Procedure for Optimal D.C. Parameter Extraction for Hot-Carrier Degradation Model Calibration and Verification

by

Steve Gia Dao

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 25, 1997
[February 1998]

© 1997 Steve Gia Dao. All Rights Reserved.

The author hereby grants to M.I.T. permission to reproduce and distribute publicly paper and electronic copies of this thesis and to grant others the right to do so.

Author .....
Department of Electrical Engineering and Computer Science
September 25, 1997

Certified by .....
James Chung
Associate Professor of Electrical Engineering
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

MAR 27 1998 Eng.

Procedure for Optimal D.C. Parameter Extraction for Hot-Carrier Degradation Model Calibration and Verification

by

Steve Gia Dao

Submitted to the Department of Electrical Engineering and Computer Science on September 25, 1997, in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering

Abstract

This study presents a methodology for optimal and accurate extraction of the **n**, **m**, and **H** parameters for D.C. hot-carrier degradation modeling. The methodology is based on a Monte Carlo simulation of the two key elements in hot-carrier reliability studies: the degradation and lifetime correlation plots. An optimized method for parameter extraction is also developed based on the extrapolation of the quantity I_{SUB}/I_D at a device lifetime of 10 years. The focus of this study is to verify the existence of an optimized parameter extraction method and to explore its sensitivity to statistical estimators and device biasing conditions within the context of balancing the device stress time with the number of device measurements subject to a constraint of fixed total time for stressing. Simulation results indicate that for a given technology, the optimized method for parameter extraction highly depends upon the level of I_{SUB}/I_D biasing as well as the choice of statistical estimators used to model process variation. The stress time sampling scheme is also found to be an influential factor of this sensitivity.

Thesis Supervisor: James Chung
Title: Associate Professor of Electrical Engineering

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Jim Chung, for suggesting this topic to me and introducing me to the challenging world of hot-carrier research. It is hard to imagine where I would academically stand had Jim not call me one afternoon asking if I would be interested in joining his research group. His initiative saved this confounded and concerned graduate student from certain perils as many searches for an intriguing thesis project often ended in failure. How can one express further gratitude and a sense of indebtness to a mentor who provided guidance at times when no solution seemed to exist, displayed an inexhaustible supply of patience at moments when even the simplest tasks somehow became difficult to accomplish, and consistently peeked into the offices of his graduate students at occasions when some were not working?

My indebtness also goes to my sponsors at Allison Transmission Division of General Motors Corporation, which provided the financial support for the master's program at M.I.T. Special kudos go to my GM mentors, Goetz Schaeffer and Fred Faude, who guided my professional growth during my internship summers, competently represented my academic goals before the GM Fellowship committee, and displayed considerable patience and understanding during the last hurdles of this thesis.

Without the personal support, camaraderie, and friendship of my research colleagues, my thesis experience would have lost much significance. I am grateful to those in the Hot-Carrier Group: Abraham Seokwon Kim, Wenjie Jiang, Huy Le, Dewi Sugiharto, and Arifur Rahman; Statistical Metrology Group: Brian Stine, Tae Park, Vikas Mehrotra, Charles Oji, Dennis Ouma, and Rajesh Divecha; and SIMOX Group: Jung Uk Yoon and Jee-Hoon Krska.

My close friends outside the research circle are on many occasions vital links to sanity amidst a tense academic atmosphere of seemingly lifeless 24-hour work days. Recognition of value far transcends the mere attempts to list names: my roommates Joonah Yoon, Li Su, Mary Chen, and Shan-Ming Woo; my spiritual sisters Inn Yuk, Betty Tsai, Beth Koehler, Erica Ross, and Jahleel Nacita; my spiritual brothers David Pearah, Leslie Loo, Jung Uk Yoon, Thomas Chen, Sang Chin, and Walter Sun; and of course, those who know too much about me and yet still call me "friend" Phuong Nguyen, Berndie Strassner, and introducing again Davy Pearah.

To not mention one's parent would cause unnecessary verbal reprimand to be inflicted on oneself. Nonetheless, recognition for my immediate family traverses beyond mere duty but rather born out of love and appreciation for their years of unconditional support. These awesome Texans would be my mother Kim Cheng, brothers Michael and David Cheng, and sisters Mary and Susan Cheng. Of equal stature in unconditional love and support are my godparents Ray and Eunice Hausler, although they reside in Illinois.

Finally, I abide in the old saying that one must leave the best for last. I owe life itself and all properties of life be it joy, happiness, indifference, sadness, or pain to my Love, Lord, and Savior Jesus Christ. For it is through knowing, loving, and surrendering my life to Christ do I fully realize and delight in the meaning of life and thus enjoy its properties! How can one be but stand in awe of all the wonders God has brought into one's life especially for such an undeserving individual? To bring an immigrant family to the greatest country, to live in the greatest state of this country, to receive one of the finest advanced education in this country, to be guided to work for a company that is more of a family, to be so blessed with the love, kindness, gentleness, and beauty of Sarah Hosfield, and above all, to receive the gift of salvation! I am silent before thee my Lord, in awe of You.

Table of Contents

List of Figures 8

List of Tables 11

Chapter 1: Introduction 12

1.1 Physics of Hot-Carrier Degradation 12

1.2 Hot-Carrier Degradation Models 13

1.2.1 Derivation of the D.C. Degradation Model 13

1.2.2 Brief Overview of A.C. Hot-Carrier Degradation and Modeling 15

1.3 Motivation and Overview of Thesis 16

Chapter 2: Current Methodology for D.C. Parameter Extraction 17

2.1 Extraction of Degradation Model Parameter **n** 18

2.2 Parameters **m** and **H** Extraction 20

2.3 **n**, **m**, and **H** Dependence on E_{ox} 21

2.4 Issues Concerning Current Methodology 26

2.4.1 D.C. Modeling Issues - Fixed E_{ox} 26

2.4.2 A.C. Modeling Issues - Varying E_{ox} 27

2.4.3 Optimization Focus of this Thesis 28

Chapter 3: New Methodology for D.C. Parameter Extraction 29

3.1 Description of Monte Carlo Simulation 29

3.1.1 Developing the Simulation Models and Parameters 31

3.1.2 Calibrating n_o, m_o, $\log H_o$ and Listing of Simulation Assumptions 33

3.1.3 Tracing through the Rest of the Monte Carlo Simulation 35

3.2 Selecting an Element for D.C. Optimization 36

3.2.1 Alternative Elements for Optimization 37

Chapter 4: Analysis of Monte Carlo Simulation 39

4.1 Effects on Degradation Plot 39

4.1.1 Effects of Turning On Only One σ 40

4.1.2 Effects of Turning On Multiple σs 46

4.1.3 Effects of Altering the Technology and I_{SUB}/I_D Bias Variables 49

4.2 Effects on Empirical Lifetime Distribution 50

4.2.1 Effects of Turning On Only One σ 51

4.2.2 Effects of Turning On Multiple σs 55

4.2.3 Effects of Altering the Technology and Optimization Variables 57

4.3 Effects on Lifetime-Correlation Plot 61

4.3.1 Effects of Turning On Only One σ 61

4.3.2 Key Elements in Choosing Values for σn, σm, σlogH, σε in Optimization Methodology 70

Chapter 5: Optimization Results Based on D.C. Modeling Issues 72

5.1 Simulation Methodology 72

5.2 Response of Optimized Parameter-Extraction Method on Balancing Device Stress Time with Number of Device Measurements 75

5.2.1 Results of Study from above Methodology 75

5.2.2 Results of Study Upon Use of Different Assumptions 78

Chapter 6: Conclusion and Recommendations 84

6.1 Summary of Findings 84

6.2 Recommendations for Further Studies 86

Appendix A Derivation of ANOVA Models 87

A.1 Testing Each Treatment Mean Against All Treatments 87

A.2 Testing of Pairwise Treatment Means 88

Appendix B Simulation Code Written in Mathematica 3.0 93

References 98

List of Figures

Figure 2.1.1: Degradation plot used for extraction of parameter **n** 18

Figure 2.2.1: Lifetime-correlation plot used to extract parameters **m** and **H** 21

Figure 2.3.1: Dependence of D.C. parameters on oxide field 23

Figure 2.3.2: ANOVA table for (A), (B) Technology 1 and (C) Technology 2 25

Figure 3.1.1: Flow chart of Monte Carlo simulation algorithm 30

Figure 4.1.1: (A) Boundary curves from Monte Carlo simulation with only σn varying, (B) Comparison of calculating boundary curves from Monte Carlo simulation versus analytical equation assuming 1σ deviation 41

Figure 4.1.2: (A) Boundary curves from Monte Carlo simulation with only σm varying, (B) Comparison of calculating boundary curves from Monte Carlo simulation versus analytical equation assuming 1σ deviation 41

Figure 4.1.3: (A) Boundary curves from Monte Carlo simulation with only $\sigma logH$ varying, (B) Comparison of calculating boundary curves from Monte Carlo simulation versus analytical equation assuming 1σ deviation 41

Figure 4.1.4: (A) Degradation from Monte Carlo simulation for varying only $\sigma\varepsilon$, (B) Regression fit of the degradation 44

Figure 4.1.5: (A) Degradation at high $\sigma\varepsilon$ and low stress time, (B) Same degradation curve showing At^n form, (C) Zoom of (B) at low stress time 45

Figure 4.1.6: Illustration of σn and σm both on 48

Figure 4.1.7: Illustration of σn and $\sigma logH$ both on 48

Figure 4.1.8: (A) Illustration of σn, σm, $\sigma logH$, $\sigma\varepsilon$ turned on, (B) Zoom of (A) for time sequence normally used to collect sample data 49

Figure 4.1.9: Effects of changing I_{SUB}/I_D bias level on Δ model 50

Figure 4.2.1: Illustration of empirical τ distribution at each I_{SUB}/I_D of lifetime-correlation plot 51

Figure 4.2.2: Lifetime distributions with only σn varying 52

Figure 4.2.3: Lifetime distributions with only σm varying 52

Figure 4.2.4: Lifetime distributions with only σlogH varying 53

Figure 4.2.5: Lifetime distributions with σn, σm, σlogH, and σε varying 56

Figure 4.2.6: Comparison of distributions from Figure 4.2.5 with changes in σn, σm, σlogH and w, I_D 57

Figure 4.2.7: Lifetime distribution with change in n_o, m_o, $logH_o$ 59

Figure 4.2.8: Lifetime distributions at different I_{SUB}/I_D levels 60

Figure 4.2.9: Lifetime distributions at different stress time sampling 60

Figure 4.3.1: Effects of varying only σn on lifetime-correlation plot 62

Figure 4.3.2: 95% prediction intervals for σn of (A) 0.005, (B) 0.02, (C) 0.05, (D) 0.1 63

Figure 4.3.3: Effects of varying only σm on lifetime-correlation plot 64

Figure 4.3.4: Effects of varying only σlogH on lifetime-correlation plot 64

Figure 4.3.5: 95% prediction intervals for σm of (A) 0.1, (B) 0.4, (C) 0.8 65

Figure 4.3.6: 95% prediction intervals for σlogH of (A) 0.1, (B) 0.5, (C) 1 66

Figure 4.3.7: Effects of varying only σε on lifetime-correlation plot 68

Figure 4.3.8: 95% prediction intervals for σε of (A) 0.01, (B) 0.05, (C) 0.1, (D) 0.5, (E) 1 69

Figure 4.3.9: Representative lifetime-correlation plot from Monte Carlo simulation with σn, σm, σlogH, and σε at conservative values 71

Figure 5.2.1: I_{SUB}/I_D distribution showing results of study using assumptions from Table 5.1.1; τ denotes the number of device measurements 76

Figure 5.2.2: (A) Effects of different σn and σm on Δ plot for large stress time range, (B) for stress time range considered in the study 77

Figure 5.2.3: I_{SUB}/I_D distributions showing optimal region for outer I_{SUB}/I_D interval 81

Figure 5.2.4: I_{SUB}/I_D distributions showing optimal region for inner low I_{SUB}/I_D interval 81

Figure 5.2.5: Log plot of Figure 5.2.3 82

Figure 5.2.6: Log plot of Figure 5.2.4.....82

Figure 5.2.7: I_{SUB}/I_D distributions showing optimal region for inner medium I_{SUB}/I_D interval ..83

Figure 5.2.8: I_{SUB}/I_D distributions showing optimal region for inner high I_{SUB}/I_D interval 83

List of Tables

Table 2.3.1: Listing of E_{ox} used for each technology in ANOVA table 25

Table 3.1.1: Initial values used in Monte Carlo simulation 35

Table 4.1.1: Values used in simulating the degradation plot 39

Table 4.3.1: Summary of fitting parameters for σn varying 62

Table 4.3.2: Summary of fitting parameters for σm varying 67

Table 4.3.3: Summary of fitting parameters for σlogH varying 67

Table 4.3.4: Summary of fitting parameters for σε varying 70

Table 5.1.1: Listing of the values used in optimization 74

Table 5.2.1: Listing of new values used to determine optimal region 79

Table A.1: Results of pairwise treatment test for Technology 1 90

Table A.2: Results of pairwise treatment test for Technology 2 91

Chapter 1

Introduction

1.1 Physics of Hot-Carrier Degradation

Due to the continued scaling of MOSFET dimensions while the power-supply voltage remains constant, the resulting high electric fields generated within the device produce hot carriers which can damage the gate oxide. The high lateral-electric field at the MOSFET drain greatly energizes mobile charge carriers within the conducting channel at the pinch-off region. Some of the energetic carriers induce impact ionization forming electron/hole pairs. Some of the energetic electrons are further energized by the high electric field and can acquire (as an ensemble) an effective temperature much higher than that of the surrounding silicon lattice. These "hot" carriers can gain sufficient energy to cross over the energy barrier of the Si-SiO_2 interface, break Si-H bonds, and create different forms of oxide damage.

NMOS hot-carrier-induced oxide damage can be separated into three distinct types. Each mechanism occurs during different gate-voltage stress regimes (for fixed drain voltage). For low gate-voltage stress (V_G~V_T, peak gate hole-current region), the generation of oxide hole traps is the major degradation mechanism [1]. For medium gate-voltage stress (V_G~$V_D/2$, peak substrate-current region), acceptor-type interface state generation is the most important degradation mechanism [2]. These acceptor-type interface states are negatively charged when occupied and neutral when empty. For high gate-voltage stress (V_G~V_D, peak gate electron-current region), the electron trapping mechanism dominates [3]. These electron traps, whose occupancy is insensitive to bias voltage, have a similar effect on the device characteristic as the acceptor-type interface states.

The generation of hot-carrier-induced oxide damage has a detrimental effect on MOSFET performance as device characteristics such as threshold voltage V_T, drain current I_D, and transconductance g_m can be adversely changed [4],[5]. The degraded device performance over time can seriously affect the operation of the circuit; thus the issue of hot-carrier reliability exists as a major concern. In order to assess the extent of the hot-carrier damage and its impact on device and circuit performance, accurate reliability simulation based on properly calibrated degradation models is needed [6].

1.2 Hot-Carrier Degradation Models

Under realistic circuit operation, devices typically undergo A.C. hot-carrier degradation. A given A.C. waveform can be partitioned in time by small time steps such that approximately D.C. conditions can be applied within each time step. This quasi-static approximation allows the use of D.C. degradation model within each time step to predict A.C. degradation and evaluate device reliability.

1.2.1 Derivation of the D.C. Degradation Model

Acceptor-type interface state generation is commonly believed to be the dominant degradation mechanism affecting NMOSFET device and digital circuit performance [7]. It is in the medium gate-bias regime that substrate current is observed to correlate very well with the observed hot-carrier degradation [8]. Thus I_{SUB} can be used as a good monitor for the amount of interface-state generation.

The substrate current is a function of the drain current and other parameters which can be extracted from experimental measurements. A general equation for the substrate current is [2]:

$$I_{SUB} = CI_D e^{-\frac{\varphi_i}{q\lambda E_m}} \tag{1.1}$$

where C is a process-dependent parameter, φ_i is the critical energy for impact ionization, λ

is the mean free path for electrons, and E_m is the maximum lateral electric field at the drain. A model for E_m [9],[10] can be substituted into Equation (1.1) yielding:

$$I_{SUB} = \left(\frac{A_i}{B_i}\right) \cdot I_D \cdot (V_D - V_{DSAT}) \cdot e^{-\frac{B_i \cdot l_c}{V_D - V_{DSAT}}} \tag{1.2}$$

where V_D is drain to source voltage and V_{DSAT} is the drain saturation voltage defined as

$$V_{DSAT} = \frac{E_{crit} \cdot L \cdot (V_G - V_T)}{E_{crit} \cdot L + V_G - V_T} \tag{1.3}$$

parameters A_i and B_i are impact ionization coefficients, V_T is the threshold voltage, l_c is the length of the effective pinch-off region, L is the effective gate channel length, and E_{crit} is the critical field for velocity saturation. Both E_{crit} and l_c are functions of the bias voltages and other physical parameters shown in (1.4).

$$\begin{aligned} E_{crit} &= E_{crit0} + E_{critg} \cdot V_G + E_{critb} \cdot V_{SUB} \\ l_c &= (l_{c0} + l_{c1} \cdot V_D) \cdot \sqrt{t_{ox}} \end{aligned} \tag{1.4}$$

where V_{SUB} is the back-body bias and t_{ox} is the gate-oxide thickness.

Equation (1.1) can be used to correlate the amount of hot-carrier-induced damage at the Si-SiO_2 interface with the measurable quantity, I_{SUB}. The amount of interface traps generated is found by the following expression [2]:

$$\Delta N_{it} = C \cdot \left(\frac{I_D}{w} \cdot e^{-\frac{\varphi_{it}}{q\lambda E_m}} \cdot t_{stress} \right)^n \tag{1.5}$$

where w is the width of the device, φ_{it} is the critical energy for interface-state generation, and t_{stress} is the amount of time the device undergoes stress. Combining (1.1) and (1.5) yields the basic D.C. degradation model for NMOSFET acceptor-type interface-state generation, which underlines most hot-carrier reliability simulation tools:

$$\Delta N_{it} = \left(\frac{I_D}{w \cdot H} \cdot \left(\frac{I_{SUB}}{I_D} \right)^m \cdot t_{stress} \right)^n \tag{1.6}$$

Equation (1.6) expresses hot-carrier-induced interface-state generation in terms of quantities which can be experimentally measured or calculated. I_D and I_{SUB} are the measurable terms, while **n**, **m**, and **H** are extractable parameters. The parameter **n** is the degradation rate coefficient, which reflects either the reaction-limited or transport-limited regimes of the interface-state generation mechanism [2]. **H** is a process-dependent constant. The parameter **m** is called the voltage acceleration factor and defined as

$$m = \frac{\varphi_{it}}{\varphi_i} \tag{1.7}$$

which is the ratio of the critical energy needed for interface-state damage and impact ionization, respectively. It is important to note that accurate extraction of these three model parameters is crucial for accurate prediction of hot-electron degradation in any reliability simulation.

1.2.2 Brief Overview of A.C. Hot-Carrier Degradation and Modeling

Although the focus of this study is on D.C. hot-carrier degradation modeling and it's parameter extraction, realistic circuits are subjected to A.C. waveforms and hence undergo A.C. degradation. Therefore, a brief discussion on A.C. modeling is warranted.

One metric for how much stress a device undergoes during A.C. hot-carrier stress is the quantity hot-electron AGE which is defined as [11]:

$$AGE = \int_0^T \frac{I_D(t)}{w \cdot H} \cdot \left(\frac{I_{SUB}(t)}{I_D(t)}\right)^m dt \tag{1.8}$$

Both I_D and I_{SUB} are time-dependent, and the AGE is normally evaluated for one period of a waveform from t=0 to t=T(period). Within one period, an A.C. waveform can be rapidly changing and hence subject the device to many different stress-bias conditions. Combining (1.8) with (1.6), a power-law dependence on AGE can be shown for hot-electron deg-

radation[11]:

$$\Delta N_{it} = [AGE]^n \tag{1.9}$$

1.3 Motivation and Overview of Thesis

Proper calibration of the hot-carrier degradation models is essential for accuracy in hot-carrier reliability simulation tools. A crucial element to insure proper calibration and verification of the models is the accurate and efficient extraction of the modeling parameters. Much work has been performed in establishing evaluation guidelines for more consistent and effective use of these models in the simulation tools [6],[13]. One study has discussed the statistical issues involved in parameter extraction and briefly shown that an optimum between accuracy and efficiency exists [13]. However, a thorough study of this trade-off has never been performed.

This thesis presents a methodology for improving the D.C. hot-carrier degradation model parameter-extraction procedure. The current unoptimized procedure is examined in order to determine the relevant optimization issues. An element for optimization is developed for D.C. parameter extraction with concluding suggestions for further studies which can improve the proposed optimized method. The goal is to present a methodology for optimal model parameter- extraction within the framework of the evaluation guidelines for accurate A.C. circuit-level reliability simulation.

Chapter 2

Current Methodology for D.C. Parameter Extraction

The effects of hot-carrier-induced damage on device performance can be quite complex. However, in order to separate the damage creation mechanism from the effects of the damage on the device characteristics, a single well-understood parameter is often used to assess the degradation. Interface trap generation ΔN_{it} can lead to changes of the threshold voltage ΔV_T, reduction of the forward-linear drain current $\Delta I_D/I_D$ and transconductance, Δg_m. Degradation of any of these device parameters can serve as monitors for ΔN_{it}; however, $\Delta I_D/I_D$ is used in this study due to its high correlation to the localized hot-carrier induced damage and its ease of measurement.

Equation (1.6) can be now written as [14]:

$$\frac{\Delta I_D}{I_D} = K\left(\frac{I_D}{w \cdot H} \cdot \left(\frac{I_{SUB}}{I_D}\right)^m \cdot t_{stress}\right)^n \tag{2.1}$$

since $\frac{\Delta I_D}{I_D} \propto \Delta N_{it}$ and where K is a proportionality constant. In this study, we will assume K=1, and thus, the reduction in drain current is set directly equal to the amount of interface damage. This assumption is permissible because the true value of K is accounted for in the extraction of the technology-dependent parameter **H**. Both I_D and I_{SUB} are measured at their initial values under stress-bias conditions, w is the width of the device, and t_{stress} denotes the amount of time the device is stressed at a given D.C. voltage. The parameters **n**, **m**, and **H** are extracted degradation model parameters obtained from experimental measurements.

2.1 Extraction of Degradation Model Parameter n

Normally, **n** is extracted first by rewriting (2.1) as a power-law relationship [22]:

$$\frac{\Delta I_D}{I_D} = A \cdot t^n \tag{2.2}$$

where A is the power-law pre-coefficient and n is the power index. As evident from (2.2), a correlation between A and **n** exists and one study has shown that A increases with decreasing value of **n**, satisfying a simple exponential relationship [23].

The NMOSFET device data used in this study comes from a 0.4 micron, LDD process with an oxide thickness of 7nm and device widths of 10 micrometers. Figure 2.1.1 illustrates the power-law relationship of the hot-carrier induced degradation and is a representative plot used to extract the parameter **n** at a fixed E_{ox}, where E_{ox} is the oxide field at the drain and is defined as:

$$E_{ox} = \frac{V_G - V_D}{t_{ox}} \tag{2.3}$$

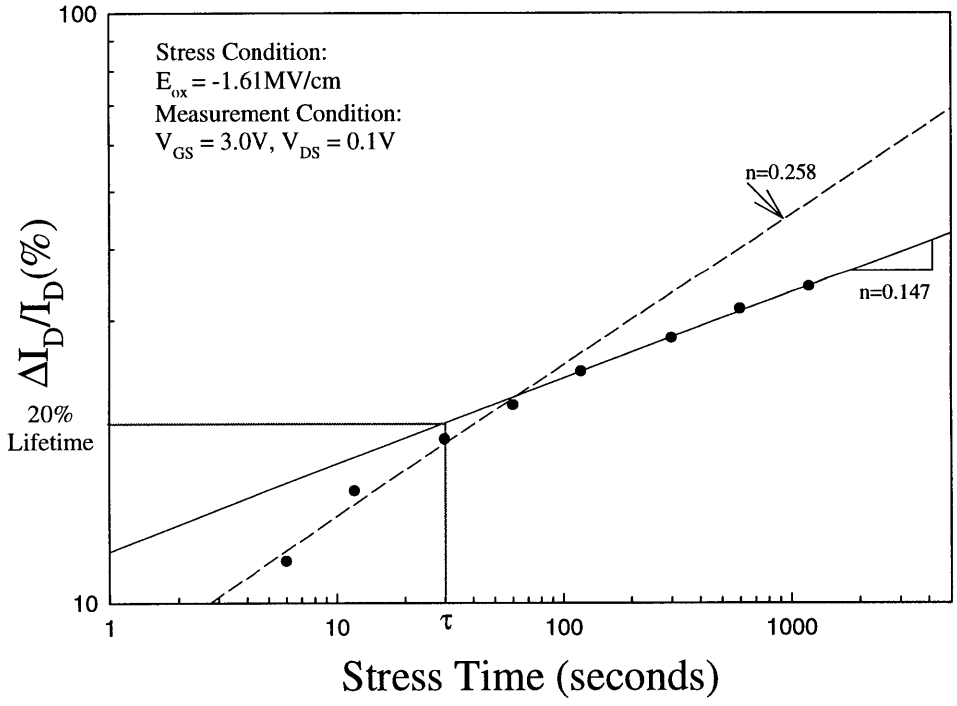


Figure 2.1.1: Degradation plot used for extraction of parameter **n**

The bias voltages for Figure 2.1.1 is V_G=2.4V and V_D=4.4V. For a constant E_{ox}, V_D is set much higher than the operating voltage and V_G follows from (2.3).

The stress time denotes how long a particular D.C. stress condition is applied to the device. The measurement time scheme of Figure 2.1.1 is {6, 12, 30, 60, 120, 300, 600, 1200, 3000} seconds with log spacing due to the log-log nature of the plot. At each time interval, the change in forward-linear drain current is measured at the operating voltage. The data points are regressed using the method of least squares, and the corresponding slope estimator becomes the parameter **n**.

For conventional devices with lightly-doped drains (LDD), the linear-current drain characteristic exhibits a two-stage effect which manifests into two different degradation rates [15]. This characteristic can be seen from the different values of the parameter **n** associated with the solid and dotted regression lines of Figure 2.1.1. Hence, the value of **n** seems to depend on the stress time. The reason for this is that LDD devices have oxide spacers used to reduce hot-carrier degradation[16]. These structures introduce additional degradation mechanisms in that trapped electrons in the spacer region increase the parasitic drain series resistance [17]. Furthermore, another study attributes this degradation, not only to the increased resistance underneath the LDD spacer region, but also in the reduction of carrier mobility in the subdiffusion and channel regions [15]. The degradation rate's saturation behavior affects the correlation between the device lifetime and I_{SUB}/I_D, which will be discussed in Section 2.2. However, this adverse effect can be eliminated by stressing the device for such a long duration that the degradation rate reaches its final asymptotic value, as shown by the solid line of Figure 2.1.1.

For each set of V_D and V_G bias conditions, the device lifetime, τ, is also calculated. The lifetime is defined as the stress time required for $\Delta I_D/I_D$ to change by a particular amount called the lifetime criterion. Equation (2.4) calculates the lifetime:

$$\tau = 10^{\frac{\Delta_{life} - \log A}{n}} \tag{2.4}$$

where Δ_{life} is the log of the lifetime criterion and logA is the intercept estimator from the linear regression. Different V_D and V_G bias conditions can be selected within a fixed E_{ox}, and each condition has an associated I_{SUB}/I_D value and lifetime, τ. The I_{SUB}/I_D and lifetime are correlated and used to extract the parameters **m** and **H**. This lifetime-correlation plot will be discussed in Section 2.2.

In summary, the operator has control of the follow variables associated with the extraction of the parameter **n**:

- the lifetime criterion Δ_{life},
- the choice of E_{ox},
- bias conditions within each E_{ox} - the spacing and quantity of V_D and V_G values,
- the stress time for each bias condition - the duration, spacing, and number of intervals for a particular sampling scheme.

2.2 Parameters m and H Extraction

For a particular E_{ox}, the gate and drain bias voltages are varied to observe the relation between the device lifetime and I_{SUB}/I_D. Normally, about 10-15 different bias conditions are sufficient to generate of a plot of normalized lifetime against normalized substrate current, as shown in Figure 2.2.1. The drain saturation current under stress conditions, I_D, is used as the normalization quantity.

