Hot-Carrier Reliability of MOSFETs At Room and Cryogenic Temperature

by

SeokWon Abraham Kim

B.S., Electrical Engineering and Computer Science, MIT, 1995
M.Eng., Electrical Engineering and Computer Science, MIT, 1995

Submitted to the

Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1999

Signature of Author 5/1/99
Department of Electrical Engineering and Computer Science
May 1, 1999

Certified by 5/1/99
James E. Chung
Associate Professor of Electrical Engineering
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Commmittee on Grauduate Students
Department of Electrical Engineering and Computer Science

Hot-Carrier Reliability of MOSFETs At Room and Cryogenic Temperature

by

SeokWon Abraham Kim

Submitted to the Department of Electrical Engineering and Computer Science, on December 20, 1998, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Electrical Engineering

ABSTRACT

Hot-carrier reliability is an increasingly important issue as the geometry scaling of MOSFET continues down to the sub-quarter micron regime. The power-supply voltage does not scale at the same rate as the device dimensions, and thus, the peak lateral E-field in the channel increases. Hot-carriers, generated by this high lateral E-field, gain more kinetic energy and cause damage to the device as the geometry dimension of MOSFETs shortens.

In order to model the device hot-carrier degradation accurately, accurate model parameter extraction is critically important. This thesis discusses the model parameters' dependence on the stress conditions and its implications in terms of the device lifetime prediction procedure.

As geometry scaling approaches the physical limit of fabrication techniques, such as photolithography, temperature scaling becomes a more viable alternative. MOSFET performance enhancement has been investigated and verified at cryogenic temperatures, such as at 77K. However, hot-carrier reliability problems have been shown to be exacerbated at low temperature. As the mean-free path increases at low temperature due to reduced phonon-scattering, hot-carriers become more energetic at low temperature, causing more device degradation.

It is clear that various hot-carrier reliability issues must be clearly understood in order to optimize the device performance vs. reliability trade-off, both at short channel lengths and low temperatures. This thesis resolves numerous, unresolved issues of hot-carrier reliability at both room and cryogenic temperature, and develops a general framework for hot-carrier reliability assessment.

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

Acknowledgment

My Lord, I thank you so much for your wonderful grace for the past four years during the doctoral program here at MIT. You have truly and faithfully led my life with your wisdom during my study. Once again, I praise you and thank you for your leadership and companionship. I am also very grateful for the opportunity to be engaged in intellectual challenges at one of the finest academic institutes in the world. Here at MIT, I have met numerous professors, colleagues, and friends from whom I have acquired a tremendous amount of knowledge, insight, and even discipline. Lord, I thank you for all of them.

One of the such great professors I met, my Lord, is my thesis supervisor Professor James Chung. During my graduate study, he taught me and guided me with great enthusiasm and patience. He commended me with enthusiasm when I made good progress, and encouraged me with patience when I was caught with impediment. Lord, I thank you for him and pray that you lavish your blessing on him as you did on me. Please, reveal yourself to him constantly so that he can always be appreciative and joyful. I pray that his life will be full of praise to you.

My God, I also thank you for my thesis committee professors, namely, Professor Duane Boning and Professor Leslie Kolodziejski. I thank them very much for taking their precious time to review my thesis and provide feedback. Lord, I pray that you pour your blessing on them as well and guide their life with your love and wisdom.

Lord, now, I thank you for my wonderful parents; my father Sam Kim, and my mother, Chong-Ye Kim. Without their constant support physically, emotionally, and spiritually, my graduate study would not have been possible. They always encouraged and trusted me under any circumstances. I cannot resist my thought of gratitude whenever I think of them. My Lord, I pray that they grow in faith everyday, realize the mission of their life, and work to expand your kingdom in this dark world for the rest of their life. It is the most important blessing that they await. I also thank you for my siblings, my brother Young Kim, my sister-in-law Kyoung-Won Kim, and their cute daughter Sunjoo. Their companionship has truly been my joy and pleasure. I also thank you for my sister Jin Kim, her husband Woong Chun, and their son Hoi-soo. Their experienced advice has been very helpful during my graduate study.

Heavenly Father, now, I thank you for my colleagues and friends whose companionship has strengthened and comforted me during the tiresome journey over the past 4 years. They are Huy Le, Arifur Rahman, Dennis Ouma, Wenjie Jiang, Jung Yoon, Jeehoon Yap, Changwoo Kang, Dong-hyun Kim, Jekwan Ryu, Dongsik Yoo, Seonghwan Cho, Junehee Lee, Seongho Cho, Jungyoor Choi, Sae-young Chung, and Won-jong Kim. Some of them have left, but most of them are still here. I pray that you will lead their graduate life with your wisdom and blessing as you led mine.

My Savior, finally, I thank you and praise you for my precious spiritual colleagues, whose prayers and companionship have been and will continue to be true comfort to me. They are Reverend Young-Ghil Lee, Reverend Hyoung-Gon Kim, Yunhee Kim, Sun-Young Lee, Kyoung-

moon Yoon, Min Han, Hackjin Kim, Hea-Kwang Chun, Song-Hyuck Oh, and Daejin Whang. I ask for your unwavering, continuous blessing all throughout their life.

My Lord, I thank you again from the most bottom of my heart for everything that you have done for me. Lead me onward through the journey of my life until I come and dwell with you. In the name of my victorious Savior Jesus Christ, I pray. Amen.

On the first day of May in year nineteen hundred ninety-nine

Biography

SeokWon Abraham Kim was born on November 15, 1970, in Jeon-joo, Korea. After immigrating from Seoul, Korea, to Atlanta in 1985, he graduated as valedictorian from South Cobb High School in Austell, GA, in 1989. He simultaneously earned his Bachelor of Science and Master of Engineering degree in Electrical Engineering and Computer Science with minor in Economics in February, 1995, at the Massachusetts Institute of Technology. His M. Eng. thesis, titled "Modeling for Hot-Electron Reliability Simulation," won the David Adler Thesis Award in the Electrical Engineering and Computer Science Department at MIT.

His research interest includes semiconductor device performance and reliability modeling. As a side project apart from his doctoral thesis, he has worked on SPICE modeling of MOSFET at cryogenic temperature. After joining the Microsystems Technology Laboratories as an undergraduate student in 1992, he has worked as a research assistant from 1993 to 1999. During his doctoral study, he had also worked at LG Semicon, in Seoul, for the summer of 1996 on parameter extraction issues for SPICE models.

Apart from the academics, SeokWon has been involved in the Graduate Student Council (GSC) at MIT. He served as chair of the Academic Project and Policy Committee at GSC from 1996 to 1997. He also served as student representative to the Committee on Graduate School Policy. He is currently an active member of Korean Church of Boston in Brookline, MA.

Table of Contents

CHAPTER 5 Dominant Degradation Asymptote (DDA) Algorithm for AC-Lifetime Prediction

CHAPTER 6 Hot-Carrier Reliability at Cryogenic Temperature

CHAPTER 7 Conclusion

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Since the invention of the first working MOSFET in 1960 by Kahng and Atalla [1.1], MOSFET technology has grown at an exponential rate that is unprecedented compared with any other industry. Although MOSFET technology began with building just a few simple test structures almost four decades ago, currently, it is possible to build more than 10^{12} MOSFETs on a single ship [1.2]. This Ultra Large Scale Integration (ULSI) has been successfully achieved by aggressive geometry scaling of MOSFET devices. The feature size of the state-of-the-art MOSFET channel length is now below 0.1 μm [1.3]. As many device and technology experts foresee that geometry scaling will approach fundamental physical limits in the near future, alternative scaling approaches are being researched. One of such approach is that of temperature scaling. In numerous studies [1.4-1.6], MOSFET performance has been shown to improve significantly at cryogenic temperatures. Temperature scaling down to liquid nitrogen temperature 77K is viewed as a viable alternative when geometry scaling approaches its limit.

Although MOSFET performance improvement has been a primary focus of semiconductor industry research for the past few decades, MOSFET reliability has also been a serious issue that has drawn more and more attention as the aforementioned, aggressive scaling methodologies have been adopted. Reliability issues, such as hot-carrier injection, electromigration, latch-up, and oxide breakdown, have been actively researched in order to optimize the MOSFET performance vs. reliability trade-off. It is clear that the performance improvement that has been achieved over the past four decades cannot be fully

maximized unless performance vs. reliability trade-off issues and their metrics are well understood and developed. Among the different reliability issues listed above, hot-carrier reliability in particular draws much attention since it is a generic problem that arises in all MOSFET technologies. Additionally, unlike the other mechanisms, hot-carrier injection worsens at cryogenic temperatures [1.7-1.8]. The increasing importance of hot-carrier reliability issues, especially at microscopic dimensions and at low temperatures, is the motivation for this dissertation.

1.2 Background in Hot-Carrier Reliability

1.2.1 Hot-Carrier Generation by E_{peak}

Hot-carriers are generated by the high lateral electric field at the drain end of the channel in a MOSFET. This generation is schematically shown in Figure 1.1. Because of aggressive channel-length scaling with today's MOSFET technologies and because the power-supply voltage has scaled at a much slower rate in order to improve performance, the lateral electric field has continually increased. In particular, the peak electric field at the drain end of the channel has become increasingly higher, accelerating the channel electrons, increasing their energy. These energetic electrons have a higher effective temperature than the electrons in the surrounding lattice, and thus, are called **hot-electrons**. These energetic hot-electrons collide with the silicon lattice, break the bonds between silicon atoms, creating electron-hole pairs. This process is called **impact ionization**. Most of the created holes from impact ionization flow out to the substrate and appear as substrate current I_{sub} as shown in Figure 1.1.

