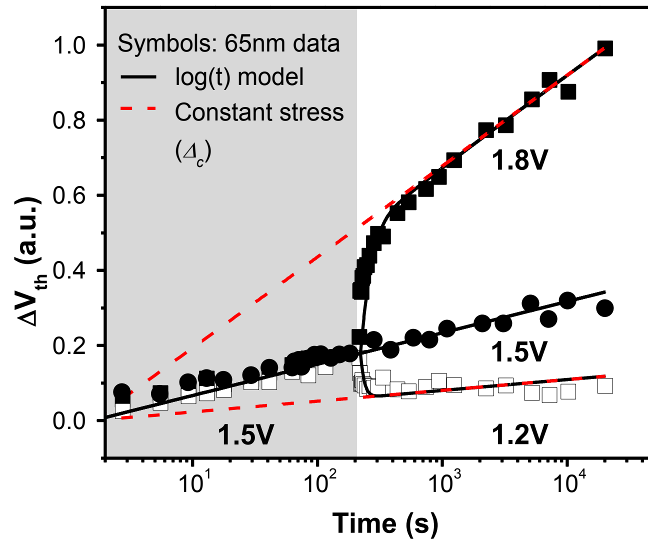


Figure 3.10. The *log*(*t*) model predicts the dynamic behavior under voltage tuning, such as the transition period, and the convergence to the constant stress condition.

Figure 3.9b, which results in a large prediction error in the long term. The recovery behavior is also governed by the *log(t)* function, shown in Eq. (3.14).

Measurements are further performed by changing the stress time under different voltages. After a stress period of 200s under 1.5V, the stress voltage changes to different values, such as 1.8V, 1.5V, or 1.2V. Depending on the second voltage, the degradation rate may increase or decrease. More interestingly, the device may experience a transition period, before the stress goes back to the equilibrium condition. Eventually, the degradation rate goes to the same as the constant stress under the second voltage. This behavior is predicted from Eq. (3.13), where the second component dominates initially, resulting in the recovery. After t>>200s, the second component decays down and first component dominates, leading to the stress under second voltage. Experimental results from

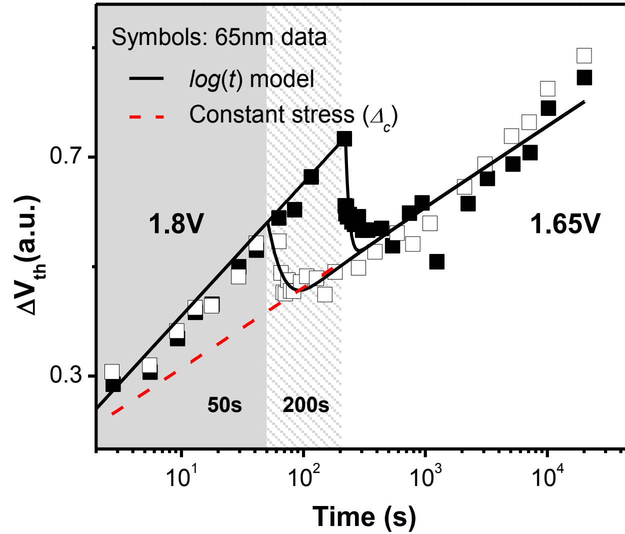


Figure 3.11. The *log*(*t*) model is sensitive to the duration of stress period under DVS, helping capture the dependence on the switching activity.

the test chip are shown in Figure 3.10, well validate these non-monotonic behaviors, supporting further study on aging under DVS.

The two components in Eq. (3.13) play an important role in long-term prediction under multiple cycles. Figure 3.11 evaluates the model prediction, with different periods under the same voltage. In this study, the device is initially stressed under 1.8V, for 50s or 200s; then the voltage is switched to 1.65V. As the stress voltage is lowered, a temporary recovery behavior is observed due to the emission of excessive amount of trapped charges. The *log(t)* model captures such behavior in both the cases. The stress profile for the case where the device is first stressed under 1.8V for 200s is higher compared to the case where 50s initial stress case. However, since in both the cases, the device is later stressed for a much longer time (~10ks) at 1.65V, the degradation converges to the constant

stress condition at 1.65V. This validation helps to predict the aging under various switching activities (α), as described in Eqs. (3.15) and (3.17).

Finally, multi cycle stress experiment is conducted to test the long-term prediction. Figure 3.12 presents the measurement under different patterns of voltages and duty cycles. Figures 3.12(a) and 3.12(b) show the multiple cycle prediction stressed at 1.8V, 1.2V and 1.8V, 0V at α=0.5. When voltage is lowered from 1.8V, a recovery is observed from the silicon data, which is well predicted from the model. Further, if the voltage is lowered down to 0V, the recovery is large as shown in Figure 3.12(b). The cycle-to-cycle model in Eq. (14) is scalable

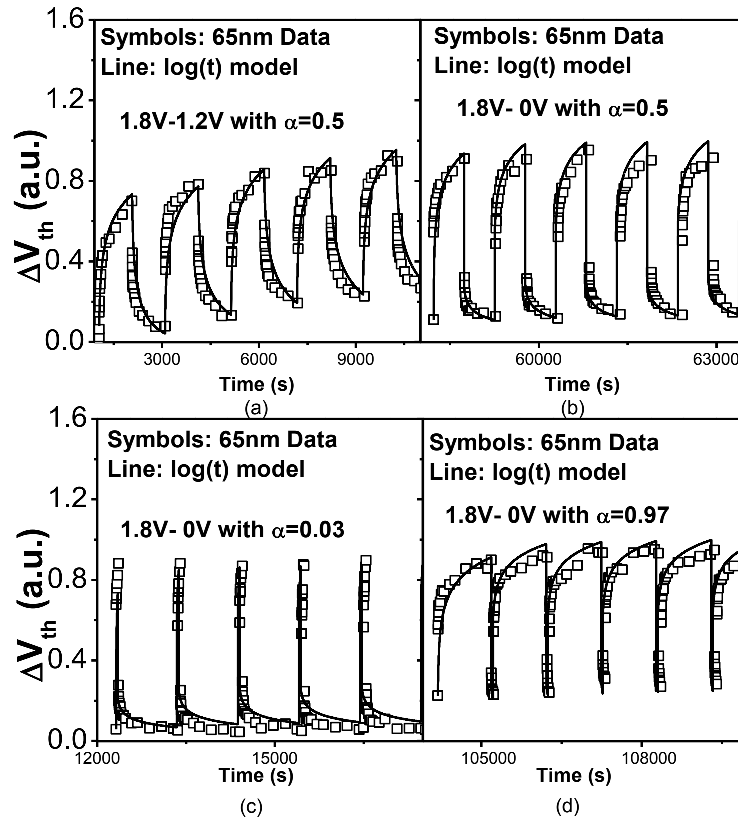


Figure 3.12. Aging prediction under stress voltages 1.8V,1.2V and 1.8V,0V; and with duty cycles 0.03, 0.97.

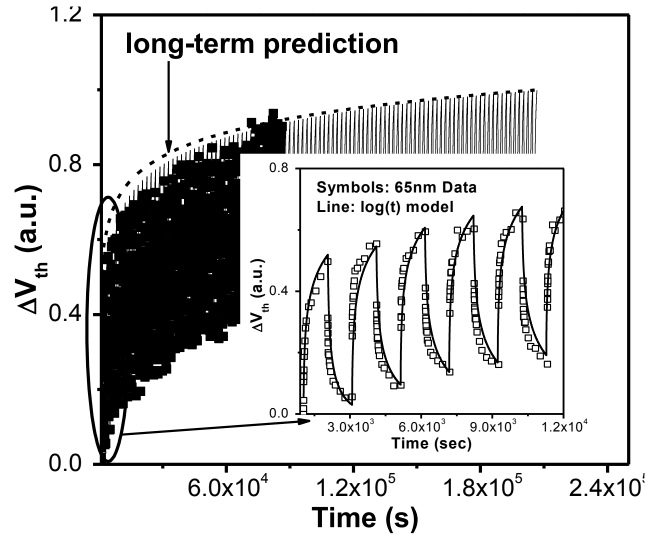


Figure 3.13. Cycle-to-cycle model predicts the dynamic degradation for 40 cycles under 1.8V, 1.2V stress and long-term model tracks the upper bound of dynamic shift.

with different duty cycles. Figs. 3.12(c) and 3.12(d) show the validation of our model with silicon data at very low (α=0.03) and very high (α=0.97) duty cycle values. From the figures, even voltage transition even for a short duration will result in the sudden shift in duration. The degradation increases rapidly when voltage is increased momentarily to 1.8V as illustrated in Figure 3.12(c). Similar behavior is observed in recovery as shown in Figure 3.12(d).

Cycle-to-cycle model is validated for various combinations of voltages and duty cycles. Experimental data is collected for 40 cycles under 1.8V, 1.2V stress and the cycle-to-cycle model captures the dynamic V_{th} shift (Fig. 3.13). The long term model presented in section 3 captures the tight upper bound of cycle-to-cycle prediction, as illustrated in Figure 3.13. Further, the degradation behavior is evaluated under wide range of duty cycles. Figure 3.14 presents the aging behavior under wide range of duty cycles as validated with the silicon data.

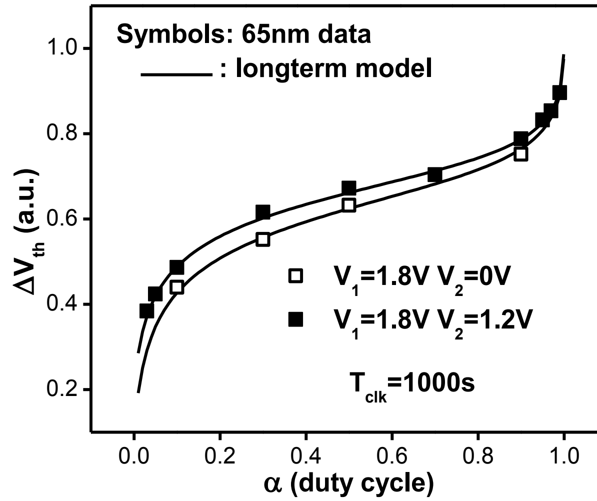


Figure 3.14. Degradation behaviour under wide range of α (0.01 to 0.99) predicted by the long-term model and validated with 65nm silicon measurement.

Moreover, the degradation rate changes rapidly when α~0 or 1, and gradual for intermediate duty cycles. This behavior is due to the sudden change in degradation at the beginning of stress and recovery phases, contributed from the fast traps. The cycle-to-cycle model well matches with the measured results and the long-term model improves the efficiency of aging prediction in DVS, without having to rely on TCAD simulations. The models provide solution to tune operating parameters in order to minimize aging at the circuit level.

3.6. Circuit Aging Prediction under Multiple V_{dd}

In the previous sections, compact models are developed to predict aging under device-to-device variations (due to variability in number of traps) and voltage tuning employed in DVS. In this section, the impact of the proposed models at the circuit level is investigated.