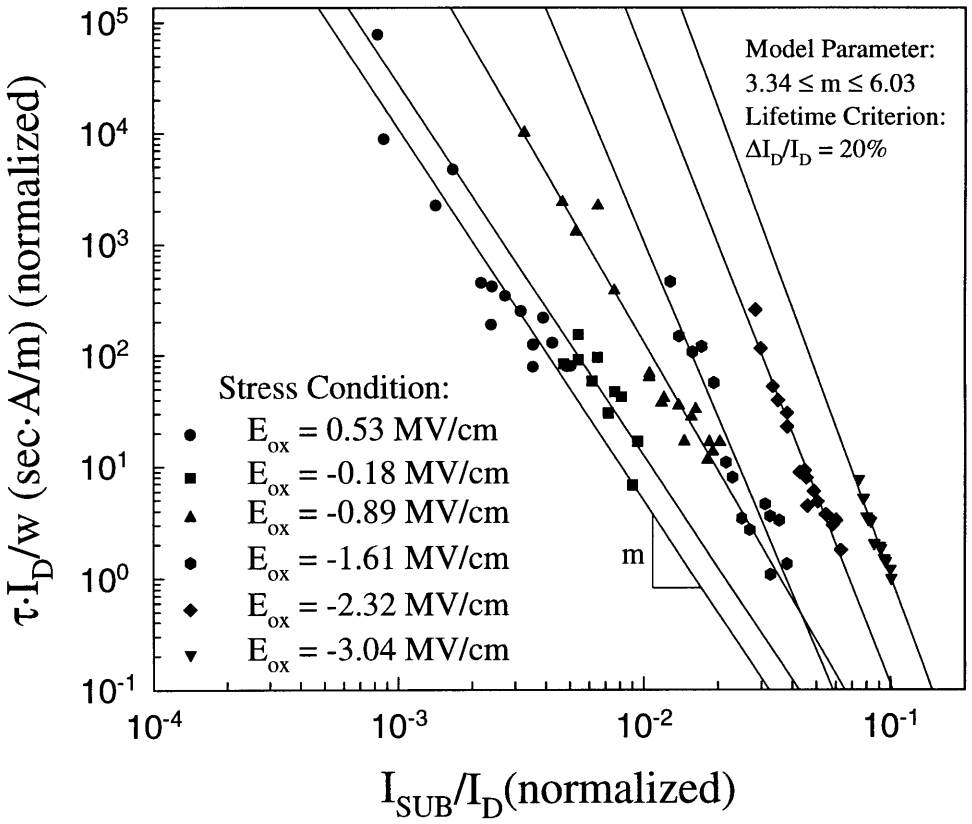


Figure 2.2.1: Lifetime-correlation plot used to extract parameters **m** and **H**

Figure 2.2.1 shows the lifetime-correlation for different families of E_{ox}. Each data point within a given E_{ox} family has a stress duration of 50 minutes, but different durations can be set by the operator.

The degradation model parameter **m** can be extracted from the lifetime-correlation plot. The parameter **m** is the slope of the regression line and **logH** which is the intercept. For a log-log plot, the intercept is where $I_{SUB}/I_D = 1$. Thus **logH** is extracted from the life-time-correlation plot and not **H**. However, **H** can be found by the following expression:

$$H = \frac{10^{\log H}}{\Delta_{life}^{1/n}} \tag{2.5}$$

Figure 2.2.1 shows a dependence of the voltage acceleration rate on the oxide electric field near the drain. Furthermore, a study has shown that this rate also has a dependence on

the lifetime criterion used if the asymptotic behavior of the degradation rate coefficient is not taken into account [13]. However, Figure 2.2.1 does not show a dependence of **m** on the chosen criterion, because each lifetime was determined when degradation had reached its final asymptotic behavior.

2.3 n, m, and H Dependence on E_{ox}

Figure 2.2.1 illustrates the family of different E_{ox} curves across varying bias conditions. The different **m** and **logH** values from Figure 2.3.1 definitely suggest a dependence of the voltage acceleration rate coefficient and process-dependent parameter on the oxide field. Proper accounting for the local oxide field dependence in extracting the degradation model parameters is essential, especially when hot-carrier evaluation under A.C. conditions require the model to be applicable over a wide range of operating voltages. This dependence can be accounted for by using different set of parameter values in the AGE model of Section 1.2.2.

The exact functional form of the E_{ox} dependence of the parameters **n**, **m**, and **H** is not known. Speculative physical explanation for the observed oxide-field dependence focuses either on the injection mechanisms or the generation of interface-state damage [17]. One speculative explanation is energy band bending of the Si substrate due to the applied oxide field, which forces the drain current path deeper into the silicon and further away from the Si-SiO_2 interface [18]. The greater the oxide field, the larger the amount of band bending, and the greater the distance hot-carriers must travel to reach the gate oxide. This results in additional critical energy (φ_{it}) required for hot electrons to cross over the Si-SiO_2 barrier height to create interface states [6].

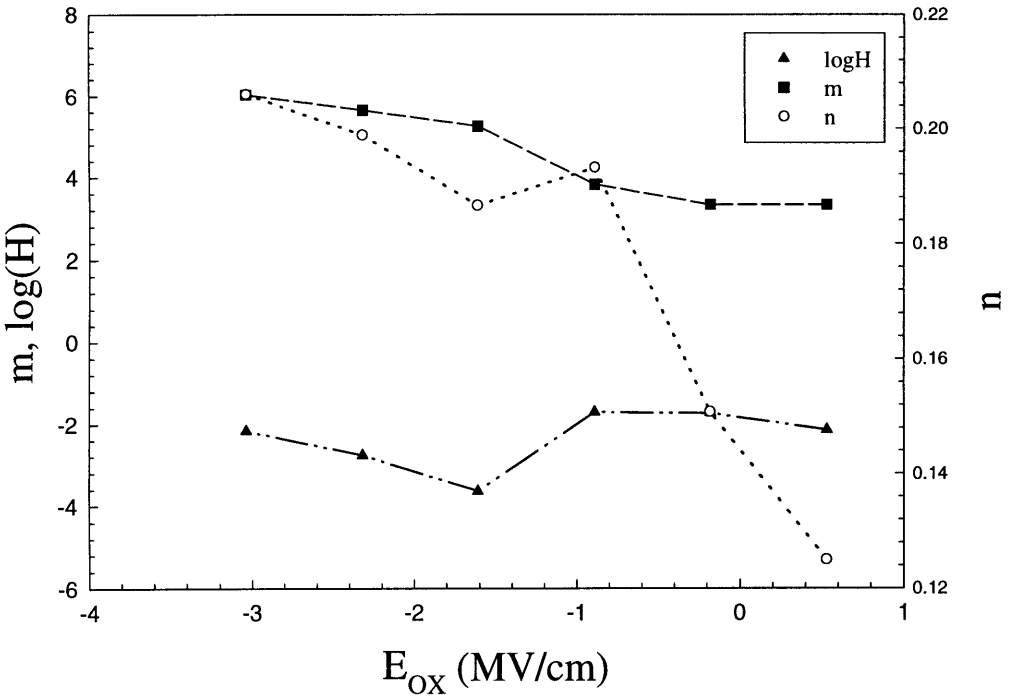


Figure 2.3.1: Dependence of D.C. parameters on oxide field

While the **m** and **H** dependence on E_{ox} is well observed and can be accounted for in A.C. hot-carrier degradation simulations, the dependence of the degradation rate coefficient **n** on the oxide field is less often taken into account [6]. Figure 2.3.1 seems to suggest that a dependence of **n** on E_{ox} exists, and one study shows a general relation between **n** and E_{ox} across a wide range of gate and drain bias conditions and device parameters [19]. This study states that improper accounting for this dependence results in significant overestimation of A.C. hot-carrier lifetime. Furthermore, the dependence is not primarily due to a combination of different degradation mechanisms such as interface-state generation or hole trapping but rather is an inherent feature of the interface-state generation mechanism itself. Other speculative explanation for the field dependence is field-dependent diffusion of interstitial hydrogen generated from the Si-H bond breaking that occurs during hot-carrier stressing. This hydrogen diffusion process has been proposed as the rate-limiting step for hot-carrier-induced interface-state generation [2].

The study in [19] was performed using conventional devices while the data from Figure 2.3.1 is for LDD structures. In order to determine the statistical dependence of the degradation rate coefficient on the oxide field for LDD devices, data from two LDD-based technologies is used. The first technology uses six E_{ox} families while the second has ten. Both technologies were stressed over a wide range of bias conditions. Table 2.3.1 lists the E_{ox} families used for each technology.

Hypothesis testing of the mean of **n** for each E_{ox} family (a treatment) is used to confirm or deny the dependence of **n** on all E_{ox} families. The hypothesis test is

H_o: $\mu_{Eox=0.53} = \mu_{Eox=-0.18} = \mu_{Eox=-0.89} = \mu_{Eox=-1.61} = \mu_{Eox=-2.32} = \mu_{Eox=-3.04}$

against

H_1: At least two of $\mu_{Eox=0.53}$, $\mu_{Eox=-0.18}$, $\mu_{Eox=-0.89}$, $\mu_{Eox=-1.61}$, $\mu_{Eox=-2.32}$, $\mu_{Eox=-3.04}$ differ.

The analysis of each treatment's mean is performed with the Analysis of Variance (ANOVA) technique. If the prob-values from the ANOVA tables of Figure 2.3.2 (A) and (C) are less than the α degree of confidence, then H_o is rejected in favor of H_1. For 95% degree of confidence, α is 0.05. The results confirm that for both technologies, there is a strong dependence of **n** on E_{ox}. Appendix A provides a derivation of the statistical models used in the ANOVA table as well as an elaboration on hypothesis testing. The analysis in Appendix A also shows the interdependence of one treatment on another and includes tables showing the amount of interdependence. For Technology 1, the treatments which have E_{ox} values less than or equal to -0.89MV/cm are statistically independent of each other. This result is confirmed by the prob value of Figure 2.3.2 (B), which is greater than the α confidence level.

(A)

Source	df	Sum of Squares	Mean Square	F Statistic	Prob Value
Model	v - 1 = 5	SS_{means} = 0.076	MS_{means} = 0.015	F(means) = 12.01	P = 8.37×10^{-9}
Error	N – v = 85	SSE = 0.108	MSE = 1.27×10^{-3}		
Total	N – 1 = 90	SS_{total} = 0.184			

(B)

Source	df	Sum of Squares	Mean Square	F Statistic	Prob Value
Model	v - 1 = 3	SS_{means} = 2.73×10^{-3}	MS_{means} = 9.10×10^{-4}	F(means) = 2.769	P = 0.462
Error	N – v = 56	SSE = 0.059	MSE = 1.05×10^{-3}		
Total	N – 1 = 59	SS_{total} = 0.061			

(C)

Source	df	Sum of Squares	Mean Square	F Statistic	Prob Value
Model	v - 1 = 9	SS_{means} = 0.136	MS_{means} = 0.015	F(means) = 13.756	P = 1.74×10^{-13}
Error	N – v = 87	SSE = 0.096	MSE = 0.0011		
Total	N – 1 = 96	SS_{total} = 0.232			

Figure 2.3.2: ANOVA table for (A), (B) Technology 1 and (C) Technology 2

Technology 1		**Technology 2**	
No. of Data Points	E_{ox} (MV/cm)	No. of Data Points	E_{ox} (MV/cm)
15	0.53	10	0.36
16	-0.18	10	-0.45
17	-0.89	9	-1.26
14	-1.61	10	-1.75
17	-2.32	9	-2.08
12	-3.04	10	-2.40
		10	-2.73
		8	-3.05
		11	-3.30
		10	-3.71

Table 2.3.1: Listing of E_{ox} used for each technology in ANOVA table

2.4 Issues Concerning Current Methodology

Section 2.1 briefly outlines the issues required for accurate extraction of the **n**, **m**, and **H** parameters. No optimized extraction procedure currently exists as the experimentalist simply chooses a stress time duration (normally 50 minutes), number of measurements (10-15 data points on lifetime-correlation plot), and the stress and bias conditions (from previous knowledge of the device's I-V curves). If the regression fit for the lifetime-correlation plot is poor, more measurements are taken. Therefore, in order to develop a methodology that optimizes the parameter extraction procedure for accuracy and efficiency, all the issues concerning the current unoptimized methodology must be examined. These issues can be divided into two categories: D.C. modeling issues and A.C. modeling issues. D.C. modeling issues are those which involve variables examined under a fixed E_{ox} while A.C. modeling issues involve a range of E_{ox} families.

2.4.1 D.C. Modeling Issues - Fixed E_{ox}

The following issues must be taken into consideration when developing an optimized procedure for D.C. parameter extraction.

- What are the individual V_G and V_D bias conditions within a particular E_{ox}? These V_G and V_D pairs also define the corresponding I_{SUB}/I_D values in the lifetime-correlation plot. One study states that the maximum stress voltage (V_D) is desirable to reduce the lifetime extrapolation error [13]. However, the stress voltage range is limited by two constraints: the upper-limit voltage should not turn on the parasitic source-bulk-drain bipolar transistors [21], and the lower-limit voltage should be set such that final asymptotic behavior is attained for the degradation rate. Another study suggests that a medium-to-high gate voltage range ($V_D/2 \geq V_G \geq V_D$), since the main hot-carrier induced mechanism in NMOSFETs is interface-state generation [3].
- Another issue associated with the bias condition is the spacing and total number of individual V_G and V_D pairs. The number of bias conditions addresses how many devices are used in the experiment and the total number of measurements needed. The number of measurements corresponds to the total number of data points appearing in the lifetime-correlation curve for a particular E_{ox} family.
- What should be the stress time scheme? This issue addresses the duration of each measurement, the number of intervals to be sampled and the spacing of these intervals for one duration.

The optimization for D.C. modeling focuses on balancing the issues addressed in the second and third bullets under the constraint of a total time allotted to extract **n**, **m**, and **H** for a particular E_{ox}. How should this total time be distributed between the duration of a measurement and the number of measurements? Longer individual device stress times reduce the lifetime extrapolation error, however, at the expense of fewer number of required measurements. The individual data points in the lifetime-correlation plot will have smaller associated error, but larger uncertainty in the regression fit due to fewer data points. Shorter individual device stress times increase the lifetime extrapolation error, however, more measurements can be achieved. Though there are more data points in the lifetime-correlation plot, which improves the regression fit, each measurement now has a larger associated error. Once an optimum between these two extremes has been determined, the gate and drain bias conditions can be varied to examine the behavioral response of an optimized parameter extraction method.

2.4.2 A.C. Modeling Issues - Varying E_{ox}

Analogous issues concerning A.C. modeling exist which are similar to those of D.C. modeling.

- What range of E_{ox} should be chosen? One study suggests an E_{ox} window around the peak substrate current since this region is found to minimize the interpolation and extrapolation errors [6]. The much observed correlation of substrate current with interface-state generation mechanism further supports this suggestion.
- The number of E_{ox} values and spacing within this range is also of concern.

Under the constraint that the total time allotted to extract a set of **n**, **m**, and **H** oxide field dependent parameters is constant, the allocation of this total time along with the total number of E_{ox} values needs to be optimized. A.C. parameter extraction optimization is complicated due to the lack of knowing the true functional dependence of **m** and **H** on the oxide field. The optimization method is further complicated if the distribution of time allotted for D.C. parameter extraction within a particular E_{ox} family is not uniform.

2.4.3 Optimization Focus of this Thesis

This thesis will only concern the optimization issues associated with D.C. modeling. Suggestions on A.C. optimization will be better discussed upon further studies of the functional dependence of **m** and **H** on the oxide field.

Chapter 3

New Methodology for D.C. Parameter Extraction

The focus of this research is to explore D.C. model parameter extraction issues. Due to the nonlinear equations involved in hot-carrier reliability modeling, the Monte Carlo simulation technique was chosen for the analysis.

3.1 Description of Monte Carlo Simulation

Most Monte Carlo simulations follow a standard algorithm which consists of developing simulation models of the system's events, generating these events in the model by random sampling from known probability distributions, collecting simulation statistics, and analyzing the results. The execution of the simulation models usually occurs on a computer using a simulation/mathematical language. Details of the Monte Carlo algorithm used in developing the methodology for optimized D.C. parameter extraction are shown in Figure 3.1.1.

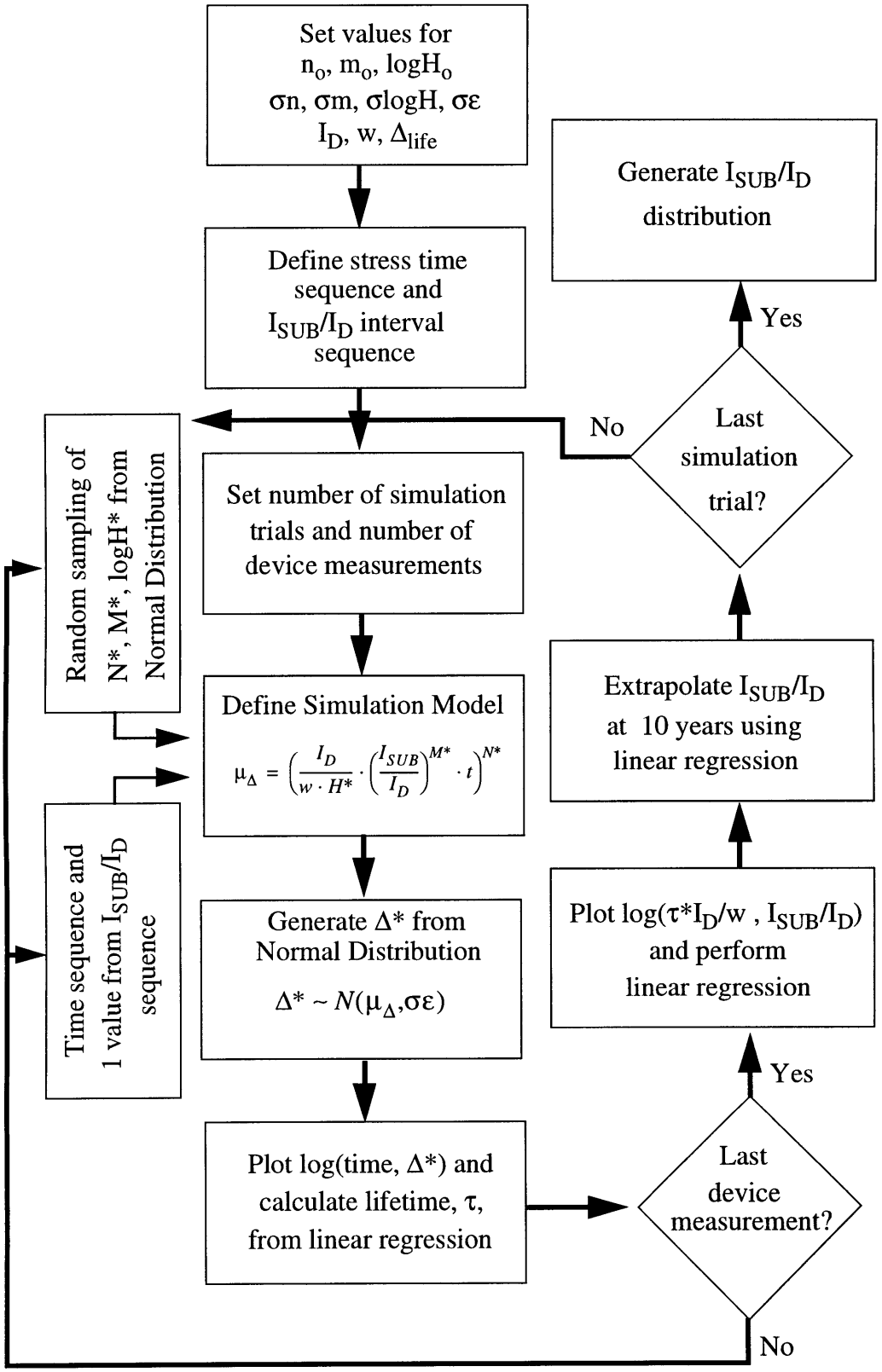


Figure 3.1.1: Flow chart of Monte Carlo simulation algorithm

3.1.1 Developing the Simulation Models and Parameters

The Monte Carlo simulation was based on the D.C. modeling process which followed a series of steps, from extracting the parameter **n** in the degradation plot, to extracting the parameters **m** and **H** in the lifetime-correlation plot. The model used in the simulation of the degradation plot was based on Equation (1.6):

$$\Delta^* \sim N(\mu_\Delta, \sigma\varepsilon)$$
$$\mu_\Delta = \left(\frac{I_D}{w \cdot H^*} \cdot \left(\frac{I_{SUB}}{I_D} \right)^{M^*} \cdot t \right)^{N^*} \tag{3.1}$$

where Δ^* represented the reduction of forward-linear drain current, $\Delta I_D/I_D$. Since **n**, **m**, and **H** were normally extracted from experimental data subject to statistical variation, they were modeled probabilistically in order to incorporate this randomness in the simulation. Modeling **n**, **m**, and **H** probabilistically also caused the quantity Δ^* to be probabilistic. Hence, simulating the degradation plot using (3.1) involved three group of variables: probabilistic, optimization, and technology. The probabilistic and optimization variables were the control variables in which their values were adjusted to analyze the model's response. The technology variables became constants for a given set of simulations and were calibrated according to known data.

The probabilistic variables were N*, M*, H*, and Δ^*. Process variations caused statistical variation in the parameters **n**, **m**, and **H**. The standard deviation, σ, for each of these variables accounted for the statistical variation due to device, die, and wafer level differences. Because the exact amount of variation, σ, for each variable was unknown and differed from one technology to the next, a range of values for σ was used to perform the analysis.

Although Δ^* is probabilistic due to N*, M*, and H*, Δ^* had its own standard deviation, $\sigma\varepsilon$, which modeled the instrumentation error and random scattering of the measured data. If exactly identical devices were to be used by the same instrument to

measure $\Delta I_D/I_D$, there would still be random variation in the measured linear drain current which was accounted for by $\sigma\varepsilon$. With the presence of $\sigma\varepsilon$, scattering would be greater for smaller values of Δ^*. Without the presence of $\sigma\varepsilon$, differences in the size and location of the stress time regions would become inconsequential.

The probabilistic variables were assumed to have a normal distribution with the exception of H*:

$$N^* \sim N(n_o,\sigma_n) \qquad M^* \sim N(m_o,\sigma_m)$$
$$\log H^* \sim N(\log H_o,\sigma_{\log H_i}) \tag{3.2}$$

where n_o, m_o, $\log H_o$, and their respective σs denoted the mean and standard deviation of each variable's normal distribution. Since **logH** was extracted from the lifetime-correlation plot and not **H**, logH* assumed a normal distribution with H* defined by the following function:

$$H^* = \frac{10^{\log H^*}}{\Delta_{life}^{1/N^*}} \tag{3.3}$$

where Δ_{life} was the lifetime definition. Δ^* was defined in (3.1). The mean was obtained from initial field data of a particular technology and σ was varied over a range centered around the mean.

The optimization variables consisted of the stress time and I_{SUB}/I_D sequences. These variables comprised the main control variables of the optimization routine. The stress time sequence had three degrees of freedom: the duration (t_{len}), number of intervals (N_{int}), and the spacing (k_{Tsp}). The duration was defined by the starting and ending stress times. The number of intervals corresponded to how many stress time values used for one duration. k_{Tsp} defined the exact value at each time interval.

Similar to this sequence was the I_{SUB}/I_D sequence for the lifetime-correlation plot. The different I_{SUB}/I_D values reflected different stress voltage conditions. This sequence

also contained three components: the bias length (I_{len}), number of intervals (N_{dev}), and the spacing (k_{Isp}). The bias length defined the starting and ending points of the sequence. N_{dev} defined how many bias conditions occurred within a fixed E_{ox} and also corresponded to the number of extrapolated device lifetimes or number of device measurements. k_{Isp} denoted the exact I_{SUB}/I_D values within the sequence. Although the experimentalist did not directly adjust I_{SUB}/I_D in changing the bias conditions, the associated V_G and V_D pairs could be readily calculated from MOSFET device models.

The technology variables consisted of n_o, m_o, $\log H_o$, w, I_D, and the lifetime criterion. Those values were determined using measured data from a particular technology, and for a set of simulation trials, these variables remained constant. The lifetime criterion was chosen to be 10% for this study. The initial stress value was used for the drain saturation current, I_D. Although I_D may have change during stressing, any such change was considered negligible. Thus I_D was assumed to remain at its fresh value throughout the simulation.

3.1.2 Calibrating n_o, m_o, $\log H_o$ and Listing Simulation Assumptions

The values for n_o, m_o, and $\log H_o$ came from the experimental measurements of a particular fixed E_{ox} for a given technology. This particular E_{ox} condition was chosen because of its high lifetime-correlation coefficient (> 90%) and its associated gate voltages were biased at the medium-to-high regime, where the major degradation mechanism was interface-state generation. Nine different bias voltages were applied to generate nine degradation curves for the same E_{ox}. Each curve had an associated degradation-rate coefficient, **n**, and the mean of these nine values became n_o. The standard deviation provided a basis to develop a range of values for σn.

The associated I_{SUB}/I_D and device lifetime extrapolated from each degradation curve were used to extract m_o and $\log H_o$ from the lifetime-correlation plot. The slope estimator

was m_o while the intercept estimator was $\log H_o$. Each estimator had an associated standard error which provided the basis for estimating a range of values for σm and $\sigma \log H$.

Using the same known data, $\sigma\varepsilon$ was derived by applying a nonlinear fit of the power-law relationship, At^n, to each degradation curve. Each fit was characterized by the mean square error (MSE) which measured the normalized square difference between the fitted curve and the measurement points. The MSE was normalized by the degrees of freedom from linear regression (which was the number of measurement points less 2) [20]. The mean of the MSE from each fitting of the nine degradation curves was defined to be $\sigma\varepsilon$. Since $\sigma\varepsilon$ had units of percentage, experimental data used in its calibration was represented with a plotting method whose ordinate scale reflected the same unit of percentage. A linear fit on a log scale plot would change the unit of measurement to log(%). Thus, a nonlinear fit on a linear scale plot was required in order to preserve the percentage units. Furthermore, a linear regression analysis on a log scale plot would result in a highly underestimated $\sigma\varepsilon$ value since the MSE would measure the square difference of a log operation.

Although individual values for σn, σm, $\sigma \log H$, and $\sigma\varepsilon$ were determined from the same set of experimental data, it was worthwhile to note that each value was highly coupled to the other. This study's analysis did not decouple the factors which uniquely influenced $\sigma\varepsilon$, σn, σm, or $\sigma \log H$ from each other. For example, the factors which caused only **n** to vary were embedded in the same experimental data used to estimate σs for **m** and **logH**. Due to a lack of filtering, high values for σn, σm, $\sigma \log H$, and $\sigma\varepsilon$ should be avoided and a range of values should be used instead with conservative choices toward the lower end. The base values served as a basis for estimating a range. Table 3.1.1 summarizes all the values discussed in this section and shows the base value for σn, σm, $\sigma \log H$, and $\sigma\varepsilon$.

Technology Variables		Probabilistic Variables	
n_o	0.278	σn : base, range	0.021, 0.005 - 0.1
m_o	3.537	σm : base, range	0.418, 0.1 - 0.8
$\log H_o$	2.214	$\sigma \log H$: base, range	0.525, 0.1 - 1.0
w	5μm	$\sigma \varepsilon$: base, range	0.015, 0.01 - 1
I_D	2.685mA		
Lifetime Criterion	10%		

Table 3.1.1: Initial values used in Monte Carlo simulation

3.1.3 Tracing through the Rest of the Monte Carlo Simulation

Once the simulation model had been established, the initial condition for the technology variables set, and a set of optimization variables selected, values for N*, M*, logH*, and Δ* were randomly sampled from the normal distribution. The computer generated the random sample using a pseudorandom number generator. This design used a Marsaglia-Zaman subtract-with-borrow generator for real numbers [24]. The advantages of this generator over most others were: implementation simplicity, speed, an extremely long period, and excellent performance on tests of randomness [25].