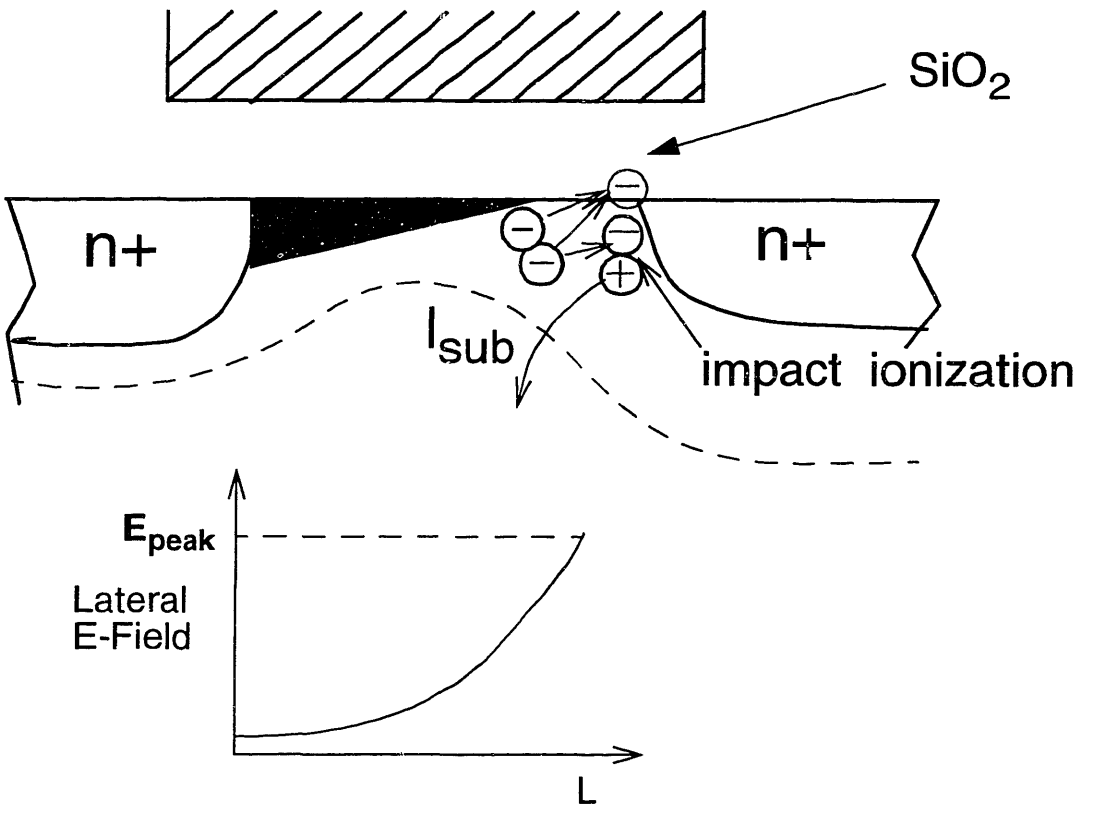

Figure 1.1: Generation of hot-electrons in the cross section of the MOSFET. Also shown is the lateral electric field along the channel.

1.2.2 Device Degradation Mechanism

The energetic hot-electrons not only can cause impact ionization, but also have sufficient energy to cross over the Si/SiO_2 energy barrier. Some of the generated hot-holes from impact ionization also can gain sufficient energy to be injected into the oxide over the Si/SiO_2 interface. Previous studies claim that, during hot-carrier stressing, hot-holes are initially injected into the oxide, increasing the density of trap centers for the subsequently injected hot-electrons [1.9-1.10]. As energetic hot-electrons cross over the Si/SiO_2 interface, they break Si-H bonds, leaving behind dangling silicon bonds. The generated interstitial hydrogen diffuses towards the gate away from the interface. Subsequent hot-

electrons that cross over the Si/SiO_2 interface get trapped at the interface and form a bond with the dangling Si atom. This structure is called an **interface-trap**, and is illustrated in Figure 1.2. In particular, what is shown in Figure 1.2 is called an acceptor-type interface trap, since it accepts the electron and becomes negatively charged.

Some of the hot-electrons travel further into the gate oxide, and get trapped in the oxide. This process is called an **electron-trap** in the oxide. Correspondingly, there can be a **hole-trap** in the oxide as well. Some hot-electrons do not get trapped in the oxide, but travel all the way out to the gate terminal and appear as gate current I_g. Usually under normal operating conditions, the order of magnitude for I_g is very small, around a hundred femto-amps (~ 10^{-13} A).

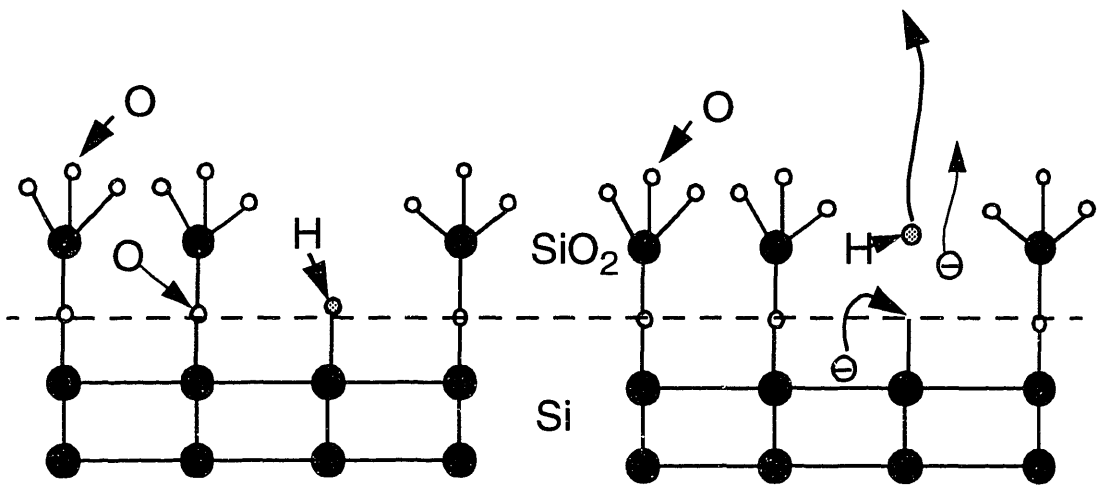

Figure 1.2: Iinterface-trap generation. The interface becomes negatively charged when the hot-electron gets trapped and forms a bond with the dangling Si atom as shown in (2b).

Although there have been numerous conflicting discussions for the past decade on which mechanism is the dominant one for MOS device degradation, recently, there has been a general consensus that interface-trap generation is the dominant degradation mechanism for the NMOS device, and electron-trap generation is the dominant degradation mechanism for the PMOS device [1.13-15].

1.2.3 Impact on Device and Circuit Performance

Hot-electron-induced interface-traps can have significant impact on MOS device performance, and hence, on circuit performance. Because of the negative charge built-up, both at the Si/SiO_2 interface and in the oxide, the threshold voltage V_T tends to increase after stress. It takes more positive gate voltage (V_g) to turn on the channel because of the negative charge built-up, and thus, V_T increases. Under severe stressing, the change in V_T, which we define as ΔV_T, can be as high as 100 mV. As V_T increases, the current drive decreases. In other words, I_d decreases for the same bias conditions. Similarly, the transconductance $g_m(\equiv\frac{\partial I_d}{\partial V_{gs}})$ also decreases. These effects are schematically shown in Figure 1.3.

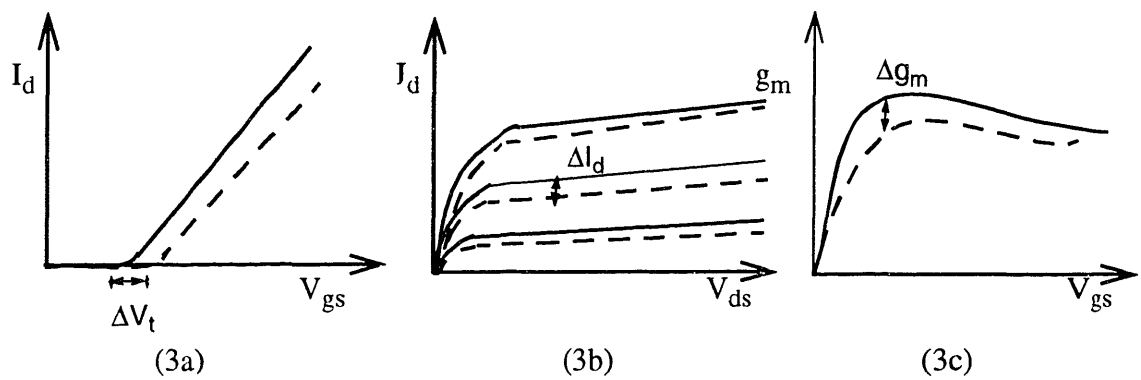

Figure 1.3: Impact of hot-electron degradation on the MOS device performance. Solid line represents a fresh device; dotted line, a degraded device.

Device degradation translates into circuit performance degradation in many ways. For digital circuits, frequency degradation has been observed due to the reduction in current drive. In analog circuits, such as a differential amplifier, gain reduction has been observed because of transconductance degradation. The offset voltage, V_{offset}, also has been observed to shift from its initial condition.