3.6.1 Aging Prediction under Variability

Device level aging due to trapping/de-trapping exhibit large variations due to randomness in number of available traps. The variability at (t>0) is also a function of transistor sizing similar to process induced variations (t=0) [4]. Aging variability increases with downsizing the devices and hence, it becomes significant with CMOS technology scaling. Further, it is important to understand

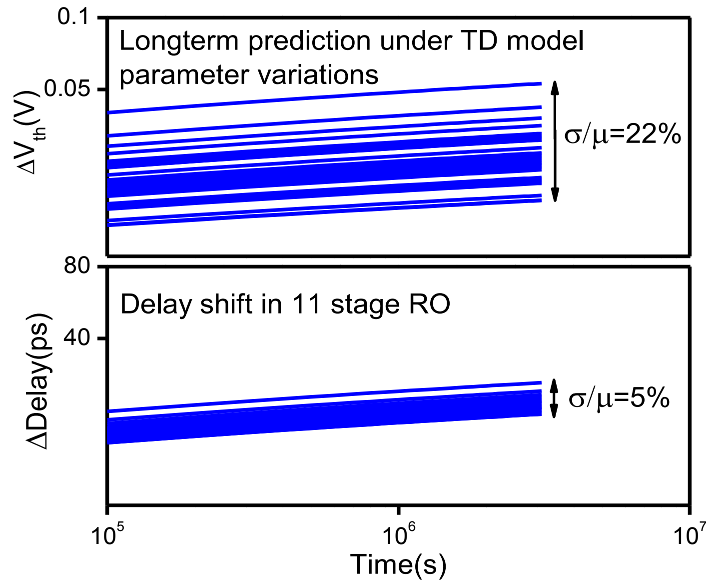


Figure 3.15. Aging prediction under device level variations of PMOS V_{th} shift and delay change in an 11 stage RO.

and estimate the translation of device aging to circuit level degradation [7], [64]. Figure 3.15 presents the long term prediction of PMOS V_{th} shift and ring oscillator frequency shift under device level variations. Though, the device level prediction has a large variation, circuit aging exhibits moderate variation due to

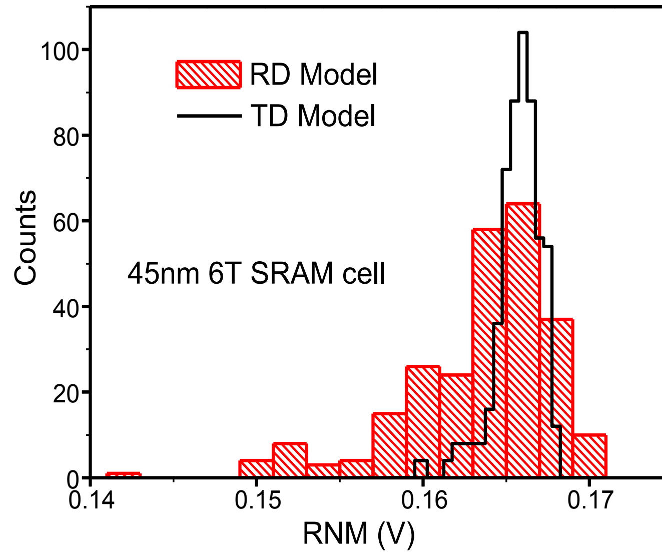


Figure 3.16. Distribution of read noise margin in a 6T SRAM cell under t^n and *log(t)* models.

the average effect of multiple transistors in the circuit. Standard deviation (σ) of V_{th} shift is 22% of mean (μ) from our measurement data. When these V_{th} shifts are induced randomly into an 11 stage Ring Oscillator (RO), σ/μ~5% is observed. Figure 3.16 presents the distribution of Read Noise Margin (RNM) in a 6T SRAM cell predicted with extrapolation using the model coefficients from the short term data. Lower bound from t^n model is much less compared to T-D model, exhibiting a wide distribution. This is due to the excess sensitivity of the model to the time exponent, n. On the other hand, T-D model predicts aging with lesser variations, thereby minimizing the product guard-banding at the design stage.

3.6.2. Multiple V_{dd} in DVS Operation

Multiple factors can be varied to achieve minimum aging and low power consumption under dynamic voltage scaling. With a DVS operation pattern under two different voltages, Eqs. (3.15) and (3.17) predict the aging under various duty cycles, clock periods and operation time. Figure 3.17 explores the dependence of the aging effect on the choice of V_{dd} values, as well as the control of α. Since the amount of degradation is an exponential function of voltage, lower stress voltage helps reduce the degradation rate. The degradation significantly decreases when the ratio of V_{low}/V_{high} drops from 1. However, the reduction rate becomes less as the V_{low} is further lowered, since the operation period with V_{high} dominates the entire degradation and therefore, further scaling of V_{low} does not help minimizing the aging. Lower α, i.e., the shorter period in V_{high} operation, is still effective in reducing the aging, as predicted by the proposed model. Besides, changing the V_{dd}

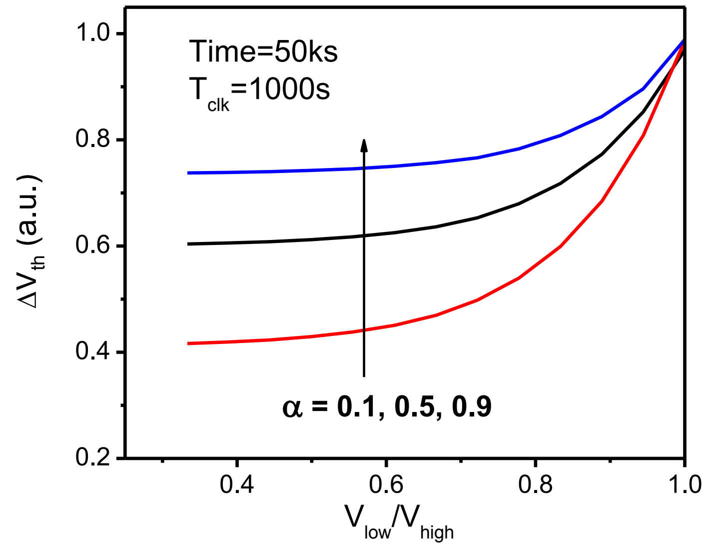


Figure 3.17. Aging prediction from cycle-to-cycle model varying the lower operating voltage in DVS.

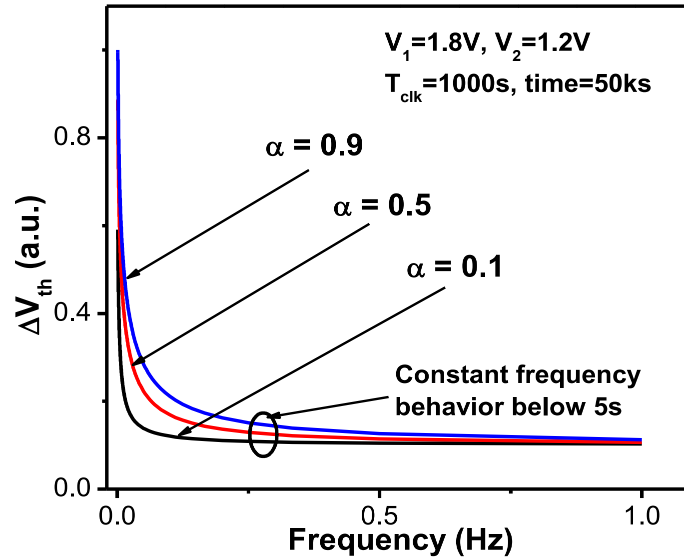


Figure 3.18. Aging prediction from cycle-to-cycle model varying the lower operating voltage in DVS.

and duty cycle, frequency can be tuned to minimize the reliability. Figure 3.18 shows the degradation behavior with wide range of frequencies. The degradation decreases as the frequency is increased and finally saturates to a constant value, as shown from experimental data in [19].

Chapter 4

FAILURE ANALYSIS OF ASYMMETRIC AGING

4.1 Circuit Aging

NBTI manifests itself as an increase in the PMOS threshold voltage (V_{th}), resulting in an increase in gate and path delays of digital circuits [42]. This phenomenon implies that many circuit paths that are not critical in the design stage may turn critical over time, causing timing violations during the operation [25]. The initial performance margins need to be increased to minimize the impact of product drift on the system [45].

It is essential to include NBTI aging in circuit timing analysis to guarantee the circuit lifetime. However, timing analysis under aging is not trivial since NBTI has a strong dependence on technology specifications and more importantly on operating conditions [7]. The analysis of NBTI effect is more complex compared to traditional reliability effects, such as Hot Carrier Injection (HCI), since NBTI exhibits both stress and recovery during dynamic circuit operation. PMOS device under NBTI undergoes degradation or recovery depending on the operating voltages in Dynamic Voltage Scaling (DVS). Conventional reliability analysis assumes constant stress conditions and do not account for the more complex operating patterns in today's digital circuits, such as operating in standby modes for long periods of time [7]. When a digital circuit is operated in standby mode, PMOS devices with low input voltage are under static NBTI stress. Hence, it is crucial to identify the critical moments in the operation pattern which have maximum NBTI effect, thereby causing increased timing violations.

4.2 Previous Works

The performance degradation in circuits has been evaluated using conventional reliability tools such as BERT, RelXpert, etc. [41]. These tools use pre-processor that generates several device and reliability parameters, and intermediate files. Using the waveforms from the first SPICE simulator run, the post-processor generates an age value for each device that quantifies the amount of degradation each device undergoes as a function of time and operating conditions. The primary disadvantage of these tools is too many SPICE parameters are required. Other limitations are change in operating voltages due to degradation over time is not considered and several iterations are required for accurate results.

Several works have proposed reliability analysis in digital circuits [44] [46-48]. Given the delay of a digital gate, aging models are proposed to estimate the delay shifts in [46]. However, it is not trivial to calculate the fresh delays of all the gates in the circuit as they depend on the type of gate, input slew rate and output capacitance. Delay models as a function of slew rate and load capacitance are proposed in [47] for combinational circuit aging analysis. The delay models are complicated as they use Taylor series expansion and Chebyshev polynomial to fit the gate delay degradation. The aging analysis in sequential circuits is more complicated as it involves Static Timing Analysis (STA). Reference [48] proposes statistical STA, performing aging analysis after updating mean and variance of delay distribution. Though this framework is effective, the critical paths which experience maximum NBTI degradation cannot be identified using this flow.

Further, NBTI-based STA framework is proposed in [44] using aging aware library. The standard cell library is updated based on the worst case degradation (assuming all PMOS gates are under stress or at switching activity leading to maximum path delay). This STA analysis does not differentiate gates of similar type with different switching activity.

Along with addressing the limitations stated above, this work focuses on aged timing analysis which is compatible and easily integrated into the standard modeling/circuit analysis flow; supports large scale simulation requiring minimum effort by the designer; and reduces the time involved in performing the aging analysis.

4.3 Asymmetric Aging

In addition to predict the aging sensitivity to operating conditions, it is essential to accurately predict the shift in path delay. Figure 4.1 shows a case of static stress followed by a dynamic operation in an inverter chain, where the PMOS devices in alternate inverters are under static stress. Previous works assume that NBTI effect is averaged out due to alternate stress and recovery phases in the logic path or do not differentiate between rising and falling transient edges [44] [49], resulting in inaccurate prediction in path delay shift. However, since NBTI only affects the low to high (t_{pLH}) delay of an inverting gate (since PMOS V_{th} is shifted), NBTI induced same edge delay change is accumulated along the logic path as illustrated in Figure 4.1. The delay accumulates only at the rising edge and falling edge is unaffected (follows the rising edge shifts of the previous gates) resulting in asymmetric aging [45], and needs to be taken into consideration while performing aged timing analysis.

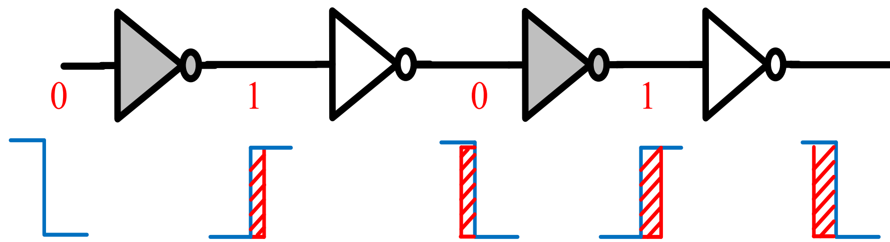


Figure 4.1. Asymmetric aging in an inverter chain, where it is statically stressed with the first input at logic low; delay accumulates only at the rising edges.

Threshold voltage shift (ΔV_{th}) and gate delay shifts (Δt_d) are determined based on the circuit operation conditions and environmental factors. Aging prediction for typical mobile and hand-held products is more complicated as it involves long standby modes due to power reductions achieved through clock gating. Traditional NBTI models handle such situations with an average activity to estimate long term degradation. While a worst case analysis by assuming that all the gates are under static stress is overly pessimistic [44], it is important to track the exact moment when the circuit experiences maximum degradation.