Once values for N*, M*, and logH* had been generated from random sampling, the mean value of Δ* was calculated according to (3.1) for a particular I_{SUB}/I_D value and at each stress time value. Δ* was generated from a normal distribution using the calculated mean values and user-defined σε according to (3.1). After generating Δ* for an entire stress time sequence, a linear regression was performed on the log-log plot. The lifetime, τ, was calculated at the lifetime definition using (2.4), where Δ_{life} was 1 (10% definition), logA was the intercept estimator, and n was the slope estimator from regression. This whole process was repeated for other I_{SUB}/I_D values until the I_{SUB}/I_D sequence expired.

Simulation of the lifetime-correlation plot involved graphing the extrapolated τ at each I_{SUB}/I_D value. All modeled process variation, measurement errors, and random scattering occurred in the degradation plot and their manifestations appeared in the dispersion of the (τ*I_D/w,I_{SUB}/I_D) points. Another linear regression was applied to this log-log data set, and the slope and intercept estimators (**m** and **logH** respectively) were used to extrapolate I_{SUB}/I_D at 10 years. Equation (3.4) showed the required calculations.

$$\frac{I_{SUB}}{I_D}(@10\text{years}) = 10^{\frac{C-\log H}{m}}$$
$$C = Log\left(\frac{315360000 \cdot I_D}{w}\right) \tag{3.4}$$

The process of random sampling N*, M*, and logH* to calculate I_{SUB}/I_D at 10 years was repeated for the number of simulation trials. Upon expiration of the trials, an empirical distribution for I_{SUB}/I_D was generated. The user could determine the number of simulation trials. The greater number of trials yielded finer resolution in the I_{SUB}/I_D distribution at the cost of a longer total simulation time.

3.2 Selecting an Element for D.C Optimization

Since the goal of hot-carrier reliability was to determine the effects of damage on the device at a distant future time, extrapolating values of I_{SUB}/I_D from the lifetime-correlation curve at 10 years lifetime was an excellent element upon which to base parameter-extraction optimization. Upon repeated simulation trials, a distribution of these I_{SUB}/I_D values at 10 years was generated. In assessing the trade-offs under a total time constraint for D.C. parameter extraction, short stress times led to large lifetime-extrapolation errors, which manifested in greater scattering of the data points in the lifetime-correlation plot. Hence, the I_{SUB}/I_D distribution at 10 years suffered even though the lifetime-correlation plot had more points. Longer stress times led to smaller extrapolation errors and reduced the scatter in the lifetime-correlation plot; however, the I_{SUB}/I_D distribution suffered due

to a fewer number of measurements. The optimum within this trade-off was defined as the stress conditions with the tightest resulting I_{SUB}/I_D distributions.

The time constraint problem could be formulated as:

$$T_{dev} = t_{len} + t_{instr} \cdot N_{int}$$
$$T_{total} = \sum_{i=1}^{N_{dev}} T_{dev}(i) \tag{3.5}$$

where T_{dev} was the total time taken to perform one stress measurement on a device, t_{len} was the stress duration, t_{instr} was the time needed for the instrument to take one reading, N_{int} was the number of intervals in the time sequence, and N_{dev} was the total number of devices for a given extraction experiment. Note, that the expression for T_{total} allowed different stress time durations for each devices. However, this study simplified the optimization problem by assuming a uniform allocation of stress time sequence for each device which consequently treated t_{instr} as a constant.

3.2.1 Alternative Elements for Optimization

Analogous to extrapolating I_{SUB}/I_D at a device lifetime of 10 years, the operating voltage at lifetime can also be determined and its distribution used as another element to base the parameter-extraction optimization. This alternative choice has further value as the maximum operating voltage is often used as the metric of comparison in many hot-carrier reliability studies [13],[26],[27]. However, this choice usually entails I_{SUB} and I_D experimental measurements to be taken at operating condition which subjects the optimization element to process variation. Or, the model for I_{SUB}/I_D can be derived from (1.2):

$$\frac{I_{sub}}{I_D} = \frac{A_i}{B_i} \cdot (V_D - V_{DSAT}) \cdot e^{-\frac{B_i \cdot l_c}{V_D - V_{DSAT}}} \tag{3.6}$$

which can be solved to obtain the operating voltage, V_D.

Another study has suggested that parameter-extraction optimization focus on the prediction interval from the regression fit of the lifetime-correlation plot [13]. The study states that a minimum prediction interval can be achieved by balancing the device stress time and the uncertainty associated with a smaller number of data points. Hence, an optimal extraction procedure which minimizes the extrapolation errors can be designed and performed [13]. The disadvantage of using statistical measures such as the prediction interval or mean square error and minimizing their magnitude as an optimization goal is that the associated statistical equations do not directly account for all the variables which need to be modeled. For example, the I_{SUB}/I_D values are used as the independent variables in the prediction-interval equation and no other terms are available to account explicitly for the device stress time. Therefore, the study chooses I_{SUB}/I_D bias conditions which are correlated to the stress-time duration. Such correlation reduces the effectiveness of the prediction interval as an optimization element.

Chapter 4

Analysis of Monte Carlo Simulation

The purpose of this analysis is to characterize the effects of the probabilistic, optimization, and technology variables on the various stages of the Monte Carlo simulation. The three major stages are: the degradation plot, empirical lifetime distribution for fixed I_{SUB}/I_D bias condition, and lifetime-correlation plot. Although all three sets of variables are examined in each stage, much of the focus will be on the probabilistic variables, since choosing the appropriate values from the ranges of Table 3.1.1 is pivotal in determining an optimal region, which is used to validate the existence of an optimal parameter extraction method.

4.1 Effects on Degradation Plot

The analysis of the degradation plot is based on the model of (1.6). Quantitative graphs are used to illustrate the individual effects of each σ as well as the simultaneous effects. Effects due to optimization and technology variables are discussed qualitatively. Table 4.1.1 lists the values used in the model of (1.6).

n_o	0.278
m_o	3.537
$\log H_o$	2.214
I_{SUB}/I_D	0.0616
I_D (mA)	2.685
w (μm)	5
lifetime	10%
time sequence (seconds)	6×10^{-7}, 1.2×10^{-6}, 3×10^{-6}, 6×10^{-6}, 1.2×10^{-5}, 3×10^{-5} ,..., 6×10^{4}, 1.2×10^{4}, 3×10^{4}

Table 4.1.1 Values used in simulating the degradation plot

4.1.1 Effects of Turning On Only One σ

In the cases of only one σn, σm, or σlogH active, Figure 4.1.1(A)-Figure 4.1.3(A) show the maximum deviation from the nominal degradation curve (represented by black squares) for selected σ values within the range from Table 3.1.1. The boundaries corresponding to each σ represent the worse-case deviation from the nominal out of 1000 simulation runs. Each selected σ values should also reflect boundary curves which clearly illustrate a deviation from one set to the next, which signifies that the choices for a particular σ should not be closely spaced together. Since the base values come from experimental measurements, the accuracy of representation is dependent on the sample size. This dependency further justifies the use of a range even though the true σ may be constant for a given technology but nonetheless remains unknown.

In Figure 4.1.1, the spread for σn is 0.005, 0.02, 0.05, and 0.1 with the base value at 0.02. σn values below 0.02 show the more realistic case of degradation usually found from current technologies, while higher σn values extend the possible degradation boundary to extreme values which show damage in excess of 1000% of the nominal for the low-to-medium stress-time range. Even at the base value of 0.02, the degradation boundary is more than double the nominal for also the low-to-medium stress-time range. In actuality, the degradation boundary for σn=0.02 should be much less than double the nominal because this σ value's effect on degradation also reflects those due to other σs. The percentage of linear current degradation with only σn active can be expressed as:

$$\log\left(\frac{\Delta I_D}{I_D}\right) = N^* \cdot \log t + N^* \cdot \log\left(\left(\frac{I_{SUB}}{I_D}\right)^{m_o} \cdot \frac{I_D}{w \cdot H_o}\right) \tag{4.1}$$

where t is the stress time and N* has a normal distribution described in (3.2). Although **H** also has an **n** dependence, it is treated as a constant. Figure 4.1.1 shows a rather large time sequence in order to illustrate the divergence of the boundary curves from the nominal curve at small stress times, while at large stress times, the boundary curves converge.

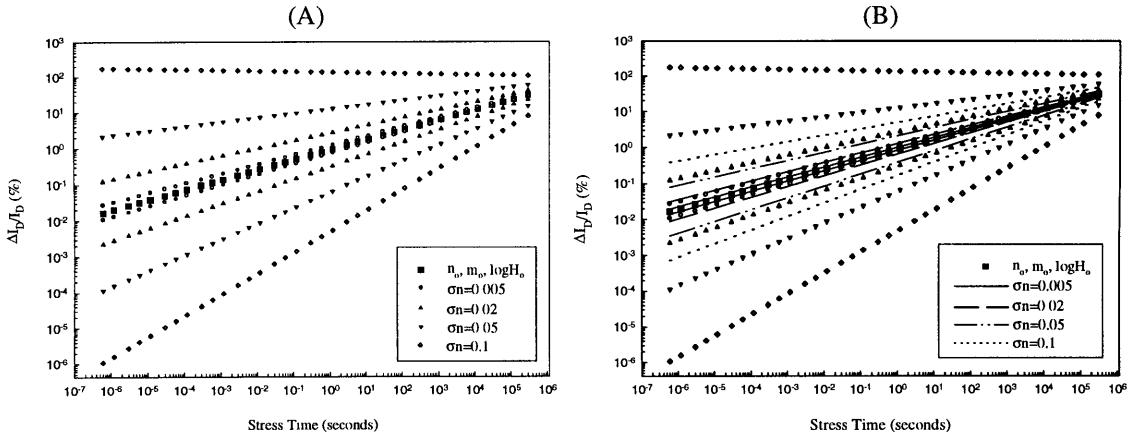


Figure 4.1.1. (A) Boundary curves from Monte Carlo simulation with only σn varying, (B) Comparison of calculating boundary curves from Monte Carlo simulation versus analytical equation assuming 1σ deviation

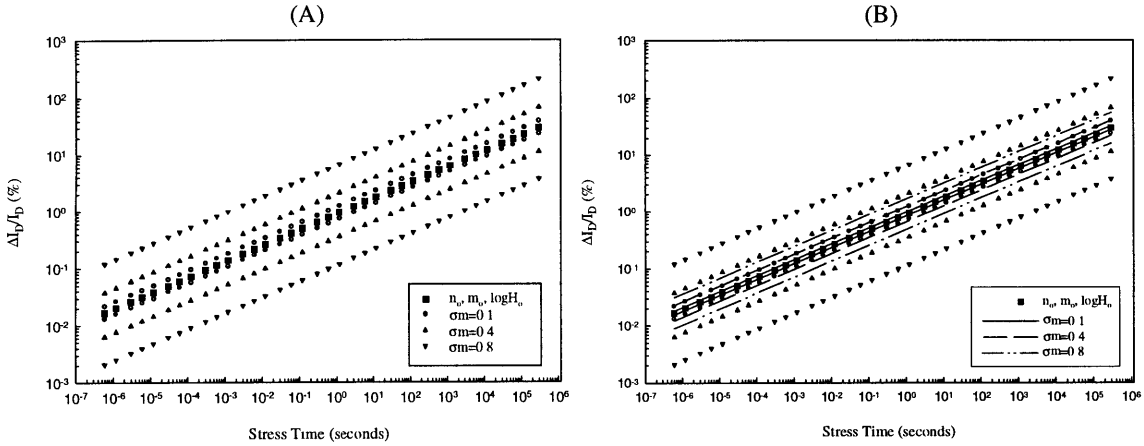


Figure 4.1.2. (A) Boundary curves from Monte Carlo simulation with only σm varying, (B) Comparison of calculating boundary curves from Monte Carlo simulation versus analytical equation assuming 1σ deviation

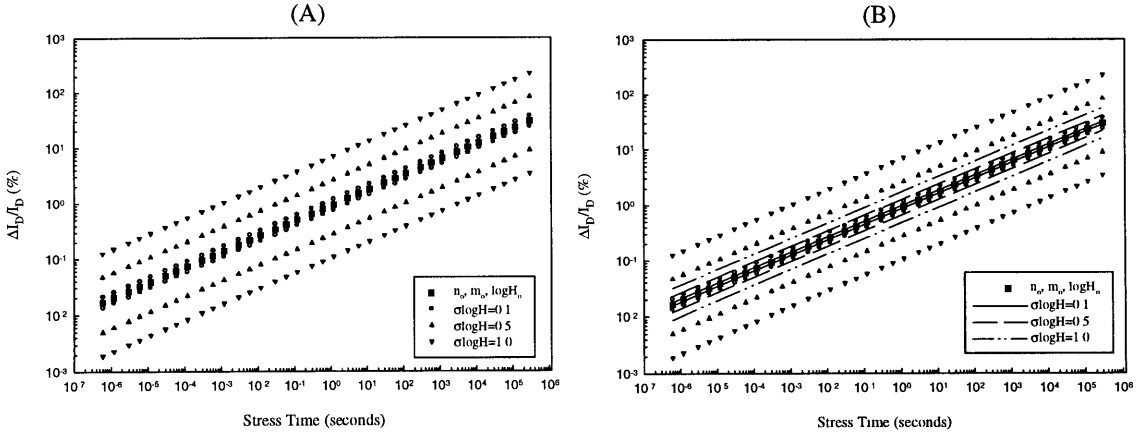


Figure 4.1.3. (A) Boundary curves from Monte Carlo simulation with only σlogH varying, (B) Comparison of calculating boundary curves from Monte Carlo simulation versus analytical equation assuming 1σ deviation

From inspection of Equation (4.1) and illustrated in Figure 4.1.1, the slope of each boundary curve changes. For the set of curves above the nominal, the upper boundaries are characterized by slopes less than n_0. For large σn values like 0.1, the slope of the upper boundary can be less than zero and hence leads to a set of boundary curves with decreasing non-monotonic slopes. For the set of curves below the nominal, the lower boundaries remain monotonic as their slopes increase above n_0 for increasing σn. Furthermore, Equation (4.1) shows that the vertical intercept for each curve changes by a factor of N*. Hence for a given σn, none of the degradation curves within the boundary ever crosses for the time frame shown, despite the variation in slope for each one. At extremely large stress times beyond what is shown, the boundary curves not only cross each other but also the nominal curve.

Figure 4.1.1(A) shows boundary curves which are determined as the worse-case deviation from the nominal curve out of 1000 simulation runs. Simulation is not the only manner in which the boundary curves can be determined. They can be calculated analytically using Equation (4.1) by substituting $(n_0+\sigma n)=N^*$. However, the analytical method is less accurate as shown by comparison of the boundary curves from simulation with the ones calculated analytically in Figure 4.1.1(B). The substitution assumes a 1σ deviation from the mean which only covers 80% of the total sampling space. A simulation is more precise in defining the boundaries since the sampling range is not confined to 1σ deviation.

The effects of varying only σm or $\sigma logH$ on the degradation plot are shown in Figure 4.1.2 and Figure 4.1.3. The spread of σm is 0.1, 0.4, 0.8 with its base value at 0.42, and the spread of $\sigma logH$ is 0.1, 0.5, 1.0 with its base value at 0.52. All the boundary curves have constant slope of 0.278 but with different intercepts. Equations (4.2) and (4.3) describe the vertical displacement of each boundary curve:

$$\log\left(\frac{\Delta I_D}{I_D}\right) = n_o \cdot \log t + n_o \cdot \log\left(\left(\frac{I_{SUB}}{I_D}\right)^{M^*} \cdot \frac{I_D}{w \cdot H_o}\right) \tag{4.2}$$

$$\log\left(\frac{\Delta I_D}{I_D}\right) = n_o \cdot \log t + n_o \cdot \log\left(\left(\frac{I_{SUB}}{I_D}\right)^{m_o} \cdot \frac{I_D}{w \cdot H^*}\right) \tag{4.3}$$

For both cases, the degradation boundaries corresponding to σs' below or near each respective base value (σm=0.4 and σlogH=0.5) best represent the cases from experimental observation.

A comparison of these two cases with that of σn shows that the effects of σn on the degradation plot is much greater at the low-to-medium stress time range. When either σm or σlogH is doubled from its base value (σm=0.8 and σlogH=1), the degradation reaches an order of magnitude greater than the nominal curve. When σn is doubled from its base value, degradation is in excess of two orders of magnitude at the low stress times and slightly one order of magnitude above at the medium stress times. However at the high stress time range, the degradation approaches the nominal for the case of σn while it remains an order of magnitude different for σm and σlogH.

The degradation due to varying only σε is more sensitive to the location of the stress time interval and the particular I_{SUB}/I_D bias level. Figure 4.1.4(A) illustrates this dependency. Equation (3.1) defines σε as the standard deviation of the mean, $\Delta I_D/I_D$, whose magnitude is a function of many parameters such as **n**, **m**, **H**, and I_{SUB}/I_D. For small I_{SUB}/I_D levels and at low stress times, the degradation can be orders of magnitude different from the nominal, especially when σε is large (greater than 0.1). However at higher stress times, the choice of σε has negligible effect on degradation even at small I_{SUB}/I_D. At the medium stress times which reflect a more realistic range used in experimental measurements (1-10,000 seconds), large σε can cause significant deviation from the nominal, which normally does not coincide with experimental observation. Small σε can also cause significant deviations if the I_{SUB}/I_D level is very low. Therefore, the

choice of σε should be balanced with the I_{SUB}/I_D level, and, to a lesser extent, the stress time.

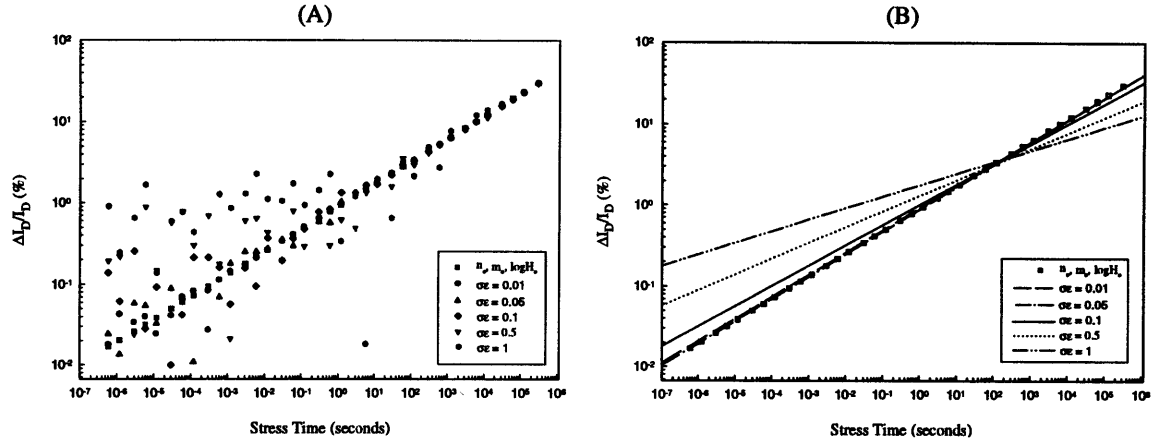


Figure 4.1.4. (A) Degradation from Monte Carlo simulation for varying only σε, (B) Regression fit of the degradation

Figure 4.1.4(B) characterizes the degradation scatter at low stress times and high σε values for given I_{SUB}/I_D level by using regression fits. At σε of 0.01, the fit virtually coincides with the nominal while fits of higher σε values increasingly depart from the nominal. The departure is characterized by a decrease in slope, increase in vertical intercept, and poor fitting coefficient. Although not apparent in Figure 4.1.4, the increase in vertical intercept is small and not noticeable on the scale shown. The poor fitting coefficient makes the fit for higher σε less reliable. Since the presence of high σε at low stress times causes the poorer fit and less reliable slope, the choice of σε does not have to be balanced with the stress time to the same extent as the I_{SUB}/I_D level if the low stress time range is avoided.

Figure 4.1.5(A) illustrates another issue for high σε at low stress times. Although the data points (black squares) are generated from a gaussian distribution about the nominal curve (hallow circles), they are distribution mostly above the nominal curve below the stress time of 1 second. Depending on the relative magnitude of the nominal curve, a high σε can frequently generate negative Δ* values at low stress times. The generation of these

points assumes a normal distribution which has equal probability of generating a point above and below the nominal. The Monte Carlo algorithm discards these negative values and only retains those with positive quantities, because subsequent regression analysis uses the log of the Δ^* values. Hence the overall distribution of points is not normal, as the number of points appearing below the nominal does not coincide in number with those lying above. Figure 4.1.5(B) and (C) show both curves on a linear-linear scale and the amount of deviation from the nominal after applying $\sigma\varepsilon$. Figure 4.1.5(C) shows the apparent deviation at low stress times. As discussed in Section 3.1.2, the need for nonlinear regression to determine the base value of $\sigma\varepsilon$ attributes the manifested effects of $\sigma\varepsilon$ to the At^n form and not to the log form. Furthermore, the nonlinear regression accounts for the significant deviations at high stress times as seen in Figure 4.1.5(B) while the log operation of Figure 4.1.5(A) minimizes the deviations which result in an underestimated $\sigma\varepsilon$.

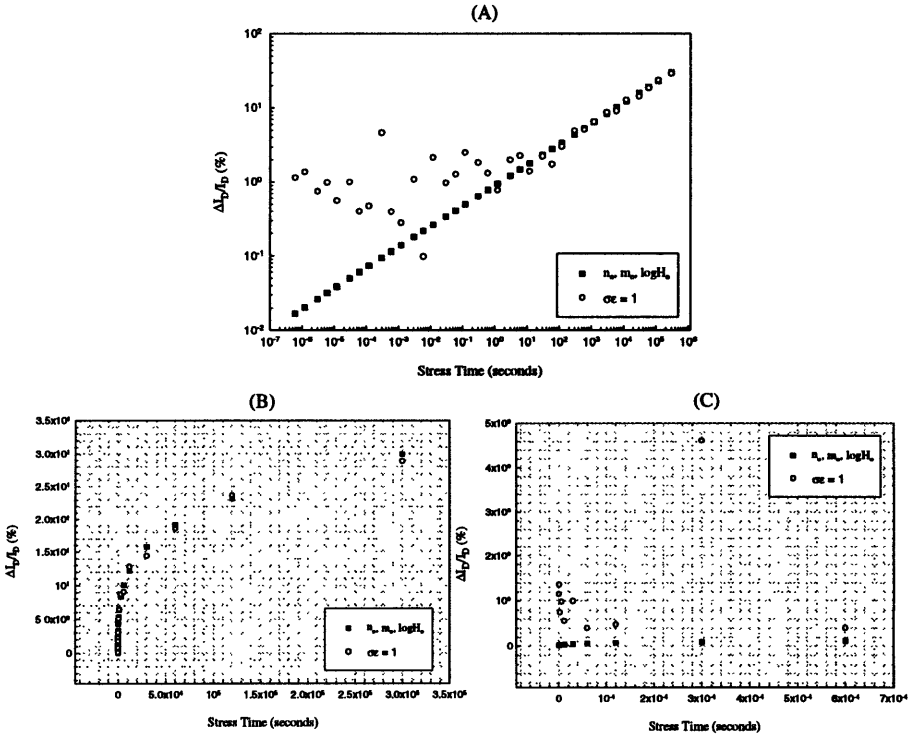


Figure 4.1.5. (A) Degradation at high $\sigma\varepsilon$ and low stress time, (B) Same degradation curve showing At^n form, (C) Zoom of (B) at low stress time

4.1.2 Effects of Turning On Multiple σs

Figure 4.1.6 and Figure 4.1.7 illustrate the cases for both σn and σm active as well as σn and σlogH. Each respective σ is chosen such that their combined effects are apparent. The combined effects of σn and σm cause the convergence of the boundary curves toward the nominal due to the change in slope. The boundary curves are also displaced further from the nominal due to greater deviation in the vertical intercept. σn changes the slope and vertical intercept in the same manner as describe for the case of only σn varying, while σm adds addition deviation to the intercept due to the power factor effect of M*, as shown in (4.4).

$$\log\left(\frac{\Delta I_D}{I_D}\right) = N^* \cdot \log t + N^* \cdot \log\left(\left(\frac{I_{SUB}}{I_D}\right)^{M^*} \cdot \frac{I_D}{w \cdot H_o}\right) \tag{4.4}$$

The effects of σn and σlogH are less intuitive and the boundary curves for different σ values are difficult to define. Unlike the previous cases, the worse-case boundary curve cannot define a limit within which lie all curves generated in the simulation for σn and σlogH values. The coupled effects of σn and σlogH cause the intercept to change in a non-monotonic direction such that the generated curves can cross each other within the stress time frame of Figure 4.1.7. This cross-over effect is the reason for the difficulty in defining a boundary region for σn of 0.02 and σlogH of 0.5. Two sets of curves are used to mark one possible boundary. The curves represented by shaded circles is chosen because it has the greatest deviation from the nominal referenced at the stress time of $6x10^{-7}$ seconds, while the curves represented by shaded triangles is chosen because it has the greatest deviation from the nominal referenced at the stress time of $3x10^{4}$ seconds. Both curves are determined from 1000 simulation trials. Since this cross-over effect has not been observed from experimental measurements, it is best to turn σlogH off during the optimization and lump the variation due to **logH** with the other σs (essentially increasing their values).

The reason for the non-monotonic, directional change of the intercept is due to H* residing in the denominator of the log-product as shown in (4.5). H* has an inverse effect on the intercept while the N* multiplier has a direct effect such that the direction of change for the intercept highly depends upon the magnitude of H* and N*. For example, increasing both H* and N* does not necessarily cause the intercept to increase. If the magnitude of increase for H* is much greater than that for N*, the intercept decreases. Thus this effect on the intercept similarly causes the degradation level to depend on the relative magnitude of N* and H*.

$$\log\left(\frac{\Delta I_D}{I_D}\right) = N^* \cdot \log t + N^* \cdot \log\left(\left(\frac{I_{SUB}}{I_D}\right)^{m_o} \cdot \frac{I_D}{w \cdot H^*}\right) \tag{4.5}$$

When both σm and σlogH are active, the behavior of the boundary curves is similar to Figure 4.1.2 and Figure 4.1.3 with the exception that the intercepts' displacement from the nominal depends on the value of M* and H*. As seen in (4.6), an increase or decrease in M* and H* usually causes the intercept to act in opposite directions while retaining the same slope, n_o. Since I_{SUB}/I_D is less than 1, an increase in M* decreases $(I_{SUB}/I_D)^{M^*}$; thus combining M* and H* further displaces the boundary curves from the nominal.

$$\log\left(\frac{\Delta I_D}{I_D}\right) = n_o \cdot \log t + n_o \cdot \log\left(\left(\frac{I_{SUB}}{I_D}\right)^{M^*} \cdot \frac{I_D}{w \cdot H^*}\right) \tag{4.6}$$

Figure 4.1.8(A) shows a realistic example of when σn, σm, σlogH, and σε are active for low-to-high stress time sequence. All the σs have been chosen at the end of the low range listed in Table 3.1.1. When σε is active in conjunction with other σs, the generated data points are distributed about a new curve based on σn, σm, and σlogH active. Due to the small σ values, this curve virtually coincides with the nominal. The degradation for a time sequence normally used in experimental measurements is shown in Figure 4.1.8(B), which is a more accurate representation of one of the technologies used in this study since the amount of deviation from the nominal for this time regime is small. At low stress times

and small σε=0.01, the degradation becomes apparent as the simulated data points exhibit a gaussian distribution about the σn, σm, and σlogH curve.

