RELIABILITY AND DATA ANALYSIS OF WEAROUT MECHANSIMS FOR CIRCUITS

A Dissertation
Presented to
The Academic Faculty

by

Shu-han Hsu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August 2020

COPYRIGHT © 2020 BY SHU-HAN HSU

RELIABILITY AND DATA ANALYSIS OF WEAROUT MECHANSIMS FOR CIRCUITS

Approved by:

Dr. Linda Milor, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Mark Davenport
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. David Keezer
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Benjamin Klein
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Ben Wang
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: [July 16, 2020]

ACKNOWLEDGEMENTS

I would like to thank Professor Linda Milor for her insightful guidance and mentoring throughout my PhD journey. I am very grateful for the opportunity to be her student, because she has shown me how to become a better researcher, as well as opened my mind to allow me to explore exciting research topics. Professor Milor is my biggest inspiration for pursing an academic career, which would not have been possible without her help, and I hope that I can be able to mentor others as she has mentored me.

I also would like to thank Professor Mark Davenport, Professor Ben Klein, Professor David Keezer, and Professor Ben Wang for their helpful suggestions and serving as committee members.

I would also like to thank Professor Azad Naeemi for his help and support.

In addition, I would like to thank Taizhi Liu, Dae-Hyun Kim, Kexin Yang, and Rui Zhang in our lab for their support, collaboration, comments, and feedback.

I would like to specially thank Dr. Yi-Da Wu for his mentoring, discussion, and help. You are an amazing researcher, and I am grateful to be able to learn from you. I also want to thank Dr. Li-Hsiang Lin for his help and discussions, which allowed me to further understand problems. Yen Pang Lai and Muya Chang, thank you for your support. Mayank Parasar, thank you for all your help with discussions. Thanks also to all my friends for being on this journey.

I would like to thank my family for their support. This thesis is dedicated to them.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS **iii**

LIST OF TABLES **vi**

LIST OF FIGURES **vii**

LIST OF SYMBOLS AND ABBREVIATIONS **xii**

SUMMARY **xiii**

CHAPTER 1. Introduction **1**
1.1 Planar MOSFET and FinFET Device Structures **3**
1.2 Statistical Analysis of Reliability Data **7**
1.2.1 Weibull Distribution 8
1.2.2 Maximum Likelihood Estimation 9
1.2.3 Quasi-Newton Method 10
1.3 Wearout Mechanisms **10**
1.3.1 Front-end Wearout Mechanism 11
1.3.2 Back-end Wearout Mechanisms 13
1.3.3 Middle-of-line Wearout Mechanism 15
1.4 Circuit Case Studies **17**
1.4.1 Process Design Kit (PDK) 17
1.4.2 Ring Oscillators 17
1.4.3 Static random-access memory (SRAM) 18
1.5 Research Objectives **19**
1.6 Thesis Overview **20**

CHAPTER 2. Failure Analysis of Circuits **21**
2.1 Separation of competing wearout mechanisms in circuits **21**
2.1.1 Methodology 23
2.1.2 Investigation of Initial Conditions vs. Sample Size 27
2.1.3 Identification of Wearout Mechanism for Each Individual Sample 30
2.2 Failure Analysis for On-line Testing **41**
2.2.1 Methodology 43
2.2.2 Error Analysis 49

CHAPTER 3. Accelerated Testing **58**
3.1 Failure Probability vs. Test Time **59**
3.2 Sample Size vs. Test Time **60**
3.3 Effects of the Number of Stages **64**
3.4 Lifetime Estimation Method for Use Conditions **66**
3.5 Testing Conditions **67**
3.5.1 Computing the Error for the Accelerated Test Conditions 67
3.5.2 Test Plans 68

3.6 Error Reduction Through Sampling **70**
3.6.1 Adjusting Sample Fractions for a Fixed Overall Sample Size 70
3.6.2 Minimum Sample Size to Achieve Fixed Tolerance 74
3.6.3 Lowest Tolerance and Total Sample Size 77
3.7 Summary **78**

CHAPTER 4. Conclusions **79**
4.1 Summary **79**
4.2 Future Work **80**

REFERENCES **81**

LIST OF TABLES

Table 1 95% Confidence Intervals for 11-stage ring oscillators at accelerated conditions (1.37V, 194.72^{o}C) 40

LIST OF FIGURES

Figure 1 – Illustration of the material degradation reaction. The material will move from the initial state to the degraded state in order to lower its Gibbs Potential Free Energy [12]. 2

Figure 2 – Materials degradation in a device can cause device parameter S to degrade with time, which can be increasing or decreasing [12]. 2

Figure 3 – Minimum feature size scaling trend for Intel logic technologies [15, 19]. 3

Figure 4 – Structure of (a) Planar MOSFET and (b) FinFET device [23]. 4

Figure 5 – Comparison of planar versus FinFET transistor electrical characteristics. (a) Channel current versus gate voltage. (b) Transistor gate delay versus operating voltage [15]. 5

Figure 6 – Normalized scaling trends in the per-bit alpha and neutron SER of SRAMs as a function of technology node [22]. 6

Figure 7 – Intel's development of the technology node for the past five generations [15]. 6

Figure 8 – Difference between the purposes of reliability and quality. 7

Figure 9 – Prediction uses [33]. 8

Figure 10 – Weibull distributions with different shape parameters, β, and same scale parameter, η. 9

Figure 11 – FEOL TDDB breakdown in the gate region of a FinFET transistor. 12

Figure 12 – Generation of traps leading to gate oxide breakdown [42]. 12

Figure 13 – BEOL TDDB breakdown occurs between the spacing of two interconnect wires [47]. 14

Figure 14 – Formation of hillock and void resulting from electromigration. 15

Figure 15 – Electromigration breakdown in an interconnect wire [52]. 15

Figure 16 – MTDDB breakdown occurs in the region between the contact and the gate. 16

Figure 17 – Ring oscillator composed of identical invertors and physical interconnections connected together. 18

Figure 18 – An SRAM is composed of six transistors, where A_1 and A_2 are the access transistors, and L_1, L_2, L_3 and L_4 are the latch transistors. 19

Figure 19 – Illustration of a test structure [12]. 22

Figure 20 – Implementation of the MLE algorithm. 22

Figure 21 – Distributions of data for 11-stage ring oscillators for FEOL TDDB with β_1=1.12, η_1=9.87 (yrs) and MOL TDDB with β_2=1.9, η_2=15.36 (yrs) (x axis: unit years), varying sample size. P is probability. The pink markers correspond to the primary wearout mechanism, while the black markers correspond to the secondary wearout mechanism. 28

Figure 22 – Parameter errors as a function of sample size for extracted competing mechanisms, FEOL TDDB with β_1=1.12, η_1=9.87 (yrs) and MOL TDDB with β_2=1.9, η2=15.36 (yrs). 29

Figure 23 – Sorting accuracy of 93.0% for FEOL TDDB (1st dist.) vs. EM failure (2nd dist.) in 14nm FinFET ring oscillators for a sample size of 100. 34

Figure 24 – FEOL TDDB selectivity in 14nm FinFET ring oscillators. 36

Figure 25 – MOL TDDB selectivity in 14nm FinFET ring oscillators. 37

Figure 26 – Test times for 11-stage ring oscillators at accelerated conditions (1.37V, 194.72OC) for a sample size of 100. 37

Figure 27 – Failure time for 100th sample out of 100 samples at FEOL TDDB=0.5 selectivity (a) as a function of voltage and temperature (b) close-up. 38

Figure 28 – Failure time for 1000th sample out of 1000 samples at FEOL TDDB=0.5 selectivity (a) as a function of voltage and temperature (b) close-up. 39

Figure 29 – Sorting accuracy of 80% for FEOL TDDB (1st dist.) vs. MOL failure (2nd dist.) in 14nm FinFET ring oscillators for a sample size of 100. 40

Figure 30 – SRAM failure samples due to FEOL TDDB for (a) a sample size of 94,000 with extracted FEOL TDDB failure parameters, η= 19.966 and β=1.119, and (b) a sample size of 100 out of 94,000 46

SRAM cells with extracted FEOL TDDB failure parameters of η=39.66 and β=0.99.

Figure 31 – Mapping between process-level Weibull parameters (η and β) and SRAM cell Weibull parameters for (a) η and (b) β for FEOL TDDB. 47

Figure 32 – Inverse mapping between process-level Weibull parameters (η and β) and SRAM cell Weibull parameters for (a) η and (b) β for FEOL TDDB. 48

Figure 33 – Standard deviation error of the extraction of ln(η) as a function of time for the (a) SRAM with η=20 yrs, β=1.12 and (b) a single device. 50

Figure 34 – Percent changes in errors in device characteristic lifetime estimation from variations in (a) temperature and (b) voltage. Voltage error differences above 5% causes the SRAM to fail upon startup and below -5% causes the SRAM to have essentially infinite characteristic lifetimes (e.g. above 300 years). 52

Figure 35 – Sensitivity of the extraction of ln ($\boldsymbol{\eta_{device}}$) to changes in (a) temperature and (b) voltage. 54

Figure 36 – Percent changes in errors for characteristic lifetime from variations in (a) channel length and (b) duty cycle (error calibrated to a duty cycle of 50%). 56

Figure 37 – Sensitivity resulting from changes in (a) channel length and (b) duty cycle. 57

Figure 38 – Example of acceleration of failure distribution [12]. 58

Figure 39 – Contour plot of the 63% failure probability (characteristic lifetime) as a function of temperature and voltage acceleration, for accelerated testing of the 1001-stage ring oscillator, with testing times ranging from two hours to six months. 60

Figure 40 – Sample size needed produce a single failure with 95% confidence as a function of voltage and temperature for a 1001-stage ring oscillator for a testing time of (a) two hours (b) two weeks (c) two months and (d) six months. 62

Figure 41 – Contour plot for sample size of four to produce at least 1 failure with 95% confidence as a function of voltage and temperature for a 1001-stage ring oscillator with testing time ranging from two hours to six months. 62

Figure 42 – Sample size needed produce a single failure with 95% confidence as a function of voltage and temperature for a 11-stage ring oscillator for a testing time of (a) two hours (b) two weeks (c) two months and (d) six months. 63

Figure 43 – Contour plot for sample size of four to produce at least 1 failure with 95% confidence as a function of voltage and temperature for a 11-stage ring oscillator with testing time ranging from two hours to six months. 63

Figure 44 – Comparison of the contour plots for the number of stages and 63% failure probability (characteristic lifetime) as a function of voltage and temperature for (a) 2 hours (b) 2 weeks (c) 2 months and (d) 6 months. 64

Figure 45 – Comparison of the number of stages and the test conditions needed to detect at least one failure for various sample sizes as a function of voltage and temperature for test times of (a) two hours (b) two weeks (c) two months and (d) six months. 65

Figure 46 – Standard deviations of estimates of the errors in $\ln(\eta)$ using Monte Carlo simulations. 68

Figure 47 – Optimal testing points for 1001-stage, 101-stage and 11-stage ring oscillators for testing times of two hours, two weeks, two months and six months. 69

Figure 48 – For a sample size of 2000, a comparison of 11-stage, 101-stage and 1001-stage ring oscillators for a testing time of two hours, two weeks, two months and six months in terms of (a) the lowest tolerance, (b) the corresponding percentage of the sample at the lower voltage point to achieve the minimum tolerance, (c) the corresponding lower voltage point to achieve the minimum tolerance, and (d) the corresponding higher voltage point to achieve the minimum tolerance. 72

Figure 49 – For a sample size of 2000, a comparison of 11-stage, 101-stage and 1001-stage ring oscillators for a testing time of two hours, two weeks, two months and six months in terms of (a) the lowest tolerance, (b) the corresponding percentage of the sample at the lower voltage point to achieve the minimum tolerance, (c) the corresponding lower voltage point to achieve the minimum tolerance, and (d) the corresponding higher voltage point to achieve the minimum tolerance. 73

Figure 50 – Relationship between percent sample size at the lower voltage and the tolerance for the 11-stage, 101-stage and 1001-stage ring 74

oscillators with a testing time of two weeks. (Tolerances above 30% are not shown for clarity.)

Figure 51 – Comparison between the total sample size and the testing time for 1001-stage, 101-stage and 11-stage ring oscillators to achieve a 10%, 20% and 30% tolerance. 75

Figure 52 – Relationship between percent tolerance and number of stages for a testing time of two weeks. 76

Figure 53 – Relationship between lowest tolerance and total sample size at a two week testing time for 11-stage ring oscillator, 101-stage ring oscillator, and 1001-stage ring oscillator. 77

LIST OF SYMBOLS AND ABBREVIATIONS

EM	Electromigration
FEOL TDDB	Front-end-of-line time-dependent dielectric breakdown (or front-end gate oxide breakdown)
FinFET	Fin field-effect transistor
GOBD	Front-end-of-line time-dependent dielectric breakdown (or front-end gate oxide breakdown)
GTDDB	Front-end-of-line time-dependent dielectric breakdown (or front-end gate oxide breakdown)
MLE	Maximum Likelihood Estimation
MOL TDDB	Middle-of-line time-dependent dielectric breakdown
MOSFET	Metal oxide semiconductor field-effect transistor
MTDDB	Middle-of-line time-dependent dielectric breakdown
Pdk	Process design kit

SUMMARY

The objective of this research is to develop methodologies for the failure analysis of circuits, as well as investigate the factors for accelerating testing for front-end-of-line time-dependent dielectric breakdown (FEOL TDDB). As the technology node enters into a new era where the planar MOSFET has now transitioned to FinFET structures due to device scaling, new reliability concerns arise. Therefore, it is critical to be able to understand and predict the failure for integrated circuits, especially as new emerging technologies, such as autonomous vehicles and wearable devices, are increasing the bar for reliability standards.

In this thesis, the separation of wearout mechanisms for circuits will be investigated, and the identification of failure modes for the failure samples will be analyzed. SRAMs and ring oscillators based on the 14nm FinFET GlobalFoundries/Samsung/IBM PDK will be used to study the failure modes. The systematic and random errors for online monitoring of SRAMS will also be examined.

Furthermore, the testing plans for acceleration testing will also be explored for ring oscillators. The effects of stage number and testing time will be discussed. Error reduction through sampling will also be used to find the best testing conditions for accelerated testing.

This work provides a way for engineers to better understand aging monitoring of circuits, and to design better testing to collect failure data. With these developments, engineers may make improved failure predictions for growing complex systems. In

addition, the circuit design and manufacturing processes can be enhanced with better yield and product performance.

INTRODUCTION

Semiconductor devices and circuits are the core components of electronic devices. However, for these electronic devices to perform practically, reliability goals must be fulfilled [1], which has becoming even more challenging due to the increasing complexity in semiconductor manufacturing [2]. Furthermore, emerging technologies, such as autonomous vehicles and wearable sensors for health monitoring, are becoming increasingly interrelated with public safety, making the need for the assessment of highly reliable complex systems increasingly important [3-6]. Particularly for autonomous vehicles, the reliability standards are raised higher than traditional vehicles [7].

Reliability is the ability of a device to conform to its electrical, visual, and mechanical specifications over a specified period of time under specified conditions [8]. The development of technology, processes, and standards are made to ensure the reliability of semiconductor devices during application [9]. Reliability engineering is built upon a vast set of disciples, such as physics, statistics, and materials, etc., to ensure the continuous improvement of every device.

Reliability is often confused with quality, but these two have different meanings. Quality refers to the device meeting its specifications, but reliability refers to the time dependence of the device degradation [10]. Degradation is a result of the Second Law of Thermodynamics, where the entropy of isolated systems will increase over time to move to a stable state in order to lower its Gibbs Potential Free Energy, as shown in Figure 1 [11]. The degradation will affect the device parameter, as shown in Figure 2 [12]. The device parameter may increase or decrease as a result over time.

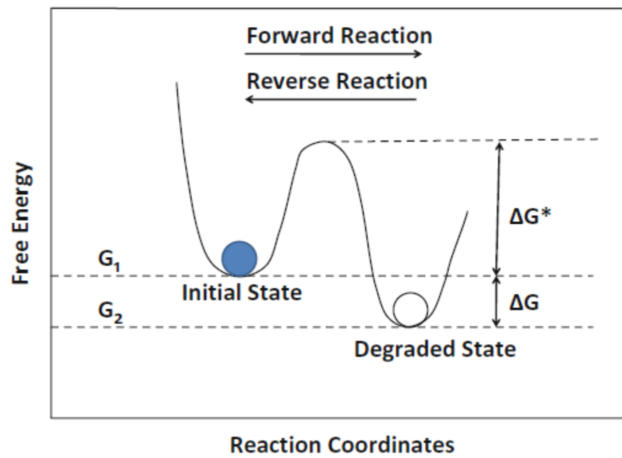


Figure 1 – Illustration of the material degradation reaction. The material will move from the initial state to the degraded state in order to lower its Gibbs Potential Free Energy [12].

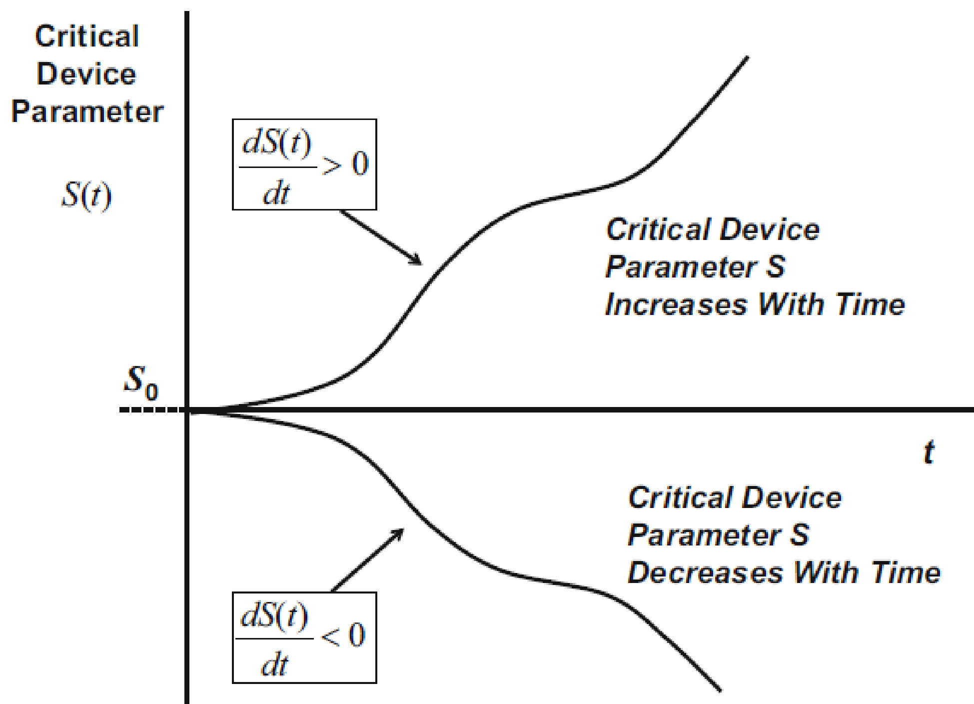


Figure 2 – Materials degradation in a device can cause device parameter S to degrade with time, which can be increasing or decreasing [12].

Here, this thesis is focused on device failure under stress, and how it impacts circuit reliability. Stress is any external agent that can cause degradation to occur in the material properties to the point where the device can no longer function properly in its intended application . All material will eventually degrade over time leading to device failure, so it is critical to pinpoint when the device cannot operate properly when designing products.

1.1 Planar MOSFET and FinFET Device Structures

In recent years, the shrinkage of device dimensions has allowed the density of integrated circuits on a chip to increase, lower costs and increase performance [13, 14], as shown in Figure 3 [15]. With each new technology node generation developed from the result of device scaling, reliability concerns arise [16], which may be due to the device structure or fabrication changes[17]. As a result, one of the most important changes from the decrease in the gate oxide length is the device structure transition from planar metal oxide semiconductor field-effect transistor (MOSFET) device to fin field-effect transistor (FinFET) structure [18].

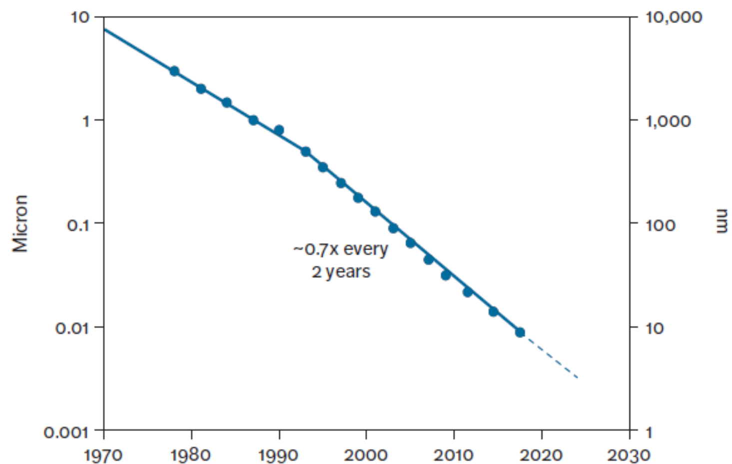
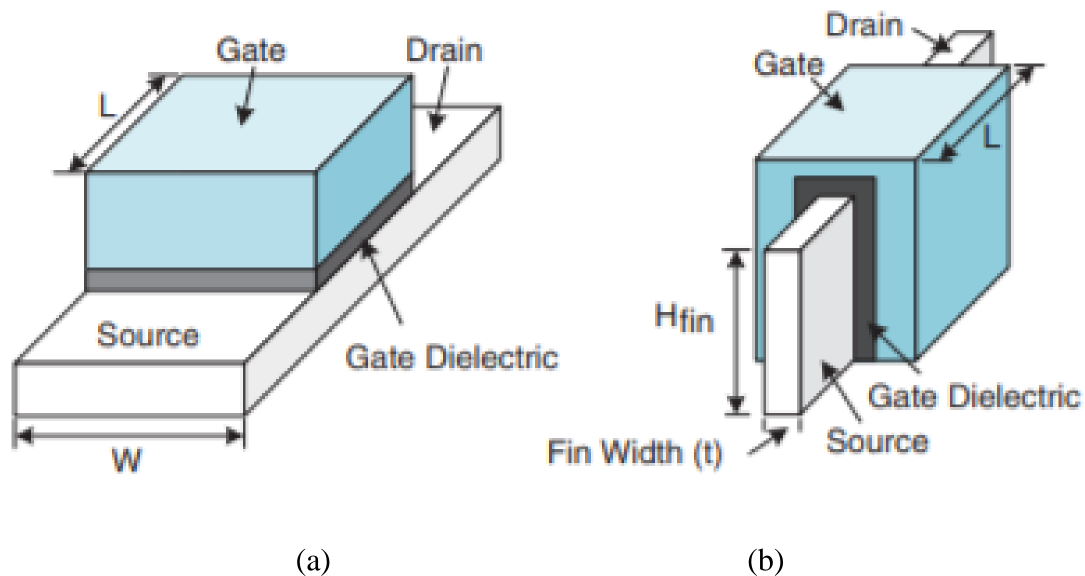


Figure 3 – Minimum feature size scaling trend for Intel logic technologies [15, 19].

The MOSFET device structure is the traditional device structure used in the semiconductor industry. It is composed of a gate that controls the current flowing from the source to drain, as shown in Figure 4 (a). However, with the scaling of process technology, the gate length has decreased dramatically, making it difficult for the gate to control the current [20].

The FinFET structure was developed to address the increased leakage current and short channel effects of planar MOSFETs resulting from the shrinking size of the devices [21]. It has also lowered the soft error rates in static random-access memory [22]. The FinFET structure has the channel elevated, so that the gate can surround it on all three sides, looking like a fin, as shown in Figure 4 (b).



(a) (b)

Figure 4 – Structure of (a) Planar MOSFET and (b) FinFET device [23].

The improvement in the electrical characteristics from switching to the FinFET structure can be seen in Figure 5. However, the different structure and fabrication process for FinFETs introduces new reliability issues [24-26]. For example, the 3D structure of FinFETs gives rise to the concern of middle-of-line time-dependent dielectric breakdown, which was not previously critical in the planar MOSFET structures[27, 28]. FinFETs took around a quarter of a century to transition from the first demonstration in research to a commercialized product [29]. The first commercially-wide available FinFET chips were produced from Intel starting from their 22nm node [30], which was introduced just recently as shown in Figure 7.

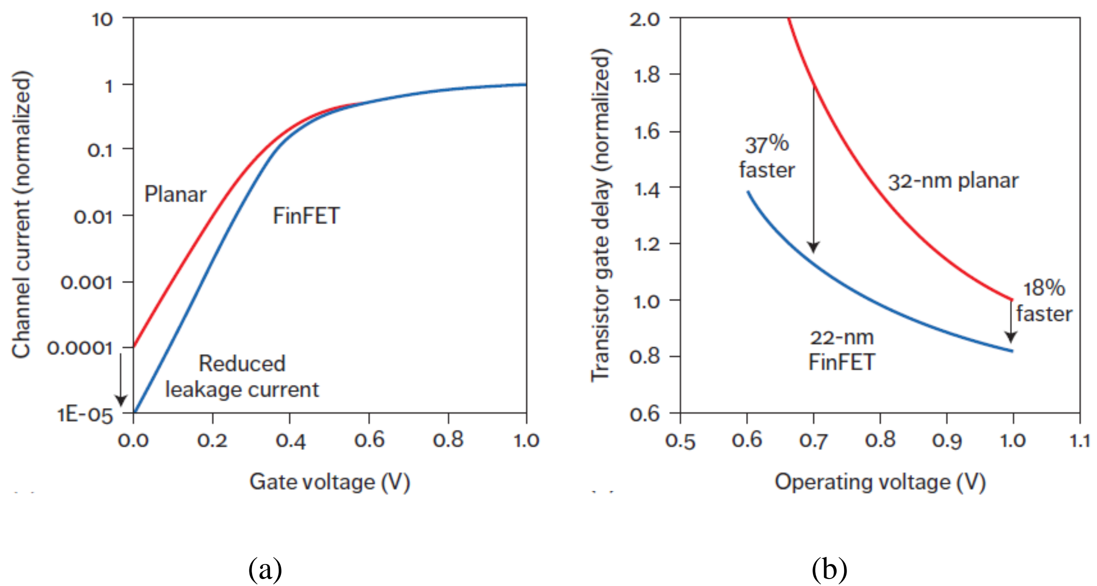


Figure 5 – Comparison of planar versus FinFET transistor electrical characteristics. (a) Channel current versus gate voltage. (b) Transistor gate delay versus operating voltage [15].