1.2.4 Concept of Charge Pumping Current

Section 1.2.2 briefly described the three hot-carrier degradation mechanisms: interface-traps, electron-traps, and hole-traps. Although all these mechanisms contribute to device parameter degradation, such as ΔV_t, Δg_m, and $\Delta I_d/I_{d0}$, generally, one mechanism dominates at a particular stress bias condition. It is not quite clear, however, how to identify which mechanism is dominant, just based on the device parameter degradation alone.

In order to distinguish between the three degradation mechanisms, an experimental technique called charge pumping has been developed [1.16], and its setup is shown in Figure 1.4. For the NMOS device shown in Figure 1.4, the source and the drain are grounded. A square wave is applied from a pulse generator to the gate whose V_{high} is greater than the V_T of the device and whose V_{low} is less than the V_{fb} (flat band voltage) of the device. The substrate is connected to a pico-ammeter to measure any net charge transport, defined as the charge pumping current.

The principle of charge pumping is as follows: When the device is pulsed into inversion by V_{high}, the surface becomes deeply depleted, and electrons flow from the source and drain regions into the channel, where some of them will be captured by interface-states at the Si/SiO_2 interface. When the device is pulsed into accumulation by V_{low}, the mobile charge drifts back to the source and drain, but the charges trapped in the

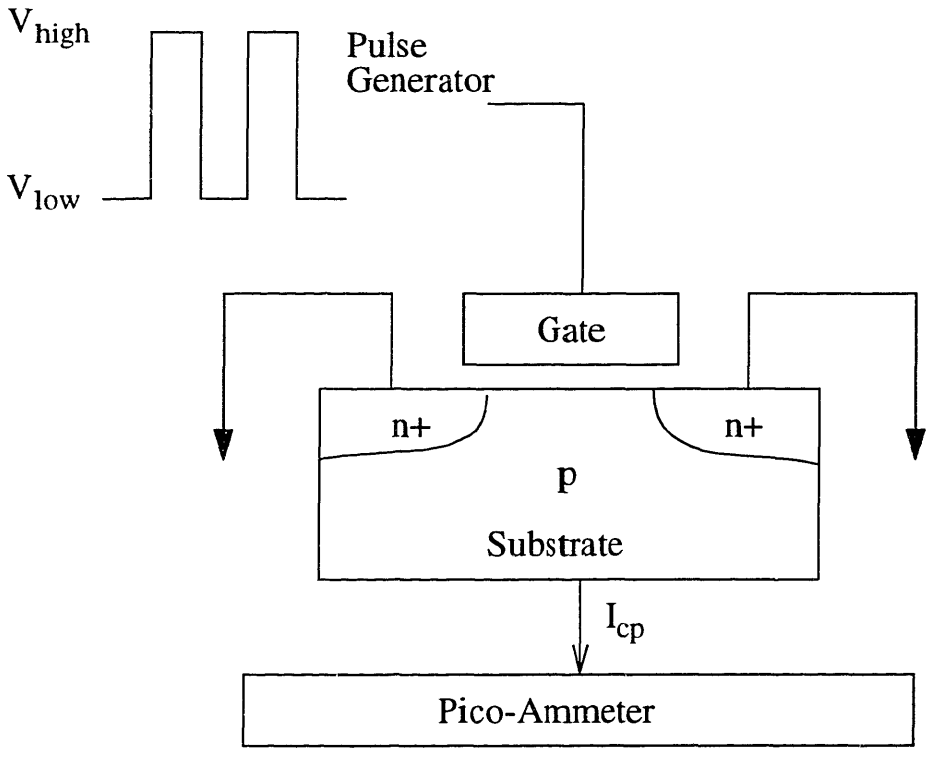

Figure 1.4: Experimental setup for charge pumping current measurement.

interface-states recombine with majority carriers from the substrate, giving rise to a net flow of negative charge into the substrate, constituting the charge pumping current. Since the current is directly proportional to the amount of interface-trapped charge, an estimate of the mean value of the interface-state density over the energy range swept by the gate pulse can be obtained by measuring the charge-pump current. As the device gets stressed and undergoes hot-carrier degradation, the charge-pump current rises due to the increased interface-trapped charge. Thus, measurement of the charge-pump current I_{cp} gives a better illumination about the degradation mechanisms than the measurement of device degradation quantities, such as ΔV_t, Δg_m, and $\Delta I_d/I_{d0}$.

1.3 Motivation for Dissertation

At the ultra short-channel regime, it is believed that MOS transistors undergo more severe hot-carrier degradation. Because the power-supply voltage scales at a much slower rate than the channel length in order to achieve higher current drivability, the lateral E-field in the MOSFET channel becomes greater, thus energizing the hot-electrons more and more. This results in increased hot-carrier damage to the device, such as increased interface-traps, which translates to more V_t shifts, and g_m and I_d reduction.

Many device and technology experts argue that temperature scaling is a viable alternative when geometry scaling reaches its physical limit [1.10-12]. At cryogenic temperatures, high current drivability is achieved through increased mobility of electrons. However, as the electron's mean-free path increases due to the reduced phonon scattering at low temperature, the electrons become more energetic and get injected more into the Si/SiO_2 interface and SiO_2 itself. Again, at the low temperatures, hot-carrier degradation becomes aggravated.

This increasing concern of hot-carrier reliability for ultra short-channel MOS transistors at cryogenic temperatures serves as motivation for this dissertation. In order to maximize fully the MOSFET performance both at ultra short-channel dimensions and at low temperatures, and still assure an acceptable device lifetime, a general assessment framework for the overall impact of hot-carrier degradation needs to be developed. In order to develop such a framework, numerous unresolved issues, which range from physical issues, such as identifying the degradation mechanism at different temperatures, to methodology issues, such as how to correlate between DC and AC degradation, must be clearly addressed. This dissertation identifies many of these unresolved issues and

addresses them in a clear, consistent method in order to better understand hot-carrier reliability at various dimensions and temperatures.

1.4 Problem Statement

In order to optimize the performance vs. reliability trade-off, an accurate MOSFET hot-carrier lifetime prediction model is necessary. Without such a model, we are forced to design the MOS transistor conservatively in order to ensure an acceptable device lifetime. In other words, the lack of an accurate lifetime prediction model prevents a precise performance vs. reliability trade-off.

Currently, there exists a hot-carrier degradation model that is widely used both in academic and industry [1.17], which allows an accurate calculation of DC device lifetime. However, the accuracy of this model, developed at U.C Berkeley, is questionable under AC operation. This raises a serious concern in hot-carrier reliability assessment because MOS devices mostly experience AC waveform stress in the typical integrated circuit environment.

Past studies have shown that the direct application of the Berkeley model to AC MOSFET operation produces an erroneous lifetime prediction [1.18-19]. The reason for this inaccuracy comes from the fact that the bias-dependence of the model parameters is not correctly taken into account. Under an AC waveform, the model parameter values constantly change as the bias conditions change. Since the existing methodology assumes constant values of the model parameters while the bias conditions change, the predicted AC-lifetime is inaccurate.

Along with lifetime prediction issues under AC waveforms, various hot-carrier degradation issues at low temperature are also addressed in this dissertation. There exist

conflicting theories as to the worst-case stress conditions at cryogenic temperatures [1.20-21]. It is still not clear whether the Berkeley model and its methodologies are still applicable at low temperature. Also, before we adopt low temperature MOSFET operation, we need to quantify how much MOSFET performance is improved vs. how much reliability is sacrificed at low temperature.

1.5 Dissertation Objectives

There are two main objectives of this dissertation. First, this dissertation develops an efficient model and algorithm that can accurately calculate the AC-lifetime of MOS transistors. This accurate AC-lifetime prediction model is indispensable in order to optimize the performance vs. reliability trade-off for ultra short-channel MOS device operation at the cryogenic temperature. The improved accuracy of the AC-lifetime prediction is achieved by first characterizing the model parameters' bias-dependencies and then by carefully incorporating such dependencies in the AC-lifetime prediction procedure. Some physical explanation for the model parameters' bias-dependencies are presented along the way.

Second, this dissertation quantifies how much hot-carrier reliability is sacrificed at low temperatures. Both the physical damage, ΔI_{cp}, and device degradation monitors, such as I_d and g_m reduction, are measured. In doing so, the DC Berkeley model and methodology are verified at low temperatures down to 77K. Degradation mechanisms are also explored at cryogenic temperatures. Based on this analysis, the MOSFET performance vs. reliability trade-off is quantified as a function of temperature.

In order to achieve the objectives, this dissertation is composed of the following sections: Chapter 2 derives and presents a widely used hot-carrier degradation model and its parameter extraction procedures; Chapter 3 discusses and characterizes the model

parameters' dependence on the bias conditions; Chapter 4 discusses an impact of the bias-dependent model parameters on AC device lifetime prediction; Chapter 5 presents a new algorithm, which can predict the AC-lifetime of MOSFETs more accurately and efficiently; and Chapter 6 discusses the hot-carrier reliability issues at cryogenic temperature. Finally, the dissertation ends with conclusions in Chapter 7.

References

[1.1]. D. Kahng and M. M. Atalla, "Silicon-silicon Dioxide Field Induced Devices," Solid-State Device Research Conference, Pittsburgh, June 1960.