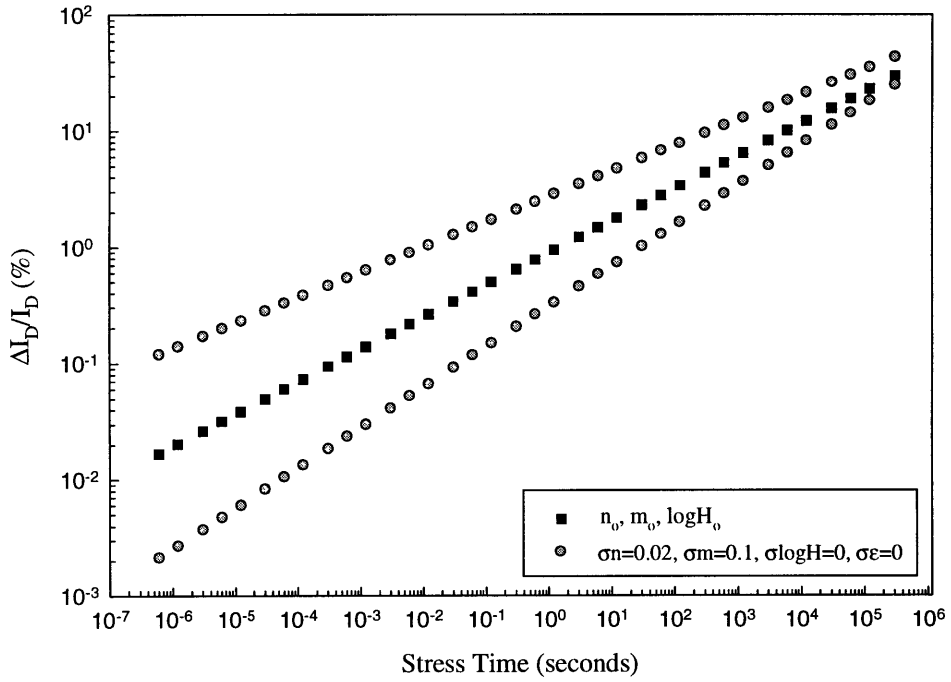


Figure 4.1.6. Illustration of σn and σm both on

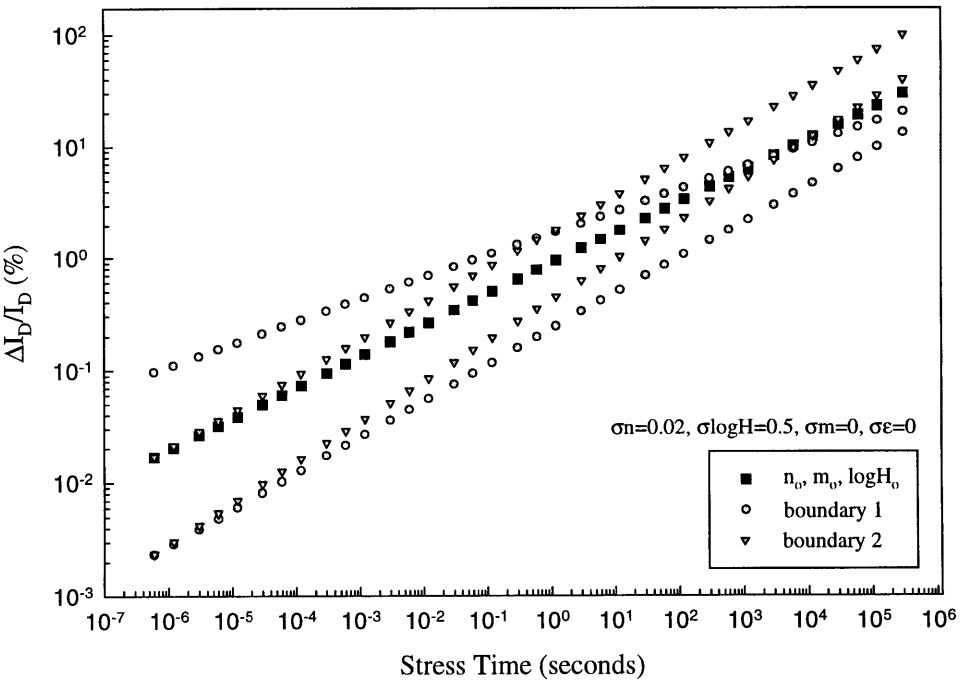


Figure 4.1.7. Illustration of σn and σlogH both on

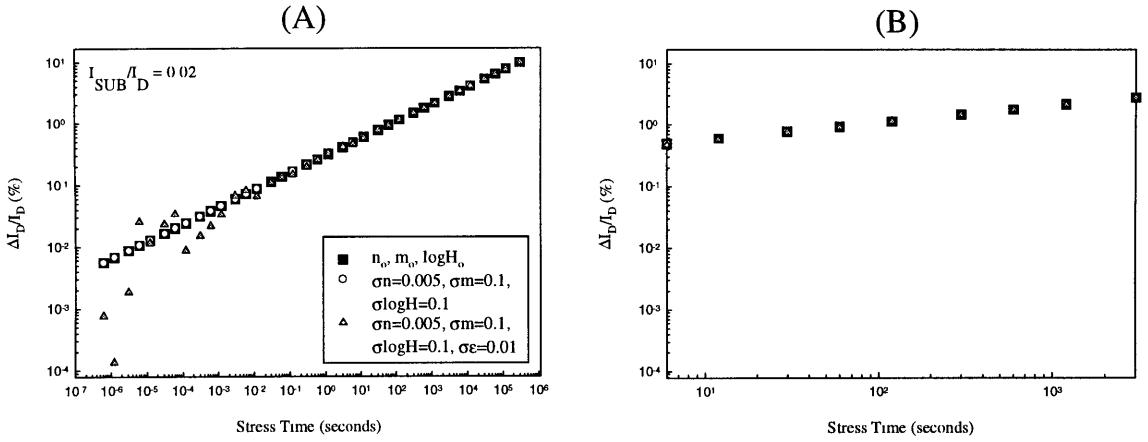


Figure 4.1.8. (A) Illustration of σn, σm, σlogH, σε turned on, (B) Zoom of (A) for time sequence normally used to collect sample data

4.1.3 Effects of Altering the Technology and I_{SUB}/I_D Bias Variables

Changing the technology and bias parameters affects the Δ* model in a deterministic manner such that boundary curves need not be used in the characterization. Both involve a change in the nominal curve. Figure 4.1.9 shows that an increase in I_{SUB}/I_D shifts the nominal degradation curve upward. The subsequent curves at the higher I_{SUB}/I_D level are spaced less apart than those in the lower level due to the log scaling. A change in either m_0, $logH_0$, w, or I_D also shifts the vertical intercept. Both w and $logH_0$ have inverse effects on Δ*, while m_0 and I_D have direct effects, assuming changes in I_D do not impact I_{SUB}/I_D. This assumption is justifiable if any changes in I_D are lumped with I_{SUB} such that I_D remains constant.

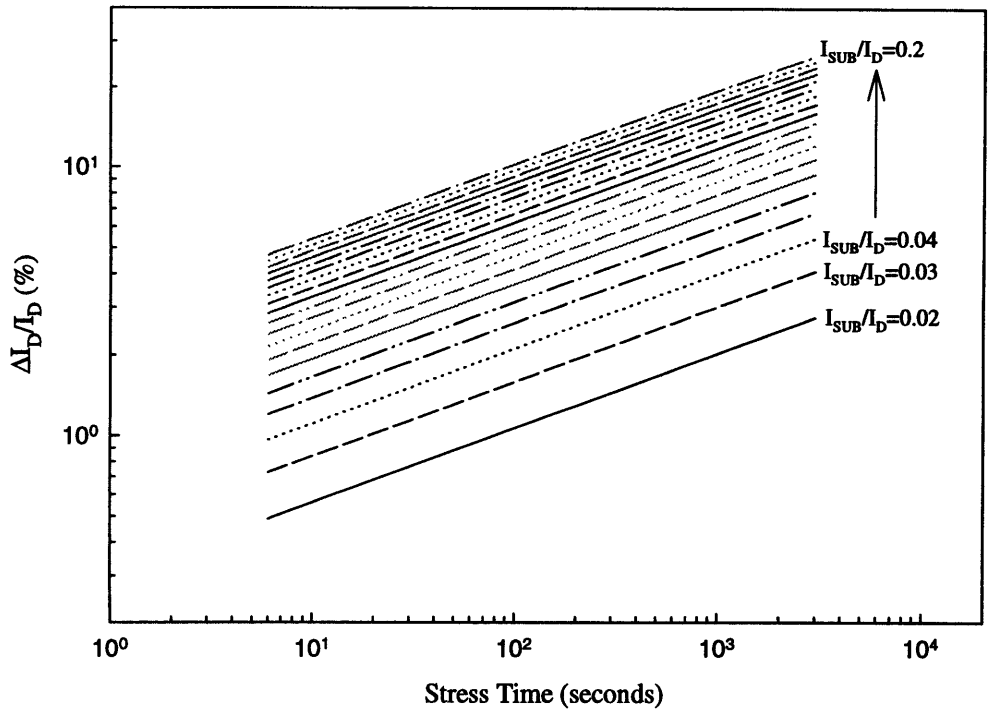


Figure 4.1.9. Effects of changing I_{SUB}/I_D bias level on Δ model

4.2 Effects on Empirical Lifetime Distribution

Figure 4.2.1 shows that at each I_{SUB}/I_D level of the lifetime-correlation plot, an empirical τ distribution can be obtained from a set of simulation trials. Larger number of trials yields better resolution for that distribution, whose shape is dependent upon the choice of σn, σm, σlogH, and σε. The focus of this section is to characterize this lifetime distribution at a particular I_{SUB}/I_D level and examine its response to changes in the probabilistic, technology, and optimization variables. Each distribution in this analysis can be modeled according to (4.7):

$$\tau^* = (\Delta^*)^{1/N^*} \cdot \left(\frac{I_{SUB}}{I_D}\right)^{-M^*} \cdot \frac{w \cdot H^*}{I_D} \tag{4.7}$$

where Δ* accounts for the σε variation according to (3.1). Furthermore, (4.7) asserts that the empirical lifetime distribution reduces to a singular point in the case of no σs turned on. Since (4.7) is a function of four random variables, each normally distributed, an exact

closed form function describing the τ distribution cannot be found; hence, the Monte Carlo approach is used to derive an empirical solution.

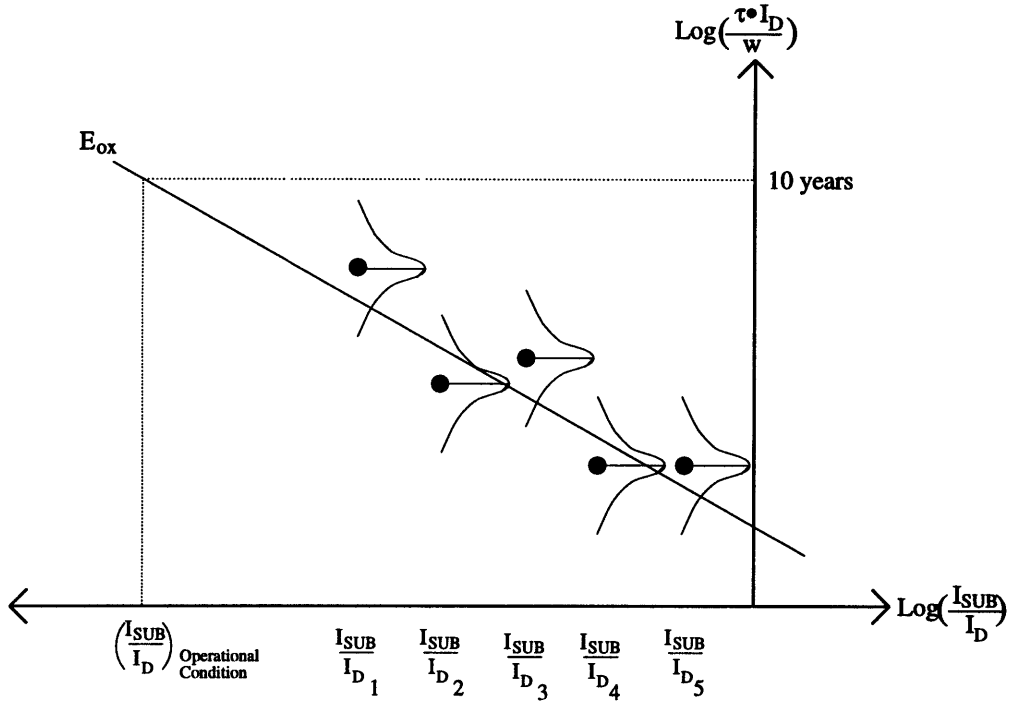


Figure 4.2.1. Illustration of empirical τ distribution at each I_{SUB}/I_D of lifetime-correlation plot

4.2.1 Effects of Turning On Only One σ

Figure 4.2.2 to Figure 4.2.4 show the empirical τ distribution for the cases of varying only σn, σm, or $\sigma logH$. The simulation parameters are equivalent to those used in characterizing the degradation plots with the exception of the stress time sequence, I_{SUB}/I_D level, and σ ranges. The stress time sequence has been reduced to a range more appropriate for experimental measurements (from 0.1 to 50 minutes). The I_{SUB}/I_D level is chosen halfway between the range observed from sample data (0.06 from range of 0.02 to 0.08). The σ ranges have been reduced toward small values, since high values result in an excessively large spread of τ, which makes comparison of τ distribution for small and large σ values not feasible on the same graph.

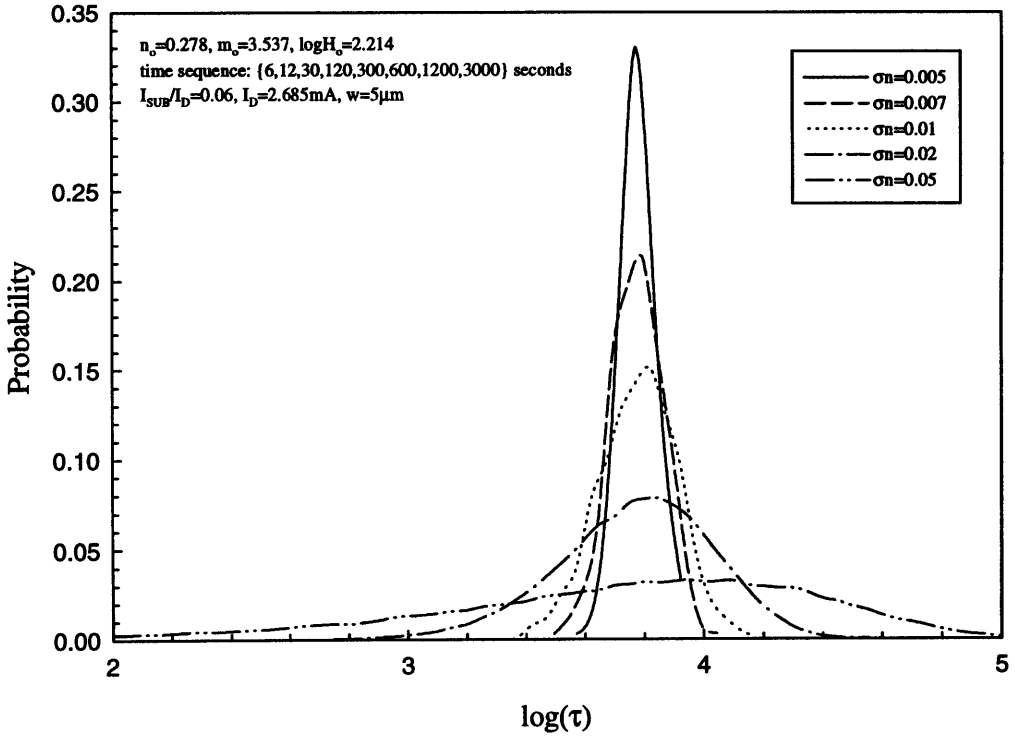


Figure 4.2.2. Lifetime distributions with only σn varying

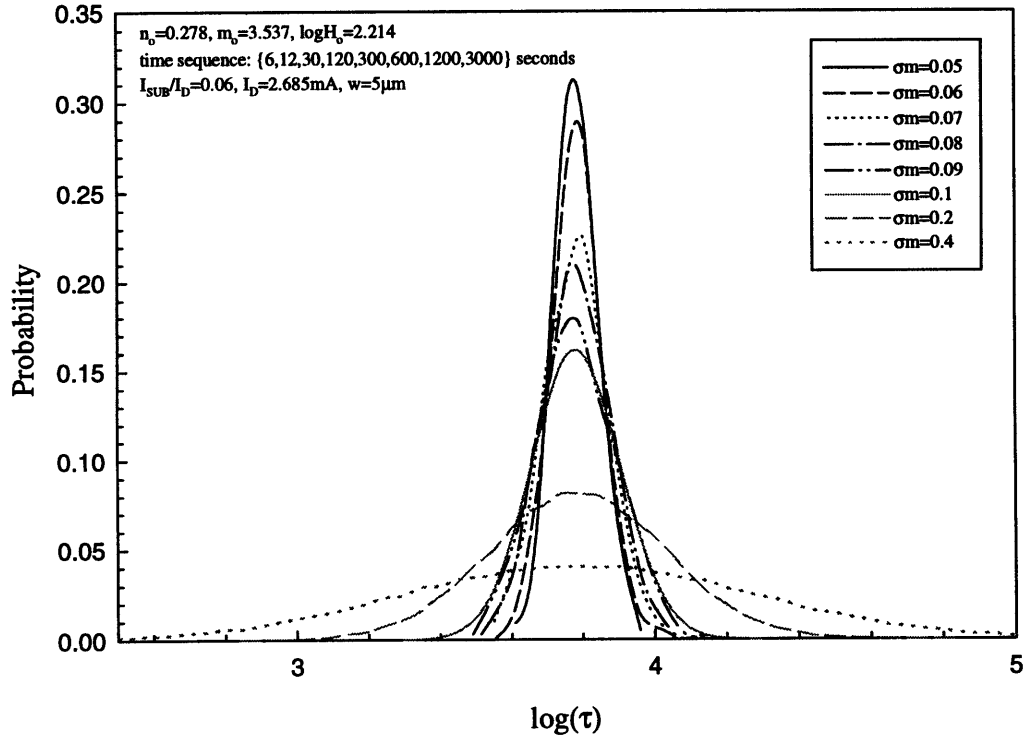


Figure 4.2.3. Lifetime distributions with only σm varying

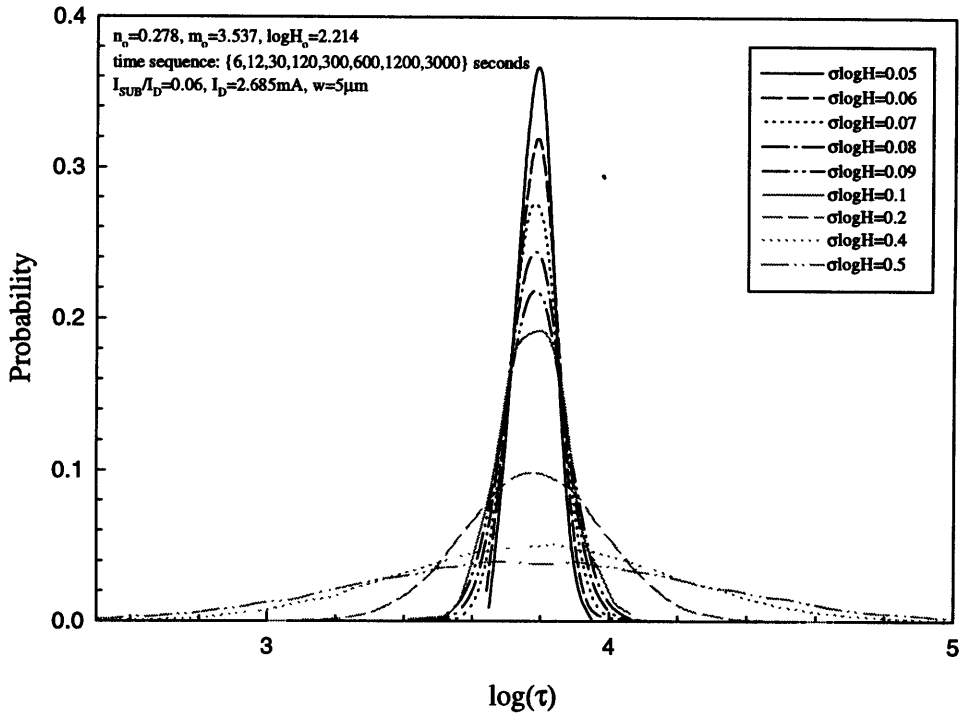


Figure 4.2.4. Lifetime distributions with only σlogH varying

Each lifetime value of a distribution is calculated according to (2.4) in the Monte Carlo simulation. The nonlinear nature of (2.4) causes the mean and median of the distributions to change for a given range of σn, σm, or σlogH values. Furthermore, the mean becomes a function of the σ value. This nonlinearity makes predicting the lifetime distribution counter intuitive and becomes less analogous to the linearization of the Δ* model in the degradation plots. However, consistency can be maintained by characterizing the lifetime distribution based on log(τ) by linearizing (2.4) to:

$$\log\tau = \frac{\log\Delta_{life} - \log A}{n} \tag{4.8}$$

As evident from Figure 4.2.3 and Figure 4.2.4, the mean and median of each distribution remain constant and no longer depend on the σ values. (4.8) also shows that the mean of log(τ) is a linear combination of the mean of σm or σlogH. For example in the case of only σm varying,

$$\mu(\log\tau) = \frac{\log\Delta_{life} - \mu(\log A)}{n_o}$$
$$\mu(\log A) = n_o \cdot \log\left(\frac{I_D}{w \cdot H_o}\right) + n_o \cdot \mu(M^*) \cdot \log\left(\frac{I_{SUB}}{I_D}\right) \tag{4.9}$$

where $\mu(M^*)=m_0$ and all other quantities are constant. Thusfar, the mean of log(t) can be determined if at most only one probabilistic variable is present, by applying the following linear rule of expected values [28]:

$$E(aX + b) = a \cdot E(X) + b \tag{4.10}$$

where operator E is the expected value (or mean) of the independent random variable X, and a and b are constants.

The mean of $\log(\tau)$ cannot be linearly determined from (4.8) for the case of σn varying since logA has **n** dependence. This scenario is analogous to having two or more probabilistic variables present. Figure 4.2.2 shows that for σn of 0.02 and below, the $\log(\tau)$ distributions appear to be gaussian with approximately constant means. But at larger σn values, the distributions' shape becomes more asymmetrical as the lower tail is wider than the upper. This is mostly attributed to a greater percentage of lifetime below the mean, which decreases with increasing σn. This nonlinear nature of $\log(\tau)$ becomes most apparent at large σn values.

Figure 4.2.3 and Figure 4.2.4 agree with intuition and are in accordance with (4.8). As σm or $\sigma \log H$ becomes larger, the mean and median of the $\log(\tau)$ distribution remain constant. The distribution is also gaussian because both M* and logH* are distributed normally. For σm varying, the mean of $\log(\tau)$ is given by (4.9). For $\sigma \log H$ varying, the mean is:

$$\mu(\log\tau) = \frac{\log\Delta_{life} - \mu(\log A)}{n_o}$$
$$\mu(\log A) = n_o \cdot \left(\log\left(\frac{I_D}{w}\right) - \mu(\log(H^*))\right) + n_o \cdot m_o \cdot \log\left(\frac{I_{SUB}}{I_D}\right) \tag{4.11}$$

where $\mu(\log(H^*))$ is given by the following expression:

$$\mu(\log(H^*)) = \log H_o - \log((\Delta_{life})^{1/n_o}) \tag{4.12}$$

log(H*) should not be confused with logH* as the former takes the log of the H* value, and the latter is the normally distributed random variable of Equation (3.2). The means from both (4.10) and (4.11) are the same and verified in Figure 4.2.3 and Figure 4.2.4.

These three case studies reveal which σ values within the ranges from Table 3.1.1 result in lifetime distributions exhibiting nonlinear properties or showing greater sensitivity toward changes in per unit σ. Using the distribution's tightness as one measure of this sensitivity, both σm and σlogH distributions exhibit considerably less sensitivity toward changes in σ of the same magnitude than σn. For instance at a σ value of 0.05, the σn distribution already exhibits considerable asymmetry. This trend is consistent with Section 4.1.1 in that greater influence by σn on the degradation curves manifests into more sensitive distributions.

4.2.2 Effects of Turning On Multiple σs

Since σε is an implicit variable within the Monte Carlo algorithm, the mean of log(τ) distribution cannot be described analytically like σm and σlogH. Its effect on both the n and logA terms of (4.8) causes the mean of increasing σε distributions to remain constant whereas the peaks shift below the mean. Figure 4.2.5 shows this property, and the distributions of Figure 4.2.5 are obtained with σn, σm, and σlogH set to 0.005, 0.05, and 0.05 respectively. The case of varying only σε is not as meaningful since σε is normally distributed about the σn, σm, and σlogH curve. Between the σε ranges of 0.01 to 0.1, the distributions show minimal differences, which are manifested in changes of the height and slight shifting of each peak below the mean. However, the distribution remarkably departs from the group in both shape and height for σε beyond 0.1. Some asymmetry can also be observed.

Figure 4.2.6 shows another distribution with increased σn, σm, and σlogH (black dash-dot-dot line) along with the distributions of Figure 4.2.5. This distribution has σε of 0.1 but with σn of 0.01, σm of 0.1, and σlogH of 0.1. Despite the doubling of σn, σm, and σlogH, this distribution has the same mean as all the distributions shown thusfar except for the cases of varying only σn greater than 0.01. Furthermore, this distribution more resembles that of σε=0.5 (black dash-dot), which suggests that increasing the other three σs have similar effects as greatly increasing only σε.

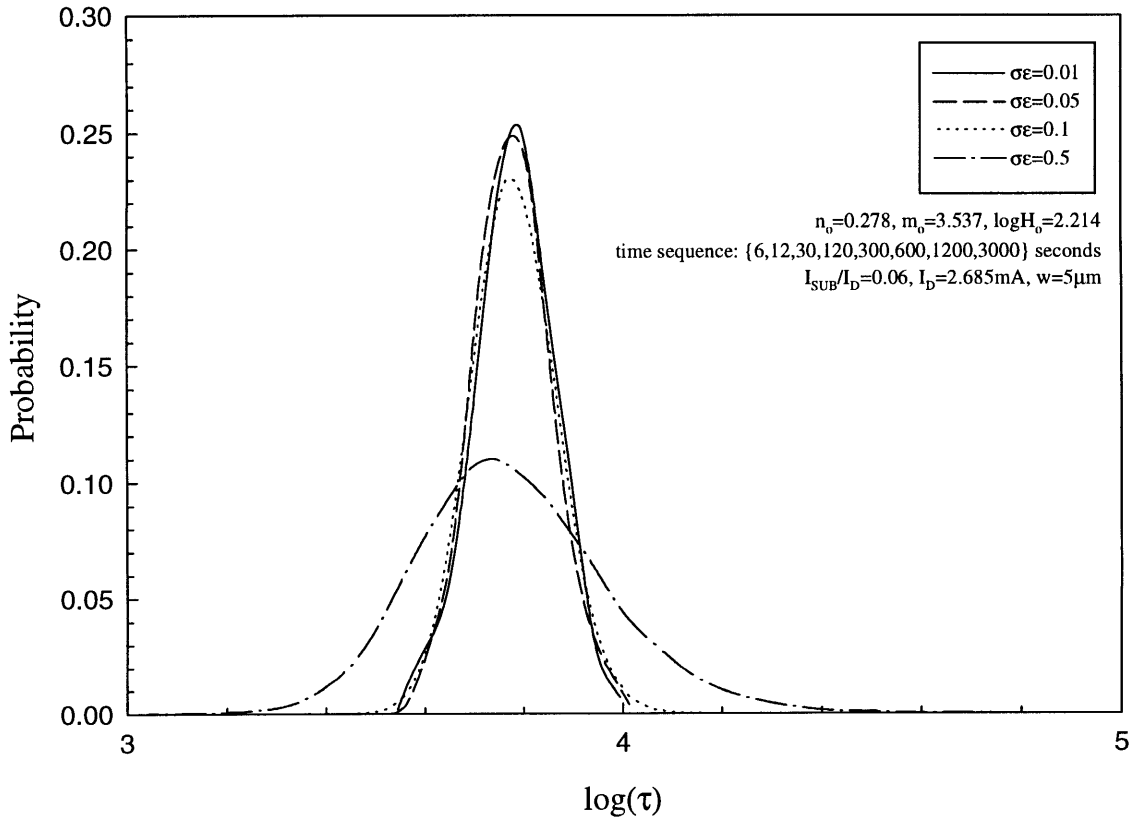


Figure 4.2.5. Lifetime distributions with σn, σm, σlogH, and σε varying

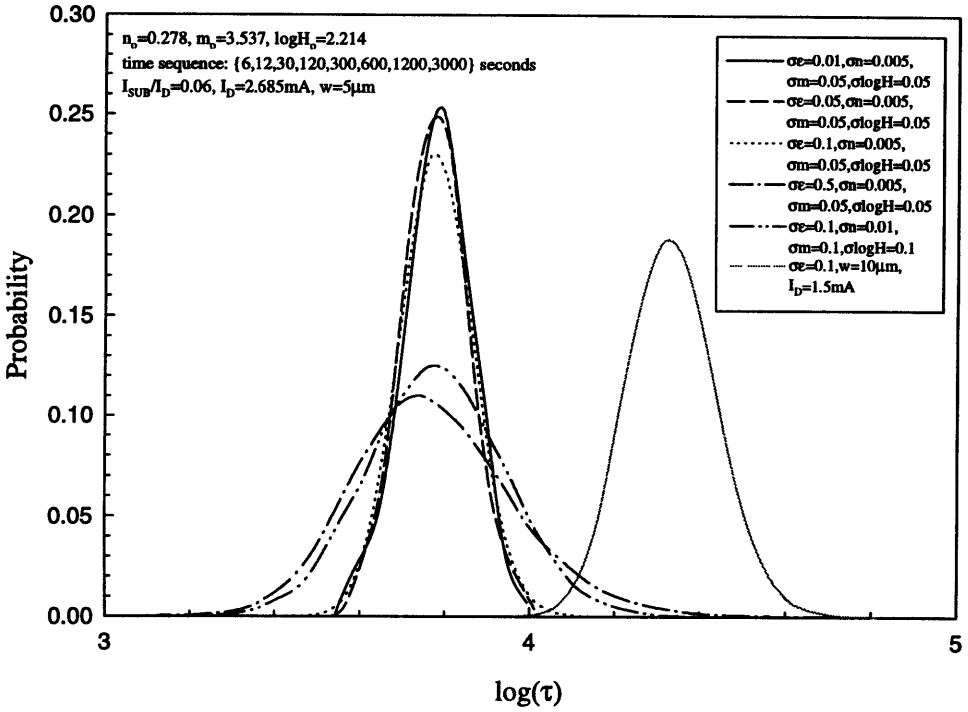


Figure 4.2.6. Comparison of distributions from Figure 4.2.5 with changes in σn, σm, σlogH and w, I_D

4.2.3 Effects of Altering the Technology and Optimization Variables

The focus of this section is to examine what effects changing w and I_D, n_o, m_o, and $logH_o$, the I_{SUB}/I_D levels, and stress time sequence have on the log(τ) distributions. The solid grey curve of Figure 4.2.6 depicts the effect of increasing the channel width to 10μm and decreasing the linear drain current to 1.5mA. The associated σn, σm, and σlogH values are 0.005, 0.05, and 0.05 respectively. Both w and I_D combine to shift the nominal degradation curve downward such that the overall $\Delta I_D/I_D$ level decreases. For the same lifetime definition of 10%, the extrapolated τ values increase which shift the mean upward. Although the peak is less than the counterpart distribution (dotted line), both distributions exhibit the same shape which indicates that changes in the probabilistic variables affect the lifetime distribution in a different manner than the technology variables. The lower peak results from a decrease in the $\Delta I_D/I_D$ level.