A typical operating pattern in a digital circuit and ΔV_{th} shift under such an operation is presented in Figure 4.2. In dynamic operation, PMOS device is under alternate stress and recovery phases, resulting in moderate degradation. When the circuit is operated in standby mode, PMOS transistor is under static stress leading

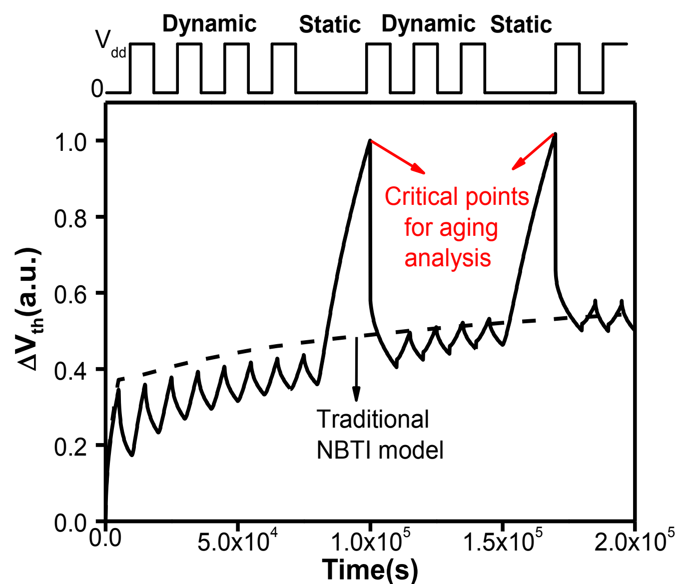


Figure 4.2. Critical points for aged timing analysis in operation pattern with standby modes.

to much larger degradation since NBTI is a strong function of activity factor (α). As illustrated in Figure 4.2, the critical point for aging analysis is the end of the standby mode, as the gate and path delays are maximum at this time instant. Conventional NBTI long term model does not capture these critical moments, leading to inaccurate aging analysis. Hence, such situations need to be appropriately accounted for during the design stage.

In this work, we propose accurate failure diagnosis method for predicting timing violations in sequential circuits under asymmetric NBTI aging. The main features of the proposed aging analysis framework include:

1. Simple models to predict gate delay shifts are proposed using long term ΔV_{th} predictive models. NBTI induced delay increase is calculated using delay sensitivity to V_{dd} in standard cell library, avoiding complicated SPICE simulations to predict aging in gate delays. The key advantage in this delay prediction method is the characterization of library cells under various aging scenarios is not required and the previously characterized cell library can be utilized.

2. Critical moments in circuit operation are identified, at which NBTI induced delay shifts are maximum and aging analysis is performed at these critical points.

3. Setup or hold violation is determined using aging aware timing analysis based on the competition between delay shifts in the logic path and clock buffer; The critical paths with reduced guard banding are recognized in this framework.

4.4 Gate Level Aging Models

Previous works use complicated models to predict delay shifts in digital gates due to NBTI [46] [47]. A simple gate delay model is proposed in this chapter that calculates the delay change due to ΔV_{th} from V_{dd} information in the cell library. The load capacitance and input slew rates impact the delay shifts, and the proposed model is independent of them. This model facilitates the designers to estimate the delay degradation with time directly from the standard cell library, without having to rely on time consuming circuit simulations to characterize the library under several aging conditions. Since the primary impact of NBTI at the device level is the increase in V_{th}, the primary effect at the circuit level is the increase in gate delay. Based on the drain current of a short channel device, the delay of a digital gate (t_d) is expressed by [57]:

$$t_d \propto \frac{CV_{dd}}{V_{dd} - V_{th}} \tag{4.1}$$

where C is the output capacitance of the gate. The change in gate delay when both V_{dd} and V_{th} are subject to change is

$$\frac{\Delta t_d}{t_d} = \frac{\Delta V_{dd}}{V_{dd}} - \frac{\Delta V_{dd} - \Delta V_{th}}{V_{dd} - V_{th}} \tag{4.2}$$

The delay change occurs when only V_{th} is changed (Δt_{dVth}) or only when V_{dd} is subject to change (Δt_{dVdd}) and is given by

$$\frac{\Delta t_{dVth}}{t_d} = \frac{\Delta V_{th}}{V_{dd} - V_{th}} \tag{4.3}$$

$$\frac{\Delta t_{dVdd}}{t_d} = \frac{-V_{th} \Delta V_{dd}}{V_{dd}\left(V_{dd} - V_{th}\right)} \tag{4.4}$$

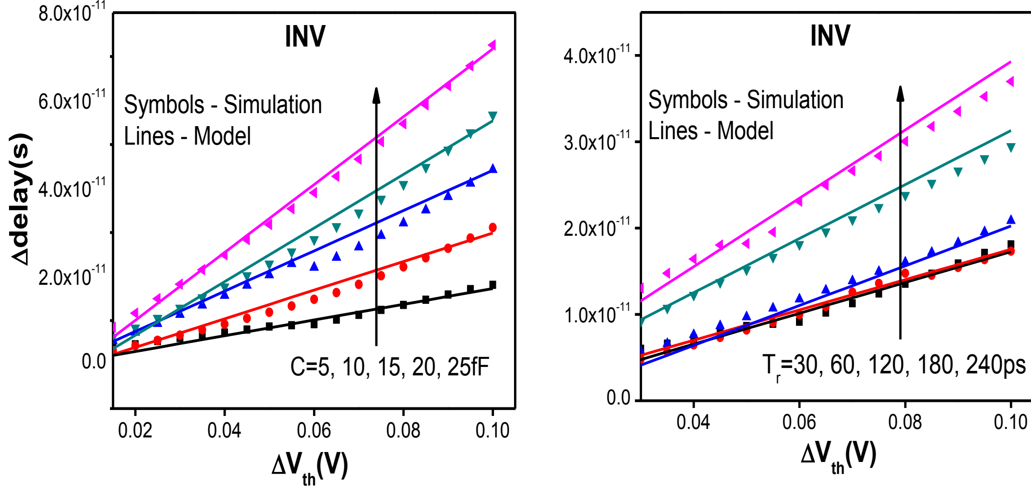


Figure 4.3: Validation of delay model in an inverter under wide output capacitance and slew rate range

The above two equations can be combined to relate Δt_{dVdd} and Δt_{dVth}:

$$\Delta t_{dVth} = -\frac{V_{dd}}{V_{th}}\left(\frac{\Delta V_{th}}{\Delta V_{dd}}\right)\Delta t_{dVdd} \tag{4.5}$$

The above model calculates the delay shift due to threshold voltage shift (with long term RD NBTI models) using the delay function of supply voltage. Figures 4.3 and 4.4 present the proposed model validation with simulation results under wide range of output capacitance (C_L) and input slew rate (T_r).

The proposed model predicts the shift in gate delay in case of inverter and NAND gates where a single PMOS exists between switching input and output, and also between V_{dd} and output. However, the situation is different in gates like NOR, where there are multiple transistors between the switching input and output.

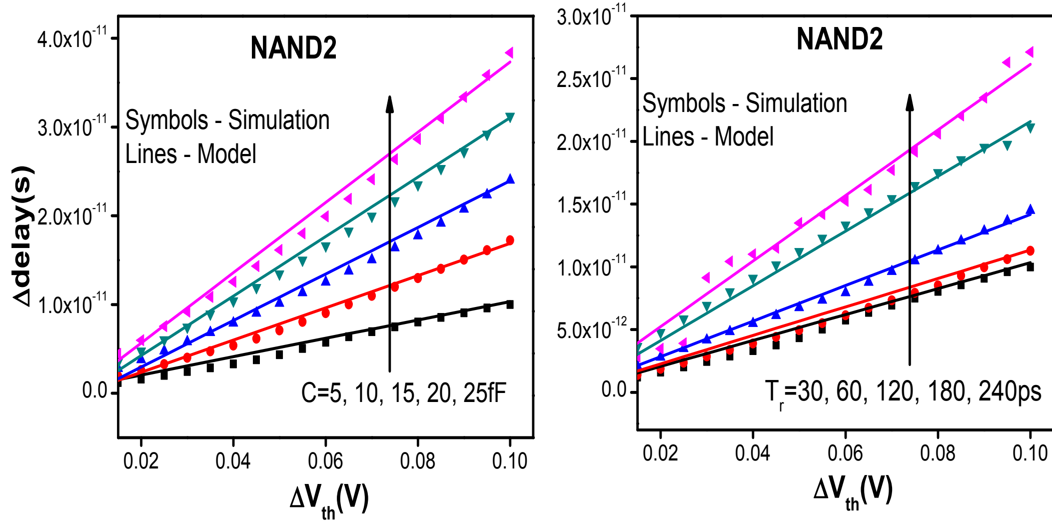


Figure 4.4: Validation of delay model in NAND2 gate.

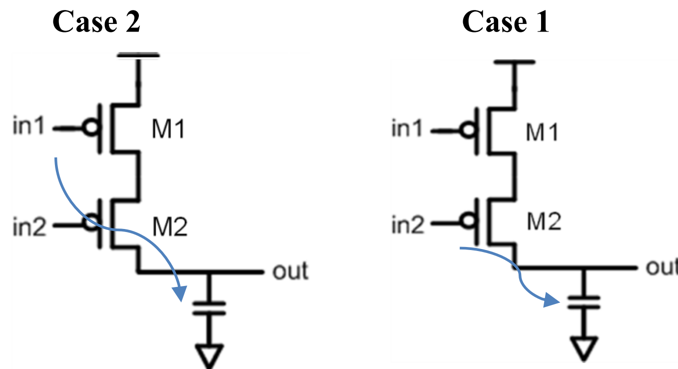


Figure 4.5: Two switching cases in NOR2 gate.

Figure 4.5 shows two switching cases in a 2 input NOR gate. In case 1, two PMOS transistors are present between the switching input (in1) and output, whereas a single transistor is present between switching input (in2) and output in case 2. The proposed model for the inverter handles the situation in case 1, since same number (two) of transistors exists between input and output, and supply voltage and output (analogous to inverter). Hence, it can use the delay sensitivity

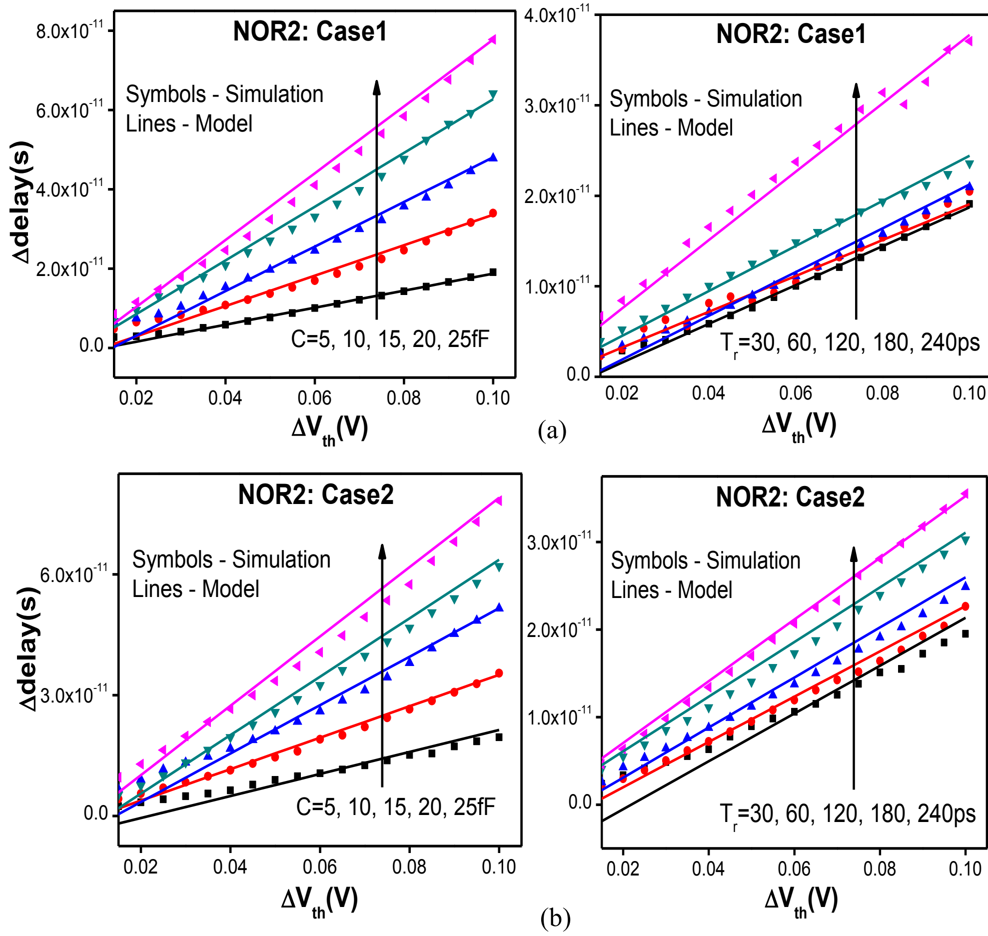


Figure 4.6: Validation of two switching cases in NOR2 gate; nominal capacitance~5fF and nominal slew=30ps.