[1.2]. H. Koga, T. Matsuki, and K. Koyoma, "Two-Dimensional Borderless Contact Pad Technology for a 0.135mm^2 4-Gigabit DRAM Cell," International Electron Device Meeting, p25, 1997.

[1.3]. H. Hu, J. Jacobs, L. Su, and D. Antoniadis, "A Study of Deep-Submicron MOSFET Scaling Based on Experiment and Simulation," IEEE Transactions on Electron Devices, Vol. 42, p669, 1995.

[1.4]. F. Gaensslen, V. Rideout, E. Walker, and J. Walker, "Very Small MOSFET's for Low-Temperature Operation," IEEE Transactions on Electron Devices, Vol.24, p218, 1977.

[1.5]. G. Baccarani, M. Wordeman, and R. Dennard, "Generalized Scaling Theory and Its Application to a 1/4 Micrometer MOSFET Design," IEEE Transactions on Electron Devices, Vol.31, p452, 1984.

[1.6]. F. Balestra and G. Ghibaudo, "Brief Review of the MOS Device Physics for Low Temperature Electronics," Soild-State Electronics, Vol.37, p1967, 1994.

[1.7]. G. Sai-Halasz, M. Wordeman, D. Kern, S. Rishton, E. Ganin, T. Chang, and R. Dennard, "Experimental Technology and Performance of 0.1 μm gate-length FETs Operated at Liquid-Nitrogen Temperature," IBM Journal of Research and Development, Vol. 34, p452, 1990.

[1.8]. G. Gildenblat and C. Huang, "N-Channel MOSFET Model for the 60-300K Temperature Range," IEEE Transactions on Computer-Aided Design, Vol. 10, p512,

1991.

[1.9]. P. Heremans, G. Bosch, R. Bellens, G. Groeseneken, and H. Maes, "Temperature Dependence of the Channel Hot-Carrier Degradation of n-Channel MOSFET's," IEEE Transactions on Electron Devices, Vol. 37, p980, 1990.

[1.10]. S. Bibyk, H. Wang, and P. Borton, "Analyzing Hot-Carrier Effects on Cold CMOS Devices," IEEE Transactions on Electron Devices, Vol. 34, p83, 1987

[1.11]. J. Tzou, C. Yao, and H. Chan, "Hot-Electron-Induced MOSFET Degradation at Low Temperatures," IEEE Electron Devices Letters, Vol. EDL-6, No.9. p450, 1985.

[1.12]. J. Bracchitta, T. Honan, and R. Anderson, "Hot-Electron-Induced Degradation in MOSFET's at 77K," IEEE Transactions on Electron Devices, Vol. ED-32, p1850, 1985.

[1.13]. B. Doyle, M. Bourcerie, J. Marchetaux, and A. Boudou, "Interface State Creation and Charge Trapping in the Medium-to-High Gate Voltage Range ($V_d/2 > V_g > V_d$) During Hot-Carrier Stressing of n-MOS Transistors," IEEE Transactions on Electron Devices, Vol. 37, p244, 1990.

[1.14]. K. Hofmann, C. Werner, W. Weber, and G. Dorda, "Hot-Electron and Hole-Emission Effects in Short n-Channel MOSFET's," IEEE Transactions on Electron Devices, Vol. 32, p691, 1985.

[1.15]. R. Woltjer, A. Hamada, and E. Takeda, "Time Dependence of p-MOSFET Hot-Carrier Degradation Measured and Interpreted Consistently Over Ten Orders of Magnitude," IEEE Transactions on Electron Devices, Vol. 40, p392, 1993.

[1.16]. R. Paulsen, and M. White, "Theory and Application of Charge Pumping for the

Characterization of Si-SiO2 Interface and Near-Interface Oxide Traps," IEEE Transactions on Electron Devices, Vol. 41, p1213, 1994.

[1.17]. C. Hu, S. Tam, F. Hsu, P.Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation - Model, Monitor, and Improvement," IEEE Transactions on Electron Devices, Vol. 32, p375, 1985.

[1.18]. S. Kim, B. Menberu, and J. Chung, "Oxide-Field Dependence of the NMOS Hot-Carrier Degradation Rate and Its Impact on AC-Lifetime Prediction," International Electron Devices Meeting, p37, 1995.

[1.19]. V. Chan, S. Kim, and J. E. Chung, "Parameter Extraction Guidelines for Hot-Electron Reliability Simulation," International Reliability Physics Symposium, p32, 1993.

[1.20]. J. Wang-Ratkovic, K. MacWilliams, R. Lacoe, and M. Song, "Impact of Channel Length and Bias Conditions on Accelerated Hot-Carrier Lifetime Testing at Cryogenic Temperatures," VLSI Symposium, 1996.

[1.21]. J. Tzou, C. Yao, R. Cheung, and H. Chan, "Hot-Electron-Induced MOSFET Degradation at Low Temperatures," IEEE Electron Device Letters, Vol. 6, p450, 1985.

Chapter 2

Hot-Carrier Degradation Model and Its Parameter Extraction

2.1 Degradation Model

As was briefly mentioned in Section 1.2.2, interface-traps have been shown to be the dominant degradation mechanism for the NMOS device at room temperature [2.1]. In other words, device I-V characteristic degradation, such as ΔV_T and $\Delta I_d/I_{d0}$, can be attributed to the interface-traps generated by hot-carriers [2.2]. Note that even for a fresh device, there are some intrinsic interface-states that are formed during the device fabrication process. We attempt to model the increase in the number of interface-traps, i.e. ΔN_{it}, which is the increase in degradation generated by hot-carriers.

Before we move on to develop the degradation model, there is one important concept that needs to be clarified. Hot-carriers are generated by the high E-field, E_{peak}, at the drain end of the channel, and thus, if we can accurately calculate E_{peak}, we should be able to calculate the number of hot-carriers generated, and hence, the number of interface-traps generated. E_{peak} can be indeed quite accurately calculated using a numerical device simulator such as MEDICI. However, such simulation requires intensive calibration to produce accurate results, and it is very difficult to experimentally verify. Our approach is to develop a semi-empirical model whose model parameters are based on a measurable quantity of the MOS devices. This model will capture the basic physical theory of hot-carrier generation and degradation, but the model parameters will be directly measurable quantities, thus facilitating experimental verification. One such useful, measurable quan-

tity for MOS devices is substrate current, I_{sub}. Although it is difficult to measure E_{peak}, it is simple to measure I_{sub}, which is produced by hot-carriers generated by the same E_{peak}. Thus, we can use I_{sub} as a monitor for hot-carrier generation and degradation.

In order to accelerate hot-carrier degradation in a controlled experiment, we apply a much higher stress voltage than the normal operating voltage to a MOSFET, and observe how much degradation occurs in a given time. Such an experimental result is shown in Figure 2.1. For an NMOS device with the given device parameters in the Figure, we have applied $V_{ds} = 5.2V$ and $V_{gs} = 0.9V$. The substrate was grounded. In Figure 2.1, we plot the percentage reduction of the drain current ($\Delta I_d/I_{d0}$), where I_{d0} is the initial drain current, vs. stress time in a log-log space. As one can see, the degradation data, $\Delta I_d/I_{d0}$, has a linear dependence on time in log-log space for this bias condition. We have assumed that

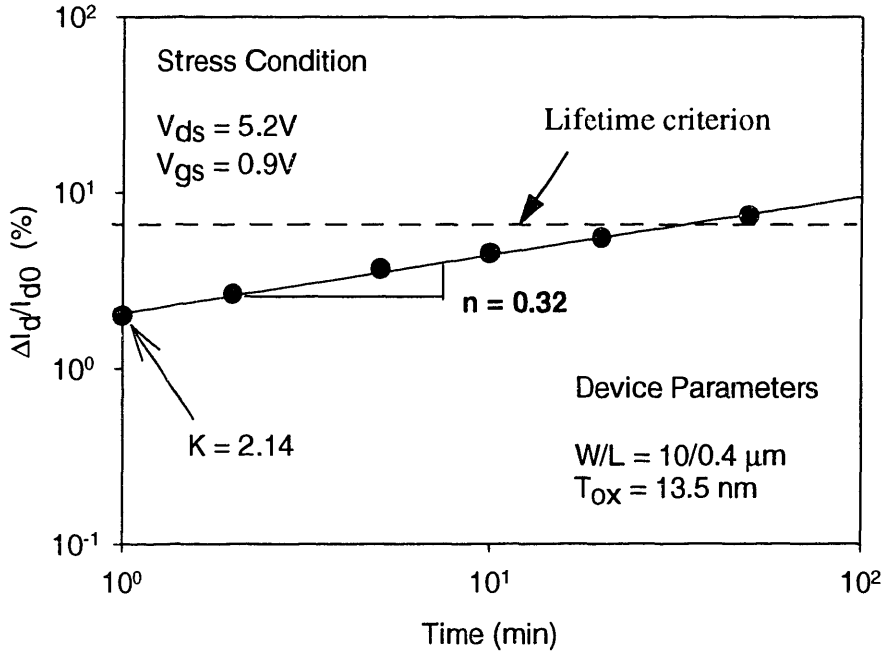

Figure 2.1: Hot-carrier degradation time dependence.

the degradation is caused by interface-trap generation as discussed above. Combining this assumption with the data, we can write the following equation:

$$\frac{\Delta I_d}{I_{d0}} \propto \Delta N_{it} = K t_{stress}^{\ n} \tag{2.1}$$

where K is the y-intercept, and **n** is the slope of the extrapolated lines in log-log space as shown in Figure 2.1.