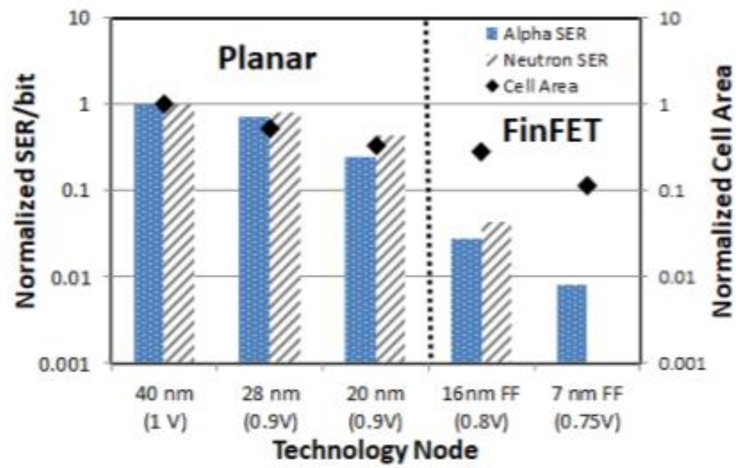


Figure 6 – Normalized scaling trends in the per-bit alpha and neutron SER of SRAMs as a function of technology node [22].

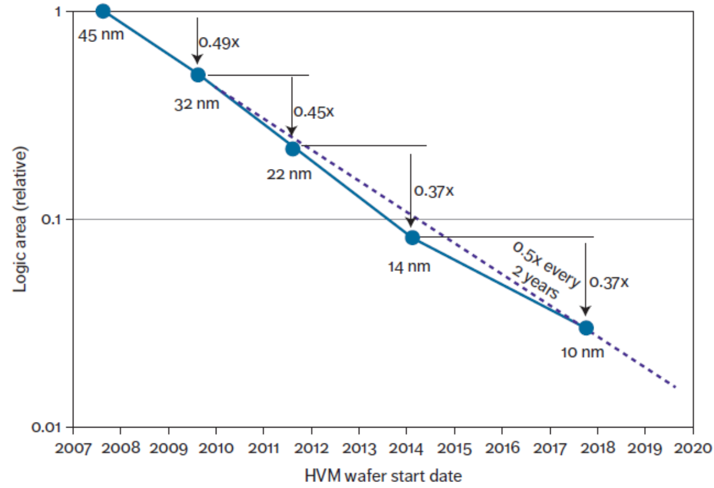


Figure 7 – Intel's development of the technology node for the past five generations [15].

1.2 Statistical Analysis of Reliability Data

To quantify reliability, reliability can also be thought of as the probability that a semiconductor device having initial satisfactory performance can continue to perform its intended function for a given time under actual usage conditions [31]. Therefore, reliability calculations are often based on statistical data collected using failure records using failure distributions. Mathematical analysis is used to predict how long devices will function. As illustrated in Figure 8, reliability aims to move the average of the failure distribution to a higher failure time, which is different than quality, which moves to reduce the variability. It is important to be able to predict the degradation of the product when it is in the field and predict the remaining life of the product over time [32]. Usages for reliability prediction are described in Figure 9 [33].

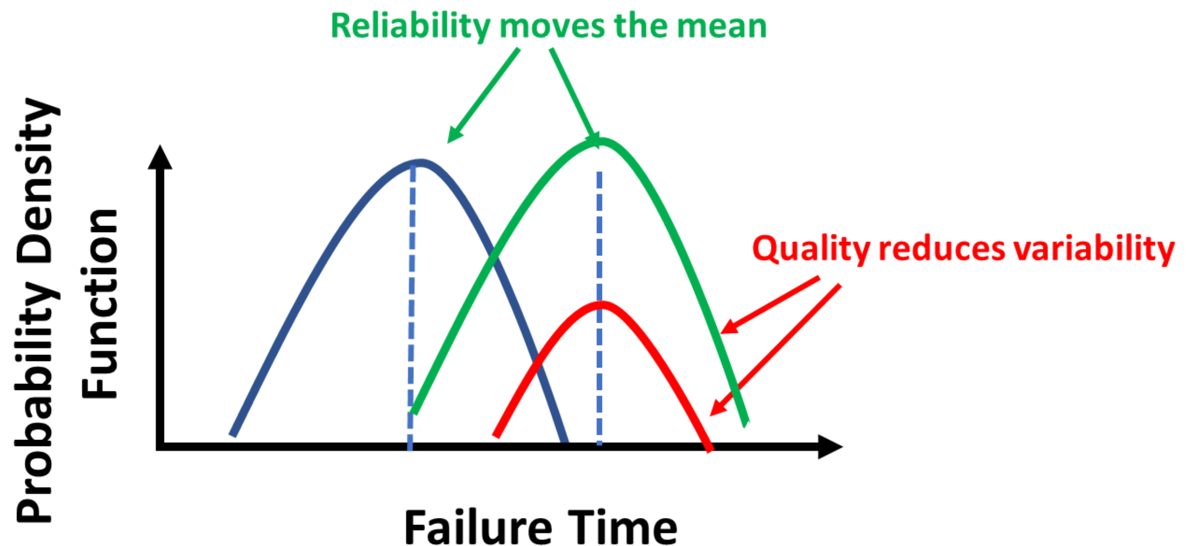


Figure 8 – Difference between the purposes of reliability and quality.

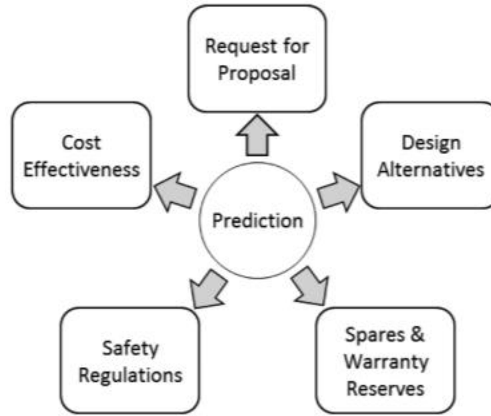


Figure 9 – Prediction uses [33].

1.2.1 Weibull Distribution

The Weibull distribution is a very flexible distribution often used to analyse reliability data, especially for system failure [12]. The distribution is named after Waloddi Weibull, who originally used the probability distribution in 1951 as a model for material breaking strength, but has now widespread use in describing lifetime distributions [34]. The probability density function for a two parameter Weibull distribution can be calculated as follows:

$$f(t) = \left(\frac{\beta}{\eta}\right)\left(\frac{t}{\eta}\right)^{\beta-1} exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \tag{1}$$

where β is the shape parameter, also known as the dispersion parameter, and η is the characteristic lifetime, also known as the scale parameter, which is the lifetime at the 63% failure probability. The exponential distribution is a special case of the Weibull distribution where $\beta = 1$. Examples of the Weibull distribution are shown in Figure 10.

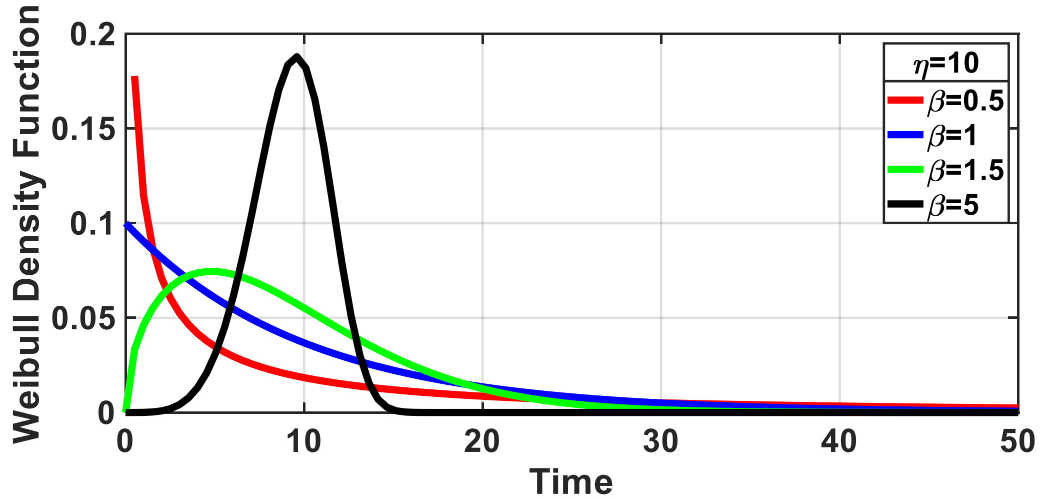


Figure 10 – Weibull distributions with different shape parameters, β, and same scale parameter, η.

The time-to-failure is when a device parameter degrades to the point that the device cannot function properly, and the characteristic lifetime can be used as an indicator of the time-to-failure. The time-to-failure of circuits can be investigated using compact modeling [35], where the failure modes can be described using the Weibull distribution. The failure modes will be described in more detail in Section 1.3.

1.2.2 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is a statistical method that determines the parameters of a model from given observations, which would be failure times in the reliability applications discussed in this proposal, by finding the parameter values that maximizes the likelihood or highest probability of getting the observations (time-to-failure data in this proposal) given the parameters. The reasoning is that the estimate which explains the data best will be the best estimator. MLE is a powerful analysis tool that can be applied to both censored (the condition when the value of a measurement or observation

is only partially known) and uncensored failure data. MLE is used in this thesis to find the competing Weibull parameters for competing wearout mechanisms in circuits.

1.2.3 Quasi-Newton Method

The quasi-Newton method is a calculation method that can find the minima or maxima of functions. It is based on Newton's method, which uses both the first and second derivative (Hessian matrix) values to find the roots of a function. Newton's method can be thought of as similar to a gradient descent method, which is a first-order method, but with the addition of using second-order information to change the step size and direction. This addition of the Hessian information helps avoid descent directions that plateau too quickly. Therefore, the quasi-Newton method is somewhere of an intermediate between Newton's method and gradient descent. The quasi-Newton method uses fewer steps to find the optimal value, taking more time to execute each step, while the gradient descent method, which is a first-order method, has the opposite properties [36].

Newton's method is generally computationally expensive and slow, because it is more difficult to calculate the second derivative. The quasi-Newton method overcomes this problem by approximating the Hessian matrix instead of computing it directly. Various algorithms are available that can be used to find the Hessian matrix.

1.3 Wearout Mechanisms

Reliability concerns in transistor devices must be addressed in order to ensure that a product can perform its required functions for a stated period of time, which can generally be divided into front-end and back-end wearout mechanisms [37]. One of the most

important front-end wearout mechanisms typically found in both planar MOSFET and FinFET structures is the front-end-of-line time-dependent dielectric breakdown (FEOL TDDB), also known as the front-end gate oxide breakdown (GOBD). For back-end wearout mechanisms, both planar MOSFET and FinFET devices have back-end-of-line time-dependent dielectric breakdown (BEOL TDDB or BTDDB) and electromigration (EM). However, FinFET devices have an extra wearout mechanism that is not found in planar MOSFET devices, which is the middle-of-line time-dependent dielectric breakdown (MOL TDDB or MTDDB), that occurs due to process scaling [38].

1.3.1 Front-end Wearout Mechanism

Front-end-of-line time-dependent dielectric breakdown (FEOL TDDB), also known as front-end gate oxide breakdown (GOBD or GTDDB), is the main front-end wearout mechanism found in transistors, as shown in Figure 11. When transistors are turned on, the gate dielectric region is subjected to voltage and thermal stress, and traps can build-up that degrade the gate oxide material [39]. As shown in Figure 11, the traps will eventually form a conduction path, leading to the breakdown of the oxide material. For the 45-65nm, technology nodes, the gate oxide can be as thin as 1.2 nm [40], where the same defect size will have a higher impact compared to the older technology node generations with thicker oxides, because the defect size is now impacting a larger portion of the oxide thickness for the thin oxides [41].

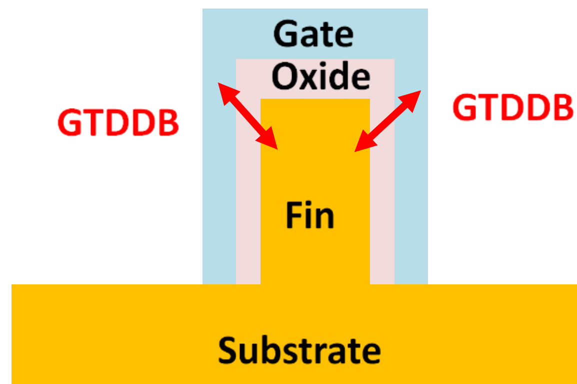


Figure 11 – FEOL TDDB breakdown in the gate region of a FinFET transistor.

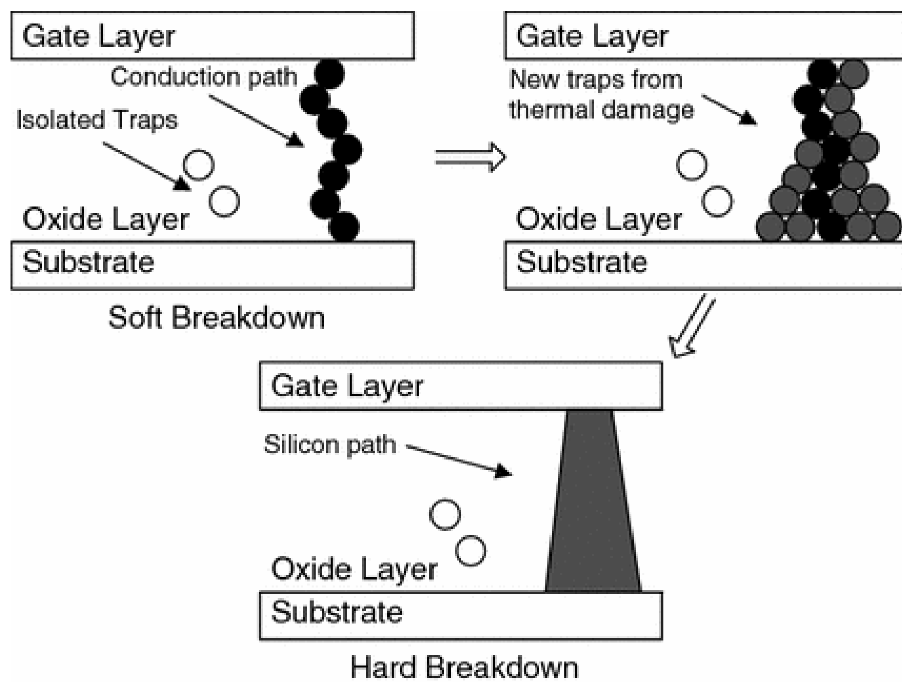


Figure 12 – Generation of traps leading to gate oxide breakdown [42].

FEOL TDDB can be modeled as [43]:

$$\eta = A_{ox}\left(\frac{1}{WL}\right)^{\frac{1}{\beta}} e^{\frac{-1}{\beta}} V^{a+bT} exp\left(\frac{c}{T}+\frac{d}{T^2}\right) s^{-1} \tag{2}$$

where β is the shape parameter, and η is the time-to-failure at the 63% probability point. W and L are the device width and length, s is the probability of stress, T is temperature, V is gate voltage, respectively. In addition, a, b, c, d, and A_{ox} are fitting parameters, which depend on the type of technology process used.

1.3.2 Back-end Wearout Mechanisms

1.3.2.1 Back-end-of-line time-dependent dielectric breakdown (BEOL TDDB or BTDDB)

Back-end-of-line time-dependent dielectric breakdown (BEOL TDDB or BTDDB) is the dielectric breakdown between adjacent metal interconnect lines[44], as shown in Figure 13. BEOL TDDB can be modelled as [45]:

$$\eta = A_{BEOL} L_i^{\frac{-1}{\beta}} exp\left(-\gamma E^m + \frac{E_a}{kT}\right) s^{-1} \tag{3}$$

where A_{BEOL} is a constant that is dependent on the technology process, L_i is the vulnerable length (the distance where the metals run parallel to each other), E_a is the activation energy (~0.5 eV), T is temperature, γ is the field acceleration factor, s is the probability of stress, and k is the Boltzmann constant. E is the electric field, which is a function of voltage, V and the distance between the conductions, S_i, i.e., E=V/S_i, and m is ½ for the $\sqrt{E}$ model [46].

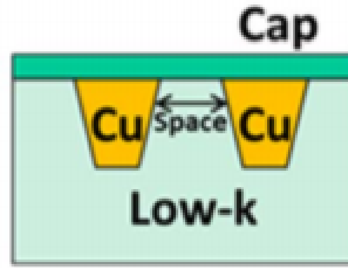


Figure 13 – BEOL TDDB breakdown occurs between the spacing of two interconnect wires [47].

1.3.2.2 Electromigration (EM)

Electromigration (EM) is the dislocation of atoms in the lattice of interconnect metals due to the momentum transfer of electrons [48]. The movement of atoms can cause voids due to the absence of atoms at one end and hillocks due to the build-up of atoms at the other end [49], as shown in Figure 14. This phenomena results in the function failure of circuits due to the loss of connections [50], as can be seen in the image in Figure 15. It can be modelled as [51]:

$$\eta = A_{EM} J^{-n} exp\left(\frac{E_a}{kT}\right) \tag{4}$$

where A_{EM} is a constant that is dependent on the technology process, T is temperature, J is current density, E_a is the activation energy (0.85 eV), n=1 (void growth), and k is the Boltzmann constant.

Figure 14 – Formation of hillock and void resulting from electromigration.

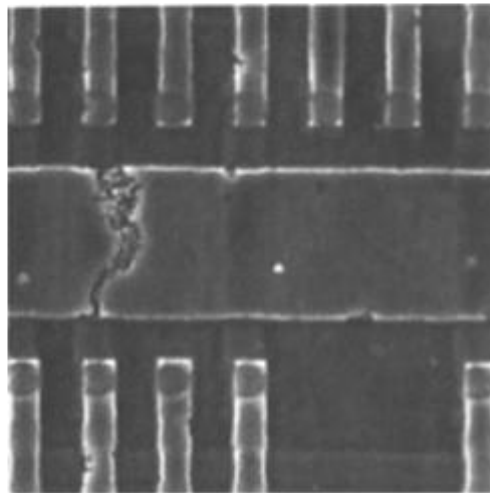


Figure 15 – Electromigration breakdown in an interconnect wire [52].

1.3.3 Middle-of-line Wearout Mechanism

Due to the increasing complexity of the technology nodes, the fabrication process for between the wafer fabrication and back-end assembly has now evolved into a separate process, called middle of the line (MEOL) process [53, 54]. The MEOL process can be performed after the front side treatment/bumping, or before the chip stacking assembly [55]. However, controlling the fabrication process between the polysilicon control gate and the diffusion contact may be challenging due to variations from the overlay, via size, line width, line edge roughness, defects, and image irregularities, which gives rise to dielectric breakdown in the middle-of-line [56].

Therefore, the middle-of-line time-dependent dielectric breakdown (MOL TDDB or MTDDB) is a growing concern due to the architecture change in FinFET transistors and dimension scaling [57], as shown in Figure 16, which can be found in advanced technology nodes [58]. MOL TDDB is similar to BEOL TDDB, but the dielectric breakdown occurs in the spacing between the gate and the source/drain contacts [59]. Therefore, the device-level model is similar to BEOL TDDB as follows [60]:

$$\eta = A_{MOL} L_i^{\frac{-1}{\beta}} exp\left(-\gamma E^m + \frac{E_a}{kT}\right) s^{-1} \quad (5)$$

where A_{MOL} is a constant the depends on the material properties of the dielectric, and m is 1 for the E model [61]. All other parameters are similar to the BEOL TDDB parameters.

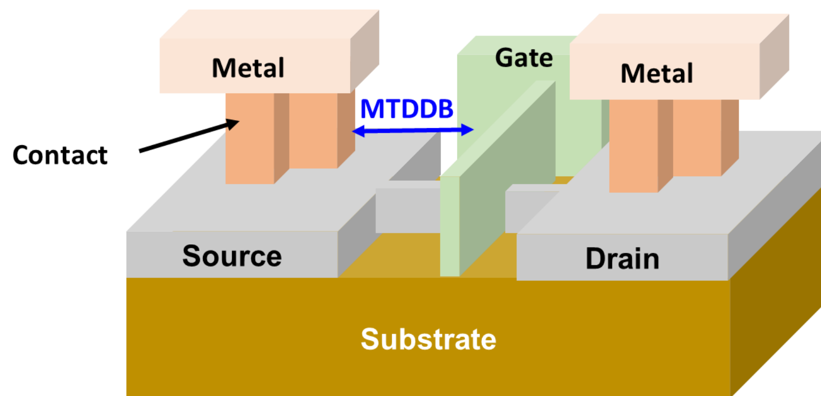


Figure 16 – MTDDB breakdown occurs in the region between the contact and the gate.

1.4 Circuit Case Studies

1.4.1 Process Design Kit (PDK)

The circuits used in this research are based on the 14nm FinFET technology node using the process design kit (PDK) jointly developed by GlobalFoundries (GF)/Samsung/IBM. A PDK is a set of files created by the foundry to model a fabrication process for the design tools that are used to design an integrated circuit. The circuits are investigated using the PDK files in Cadence Virtuoso, SPICE simulation, and Mentor Calibre for design rule checks of the layouts. The types of circuits used in this thesis are focused on ring oscillators and static random-access memories.

1.4.2 Ring Oscillators

Ring oscillators are a type of circuit that are often used in process validation [62, 63], such as in monitoring the gate delay and speed-power product of fabricated circuits due to easy implementation. The simplest type of a ring oscillator structure is composed of identical invertors and physical interconnections, as shown in Figure 17, where the output signal oscillates with a certain period depending on the gate delay. The gate delay is the length of time when the input to a logic gate becomes stable and valid to change, to the time that the output of that logic gate is stable and valid to change.

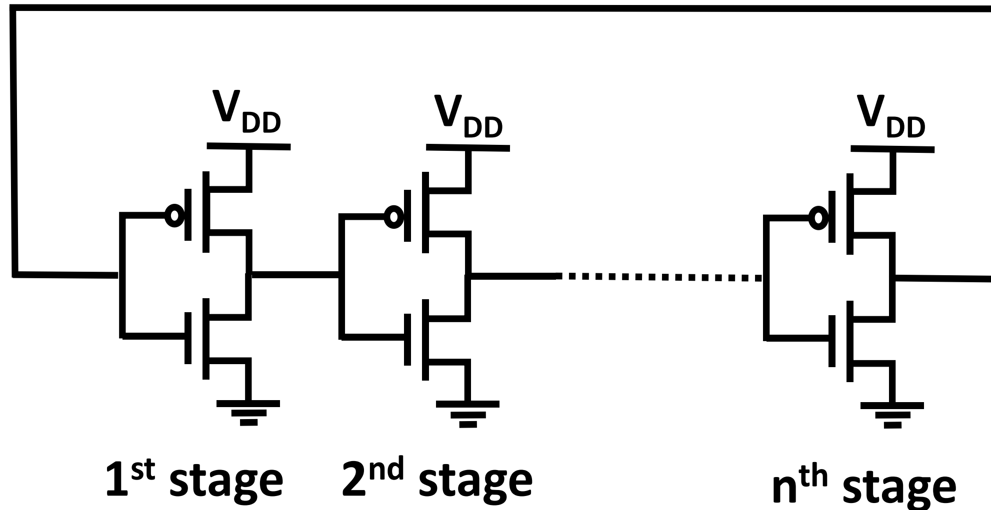


Figure 17 – Ring oscillator composed of identical invertors and physical interconnections connected together.

To oscillate, the ring oscillator requires an odd number of stages. A stage number is the number of invertors (nmos and pmos connected together in a series) in a ring oscillator. The duty cycle of a ring oscillator is the ratio of the time the circuit is on compared to the time the circuit is off, which can be used to describe the percentage of time a signal is active in the ring oscillator.

1.4.3 Static random-access memory (SRAM)

Static random-access memory (SRAM) is a type of circuit that retains data bits in its memory as long as power is supplied [64], occupying a major portion of the total area and power of system-on-chip ICs [65]. It is also the most common embedded-memory option for CMOS ICs [66]. The need for low power consumption and high performance for ultra-low power circuits, such as mobile and wearable devices, are the driving force for the demand for SRAMs [67].

As shown in Figure 18, an SRAM is composed of six transistors, where A_1 and A_2 are the access transistors, and L_1, L_2, L_3 and L_4 are the latch transistors. The data bit is stored in the latch transistors, which is the basic memory cell. The access transistors are used to control the access to the memory cell when reading and writing the data.

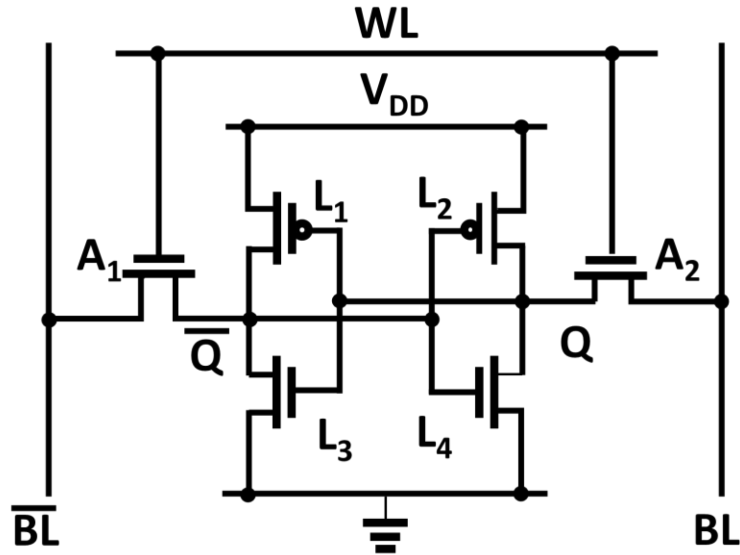


Figure 18 – An SRAM is composed of six transistors, where A_1 and A_2 are the access transistors, and L_1, L_2, L_3 and L_4 are the latch transistors.

1.5 Research Objectives

This thesis aims to implement data analysis techniques to detect and identify competing wearout mechanisms in circuits. A methodology for online monitoring is also developed to detect FEOL TDDB failures. Test plans and sampling for accelerated testing are also studied to lower errors.

1.6 Thesis Overview

This thesis is organized as follows. Chapter 2 will discuss the failure analysis for ring oscillators and SRAMs. Chapter 3 will describe the investigation of test conditions and sampling for accelerated testing. The conclusions and future work will be explained in Chapter 4.

FAILURE ANALYSIS OF CIRCUITS

This thesis focuses on two types of failure analysis of circuits. The first type is when the testing time is sufficient, so the entire failure data set can be collected. For these types of data sets, the failure data may be composed of more than one wearout mechanism, so the failure modes may need to be separated. The other type is for online monitoring, where the samples are monitored as they fail [68]. Generally, this type of failure data set will be incomplete, because not all samples may have failed. Through online monitoring, the goal is to prevent the user from experiencing the effects of failure and provide , notice of impending failure to allow corrective measures to be taken, which may be necessary in safety critical applications [69, 70].