As stated in Section 4.1.3, the effects of changing either technology or I_{SUB}/I_D bias parameters cause a shift in the nominal degradation curve, which manifests itself into shifting the mean of the subsequent log(τ) distributions. Figure 4.2.7 and Figure 4.2.8 illustrate the lifetime distributions at different means when different n_o, m_o, $logH_o$ and I_{SUB}/I_D levels are considered. Both figures were generated using the same $\sigma\varepsilon$, σn, σm, and $\sigma logH$. Their shapes remain gaussian and similar to the w, I_D distribution of Figure 4.2.6. Of all the parameters which influence the nominal curve, changing the base technology constants n_o, m_o, $logH_o$ has the most dramatic affect, as the mean shifts orders of magnitude below the other distributions. Figure 4.2.8 shows that as the I_{SUB}/I_D level increases, subsequent peaks also increase because the $\Delta I_D/I_D$ level has increased.

Figure 4.2.9 verifies the intuitive insight one expects the effects of considering different stress time schemes to have on the lifetime distribution. It shows that at low stress times (0.01 - 0.5 minutes), the extrapolated τ values are extremely spread out, resulting in a wide distribution. The poor quality of this distribution compared to the other three in Figure 4.2.9 indicates the large uncertainty in the lifetime extrapolation, which results from large scattering of the Δ^* data points at extremely low stress times. This large scattering is a result of using a low stress time range and further aggravated with a high $\sigma\varepsilon$ value.

The four sets of stress time range are chosen to cover a wide range of time sampling duration. The longest duration involves 0.01 to 5000 minutes and three other ranges as subsets of this: 0.01 to 0.5 minutes, 0.1 to 50 minutes, and 10 to 5000 minutes. All sets have the same time spacing pattern: 1, 2, 5, 10, 20, 50, etc. For the choice of $\sigma\varepsilon$=0.1, σn=0.005, and σm=$\sigma logH$=0.05, only the subsets involving low stress time ranges (0.01 to 0.5 and 0.1 to 50 minutes) exhibit noticeable change in distribution from the other two. Even for the range of 0.1 to 50 minutes, the change is small compared with the lowest subset. However, none of the stress time schemes affects the distributions' mean. The log(τ)

distributions' response to different stress time durations greatly depends on the choice of $\sigma\varepsilon$, σn, σm, and $\sigma \log H$, which will be a focal issue in the optimization methodology of Chapter 5.

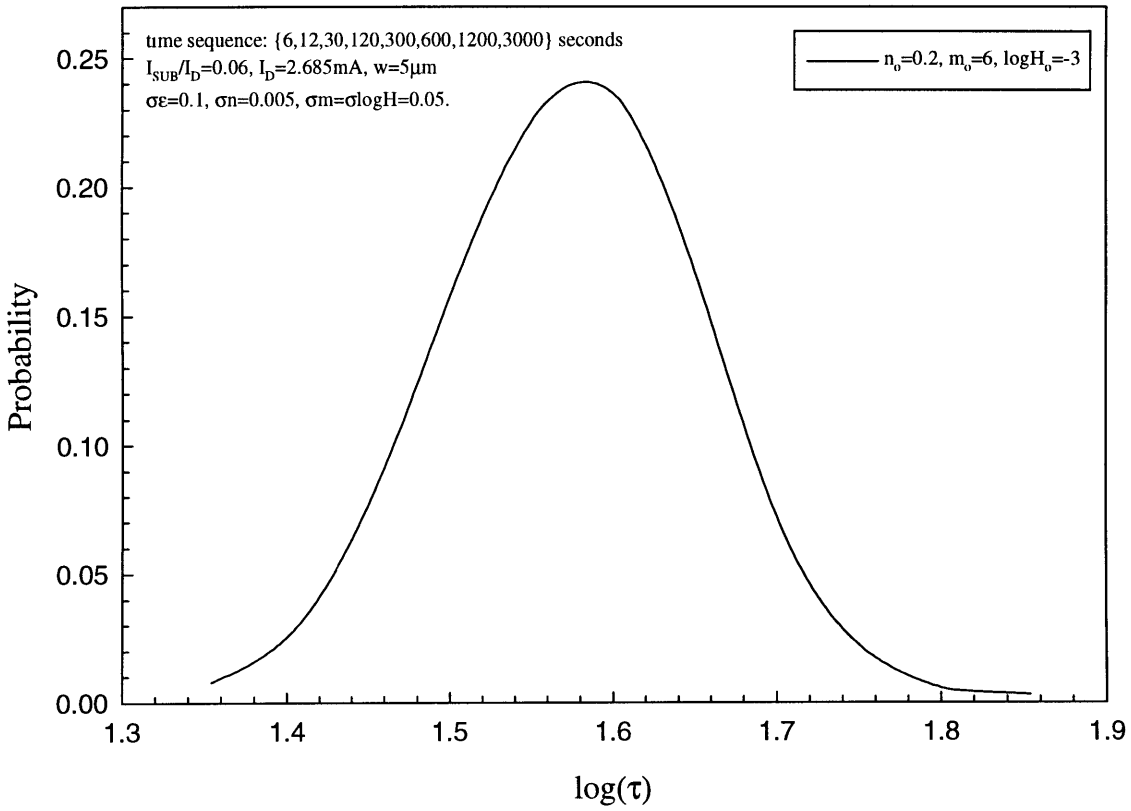


Figure 4.2.7. Lifetime distribution with change in n_o, m_o, $logH_o$

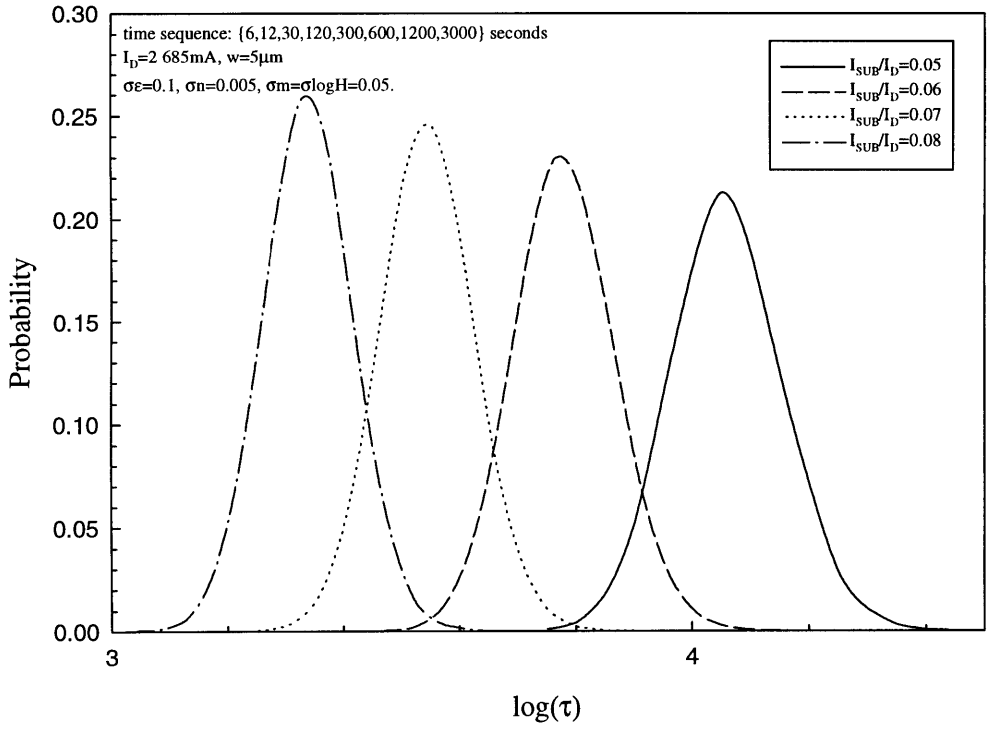


Figure 4.2.8. Lifetime distributions at different I_{SUB}/I_D levels

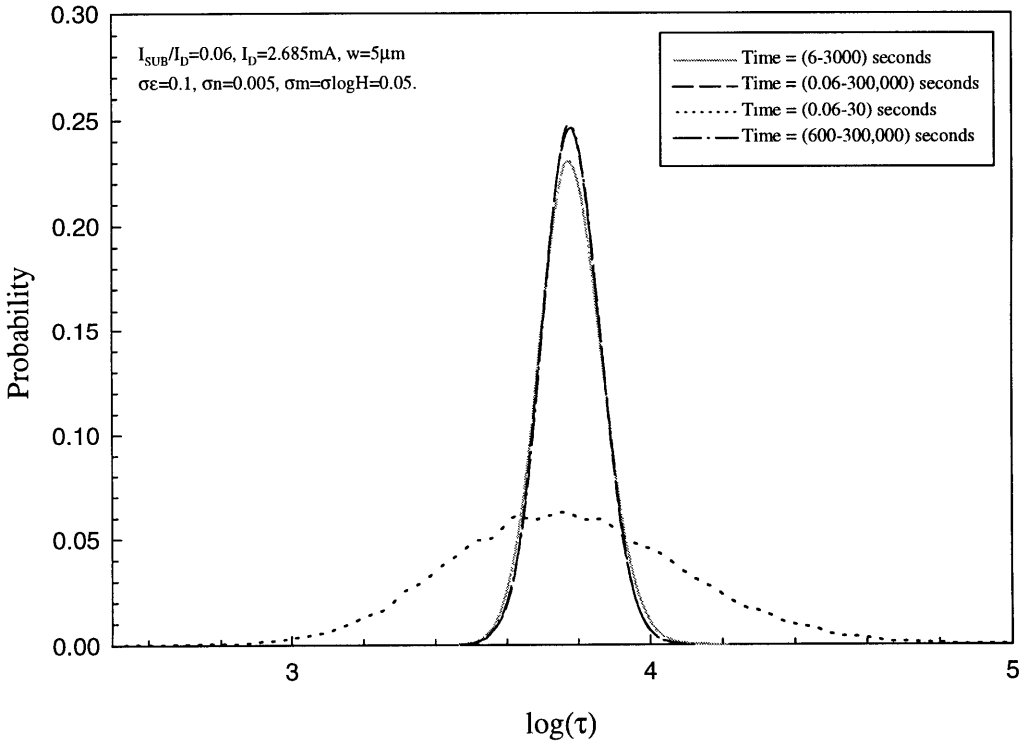


Figure 4.2.9. Lifetime distributions at different stress time sampling

4.3 Effects on Lifetime-Correlation Plot

Section 4.2 analyzes the effects that different parameters have on the lifetime distribution at a fixed I_{SUB}/I_D level. The focus of this section is to examine the response of the lifetime-correlation plot to the same parametric changes but for varying I_{SUB}/I_D values. An I_{SUB}/I_D interval from 0.02 to 0.08 with 0.0067 spacing was used and chosen as inclusive of the values from experimental observation. The assumptions used to generate the lifetime-correlation plots are the same as the ones used to generate the degradation plots and are listed in Table 4.1.1. Since the usual method to extract the **m** and **logH** parameters from the lifetime-correlation plot is linear regression, a regression curve is used to characterize the effects of different σn, σm, $\sigma logH$, and $\sigma\varepsilon$. Larger σ values result in greater scattering of the $(I_{SUB}/I_D, \tau * I_D/w)$ data points, which leads to a less reliable slope and poorer fitting coefficient. The regression fit used in subsequent characterizations is determined from 1000 simulation trials as the one which has poor fitting coefficient and poor slope. The degree of poorness for the slope is referenced from that of the nominal. The importance of satisfying both poor slope and poor fitting coefficient is critical because there exists cases in which a high fitting coefficient occurs but with poor slope or accurately extracted slope but with poor fitting coefficient.

4.3.1 Effects of Turning On Only One σ

The case of varying only σn is shown in Figure 4.3.1. The slope of the nominal fit (black line) is used as the reference for subsequent extracted **m** values and also has a fitting coefficient of 1. Each σn curve represents a fit in which both the extract **m** value and fitting coefficient have the greatest deviation from their reference values. Only the lower boundary of the conic section is shown (those whose slope is more negative than the reference m_o of -3.537). The upper boundary (those with slope more positive than the reference) is symmetrical about the nominal line, and both boundaries form the conic section.

Figure 4.3.2 offers a statistical measure of different σs' effect using a 95% prediction interval, which indicates that for repetitive trials of the same σn, subsequent regression fits lie within this interval 95% of the time. Table 4.3.1 summarizes the extracted **m, logH** and fitting coefficient for each case. Figure 4.3.2(A) and (B) show relatively thin prediction intervals for σn up to 0.02 and unrealistically large intervals beyond 0.02. The fitting coefficient for σn=0.02 is nonetheless above 90%, which is highly acceptable. This breakpoint of σn=0.02 is consistent with the conclusion drawn from the analysis of the degradation plot and lifetime distribution.

σn	m	logH	R^2
0.005	-3.804	1.843	0.98
0.02	-4.759	0.584	0.95
0.05	-8.147	-4.709	0.45
0.1	-33.061	-44.585	0.12

Table 4.3.1. Summary of fitting parameters for σn varying

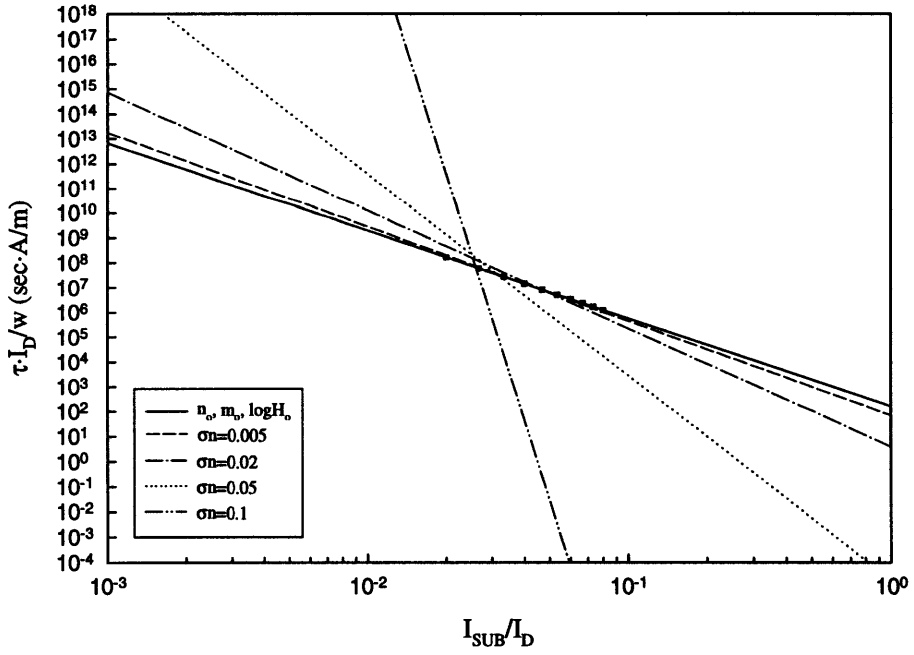


Figure 4.3.1. Effects of varying only σn on lifetime-correlation plot

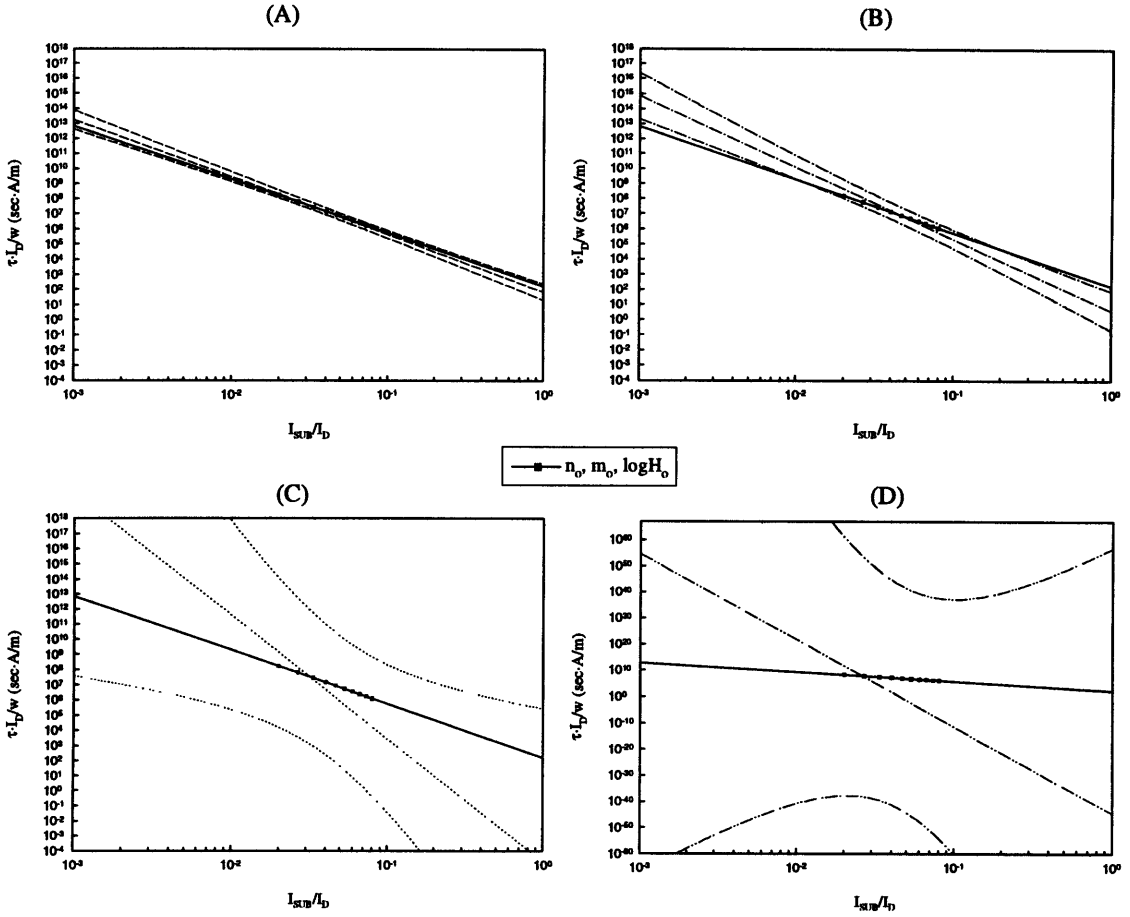


Figure 4.3.2. 95% prediction intervals for σn of (A) 0.005, (B) 0.02, (C) 0.05, (D) 0.1

Figure 4.3.3 to Figure 4.3.6 reveal that the effects of varying either σm or σlogH are not as dramatic as those of σn. Both cases exhibit similar characteristics as the extracted **m** and fitting coefficients are closely related at σm and σlogH beyond 0.1. Once again these similarities are in par with what has been discussed earlier in the degradation and lifetime distribution plots. Both cases indicate that the fit becomes unacceptable beyond the break-point of 0.1 as indicated by a large increase in negative slope, poor fitting coefficient, and widening of the 95% prediction interval. Table 4.3.2 and Table 4.3.3 summarize these values.