Now, let us define a critical energy ϕ_{it} as the amount of energy that a hot-electron must possess in order to create the interface-trap ΔN_{it}. Then, we can expand on what K is in Equation 2.1 as follows [2.2]:

$$\Delta N_{it} = K_1[(\text{Electron Flux})(\text{Probability of electrons gaining } \phi_{it})(t_{stress})]^{\mathbf{n}} \tag{2.2}$$

where K_1 is a technology-dependent constant. Now, let us look at the second parenthesized term more closely. If ϕ_{it} is the critical energy for ΔN_{it}, then $\frac{\phi_{it}}{qE_{peak}}$ is the average distance that an electron must travel under E_{peak} to gain ϕ_{it}. Now, by dividing this quantity by the electron mean-free path λ, and using Boltzman statistics, we can obtain a probability that an electron gains ϕ_{it} without undergoing an energy-losing collision under E_{peak} as $\exp\left(-\frac{\phi_{it}}{q\lambda E_{peak}}\right)$. Since $\frac{I_d}{W}$ represents the normalized electron flux in the channel where W is the width of the device, we can derive the following equation by substituting these quantities into Equation 2.2:

$$\Delta N_{it} = K_1\left(\frac{I_d}{W}\exp\left(-\frac{\varphi_{it}}{q\lambda E_{peak}}\right)t_{stress}\right)^n \tag{2.3}$$

Although Equation 2.3 can be used to calculate ΔN_{it}, determining the E_{peak} value is very difficult as discussed before. However, recall that the same E_{peak} also generates substrate current I_{sub}. Now, if we define ϕ_i as the impact ionization energy required for substrate current generation, then by similar argument, we can write down the following equation [2.3]:

$$I_{sub} = K_2 I_d \exp\left(-\frac{\varphi_i}{q\lambda E_{peak}}\right) \tag{2.4}$$

where K_2 is a technology-dependent constant for I_{sub}.

Now, by solving Equation 2.4 for E_{peak} and substituting it into Equation 2.3, we derive the following degradation model:

$$\Delta N_{it} = \left(\frac{I_d}{WH}\left(\frac{I_{sub}}{I_d}\right)^m t_{stress}\right)^n \tag{2.5}$$

where $m = \frac{\varphi_{it}}{\varphi_i}$, and **H** is a lumped technology-dependent constant. Equation 2.5, as discussed before, is a semi-empirical model which captures the basic physical theory of hot-carrier degradation, but whose model parameters are based on measurable quantities, such as I_d and I_{sub}. Now, if we assume that the device I-V characteristic degradation, such as $\Delta I_d/I_{d0}$, is attributable to the interface-trap generation, then we can calculate $\Delta I_d/I_{d0}$ with Equation 2.6 [2.2, 2.4].

$$\frac{\Delta I_d}{I_{d0}} = \left(\frac{I_d}{WH}\left(\frac{I_{sub}}{I_d}\right)^m t_{stress}\right)^n \tag{2.6}$$

Notice that, besides directly measurable parameters such as I_d and I_{sub} of the MOSFET,

there are three degradation model parameters that we must extract in order to use Equation 2.6, namely **n**, **m**, and **H**. Once we extract these parameters experimentally, then we can calculate $\Delta I_d/I_{d0}$ as a function of stress time at a given stress bias condition.

2.2 Degradation Model Parameter Extraction

2.2.1 Degradation Rate Coefficient n

The degradation rate coefficient **n** is a time-acceleration factor, which tells us how fast the device degrades. For a fixed DC bias condition V_{gs}, V_{ds}, and V_{bs}, Equation 2.6 can be rewritten as follows:

$$\frac{\Delta I_d}{I_{d0}} = K(t_{stress})^n \tag{2.7}$$

where K is a constant since I_d and I_{sub} are constant for a fixed DC bias condition. Now, by taking the logarithm on both sides of Equation 2.7, Equation 2.8 is derived as follows:

$$\log\left(\frac{\Delta I_d}{I_{d0}}\right) = n\log(t_{stress}) + n\log K \tag{2.8}$$

Thus, by plotting $\left(\frac{\Delta I_d}{I_{d0}}\right)$ vs. t_{stress} in log-log space, we can extract the degradation rate coefficient **n** by taking the slope of the fitted line through the data. The y-intercept at y=1 yields $n\log K$. This is precisely what is done in Figure 2.1. For the stressed device, V_{ds} is fixed at 5.2V, V_{gs} at 0.9V, and V_{bs} at 0V. After having plotted $\left(\frac{\Delta I_d}{I_{d0}}\right)$ vs. t_{stress} in the log-log space, we extract the parameter **n** by linearly regressing the data.

2.2.2 The m and H Coefficients

The other two parameters that we have to extract for the degradation model (Equation 2.6) are the **m** and **H** parameters. Mathematically, $m = \frac{\varphi_{it}}{\varphi_i}$ as was found in deriving Equation 2.6. Physically, **m** is a voltage-acceleration factor. We typically stress devices at a much higher voltage than the normal operating voltage in order to expedite the degradation. In an accelerated stress experiment, we are able to observe measurable degradation in a reasonable measurement interval. However, what we really want to know is how much degradation would occur at a normal operating voltage over the normal operating lifetime. The **m** is the parameter that allows us to extrapolate the amount of degradation at a high stress voltage to the degradation at an operating voltage, and thus is called the voltage-acceleration factor. The parameter **H** is a technology-dependent constant.

The **m** and **H** parameter extraction procedure proceeds as follows. First, we define the MOS device lifetime criterion. For example, we can define the device lifetime criterion as $\Delta I_d/I_{d0} = 10\%$ in the linear region as done in Figure 2.1. In other words, we consider the device to be malfunctioning when its drain current in the linear region degrades by 10% or more. Now, in Equation 2.6, we substitute the following quantity:

$$0.1 = \left(\frac{I_d}{WH} \left(\frac{I_{sub}}{I_d} \right)^m \tau \right)^n \tag{2.9}$$

where τ is the lifetime of the device. Now, we can rearrange Equation 2.9 as follows:

$$\frac{\tau I_d}{W} = \left(\frac{I_{sub}}{I_d} \right)^{-m} H(0.1)^{\frac{1}{n}} \tag{2.10}$$

Now, by plotting $\frac{\tau I_d}{W}$ vs. $\frac{I_{sub}}{I_d}$ in log-log space, we can extract **m** from the slope of the fitted line and **H** from the y-intercept. The parameter **n** has already been determined from Figure 2.1. This procedure is shown below in Figure 2.2.

Notice that each data point in Figure 2.2 corresponds to a stressed device. For Figure 2.2, 7 distinct devices were stressed at different bias conditions to extract the **m** and **H** parameters. Generally, the higher I_{sub}/I_d corresponds to the higher stress V_{ds}. When the stress V_{ds} is higher, the device lifetime is shorter, which is why we see a negative slope in Figure 2.2.

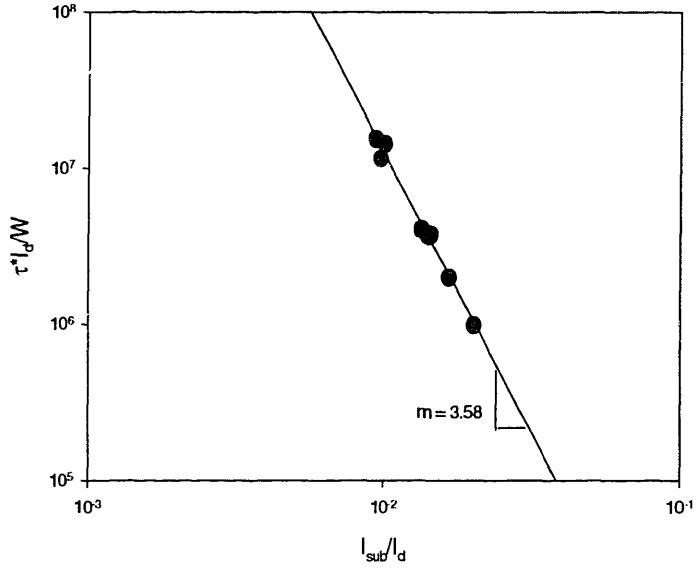

Figure 2.2: The parameter **m** and **H** extraction.

References

[2.1]. B. Doyle, M. Bourcerie, J. Marchetaux, and A. Boudou, "Interface State Creation and Charge Trapping in the Medium-to-High Gate Voltage Range ($V_d/2 > V_g > V_d$) During Hot-Carrier Stressing of n-MOS Transistors," IEEE Transactions on Electron Devices, Vol. 37, p244, 1990.

[2.2]. C. Hu, S. Tam, F. Hsu, P.Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation - Model, Monitor, and Improvement," IEEE Transactions on Electron Devices, Vol. 32, p375, 1985.

[2.3]. P.Ko, R. Muller, and C. Hu, "A Unified Model for Hot-Electron Currents in MOSFETs," IEEE International Electron Device Meetings, p600, 1981.

[2.4]. S. Tam, P. Ko, and C. Hu, "Lucky-Electron Model of Channel Hot-Electron Injection in MOSFETs," IEEE Transactions on Electron Devices, Vol. 31, p1116, 1984.