1.7 Separation of competing wearout mechanisms in circuits

The standard method for analyzing failure samples is through examining test structures, with a typical test structure shown in Figure 19. However, actual products use circuits, which are more complex than test structures. One of the main differences between circuits and test structures is that test structures only have a single wearout mode, but circuits have confounded wearout modes. Therefore, the results from test structures may not be reflective of actual circuits.

The collection of lifetime data on large numbers of circuits is challenging, due to test setup cost and the need for large numbers of sample circuits. Ring oscillators are an intermediate between circuits and test structures. Because ring oscillators have behaviors similar to circuits, they are used instead of test structures in this thesis to test for failure modes.

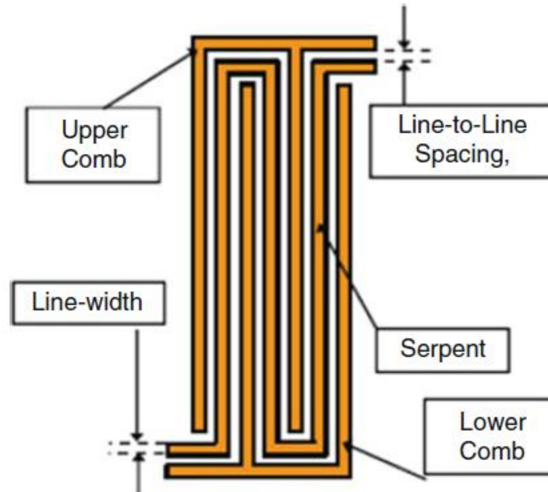


Figure 19 – Illustration of a test structure [12].

In addition, invasive diagnostic methods are generally used for failure analysis, such as transmission electron microscopy (TEM), e-beam, or scanning electron microscopy (SEM), to study the failure modes, which require samples to be cut open, for example, using focused ion beam (FIB) techniques [71, 72]. For advanced technology nodes, the metrology for TEM and other failure analysis techniques can become intricate and complex, requiring significant time to prepare and analyze samples [73]. This causes a wait time to receive the failure results and high costs, which can impact product costs if done too often.

Furthermore, electron radiation from TEM and other failure analysis techniques may alter the composition or microstructure of the sample, making it difficult to interpret the results authentically [74, 75]. Also, for technology nodes for 10nm and below, when using the energy dispersive X-ray spectrometer in TEM and other failure analysis techniques for elemental identification, the peak for the composition analysis may overlap due to more and more elements used in FEOL TDDB and MOL TDDB processes[76]. Therefore, it is necessary to find a quick and non-invasive method to separate the causes of failure, so that efforts for process improvement can be prioritized.

1.7.1 Methodology

To analyze the confounded wearout modes in a circuit, data analysis techniques are used to separate the competing wearout mechanisms. The advantages for using data analysis include quicker, faster, scalable, and more cost-effective analysis of complex data [77, 78], and may be implemented in the monitoring of semiconductor manufacturing [79]. The wearout mechanisms are modeled as competing wearout mechanisms, which occur when failures are due to more than one degradation mode and are independent of each other.

For competing wearout mechanisms, suppose mechanism 1 (primary breakdown mode) has a probability density function, $f_1(t)$ and cumulative distribution function, $F_1(t)$. The survival function is $R_1(t) = 1 - F_1(t)$. Similarly, mechanism 2 (secondary breakdown mode) has the probability density function, $f_2(t)$; cumulative distribution function, $F_2(t)$; survival function, $R_2(t)$. Thus, the competing failure probability density function, $f(t)$, can be described as below [80].

$$
\begin{aligned}
f(t) &= P\{T_1 = t, T_2 \geq t\} \cup P\{T_1 \geq t, T_2 = t\} \\
&= P(T_1 = t, T_2 \geq t) + P(T_1 \geq t, T_2 = t) \\
&= P(T_1 = t)\, P(T_2 \geq t) + P(T_1 \geq t)\, P(T_2 = t) \\
&= f_1(t) * R_2(t) + f_2(t) * R_1(t)
\end{aligned}
\tag{6}
$$

The competing failure probability density function is different than the mixed Weibull probability density function, as shown below [81]:

$$f(t) = a * f_1(t) + b * f_2(t) \quad (7)$$

where a and b are the mixed weights. The mixed Weibull probability density function occurs when the breakdown is due to both mechanisms at the same time and should not be confused with the competing Weibull probability density function [82]. The competing Weibull probability density function describes the breakdown at a specific failure time due to only one mechanism, but the cause can be from either mechanism 1 or 2, but not both.

MLE is employed to estimate the competing Weibull parameters. The parameter values are found by maximizing the likelihood that the process described by the model has produced the data that were observed using the likelihood function. The likelihood function describes how particular values of statistical parameters are for a given set of failure observations, and for uncensored data can be simplified from [80] as:

$$\mathcal{L}(\theta) = C \prod_{i=1}^{N} f(t_i) \quad (8)$$

where θ is the set of competing Weibull parameters, β_1, β_2, η_1, η_2. Generally, because the goal is to obtain θ, not the actual value of $\mathcal{L}(\theta)$, it is easier to work with the log likelihood function compared to the likelihood function. The log likelihood function can be written as:

$$\ln \mathcal{L}(\theta) = \sum_{i=1}^{N} \ln f(t_i) + \ln C \quad (9)$$

where C is a constant.

The derivatives for MLE can also be simplified from [80] as:

$$\frac{\partial \ \ln \mathcal{L}(\theta)}{\partial \beta_1} = \sum_{i=1}^{N} \left\{ \frac{dR_1(t_i)}{d\beta_1} f_2(t_i) + \frac{df_1(t_i)}{d\beta_1} R_2(t_i) \right\} / f(t_i) \tag{10}$$

$$\frac{\partial \ \ln \mathcal{L}(\theta)}{\partial \beta_2} = \sum_{i=1}^{N} \left\{ R_1(t_i) \frac{df_2(t_i)}{d\beta_2} + f_1(t_i) \frac{dR_2(t_i)}{d\beta_2} \right\} / f(t_i) \tag{11}$$

$$\frac{\partial \ \ln \mathcal{L}(\theta)}{\partial \eta_1} = \sum_{i=1}^{N} \left\{ \frac{dR_1(t_i)}{d\eta_1} f_2(t_i) + \frac{df_1(t_i)}{d\eta_1} R_2(t_i) \right\} / f(t_i) \tag{12}$$

$$\frac{\partial \ \ln \mathcal{L}(\theta)}{\partial \eta_2} = \sum_{i=1}^{N} \left\{ R_1(t_i) \frac{df_2(t_i)}{d\eta_2} + f_1(t_i) \frac{dR_2(t_i)}{d\eta_2} \right\} / f(t_i) \tag{13}$$

$$\frac{dR_k(t)}{d\beta_k} = -R_k(t) \left(\frac{t}{\eta_k} \right)^{\beta_k} \ln \left(\frac{t}{\eta_k} \right) \tag{14}$$

$$\frac{df_k(t)}{d\beta_k} = \left(\frac{t}{\eta_k} \right)^{\beta_k - 1} R_k(t) \left\{ \left(\frac{1}{\eta_k} \right) + \left(\frac{\beta_k}{\eta_k} \right) \ln \left(\frac{t}{\eta_k} \right) \right\} - f_k(t) \left(\frac{t}{\eta_k} \right)^{\beta_k} \ln \left(\frac{t}{\eta_k} \right) \tag{15}$$

$$\frac{dR_k(t)}{d\eta_k} = R_k(t) \left(\frac{\beta_k}{\eta_k} \right) \left(\frac{t}{\eta_k} \right)^{\beta_k} \tag{16}$$

$$\frac{df_k(t)}{d\eta_k} = \left(\frac{\beta_k}{\eta_k} \right)^2 \left(\frac{t}{\eta_k} \right)^{\beta_k - 1} R_k(t) \left\{ -1 + \left(\frac{t}{\eta_k} \right)^{\beta_k} \right\} \tag{17}$$

where k=1 or 2.

The quasi-Newton method using the Davidon-Fletcher-Powell algorithm, also referred to as the variable metric method, is used to optimize equation (9). The algorithm was originally proposed by Davidon in 1959 and later developed by Flectcher and Power in 1963 [83]. Because the evaluation and use of the Hessian matrix is impractical, time-consuming and costly, the Davidon-Fletcher-Powell algorithm approximates the inverse Hessian matrix instead. An initial matrix H_0 is chosen (usually $H_0=I$, where I is the identity matrix, also called unit matrix), and the inverse Hessian is updated by the sum of two symmetric rank one matrices, which allows the algorithm to run faster than a rank two calculation. The updates continue until the optimization point is reached.

The Davidon-Fletcher-Powell algorithm was chosen, because this algorithm is suitable for a data set on the order of 10 to 1000 samples. Industrial data sets for reliability failure times are generally on the order of 10 to 1000 samples. Data sets larger than 1000 samples may take too long to monitor or may be too costly. Therefore, the Davidon-Fletcher-Powell algorithm can be applied to the failure sets and can be used to analyze the data sets quickly. The implementation is shown in Figure 20 [80].

Algorithm Procedure [80]
1. Initial condition: $\theta^0 = [\beta_1, \beta_2, \eta_1, \eta_2]^T$, $L^0 = lnL(\theta^0)$
2. Set optimization direction:
$d^i = -S_i g^i$, where $S_i = I(the\ unit\ matrix)$,
$g^i = -\nabla \ln L\left(\theta^i\right)^T$, i=0
3. Line search: $\theta^{i+1} = \theta^i + \alpha_i d^i$, where α_i is the optimal step length
4. Calculate parameters for Hessian matrix and new direction:
$p^i = \alpha_i d^i,\ g^{i+1} = -\nabla \ln L\left(\theta^{i+1}\right)^T$, $q_i = g^{i+1} - g^i$
5. Estimate inverse Hessian matrix:
$$S_{i+1} = S_i + \frac{p^i p^{iT}}{p^{iT} q^i} - \frac{S_i q^i q^{iT} S_i}{q^{iT} S_i q^i}$$
6. Find set of Weibull parameters: Set i=i+1. If i=4 (number of Weibull parameters) then go to Step 7; otherwise go back to step 2.
7. Iteration procedure for optimization and stop condition:
If $\left|\ln L(\theta^i) - L^0\right| < \varepsilon$, then stop. Otherwise, set $L^0 = \ln L\ (\theta^i)$, i=0 and go back to step 2.

Figure 20 – Implementation of the MLE algorithm.

1.7.2 Investigation of Initial Conditions vs. Sample Size

The competing wearout mechanisms for 11-stage ring oscillators based on the 14nm pdk FinFET technology node jointly developed by GlobalFoundries(GF)/Samsung/ IBM were studied using the above method to extract the competing Weibull parameters from the confounded failure data. These ring oscillators have Weibull parameters of FEOL TDDB with β_1=1.12, η_1=9.87 (yrs) and MOL TDDB with β_2=1.9, η_2=15.36 (yrs).

Before using the MLE algorithm to extract the overall competing Weibull parameters, the failure distributions of the competing wearout mechanisms were modeled by first picking a point randomly from each individual distribution. Next, the smaller value is set as the lifetime, because it is the mechanism that fails first at that time point. Then,

the points are plotted as ordered pairs: (ln(t_1 ln(-ln(1-($\frac{1}{2N}$))), (ln(t_2), ln(-ln(1-($\frac{3}{2N}$))), etc. This was done for sample sizes N of 10, 100 and 100, as shown in Figure 21.

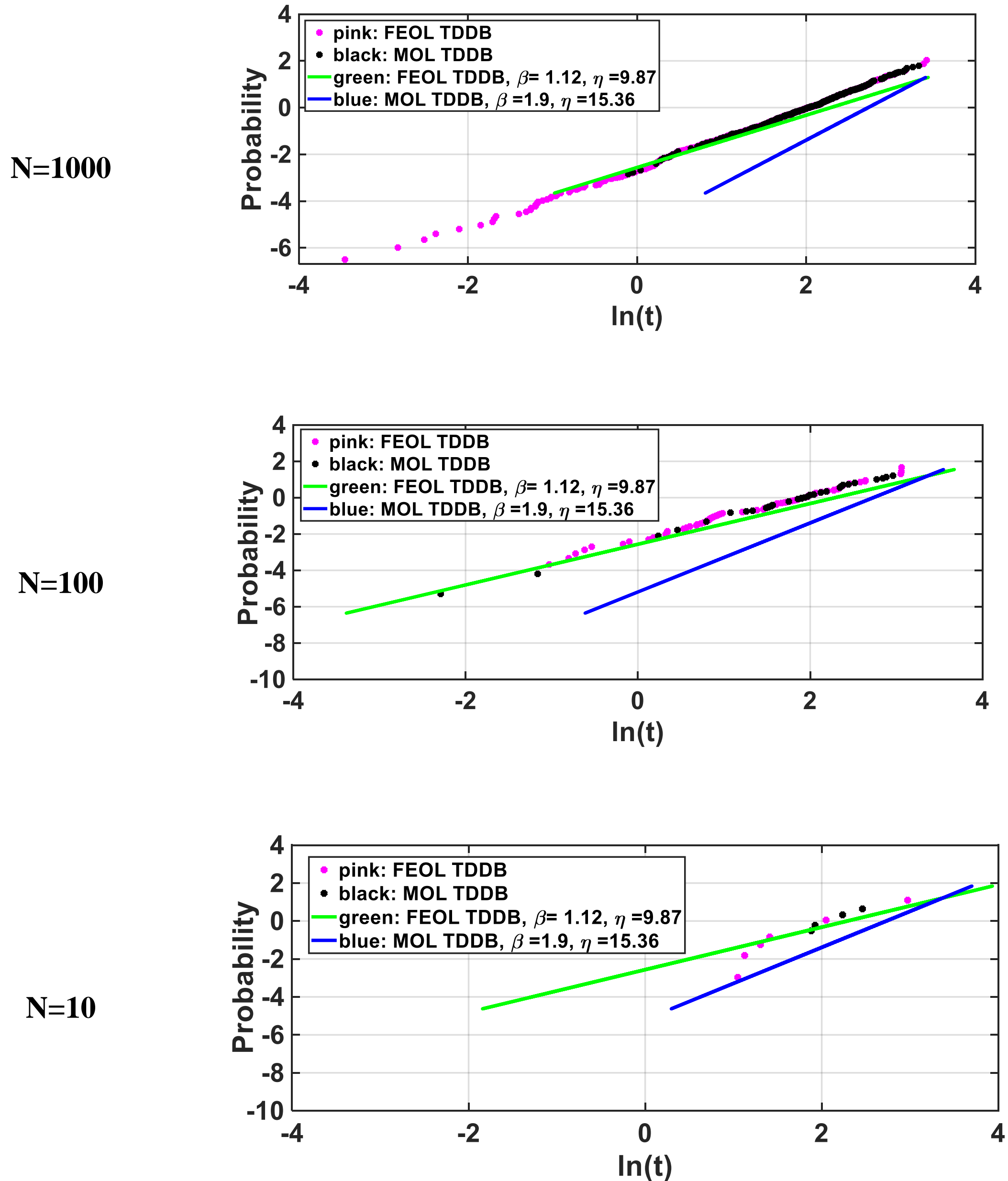


Figure 21 – Distributions of data for 11-stage ring oscillators for FEOL TDDB with β_1=1.12, η_1=9.87 (yrs) and MOL TDDB with β_2=1.9, η_2=15.36 (yrs) (x axis: unit years), varying sample size. P is probability. The pink markers correspond to the primary wearout mechanism, while the black markers correspond to the secondary wearout mechanism.

The error results of the extracted competing Weibull parameters are shown in Figure 22. The initial conditions were set at the original values to exclude the effects of randomness. As the sample size is increased, the errors for all four parameters decrease, and when the sample size is increased from an order of magnitude from 10 to 1000, the parameter errors generally decrease around 5-fold. The shape parameter of the second competing Weibull mechanism, β_2, has the highest error in all cases. The β_2 value is harder to separate since it is closer to the β_1 value, and there are fewer samples originating from the second mechanism since it fails later.

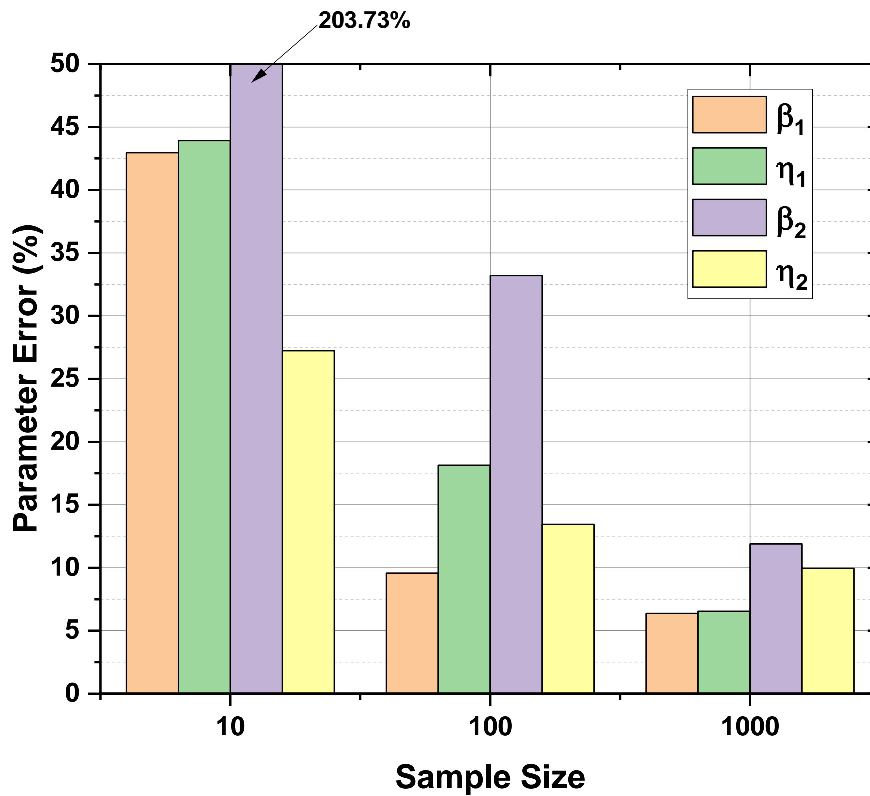


Figure 22 – Parameter errors as a function of sample size for extracted competing mechanisms, FEOL TDDB with β_1=1.12, η_1=9.87 (yrs) and MOL TDDB with β_2=1.9, η2=15.36 (yrs).

For the cases where the initial conditions are not at the initial values, the shape and scale parameters, η and β, were both set at the same deviation from the original values at the same time. In other words, the β and η of both mechanisms were both set at 5% deviation from the initial condition, and MLE was employed to obtain the estimation results. Then, this procedure was repeated by setting both β and η of both mechanisms at 10% deviation from the initial condition, and the process was repeated again increasing the deviation by 5% each time up to the 15% deviation from the initial condition. The entire procedure was also repeated for deviation from -5% to -15% from the initial condition.

For a sample size of 10, the wearout mechanisms were able to be distinguished up to a deviation error of $\pm 5\%$ from the actual value for the initial condition, while the sample sizes of 100 and 1000 could be separated for up to a deviation error of $\pm 15\%$ from the actual value for the initial condition.

1.7.3 Identification of Wearout Mechanism for Each Individual Sample

The previous section described a methodology to determine which competing wearout mechanisms are present in a set of failure samples. However, it is not known which samples belong to which degradation mode. Here, a methodology is developed to identify the probabilistic origin of failure for each monitored sample, determine the region of error indicating the time period where the cause of failure is unknown, and analyze the sorting accuracy. In doing so, only the necessary samples for physical failure analysis are selected instead of all failure samples, saving time and money.

As mentioned previously, the competing failure probability density function, f(t), can be described as:

$$f(t) = f_1(t) * R_2 (t) + f_2 (t) * R_1(t) \quad (7)$$

Therefore, the competing probability density function contribution from mechanism 1, called a_1, is defined below:

$$a_1 = f_1(t) * R_2(t) \quad (18)$$

which is the probability density function portion of the overall system showing that mechanism 1 has failed but mechanism 2 is still working. Similarly, the competing probability density function contribution from mechanism 2, called a_2, can be defined as:

$$a_2 = f_2(t) * R_1(t) \quad (19)$$

If the overall Weibull parameters for each set is known from the algorithm analysis in the previous section, each failure sample can be further sorted into its respective failure distribution. For each failure time point, the time-to-failure value can be inputted into each competing probability density function contribution, equation (18) and equation (19), for each distribution. Since a higher value represents the higher probability of the sample belonging to that respective distribution, one can compare the relative values, or ratio of equation (18) and equation (19), to sort the samples.

An interesting phenomenon is that for the case of competing wearout mechanisms, the relative values or the ratio of the hazard function is also the same as comparing eq. (18) to eq. (19). The hazard function for mechanism 1 is:

$$h_1(t) = \frac{f_1(t)}{R_1(t)} \quad (20)$$

and similarly, the hazard function for mechanism 2 is:

$$h_2(t) = \frac{f_2(t)}{R_2(t)} \tag{21}$$

The hazard function, also known as the instantaneous failure rate, shows the conditional probability of a failure given that the system is currently working. When multiplying both sides of eq. (20) or eq. (20) by R_1 (t)*R_2 (t), they can be rewritten as:

$$R_1(t) * R_2(t) * h_1(t) = f_1(t) * R_2(t) \tag{22}$$

and:

$$R_1(t) * R_2(t) * h_2(t) = f_2(t) * R_1(t) \tag{23}$$

where the right sides of eqs. (22) and (23) equal eqs. (18) and (29), respectively. Since only the relative values or ratio, not the absolute value, is needed, using the hazard function to sort the samples has the same results as using the competing probability density function contributions.

Looking back at equation (12), at any time point, the competing failure probability is always composed of two contributions, f_1(t)* R_2(t) and f_2(t) * R_1(t). Therefore, x, which is the percentage of failures from mechanism 1 at a given time t, can be found by:

$$x = \frac{f_1(t)*R_2(t)}{f_1(t)*R_2(t) + f_2(t)*R_1(t)} \tag{24}$$

and y, which is the percentage of failures from mechanism 2 at a given time t, is:

$$y = \frac{f_2(t)*R_1(t)}{f_1(t)*R_2(t) + f_2(t)*R_1(t)} \tag{25}$$

Plotting equation (24) and (21) for all failure times will show the region where error will most likely be highest, which occurs near x=y=0.5, meaning that there is a 50% probability that the sorting could be right for either distribution. This is shown as an example in Figure 23, which looks at FEOL TDDB with β_1=1.64, η_1=9.87 and EM with β_2=1.14, η_2=25.1296 for a sample size of 100 in 14nm FinFET ring oscillators. The plot will also show the region where one distribution has a 100% probability of showing up (a.k.a. the other distribution having a 0% probability of showing up), meaning that this region can have failure samples sorted to their relative distributions without any inaccuracies. When the distribution's 100% probability lowers, any future time point may be sorted incorrectly, which is called the region of error. This region identifies the time periods that are most important for physical failure analysis, where it may be necessary to perform further in-depth diagnosis.

No difference was found in the analysis of sorting errors between the original and MLE extracted parameters. The region of no error was found to increase when either the sample size or β ratio of the dominant wearout mechanism to secondary wearout mechanism is decreased. The sorting accuracy was also found to increase as the β ratio increases but varies slightly with a difference in sample size. When the percentage of each failure distribution is near 50%, there is a higher probability of the samples being sorted to the wrong distribution, due to the risk of the wrong categorization being around 50% too. This information can be used to signal that the samples near this area are the only ones that one may need to perform physical failure analysis, not the entire lot, which saves analysis costs. Therefore, the above methodology is a quick procedure to perform preliminary

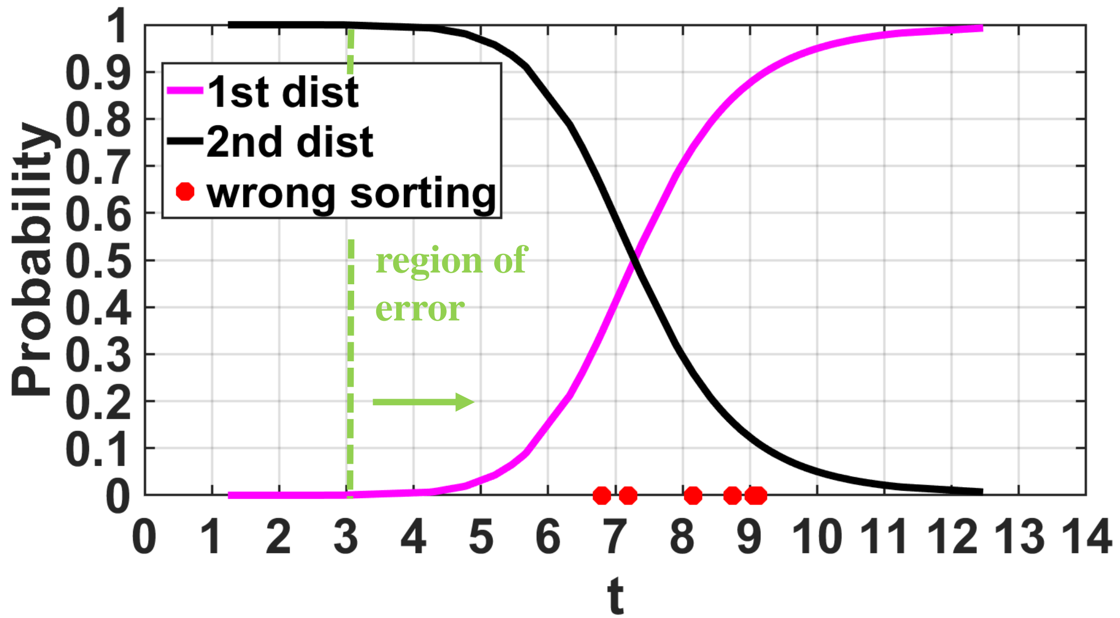


Figure 23 – Sorting accuracy of 93.0% for FEOL TDDB (1st dist.) vs. EM failure (2nd dist.) in 14nm FinFET ring oscillators for a sample size of 100.

screening to identify the wearout mechanism in individual samples for confounded wearout modes.

1.7.3.1 Case Study

A case study using 11-stage ring oscillators based on the 14nm GlobalFoundries/IBM/Samsung FinFET technology node was used to investigate the effects of the competing wearout mechanisms. The parameters for the 14nm FinFET technology node were extracted from experimental data [84-86]. The 11-stage ring oscillators have FEOL TDDB wearout parameters of β=1.64, η=20 yrs, and MOL TDDB wearout parameters of β=1, η=10 yrs.