to supply voltage and predict change in delay due V_{th} shift. However, in case 2, the contribution of two PMOS devices towards the gate delay is different. The V_{th} shift in M2 has a larger impact on gate delay than M1, since M2 is in the path between switching input and output, and also in the path between V_{dd} and output. The shift in delay due to V_{th} shift in this case is modeled by:

$$\Delta t_{dVth} = -\frac{V_{dd}}{V_{th}}\left(\frac{k\Delta V_{th1} + \Delta V_{th2}}{2\Delta V_{dd}}\right)\Delta t_{dVdd} \tag{4.6}$$

where k (~0.25) denotes the contribution of V_{thM1} towards the delay shift compared to that from Vth_{M2} in case 2. For a NOR gate with N inputs, the delay shift can be predicted by

$$\Delta t_{dVth} = -\frac{V_{dd}}{V_{th}}\left(\frac{k\sum_{i=0}^{N-m}\Delta V_{thi} + \sum_{j=0}^{m}\Delta V_{thj}}{N\Delta V_{dd}}\right)\Delta t_{dVdd} \tag{4.7}$$

where m transistors exist in the path of the switching input and output that have more contribution towards delay. The delay model proposed is validated in both the switching cases of NOR2 gate (Figure 4.6). The proposed prediction of delay shift is simple and accurate, enabling reliable failure assessment under NBTI effect.

4.5 Asymmetric Aging Analysis

The device and gate level aging models presented in the previous section are crucial for the asymmetric aging analysis at the circuit level. The experimental setup, simulation method and prediction of timing violations due to aging in logic path and clock buffer are demonstrated.

4.5.1 Simulation Method

Aging analysis in digital circuits can be implemented at the SPICE level where each transistor in the circuit is replaced by the sub-circuit model of the aged device [52]. For larger circuits, replacing every transistor in the circuit is not a practical approach and a gate level timing analysis is required. Figure 4.7 presents the experimental setup and static timing analysis framework implemented in this work. NBTI aging aware library is used to calculate the delay shift in digital gates. Our framework uses delay information under different V_{dd} in standard library and predicts the delay shift due to change in V_{th} using a simple gate delay model, as illustrated in Figure 4.7.

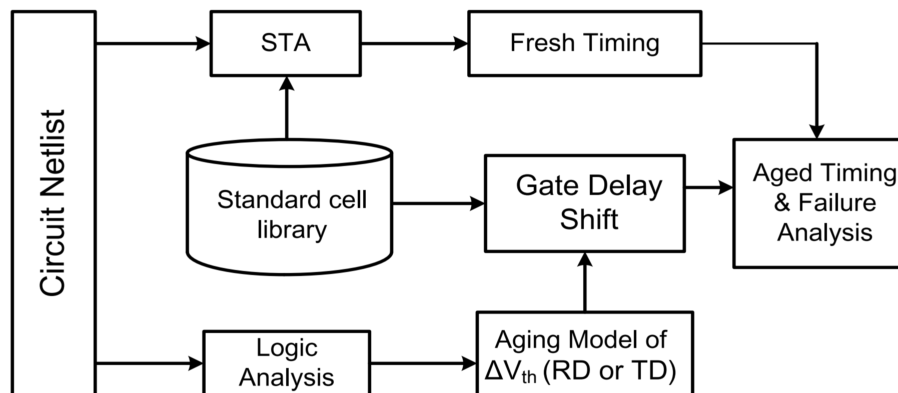


Figure 4.7. Statistical timing analysis framework for failure diagnosis under NBTI. .

For a given digital circuit, we begin with the standard Static Timing Analysis (STA) which generates fresh timing report with timing information of all the paths in the circuit, without considering the NBTI effect. Logic analysis is performed on the circuit to obtain activity factors (α) in case of AC stress and node voltages in case of static stress. Based on the stress conditions, PMOS V_{th} shift and gate delay shifts are computed using delay information from standard cell library under different slew rates and load capacitances. Aged timing report is then obtained by updating gate and path delay shifts in fresh timing report, thus identifying the paths violating timing requirements.

Our proposed framework is general and can be extended to other aging mechanisms such as Positive Bias Temperature Instability (PBTI). The strength of the implementation is that it is integrated into the standard STA flow. The sizing and complexity of the circuit on which the aging analysis can be run using this method directly depends on the STA capability being used. In this work, the implementation of this aging flow is performed using PrimeTime, commercial STA tool from Synopsys. The library used in the entire aging analysis is a 45nm Nangate standard cell library characterized with the Predictive Technology Model (PTM) [58]. The aging aware delay model captures the shift in the rising delays of each gate in the circuit. The non-critical paths in the fresh circuit may turn critical over time due to aging, depending on the size of the paths and types of gates.

The aging timing analysis illustrated above has to be performed at the critical instants in the operation. Aging at these critical moments is prominent and has maximum impact on the circuit performance as mentioned in previous

chapter. NBTI increases the path delay of a logic path which depends on the type, size and number of gates in the path. The increase in path delay results in the decrease of setup slack, may lead to a timing violation and logic failure. When a digital circuit is operated in alternate standby and normal modes, it is under alternate static and dynamic stress phases respectively. The degradation in these phases is estimated using our long term models and the transition from one phase to the other is handled by the boundary condition from TD model. The degradation at the end of a particular standby mode depends on the switching activity of the previous dynamic phase. Figure 4.8 illustrates the path delay shifts in inverter, NOR2 and NAND2 gates with activity factor (α). When α is either close to 0 or 1, the degradation changes rapidly with a slight change in α and the aging sensitivity to α becomes less at intermediate α values [53] [54]. The degradation is maximum in NOR gates (PMOS stack) and at α=1 (static stress). Hence, the end of standby mode (or the onset of next dynamic phase) is the critical moment (shown in Figure 4.2). Aging analysis at these moments facilitate to obtain the upper bound of path delays and the number of timing violations in the sequential circuit.

4.5.2 Timing Violations in Sequential Circuits

NBTI induced shifts in the logic paths and clock buffers of sequential circuits result in timing violations. Figure 4.8 demonstrates a setup violation in an inverter chain example between two sequential elements. The output of the final inverter is passed through second sequential element at the rising edge of the clock. The path delay along the inverter chain needs to be smaller than required

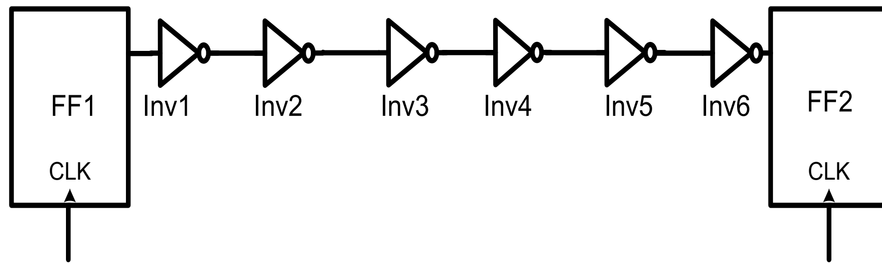


Figure 4.8. Inverter chain between two sequential elements.

data arrival time to avoid any setup violations. Even though the fresh circuit meets the timing requirements, asymmetric aging increases the path delay along the inverter chain due to gate delay increase in alternate inverters, and may result in setup violation [55].

A case of setup violation in the same inverter chain is illustrated in Table 3. Individual gate and accumulated path delays are present in the fresh timing report. The required data arrival time to avoid setup violation is 285ps, when operated at a clock with a period of 340ps. In the fresh circuit, the accumulated path delay is 280.6ps with a positive setup slack, indicating that there is no timing violation in the circuit. The gate and path delays shift when stressed for $2x10^6$s (~1 month: stress time used in our entire analysis at an accelerated temperature of 105^oC). Rising edge delays of INV1, INV3 and INV5 are increased as highlighted in Table 3, resulting in an increase in overall accumulated delay along the chain to 286.7ps, with a negative slack of -1.7ps. Negative setup slack indicates that data reaches later than required arrival time, leading to a setup violation.

NBTI induced delay shifts in the logic path result in setup violation as presented in Table 3. Similarly, NBTI induced delay shifts are present in the clock

Table 3: Demonstration of setup violation in inverter chain.

Gates in path	Fresh		Tr. type	Aged ($2x10^6$s)			
	Gate delay (ps)	**Path delay (ps)**		**RD Model Gate delay (ps)**	**TD Model Gate Delay (ps)**	**RD Model Path delay (ps)**	**TD Model Path delay (ps)**
DFF1	192.6	192.6	Fall	192.6	192.6	192.6	192.6
Inv1	20.7	213.3	Rise	22.8	22	215.4	214.6
Inv2	11.1	224.4	Fall	11.1	11.1	226.5	225.7
Inv3	17.2	241.6	Rise	19.2	18.5	245.7	244.2
Inv4	11	252.6	Fall	11	11	256.7	255.2
Inv5	17.1	269.7	Rise	19.1	18.4	275.8	273.6
Inv6	10.9	280.6	Fall	10.9	10.9	286.7	284.5
Required data arrival time		285	Required data arrival time			285	285
Setup slack		+4.4	Setup slack			**-1.7**	**0.5**

tree as shown in the Figure 4.9 (logic path from s5378, an ISCAS89 circuit). Since the clock tree is symmetrical to avoid skew, clock ticks at the same moment in both sequential elements even under aging (ignoring variations in degradation). When the STA tool is invoked, fresh timing report is generated which consists delays for all the paths in the circuit. STA tool is invoked to generate fresh timing

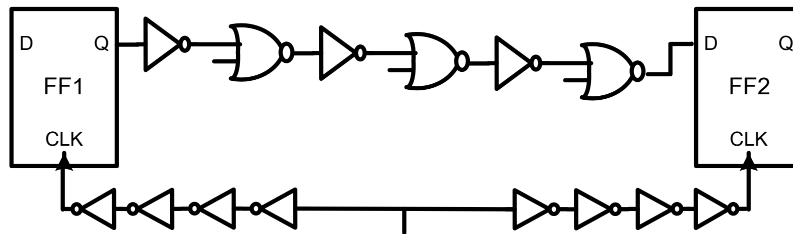


Figure 4.9. Logic path in a ISCAS89 circuit; NBTI does not result in hold violation due to symmetric clock tree at two FFs.

report and 20% paths with maximum accumulated delay are identified for setup analysis. Similarly, paths with minimum delay are identified for hold analysis; PrimeTime also supports logic analysis, which generates intermediate node voltages or α values depending on the type of circuit operation. ΔV_{th} is predicted using the R-D model based on the stress conditions. Gate delay shifts are computed by delay models in section 2 and aged timing report is generated. Figure 4.10a shows the distribution of shift in path delays when the proposed framework is implemented in s5378 circuit with 179 paths. The distribution of