Chapter 3

Oxide-Field Dependence of the Degradation Model Parameters

3.1 Worst-case Stress Condition

There have been numerous studies [3.1-3.3] that have investigated the worst-case hot-carrier stress conditions. Currently, there is a general consensus that the peak I_{sub} bias condition is the worst-case stress condition, although there are some recent reports that contest this issue [3.4-3.5]. In Figure 3.1, we have plotted I_{sub} and the corresponding degradation $\Delta I_d/I_{d0}$ against V_{gs} for a fixed V_{ds} bias condition. As one can see, the degradation occurred most at the peak I_{sub} condition. The stress time was 50 minutes.

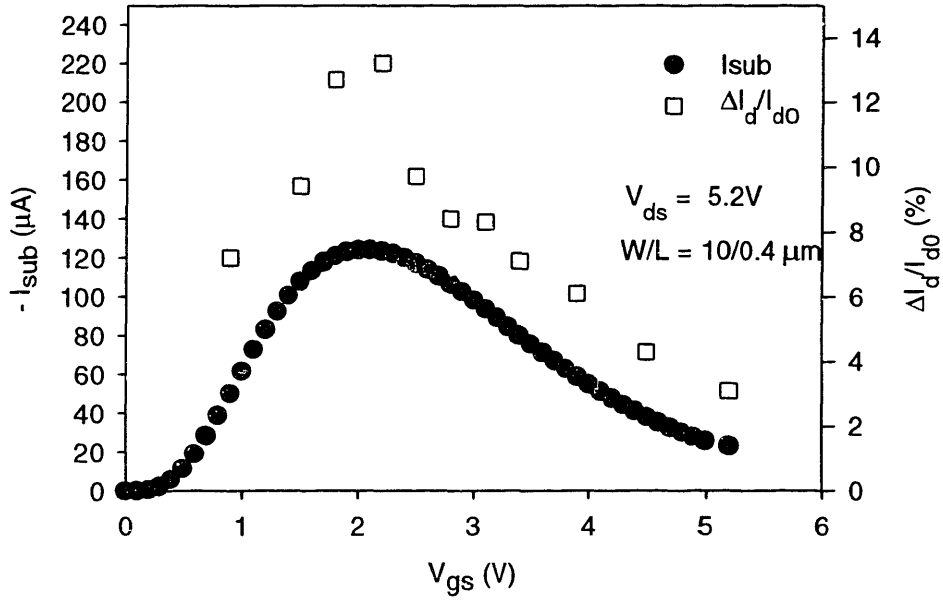

Figure 3.1: Comparison of the peak I_{sub} condition with the worst degradation condition.

An explanation for the behavior shown in Figure 3.1 is as follows: as V_{gs} decreases from 5V (for a fixed V_{ds}), the device is moving towards the stronger saturation regime from the linear regime, and hence, the E_{peak} increases. The increased E_{peak} generates more I_{sub}, and we observe that I_{sub} increases accordingly as V_{gs} decreases. Even though the E_{peak} continues to increase as V_{gs} approaches 0V, the drain current I_d becomes very small as V_{gs} approaches 0V. In other words, the number of channel carriers is not large enough to generate hot-carriers even though the E_{peak} is very strong when V_{gs} is very small. Thus, we observe that I_{sub} decreases as V_{gs} decreases towards 0V.

Figure 3.1 clearly shows that the degradation correlates well with I_{sub}. Since the I_{sub} behavior takes into account both the number of channel carriers and the E_{peak} as a function of bias conditions, $\Delta I_d/I_{d0}$ correlates well with I_{sub} as inspected.

3.2 Oxide-Field Dependence of the Degradation Rate Coefficient n

3.2.1 Bias Dependence of n

The extraction procedure for the parameter **n** was explained in Section 2.2.1. In order to characterize the parameter **n**, NMOS devices were stressed at various bias conditions, and the result is shown in Figure 3.2. For each stress performed, the NMOS devices were biased at V_{ds} = 5.2V. Each device, however, was stressed at a different V_{gs} bias condition as shown in the Figure. An interesting observation in Figure 3.2 is that the degradation rate coefficient **n** changes as we vary the stress bias condition. From a low V_{gs} that is barely above the threshold voltage V_{th} (V_{gs} = 0.9V) to a high V_{ds} (V_{gs} = 5.2V) at which the device is biased in the linear regime, the parameter **n** initially increases, reaches a maximum, and decreases. In order to further investigate this bias dependence of **n**, stress conditions were repeated for multiple devices, and the results are shown in Figure 3.3.

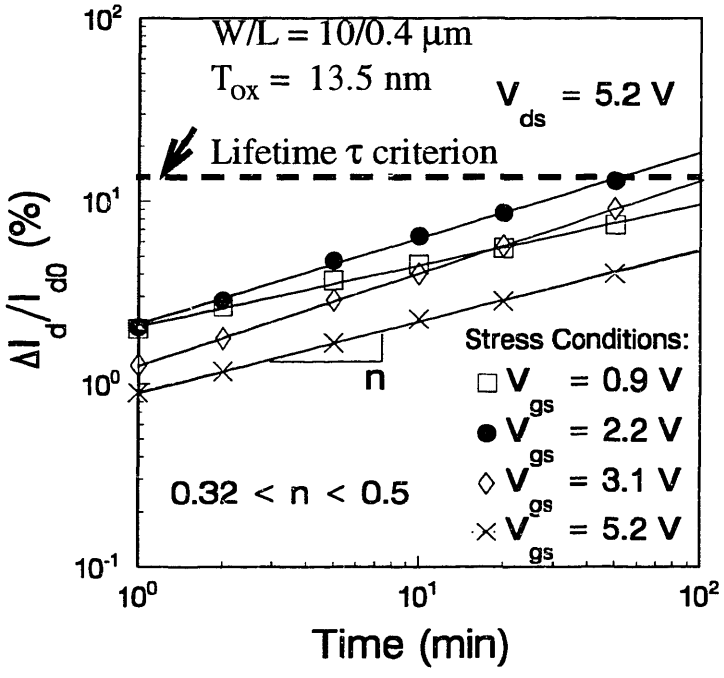

Figure 3.2: $\Delta I_d/I_{d0}$ vs. time for various V_{gs} bias conditions.

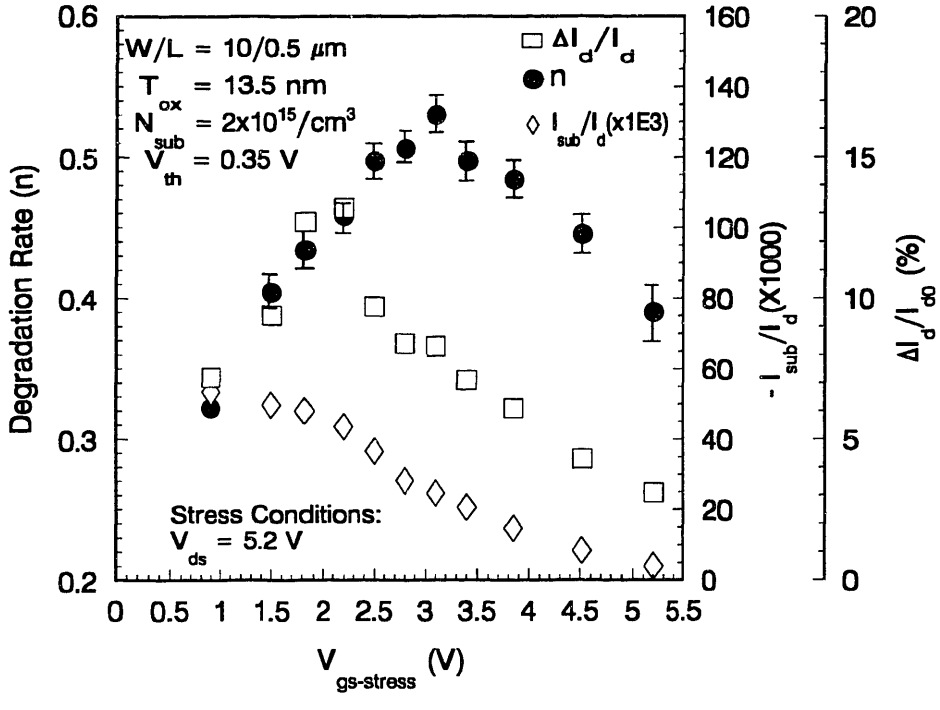

Figure 3.3: The **n**, $\Delta I_d/I_{d0}$, and I_{sub}/I_d vs. V_{gs}. T_{stress} = 50 minutes.

Figure 3.3 shows the parameter **n**, the degradation $\Delta I_d/I_{d0}$, and the ratio I_{sub}/I_d plotted against V_{gs}. There are several key observations to be made from Figure 3.3. First, the parameter **n** has a strong V_{gs} bias dependence and varies with a bell shape, which is similar to the degradation behavior of $\Delta I_d/I_{d0}$. The bias-dependence of the parameter **n** raises an interesting issue with respect to MOSFET lifetime prediction because MOSFETs in a circuit experience time-varying AC circuit waveforms, and thus, time-varying bias conditions. Since the parameter **n** changes with bias conditions, it is no longer clear which value of **n** should be used in Equation 2.6 in order to calculate the MOSFET lifetime when it undergoes various bias conditions under AC circuit operation.