The selectivity of a mechanism is the probability that the failure is caused by that mechanism when there are multiple failure possibilities. If there are two wearout

mechanisms, x and y with Weibull distributions, η_x, β_x and η_y, β_y, respectively, selectivity can be computed as follows [87]:

$$selectivity = \frac{P_{x_fail_first}}{P_{fail}} \tag{26}$$

where

$$P_{x_{fail_{first}}} = 1 - \exp\left(-\left(\frac{t_{stop}}{\eta_y}\right)^{\beta_y}\right) - \int_0^{\left(\frac{t_{stop}}{\eta_y}\right)^{\beta_y}} \exp\left(-u - \left(\frac{\eta_x}{\eta_y}\right)^{\beta_x} u^{\frac{\beta_x}{\beta_y}}\right) du \tag{27}$$

with t_{stop}=testing time

and

$$P_{fail} = \left(1 - exp\left(-\left(\frac{t_{stop}}{\eta_x}\right)^{\beta_x}\right)\right)\left(1 - exp\left(-\left(\frac{t_{stop}}{\eta_x}\right)^{\beta_x}\right)\right) \tag{28}$$

The selectivity for FEOL TDDB and MOL TDDB was found for various voltages and temperatures, as shown in Figure 24 and Figure 25. FEOL TDDB selectivity is higher at higher voltages and temperatures, while MOL TDDB is preferred at lower voltages and temperatures.

To shorten the test time, the selectivity maps were used to find a region where the FEOL TDDB and MOL TDDB had a selectivities of 0.5, and the ring oscillators were accelerated to 1.37 V, 199.72°C. The test times for the ring oscillators to reach 100 samples are shown in Figure 26, where it can be seen that there is an even mix of samples from both

wearout mechanisms. The 100th sample failure time out of 100 samples at FEOL TDDB selectivity=0.5 at various voltages and temperatures are shown in Figure 27, while the 1000th sample failure time out of 1000 samples at the same conditions are shown in Figure 28. The testing times for the 100th sample failure and 1000th sample failure can decrease by three orders depending on the accelerated testing conditions. For the same accelerated conditions of 1.37 V, 199.72ºC, the 100th sample failure time out of 100 samples is 9.3 days, while the 1000th sample failure time out of 1000 samples is 13.34 days.

The confidence interval for the extracted competing wearout mechanism are shown in Table I. Because the two competing wearout mechanisms have parameters that are close together, it is harder to separate the all the parameters, which results in a larger confidence interval level range. Also, because the selectivity is at 0.5, meaning that there is an equal probability of either sample failing at a failure point, the sorting accuracy is also lower, which is 61%, as shown in Figure 29.

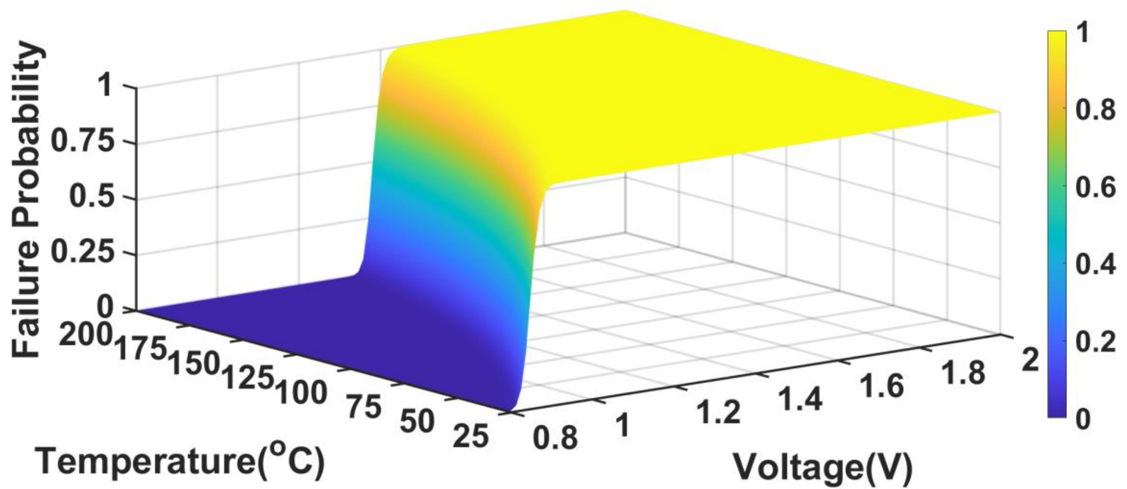


Figure 24 – FEOL TDDB selectivity in 14nm FinFET ring oscillators.

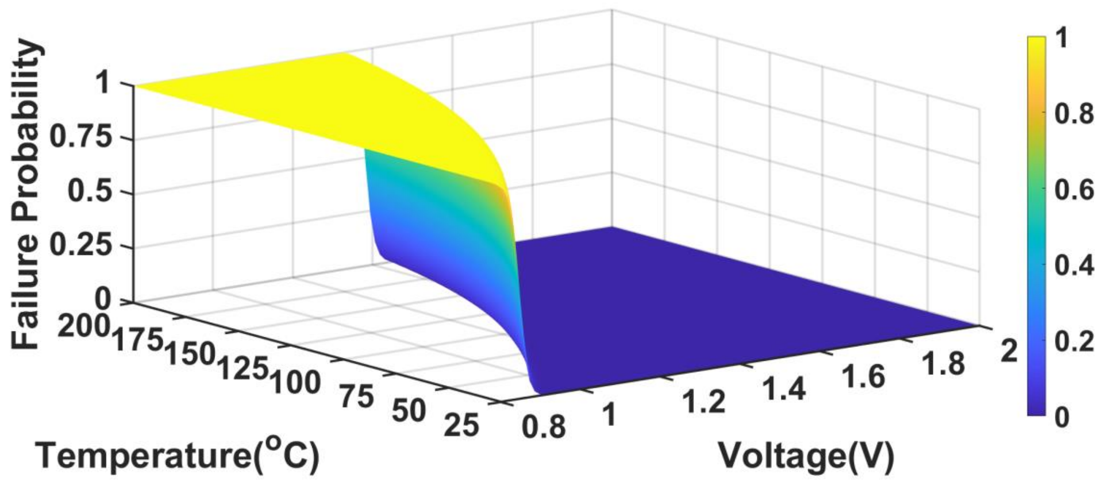


Figure 25 – MOL TDDB selectivity in 14nm FinFET ring oscillators.

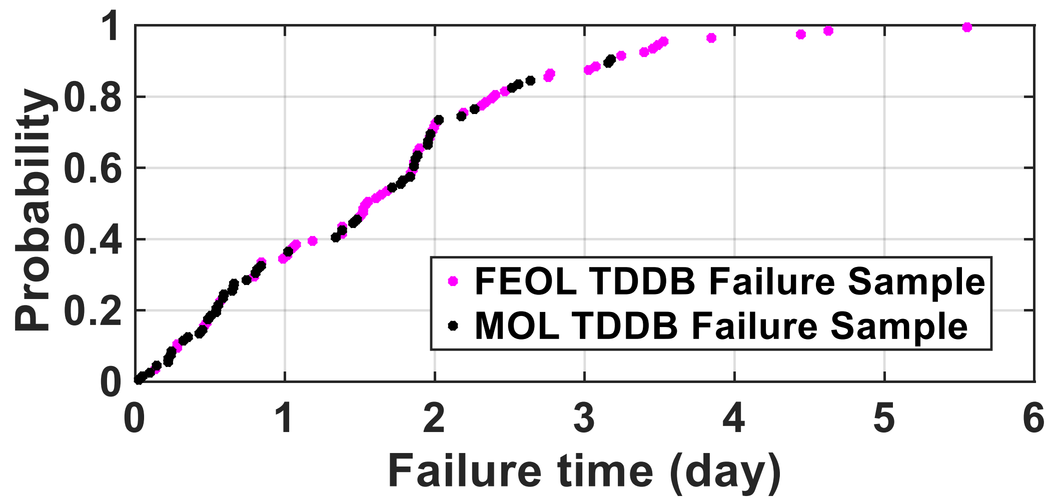
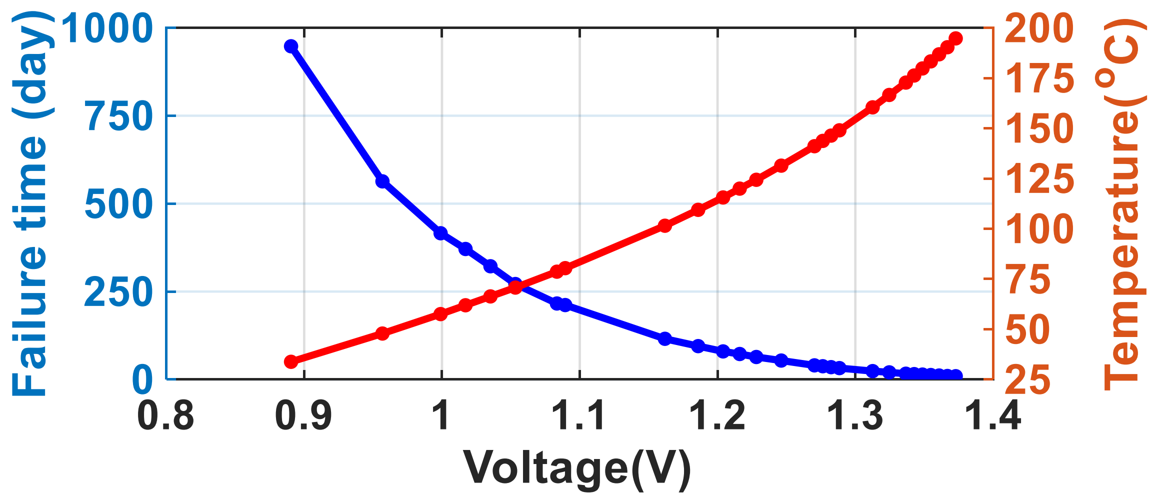
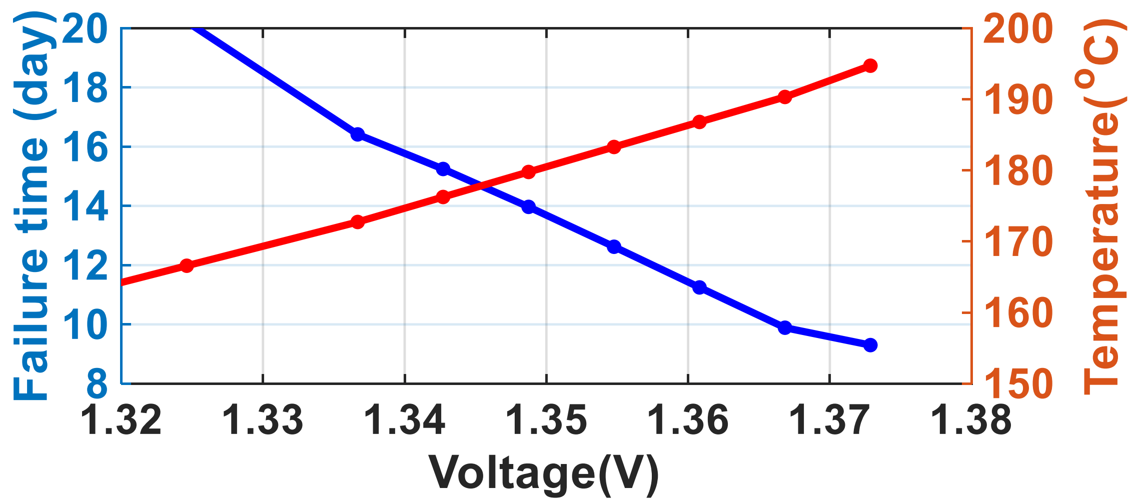


Figure 26 – Test times for 11-stage ring oscillators at accelerated conditions (1.37V, 194.72OC) for a sample size of 100.

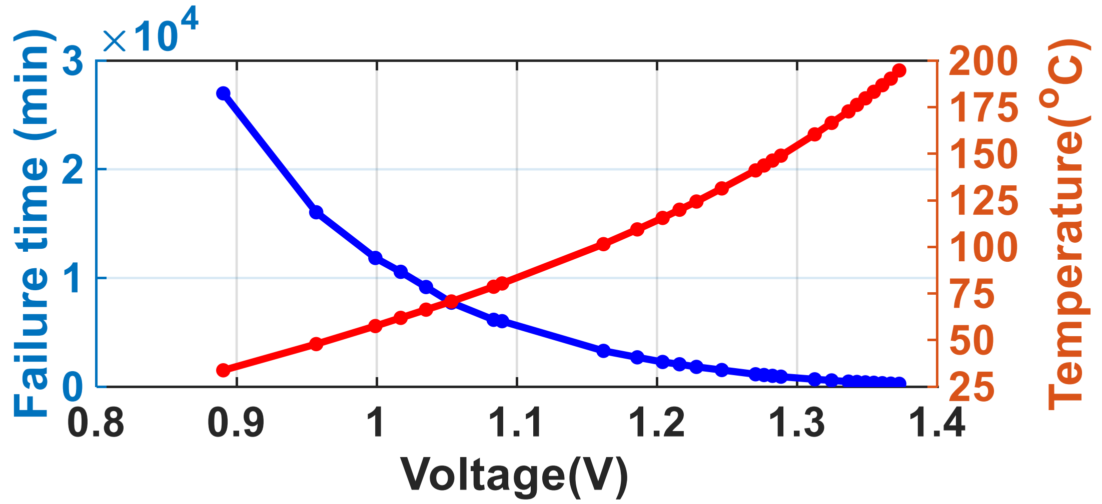


(a)

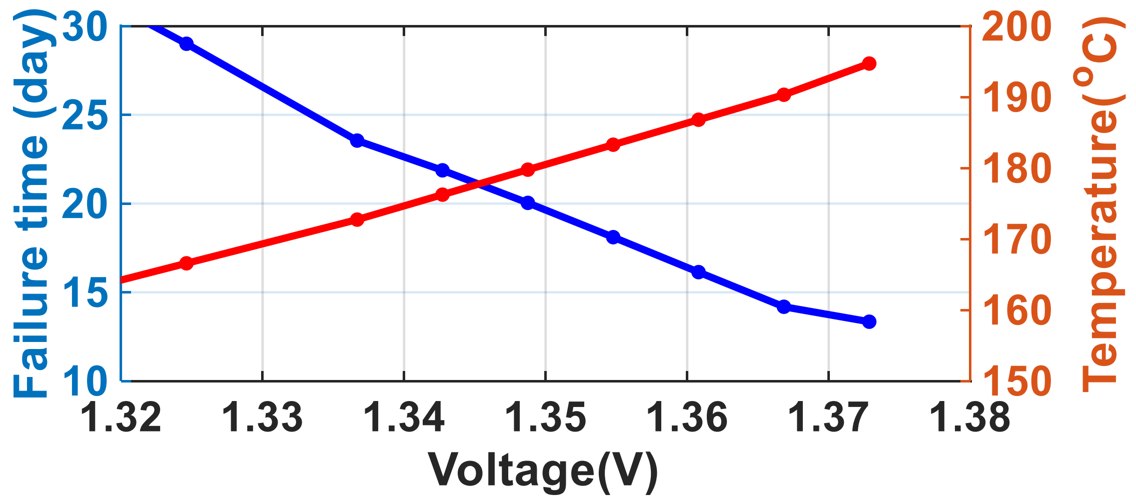


(b)

Figure 27 – Failure time for 100th sample out of 100 samples at FEOL TDDB=0.5 selectivity (a) as a function of voltage and temperature (b) close-up.



(a)



(b)

Figure 28 – Failure time for 1000th sample out of 1000 samples at FEOL TDDB=0.5 selectivity (a) as a function of voltage and temperature (b) close-up.

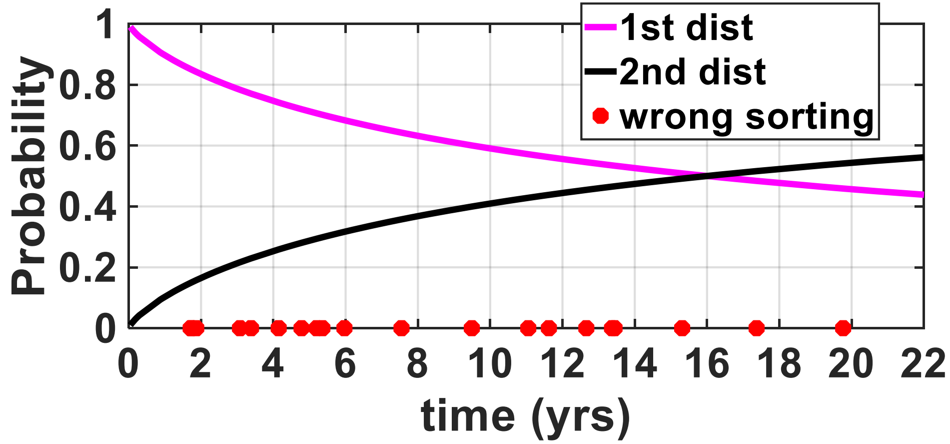


Figure 29 – Sorting accuracy of 80% for FEOL TDDB (1st dist.) vs. MOL failure (2nd dist.) in 14nm FinFET ring oscillators for a sample size of 100.

Table 1

95% Confidence Intervals for 11-stage ring oscillators at accelerated conditions (1.37V, 194.72°C)

	FEOL TDDB		**MOL TDDB**	
	Beta1	**Eta1**	**Beta2**	**Eta2**
Original Values	***1.64***	***2.88 (days)***	***1.9***	***3.13 (days)***
Sample Size = 100	2.39 ± 1.39	3.19 ± 2.09	1.01 ± 0.48	2.46 ± 3.00

1.7.3.2 Application to Trojan Detection

The methodology of extracting wearout parameters with MLE can also be applied to detect Trojans and to select suspicious samples for failure analysis. Instead of extracting parameters for two confounded distributions, we assume a known distribution for mechanism 1, and use MLE to extract the parameters for mechanism 2 based on the data. Since hardware Trojans are triggered by unlikely events and accelerate a specific wearout mode depending on its design, a worst-case scenario is used as a study, where the original GTDDB parameters are β_1=1.64, η_1=10 yrs, and Trojan affected samples have altered GTDDB parameters to β_2=1.64, η_2=5 yrs in a 14nm FinFET 501-stage ring oscillator. For a sample size of 100, the sorting accuracy was found to be 80%.

1.8 Failure Analysis for On-line Testing

For online testing, product failures are observed one by one as they fail with the increase of time. Because the time to observe the failures of all products may be too long, on-line testing generally has a limited set of samples, instead of a collection of all samples. Therefore, this thesis researches a methodology to determine the wearout parameters of failure samples as the samples fail immediately, through using on-line data collected during operations.

SRAMs are major components of systems-on-chips and are also used for memory in systems that require very low power consumption and easy access to data [88]. It is static and volatile, where data retention exists as long as the device is powered without any form of a refresh. When the power is cut, data will be lost. Because it is random access, the next memory location that can be read or written does not depend on the previous access

location. The static property of SRAM is due to the feedback mechanism used to maintain the stored bit state.

To ensure that memory operation is stable during operation, the reliability of SRAMs need to be considered. Oxide layer failures cause transistor malfunctions that translate to the circuit level, such as the flipping of cell data due to voltages lower than the nominal one, making FEOL TDDB an important cause of concern [89]. In this study, we use data on failures in the SRAM to estimate the wearout model parameters of FEOL TDDB, which are based on a time-to-failure stamp. To detect failure rates, it is necessary to monitor actual failures and to link these failures to lifetime models. Because of the large number of identical cells, the SRAM can be used to detect the characteristics of wearout due to FEOL TDDB. Therefore, the SRAM data is used as a vehicle to appropriately estimate the model parameters. The parameters to be extracted are based on the two parameter Weibull distribution for FEOL TDDB, which are the characteristic lifetime, η, and the shape parameter, β. The model parameters are extracted from time-to-failure data from the cells in the SRAM.

In addition, this study also investigates the accuracy in extracting the model parameters by considering both random and systematic errors. The random errors occur due to the availability of samples (failed SRAM cells). Systematic errors occur from usage variations, such as supply voltage and operating temperature fluctuations, as well as variations due to process parameters and workload. The analysis of systematic errors is used to determine when and if sensor data is needed to supplement analytical wearout models when estimating wearout model parameters.

1.8.1 Methodology

The lifetime distribution of a device due to wearout by front-end gate oxide breakdown (FEOL TDDB) can be found by [87]:

$$P(t) = 1 - exp\left(-\left(\frac{t}{\eta}\right)^{\beta}\right) \tag{29}$$

where η is the characteristic lifetime and β is the shape parameter.

The characteristic lifetime of the SRAM, η_{SRAM}, is a combination of Weibull distributions for the components, and is the solution of [90]:

$$1 = \sum_{i=1}^{n} \left(\frac{\eta_{SRAM}}{\eta_i}\right)^{\beta_i} \tag{30}$$

where η_i, i=1, …, n are the characteristic lifetime of all the circuit components, and β_i are the corresponding shape parameters. Similarly, it can be found that [91]:

$$\beta_{SRAM} = \sum_{i=1}^{n} \beta_i \left(\frac{\eta_{SRAM}}{\eta_i}\right)^{\beta_i} \tag{31}$$

When investigating the failure of the SRAM, the probability of stress for the circuit needs to be considered, because the usage scenario needs to be taken into account, where the circuit may be on and off at different times. For example, if the SRAM stores logic “1” 50% of the time and logic “0” 50% of the time, then s=0.5 in Equation (2) for all cells’ four transistors in the latch. s≈0 for the access transistors because it is only turned on when the

cell is accessed. 50% is set as the baseline for comparison. If the duty cycle is changed, then the SRAM will degrade at a different rate.

The SRAM failures due to FEOL TDDB during operation are calculated using Monte Carlo simulation. The random variable is the failure probability in Equation (18). The resulting data are time stamps for the failures of SRAM cells. By using the sequence of time stamps for SRAM failures due to FEOL TDDB, the Weibull parameters are extracted using generalized maximum likelihood estimation [92].

The original SRAM cell parameters for FEOL TDDB degradation in this study are η=20 years and β=1.12. As shown in Figure 30 (a), the SRAM failure samples for a sample size of 94,000 (an SRAM with 94k cells) due to FEOL TDDB are plotted on a Weibull plot. The FEOL TDDB extracted parameters are η=19.966 and β=1.119. Figure 30 (b) shows the case where only the first 100 failed samples are available. The FEOL TDDB extracted parameters are η=39.66, β=0.99 for the first 100 samples, which are far from the actual parameters, η= 20 and β=1.12.

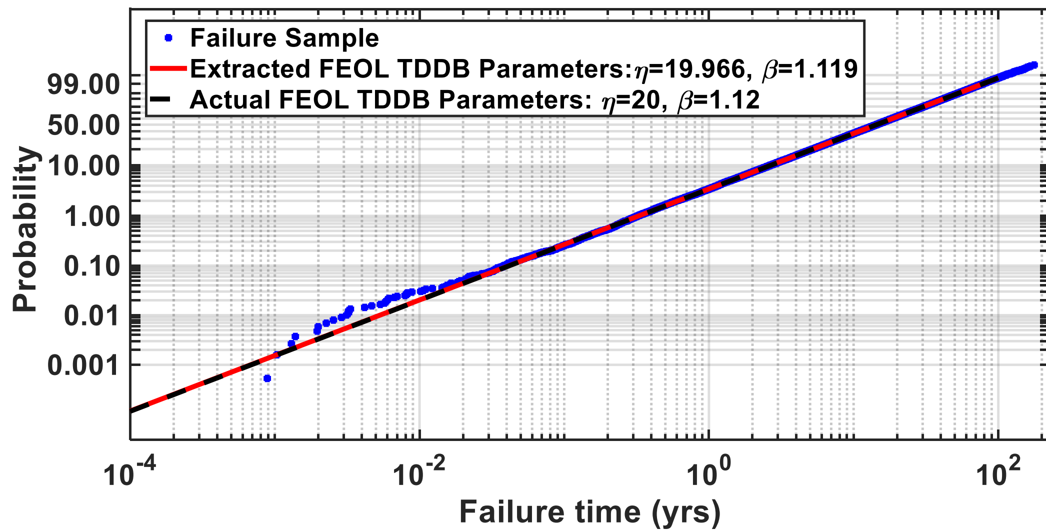
When data is collected during operations, information will generally be available for only part of the samples. As the sample size is increased, the FEOL TDDB extracted parameters become closer to the actual parameters. However, the monitoring time may be too long or the cost may be too great to be able to collect all the failure information from every sample.

The wearout model, Equation (2), are for single devices with s=1, where the probability of stress is considered to be the always on. The degradation for the devices will be constant with time. However, the observed data is from an SRAM cell, where the

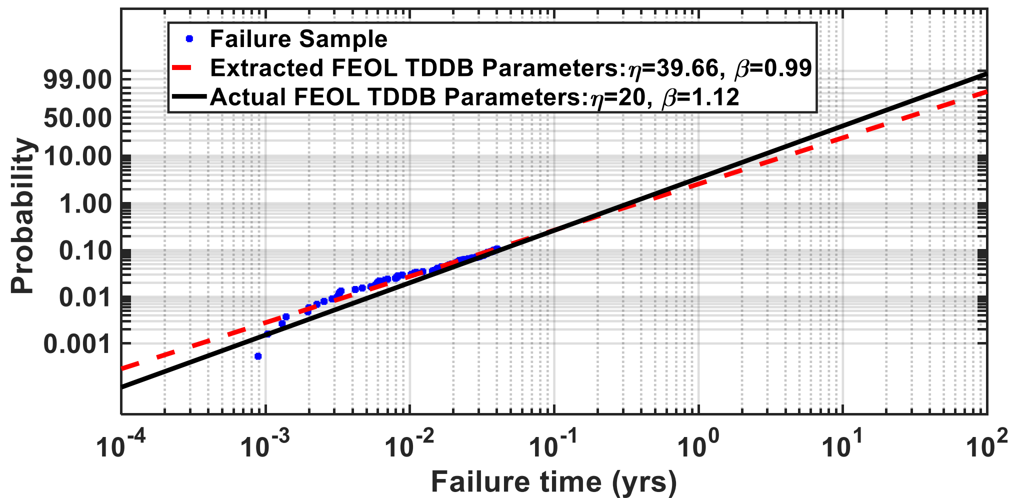
probability of stress may be different, because the circuit may be on or off depending on the time.

The probability of stress in an SRAM is based on the duty cycle. A duty cycle of 0 refers to the cell storing 0 all the time, and a duty cycle of 0.5 means that the cell stores 0 half of the time. Generally, duty cycle distributions can be found to be around 30-50% [94]. In this case, the wearout model parameters observed from SRAM data (η and β) are not the same as the single devices, but are for collections of devices, as computed with Equation (30) and (31). Therefore, the observed results need to be mapped to the device model.

Simulation is used to find the mapping between the process-level Weibull parameters and SRAM cell Weibull parameters, as shown in Figure 31. However, because only the SRAM parameters are observed during testing, the maps in Figure 31 need to be inverted to map the SRAM parameters into device model parameters, as shown in Figure 32. Therefore, Figure 32 is used to find the device model parameters from the observed failure data of the SRAMs.