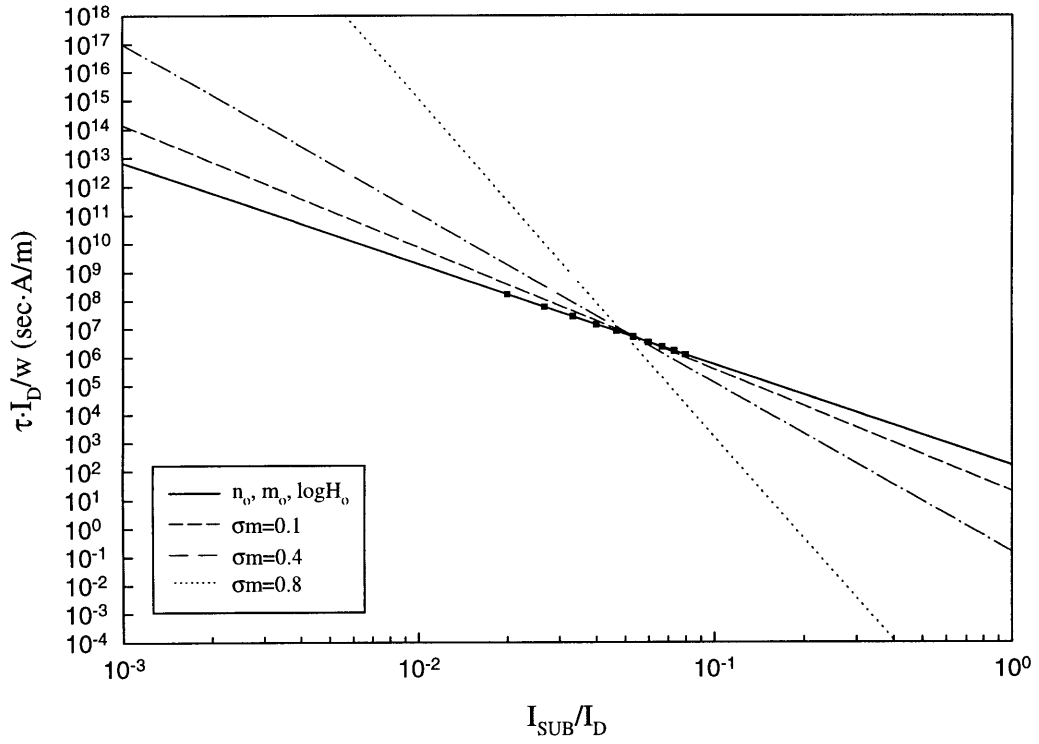


Figure 4.3.3. Effects of varying only σm on lifetime-correlation plot

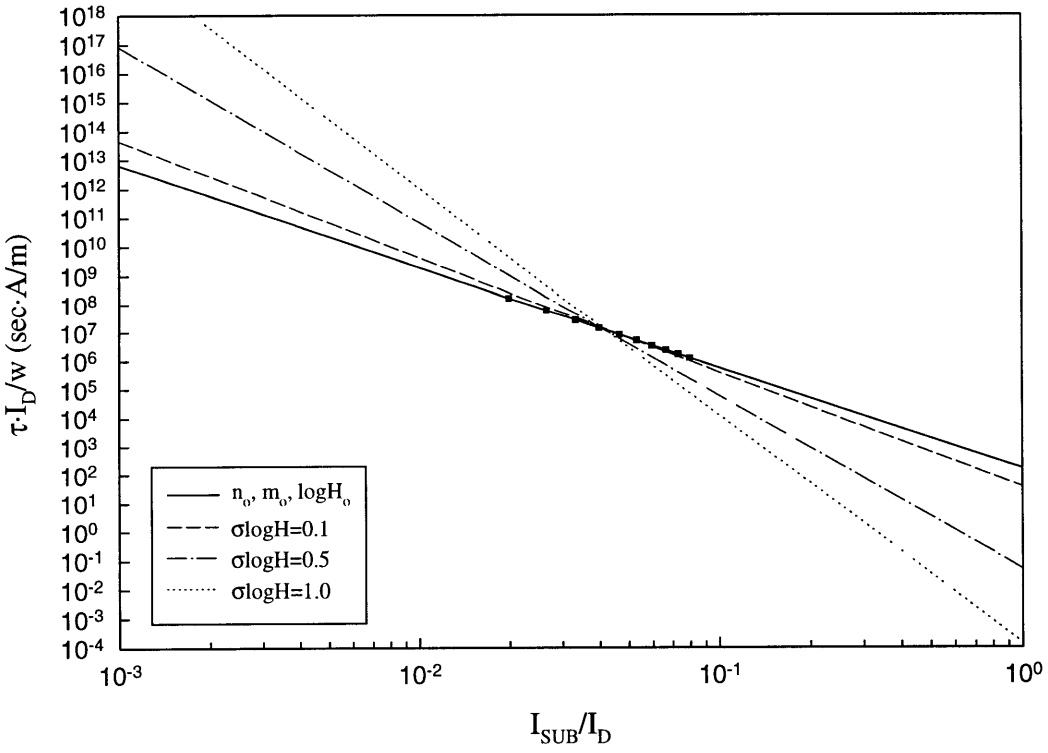


Figure 4.3.4. Effects of varying only σlogH on lifetime-correlation plot

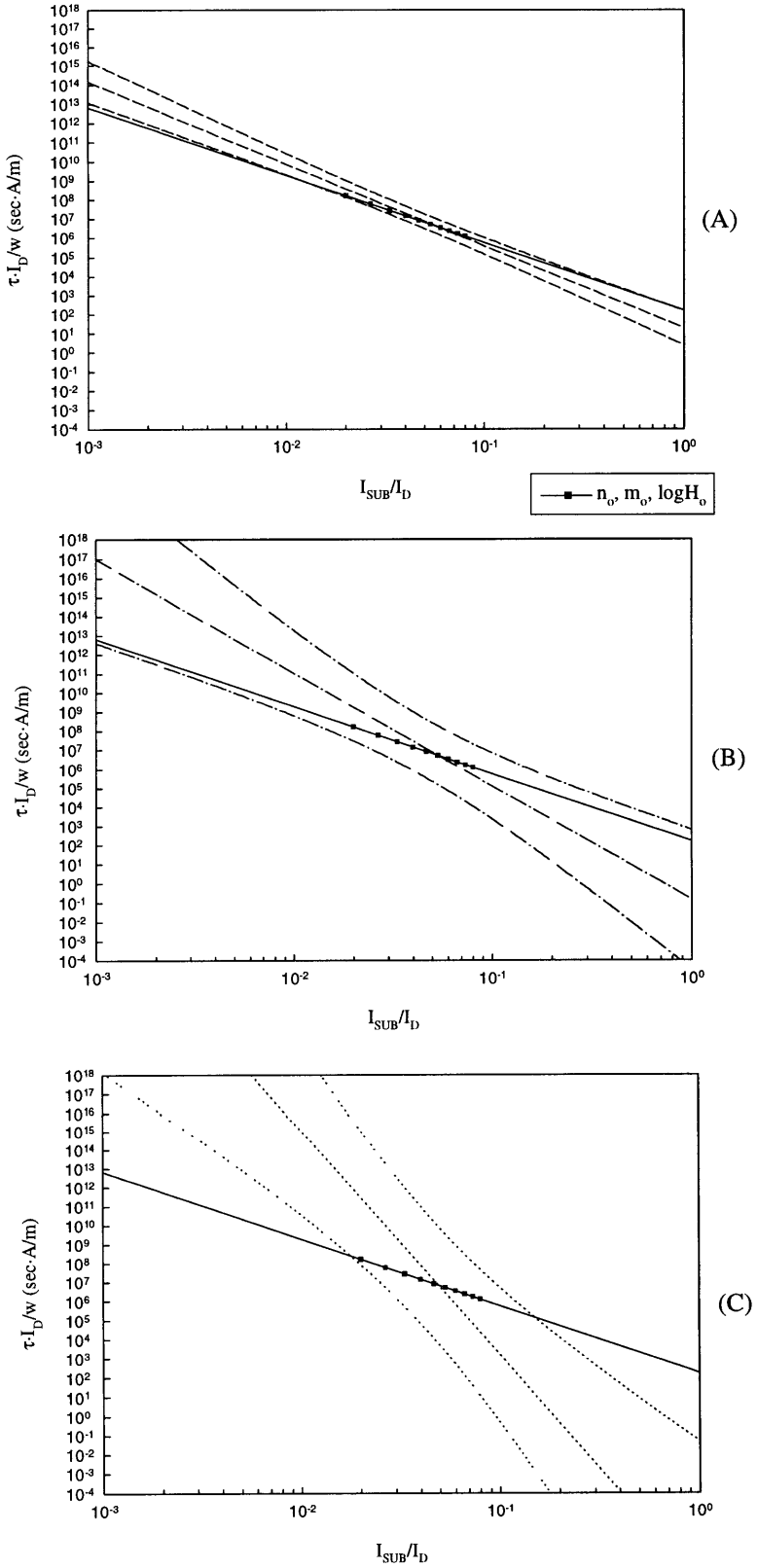


Figure 4.3.5. 95% prediction intervals for σm of (A) 0.1, (B) 0.4, (C) 0.8

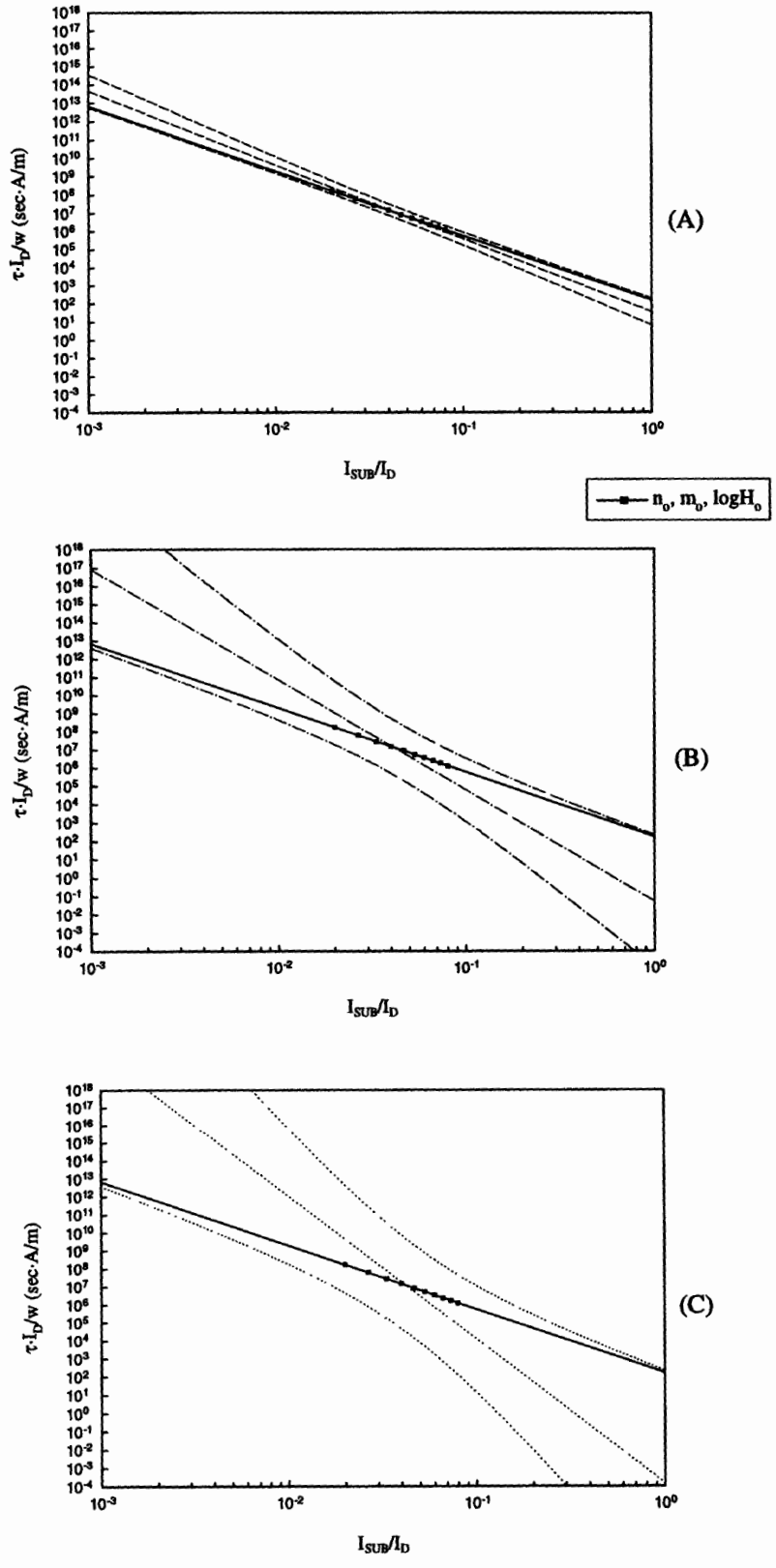


Figure 4.3.6. 95% prediction intervals for σlogH of (A) 0.1, (B) 0.5, (C) 1

σm	m	logH	R^2
0.1	-4.287	1.294	0.97
0.4	-5.940	-0.835	0.77
0.8	-11.848	-8.642	0.77

Table 4.3.2. Summary of fitting parameters for σm varying

σlogH	m	logH	R^2
0.1	-4.034	1.564	0.97
0.5	-6.078	-1.304	0.79
1	-7.947	-3.858	0.68

Table 4.3.3. Summary of fitting parameters for σlogH varying

The effects of adding variation about the nominal degradation curve due to only σε, and how it manifests itself onto the lifetime-correlation plot can be seen in Figure 4.3.7 and characterized by the prediction intervals of Figure 4.3.8. The variations due to σε do not become significant below the value of 0.1 as indicated by the fitting coefficients of Table 4.3.4. Although their prediction intervals are similarly thin, there is significant deviation of the extracted slope from the reference. Only at the smallest value of σε=0.01 does the prediction interval encompass the entire nominal line. At the extreme value of σε=1.0, the poor fit is analogous to the case of σn=0.1. Nonetheless the simulation shows that any choice of σε below 0.1 yields acceptable correlation fits.

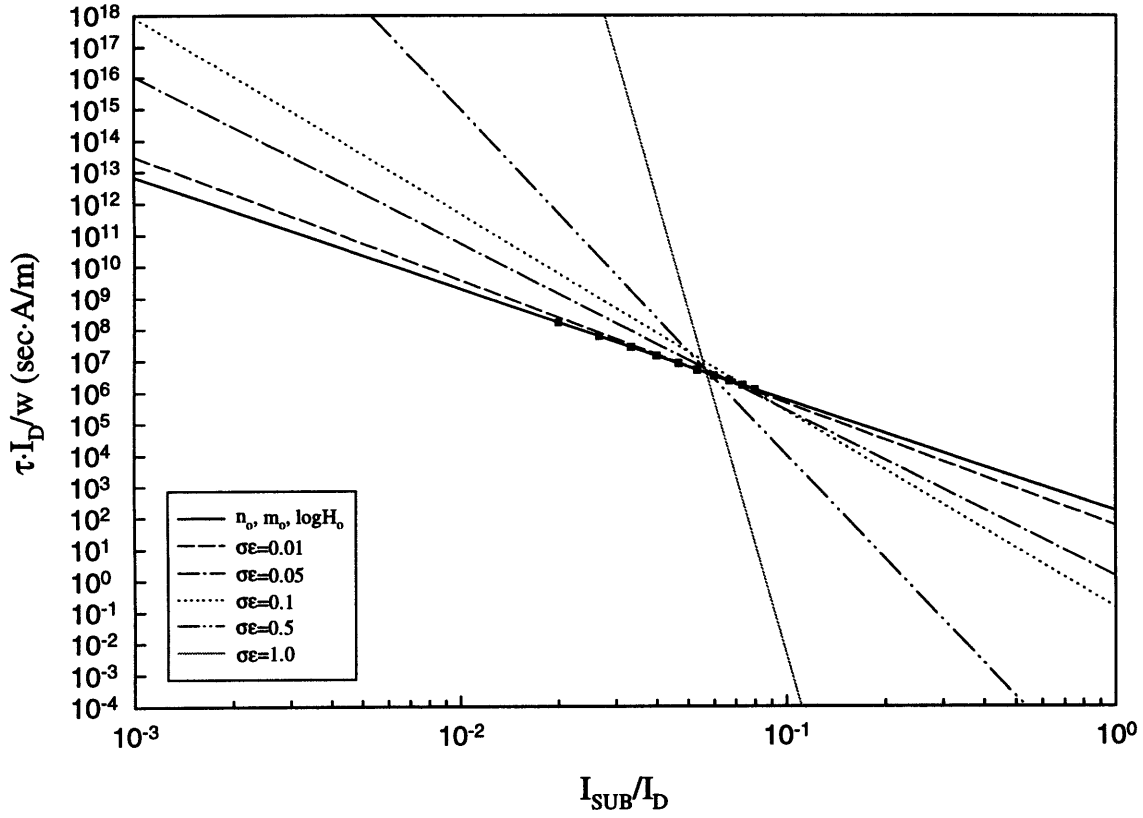


Figure 4.3.7. Effects of varying only σε on lifetime-correlation plot

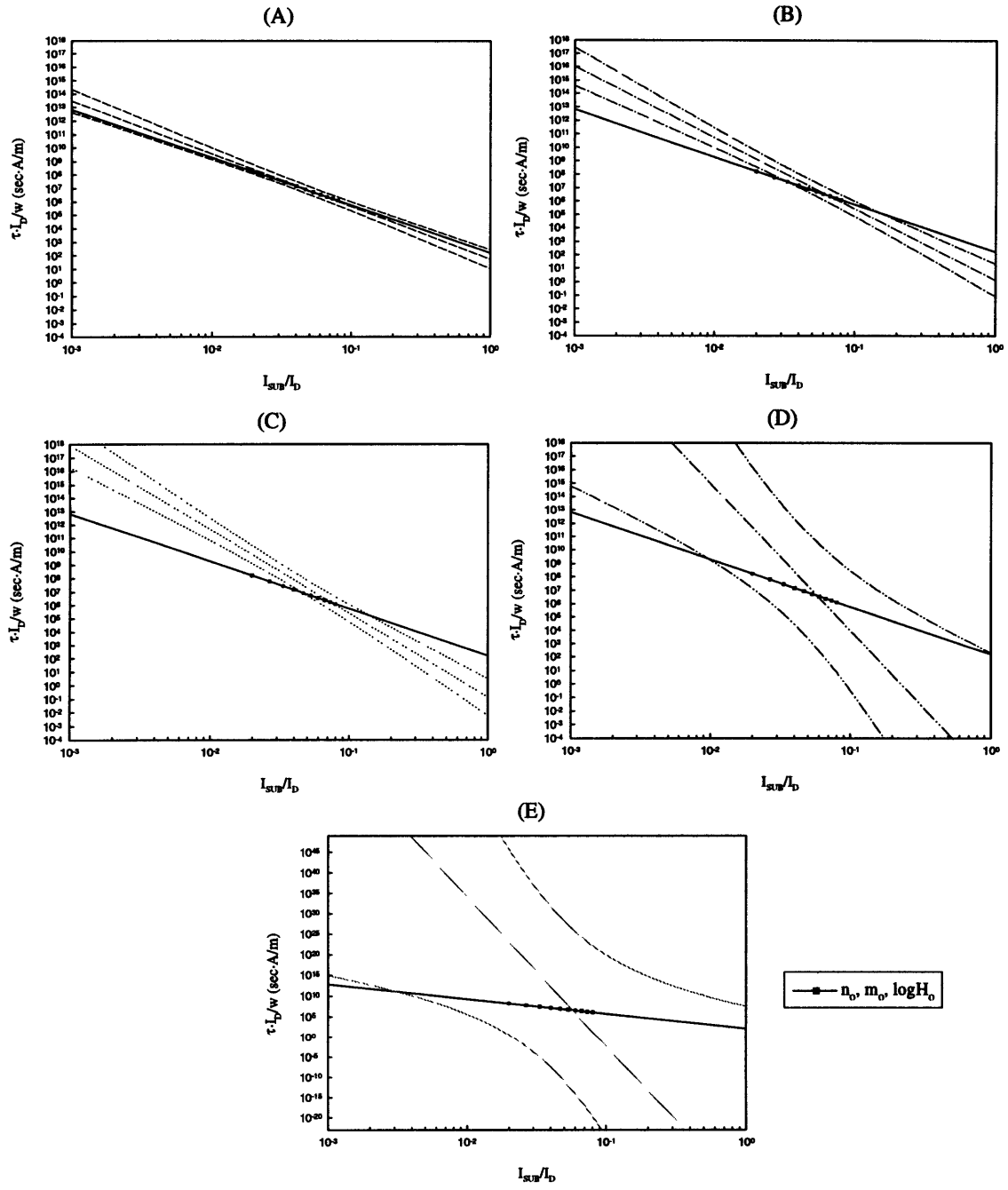


Figure 4.3.8. 95% prediction intervals for σε of (A) 0.01, (B) 0.05, (C) 0.1, (D) 0.5, (E) 1

σε	m	logH	R^2
0.01	-3.914	1.738	0.98
0.05	-5.300	0.128	0.96
0.1	-6.269	-0.870	0.96
0.5	-10.985	-6.995	0.64
1.0	-36.490	-38.839	0.44

Table 4.3.4. Summary of fitting parameters for σε varying

4.3.2 Key Elements in Choosing Values for σn, σm, σlogH, σε in Optimization Methodology

The analysis of this chapter evaluates the impact of various σ values on the three key facets of D.C. parameter extraction and modeling. In the case of turning on only one σ, the breakpoints for σn is 0.02, σm is 0.1, σlogH is 0.1, and σε is 0.1. Values beyond the breakpoints produce curves which largely deviate from the nominal degradation curve, wide lifetime distributions with highly shifted means, and lifetime-correlation curves with poor fits and unreliable extracted slopes. This manifests into unrealistically, unacceptably, and inaccurately generated test cases. Furthermore, simulations which mimic realistic scenarios require simultaneously active σs, whose combined effects further complicate the determination of a set of breakpoints. Setting each σ to its respective breakpoint value results in the above unwanted condition. Setting each to its lowest value from the range of Table 3.1.1 may be conservative; however, Figure 4.3.9 shows that such condition results in simulated data which most resembles experimental data. Further Monte Carlo analysis involving all σs indicates that a relatively tight lifetime distribution and high fitting coefficient (above 90%) can still be attained for values up to σε=0.1, σn=0.005, σm=0.1, and σlogH=0.1. σn has the most adverse impact overall followed by σε while σm and σlogH have equal influence.

The choices for σn, σm, and σlogH are independent of the technology parameters such as w, I_D, n_o, m_o, and $logH_o$. A change in the technology parameters does not require a different set of probabilistic variables with the exception of σε. The change tends to shift the position of the nominal line. If the shift is such that the nominal line lies at low $\Delta I_D/I_D$ levels, noticeable sensitivity can be observed among the three key facets due to high σε. The situation is further aggravated if the I_{SUB}/I_D level and stress time duration are also low. One typical manifestation is that large amounts of scatter appear in the lifetime-correlation plot due to large uncertainty in extrapolated lifetime. This results in a poor regression fit and inaccurately extracted slope. Hence, both I_{SUB}/I_D and stress time must consider the choice of σε in order to avoid the unwanted condition.

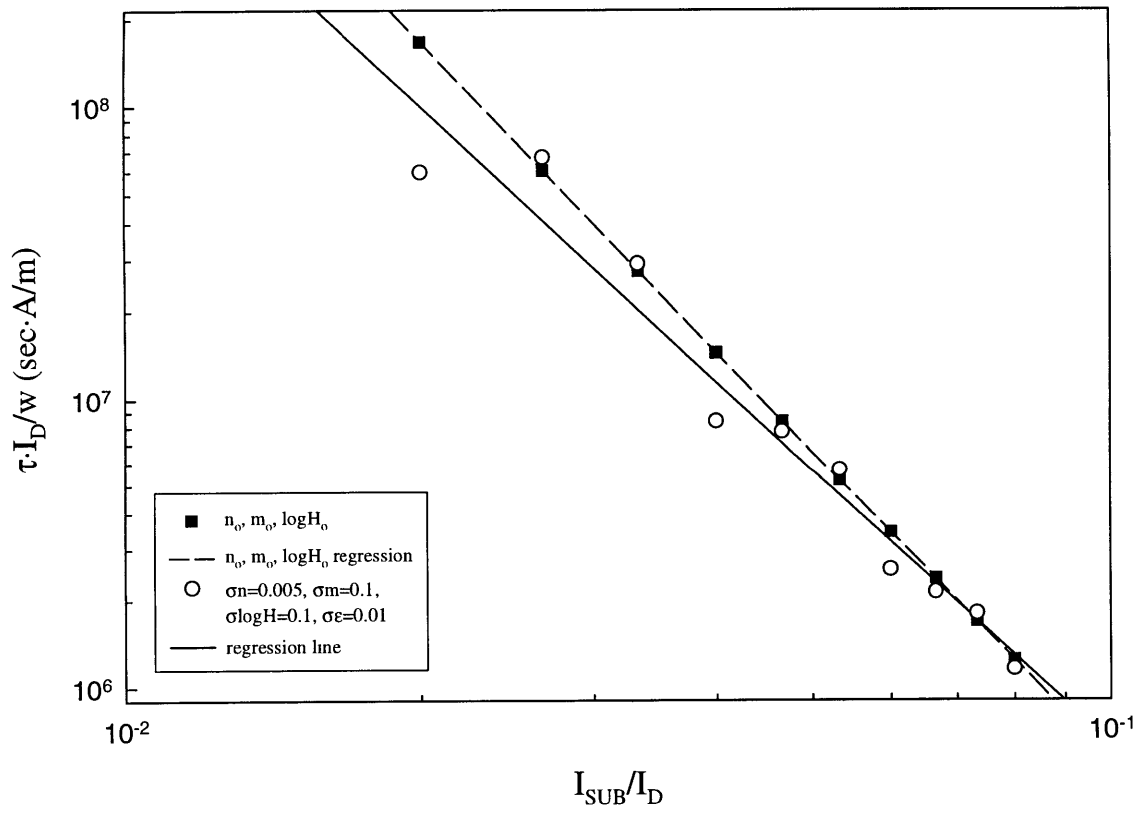


Figure 4.3.9. Representative lifetime-correlation plot from Monte Carlo simulation with σn, σm, σlogH, and σε at conservative values

Chapter 5

Optimization Results Based on D.C. Modeling Issues

The two main optimization issues concerning D.C. modeling are balancing the stress time sequence with the number of device measurements for a given total time constraint and selecting the appropriate V_G and V_D bias conditions within the context of the first issue. It is believed that the dominant issue in locating an optimal region concerns the former; thus, this study focuses on the trade-offs between longer individual device stress times and fewer numbers of device measurements versus shorter stress times and greater numbers of device measurements. From experimental observation, the V_{GD} bias window is quite narrow around -2.5V to -0.5V in order to maintain acceptor-type interface state generation as the dominant degradation mechanism. For a given V_{GD}, the degrees of freedom in selecting individual V_G and V_D values are limited due to the narrow window. The small granularity of V_G and V_D choices does not cause sufficient variation in the I_{SUB}/I_D bias range that affects the individual distributions of I_{SUB}/I_D extrapolated at 10 years. Hence, the optimal region is unlikely to be affected due to the individual distributions' insensitivity. However, it has been found that the stress time and number of device measurements significantly influence the existence and location of this optimal region, which also depends on the I_{SUB}/I_D bias range for a given technology.

5.1 Simulation Methodology

The results of this section was based on the algorithm outlined in Chapter 3 and also summarized in the flowchart of Figure 3.1.1. For each simulation trial, a specific I_{SUB}/I_D was extrapolated at 10 years according to Equation (3.4). Table 5.1.1 summarized the assumptions used in the simulation. Since the bias conditions, V_G and V_D, were not the more important issue in optimizing D.C. parameter extraction, no models were used for

I_{SUB} and I_D. Thus, values for the I_{SUB}/I_D sequence were chosen explicitly based on insight from experimental observations. Linear spacing of this sequence was based on a constant defined by the following expression:

$$k_{Isp} = \frac{(I_{SUB}/I_D)_{max} - (I_{SUB}/I_D)_{min}}{N_{dev} - 1} \tag{5.1}$$

where N_{dev} was the number of device measurements. Conservative values mentioned in Section 4.3.2 were used for the probabilistic variables. These were the lowest values from the σ ranges of Table 3.1.1 and produced simulated degradation and lifetime-correlation curves which best approximated those from experimental measurements.

The optimization of stress time and number of device measurements was based on the two equations of (3.5). A time constraint of 500 minutes had been chosen for the stress duration, T_{total}. Each stress time sequence consisted of four samples with log spacing. Such spacing was favored in experimental extraction procedure due to the log-log scale of degradation and lifetime-correlation plots. A large total time constraint and relatively small number of time samples were chosen in order to simplify the number of variables used to calculate T_{dev}, which was the time taken to perform one device stress measurement. The typical time for the data acquisition instrument to take a measurement, to analyze it, and to record it was approximately one minute. The t_{instr} overhead, which amounted to four minutes, could be neglected from (3.5) since the total time constraint was large. Since each device was subject to the same stress time sequence for a set of simulations, Equation (3.5) was reduced to the following expression which underlined the remainder of this study:

$$T_{total} = N_{dev} \cdot T_{dev} \tag{5.2}$$

In order to obtain an empirically exact distribution for each set of simulations, large number of trials was required for best resolution. Hence each I_{SUB}/I_D distribution was

obtained with 5000 to 50,000 trials. If the I_{SUB}/I_D population had a tight distribution, then the number of trials could be small. However for more ill-defined and wider distributions, large number of trials was required. However, a large number of trials for each distribution was discouraged since it entailed long simulation time. Furthermore for tight distributions, a saturation point existed for the number of trials needed to obtain precision.

n_o	0.278
m_o	3.537
$logH_o$	2.214
I_D(mA)	2.685
w (μm)	5
lifetime	10%
σn	0.005
σm	0.1
σlogH	0.1
σε	0.01
I_{SUB}/I_D sequence	0.02 - 0.08
Time constraint (minutes)	500
Number of time samples	4

Table 5.1.1. Listing of the values used in optimization

5.2 Response of Optimized Parameter-Extraction Method on Balancing Device Stress Time with Number of Device Measurements

5.2.1 Results of Study from above Methodology

Seven combinations of number of device measurements and stress time duration were chosen to represent each I_{SUB}/I_D distribution. The stress time duration coincides with the log spacing scheme, and the product of this duration with the number of device measurements is 500 minutes. Figure 5.2.1 shows that no optimal region exists as the I_{SUB}/I_D distribution becomes tighter with increasing number of device measurements. This seems to indicate that the uncertainty in the extrapolated lifetime does not manifest into greater scattering of the (I_{SUB}/I_D,$\tau*I_D/w$) points. But detailed examinations reveal that although the uncertainty does cause scattering in the lifetime-correlation plot, it is not sufficiently significant to cause a noticeable difference at low stress times. The behavior of the I_{SUB}/I_D distribution for large number of device measurements is analogous to the model used to calculate a prediction interval, which becomes more narrow as the number of sample points increases. Even with a large sample, the interval's width could decrease if the sample points are sufficiently scattered since the variance would increase. Hence, the scatter introduced by the lifetime extrapolation errors at low stress times for the study of Figure 5.2.1 is not the dominant factor. The lack of scatter is due to the conservative choice of $\sigma\varepsilon$, σn, σm, and $\sigma \log H$. Figure 5.2.2(B) shows a typical degradation profile for the stress time considered in this study. The thin black curve pertaining to $\sigma n=0.005$ and $\sigma m=0.1$ illustrates that the curve about which $\sigma\varepsilon$ is distributed has equal variation across the entire stress time range. The small choice of $\sigma\varepsilon=0.01$ further aggravates the problem as any deviation from the $\sigma n=0.005$ and $\sigma m=0.1$ curve becomes insignificant. Hence $\Delta*$ at the high stress times sees the same variation as at the low stress times while greater number of device measurements at low stress times tightens the subsequent I_{SUB}/I_D distribution.