Second, although the shape of the **n** parameter's V_{gs} dependence is similar to that of the degradation $\Delta I_d/I_{d0}$ is, the peak location is different. For example, in Figure 3.3, the maximum degradation at a stress time of 50 minutes occurs at $V_{gs} = 2.2V$, whereas the peak **n** parameter value occurs at $V_{gs} = 3.1V$. In other words, although more degradation occurred at $V_{gs} = 2.2V$ in 50 minutes, the degradation rate is faster at $V_{gs} = 3.1V$. This suggests that, if the device is stressed long enough, eventually, the maximum degradation would occur at $V_{gs} = 3.1V$. A complication arises with respect to MOSFET lifetime prediction under AC circuit operation because, at a future time point of interest, such as 10 years, it is not clear which bias condition would dominate the degradation. In other words, it is not clear whether to use $\mathbf{n} = 0.45$ ($V_{gs} = 2.2V$) or $\mathbf{n} = 0.54$ ($V_{gs} = 3.1V$) to calculate the total degradation at 10 years in Equation 2.6. This issue will be discussed in greater detail in Chapter 4.

Third, neither the parameter **n** nor the degradation $\Delta I_d/I_{d0}$ correlates with I_{sub}/I_d.

The I_{sub}/I_d ratio represents the normalized strength of E_{peak}, and thus, one would speculate that the degradation $\Delta I_d/I_{d0}$ would correlate with I_{sub}/I_d. Indeed, that is what we observe in Figure 2.2. The higher is I_{sub}/I_d, the more degradation occurs, and hence, the shorter is the lifetime. However, it is clear in Figure 3.3 that this trend is only valid in a relatively high V_{gs} region. For example, in Figure 3.3, the degradation $\Delta I_d/I_{d0}$ is larger at $V_{gs} = 2.2V$ than at $V_{gs} = 1.5V$ although the I_{sub}/I_d is higher at $V_{gs} = 1.5V$ than at $V_{gs} = 2.2V$. This issue of model validity will be discussed in greater detail in Section 3.4.

3.2.2 Charge Pumping Current Measurement

In order to further investigate the nature of the bias-dependent degradation rate **n**, the charge-pump current I_{cp} has been measured as a monitor for hot-carrier degradation. As explained in Section 1.2.4, I_{cp} serves as a direct monitor for the hot-carrier induced interface-traps. The rise in I_{cp} monitors the increase in the number of interface-traps [3.9-3.12]. Multiple NMOS devices were stressed at a wide range of bias conditions as in Figure 3.2 with I_{cp} as the degradation monitor, and the result is plotted in Figure 3.4. As one can see, the extracted degradation rate **n** based on ΔI_{cp}, rather than $\Delta I_d/I_{d0}$ as in Figure 3.2, shows a very similar bias dependence. This is further illuminated in Figure 3.5, which plots ΔI_{cp}, the extracted **n** based on ΔI_{cp}, and the I_{sub}/I_d ratio against V_{gs}. The charge-pump experiment illustrates that the rate of physical hot-carrier damage, the interface-trap generation, is bias-dependent. In other words, it is not just a manifestation of the hot-carrier damage, $\Delta I_d/I_{d0}$, ΔV_t, and Δg_m, that is bias-dependent, but the damage itself is inherently bias-dependent.

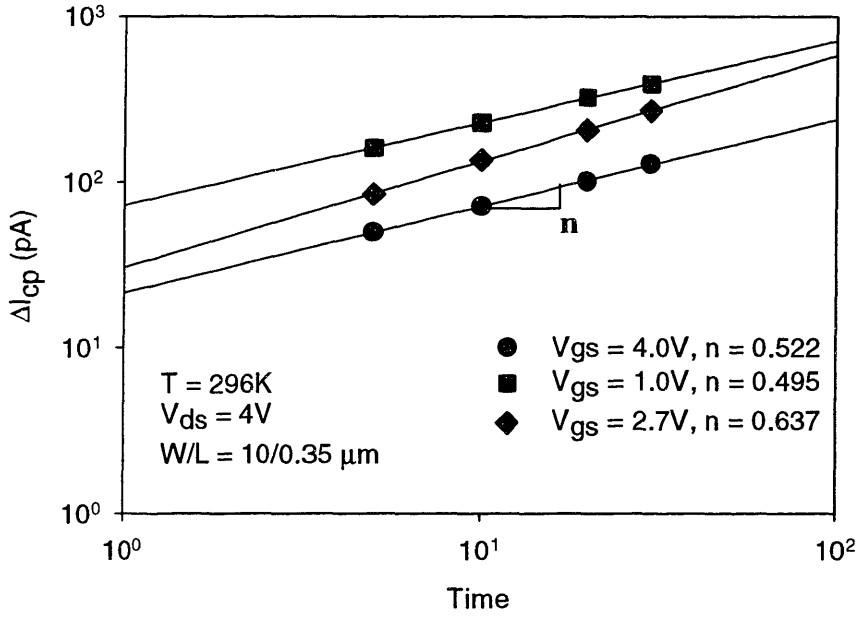

Figure 3.4: The bias dependence of the **n** extracted from the ΔI_{cp}.

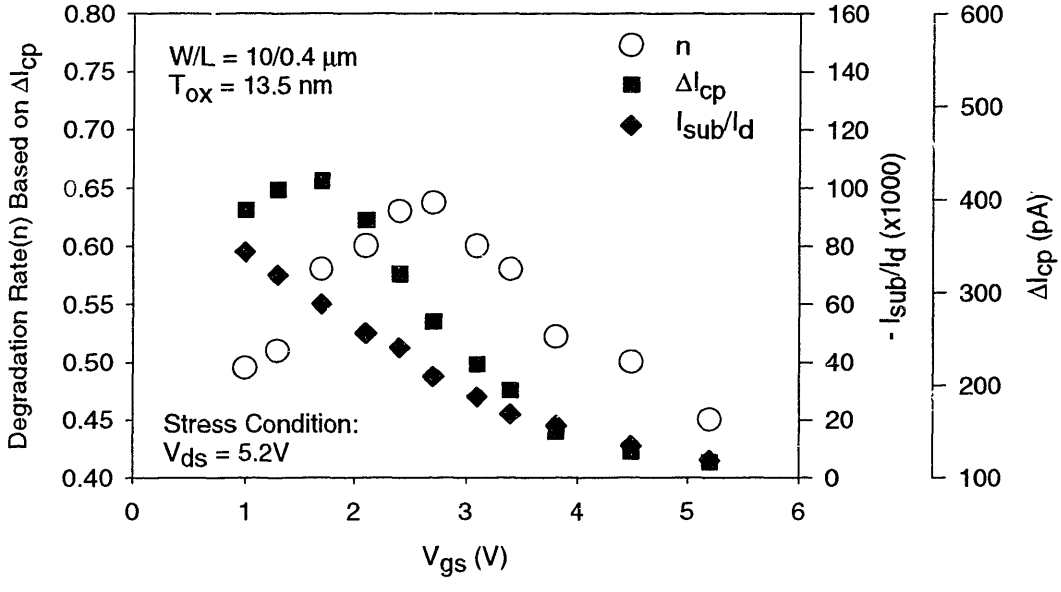

Figure 3.5: ΔI_{cp}, **n**, and I_{sub}/I_d as a function of stressing V_{gs} for a fixed V_{ds}.

3.2.3 Physical Theory and Model of the Bias-Dependent n Coefficient

The bias-dependence of the degradation rate **n**, extracted based on device degradation monitor, such as $\Delta I_d/I_{d0}$ and ΔV_t, has been observed in previous studies [3.13-15]. These studies attributed the bias dependence of **n** to different degradation mechanisms in different bias regions. According to [3.6-3.7], there exist distinct bias regions where different degradation mechanisms are dominant. For the low gate bias region, $V_{gs} \sim V_t$, hole-trapping is the prevalent degradation mechanism. For the mid-gate bias region where $V_{gs} \sim \frac{V_{ds}}{2}$, interface-trap generation is the dominant mechanism, and for the high-gate bias region, $V_{gs} \sim V_{ds}$, electron-trapping is dominant. Thus, one can speculate that the different degradation rates at different bias regions can be explained by different degradation mechanisms.

However, Figure 3.5 shows that the parameter **n**, extracted based on ΔI_{cp}, still has a bias dependence. In other words, the interface-trap generation rate is inherently bias dependent. In order to observe whether this is a universal phenomenon over a wide range of bias conditions, such as V_{gs} and V_{ds}, and device parameters, such as T_{ox}, stress experiments were performed on multiple dimensions of devices with varying L_{eff} and T_{ox}. The parameter **n** was extracted based on $\Delta I_d/I_{d0}$.