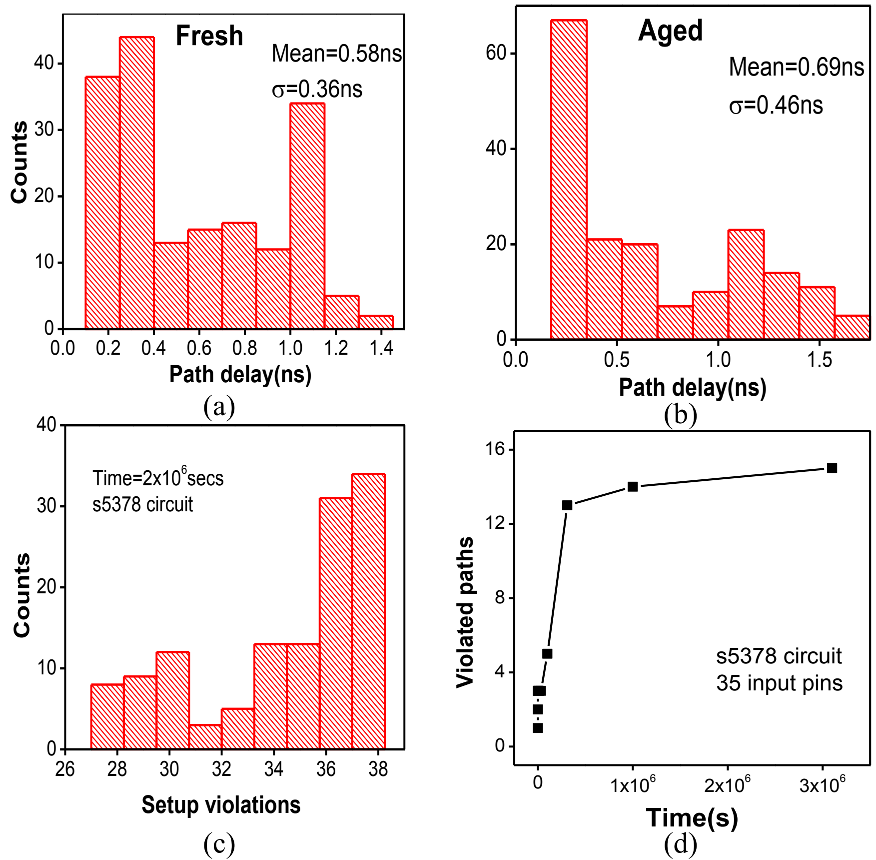


Figure 4.10. Distribution of path delays in (a) fresh (b) aged s5378 circuit; distribution of (c) setup violations under different input patterns (d) timing violations increase with time.

path delays has a mean (μ) of 0.58ns and variance (σ) of 0.36ns as shown in Figure 4.10a. Such a wide distribution is due to the variety of gate types, path structures and large number of gates in the circuit. Under aging, the mean increases by approximately 20% when the circuit is stressed for $2x10^6$s. Since both the minimum and maximum path delays are shifted by same percentage [22], standard deviation increases due to aging and shifts by 25% in s5378 circuit (Figure 4.10b). Though the fresh circuit meets the timing requirements, increase in path delays lead to timing violations.

Further, the internal node voltages depend upon the input voltages when the circuit is in standby mode. The input of any PMOS transistor can be either 0 (gnd) or 1 (V_{dd}) depending on the input voltage pattern. If the input is 0, the PMOS V_{th} shifts, resulting in gate delay increase and the delay is predominant in case of gates such as NOR4. The input pattern dependence is shown in Figure 4.10c, showing the distribution of setup violations due to aging with 2^7 different input patterns. The distribution is wide showing the importance of input voltages on aging in failure diagnosis under NBTI. Figure 4.10d shows the timing violation of paths in s5378 circuit with stress time. Timing violations occur even when the stress time is low and the increase in number of violations gradually decreases for longer stress times. This behavior is similar to the ΔV_{th} shift with stress time.

Table 4. Setup violations in ISCAS89 circuits.

Design	Clock period (ns)	t=1year	t=5years	t=10years
S27	0.48	1	1	1
S382	0.9	1	2	2
S420	0.87	1	1	2
S444	1.05	0	1	1
S510	0.95	1	1	1
S641	2.76	3	4	4
S713	2.91	4	4	5
S820	2.3	1	2	2
S832	2.4	2	2	2

Our methodology is comprehensively demonstrated in different ISCAS89 circuits and is summarized in Table 4. As the number of gates and complexity of the circuit increases, there is an increase in number of timing violations. The framework can be implemented in any large scale circuit using commercially available STA tools.

4.6 Impact on VLSI Design

The proposed asymmetric aging analysis facilitates to observe the impact of NBTI on VLSI design. In this section, impact of NBTI in clock cycle paths of DDR memory controller is assessed using our framework.

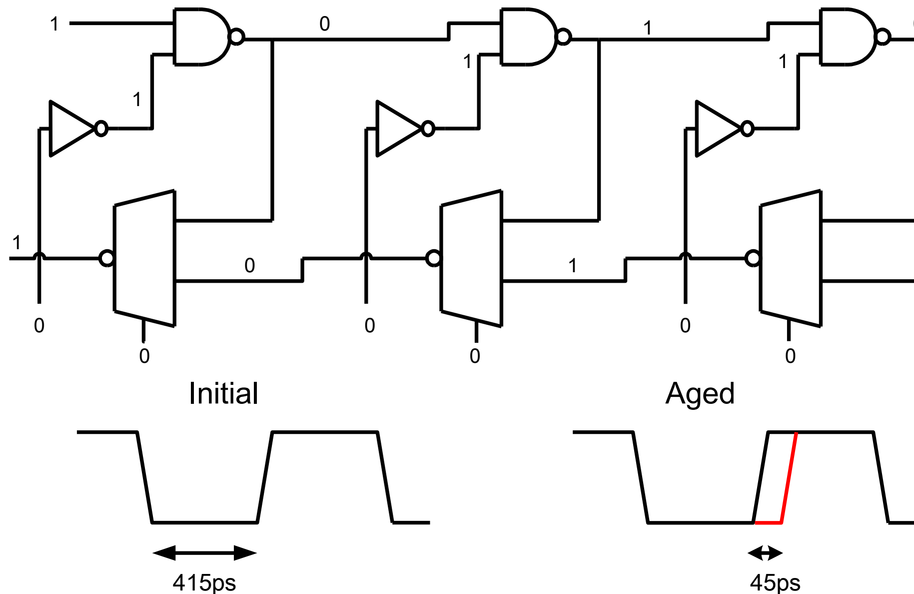


Figure 4.11. The delay shift of the rising edge due to asymmetric aging in half cycle path of DDR memory controller.

In digital circuits operation, the stress is dynamic in nature which leads to moderate degradation compared to static stress. Advanced circuits with clock gating techniques or half cycle paths (e.g. DDR) that transmit data both in rising and falling edges experience a larger degradation since the stress in static in nature. NBTI impact on these circuits can be understood with our methodology for appropriate guard-banding protection. Figure 4.11 shows a half cycle path in Dual Data Rate (DDR) memory that requires both rising and falling edges to trigger data transfer. Output (OUT) of the half cycle path is at 50% duty cycle in the fresh circuit (t=0). Several gates in the circuit are under static DC stress. When

the circuit is under static stress for $2x10^6$s, asymmetric aging leads to increase in propagation delay from 416ps to 460ps of the rising edge. The time for which the circuit spends in low phase decreases by ~10%. Similar results are presented in [45] in half cycle paths validated with silicon evidence. Since the circuit is sensitive to both edges, shift in rising edge has to be compensated by increasing the pulse width to 460ps (Figure 4.11). This results in reduction of clock frequency over time.

In summary of our framework, accurate failure diagnosis method due to asymmetric aging in digital circuits is proposed in this work. A simple analytical model to predict the aging induced delay shifts in a digital gate is presented using device level long term ΔV_{th} models. The aging prediction uses the library cell delays without relying on re-characterization of the standard cell library. The proposed reliability analysis is demonstrated in ISCAS89 circuits (with 45nm cell library) and can be applied to any large scale circuits. The mean and variance of path delays shift by ~20% due to aging in larger benchmark circuits. The clock frequency of half cycle paths in DDR memory controller decreases by 10% over lifetime and NBTI results in accessing incorrect bit cell in SRAM caches, leading to read or write failure. Our aging analysis framework enables resilient design techniques to mitigate NBTI aging in digital circuits.

Chapter 5

STATISTICAL AGING PREDICTION

5.1 Statistical Circuit Aging

Statistical circuit aging prediction is not trivial since NBTI has strong dependence on device physical parameters and operation conditions. The situation is more complex in today's digital circuits as they operate under multiple V_{DD} in Dynamic Voltage Scaling (DVS), leading to supply voltage variations. Hence, it is critical to understand, simulate and mitigate the NBTI effect in early design stages to ensure reliable circuit operation for desired lifetime.

Threshold voltage and gate delay shifts can be determined using TD

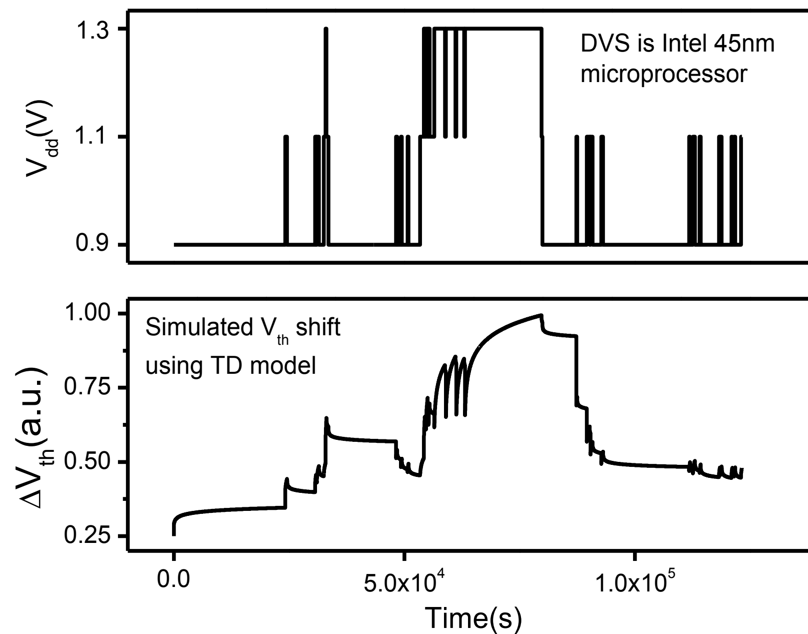


Figure 5.1. DVS operation pattern in a Intel 65nm processor and predicted V_{th} shift

model based on the circuit operation conditions. Today's circuit operation involves multiple V_{DD} under Dynamic Voltage Scaling (DVS). Figure 5.1 presents the operation pattern in a 65nm dual core Intel processor and threshold voltage shift under such an operation. When the supply voltage is changed from a higher V_{DD} to lower V_{DD}, the degradation undergoes recovery. TD model correctly predicts the recovery under DVS, whereas RD model predicts a very low degradation when operated under lower V_{DD}, resulting in aging overestimation. Further, the gate delay and circuit delay shifts may turn certain paths into critical paths, resulting in timing violations. Therefore, it is essential to include accurate NBTI models in the aged timing analysis to guarantee circuit lifetime. TD model based aging analysis realizes realistic failure rate in digital circuits avoiding overly pessimistic prediction from RD model.

This chapter leverages trapping/detrapping based compact models for transistor degradation to achieve accurate reliability prediction under device to device and supply voltage variations.

5.2 Timing Violations under TD and RD Models

A setup violation in the INV chain in Figure 4.8 is illustrated in Table 5. Since NBTI only shifts low to high delay of a signal, only the rising edges are shifted resulting in asymmetric aging. This is considered in our analysis as only delays of Inv1, Inv3 and Inv5 (rising transitions) are shifted, as highlighted in Table 5. Power law time dependence in RD model overestimates aging compared to log(t) dependence in TD model. Path delay for the circuit increases to 286.7ps when RD model is used and TD model predicts path delay as 284.5ps. The required data arrival time is unchanged under the assumption that clock edge is not delayed.