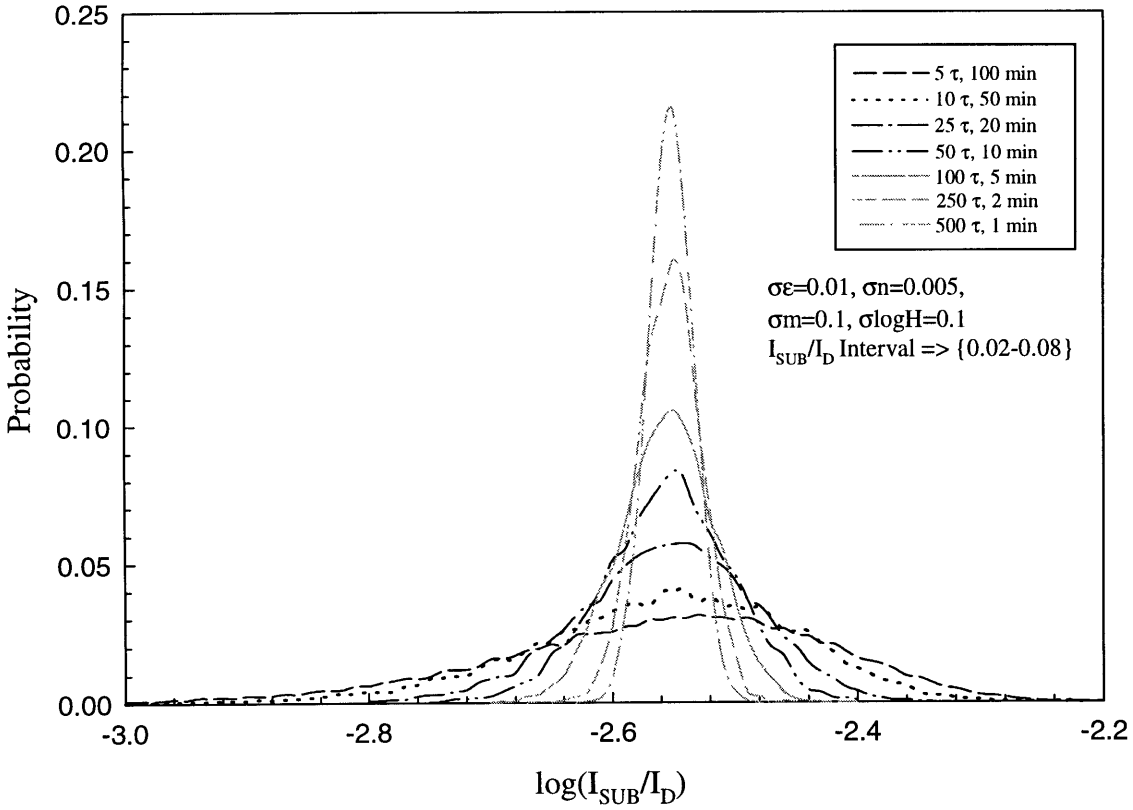
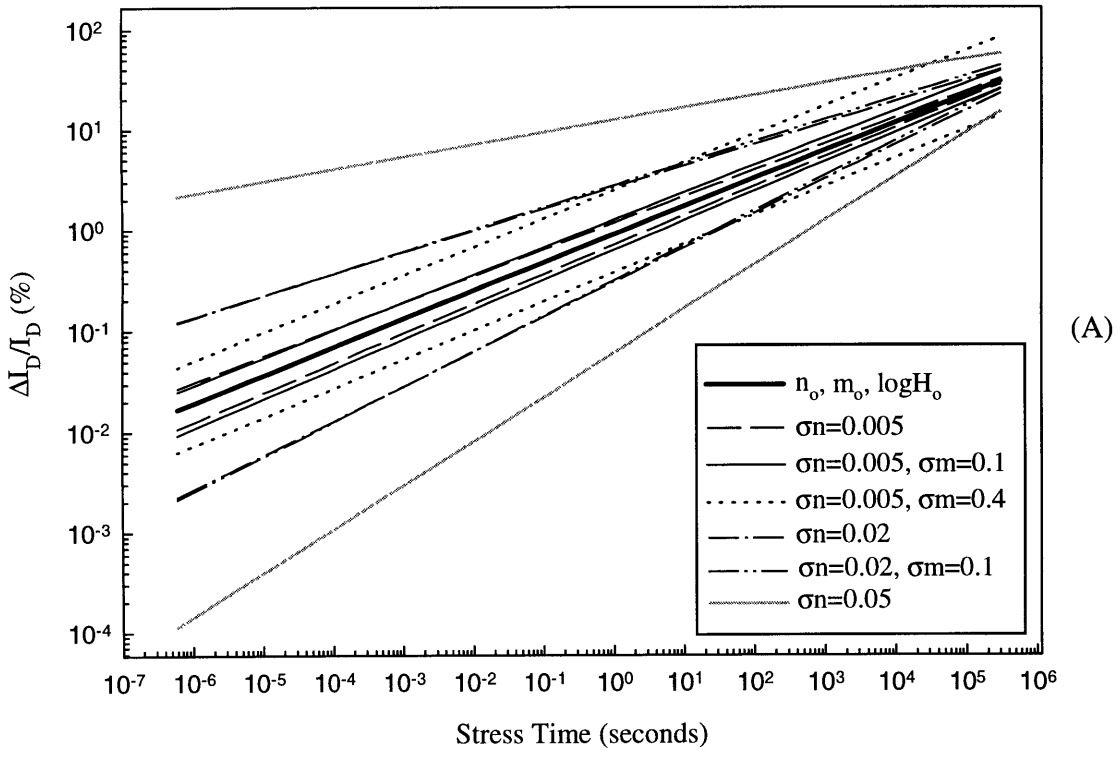


Figure 5.2.1. I_{SUB}/I_D distribution showing results of study using assumptions from Table 5.1.1; τ denotes the number of device measurements

While Figure 5.2.1 shows a $\log(I_{SUB}/I_D)$ distribution, the corresponding I_{SUB}/I_D distribution shows similar results. Its appearance is also gaussian-like revealing the same trend at larger number of device measurements. The reason for using a log distribution even though the linear I_{SUB}/I_D value is extrapolated at 10 years will become apparent in the later discussion.



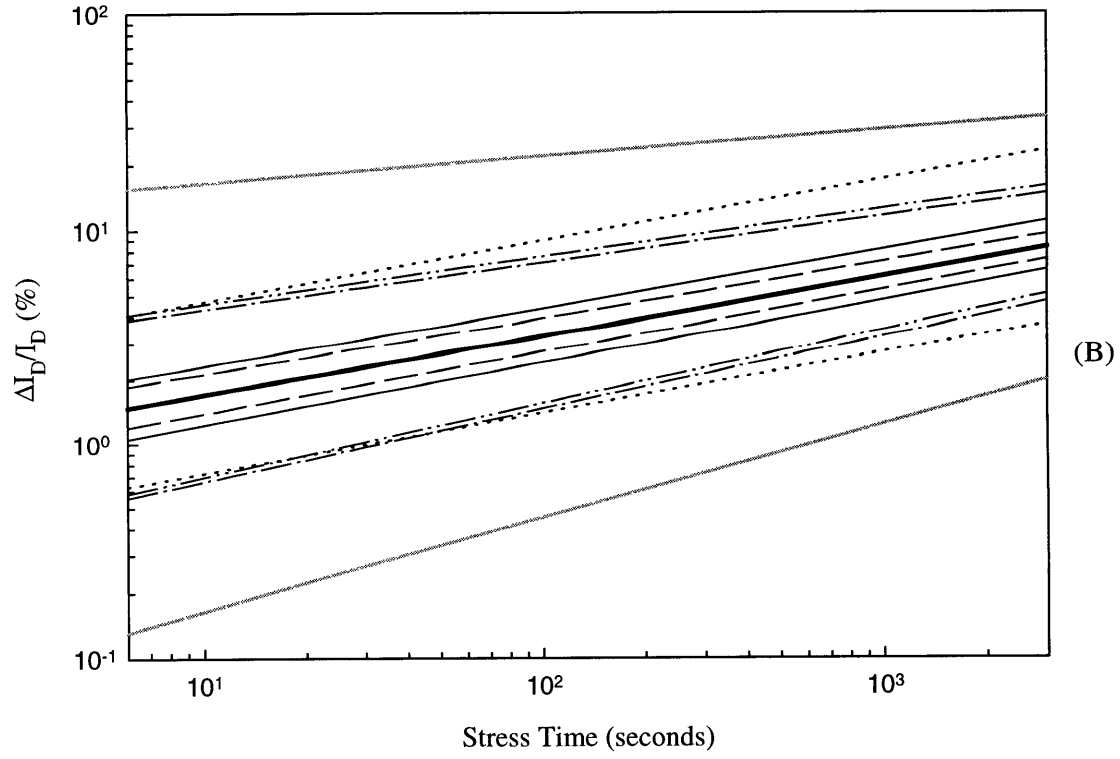


Figure 5.2.2. (A) Effects of different σn and σm on Δ plot for large stress time range, (B) for stress time range considered in the study

5.2.2 Results of Study Upon Use of Different Assumptions

Figure 5.2.1 reveals that the existence of an optimal region highly depends upon the choice of σ used to model the variations. Thus for a particular technology with quantifiable amounts of variation, the methodology outlined thusfar can be used to verify the existence of an optimal region. However it is of great importance to determine under what conditions such an optimal region exist and identify the key elements which affect its existence. The degradation profiles of Figure 5.2.2(A) and (B) can be used to identify prudent choices for the σs which may result in an optimal region. The extended stress time range of Figure 5.2.2(A) clearly illustrates variation in Δ^* across different stress time regimes for various σn and σm values. Figure 5.2.2(B) shows the variational effects at the studied time frame. The case of σn=0.02 and σm=0.1 (dash-dot-dot curve) presents the best candidate for differential variation between low and high stress time regimes, even for the small time range of this study. σlogH is not accounted for due to its adverse unrealistic effects when combined with σn. These effects are shown in Figure 4.1.7. Although σm can also be turned off and all variations attributed to σn and $\sigma\varepsilon$, experimental observation shows a need to account for variations across multiple extraction parameters. A large value of 0.2 is chosen for $\sigma\varepsilon$ in order to cause additional extrapolation error at low stress times. This manifests into additional scatter in the lifetime-correlation plot such that the larger number of measurements has less influence in defining the I_{SUB}/I_D distribution.

The I_{SUB}/I_D bias interval also has been extended in order to examine the optimal region's response to different I_{SUB}/I_D bias conditions. This is accomplished by partitioning the large interval into smaller subintervals. A range of 0.02 to 0.2 is used for the outer interval with subintervals at 0.02 to 0.08, 0.08 to 0.14, and 0.14 to 0.2. The spacing is still in accordance with (5.1). Table 5.2.1 summarizes the new assumptions used for the remainder of the study.

n_o	0.278
m_o	3.537
$\log H_o$	2.214
I_D(mA)	2.685
w (μm)	5
lifetime	10%
σn	0.02
σm	0.1
σlogH	0
σε	0.2
I_{SUB}/I_D range	0.02 - 0.2
Time constraint (minutes)	500
Number of time samples	4

Table 5.2.1. Listing of new values used to determine optimal region

Figure 5.2.3 and Figure 5.2.4 suggest an optimal region exists for the outer I_{SUB}/I_D interval of 0.02 to 0.2 and inner interval of 0.02 to 0.08. The former's optimal region lies around 10 to 50 device measurements at stress times from 10 to 50 minutes. The latter has an optimal region closer to 5 to 10 device measurements at stress times from 50 to 100 minutes. If a log operation is applied to the individually extrapolated I_{SUB}/I_D values and plotted in Figure 5.2.5 and Figure 5.2.6, the optimal regions become visibly clear. The more well-defined optimal region from the outer I_{SUB}/I_D interval lies around 50 device measurements with 10 minute stress duration. The corresponding one for the inner interval lies definitely above 25 device measurements and below 100 device measurements. These boundaries are well balanced around the 50τ distribution.

The asymmetry of the linear I_{SUB}/I_D distribution compresses the width of the optimal region because it is located at a relatively low I_{SUB}/I_D range for both cases. The reason for the asymmetric distributions is based on the nonlinear function used to extrapolate I_{SUB}/I_D at 10 years. This property is analogous to one affecting the lifetime distributions discussed in Section 4.2. The gaussian-appearance of the $\log(I_{SUB}/I_D)$ distributions is consistent with those of the lifetime distribution. The more ill-defined distributions at large number of device measurements (such as 500τ) is due to a lack of resolution from fewer simulation trials.

No optimal region exists for the inner medium and high I_{SUB}/I_D intervals, as shown by Figure 5.2.7 and Figure 5.2.8. The higher I_{SUB}/I_D values place the nominal degradation curve at high $\Delta I_D/I_D$ levels such that a $\sigma\varepsilon=0.2$ has similar effects as that of $\sigma\varepsilon=0.01$ from Section 5.2.1. The exact $\Delta I_D/I_D$ levels can be seen in Figure 4.1.9. Since these nominal curves are spaced relatively closed to one another, the three σs generate degradation curves which highly interact across the different nominal curves. This leads to more gaussian-appearing distribution shapes as the number of device measurements increases. Because these two cases are analogous to that of $\sigma\varepsilon=0.01$, the dominant mechanism which defines the I_{SUB}/I_D distributions is the number of device measurements. Unlike the $\sigma\varepsilon=0.01$ case, the distributions from 5 to 50 device measurements are asymmetric and largely spread out across low I_{SUB}/I_D values.

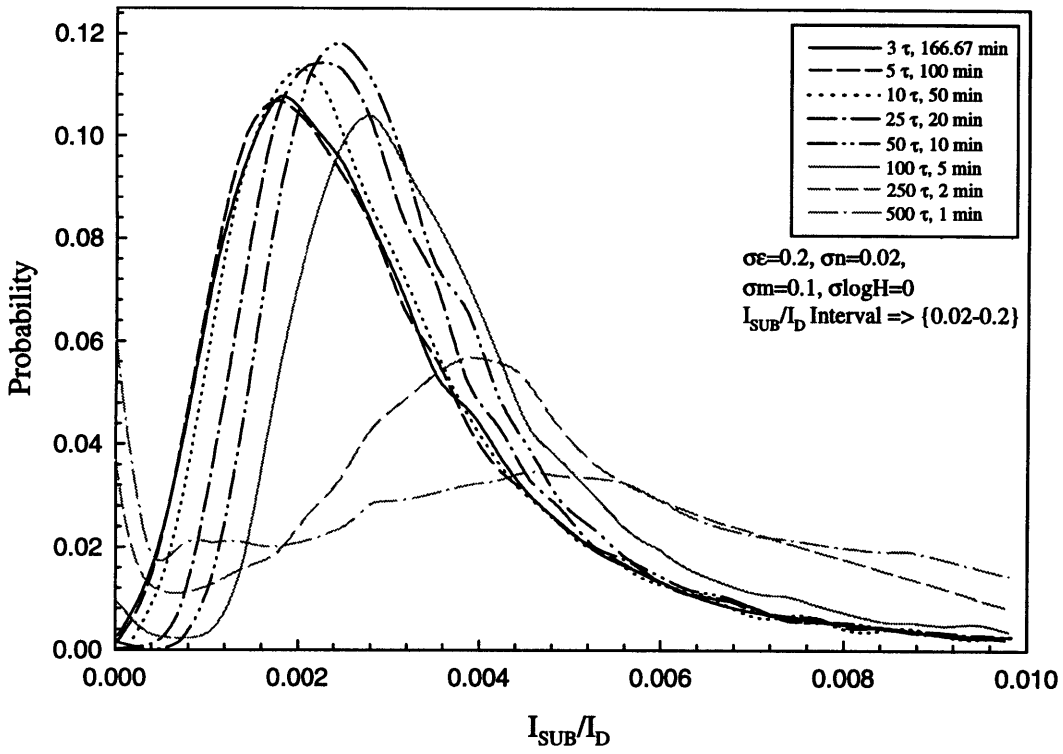


Figure 5.2.3. I_{SUB}/I_D distributions showing optimal region for outer I_{SUB}/I_D interval

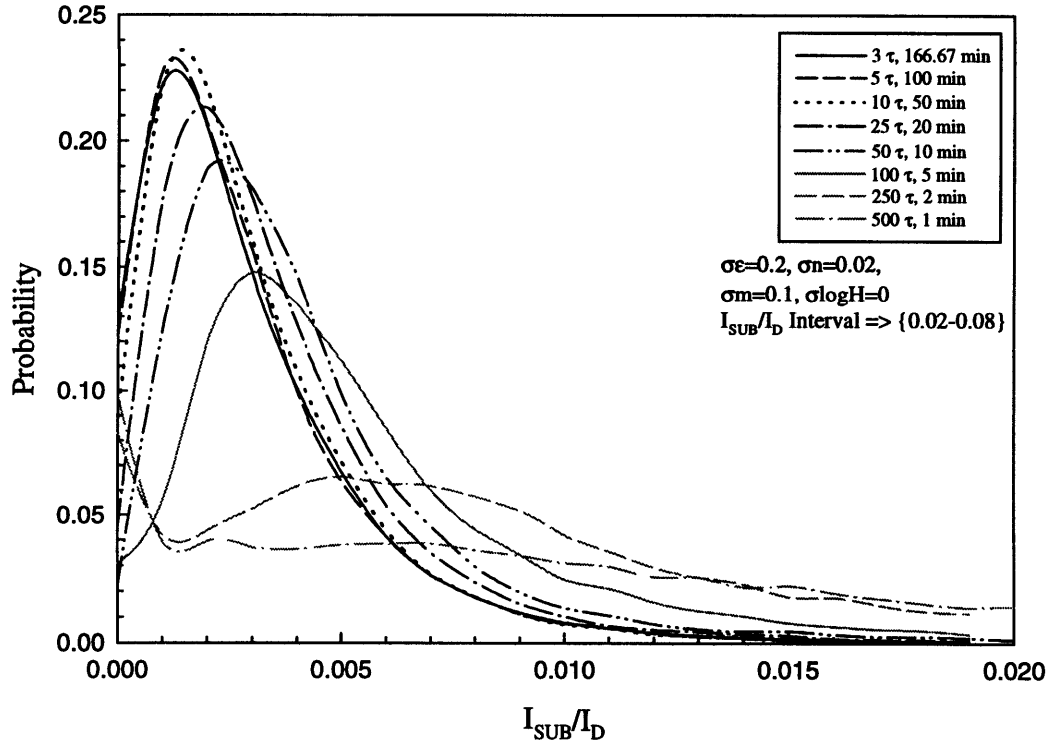


Figure 5.2.4. I_{SUB}/I_D distributions showing optimal region for inner low I_{SUB}/I_D interval

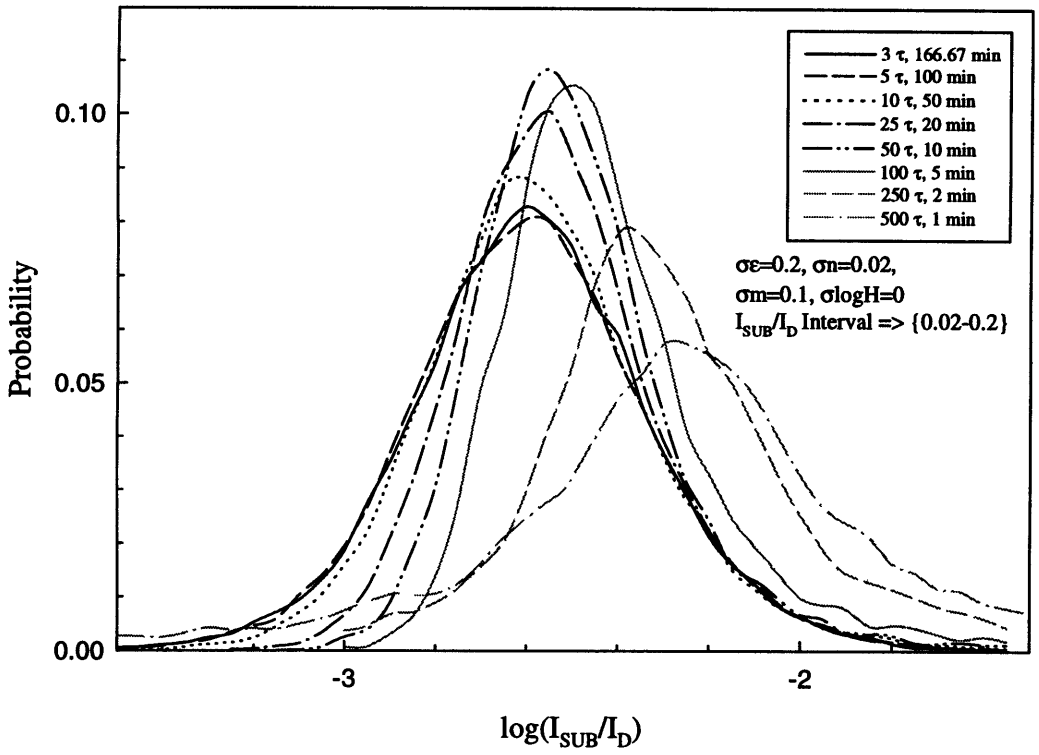


Figure 5.2.5. Log plot of Figure 5.2.3

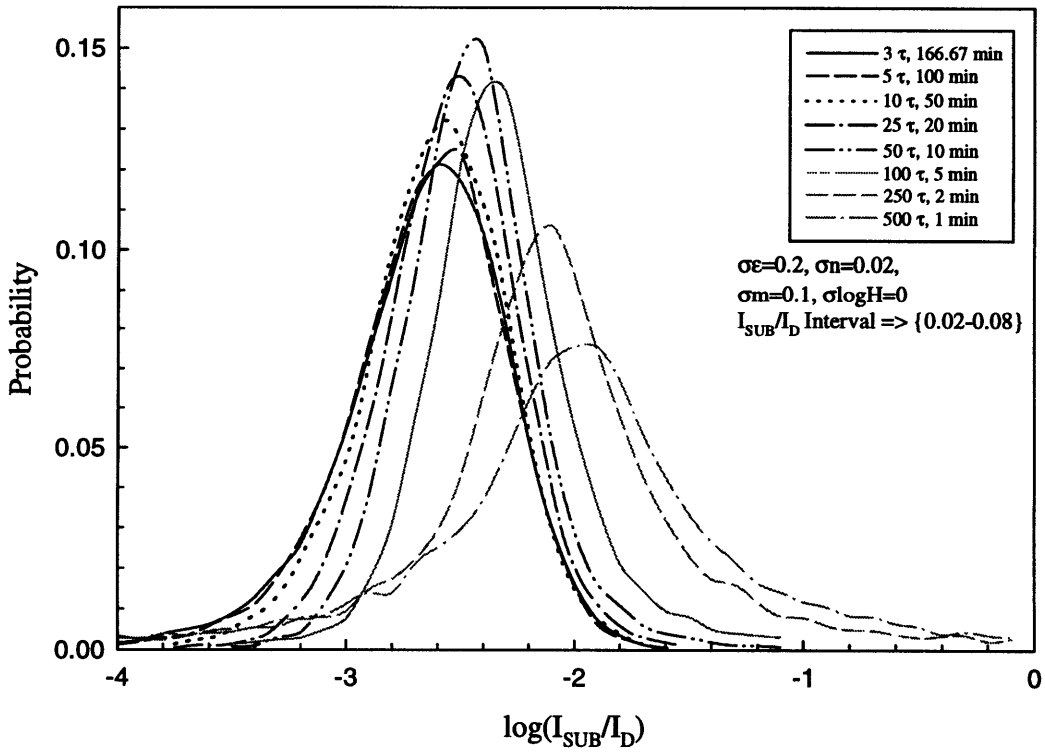


Figure 5.2.6. Log plot of Figure 5.2.4

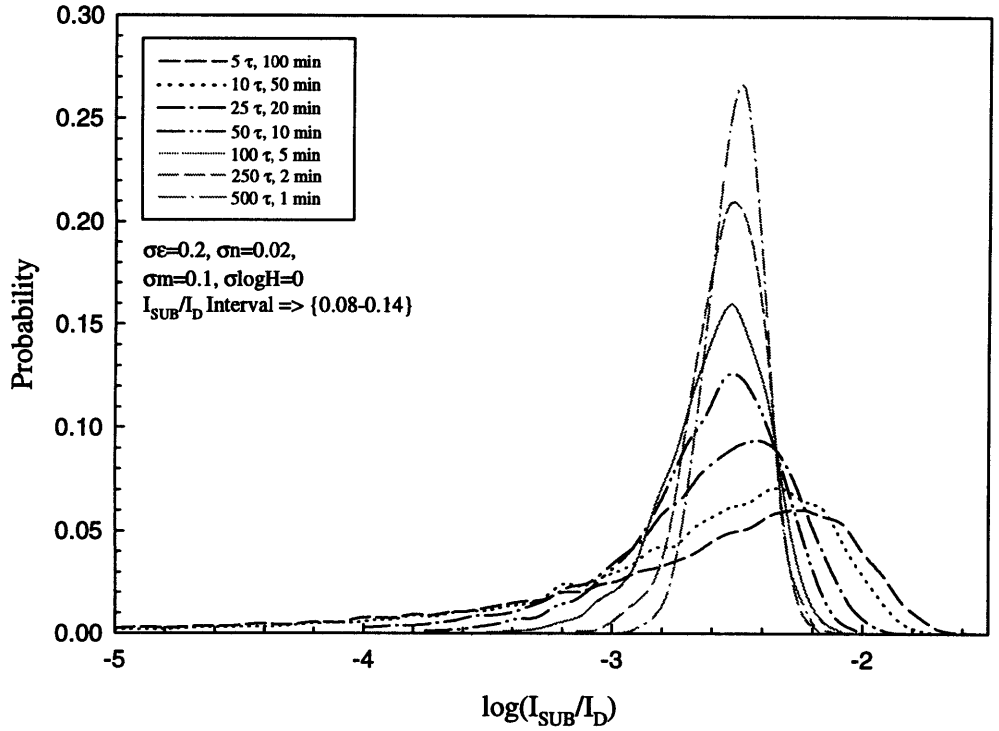


Figure 5.2.7. I_{SUB}/I_D distributions showing optimal region for inner medium I_{SUB}/I_D interval

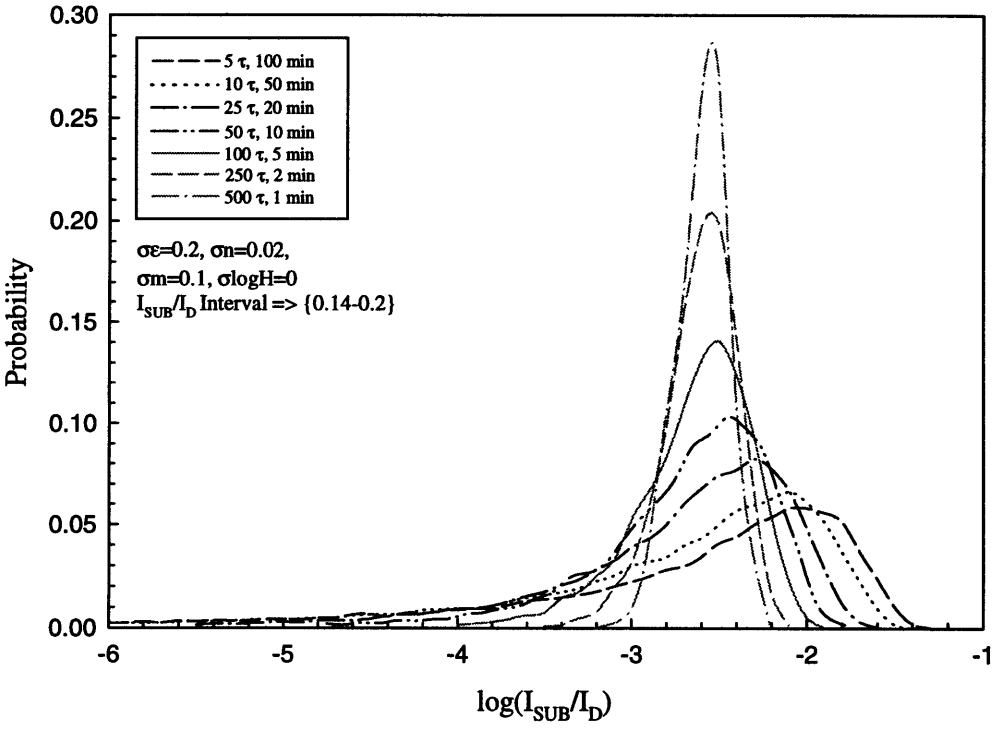


Figure 5.2.8. I_{SUB}/I_D distributions showing optimal region for inner high I_{SUB}/I_D interval

Chapter 6

Conclusion and Recommendations

6.1 Summary of Findings

A methodology has been developed for optimal and accurate extraction of **n**, **m**, and **H** parameters for D.C. hot-carrier degradation modeling. This methodology is based on Monte Carlo simulations of the degradation and lifetime-correlation plots. Boundary curves in the degradation plot define the limits within which lie the set of possible Δ^* curves deviating from the nominal. The extracted slope and fitting coefficient of the lifetime-correlation curves measure the impact of different probabilistic variables. Lifetime distributions at a fixed I_{SUB}/I_D level are also used to characterize the effects of different probabilistic, technology, and optimization variables. This intermediate analysis links the degradation plot to the lifetime-correlation plot. The purpose of these characterization tools is to identify appropriate and realistic choices for the σs, which are used to model the process variation manifested in **n**, **m**, **H**, and instrumentation error. Conservative values resulting in the most realistic degradation scenario are chosen to perform an optimization study focused on addressing two main issues: balancing device stress time with the number of device measurements under a total time for stressing constraint, and the former issue's sensitivity towards different V_G and V_D bias conditions. An element for optimization based on the extrapolated I_{SUB}/I_D values at device lifetime of 10 years has been developed. The corresponding I_{SUB}/I_D distribution is used to define an optimal region in order to evaluate the two main issues.