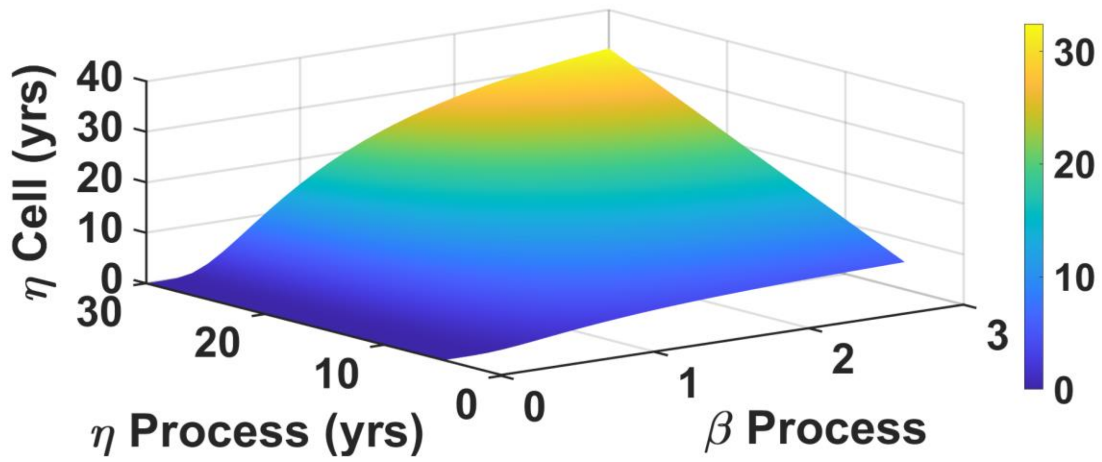


(a)

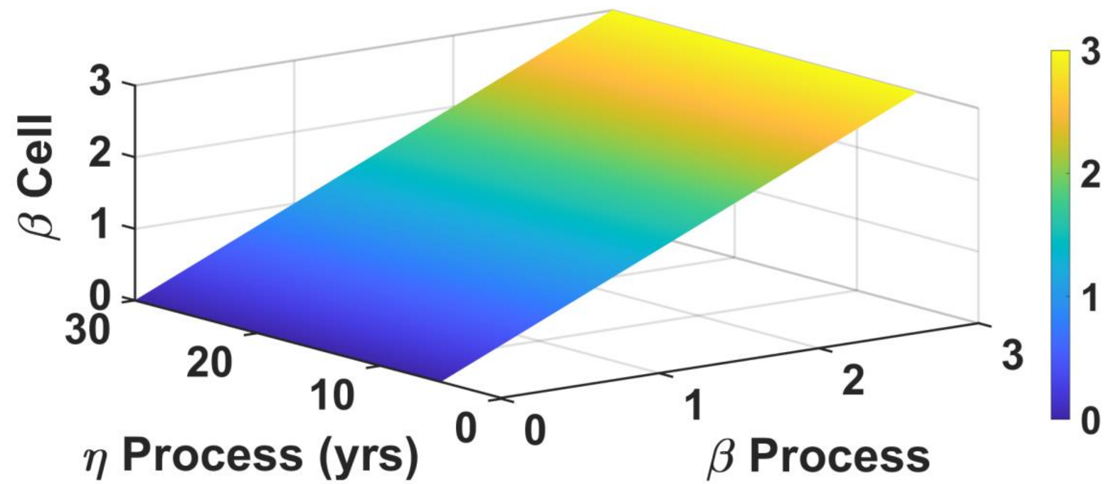


(b)

Figure 30 – SRAM failure samples due to FEOL TDDB for (a) a sample size of 94,000 with extracted FEOL TDDB failure parameters, η= 19.966 and β=1.119, and (b) a sample size of 100 out of 94,000 SRAM cells with extracted FEOL TDDB failure parameters of η=39.66 and β=0.99.

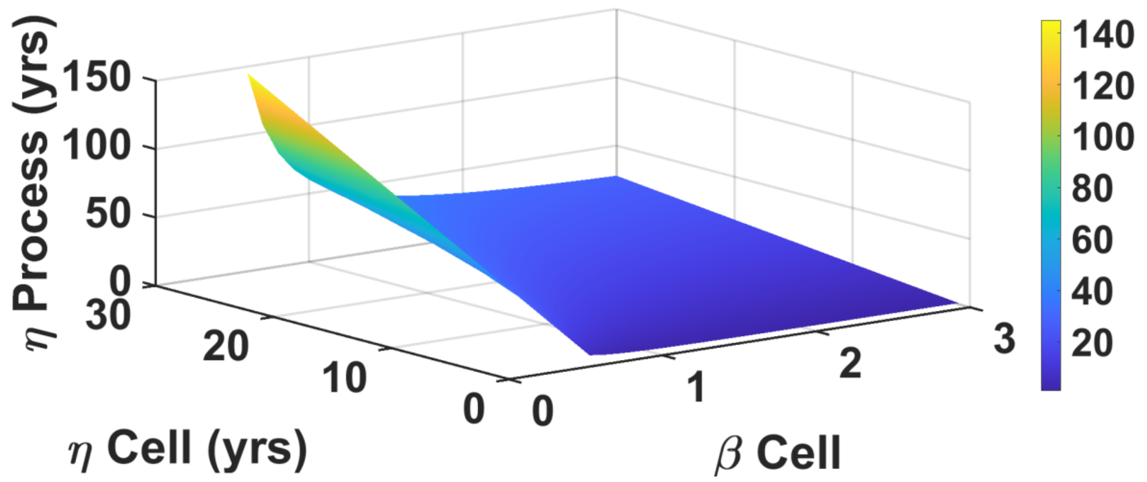


(a)

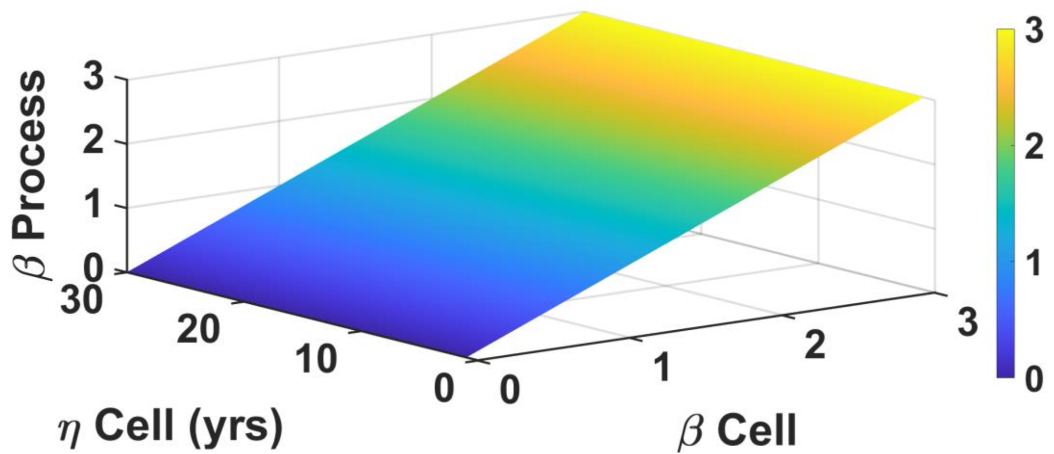


(b)

Figure 31 – Mapping between process-level Weibull parameters (η and β) and SRAM cell Weibull parameters for (a) η and (b) β for FEOL TDDB.



(a)



(b)

Figure 32 – Inverse mapping between process-level Weibull parameters (η and β) and SRAM cell Weibull parameters for (a) η and (b) β for FEOL TDDB.

1.8.2 Error Analysis

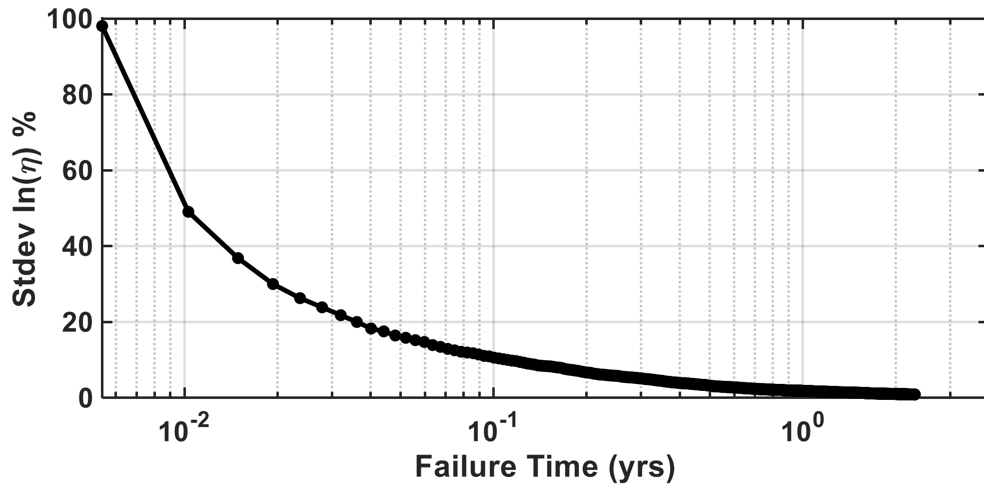
The errors in extracting the parameters are composed of random and systematic errors. The random error occurs from the limited sample size used to extract the parameters and properties of the map in Figure 32. The systematic error occurs due to variations in temperature, supply voltage, process parameters and use scenario. Here, the die-to-die variation in channel length is the primary source of process parameter variations. The use scenario depends on the duty cycle of the SRAM.

1.8.2.1 Random Error

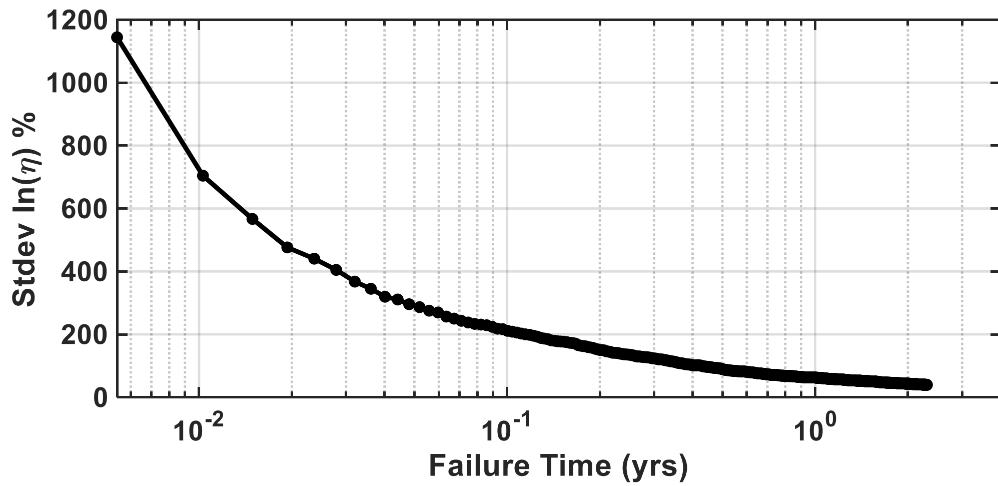
By using model parameters and Equation (29), the expected number of samples as a function of time for an SRAM with 94k cells can be calculated. Figure 33 shows the relative standard deviation (standard deviation/mean) for ln(η) as a function of time, and the result for β is similar. Figure 33 (a) is the relative error in estimating SRAM parameters, which were mapped to device model parameters with the functions in Figure 32. The errors in estimating the parameters of the SRAM can be found by:

$$\sigma^2(\ln(\eta_{device})) = \left(\frac{\partial(\ln(\eta_{device}))}{\partial(\ln(\eta_{SRAM}))}\right)^2 \sigma^2(\ln(\eta_{SRAM})) + \left(\frac{\partial(\ln(\eta_{device}))}{\partial(\beta_{SRAM})}\right)^2 \sigma^2(\beta_{SRAM}) \quad (32)$$

The standard deviation errors for the device are larger than those of the SRAM, because the mapping from the cell to process-level parameters introduces large sensitivities. For the SRAM cell, a 30% error is observed at 0.019 years, a 20% error is seen at 0.036 years, and a 10% error is found at 0.111 yrs. For the device in Figure 33 (b), the standard deviation falls to 30% in 4.7 years.



(a)



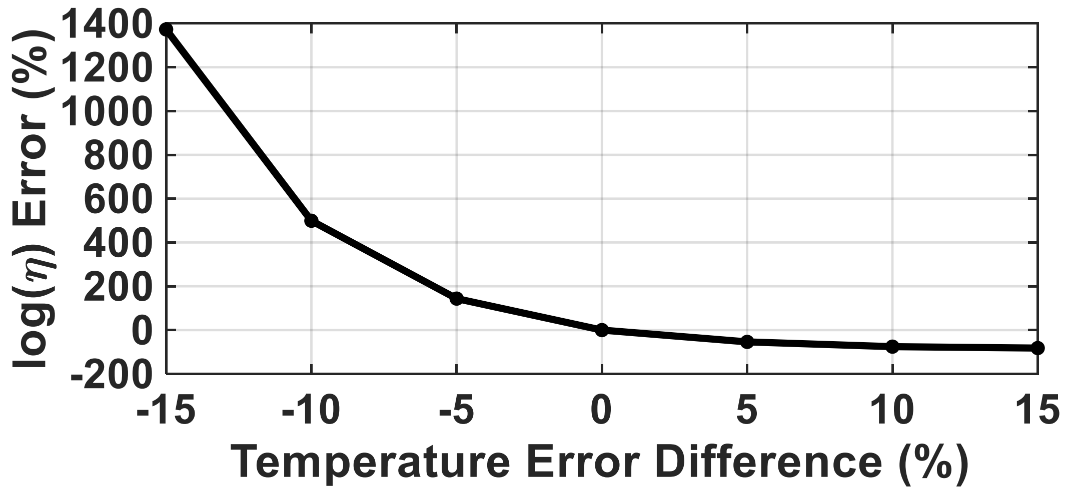
(b)

Figure 33 – Standard deviation error of the extraction of ln(η) as a function of time for the (a) SRAM with η=20 yrs, β=1.12 and (b) a single device.

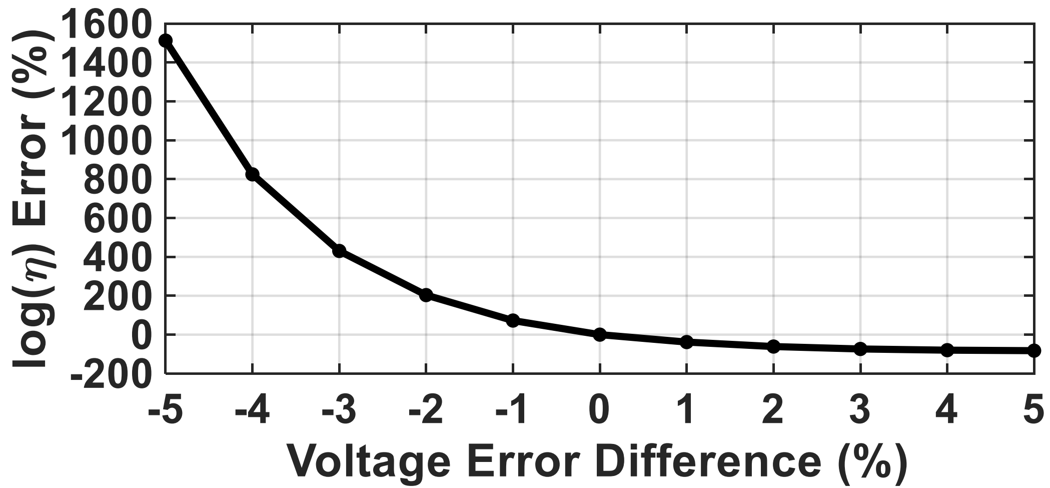
1.8.2.2 Systematic Errors

The percent changes in the SRAM characteristic lifetime errors due to the percent changes in operating temperature and voltage are shown in Figure 34. When the operating temperature overshoots by 15% or voltage overshoots by 5%, the characteristic lifetime errors drop 82%. However, when the operating temperature undershoots by 15%, the SRAM characteristic lifetime errors can increase by 1373%. Similarly, when the operating voltage undershoots by 5%, the SRAM characteristic lifetime errors can increase by 1512%. This signals that undershooting the operating conditions has a larger effect on changing the SRAM lifetime compared to overshooting, which can be as large as 16.7 times larger for temperature with a 15% error in operating conditions, and 18.44 times larger for voltage with a 5% error in operating conditions, respectively.

These errors translate into systematic errors in the estimation of device wearout parameters. The systematic errors due to shifts in temperature and voltage are very large, especially for shifts towards lower temperatures and voltages. For positive shifts in voltage and temperature, the systematic errors are not as large.



(a)



(b)

Figure 34 – Percent changes in errors in device characteristic lifetime estimation from variations in (a) temperature and (b) voltage. Voltage error differences above 5% causes the SRAM to fail upon startup and below -5% causes the SRAM to have essentially infinite characteristic lifetimes (e.g. above 300 years).

To see how the percent changes in characteristic lifetime due to variations in systematic errors translate into actual errors in device model parameters, the sensitivity of device model parameters to temperature and voltage can be computed as follows:

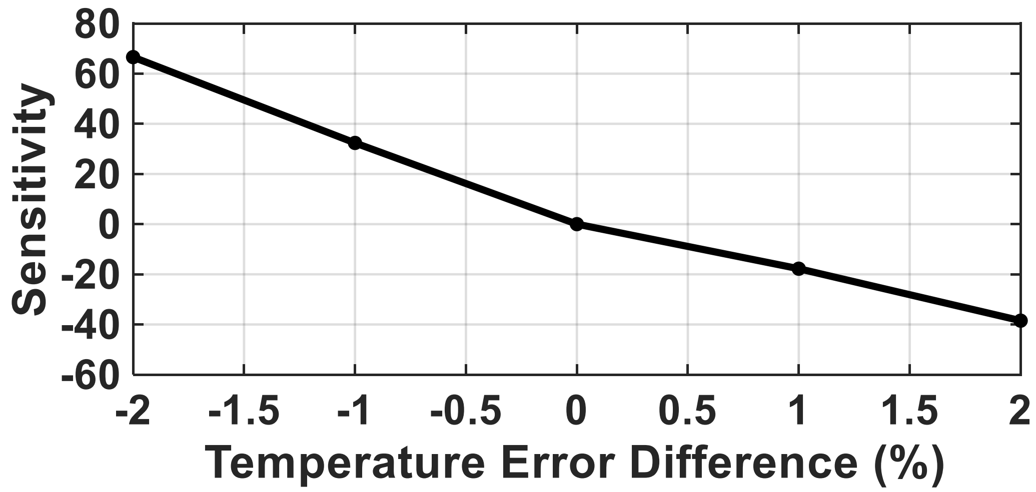
$$\frac{\Delta ln\,(\eta_{device})}{ln\,(\eta_{device})} = \frac{\partial(ln\,(\eta_{device}))/ln\,(\eta_{device})}{\partial(ln\,(\eta_{SRAM}))/ln\,(\eta_{SRAM})} \cdot \frac{\partial(ln\,(\eta_{SRAM}))/ln\,(\eta_{SRAM})}{\partial(T)/T} + \frac{\partial(ln\,(\eta_{device}))/ln\,(\beta_{device})}{\partial(\beta_{SRAM})/\beta_{SRAM}} \cdot \frac{\partial(\beta_{SRAM})/\beta_{SRAM}}{\partial(T)/T} \tag{33}$$

and

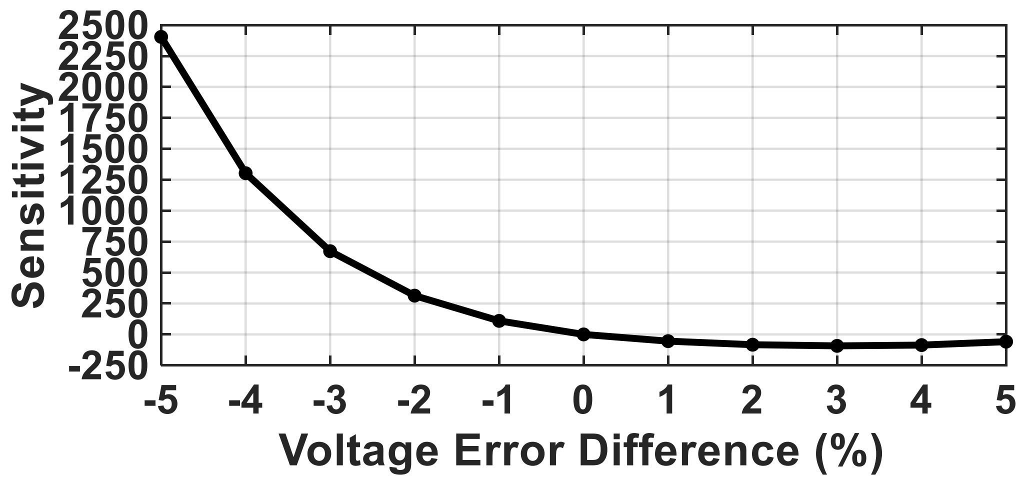
$$\frac{\Delta \ln\,(\eta_{device})}{\ln\,(\eta_{device})} = \frac{\partial(\ln\,(\eta_{device}))/\ln\,(\eta_{device})}{\partial(\ln\,(\eta_{SRAM}))/\ln\,(\eta_{SRAM})} \cdot \frac{\partial(\ln\,(\eta_{SRAM}))/\ln\,(\eta_{SRAM})}{\partial(V)/V} + \frac{\partial(\ln\,(\eta_{device}))/\ln\,(\eta_{device})}{\partial(\beta_{SRAM})/\beta_{SRAM}} \cdot \frac{\partial(\beta_{SRAM})/\beta_{SRAM}}{\partial(V)/V} \tag{34}$$

There are various sensors for monitoring temperature and voltage [93-97], which are widely used and embedded in system-on-chips (SoCs). These sensors are used to slow down operations when the temperature is too high to prevent overheating. A typical limit for temperature is 85°C, which is the limit for Raspberry Pi SOCs[98]. For a 45nm process, the accuracy for a temperature and voltage sensor is 4.13 °C and 10.67 mV, respectively in the range from 0.91V~1.09V and 0 °C~120 °C [94]. Thus, temperature variations can be detected within 2%.

The sensitivity of extraction of wearout parameters to variation in temperature and voltage is shown in Figure 35. When the temperature or voltage variations are positive, the random errors dominate the temperature/voltage variations, and vise versa. The positive temperature/voltage changes make the samples fail in a very short time, so the random



(a)



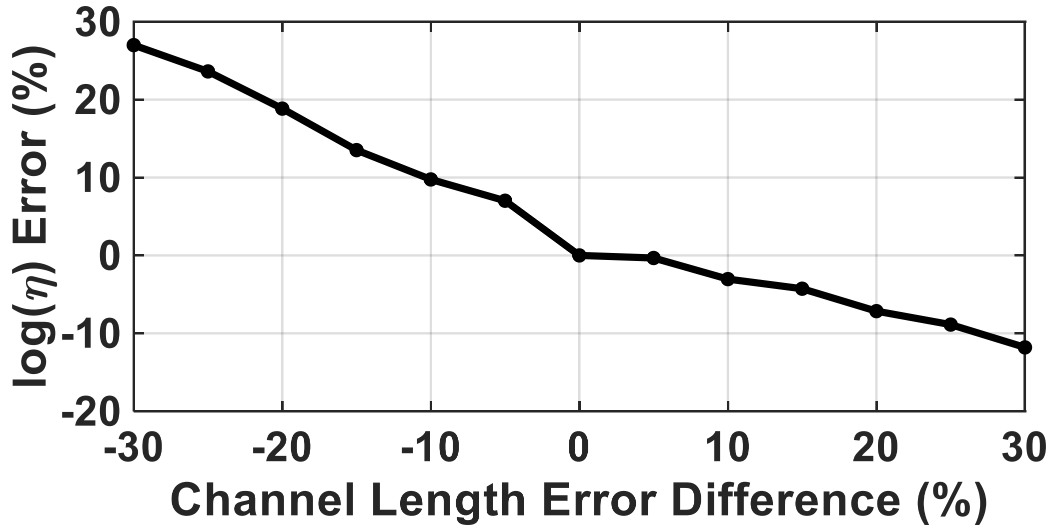
(b)

Figure 35 – Sensitivity of the extraction of ln (η_{device}) to changes in (a) temperature and (b) voltage.

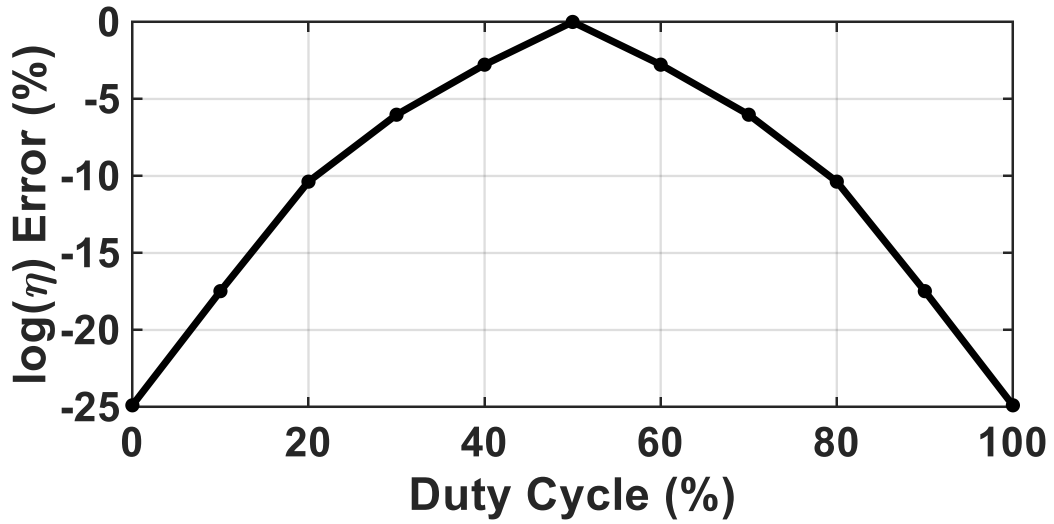
effects from sample size is hard to observe for these cases. Therefore, process monitors are needed to make sure the temperature and voltage do not have negative changes, which makes the errors increase greatly.

Besides the effects of the environmental parameters, such as temperature and voltage, other systematic variations also need to be considered. The process parameters and duty cycles may also be causes of systematic variations. As can be seen from Figure 36, errors in process parameters and duty cycle cause smaller changes in errors in the lifetime. From Figure 37, these errors cause smaller errors in the extracted parameters than voltage and temperature. The errors in process parameters and duty cycle are also smaller than random variations. If cost and space are an issue, process and duty cycle monitors can be excluded, because they are not as dominant.