Table 5. Fresh and aged analysis of circuit in Figure 4.8.

Gates in path	**Fresh**		**Tr. type**	**Aged ($2x10^6$s)**			
	Gate delay (ps)	**Path delay (ps)**		**RD Model Gate delay (ps)**	**TD Model Gate Delay (ps)**	**RD Model Path delay (ps)**	**TD Model Path delay (ps)**
DFF1	192.6	192.6	Fall	192.6	192.6	192.6	192.6
Inv1	20.7	213.3	Rise	22.8	22	215.4	214.6
Inv2	11.1	224.4	Fall	11.1	11.1	226.5	225.7
Inv3	17.2	241.6	Rise	19.2	18.5	245.7	244.2
Inv4	11	252.6	Fall	11	11	256.7	255.2
Inv5	17.1	269.7	Rise	19.1	18.4	275.8	273.6
Inv6	10.9	280.6	Fall	10.9	10.9	286.7	284.5
Required data arrival time		285	Required data arrival time			285	285
Setup slack		+4.4	Setup slack			**-1.7**	**0.5**

When RD model is used, the accumulated delay along the logic path is higher than required data arrival time, causing a setup violation with negative slack of -1.7ps. TD model predicts a positive slack of 0.5ps and the circuit does not encounter any timing violation.

The proposed trapping/detrapping based BTI models and asymmetric aging analysis can be implemented in digital circuits using commercial STA tools and PrimeTime tool is used to perform aged timing analysis in ISCAS89 benchmark circuits. A 45nm Nangate open cell library characterized by Predictive Technology Model (PTM) is used in our analysis. STA tool is invoked to generate fresh timing report and 20% paths with maximum accumulated delay are identified for setup analysis. PrimeTime also supports logic analysis, which generates intermediate node voltages or α values depending on the type of circuit operation. ΔV_{th} is predicted using both TD and RD models based on the stress conditions. Gate delay shifts are computed by aging aware library and aged timing report is generated. Figure 5.2a shows the distribution in shift in path delays when the proposed framework is implemented in s5378 circuit with 179 paths. Trapping/detrapping based aging model exhibits a narrow distribution with μ of 19ps and σ of 18.9 ps. RD model predicts a wider distribution of delay shifts and over-estimates aging by 50% as illustrated in Figure 5.2a. The observed behavior is due to fast changing degradation with time in RD model compared to gradual change in TD model due to logarithmic time dependence.

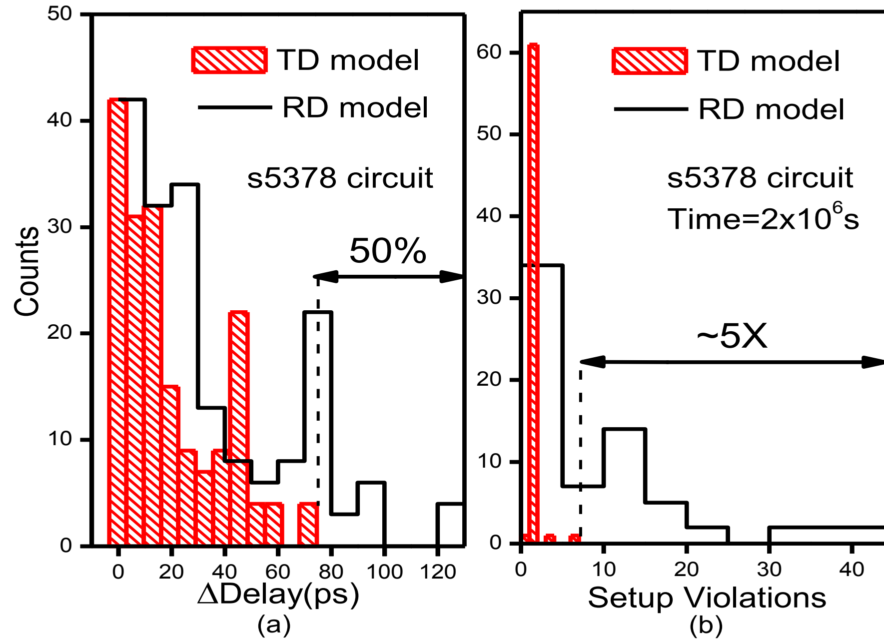


Figure 5.2. Distribution of (a) path delays and (b) violations under variations in TD and RD model parameters in s5378 circuit.

As mentioned in the previous chapter, the number of timing violations in a given circuit depends on the number of gates and structure of paths. The variations present at the device level are also significant to determine the failure rate due to aging. The predicted ΔV_{th} variations using the extracted model parameters are used to calculate the variations in gate delay shifts. Aged timing under variations is obtained and number of violations can be predicted under both models. Figure 5.2b presents the distribution of the number of setup violations in s5378 circuit due to model parameter variations in both TD and RD models. When the circuit is under static operation for $2x10^6$s, RD model has a wider distribution and predicts approximately 5X more number of violations compared to TD model. The main variation parameter in RD model is the time exponent (n) and a small change in n value leads to huge difference in aging prediction. Hence, the exponential

Table 6: Setup Comparisons from both models.

Design	Clock period (ns)	t=1year		t=5years		t=10years	
		RD	TD	RD	TD	RD	TD
S27	0.48	1	1	1	1	1	1
S382	0.9	1	1	2	2	2	2
S386	0.87	2	2	3	2	3	2
S444	1.05	1	1	1	1	1	1
S510	0.95	1	1	2	1	2	1
S641	2.76	3	3	4	4	4	4
S820	2.3	4	1	4	1	4	1
S832	2.4	4	3	4	3	4	3

sensitivity of predicted ΔV_{th} to time exponent leads to a wide failure distribution. On the other hand, the main variation parameter in the TD model is ϕ and ΔV_{th} has a linear dependence on it. Hence, the predicted degradation has a narrow distribution under variations due to less sensitivity of aging model to ϕ. Trapping/detrapping based NBTI model correctly predicts the critical paths under aging, avoiding the need to protect large number of paths during the design stage.

Comprehensive demonstration of our aging analysis using both TD and RD models is demonstrated in different ISCAS89 benchmark circuits and summarized in Table 6. RD model overestimates the failure due to aging in majority of the circuits. Therefore, trapping/detrapping based BTI models correctly predict the degradation and failure rate in circuits. Our aging prediction guides designers to correctly predict the guard band under device to device variations.

5.3 Supply Voltage Variations

Along with variations at the device level, supply voltage variations also impact aging due to strong dependence of NBTI to V_{DD}. Our long-term trapping model predicts the V_{th} shift as a function of V_{DD} and by injecting ΔV_{th} to PMOS devices, shift in the circuit delay can be estimated. When a circuit is operated under multiple V_{DD}, the degradation can be computed by using the boundary condition.

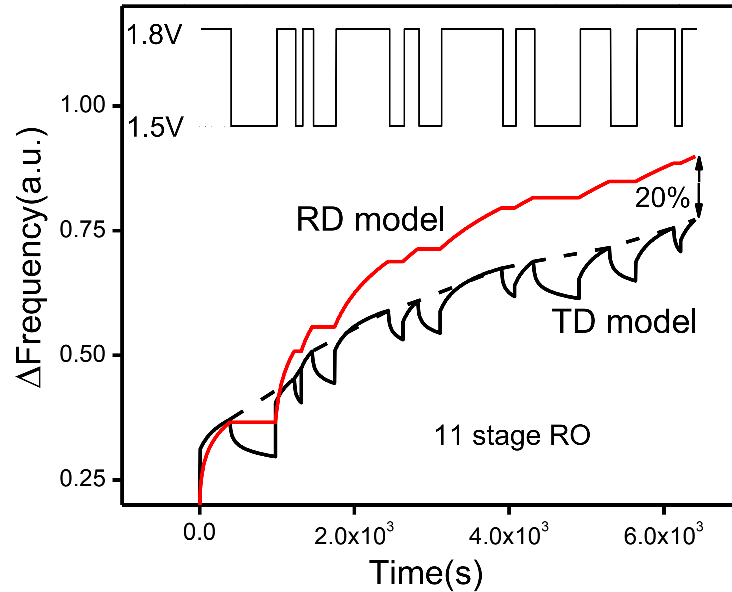


Figure 5.3. Prediction of delay shift in an 11 stage RO under sequence of V_{DD}=1.8V and 1.5V; RD model overestimates circuit aging.

Figure 5.3 presents the absolute shift in frequency of an 11 stage RO when operated under two supply voltages of 1.8V and 1.5V for random operation times. RD model does not predict any recovery when operated under a lower V_{DD} and hence, predicts monotonic RO frequency shift. However, TD model correctly predicts the recovery under a lower V_{DD} and estimates the upper bound of

degradation accurately. The situation is more complex when the circuit is operated under more than 2 supply voltages, which can be handled by TD aging model using the appropriate boundary condition. Therefore, our proposed aging prediction under trapping/detrapping theory facilitates robust long term device and circuit reliability prediction under V_{DD} variations inherent in DVS operation.

In conclusion, this chapter presents statistical timing analysis that is performed under device level variations and demonstrated in ISCAS89 benchmark circuits; RD model overestimates aging and TD model correctly predicts aging under device to device and supply voltage variations. With solid verification with measured 65nm device data, the proposed statistical aging helps designers to accurately monitor and manage circuit lifetime.

Chapter 6

SUMMARY AND FUTURE WORK

6.1 Thesis Conclusions

This work is a study of trapping/detrapping mechanism causing NBTI effect and efficient aging prediction at the circuit level. The specific contributions of this work include:

- Chapter 1 gives an overview of circuit aging and the importance of aging with scaling of the technology.
- Chapter 2 presents the degradation models of NBTI using Reaction-Diffusion model and its short comings. Several key features of NBTI effect especially fast recovery and variation in time exponent are not explained by RD model and shows the necessity of new trapping.de-trapping physics based aging model.
- Chapter 3 introduces the trap properties in T-D mechanism responsible for NBTI effect and proposes a new statistical aging model. The proposed model is scalable with technology and is validated with 65nm measured aging data. Further, dynamic V_{th} shift in DVS operation is predicted from cycle-to-cycle models and long-term model is developed to estimate the upper bound of multiple cycle aging.
- New framework is proposed in chapter 4 that leverages compact aging models and standard STA tools to predict the delay shift accumulated in the logic paths. Our simulation methodology takes asymmetric

aging into account and efficiently predicts setup and time violations due to delay shift in logic paths and clock buffers respectively

- Chapter 5 shows the prediction mismatch when RD and TD models are applied to our framework. RD model significantly overestimates aging under device to device and supply voltage variations. Further, RD model does not predict under lower V_{DD} operation leading to additional over prediction. Trapping/de-trapping aging model accurately predicts lifetime both at the device and circuit level.

6.2 Future Work

Several research directions for further investigation arise from this thesis. (1) Due to the introduction of high-k dielectrics from 45nm technology, the experimental data show that the power-law exponent is 0.07-0.10. Since the high-k gate stack actually consists of two layers, namely a high-k dielectric layer and an interfacial SiO2 layer, the fast stress and recovery component most likely associates with the defects in the SiO2 interfacial layer induced by the overlaying high-k film. Thus, stress-induced changes in the threshold voltage, i.e., $_V_{th}$, should be modeled by both the hole trapping/detrapping mechanism in the case of positive bias temperature instability (PBTI), which arises in this case. Compact models for both NBTI and PBTI are necessary in order to reduce guard-banding considering trapping/detrapping, would be more accurate to characterize transistor degradation with time. (2) Statistical behavior at the device level is well studied in this thesis. Further, aging variability impact on performance of digital circuits needs to be analyzed. (3) The entire aging analysis is made on analog circuits in this thesis. What's the impact of NBTI effect on analog circuits? Can this simulation framework also be used for analog circuits performance analyzing? These questions need to be answered in future work.