Simulation results indicate that the existence of the optimal region highly depends upon the choice of σs and the I_{SUB}/I_D bias regime. For the conservative choice of values, no optimal region exists because increasing the number of device measurements result in

tighter and more well-defined I_{SUB}/I_D distributions. Even though increasing the number of device measurements entails shorter stress times, the corresponding I_{SUB}/I_D distributions do not broaden. The scatter introduced by lifetime extrapolation error in the degradation plot due to low stress times has negligible influence on the extrapolation of I_{SUB}/I_D values at a device lifetime of 10 years. The more dominant influence on the I_{SUB}/I_D extrapolation is the larger number of sample points in the lifetime-correlation plot, which consequently improves the I_{SUB}/I_D extrapolation. The small values of σ cause equally distributed variation across the examined stress time regime. As a result, larger values for $\sigma\varepsilon$, σn, and σm are chosen such that the degradation profile shows greater variation in the low stress time regime than in the high regime. The I_{SUB}/I_D bias interval has been extended and subintervals developed in order to study the optimal region's sensitivity toward different bias regimes. Results indicate that an optimal region for the outer extended and inner low bias intervals exists. Although they do not exactly coincide, both regions have close proximity towards one another. No optimal region exists for the inner medium and high intervals, because both cases are similar to the case with conservative σ choices in that larger number of device measurements improves the I_{SUB}/I_D distributions. The low I_{SUB}/I_D interval places the nominal degradation curve at sufficiently low $\Delta I_D/I_D$ levels such that variational differences from low-to-high stress time regimes can be significantly accounted for in the simulation. The medium and high I_{SUB}/I_D intervals effectively view the larger $\sigma\varepsilon$ as a relatively small value similar to the conservative choice case. Furthermore, the optimal region of the inner low interval is better defined than the outer interval. The existence of an optimal region for this outer interval is due to the inclusion of the low I_{SUB}/I_D regime. However, the inclusion of higher I_{SUB}/I_D regimes degrades the optimal region's quality.

The optimal region's sensitivity toward I_{SUB}/I_D bias levels raises a hypothesis that it also has a dependence on the technology variables, which affect the $\Delta I_D/I_D$ level of the

nominal degradation curve. However, this study does not quantitatively validate this hypothesis nor does it attempt to quantify the level of dependence. Experience surmises that the dependence is less significant than the choice of σs and I_{SUB}/I_D bias levels. This study does not directly examine the effects of explicit V_G and V_D bias conditions on the optimal region due to the belief that the small degree of freedom within the narrow V_{GD} window limits the possibilities of affecting the optimal region's sensitivity. It is conjectured that any influence is negligible.

6.2 Recommendations for Further Studies

It is noteworthy that of all the cases in which an optimal region fails to exist, not one case shows an improvement of I_{SUB}/I_D distribution as the number of device measurements decreases. One possible reason is a mismatch in the weighing of each sample point in the lifetime-correlation plot with the corresponding lifetime extrapolation error. Currently any "weighing" is implicitly accounted for by the scattering of the entire sample point space. Improvements to the methodology is possible if a weighing scheme is used in the extrapolation of I_{SUB}/I_D values at 10 years. This scheme should assign greater weights to the sample points corresponding to the I_{SUB}/I_D distribution set with fewer number of device measurements. The reason is that the lifetime extrapolation error has been reduced due to the longer stress times. The determination and implementation of these weights into this study's methodology become the primary focus of the next study. One possible implementation is to employ a weighted least squares regression on the I_{SUB}/I_D extrapolation. Nonetheless, choosing the appropriate weights is a complex and challenging task since they should model actual physical quantities.

Appendix A

Derivation of ANOVA Models

A.1 Testing Each Treatment Mean Against All Treatments

The following analysis derives the models used in the ANOVA table of section 2.3. Each E_{ox} treatment contains a certain number of extracted **n** parameters. Let n_ν be the number of extracted **n** parameters for a given treatment and ν denoting the specific treatment, then

$$N = \sum_{\nu = 1, 2, 3, 4, 5, 6} n_\nu \tag{A.1.1}$$

and the mean of **n** over all treatments be

$$\bar{n} = \frac{\sum_{\nu = 1, 2, 3, 4, 5, 6} \sum_{k = 1}^{n_\nu} x_{\nu, k}}{N} \tag{A.1.2}$$

where $x_{\nu,k}$ is the individual **n** parameter values within each ν treatment [20].

A measure of the variability between the sample treatment means is called the treatment sum of squares and denoted by SS_{means}:

$$SS_{means} = \sum_{\nu = 1, 2, 3, 4, 5, 6} n_\nu \cdot (\bar{x}_\nu - \bar{n})^2 \tag{A.1.3}$$

where $\bar{x}_\nu$ is the mean of the **n** values for a specific treatment.

A measure of the variability of the observed values of **n** around their respective treatment means is the error sum of squares and denoted by SSE:

$$SSE = \sum_{\nu = 1, 2, 3, 4, 5, 6} \sum_{k = 1}^{n_\nu} (x_{\nu, k} - \bar{x}_\nu)^2 \tag{A.1.4}$$

with treatment mean square (MS_{means}) and error mean square (MSE) to be:

$$MS_{means} = \frac{SS_{means}}{\nu_t - 1} \qquad MSE = \frac{SSE}{N - \nu_t} \tag{A.1.5}$$

where ν_t is the total number of treatments.

To determine any statistically significant differences between the treatment means, the amount of between treatment variability is compared to the amount of within treatment variability. An F-test for difference between treatment means is used [20] and defined as:

$$F(means) = \frac{MS_{means}}{MSE} \tag{A.1.6}$$

in which the null hypothesis, H_o, is rejected in favor of the alternate hypothesis, H_1, if $F(means) > F^{(\nu_t - 1, N - \nu_t, \alpha)}$, where α is the degree of confidence. For 95% confidence, α is 0.05. Furthermore, the null hypothesis can also be rejected if the prob-value which is the area to the right of F(means) under the curve of the F-distribution having ν_t-1 and N-ν_t degrees of freedom is less than α. A large value of F(means) results when SS_{means}, which measures the between treatment variability, is large in comparison to SSE, which measures the within treatment variability. F(means) being sufficiently large implies that the null hypothesis should be rejected. A large F(means) also entails a small prob-value. For Technology 1 which has six E_{ox} families and 91 **n** samples, $F^{(5,85,0.05)}$=2.32 with 95% confidence level. This value is significantly less than $F(means)_{Tech1}$ = 12.01. For Technology 2 which has 10 E_{ox} families and 97 **n** samples, $F^{(9,87,0.05)}$ =1.99 with 95% confidence level. $F(means)_{Tech2}$ = 13.76 which is significantly greater than $F^{(9,87,0.05)}$.

A.2 Testing of Pairwise Treatment Means

Once an overall dependence of **n** on E_{ox} has been found for all treatment means, a pairwise comparison of specific treatment means can be performed to determine whether statistically significant differences exist between some of these treatments. This can be

done by testing the null hypothesis H_o: $\mu_i = \mu_j$ against H_1: $\mu_i \neq \mu_j$ where i and j can be any specific treatment. A t-test based on the Student-t distribution is used:

$$t = \frac{\bar{n}_i - \bar{n}_j}{\sqrt{MSE} \cdot \sqrt{\frac{1}{N_i} + \frac{1}{N_j}}} \tag{A.2.1}$$

where N_i and N_j are the total number of **n** values in treatments i and j. If $|t| > t^{(N-v_n, \alpha/2)}$, then H_o is rejected in favor of H_1.

Table A.1 lists the results of each pairwise test and the associated |t| value for Technology 1. $t^{(85,0.025)}$=1.99 for Technology 1 and is used for the comparison. The results state that for tests 1-9, the mean of each examined E_{ox} has some form of statistical dependence on each other. However for tests 10-15, each mean is statistically independent of the other. The dependence and independence fall into two distinct groups: those with E_{ox} values greater than -0.89MV/cm are statistically dependent while those with E_{ox} values less than or equal to -0.89MV/cm are statistically independent. Further hypothesis testing within each group suggest that the group with E_{ox} less than or equal to -0.89MV/cm is independent of **n** while the other group shows a strong dependence of **n** on E_{ox}. The ANOVA table for the group showing statistical independence appears in Figure 2.3.2 (B). The $F^{(3,56,0.05)}$ value with four treatments, 60 **n** samples, and 95% confidence level is 2.77.

Test Number	Test	|t| value	Significance
1	$\mu_{Eox=0.53}$, $\mu_{Eox=-0.18}$	2.05	yes
2	$\mu_{Eox=0.53}$, $\mu_{Eox=-0.18}$	5.44	yes
3	$\mu_{Eox=0.53}$, $\mu_{Eox=-1.61}$	4.69	yes
4	$\mu_{Eox=0.53}$, $\mu_{Eox=-2.32}$	5.89	yes
5	$\mu_{Eox=0.53}$, $\mu_{Eox=-3.04}$	5.91	yes
6	$\mu_{Eox=-0.18}$, $\mu_{Eox=-0.89}$	3.42	yes
7	$\mu_{Eox=-0.18}$, $\mu_{Eox=-1.61}$	2.75	yes
8	$\mu_{Eox=-0.18}$, $\mu_{Eox=-2.32}$	3.87	yes
9	$\mu_{Eox=-0.18}$, $\mu_{Eox=-3.04}$	4.06	yes
10	$\mu_{Eox=-0.89}$, $\mu_{Eox=-1.61}$	0.51	no
11	$\mu_{Eox=-0.89}$, $\mu_{Eox=-2.32}$	0.47	no
12	$\mu_{Eox=-0.89}$, $\mu_{Eox=-3.04}$	0.96	no
13	$\mu_{Eox=-1.61}$, $\mu_{Eox=-2.32}$	0.96	no
14	$\mu_{Eox=-1.61}$, $\mu_{Eox=-3.04}$	1.39	no
15	$\mu_{Eox=-2.32}$, $\mu_{Eox=-3.04}$	0.53	no

Table A.1: Results of pairwise treatment test for Technology 1

No such distinct demarcation exists for Technology 2 as the t-test for pairwise comparison shows only some E_{ox} family having statistical difference which does not fall into any distinct groupings. The $t^{(87,0.025)}$ value with 10 treatments, 97 **n** samples, and 95% confidence level is 1.99. Table A.2 shows the result of each pairwise test.

Test Number	Test	|t| value	Significance
1	$\mu_{Eox=0.36}$, $\mu_{Eox=-0.45}$	3.81	yes
2	$\mu_{Eox=0.36}$, $\mu_{Eox=-1.26}$	2.50	yes
3	$\mu_{Eox=0.36}$, $\mu_{Eox=-1.75}$	3.34	yes
4	$\mu_{Eox=0.36}$, $\mu_{Eox=-2.08}$	4.05	yes
5	$\mu_{Eox=0.36}$, $\mu_{Eox=-2.40}$	4.06	yes
6	$\mu_{Eox=0.36}$, $\mu_{Eox=-2.73}$	3.53	yes
7	$\mu_{Eox=0.36}$, $\mu_{Eox=-3.05}$	5.81	yes
8	$\mu_{Eox=0.36}$, $\mu_{Eox=-3.30}$	6.30	yes
9	$\mu_{Eox=0.36}$, $\mu_{Eox=-3.71}$	9.97	yes
10	$\mu_{Eox=-0.45}$, $\mu_{Eox=-1.26}$	1.21	no
11	$\mu_{Eox=-0.45}$, $\mu_{Eox=-1.75}$	0.57	no
12	$\mu_{Eox=-0.45}$, $\mu_{Eox=-2.08}$	0.34	no
13	$\mu_{Eox=-0.45}$, $\mu_{Eox=-2.40}$	0.25	no
14	$\mu_{Eox=-0.45}$, $\mu_{Eox=-2.73}$	0.28	no
15	$\mu_{Eox=-0.45}$, $\mu_{Eox=-3.05}$	0.22	yes
16	$\mu_{Eox=-0.45}$, $\mu_{Eox=-3.30}$	2.40	yes
17	$\mu_{Eox=-0.45}$, $\mu_{Eox=-3.71}$	6.16	yes
18	$\mu_{Eox=-1.26}$, $\mu_{Eox=-1.75}$	0.65	no
19	$\mu_{Eox=-1.26}$, $\mu_{Eox=-2.08}$	1.51	no
20	$\mu_{Eox=-1.26}$, $\mu_{Eox=-2.40}$	1.45	no
21	$\mu_{Eox=-1.26}$, $\mu_{Eox=-2.73}$	0.94	no
22	$\mu_{Eox=-1.26}$, $\mu_{Eox=-3.05}$	3.31	yes

Table A.2: Results of pairwise treatment test for Technology 2

Test Number	Test	|t| value	Significance
23	$\mu_{Eox=-1.26}$, $\mu_{Eox=-3.30}$	3.57	yes
24	$\mu_{Eox=-1.26}$, $\mu_{Eox=-3.71}$	7.20	yes
25	$\mu_{Eox=-1.75}$, $\mu_{Eox=-2.08}$	0.90	no
26	$\mu_{Eox=-1.75}$, $\mu_{Eox=-2.40}$	0.82	no
27	$\mu_{Eox=-1.75}$, $\mu_{Eox=-2.73}$	0.29	no
28	$\mu_{Eox=-1.75}$, $\mu_{Eox=-3.05}$	2.76	yes
29	$\mu_{Eox=-1.75}$, $\mu_{Eox=-3.30}$	2.99	yes
30	$\mu_{Eox=-1.75}$, $\mu_{Eox=-3.71}$	6.73	yes
31	$\mu_{Eox=-2.08}$, $\mu_{Eox=-2.40}$	0.10	no
32	$\mu_{Eox=-2.08}$, $\mu_{Eox=-2.73}$	0.62	no
33	$\mu_{Eox=-2.08}$, $\mu_{Eox=-3.05}$	1.84	no
34	$\mu_{Eox=-2.08}$, $\mu_{Eox=-3.30}$	1.99	yes, weak
35	$\mu_{Eox=-2.08}$, $\mu_{Eox=-3.71}$	5.65	yes
36	$\mu_{Eox=-2.40}$, $\mu_{Eox=-2.73}$	0.53	no
37	$\mu_{Eox=-2.40}$, $\mu_{Eox=-3.05}$	1.98	no, weak
38	$\mu_{Eox=-2.40}$, $\mu_{Eox=-3.30}$	2.15	yes
39	$\mu_{Eox=-2.40}$, $\mu_{Eox=-3.71}$	5.91	yes
40	$\mu_{Eox=-2.73}$, $\mu_{Eox=-3.05}$	2.48	yes
41	$\mu_{Eox=-2.73}$, $\mu_{Eox=-3.30}$	2.69	yes
42	$\mu_{Eox=-2.73}$, $\mu_{Eox=-3.71}$	6.44	yes
43	$\mu_{Eox=-3.05}$, $\mu_{Eox=-3.30}$	0.01	no
44	$\mu_{Eox=-3.05}$, $\mu_{Eox=-3.71}$	3.59	yes
45	$\mu_{Eox=-3.30}$, $\mu_{Eox=-3.71}$	3.90	yes

Table A.2: Results of pairwise treatment test for Technology 2

Appendix B

Simulation Code Written in *Mathematica 3.0*

■ Monte Carlo Simulation to generate ISUB/ID distribution at device lifetime of 10 years.

■ Purpose: to observe the effect of varying σn, σm, σlogH, σε, stress time sequence, and ISUB/ID interval on the distribution.

Note: the parameter **m** is positive as defined by the Δ equation although the extracted value is negative;
for 10% lifetime definition, log(Δ) = log(10) = 1;
and σε is measured in %.

Load the necessary *Mathematica* 3.0 packages needed for simulation.

```
<< Graphics'Graphics'
<< Statistics'ContinuousDistributions'
<< Graphics'MultipleListPlot'
<< Statistics'LinearRegression'
<< Statistics'NonlinearFit'
<< Statistics'DataManipulation'
<< Graphics'Legend'
```

The technology variables are set as:

```
no = 0.278;
mo = 3.537;
logHo = 2.214;
Ho = 10^(logHo) / (lifetime^(1/no));
w = 5*10^(-6);
Id = 2.685*10^(-3);
lifetime = 0.1;
```

The probabilistic variables are set as:

```
σε = 0.01;
σn = 0.005;
σm = 0.1;
σlogH = 0.1;
```

The time and ISUB/ID biasing sequences are (time is in seconds):

```
time = 60*{0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50};

IsubIdmin = 0.02;
IsubIdmax = 0.08;
IsubId = Table[i, {i, IsubIdmin, IsubIdmax, (IsubIdmax - IsubIdmin) / (TauCollect - 1)}];
```

The number of simulation trials and number of device measurements(NumberDevice) are:

```
NumberTrial = 10000;
NumberDevice = 50;
```

The following initializes some arrays needed to store intermediate simulation variables and defines intermediate constants.

```
ntable = Table[i, {i, TauCollect}];
mtable = Table[i, {i, TauCollect}];
htable = Table[i, {i, TauCollect}];
tautable = Table[i, {i, TauCollect}];
isubidtable = Table[i, {i, NumberRun}];

logΔ = Log[10, (lifetime*100)];
tau10years = Log[10, 315360000 * Id/w];
```

Explanation of the intermediate variables:
ntable is an array which holds the **n** values generated by σn;
mtable is an array which holds the m values generated by σm;
htable is an array which holds the calculated **H** values from the **logH** generated by σlogH;
tautable is an array which holds the extrapolated lifetime from the degradation plot;
isubidtable is an array which holds the extrapolated ISUB/ID values at device lifetime of 10 years from the lifetime correlation plot;
logD is a constant of the lifetime definition;
tau10years is a constant for τID/w at 10 years;
TrialIteration is a variable which records the simulation trial;
DeviceIteration is a variable which records the number of the device being operated on by the simulator;

delta[t_,Is] defines the degradation model;
deltatable is an array which stores the evaluated Δ values at a particular ISUB/ID for the entire stress time sequence;
$\sigma\epsilon$ intoduces scatter to the Δ values and is stored in the deltascatter array;
deltascatterplot is an array which stores the (time,Δ) pairs to be used in regression analysis;
deltascatterplotlog stores the log values of deltascatterplot;
regressionscatter is a variable which holds the results of the regression analysis
flag is a variable which checks if the ($\Delta + \sigma\epsilon$) values are negative since a log operation cannot be perform if true;
tauplot is an array which stores the (ISUB/ID,τID/w) pairs to be used in regression analysis;
tauplotlog stores the log values of tauplot;
regressiontau is a variable which holds the results of the regression analysis;
regressiontauline defines the lifetime correlation equation;
IsubIdExtrapolate equates the lifetime corrrelation equation with τID/w coordinate at 10 years;
logisubid solves for ISUB/ID at 10 years;
i is an array counter.

```
For [TrialIteration = 1, TrialIteration < (NumberTrial + 1), TrialIteration++,

For [DeviceIteration = 1, DeviceIteration < (NumberDevice + 1), DeviceIteration++,
mtable[[DeviceIteration]] = Random[NormalDistribution[mo, sigmam]];
 ntable[[DeviceIteration]] = Random[NormalDistribution[no, sigman]];
 logH = Random[NormalDistribution[logHo, sigmalogH]];
 htable[[DeviceIteration]] = (10^(logH)) / (lifetime^(1/ntable[[DeviceIteration]]));

delta[t_, Is_] =
   ((((Is)^mtable[[DeviceIteration]]) * Id/(w*htable[[DeviceIteration]]))^
      ntable[[DeviceIteration]]) *
    t^ntable[[DeviceIteration]];
 deltatable =
   100*Table[delta[time[[i]], IsubId[[DeviceIteration]]], {i, Length[time]}];
deltascatter =
   Table[Random[NormalDistribution[deltatable[[i]], σε]], {i, Length[deltatable]}];

If [Min[deltascatter] < 0, Flag = 0, Flag = 1];
If[Flag == 0, DeviceIteration = DeviceIteration - 1,
  deltascatterplot = Table[{time[[i]], deltascatter[[i]]}, {i, Length[time]}] // N;
 deltascatterplotlog = Log[10, deltascatterplot];
  regressionscatter = Regress[deltascatterplotlog,
     {1, f}, f, RegressionReport -> {BestFit, BestFitParameters,
                    ParameterCITable, RSquared}];

 tautable[[DeviceIteration]] =
    10^((logΔ - Extract[regressionscatter, {3, 2, 1, 1, 1}]) /
       (Extract[regressionscatter, {3, 2, 1, 2, 1}]));

  ]; (*end of 2nd If statement*)
];    (*end of DeviceIteration FOR LOOP*)

tauplot = Table[{IsubId[[i]], tautable[[i]]*Id/w}, {i, TauCollect}];
tauplotlog = Log[10, tauplot];
regressiontau =
  Regress[tauplotlog, {1, f}, f, RegressionReport -> {BestFit, BestFitParameters,
                    ParameterCITable, EstimatedVariance, RSquared}];
regressiontauline =
  Extract[regressiontau, {3, 2, 1, 2, 1}] * x + Extract[regressiontau, {3, 2, 1, 1, 1}];

IsubIdExtrapolate = regressiontauline == tau10years;
logisubid = Solve[IsubIdExtrapolate, x];
isubidtable[[iteration]] = 10^(Part[logisubid, 1, 1, 2]);
] (*end of TrialIteration FOR LOOP*)
```

The following finds the mean and standard deviation of the ISUB/ID distribution.

```
Print["Mean of Isub/Id Distribution = ", Mean[isubidtable]]
Print["StandardDeviation of Isub/Id Distribution = ", StandardDeviation[isubidtable]]
```

The following calculates the (x,y) coordinates to plot the distribution by counting the frequency of values occuring within each bin. The y coordinate represents the probability of occurance.

```
binsize = 0.05;
Print["Binsize = ", binsize];
Print["Isub/Id values are:"]
isubidx = Table[i, {i, Min[isubidtable], Max[isubidtable], binsize}]
Print["Count of Isub/Id for each bin is:"]
isubidy = BinCounts[isubidtable, {Min[isubidtable], Max[isubidtable], binsize}]
Print["Coordinates of Isub/Id distribution plot are:"]
isubidxy =
 Table[{isubidx[[i]], isubidy[[i]] / Length[isubidtable]}, {i, Length[isubidx]}] // N
```

The following plots the ISUB/ID distribution as well as the mean value.

```
plot1 = ListPlot[isubidxy,
     AxesLabel -> {"Isub/Id", "Prob"}, Prolog -> AbsolutePointSize[5],
     PlotRange -> {{Min[isubidtable], Max[isubidtable]},
    {0, (0.05 + Max[isubidy / Length[isubidtable]])}},
     AxesOrigin -> {Min[isubidx], 0},
     PlotLabel -> "Isub/Id Distribution at 10 yrs"]
plot2 = ParametricPlot[{Mean[isubidtable], x},
  {x, 0, (0.05 + Max[isubidy / Length[isubidtable]])}, PlotStyle -> {Dashing[{0.025}]}]
Show[plot1, plot2]
```

References

[1] B. Doyle, M. Bourcerie, C. Bergonzoni, R. Benecchi, A. Bravis, K. Mistry, and A. Boudou, "The Generation and Characterization of Electron and Hole Traps Created by Hole Injection During Low Gate Voltage Hot-Carrier Stresssing of N-MOS Transistors," *IEEE Trans. on Electron Devices*, vol. 37, no. 8, pp. 1869-1876, August 1990.

[2] C. Hu, S. Tam, F. Hsu, P. Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement," *IEEE Trans. on Electron Devices*, vol. 32, no. 2, pp. 375-385, February 1985.

[3] B. Doyle, M. Bourcerie, J. Marchetaux, and A. Boudou, "Interface State Creation and Charge Trapping in the Medium-to-High Gate Voltage Range ($V_D/2 \geq V_G \geq V_D$) During Hot-Carrier Stressing of n-MOS Transistors," *IEEE Trans. on Electron Devices*, vol. 37, no. 3, pp. 744-754, March 1990.

[4] W. Weber, L.Risch, W. Krautschneider, and Q. Wang, "Hot-Carrier Degradation of CMOS-Inverters", *International Electron Devices Meeting Tech. Digest*, pp. 208-211, 1988.

[5] V-H. Chan, and J. Chung, "The Impact of NMOSFET Hot-Carrier Degradation on CMOS Analog Subcircuit Performance," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 6, pp. 644-649, June 1995.

[6] W. Jiang, H. Le, S. Dao, S. Kim, B. Stine, and J. Chung, "Key Hot-Carrier Degradation Model Calibration and Verification Issues for Accurate AC Circuit-Level Reliability Simulation," *Proceedings of IEEE International Reliability Physics Symposium*, pp. 300-306, 1997.

[7] K. Quader, P. Ko, and C. Hu, "Projecting CMOS Circuit Hot-Carrier Reliability from DC Device Lifetime," *International Electron Devices Meeting Tech. Digest*, pp. 511-514, 1993.

[8] P. Lee, "Modeling and Simulation of Hot-carrier Effects in MOS Devices and Circuits," Ph.D Thesis, 1990.

[9] M. Kuo, K. Seki, P. Lee, J. Choi, P. Ko, and C. Hu, "Simulation of MOSFET Lifetime under AC Hot-Electron Stress," IEEE Trans. Electron Devices, vol. 35, no. 7, pp. 1004-1011, July 1988.

[10] BTABERT User's Manual.

[11] P. Lee, M. Kuo, K. Seki, P. Ko, and C. Hu, "Circuit Aging Simulator (CAS)," *International Electron Devices Meeting Tech. Digest*, pp. 134-137, 1988.

[12] B. Menberu, "Analysis of Hot-Carrier AC Lifetime Model for MOSFET," S.M. Thesis, 1996.

[13] V-H. Chan, J. Kim, and J. Chung, "Parameter Extraction Guidelines for Hot-Electron Reliability Simulation," *Proceedings of IEEE International Reliability Physics Symposium*, pp. 32-37, 1993.

[14] J. Chung, P. Ko, C. Hu,, "A Model for Hot-Electron Induced MOSFET Linear Current Degradation Based on Mobility Reduction due to Interface-State Generation," *IEEE Trans. on Electron Devices*, vol. 38, no. 6, pp. 1362-1370, June 1991.

[15] V-H Chan and J. Chung, "Two-Stage Hot-Carrier Degradation and Its Impact on Submicrometer LDD NMOSFET Lifetime Prediction," *IEEE Trans. on Electron Devices*, vol. 42, no. 5, pp.957-962, May 1995.

[16] D. Baglee, C. Duvvury, M. Smayling, and M. Duane, "Lightly Doped Drain Transistors for Advanced VLSI Circuits," *IEEE Trans. on Electron Devices*, vol. 32, no. 5, pp. 896-902, May 1985.

[17] R. Woltjer and G. Paulzen, "Universal Description of Hot-Carrier-Induced Interface States in NMOSFETs," *International Electron Devices Meeting Technical Digest*, pp. 535-538, 1992.

[18] J. Choi, P. Ko, and C. Hu, "Effect of Oxide Field on Hot-Carrier-Induced Degradation of Metal-Oxide-Semiconductor Field-Effect Transistors," *Applied Physics Letter*, vol. 50, no. 17, pp. 1188-1190, April 1987.

[19] S. Kim, B. Menberu, T. Kopley, and J. Chung, "Oxide-Field Dependence of the NMOS Hot-Carrier Degradation Rate and Its Impact on AC-Lifetime Prediction," *International Electron Devices Meeting Technical Digest*, pp. 37-40, 1995.

[20] B. Bowerman and R. O'Connell, *Linear Statistical Models*, Duxbury Press, second edition, 1990.

[21] G. Krieger, P. Cuevas, and M. Misheloff, "The Effect of Impact Ionization Induced Bipolar Action on N-Channel Hot-Electron Degradation," *IEEE Electron Device Letters*, vol. 9, no. 1, pp. 26-28, January, 1988.

[22] E. Takeda, and N. Suzuki, "An Empirical Model for Device Degradation Due to Hot-Carrier Injection," *IEEE Electron Device Letters*, vol. EDL-4, no. 4, pp. 111-113, April 1983.

[23] S. Sun, M. Orlowski, and K-Y. Fu, "Parameter Correlation and Modeling of the Power-Law Relationship in MOSFET Hot-Carrier Degradation," *IEEE Electron Device Letters*, vol. 11, no. 7, pp. 297-299, July 1990.

[24] S. Wolfram, *The Mathematica Book*, Cambridge University Press, third edition, 1996.

[25] G. Marsaglia and A. Zaman, "Some portable very-long-period random number generators," Computers in Physics, vol. 8, no. 1, pp. 117-121, January/February 1994.

[26] E. Snyder, A. Kapoor, and C. Anderson, "The impact of statistics on hot-carrier lifetime estimates of n-channel MOSFETs," *Proceedings of SPIE Microelectronic Processing Conference*, September 1992.

[27] J. Chen, B. McGaughy, C. Hu, "Statistical Variation of NMOSFET Hot-Carrier Lifetime and its Impact on Digital Circuit Reliability," *International Electron Devices Meeting Technical Digest*, pp. 29-32, 1995.

[28] J. Devore, *Probability and Statistics for Engineering and the Sciences*, Duxbury Press, Fourth Edition, 1995.