Overall, systematic errors are larger than random errors when extracting device wearout parameters. Of the systematic errors, changes in supply voltage and temperature produce the largest errors. All four conditions should be monitored with sensors during operation to update the models accordingly. With appropriate sensors of operating conditions, the SRAM can be used to estimate wearout model parameters for individual chips using data from operation.

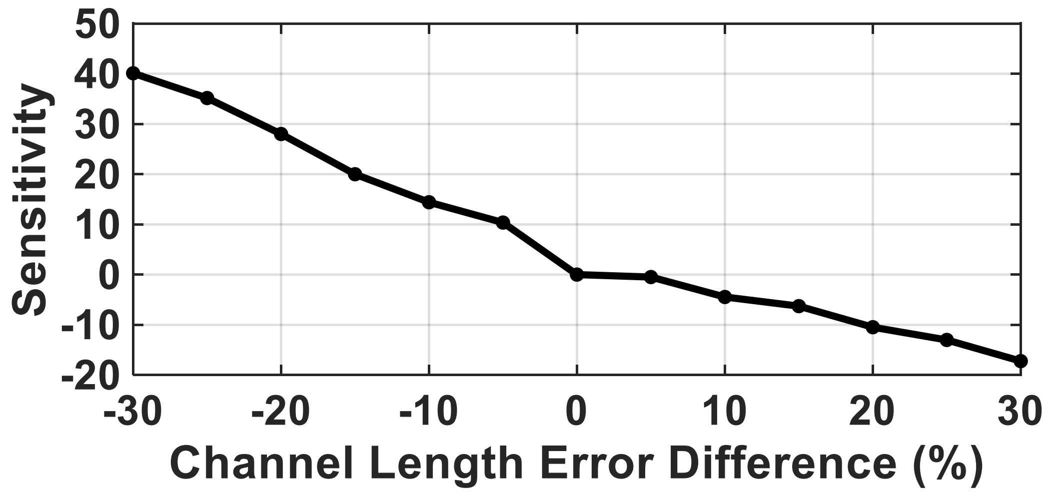


(a)

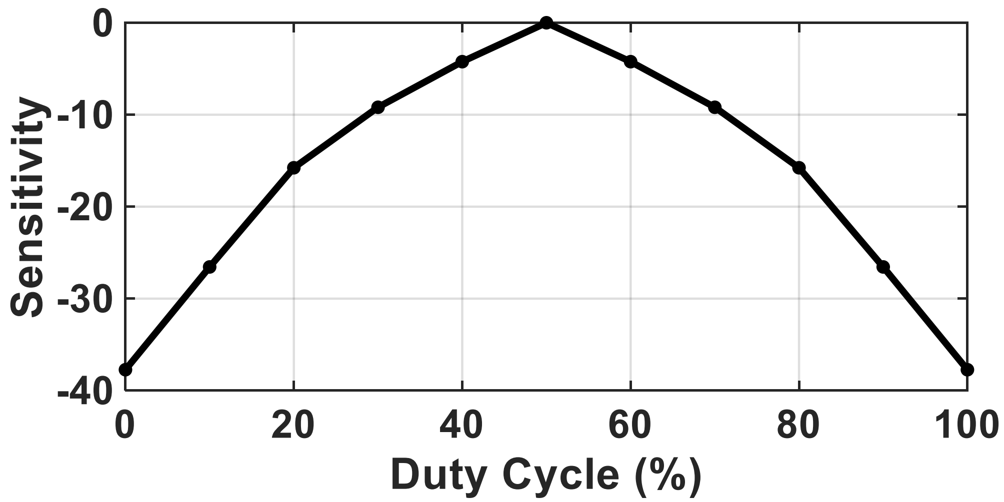


(b)

Figure 36 – Percent changes in errors for characteristic lifetime from variations in (a) channel length and (b) duty cycle (error calibrated to a duty cycle of 50%).



(a)



(b)

Figure 37 – Sensitivity resulting from changes in (a) channel length and (b) duty cycle.

ACCELERATED TESTING

To evaluate product lifetime, accelerated life tests are performed to collect failure data by applying high stresses to samples, such as high voltages and temperatures, to accelerate the normal degradation rate [99]. The collected failure information is then extrapolated to predict lifetime, as shown in Figure 38. Accelerated testing are utilized when there are measures that relate system health to function or operation [100], which can be used to understand the stress and dependence of failure mechanisms [101]. Health is the extent of deviation or degradation from its expected typical operating performance [102]. The information is then used to make improvements early in the design cycle through better design rules or materials selection criteria, as well as process and manufacturing control [103].

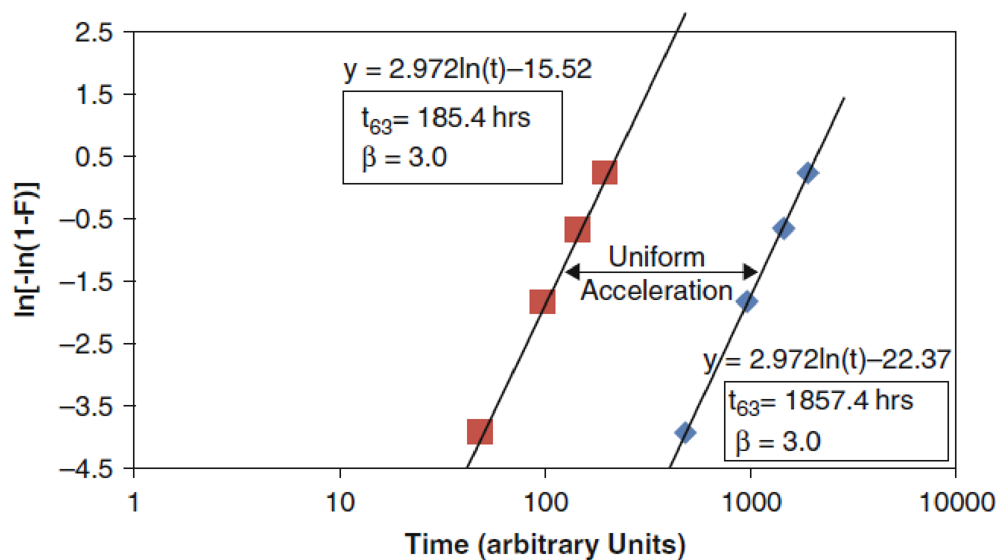


Figure 38 – Example of acceleration of failure distribution [12].

However, this process produces errors in lifetime estimation, leading to large costs due to inaccuracies. The effects of various parameters and stresses for 14nm FinFET ring oscillator circuits are investigated in order to understand how to select the correct testing criteria for accelerated testing. The study focuses on FEOL TDDB, because it is one of the most important front-end wearout mechanism in circuits.

1.9 Failure Probability vs. Test Time

The effects of testing time on the characteristic lifetime for FEOL TDDB of a 1001-stage ring oscillator for various temperature and voltage conditions are shown in Figure 39. The characteristic lifetime, which is the time when 63% fail, shifts towards higher temperatures and voltages as the testing time is decreased, because the degradation time is shorter for the 1001-stage ring oscillator. Since 63% failure corresponds to the characteristic lifetime, the contours indicate the test conditions where 63% fail prior to the end of testing.

When looking at the characteristic lifetime comparison in Figure 39 for 0.85 V, there is a 114 ^{o}C difference in temperature when changing from a two hour testing period to a two week testing period, but only a 28 ^{o}C difference when changing from a two week to a two month testing period. Similarly, the change in temperature from a testing time of two months to six months is about the same as the change from two weeks to two months. Since the change in temperature is much more significant when going from a testing time of two hours to two weeks, the two week testing time is a better choice if the testing conditions need to be run at lower temperatures because of cost.

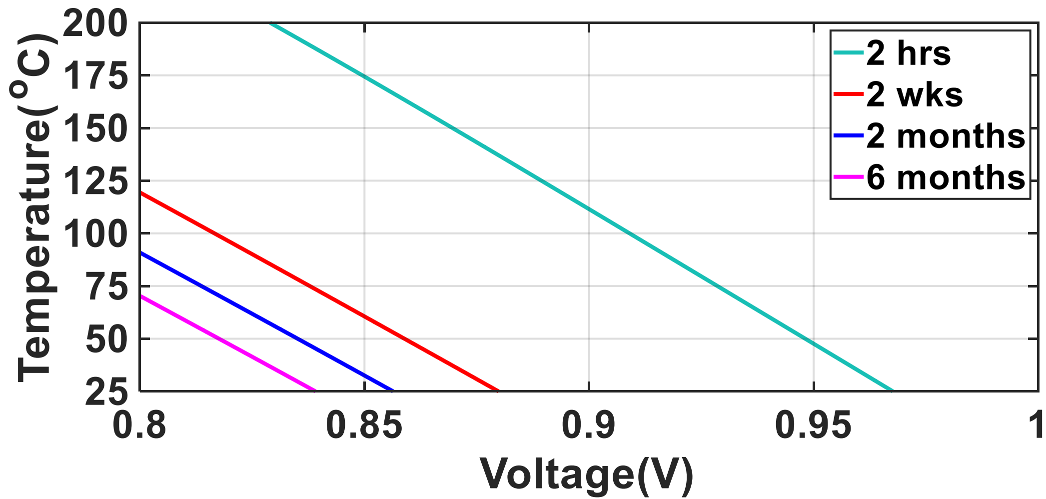


Figure 39 – Contour plot of the 63% failure probability (characteristic lifetime) as a function of temperature and voltage acceleration, for accelerated testing of the 1001-stage ring oscillator, with testing times ranging from two hours to six months.

For the 101-stage and 11-stage ring oscillators, the failure probability over various voltages and temperatures has a similar trend with the 1001-stage ring oscillator. The changes in temperature for different testing times also have similar results for the 101-stage and 11-stage ring oscillators. Therefore, if the temperature conditions need to be lower for the 101-stage and 11-stage ring oscillator, the two week testing period is the most suitable in terms of cost. Also note that smaller ring oscillators require more acceleration.

1.10 Sample Size vs. Test Time

The sample sizes required to produce at least one failed sample with 95% confidence for test times of two hours, two weeks, two months, and six months for 1001-stage and 11-stage ring oscillators for various voltages and temperatures are shown in Figure 40 to Figure 43. The results for the 101-stage ring oscillator is similar to the 1001-stage and 11-stage ring oscillators. When comparing the failure probabilities to the sample sizes, the high voltage and temperature regions only require a sample size of 1 to detect at least one

failure due to the failure probability being 100%. However, as the temperatures and voltages are decreased, the increase in sample size needed for a 95% confidence level increases by three orders of magnitude for a two hour testing time, whereas the sample size increase is only one order of magnitude when the testing time is longer at six months.

The change in temperature conditions at fixed voltages for the same sample size for each type of ring oscillator for various testing times follows the same trend as the characteristic lifetime. For both ring oscillators, there is a significant difference in the voltage for the required sample size at a 95% confidence level when changing from a two hour to a two week testing period, but a much smaller difference when changing from a two week to a two month testing period, as well as from a two month to a six month testing period. Therefore, the two week testing period is also optimal when considering the required sample size to produce a failed sample for a fixed confidence level and the voltage requirement if a lower temperature test is required.

Also, for longer testing times, such as six months, the 11-stage ring oscillator requires at least an order of magnitude more samples compared to the 101-stage ring oscillator, and a similar trend is seen when comparing the 1001-stage ring oscillator to the 101-stage ring oscillator. When comparing the changes in a sample size to produce four failures for all ring oscillators, the corresponding voltage and temperature conditions shift higher as the number of stages decreases.

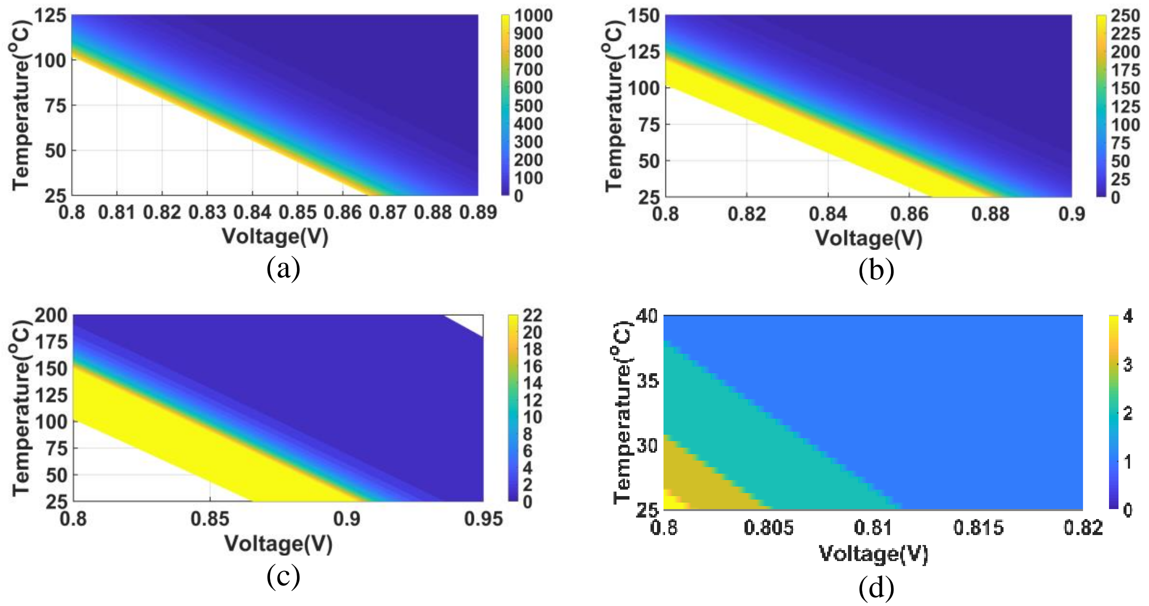


Figure 40 – Sample size needed produce a single failure with 95% confidence as a function of voltage and temperature for a 1001-stage ring oscillator for a testing time of (a) two hours (b) two weeks (c) two months and (d) six months.

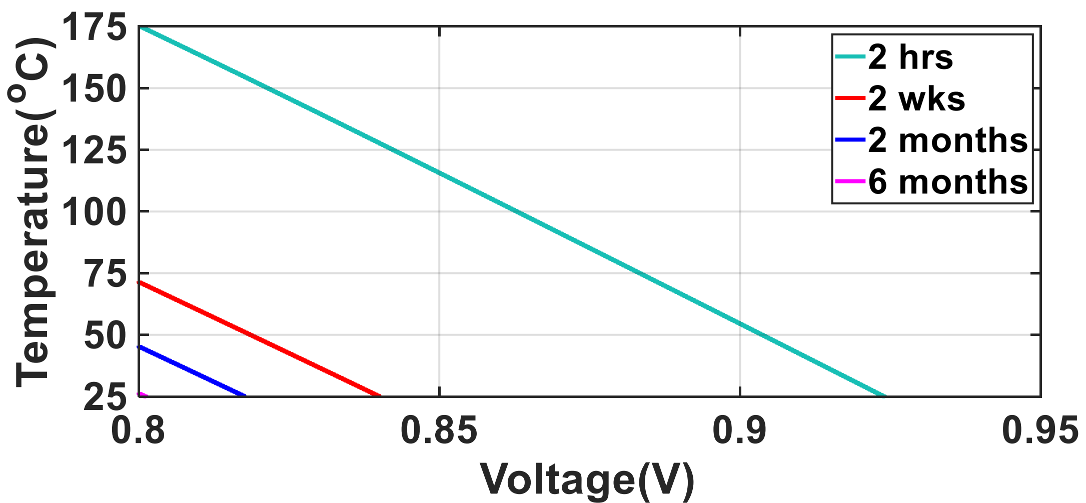


Figure 41 – Contour plot for sample size of four to produce at least 1 failure with 95% confidence as a function of voltage and temperature for a 1001-stage ring oscillator with testing time ranging from two hours to six months.

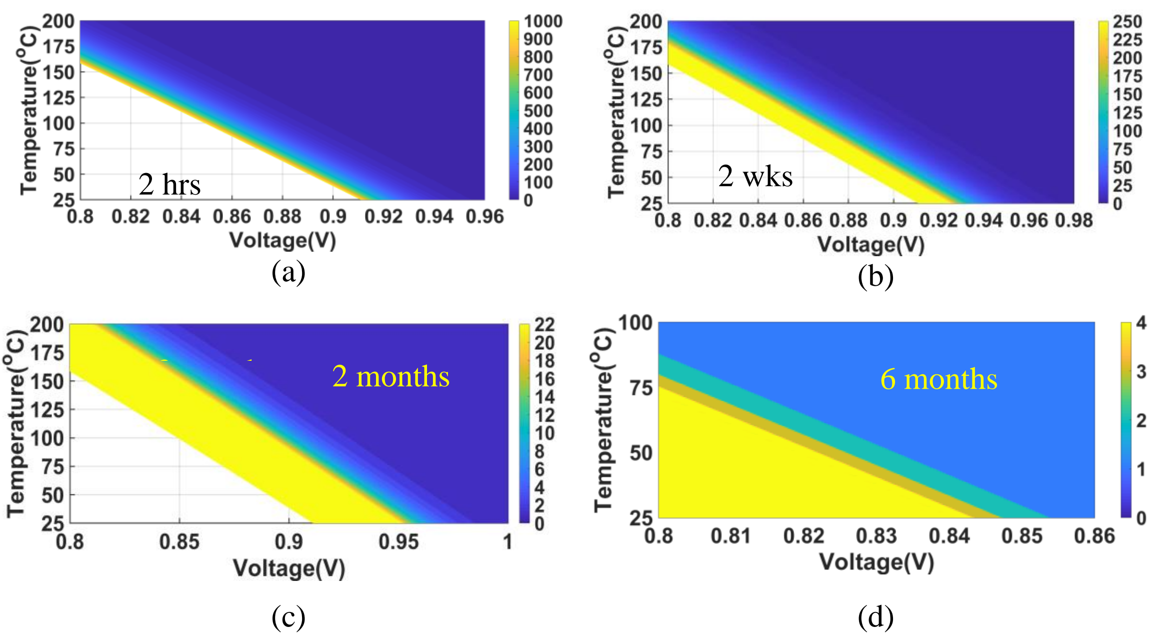


Figure 42 – Sample size needed produce a single failure with 95% confidence as a function of voltage and temperature for a 11-stage ring oscillator for a testing time of (a) two hours (b) two weeks (c) two months and (d) six months.

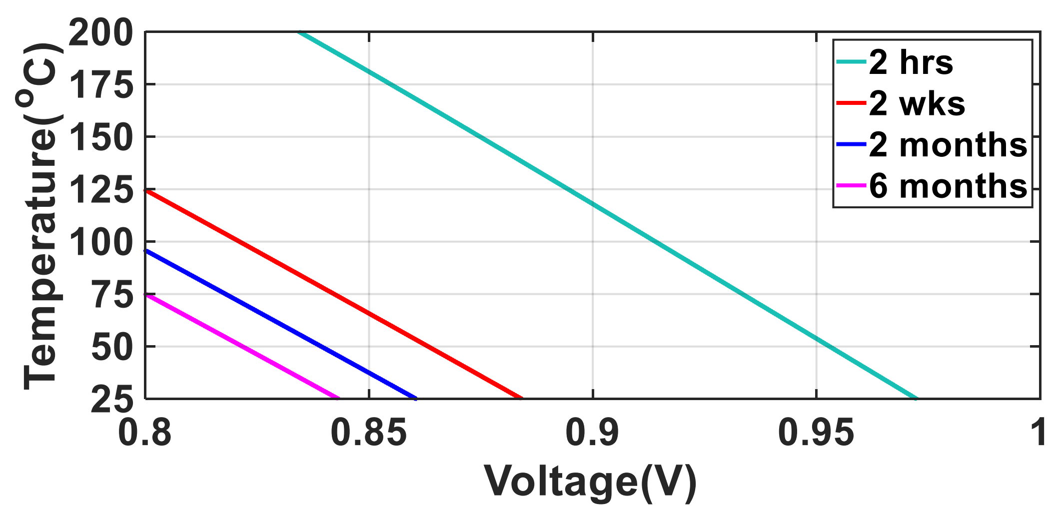


Figure 43 – Contour plot for sample size of four to produce at least 1 failure with 95% confidence as a function of voltage and temperature for a 11-stage ring oscillator with testing time ranging from two hours to six months.

1.11 Effects of the Number of Stages

1.11.1.1 Number of Stages vs. Characteristic Failure Lifetime

The characteristic lifetime, which occurs when there is a failure probability of 63%, shifts to lower temperatures and voltages as the number of stages in the ring oscillators increases, as shown in Figure 44. This is due to the increase in area as the number of stages is increased, which allows for more area under stress. When the testing time is increased, the temperature and voltage required for 63% failure probability is also lowered for each type of ring oscillator.

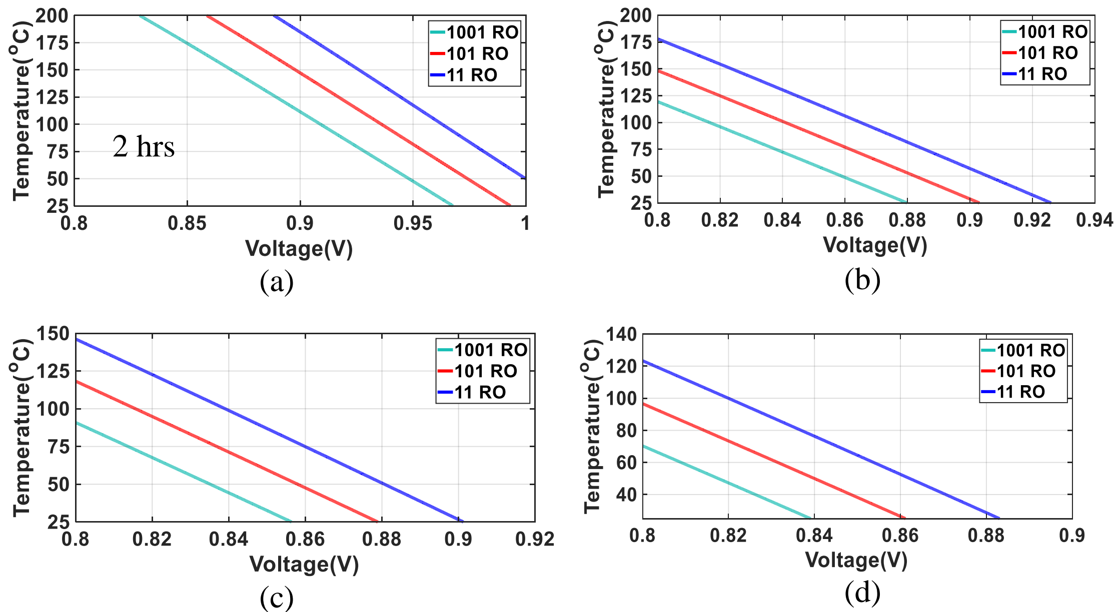


Figure 44 – Comparison of the contour plots for the number of stages and 63% failure probability (characteristic lifetime) as a function of voltage and temperature for (a) 2 hours (b) 2 weeks (c) 2 months and (d) 6 months.

1.11.1.2 Number of Stages vs the Minimum Required Sample Size

A comparison of number of stages and sample size for the three different types of ring oscillators over different voltage and temperature regions is shown in Figure 45. The sample size needed for a 95% confidence level to detect at least one failure can decrease by an order of magnitude as the testing time increases from two hours to six months for all three different types of ring oscillators.

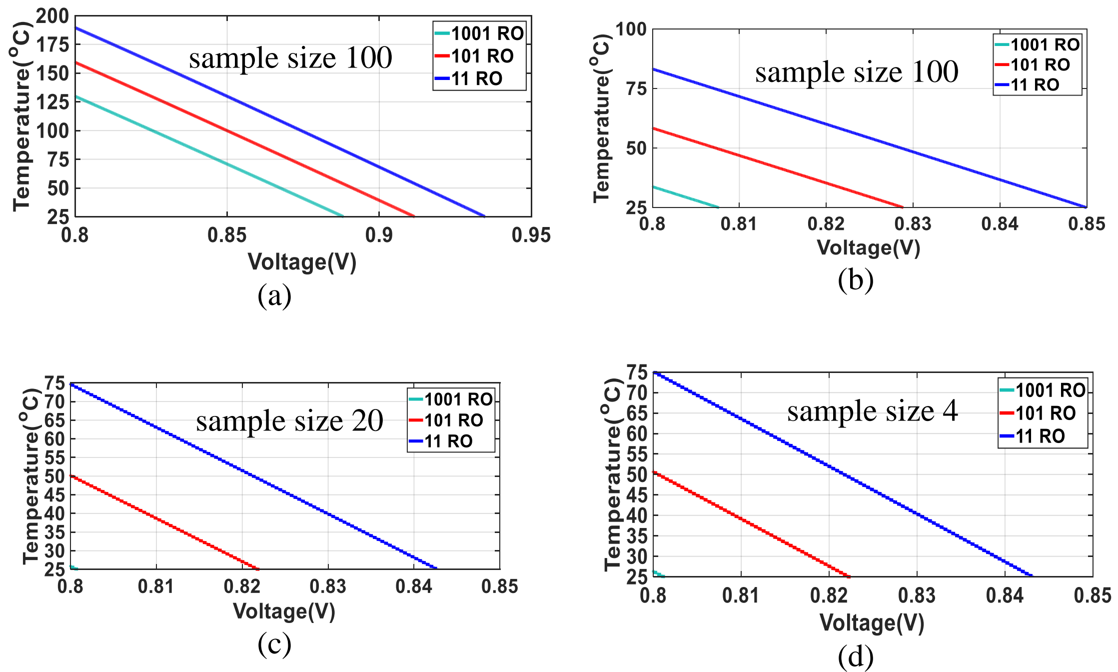


Figure 45 – Comparison of the number of stages and the test conditions needed to detect at least one failure for various sample sizes as a function of voltage and temperature for test times of (a) two hours (b) two weeks (c) two months and (d) six months.

1.12 Lifetime Estimation Method for Use Conditions

The extraction of wearout distribution parameters at accelerated test conditions depends on the number of available failed samples. The number of failure samples can be viewed as a function of acceleration and increases as the circuit is stressed more. Based on the available sample size, the variance in parameter extraction is computed at all possible test conditions, σ_i^2. This study focuses on the variance of ln(η) since any errors in extracting β are similarly reduced.