REFERENCES

[1] S. V. Kumar, C. H. Kim and S. S. Sapatnekar, “Adaptive techniques for overcoming performance degradation due to aging in digital circuits”, *Asia and South Pacific Design Automation Conference*, pp. 284-289, 2009.

[2] M. A. Alam, “A critical examination of mechanics of dynamic NBTI for PMOSFETs”, *International Electron Devices Meeting*, pp. 14.4.1-14.4.4, 2003.

[3] W. Wang, V. Reddy, A. T. Krishnan, R. Vattikonda, S. Krishnan and Y. Cao, “Compact modeling and simulation of circuit relaibility for 65nm CMOS technology”, *IEEE Transactions on Device and Materials Reliability*, vol. 7, no. 4, pp. 509-517, 2007.

[4] M. Denais, et al., “On-the fly charecterization of NBTI in ultra-thin gate oxide PMOSFET’s”, *International Electron Devices Meeting*, pp. 109-112, 2004.

[5] X. Zhou, K. Y. Lim and D. Lim, “A simple and unambiguous definition of threshold voltage and its implications in deep-submicron MOS device modeling”, *IEEE Transactions on Electron Devices*, pp. 807-809, 1999.

[6] V. Reddy, “Impact of negative bias temperature instability in digital circuit reliability”, *International Reliability and Physics Symposium*, pp. 248-253, 2002.

[7] R. Zheng, et al., “Circuit aging prediction for low-power operation”, *Customs Integrated Circuits Conference*, pp. 427-430, 2009.

[8] D. K. Schroder and J. A. Babcock, “Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing,” *Journal of Applied Physics*, vol. 94, no. 1, pp. 1-18, 2003.

[9] K. Kunhyuk, M. A. Alam, K. Roy, “Estimation of NBTI degradation using IDDQ measurement”, *International Reliability and Physics Symposium*, pp. 10-16, 2007.

[10] M. A. Alam, S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectronics Reliability*, vol. 45, pp. 71--81, 2005.

[11] K. L. Chen, S. A. Saller, I. A. Groves, and D. B. Scott, “Reliability effects on MOS transistors due to hot carrier injection”, *IEEE Transactions on Electron Devices*, vol. ED-32, no. 2, pp. 386-393, 1985.

[12] C. Hu, S. C. Tam, F. C. Hsu, E K. Ko, T. Y. Chan, and K. W. Terrill, "Hot-electron induced MOSFET degradation model-model, monitor, and improvement," *IEEE Transactions on Electron Devices*, vol. ED-32, no. 2, pp.375-385, 1985.

[13] S. Bharadwaj, W. Wang, R. Vattikonda, Y. Cao and S. Vrudhula, “Predictive modeling of the NBTI effect for reliable design,” *Customs Integrated Circuits Conference*, pp. 189-192, 2006.

[14] Y. Cao, W. Zhao, “Predictive technology model for nano-CMOS design exploration”, International *Conference on Nano-Networks*, 2006.

[15] N. Kimizukaand, T. Yamamoto, T. Mogamiand, K. Yamaguchi, K. Imaiand, and T. Horiuchi, “The impact of bias temperature instability for direct-tunnelingultra-thin gate oxide on mosfet scaling”, *VLSI Technology, 1999. Digest of Technical Papers. 1999 Symposium* on, pages 73–74, 1999.

[16] A. T. Krishnan, C. Chancellor, S. Chakravarthi, P. E. Nicollian, V. Reddy, and A. Varghese, “Material dependence of hydrogen diffusion: Implication for nbti degradation”, *IEEE International Electron Devices Meeting*, Dec. 2005.

[17] C. Constantinescu, “Trends and challenges in vlsi circuit reliability”, *IEEE Computer Society*, pages 14–19, 2003.

[18] http://www.itrs.net, *International technology roadmap for semiconductor (itrs)*.

[19] G. Chen, K. Y. Chuah, M. F. Li, Daniel SH Chan, C. H. Ang, J. Z. Zheng, Y. Jin, and D. L. Kwong, “Dynamic NBTI of PMOS transistors and its impact on device lifetime”, *International Reliability Physics Symposium*, pages 196–202, 2003.

[20] B. C. Paul, K. Kang, H. Kufluoglu, M. A. Alam, and K. Roy, “Impact of NBTI on the temporal performance degradation of digital circuits”, *IEEE Electronic Device Letters*, 26(8):560–562, Aug. 2005.

[21] J. KO and S. CM, “Negative bias stress of mos devices at high electric fields and degradation of mos devices”, *Journal of Applied Physics*, 48(5):2004–2014, May 1977.

[22] G. Bersuker, J. H. Sim, C. S. Park, C. D. Young, S. Nadkarni, R. Choi, and B. H. Lee, “Intrinsic threshold voltage instability of the hfo2 nmos transistors”, *International Reliability Physics Symposium*, pages 179–183, March 2006.

[23] R. Vattikonda, W. Wang, and Y. Cao, “Modeling and minimization of pmos nbti effect for robust nanometer design”, *IEEE/ACM Design Automation Conference*, pages 1047–1052, Jul. 2006.

[24] M. Agarwal, B. C. Paul, Ming Zhang, and S. Mitra, “Circuit failure prediction and its application to transistor aging”, *VLSI Test Symposium*, pages 277–286, May 2007.

[25] Y. Wang, H. Luo, K. He, R. Luo, H. Yang, and Y. Xie, “Temperature-aware nbti modeling and the impact of input vector control on performance degradation”, *Design, Automation and Test in Europe Conference and Exhibition*, pages 1–6, April 2007.

[26] S. Chakravarthi, A. T. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, “A comprehensive framework for predictive modeling of negative bias temperature instability”, *IEEE International Reliability Physics Symposium*, pages 273–282, 2004.

[27] K. Bernstein et al., "High-Performance CMOS Variability in the 65-nm Regime and Beyond", *IBM J. RES. & DEV.* Vol 50, 433--449, 2006.

[28] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, And S. Krishnan, “Impact of negative bias temperature instability on digital circuit reliability”, *IEEE International Reliability Physics Symposium*, pages 248–254, Apr. 2002.

[29] S. Mahapatra, D. Saha, D. Varghese, and P. B. Kumar, “On the generation and recovery of interface traps in mosfets subjected to nbti, fn, and hci stress”, *IEEE Tran. on Electron Devices*, 53(7):1583–1592, Jul. 2006.

[30] S. Ogawa and N. Shiono, “Generalized diffusion-reaction model for the low-

field charge build up instability at the Si-SiO2 interface", *Physical Review*, 51(7):4218–4230, Feb. 1995.

[31] M. A. Alam, "On the reliability of micro-electronic devices: An introductory lecture on negative bias temperature instability", *Nanotechnology 501 Lecture Series*, Sep. 2005.

[32] F. C. Hsu and S. Tam, "Relationship between mosfet degradation and hot-electron induced interface-state generation", *IEEE Electron Device Letters*, 5(2):50–52, Feb. 1984.

[33] S. Mahapatra, P. B. Kumar, and M. A. Alam, "Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-mosfets", *IEEE Tran. on Electron Devices*, 51(9):1371–1379, Sep. 2004.

[34] W. Zhao, F. Liu, K. Agarwal, D. Acharyya, S. Nassif, K. Nowka, and Y. Cao, "Rigorous extraction of process variations for 65nm cmos design", *ESSDERC/ESSCIRC*, 2007.

[35] J. Keane, W. Zhang and C. H. Kim, "An on-chip monitor for statistically significant circuit aging characterization," IEDM, pp. 4.2.1-4.2.4, 2010.

[36] T. Sato, et al., "A device array for efficient bias temperature instability measurements," Proc. ESSDERC., pp. 143-146, 2011.

[37] T. Grasser, et al., "Switching oxide traps as the missing link between negative bias temperature instability and random telegraph noise," Int. Electron Devices Meeting Tech. Dig., pp. 729-732, 2009.

[38] V. Huard, C. Parthasarathy, C. Guerin, T. Valentin, E. Pion, M, Mammasse, N. Planes, L. Camus, "NBTI degradation: From transistor to SRAM errors", IRPS, pp. 289-300, 2008.

[39] G. I. Wirth, R. da Silva, B. Kaczer, "Statistical model for MOSFET bias temperature instability component due to charge trapping," IEEE Tran. Electron Devices, vol.58, no.8, pp.2743-2751, Aug. 2011.

[40] S. Pae, J. Maiz, C. Prasad, "Effect of NBTI degradation on transistor variability in advanced technologies," Integrated Reliability Workshop, vol., no., pp.18-21, 15-18 Oct. 2007.

[41] "Reliability Simulation in Integrated Circuit Design," available at: http://www.cadence.com/.

[42] M. Agarwal, B. C. Paul, M. Zhang, and S. Mitra, "Circuit Failure Prediction and its Application to Transistor Aging", IEEE VLSI Test Symp., pp. 277-286, 2007.

[43] R. Vattikonda, W. Wang, and Y. Cao, "Modeling and minimization of pmos nbti effect for robust nanometer design", Design Automation Conference, pp. 1047-1052, Jul. 2006.

[44] W. Wang, Z. Wei, S. Yang and Y. Cao, "An Efficient Method to Identify Critical Gates under Circuit Aging", Int. Conference on Computer Aided Design, pp. 735-740, Nov. 2007.

[45] A. T. Krishnan, F. Cano, C. Chancellor, V. Reddy, Q. Zhangfen, P. Jain, J. Carulli, J. Masin, S. Zuhoski, S. Krishnan, and J. Ondrusek, "Product Drift from NBTI: Guardbanding, Circuit and Statistical Effects", Int. Electron Devices Meeting, pp. 4.3.1-4.3.4, 2010.

[46] B. C. Paul, Kunhyuk Kang, H. Kufluoglu, M. A. Alam, K. Roy, "Impact of NBTI on the temporal performance degradation of digital circuits," IEEE Electron Device Letters, vol.26, no.8, pp. 560- 562, Aug. 2005.

[47] W. Wang, S. Yang, S. Bhardwaj, R. Vattikonda, S. Vrudhula, F. Liu, Y. Cao, "The Impact of NBTI Effect on Combinational Circuit: Modeling, Simulation, and Analysis," IEEE Transactions on VLSI Systems, vol. 18, no. 2, pp. 173-183, 2010.

[48] H. Sangwoo, K. Juho, "NBTI-aware statistical timing analysis framework," IEEE International SOC Conference, pp.158-163, 2010.

[49] T.-B. Chan, J. Sartori, P. Gupta, and R. Kumar, “On the Efficacy of NBTI Mitigation Techniques,” Design, Automation, and Test in Europe, Mar. 2011.

[50] T. Grasser, W. Gos, V. Sverdlov, and B. Kaczer, “The Universality of NBTI Relaxation and its Implications for Modeling and Characterization”, Int. Reliability Physics Symposium, pp. 268–280, 2007.

[51] Y. Tsai et al., “Prediction of NBTI Degradation for Circuit under AC Operation", Int. Reliability Physics Symposium, pp. 665 –669, May 2010.

[52] R. Vattikonda, Y. Luo, A. Gyure, X. Qi, S. Lo, M. Shahram, Y. Cao, K. Singhal, and D. Toffolon, “A New Simulation Method for NBTI Analysis in SPICE Environment,” Int. Symposium on Quality Electronic Design, pp. 41–46, 2007.

[53] T. Grasser, B. Kaczer, P. Hehenberger, W. Gos, R. O. Connor, H.Reisinger, W. Gustin, and C. Schlunder, “Simultaneous Extraction of Recoverable and Permanent Components Contributing to Bias-Temperature Instability,” Proc. Int. Electron Devices Meeting, pp.801–804, 2007.

[54] V. Huard, C. Parthasarathy, N. Rallet, C. Guerin, M. Mammase, D. Barge, and C. Ouvrard, “New Characterization and Modeling Approach for NBTI Degradation from Transistor to Product Level,” in Proc. Int. Electron Devices Meeting, 2007, pp. 797–800.