To find the test points that minimize the variance at use conditions, a weighted regression equation is used to estimate the lifetime at use conditions based on the accelerated test points. Weighted regression is used, because the errors, or variance, are different at each test condition [104]. The deviation between the observed and expected values of y_i is multiplied by a weight, w_i, inversely proportional to the variance at that testing condition for n samples, σ_i^2:

$$w_i = \frac{1}{\sigma_i^2} \tag{35}$$

where i=1, …b are the test points. The lifetime at use conditions is computed by regression by solving

$$(X^T W X)\beta = X^T W y \tag{36}$$

where W is a diagonal matrix of elements w_i, and β is a vector of regression coefficients. The linear regression equation for FEOL TDDB is approximated as

$$\ln(\eta) = a_0 + a_1 \ln(V) + a_2 \left(\frac{1}{T}\right). \quad (37)$$

To find the optimal testing conditions, we minimize the variance at use conditions:

$$(\hat{y}) = x_0^T (X^T W X)^{-1} x_0 s^2 \quad (38)$$

where x0 are at use conditions and $s^2 = 1$ (variance of the residual mean square which follows a normal distribution).

1.13 Testing Conditions

1.13.1 Computing the Error for the Accelerated Test Conditions

To study the optimal test conditions, three types of ring oscillators with different numbers of stages, the 1001-stage, 101-stage, and 11-stage ring oscillators were examined. Furthermore, to investigate the effects of testing time, each type of ring oscillator was also studied for testing times of two hours, two weeks, two months, and six months.

The acceleration maps for the circuits were obtained by finding the lifetime distribution at each voltage and temperature. An available number of tested circuits was first assumed, with a specified maximum test time. A minimum lifetime is also specified for samples that failed too fast (smaller than one minute), because the time to failure could not be measured. These are declared to be "dead on arrival" and dropped from the sample. At each voltage and temperature, the probability between the "dead on arrival" time and the maximum test time is determined. The total available number of tested circuits is multiplied by this probability to determine a failed sample size at each voltage and temperature.

The errors (variance in the parameter estimates) in characteristic lifetime, σ_i, as a function of sample size can be calculated using Monte Carlo simulation with the generalized maximum likelihood method for estimation [92]. As shown in Figure 46, the standard deviation for estimating ln(η) decreases with the increase in sample size, signifying that a larger sample size increases the accuracy in estimating ln(η).

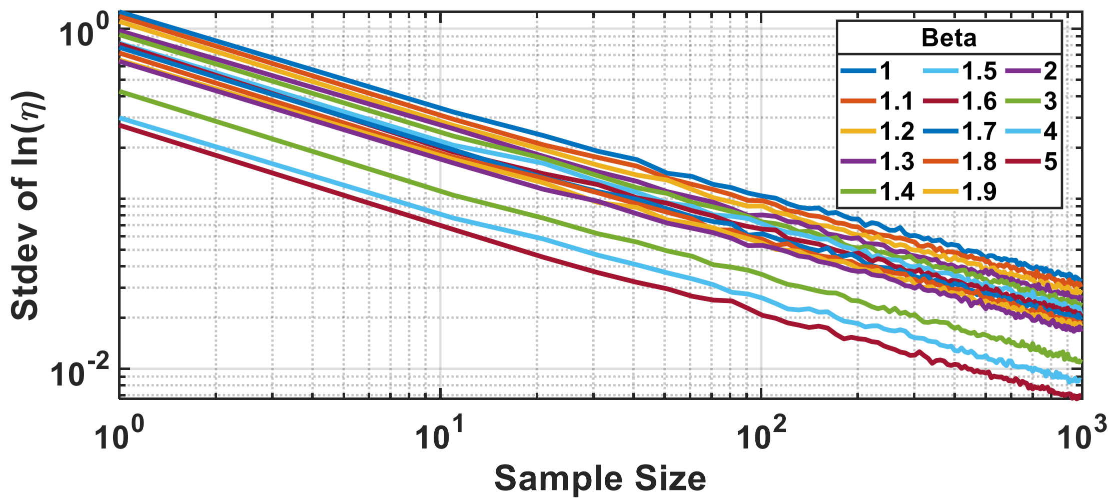


Figure 46 – Standard deviations of estimates of the errors in ln(η) using Monte Carlo simulations.

1.13.2 Test Plans

Experiments were designed by varying the distance between two different test points in voltage or temperature. The minimal variance of ln(η) at use conditions will also be referred to as the tolerance level. The variance at use conditions is minimized with two-point test plans with duplicates and voltage acceleration only for the ring oscillators. For all acceleration test domains, either the two- or three-point test plans with duplicates minimize errors at use conditions for ring oscillators. After considering all possible two- and three-point test plans, it was found that the two-point test plans were best for the ring

oscillators. Therefore, two test points with duplicates of (V_L, T_0) and (V_H, T_0) were used to find the minimum total variance as [104]:

$$\sigma^2 = \frac{(\ln(V_L) - \ln(V_{nom}))^2 \sigma_H^2}{(\ln(V_H) - \ln(V_L))^2} + \frac{(\ln(V_H) - \ln(V_{nom}))^2 \sigma_L^2}{(\ln(V_H) - \ln(V_L))^2} \tag{39}$$

where V_L is the lower value of the accelerated voltage, V_H is the higher value of the accelerated voltage, V_{nom}=0.8 V is the voltage at use condition, and T_0=25℃ is the temperature at use condition. σ_H^2 and σ_L^2 are the variances for 100 samples at each of the test points V_H and V_L, respectively.

The resulting test plans are shown in Fig. 2. These optimal test plans assumed an available sample size of 2000. When the test time is shorter, more acceleration is needed to induce the same number of failures. The size of the ring oscillator does not affect the optimal test plans very much.

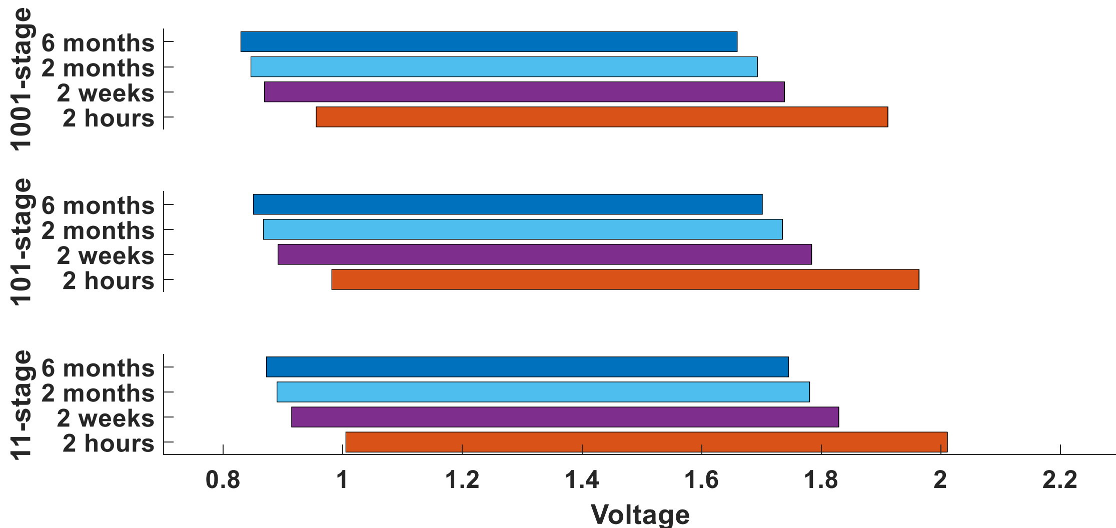


Figure 47 – Optimal testing points for 1001-stage, 101-stage and 11-stage ring oscillators for testing times of two hours, two weeks, two months and six months.

1.14 Error Reduction Through Sampling

Under ideal circumstances, an infinite number of samples for testing can eliminate all errors. However, the available sample size for testing is limited, so other methods must be used to practically achieve a balance between the sample size and error. To do this, first, the sample fractions for a fixed overall sample size are adjusted. Second, a reasonable tolerance level is set, and the overall sample size needed to achieve this tolerance is found.

1.14.1 Adjusting Sample Fractions for a Fixed Overall Sample Size

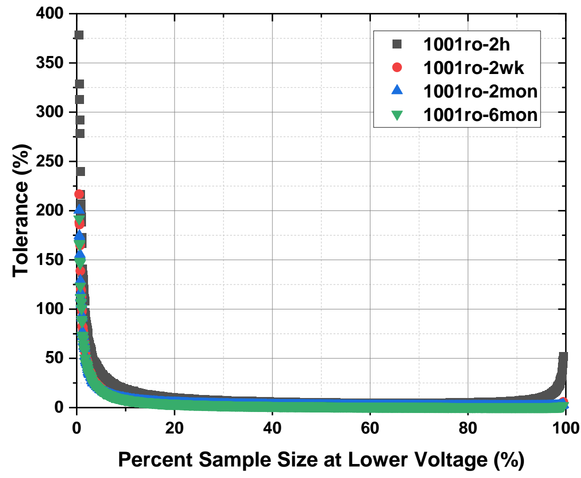
For a fixed overall total sample size of 2000 samples, the errors, also known as tolerance, depends on the percentage of samples at each of the test points and the testing time. As shown in Figure 48, the tolerance is about the same for a wide range of sampling fractions, so it is just necessary to avoid placing almost all samples at either the high- or low-test conditions. For a two-point voltage test plan, the tolerance for the 1001-stage ring oscillator reduces the most when changing from a testing time of two hours to two weeks for the same percentage of samples at the lower testing point. The results are similar for the other two ring oscillators.

The comparisons of the lowest tolerances for the 11-stage, 101-stage and 1001-stage ring oscillators with various testing times of two hours, two weeks, two months and six months, as well as the corresponding percentage of sample size at the lower and higher voltage points, are shown in Figure 49. For the same testing time, the lowest tolerance decreases with the increase in the number of stages and testing time. Because the lower acceleration voltage can be decreased as the testing time increases, more samples can be placed at the lower voltage point with the increase in testing time for all ring oscillators.

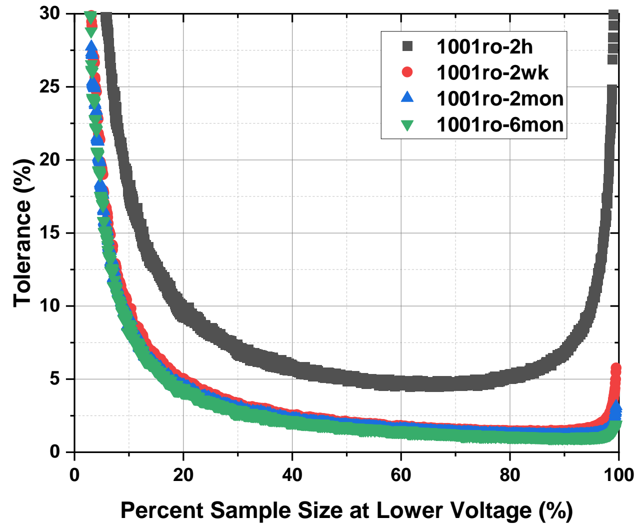
At the minimum tolerance, the V_L tolerance dominates that of V_H, since most of the samples are placed at the lower voltage point, meaning that changes in V_H are not as sensitive as V_L for various testing times. Also, the higher voltage point value has little sensitivity to testing time for the same number of stages. For both the higher and lower voltage points, the smaller ring oscillators require higher acceleration for the same testing times.

For all testing times and numbers of stages, the tolerance is lowest when most of the samples are placed at the lower voltage, since extrapolating from a lower voltage point gives smaller errors at use conditions. There still needs to be some samples placed at the higher voltage point in order to obtain the acceleration factor. However, an insufficient number of samples at either voltage point can result in the high tolerance values as seen in the left and right ends at Figure 50.

The minimum tolerance can be significantly reduced from 216.7% to as low as around 1.3% when the percentage of samples at the lower voltage point is changed from 0.5% to 83% for a testing time of two weeks for a 1001-stage ring oscillator. For the 1001-stage ring oscillator, the tolerance when 50% of the samples are placed at V_L is 2.1%. However, the difference when placing 83% of the samples at V_L instead of 50% causes only a minor change of 0.8%, which is close to negligible. The results are similar for the 101-stage and 1001-stage ring oscillators.

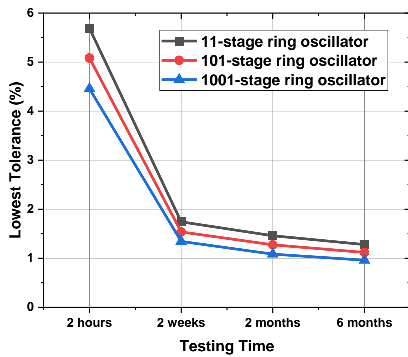


(a)

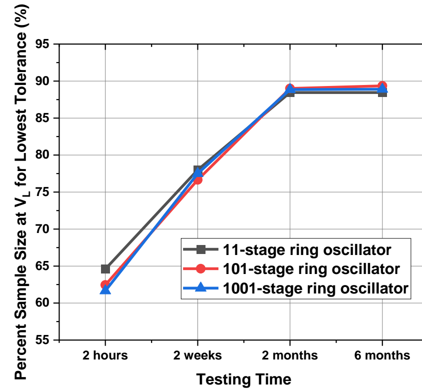


(b)

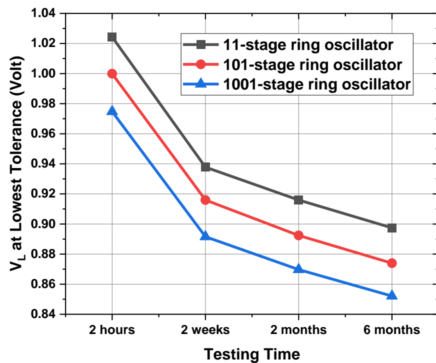
Figure 48 Relationship between percent sample size at the lower voltage and tolerance for the 1001-stage ring oscillator with overall sample size of 2000 and testing times of two hours, two weeks, two months and six months for (a) the overall tolerance range and (b) with a close up of the tolerance range below 30%.



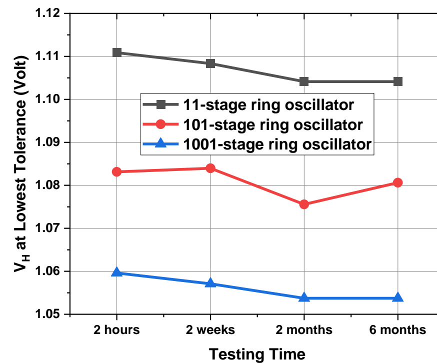
(a)



(b)



(c)



(d)

Figure 49 – For a sample size of 2000, a comparison of 11-stage, 101-stage and 1001-stage ring oscillators for a testing time of two hours, two weeks, two months and six months in terms of (a) the lowest tolerance, (b) the corresponding percentage of the sample at the lower voltage point to achieve the minimum tolerance, (c) the corresponding lower voltage point to achieve the minimum tolerance, and (d) the corresponding higher voltage point to achieve the minimum tolerance.

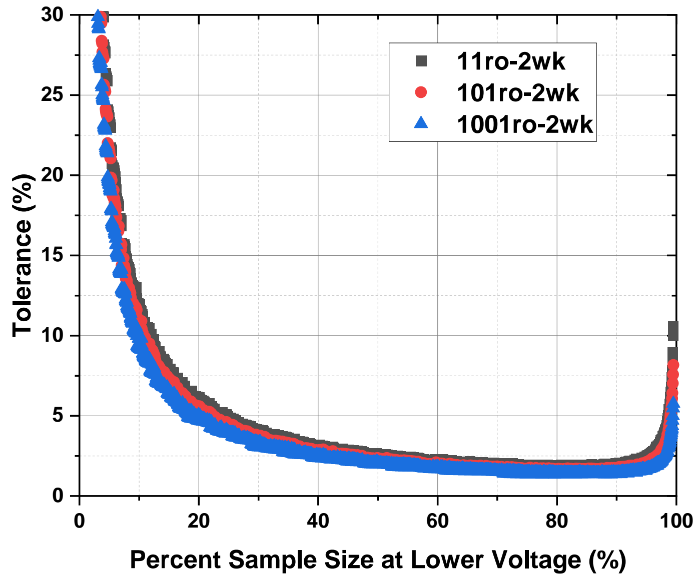


Figure 50 – Relationship between percent sample size at the lower voltage and the tolerance for the 11-stage, 101-stage and 1001-stage ring oscillators with a testing time of two weeks. (Tolerances above 30% are not shown for clarity.)

1.14.2 Minimum Sample Size to Achieve Fixed Tolerance

The minimum sample sizes necessary to achieve a 10%, 20% and 30% tolerance for the 1001-stage, 101-stage and 11-stage ring oscillators with different testing times are shown in Figure 51. For all ring oscillator sizes, the total number of samples decreases around three-fold when changing from a two hour to two week testing time period, but the sample size does not vary much when changing from two weeks to two months or from two months to six months. Also, at the 10% tolerance level for all ring oscillators, there is around an order of magnitude difference in the sample size when changing from the two hour to two week testing time frame. This also signifies that the two week testing period produces the best step-up in tolerance compared to the two- and six-month test times.

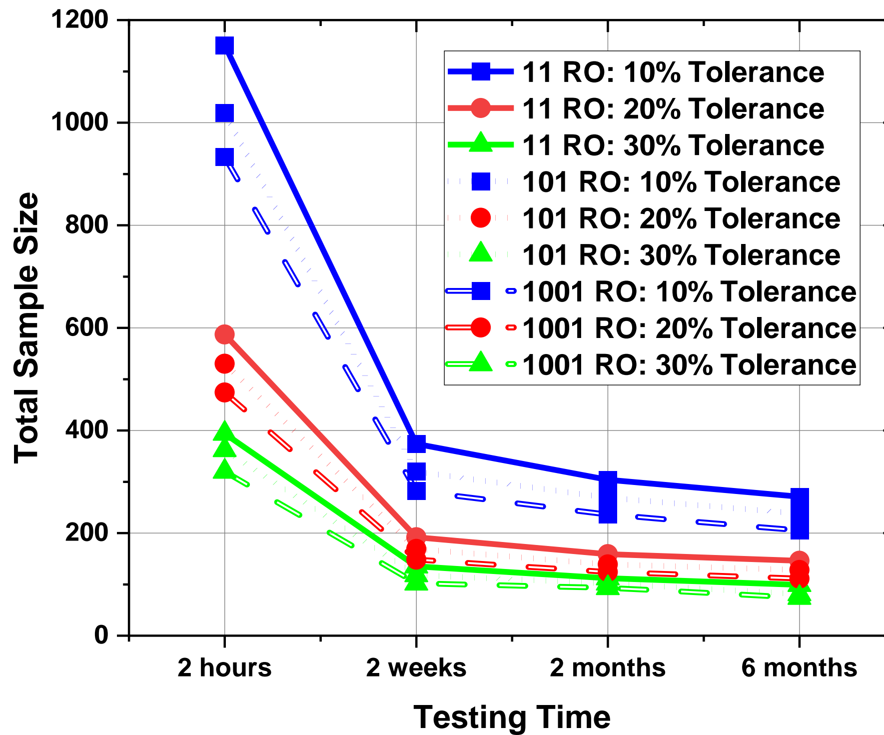


Figure 51 – Comparison between the total sample size and the testing time for 1001-stage, 101-stage and 11-stage ring oscillators to achieve a 10%, 20% and 30% tolerance.

However, if the test cost is dependent only on the cost of test time and sample size, two hours is the optimal for the four test times considered.

For the same number of stages and testing time, the required minimum sample sizes are around two times smaller when changing from a 10% to a 20% tolerance level. However, the difference is around 1.3 to 1.5 times smaller when changing from a 20% to a 30% tolerance level. Therefore, the trade-off between the increase in the total minimum sample size and tolerance level is larger as the accuracy increases.

In Figure 52, the minimum total sample sizes for achieving various tolerances for the two week testing time is shown. The trade-off between the increase in the total minimum sample size and tolerance level increases as the tolerance level is lowered for the same number of stages and testing time. For a 5% tolerance level, the total minimum sample sizes need to be above 500 for all numbers of stages. However, when the tolerance is changed to 30%, the required minimum total sample sizes can be reduced to around 100, and the 20% tolerance level can be achieved with sample sizes under 200. For tolerances lower than 20%, the change in sample size increases considerably higher. Decreasing the tolerance level further below 10% requires more than a 100 sample size increase for ring oscillators. For the same tolerance level, the smaller ring oscillators require a larger total minimum sample size to reach the same accuracy.

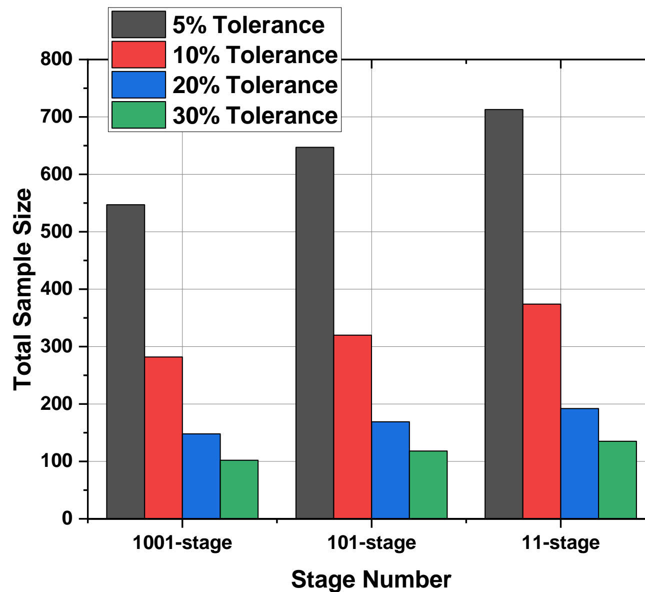


Figure 52 – Relationship between percent tolerance and number of stages for a testing time of two weeks.

1.14.3 Lowest Tolerance and Total Sample Size

The relationship between the lowest tolerance and total sample size for a two week testing period is compared for the 11-stage, 101-stage and 1001-stage ring oscillators in Figure 53. For all three ring oscillators, the lowest tolerance decreases as the total sample size increases, because the errors at both the V_L and V_H decrease as more sampling is available. For the same total sample size, the lowest tolerance is about the same for all three ring oscillators.

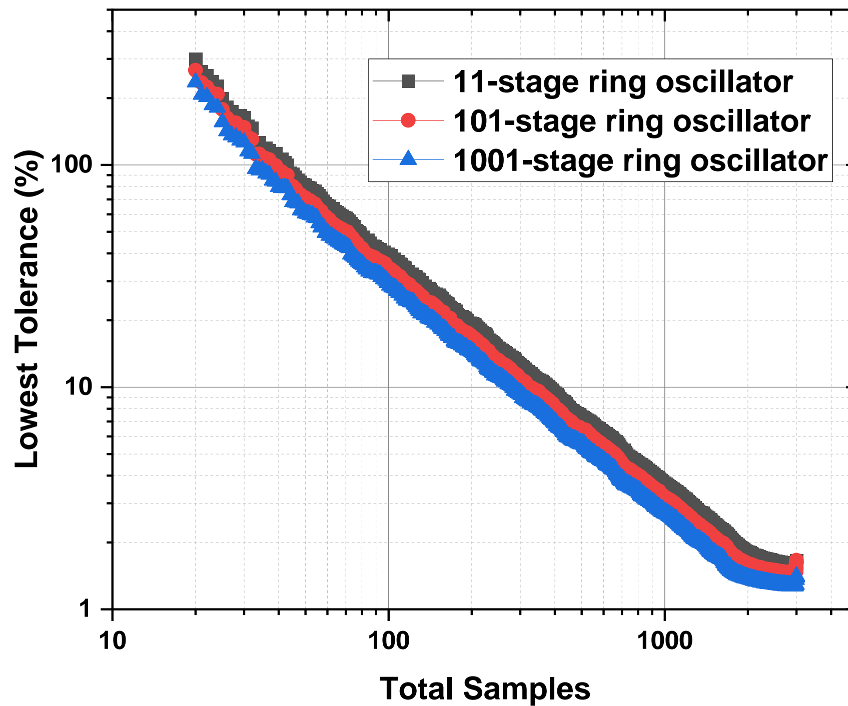


Figure 53 – Relationship between lowest tolerance and total sample size at a two week testing time for 11-stage ring oscillator, 101-stage ring oscillator, and 1001-stage ring oscillator.

If the silicon area and cost are the same, the 1000 11-stage will have the lowest tolerance level, followed by 100 101-stage and 10 1001-stage ring oscillators for the two week testing period, because the tolerance level will decrease by an order of magnitude for each comparison described, respectively. Similarly, 10000 11-stage ring oscillators will produce lower errors compared to 1000 101-stage ring oscillators, which is lower than 100 1001-stage ring oscillators. Therefore, it is best to opt for a larger number of smaller ring oscillators compared to a smaller number of large ring oscillators to have more accurate results if silicon area and cost have the same consideration However greater acceleration will be required for smaller ring oscillators.

1.15 Summary

The minimum errors for accelerated testing of ring oscillators can be achieved by using the two-point test plan. For a fixed sample size, the lowest tolerance occurs when most of the samples are placed at the lower voltage point for all ring oscillators with different numbers of stages and testing times. For a fixed tolerance, the lowest sample size also occurs when most of the samples are placed at VL for all ring oscillators with different numbers of stages and testing times.

Since large ring oscillators require more area, errors can be reduced if a large number of small ring oscillators are implemented, rather than a few large ones. Shorter test times minimize the tradeoff between accuracy and tester usage, indicating that testing a large sample for a shorter time is also more optimal. The optimal solution of testing larger numbers of small ring oscillators with shorter test times is limited by the required acceleration which can potentially distort results.

CONCLUSIONS

1.16 Summary

This thesis investigates the failure analysis and accelerated testing of circuits. The 14nm FinFET PDK developed by GlolbalFoundries/Samsung/IBM was used to investigate the circuits. SRAMs and ring oscillators were used as case studies.

A methodology was developed to separate competing wearout mechanisms and identify the samples belonging to their respective degradation modes using data analysis techniques for ring oscillators. In addition, the online monitoring of SRAMs for FEOL TDDB were studied to investigate the impact of systematic and random errors on extracting wearout parameters. It was found that systematic errors are dominant over random errors. Temperature and voltage variations produce the largest influences in sensitivity to device model parameters, and monitors should be implemented to screen these variables.

The experimental plans for accelerated testing were found for FEOL TDDB in ring oscillators for a 1001-stage, 101-stage and 11-stage ring oscillator for testing times of 2 hours, 2 weeks, 2 months and 6 month, with the two point test plans having the lowest errors. Errors were reduced further by investigating the effects on sampling for a fixed overall sample size, as well as finding the minimum total sample size to achieve a fixed tolerance. Placing the majority of the samples at the lower voltage point was found to have the lowest errors.

1.17 Future Work

There are several directions that may be taken to expand the research in this thesis. The methodology for separating competing wearout mechanisms can be implemented for online monitoring. The accelerated testing plans can be used to test FEOL TDDB degradation in ring oscillators circuits based on the TSMC 130nm technology node. Furthermore, the incorporation of FEOL TDDB breakdown can be used in statistical timing analysis to calculate the delay in order to design better circuits, by incorporating the reliability budget considerations to achieve higher performance.

REFERENCES

[1] J. Ahn, M. F. Lu, N. Navale, D. Graves, G. Refai-Ahmed, P. Yeh, and J. Chang, "Product-level reliability estimator with budget-based reliability management in 16nm technology," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2-6 April 2017 2017, pp. 3A-3.1-3A-3.6.

[2] K. Schuegraf, M. C. Abraham, A. Brand, M. Naik, and R. Thakur, "Semiconductor Logic Technology Innovation to Achieve Sub-10 nm Manufacturing," *IEEE Journal of the Electron Devices Society,* vol. 1, no. 3, pp. 66-75, 2013.

[3] R. Mariani, "An overview of autonomous vehicles safety," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 11-15 March 2018 2018, pp. 6A.1-1-6A.1-6.

[4] M. Duncan and P. Roche, "Paving the way towards autonomous driving — Tackling soft errors to security challenges," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2-6 April 2017 2017, pp. 2E-1.1-2E-1.8.

[5] D. Hemapriya, P. Viswanath, V. M. Mithra, S. Nagalakshmi, and G. Umarani, "Wearable medical devices — Design challenges and issues," in *2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT)*, 16-18 March 2017 2017, pp. 1-6.

[6] E. Landman, S. Cohen, N. Brousard, R. Gewirtzman, I. Weintrob, E. Fayne, Y. David, Y. Bonen, O. Niv, S. Tzroia, A. Burlak, and J. W. McPherson, "Degradation Monitoring — from a Vision to Reality," in *2019 IEEE International Reliability Physics Symposium (IRPS)*, 31 March-4 April 2019 2019, pp. 1-4.

[7] A. Strasser, P. Stelzer, C. Steger, and N. Druml, "Live State-of-Health Safety Monitoring for Safety-Critical Automotive Systems," in *2019 22nd Euromicro Conference on Digital System Design (DSD)*, 28-30 Aug. 2019 2019, pp. 102-107.

[8] E. A. Elsayed, *Reliability Engineering*. Hoboken, UNITED STATES: John Wiley & Sons, Incorporated, 2012.

[9] P. D. T. O. Connor, "Commentary: reliability-past, present, and future," *IEEE Transactions on Reliability,* vol. 49, no. 4, pp. 335-341, 2000.

[10] R. Radojcic and P. Giotta, "Semiconductor device reliability vs process quality," *Microelectronics Reliability,* vol. 32, no. 3, pp. 361-368, 1992/03/01/ 1992.

[11] J. W. McPherson, "Understanding the underlying degradation physics for proper time-to-failure distribution selection," in *2015 IEEE International Reliability Physics Symposium*, 19-23 April 2015 2015, pp. ER.1.1-ER.1.7.

[12] J. W. McPherson, *Reliability Physics and Engineering Time-To-Failure Modeling*, 3rd ed. 2019.. ed. Cham : Springer International Publishing : Imprint: Springer, 2019.

[13] M. Ieong, B. Doris, J. Kedzierski, K. Rim, and M. Yang, "Silicon Device Scaling to the Sub-10-nm Regime," *Science,* vol. 306, no. 5704, p. 2057, 2004.

[14] G. G. Shahidi, "Chip Power Scaling in Recent CMOS Technology Nodes," *IEEE Access,* vol. 7, pp. 851-856, 2019.

[15] M. T. Bohr and I. A. Young, "CMOS Scaling Trends and Beyond," *IEEE Micro,* vol. 37, no. 6, pp. 20-29, 2017.

[16] J. W. McPherson, "Scaling-induced reductions in CMOS reliability margins and the escalating need for increased design-in reliability efforts," in *Proceedings of the IEEE 2001. 2nd International Symposium on Quality Electronic Design*, 28-28 March 2001 2001, pp. 123-130.

[17] J. W. McPherson, "Reliability challenges for 45nm and beyond," in *2006 43rd ACM/IEEE Design Automation Conference*, 24-28 July 2006 2006, pp. 176-181.

[18] K. Ahmed and K. Schuegraf, "Transistor wars," *IEEE Spectrum,* vol. 48, no. 11, pp. 50-66, 2011.

[19] B. H. Lee, J. Oh, H. H. Tseng, R. Jammy, and H. Huff, "Gate stack technology for nanoscale devices," *Materials Today,* vol. 9, no. 6, pp. 32-40, 2006/06/01/ 2006.

[20] R. S. Pal, S. Sharma, and S. Dasgupta, "Recent trend of FinFET devices and its challenges: A review," in *2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*, 3-4 March 2017 2017, pp. 150-154.

[21] D. Hisamoto, T. Kaga, and E. Takeda, "Impact of the vertical SOI 'DELTA' structure on planar device technology," *IEEE Transactions on Electron Devices,* vol. 38, no. 6, pp. 1419-1424, 1991.

[22] B. Narasimham, S. Gupta, D. Reed, J. K. Wang, N. Hendrickson, and H. Taufique, "Scaling trends and bias dependence of the soft error rate of 16 nm and 7 nm FinFET SRAMs," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 11-15 March 2018 2018, pp. 4C.1-1-4C.1-4.

[23] X. Guo, V. Verma, P. Gonzalez-Guerrero, S. Mosanu, and M. R. Stan, "Back to the Future: Digital Circuit Design in the FinFET Era," *Journal of Low Power Electronics,* vol. 13, no. 3, pp. 338-355, // 2017.

[24] C. Su, M. Armstrong, L. Jiang, S. A. Kumar, C. D. Landon, S. Liu, I. Meric, K. W. Park, L. Paulson, K. Phoa, B. Sell, J. Standfest, K. B. Sutaria, J. Wan, D. Young, and S. Ramey, "Transistor reliability characterization and modeling of the 22FFL

FinFET technology," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 11-15 March 2018 2018, pp. 6F.8-1-6F.8-7.

[25] Y. Lee, P. J. Liao, K. Joshi, and D. S. Huang, "Circuit-based reliability consideration in FinFET technology," in *2017 IEEE 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 4-7 July 2017 2017, pp. 1-7.

[26] A. Rahman, J. Dacuna, P. Nayak, G. Leatherman, and S. Ramey, "Reliability studies of a 10nm high-performance and low-power CMOS technology featuring 3rd generation FinFET and 5th generation HK/MG," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 11-15 March 2018 2018, pp. 6F.4-1-6F.4-6.

[27] R. G. Southwick, E. Wu, S. Mehta, and J. H. Stathis, "Time dependent dielectric breakdown of SiN, SiBCN and SiOCN spacer dielectrics," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2-6 April 2017 2017, pp. DG-1.1-DG-1.5.

[28] X. Federspiel, J. Jasse, D. Ney, D. Roy, and M. Rafik, "Middle of the Line Dielectrics Reliability and Percolation Modelling through 65nm to 28nm Nodes," in *2019 IEEE International Integrated Reliability Workshop (IIRW)*, 13-17 Oct. 2019 2019, pp. 1-4.

[29] W. P. Maszara and M. Lin, "FinFETs - Technology and circuit design challenges," in *2013 Proceedings of the ESSCIRC (ESSCIRC)*, 16-20 Sept. 2013 2013, pp. 3-8.

[30] D. James, "Intel Ivy Bridge unveiled — The first commercial tri-gate, high-k, metal-gate CPU," in *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference*, 9-12 Sept. 2012 2012, pp. 1-4.

[31] J. J. Naresky, "Reliability Definitions," *IEEE Transactions on Reliability,* vol. R-19, no. 4, pp. 198-200, 1970.

[32] C. Bailey, H. Lu, S. Stoyanov, C. Yin, T. Tilford, and S. Ridout, "Predictive reliability and prognostics for electronic components: Current capabilities and future challenges," in *2008 31st International Spring Seminar on Electronics Technology*, 7-11 May 2008 2008, pp. 67-72.

[33] J. G. Elerath and M. Pecht, "IEEE 1413: A Standard for Reliability Predictions," *IEEE Transactions on Reliability,* vol. 61, no. 1, pp. 125-129, 2012.

[34] R. B. Abernethy, *The new Weibull handbook : Reliability & statistical analysis for predicting life, safety, risk, support costs, failures, and forecasting warranty claims, substantiation and accelerated testing, using Weibull, Log normal, crow-AMSAA, probit, and Kaplan-Meier models*, 5th ed.. ed. North Palm Beach, Fla. : R.B. Abernethy, 2006.

[35] X. Li, J. Qin, and J. B. Bernstein, "Compact Modeling of MOSFET Wearout Mechanisms for Circuit-Reliability Simulation," *IEEE Transactions on Device and Materials Reliability,* vol. 8, no. 1, pp. 98-121, 2008.

[36] Q. V. Le, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, and A. Y. Ng, "On optimization methods for deep learning," presented at the Proceedings of the 28th International Conference on International Conference on Machine Learning, Bellevue, Washington, USA, 2011.

[37] C. Chen, F. Ahmed, and L. Milor, "A comparative study of wearout mechanisms in state-of-art microprocessors," in *2012 IEEE 30th International Conference on Computer Design (ICCD)*, 30 Sept.-3 Oct. 2012 2012, pp. 271-276.

[38] P. J. Roussel, A. Chasin, S. Demuynck, N. Horiguchi, D. Linten, and A. Mocuta, "New methodology for modelling MOL TDDB coping with variability," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 11-15 March 2018 2018, pp. 3A.5-1-3A.5-6.

[39] S. Lombardo, J. H. Stathis, B. P. Linder, K. L. Pey, F. Palumbo, and C. H. Tung, "Dielectric breakdown mechanisms in gate oxides," *Journal of Applied Physics,* vol. 98, no. 12, p. 121301, 2005.

[40] Y. Lee, N. Mielke, M. Agostinelli, S. Gupta, R. Lu, and W. McMahon, "Prediction of Logic Product Failure Due To Thin-Gate Oxide Breakdown," in *2006 IEEE International Reliability Physics Symposium Proceedings*, 26-30 March 2006 2006, pp. 18-28.

[41] J. W. McPherson, "Reliability Trends with Advanced CMOS Scaling and The Implications for Design," in *2007 IEEE Custom Integrated Circuits Conference*, 16-19 Sept. 2007 2007, pp. 405-412.

[42] H. Solar Ruiz and R. Berenguer Pérez, "CMOS Performance Issues," in *Linear CMOS RF Power Amplifiers: A Complete Design Workflow*. Boston, MA: Springer US, 2014, pp. 57-73.

[43] K. Yang, T. Liu, R. Zhang, and L. Milor, "A comparison study of time-dependent dielectric breakdown for analog and digital circuit's optimal accelerated test regions," in *2017 32nd Conference on Design of Circuits and Integrated Systems (DCIS)*, 22-24 Nov. 2017 2017, pp. 1-6.

[44] E. Y. Wu, B. Li, J. H. Stathis, and C. LaRow, "Time-dependent clustering model versus combination-based approach for BEOL/MOL and FEOL non-uniform dielectric breakdown: Similarities and disparities," in *2014 IEEE International Reliability Physics Symposium*, 1-5 June 2014 2014, pp. 5B.2.1-5B.2.7.

[45] C. Chen and L. Milor, "System-level modeling and microprocessor reliability analysis for backend wearout mechanisms," in *2013 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 18-22 March 2013 2013, pp. 1615-1620.

[46] F. Chen, O. Bravo, K. Chanda, P. McLaughlin, T. Sullivan, J. Gill, J. Lloyd, R. Kontra, and J. Aitken, "A Comprehensive Study of Low-k SiCOH TDDB Phenomena and Its Reliability Lifetime Model Development," in *2006 IEEE International Reliability Physics Symposium Proceedings*, 26-30 March 2006 2006, pp. 46-53.

[47] P. Chen, S. Lee, A. S. Oates, and C. W. Liu, "BEOL TDDB reliability modeling and lifetime prediction using critical energy to breakdown," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 11-15 March 2018 2018, pp. 6B.5-1-6B.5-6.

[48] K. N. Tu, "Recent advances on electromigration in very-large-scale-integration of interconnects," *Journal of Applied Physics,* vol. 94, no. 9, pp. 5451-5473, 2003/11/01 2003.

[49] C. Y. Liu, C. Chen, and K. N. Tu, "Electromigration in Sn–Pb solder strips as a function of alloy composition," *Journal of Applied Physics,* vol. 88, no. 10, pp. 5703-5709, 2000.

[50] S. Hsu, K. Yang, R. Zhang, and L. Milor, "Lifetime Estimation Using Ring Oscillators for Prediction in FinFET Technology," in *2018 International Integrated Reliability Workshop (IIRW)*, 7-11 Oct. 2018 2018, pp. 1-4.

[51] W. Ahn, H. Zhang, T. Shen, C. Christiansen, P. Justison, S. Shin, and M. A. Alam, "A Predictive Model for IC Self-Heating Based on Effective Medium and Image Charge Theories and Its Implications for Interconnect and Transistor Reliability," *IEEE Transactions on Electron Devices,* vol. 64, no. 9, pp. 3555-3562, 2017.

[52] A. Scorzoni, B. Neri, C. Caprile, and F. Fantini, "Electromigration in thin-film interconnection lines: models, methods and results," *Materials Science Reports,* vol. 7, no. 4, pp. 143-220, 1991/12/01/ 1991.

[53] D. J. Na, K. O. Aung, W. K. Choi, T. Kida, T. Ochiai, T. Hashimoto, M. Kimura, K. Kata, S. W. Yoon, and A. C. B. Yong, "TSV MEOL (Mid End of Line) and packaging technology of mobile 3D-IC stacking," in *2014 IEEE 64th Electronic Components and Technology Conference (ECTC)*, 27-30 May 2014 2014, pp. 596-600.

[54] J. H. Lau, "Supply chains for 3D IC integration manufacturing," in *2012 14th International Conference on Electronic Materials and Packaging (EMAP)*, 13-16 Dec. 2012 2012, pp. 1-7.

[55] D. Kim, D. Lee, Y. Seo, J. Park, S. Han, B. Jang, J. Kim, Y. Chung, S. Seo, and C. Lee, "Optimization and challenges on TSV MEOL integration," in *2014 IEEE 64th Electronic Components and Technology Conference (ECTC)*, 27-30 May 2014 2014, pp. 582-589.

[56] F. Chen, S. Mittl, M. Shinosky, R. Dufresne, J. Aitken, Y. Wang, K. Kolvenback, W. K. Henson, and D. Mocuta, "New electrical testing structures and analysis method for MOL and BEOL process diagnostics and TDDB reliability assessment," in *2013 IEEE International Reliability Physics Symposium (IRPS)*, 14-18 April 2013 2013, pp. PI.1.1-PI.1.5.

[57] X. Federspiel, D. Nouguier, D. Ney, and T. Ya, "Conductivity and reliability of 28nm FDSOI middle of the line dielectrics," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2-6 April 2017 2017, pp. DG-9.1-DG-9.4.

[58] T. Kauerauf, A. Branka, G. Sorrentino, P. Roussel, S. Demuynck, K. Croes, K. Mercha, J. Bömmels, Z. Tőkei, and G. Groeseneken, "Reliability of MOL local interconnects," in *2013 IEEE International Reliability Physics Symposium (IRPS)*, 14-18 April 2013 2013, pp. 2F.5.1-2F.5.5.

[59] E. Y. Wu, "Facts and Myths of Dielectric Breakdown Processes—Part I: Statistics, Experimental, and Physical Acceleration Models," *IEEE Transactions on Electron Devices,* vol. 66, no. 11, pp. 4523-4534, 2019.

[60] D. Kim and L. Milor, "Memory reliability estimation degraded by TDDB using circuit-level accelerated life test," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2-6 April 2017 2017, pp. RT-1.1-RT-1.6.

[61] K. Yiang, H. W. Yao, and A. Marathe, "TDDB Kinetics and their Relationship with the E- and $\sqrt{}$E-models," in *2008 International Interconnect Technology Conference*, 1-4 June 2008 2008, pp. 168-170.

[62] C. V. Dam and M. Hauser, "Ring oscillator reliability model to hardware correlation in 45nm SOI," in *2013 IEEE International Reliability Physics Symposium (IRPS)*, 14-18 April 2013 2013, pp. CM.1.1-CM.1.5.

[63] H. Dao, "Process evaluation, validation, and monitoring with ring oscillator scribelane modules," in *2016 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 16-19 May 2016 2016, pp. 218-219.

[64] A. Bhaskar, "Design and analysis of low power SRAM cells," in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 21-22 April 2017 2017, pp. 1-5.

[65] C. Bae, S. Pae, C. Yu, K. Kim, Y. Kim, and J. Park, "SRAM stability design comprehending 14nm FinFET reliability," in *2015 IEEE International Reliability Physics Symposium*, 19-23 April 2015 2015, pp. MY.13.1-MY.13.5.

[66] M. Qazi, M. Sinangil, and A. Chandrakasan, "Challenges and Directions for Low-Voltage SRAM," *IEEE Design & Test of Computers,* vol. 28, no. 1, pp. 32-43, 2011.

[67] F. B. Yahya, H. N. Patel, V. Chandra, and B. H. Calhoun, "Combined SRAM read/write assist techniques for near/sub-threshold voltage operation," in *2015 6th Asia Symposium on Quality Electronic Design (ASQED)*, 4-5 Aug. 2015 2015, pp. 1-6.

[68] S. Chun, J. D. S. Spüntrup, and J. N. Burghartz, "Design of online aging sensor architecture for mixed-signal intergrated circuit," in *2013 International Semiconductor Conference Dresden - Grenoble (ISCDG)*, 26-27 Sept. 2013 2013, pp. 1-4.

[69] M. Baybutt, C. Minnella, A. Ginart, P. W. Kalgren, and M. J. Roemer, "Improving digital system diagnostics through Prognostic and Health Management (PHM) technology," in *2007 IEEE Autotestcon*, 17-20 Sept. 2007 2007, pp. 537-546.

[70] T. Sutharssan, S. Stoyanov, C. Bailey, and Y. Rosunally, "Data analysis techniques for real-time prognostics and health management of semiconductor devices," in *18th European Microelectronics & Packaging Conference*, 12-15 Sept. 2011 2011, pp. 1-7.

[71] H. Choi, S. Heo, H. Hong, S. Yang, Y. Han, Y. Cho, S. Won, and T. Park, "High resolution short defect localization in advanced FinFET device using EBAC and EBIRCh," in *2017 IEEE 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 4-7 July 2017 2017, pp. 1-4.

[72] Y. Greenzweig, Y. Drezner, A. Raveh, O. Sidorov, and R. H. Livengood, "E-beam invasiveness on 65 nm complementary metal-oxide semiconductor circuitry," *Journal of Vacuum Science & Technology B,* vol. 29, no. 2, p. 021202, 2011.

[73] C. H. Chu, P. S. Kuo, T. F. Chang, J. X. Yang, J. Chang, and R. Yang, "Failure analysis of IC contains FinFET," in *2017 IEEE 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 4-7 July 2017 2017, pp. 1-4.

[74] B. Liu, Y. Hua, Z. Dong, P. K. Tan, Y. Zhao, Z. Mo, J. Lam, and Z. Mai, "The Overview of the Impacts of Electron Radiation on Semiconductor Failure Analysis by SEM, FIB and TEM," in *2018 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 16-19 July 2018 2018, pp. 1-6.

[75] Y. Pan, Y. Zhao, P. K. Tan, Z. Mai, F. Rivai, and J. Lam, "Problems of and Solutions for Coating Techniques for TEM Sample Preparation on Ultra Low-k Dielectric Devices after Progressive-FIB Cross-section Analysis," in *2018 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 16-19 July 2018 2018, pp. 1-5.

[76] H. Yang, Y. Li, H. Zhu, X. Yin, and A. Du, "Advanced TEM Application in 10nm below Technology Node Device Analysis," in *2019 IEEE 26th International Symposium on Physical and Failure Analysis of Integrated Circuits (IPFA)*, 2-5 July 2019 2019, pp. 1-4.

[77] G. Villareal, J. Na, J. Lee, and T. Ho, "Advantages of using big data in semiconductor manufacturing," in *2018 29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 30 April-3 May 2018 2018, pp. 139-142.

[78] J. Moyne and J. Iskandar, "Big Data Analytics for Smart Manufacturing: Case Studies in Semiconductor Manufacturing," *Processes,* vol. 5, no. 3, p. 39, 2017.

[79] D. Gkorou, A. Ypma, G. Tsirogiannis, M. Giollo, D. Sonntag, G. Vinken, R. v. Haren, R. J. v. Wijk, J. Nije, and T. Hoogenboom, "Towards Big Data Visualization for Monitoring and Diagnostics of High Volume Semiconductor Manufacturing," presented at the Proceedings of the Computing Frontiers Conference, Siena, Italy, 2017.

[80] T. Ishioka and Y. Nonaka, "Maximum likelihood estimation of Weibull parameters for two independent competing risk," *IEEE Transactions on Reliability,* vol. 40, no. 1, pp. 71-74, 1991.

[81] D. B. Kececioglu and W. Wendai, "Parameter estimation for mixed-Weibull distribution," in *Annual Reliability and Maintainability Symposium. 1998 Proceedings. International Symposium on Product Quality and Integrity*, 19-22 Jan. 1998 1998, pp. 247-252.

[82] W. Wang and M. Jiang, "Competing failure or mixed failure models," in *2014 Reliability and Maintainability Symposium*, 27-30 Jan. 2014 2014, pp. 1-6.

[83] U. Kusters, "Elements of statistical computing—numerical computation, R. A. Thisted. Chapman and Hall, New York and London, 1988. ISBN 0-412-01371-1, cloth £28.50, pp. xx + 427," *Journal of Applied Econometrics,* vol. 6, no. 2, pp. 220-224, 1991.

[84] Z. Chbili and A. Kerber, "Self-heating impact on TDDB in bulk FinFET devices: Uniform vs Non-uniform Stress," in *2016 IEEE International Integrated Reliability Workshop (IIRW)*, 9-13 Oct. 2016 2016, pp. 45-48.

[85] K. B. Yeap, F. Chen, H. W. Yao, T. Shen, S. F. Yap, and P. Justison, "A Realistic Method for Time-Dependent Dielectric Breakdown Reliability Analysis for Advanced Technology Node," *IEEE Transactions on Electron Devices,* vol. 63, no. 2, pp. 755-759, 2016.

[86] T. Shen, K. B. Yeap, C. Christiansen, and P. Justison, "Field acceleration factor extraction in MOL and BEOL TDDB," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2-6 April 2017 2017, pp. DG-2.1-DG-2.5.

[87] K. Yang, R. Zhang, T. Liu, D. Kim, and L. Milor, "Optimal Accelerated Test Regions for Time- Dependent Dielectric Breakdown Lifetime Parameters Estimation in FinFET Technology," in *2018 Conference on Design of Circuits and Integrated Systems (DCIS)*, 14-16 Nov. 2018 2018, pp. 1-6.

[88] B. H. Calhoun, Y. Cao, X. Li, K. Mai, L. T. Pileggi, R. A. Rutenbar, and K. L. Shepard, "Digital Circuit Design Challenges and Opportunities in the Era of Nanoscale CMOS," *Proceedings of the IEEE,* vol. 96, no. 2, pp. 343-365, 2008.

[89] S. Khan and S. Hamdioui, "Trends and challenges of SRAM reliability in the nano-scale era," in *5th International Conference on Design & Technology of Integrated Systems in Nanoscale Era*, 23-25 March 2010 2010, pp. 1-6.

[90] D. Kim and L. S. Milor, "Memory yield and lifetime estimation considering aging errors," in *2015 IEEE International Integrated Reliability Workshop (IIRW)*, 11-15 Oct. 2015 2015, pp. 130-133.

[91] K. Yang, T. Liu, R. Zhang, and L. Milor, "Circuit-level reliability simulator for front-end-of-line and middle-of-line time-dependent dielectric breakdown in FinFET technology," in *2018 IEEE 36th VLSI Test Symposium (VTS)*, 22-25 April 2018 2018, pp. 1-6.

[92] D. Kim, S. Hsu, and L. Milor, "Optimization of Experimental Designs for System-Level Accelerated Life Test in a Memory System Degraded by Time-Dependent Dielectric Breakdown," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 27, no. 7, pp. 1640-1651, 2019.

[93] T. Kishimoto, T. Ishihara, and H. Onodera, "On-chip reconfigurable monitor circuit for process variation and temperature estimation," in *2018 IEEE International Conference on Microelectronic Test Structures (ICMTS)*, 19-22 March 2018 2018, pp. 111-116.

[94] Y. Miyake, Y. Sato, S. Kajihara, and Y. Miura, "Temperature and Voltage Measurement for Field Test Using an Aging-Tolerant Monitor," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 24, no. 11, pp. 3282-3295, 2016.

[95] A. K. M. M. Islam, J. Shiomi, T. Ishihara, and H. Onodera, "Wide-Supply-Range All-Digital Leakage Variation Sensor for On-Chip Process and Temperature Monitoring," *IEEE Journal of Solid-State Circuits,* vol. 50, no. 11, pp. 2475-2490, 2015.

[96] D. A. Kamakshi, H. N. Patel, A. Roy, and B. H. Calhoun, "A 28 nW CMOS supply voltage monitor for adaptive ultra-low power IoT chips," in *2017 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, 16-19 Oct. 2017 2017, pp. 1-2.

[97] C. Chung and M. Sun, "An all-digital voltage sensor for static voltage drop measurements," in *2016 IEEE Sensors Applications Symposium (SAS)*, 20-22 April 2016 2016, pp. 1-4.

[98] R. Pi. "Frequency management and thermal control documentation." https://www.raspberrypi.org/documentation/hardware/raspberrypi/frequency-management.md.

[99] L. A. Escobar and W. Q. Meeker, "A Review of Accelerated Test Models," *Statistical Science,* vol. 21, no. 4, pp. 552-577, 2006.

[100] J. A. McLinn, "Understanding accelerated degradation," in *2016 Annual Reliability and Maintainability Symposium (RAMS)*, 25-28 Jan. 2016 2016, pp. 1-8.

[101] "IEEE Guide for Selecting and Using Reliability Predictions Based on IEEE 1413," *IEEE Std 1413.1-2002,* pp. 1-106, 2003.

[102] V. A. Sotiris, P. W. Tse, and M. G. Pecht, "Anomaly Detection Through a Bayesian Support Vector Machine," *IEEE Transactions on Reliability,* vol. 59, no. 2, pp. 277-286, 2010.

[103] J. McLeish, "Physics of Failure Based Simulated Aided/Guided Accelerated Life Testing," in *2018 Annual Reliability and Maintainability Symposium (RAMS)*, 22-25 Jan. 2018 2018, pp. 1-5.

[104] S. Hsu, K. Yang, and L. Milor, "Reliability and Accelerated Testing of 14nm FinFET Ring Oscillators," in *2019 XXXIV Conference on Design of Circuits and Integrated Systems (DCIS)*, 20-22 Nov. 2019 2019, pp. 1-7.