[55] W. Wang, S. Yang, S. Bhardwaj, R. Vattikonda, S. Vrudhula, F. Liu, and Y. Cao, “The Impact of NBTI on the Performance of Combinational and Sequential circuits,” Design Automation Conference, pp. 364–369, Jun. 2007.

[56] K. Kang, S. P. Park, K. Roy, and M. A. Alam, “Estimation of Statistical Variation in Temporal NBTI Degradation and its Impact on Lifetime Circuit Performance,” in Proc. IEEE/ACM Int. Conf. Computer Aided Design, pp. 730–734, 2007.

[57] T. Sakurai and A. R. Newton, "Alpha-Power Law Mosfet Model and its Application to CMOS Inverter Delay and Other Formulas," IEEE J. Solid-State Circuits, vol. 25, no. 2, pp. 584-594, Apr. 1990.

[58] W. Zhao and Y. Cao, "New Generation of Predictive Technology Model for Sub-45 nm Early Design Explorations," IEEE Trans. Electron Devices, vol. 53, no. 11, pp. 2816-2823, Nov. 2006.

[59] G. I. Wirth, J. Koh, R. da Silva, R. Thewes, and Ralf Brederlow, "Modeling of statistical low-frequency noise of deep-submicron MOSFETs," *Trans. on Electron Dev.*, vol. 52, pp. 1576-1588, 2005.

[60] G. Wirth, R. da Silva, P. Srinivasan, J. Krick and R. Brederlow. "Statistical model for MOSFET low-frequency noise under cyclo-stationary conditions," *IEDM*, p.30.5.1-4, 2009.

[61] R. da Silva and G. Wirth. "Logarithmic behavior of the degradation dynamics of metal oxide semiconductor devices," *J. Statistical Mechanics*, v. P04025, p. 01-12, 2010.

[62] T. Grasser, et al., "The paradigm shift in understanding the bias temperature instability: From reaction-diffusion to switching oxide traps," *TED*, vol. 58, no. 11, pp. 3652-3666, Nov. 2011.

[63] J. Velamala, K. Sutaria, T. Sato and Y. Cao, "Aging statistics based on trapping/detrapping: Silicon evidence, modeling and long-term prediction," *IRPS*, 2F.2.1-2F.2.5, 2012.

[64] W. Wang, V. Reddy, B. Yang, V. Balakrishnan, S. Krishnan and Y. Cao, "Statistical prediction of circuit aging under process variations," *CICC*, pp. 13-16, 2008.

[65] A. Agarwal, B. C. Paul, H. Mahmoodi, A. Datta, and K. Roy, "A processtolerant cache architecture for improved yield in nanoscale technologies," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 1, pp. 27–38, Jan. 2005.

[66] K. Kuhn, "Reducing variation in advanced logic technologies: Approaches

to process and design for manufacturability of nanoscale CMOS," in *IEDM Tech. Dig.*, 2007, pp. 471–474.

[67] K. Mistry *et al.*, "A 45 nm logic technology with high-k + metal gate transistors, strained silicon, 9 Cu interconnect layers, 193 nm dry patterning, and 100% Pb-free packaging," in *IEDM Tech. Dig.*, 2007, pp. 247–250.

[68] J. Maiz, "Reliability challenges: Preventing them from becoming limiters to technology scaling," in *Proc. Int. Integr. Rel. Workshop*, 2006.

APPENDIX A

SCRIPT OF AGED TIMING ANALYSIS IN SYRA

```
sub Dcal ($inst,$Vgs1,$time,$T,$nfile)

{

##%%%%%%%%%%%%  Constant parameters

$Eah2 = 0.49;

$E0 = 0.0825;

$q = 1.6*(1e-19);

$eox = 3.9*(8.85*1e-21);

$T0 = 1e-8;

$k = 0.0259/300;

$kT = $k*($T+273);

$B = exp(-$Eah2/$kT);

$C = $B/$T0;

$n=$nfile;

###%%%%%%%%%%%%  Technology parameters   (65nm)

$K = 7.5;

$Tox = 1.2;

$Vth = 0.4;

$Eox = ($Vgs1-$Vth)/$Tox;

$Cox = $eox/$Tox;

#print " $Kv1 = (($q*$Tox)/$eox)**3*($K**2*$Cox*($Vgs1-
$Vth)*(exp($Eox/$E0))**2*sqrt($C)) \n";

$Kv1 = (($q*$Tox)/$eox)**3*($K**2*$Cox*($Vgs1-
$Vth)*(exp($Eox/$E0))**2*sqrt($C));
```

```
##### TD model parameters ####
$A=0.01;
$B=0.005;
$C1=1e-2;
$A2=5.5;
$K=2.6;
$B2=0.55e-1;
#$K1=0.06e2;
#$delvth = ($Kv1*(($time)**$n));
$delvth=$n*exp(-1*$B2*($A2-
$K*$Vgs1/$Tox)/($kT))*($A+$B*log(1+$C1*$time));
#print "$Kv1\n";
return ($delvth);
}
#Sub-routine to round the integers to 6 decimal places.
sub round {
my @tmp = split(/\./,$_[0]);
my @tmp2 = split('',$tmp[1]);
splice(@tmp2,6);
$tmp[1] = join('',@tmp2);
$tmp = join('.',@tmp);
return $tmp;
}
```

```
sub update_log
{
$logfile = $switches{log};
$oper = $switches{oper};
$desi = $D;
if ( $oper =~ /dynamic/ )
{
`rm -rf timing_${desi}_dynamic.log`;
`touch timing_${desi}_dynamic.log` ;
$newfile = "timing_${desi}_dynamic.log" ;
$dfile = "delay_${desi}.txt";
update_sub_log ( $logfile  , $newfile , $dfile  );
}elsif ( $oper =~ /static/ )
{
open (PF , "patterns_file") || die ("ERROR\n");
foreach my $p ( <PF> )
{
chomp($p);
$P = $p;
$P =~ s/output_${D}_//g;
$P =~ s/\.list//g;
print "#############   $P #############\n";
`rm -rf timing_${desi}_static_${P}.log`;
```

```
`touch timing_${desi}_static_${P}.log`;
$newfile = "timing_${desi}_static_${P}.log" ;
$dfile = "delay_${desi}_${P}.txt";
update_sub_log ( $logfile  , $newfile , $dfile  );
}
}
close(PF);
}
sub update_sub_log
{
open(FH_lis, "$logfile") || die ("cannot open specified file\n");
open(FH_out, ">>$newfile") || die("cannot open output file\n");
open (FH_delay, "$dfile") || die ("cannot open delay file\n");
$i=0;
my @type;
my @instance;
my @delay;
@DELAY= <FH_delay>;
foreach my $d ( @DELAY )
{
chomp($d);
@linesplit1=split(/\t/,$d);
@type[$i]=@linesplit1[0];
```

```
@instance[$i]=@linesplit1[2];

@delay[$i]=@linesplit1[3];

#print " @instance[$i] @type[$i] @delay[$i]\n";

$i=$i+1;

}

#print " @instance[$i] @type[$i] @delay[$i]\n";

#$line1=<FH_delay>;chop($line1);

#while ($line1 ne ""){

#@linesplit1=split(/\t/, $line1);

#@type[$i]=@linesplit1[0];

#@instance[$i]=@linesplit1[2];

#@delay[$i]=@linesplit1[3];

#print " @instance[$i] @type[$i] @delay[$i]\n";

#$i=$i+1;

#$line1=<FH_delay>;

#chop($line1)

#}

$gate_count=$i;

#print "Number of gates are $gate_count\n";



my @linesplit2;

my $linesplit2;
```

```
my @linesplit3;
my @linesplit3;
my $linesplit4;
my $linesplit4;
my $newdelay;
$count=0;
$line2=<FH_lis>;
$delay_acc=0;
#chop($line2);
while ($line2 ne "")
{
@temp1 = split(/ +/ , $line2 );
if($line2=~'/ZN' || $line2=~'/Q')
{
@linesplit2=split('/', $line2);
$current_instance=@linesplit2[0];
#print "$current_instance  ";
for($j=1;$j<=$gate_count;$j++)
{
if (  $temp1[$#temp1] =~ /r/ )
{
if($current_instance=~/@instance[$j]/)
{
```

```
$count=$j;
$newdelay=@delay[$j];
}
}else { $newdelay = 0 ; }


}
$delay_acc=$delay_acc+$newdelay;
@linesplit3=split(' ', $line2);
if(@linesplit3[2] eq "<-")
{
$current_delay=$linesplit3[3];
$current_delay=$current_delay+$newdelay;


$current_acc=$linesplit3[4];
$current_acc=$current_acc+$delay_acc;
}elsif(@linesplit3[2] ne "<-")
{
$current_delay=$linesplit3[2];
#print "$current_delay \t $newdelay \n";
$current_delay=$current_delay+$newdelay;
#print " CHangesd delay is $current_delay\n";
$current_acc=$linesplit3[3];
$current_acc=$current_acc+$delay_acc;
```

```
}
#print "$current_instance ";
@linesplit4=split(' ',$line2);
if(@linesplit4[2] eq "<-")
{
$linesplit4[3]=$current_delay;
$linesplit4[4]=$current_acc;
# print  "$linesplit4[3] $newdelay $delay_acc $linesplit4[4]\n";
print     FH_out     "         $linesplit4[0]     $linesplit4[1]     $linesplit4[2]
$linesplit4[3]   $linesplit4[4] $linesplit4[5]\n";
}
if(@linesplit4[2] ne "<-"){
$linesplit4[2]=$current_delay;
$linesplit4[3]=$current_acc;
#print "$linesplit4[2] $newdelay $delay_acc $linesplit4[3]\n";
print FH_out "  $linesplit4[0] $linesplit4[1]                       $linesplit4[2]
$linesplit4[3] $linesplit4[4]\n";
#print   "   $linesplit4[0]  $linesplit4[1]                              $linesplit4[2]
$linesplit4[3] $linesplit4[4]\n";
}

#print "$current_delay ";
#print "$newdelay\n";
```

```
# $line= join(' ',@linesplit3);

}

elsif($line2=~'/D' || $line2=~'/A'){

@linesplit5=split(' ',$line2);

if(@linesplit5[2] eq "<-"){

$linesplit5[4]=$current_acc;

#print  "$linesplit5[3] $newdelay $delay_acc $linesplit5[4]\n";

print     FH_out     "          $linesplit5[0]     $linesplit5[1]     $linesplit5[2]
$linesplit5[3]   $linesplit5[4] $linesplit5i[5]\n";

}

if(@linesplit4[2] ne "<-"){

$linesplit5[3]=$current_acc;

#print "$linesplit5[2] $newdelay $delay_acc $linesplit5[3]\n";

print FH_out "  $linesplit5[0] $linesplit5[1]                        $linesplit5[2]
$linesplit5[3] $linesplit5[4]\n";

}

}

elsif($line2=~/data arrival time/){
```

```
@tt = split (/ /, $line2 );
if ( $tt[$#tt] > 0 ) {
@linesplit6=split(' ',$line2);
$linesplit6[3]=$current_acc;
print FH_out "  $linesplit6[0] $linesplit6[1] $linesplit6[2]
$linesplit6[3]\n"
}else {
@linesplit7=split(' ',$line2);
$linesplit7[3]=-$current_acc;
print FH_out "  $linesplit7[0] $linesplit7[1] $linesplit7[2]
$linesplit7[3]\n"
}
}
elsif($line2=~'slack'){


@linesplit8=split(' ',$line2);
$slack=@linesplit8[2];
#print "slack is $slack\n";
$linesplit8[2]=$linesplit8[2]-$delay_acc;
print FH_out "  $linesplit8[0] $linesplit8[1]
$linesplit8[2]\n"
}
elsif($line2=~'Startpoint:'){
```

```
print FH_out $line2;

$delay_acc=0;

}

elsif($line2!~'/ZN' || $line2!~'/Q'){

print FH_out $line2;

}

$line2=<FH_lis>;

}

}
```