Rather than plotting **n** against V_{gs}, as in Figure 3.5, at different V_{ds} bias conditions for the MOS transistors with various ranges of device parameters, we have plotted **n** against E_{ox} in Figure 3.6, where E_{ox} is the vertical electric field in SiO_2 at the drain end of the channel. At the drain end of the channel, it is approximated that

$$E_{ox} \cong \frac{V_{gs} - V_{ds}}{T_{ox}} \tag{3.1}$$

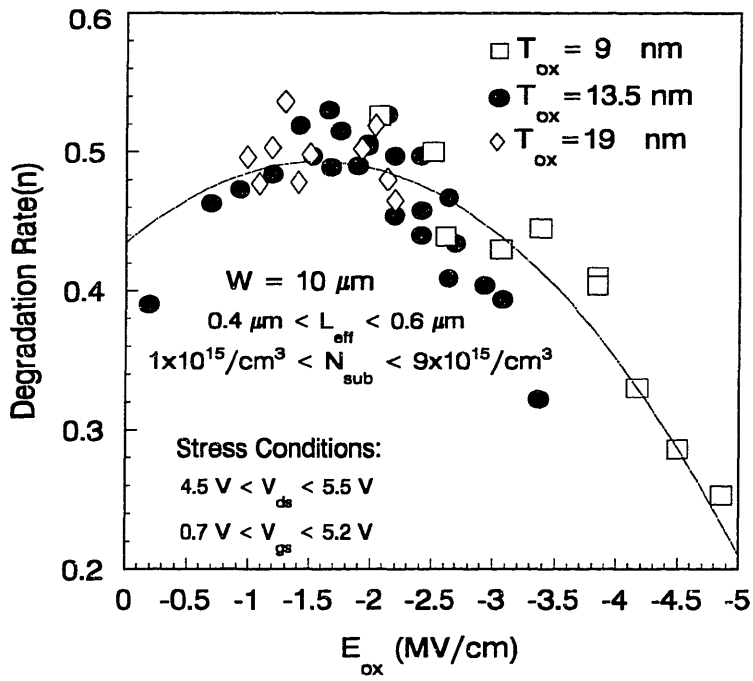

Figure 3.6: An empirical model of **n** as a second order polynomial function of E_{ox}.

Notice that E_{ox} is a generalized parameter that can be extracted over a wide range of bias conditions, V_{gs} and V_{ds}, and device parameter T_{ox}. For the data shown in Figure 3.6, the experimental devices' oxide thickness ranged from 9 nm to 19 nm, channel length from 0.4 μm to 0.6 μm, and the channel doping varied from $1x10^{15}$ cm^{-3} to $9x10^{15}$ cm^{-3}. The bias conditions also varied; V_{ds} from 4.5V to 5.5V, and V_{gs} from 0.7V to 5.2V.

The first observation in Figure 3.6 is that the degradation rate **n** seems to increase, reaches its maximum, and then decreases as E_{ox} increases. One can fit the data with a second order polynomial function as shown in the figure to develop an empirical model of **n** as a function of E_{ox}. However, it is not physically intuitive as to why **n** has to vary with E_{ox} in a bell-shaped curve. One hypothesis is that this is due to the oxide-field dependent

diffusion of interstitial hydrogen that is generated from silicon-hydrogen bond breaking, which occurs during hot-carrier injection as shown in Figure 1.2. This hydrogen diffusion process has been proposed as the rate-limiting step for hot-carrier-induced interface-trap generation [3.16-17], and thus, one may speculate that the diffusion of interstitial hydrogen is dependent upon the oxide field.

However, this theory still does not completely explain the observed behavior in Figure 3.6. There is no clear answer why the diffusion rate of interstitial hydrogen has such a bell-shaped dependence on E_{ox}. In other words, it is not clear why the diffusion rate of interstitial hydrogen first increases with E_{ox}, reaches its maximum, and eventually decreases with E_{ox}.

Another hypothesis to explain the shown behavior in Figure 3.6 is that the number of hot-electrons *generated* and *injected* into the Si/SiO_2 interface plays a more dominant role in determining the overall degradation rate than does the interstitial hydrogen diffusion. The more hot-electrons that are injected into the Si/SiO_2 interface over the energy barrier in a shorter period of time, the greater the degradation rate. This theory is plausible because, in order to create the interface-traps, the hot-electrons first need to be generated, then need to be injected into the Si/SiO_2 interface over the energy barrier, and then travel through the SiO_2. Some of these hot-electrons get trapped in the SiO_2. Only a small fraction travels all the way out to the gate terminal and appears as gate current, I_g.

Recall that in the saturation regime where E_{peak} is high enough to generate hot-electrons, V_{ds} is greater than V_{gs} as shown in Figure 3.7 below. Hence, the vertical oxide-field is pointed from the drain to the gate. This oxide field is in the direction that opposes

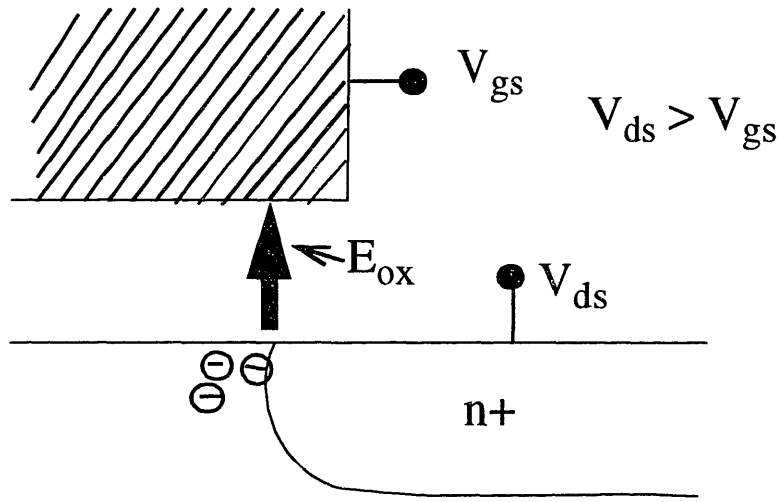

Figure 3.7: Direction of E_{ox} that opposes the hot-electron injection.

hot-electron injection. Thus, the greater the E_{ox}, the more difficult it becomes for hot-electrons to get injected across the Si/SiO_2 interface and the SiO_2. Therefore, as E_{ox} increases, the rate of interface-trap generation decreases. This interpretation is shown in Figure 3.8. The same data shown in Figure 3.6 is replotted in Figure 3.8. However, instead of fitting the data as a second order polynomial function, we observe that the degradation rate exponentially decreases as E_{ox} increases beyond a certain value. This is in accordance with the hypothesized theory above. As shown in the figure, we can model the degradation rate as

$$n = A \cdot \exp(-B \cdot E_{ox}) \quad \text{for } E_{ox} > C \tag{3.2}$$

where A and B are extracted constants to minimize the error of the data fit, and C is the E_{ox} value beyond which Equation 3.2 is valid. In Figure 3.8, C is approximately 2MV/cm.

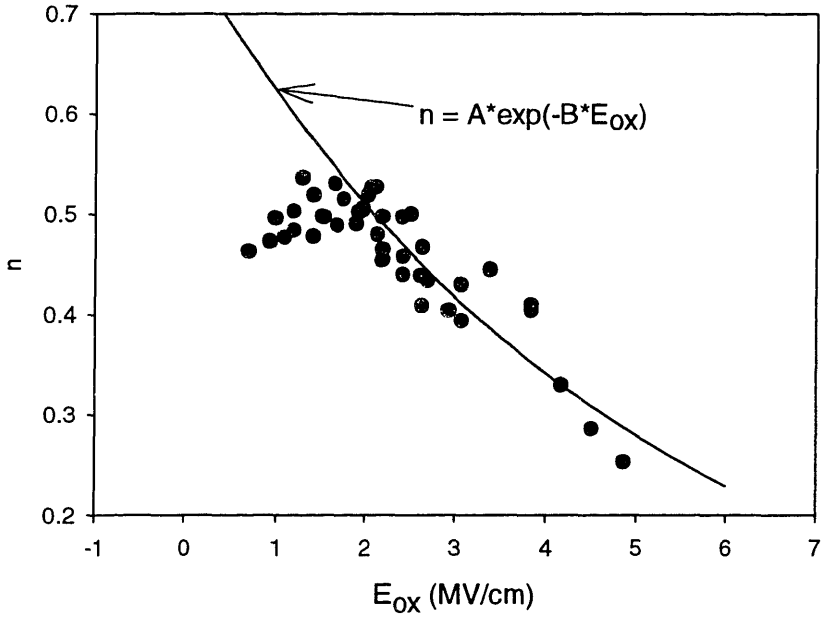

Figure 3.8: Decaying degradation rate as E_{ox} increases beyond certain value.

Now, the obvious question is why the degradation rate decreases at small value of E_{ox}. More specifically, in Figure 3.8, the degradation rate **n** peaks around E_{ox}=1.5MV/cm, and decreases as E_{ox} decreases. Thus, the oxide-field opposition of hot-electron injection does not explain the observed data at low oxide-fields.

Recall from Equation 3.1, however, that when E_{ox} decreases, V_{gs} rises toward V_{ds}. In other words, the device is moving away from the saturation regime toward the linear regime. As a result, the E_{peak} decreases, and thus I_{sub} decreases as well. Summarized in another way, at the low oxide-field, the opposition force against hot-electron injection into Si/SiO_2 interface may not be strong, but there are not as many hot-electrons *generated* to create the interface-trap under this bias condition. Thus, the overall degradation rate decreases under the low oxide-field.

Based on this argument, we substitute I_{sub}/I_d in place of the constant A in Equation 3.2 as follows:

$$n = \frac{I_{sub}}{I_d} \cdot \exp(-B \cdot E_{ox}) \tag{3.3}$$

Notice that, as we correct the model for the low oxide-field, the condition, $E_{ox} > C$ in Equation 3.2, is released. Equation 3.3 argues that the rate of interface-trap generation is determined by both the number of hot-electrons *generated* and the number of hot-electrons *injected* into the Si/SiO_2 interface. I_{sub}/I_d represents the normalized number of hot-electrons generated, and E_{ox} represents the opposition force against injection of the generated hot-electrons.

In order to verify the model, Equation 3.3 is rearranged as follows:

$$\ln\left(\frac{n \cdot I_d}{I_{sub}}\right) = -B \cdot E_{ox} \tag{3.4}$$

Now, with the same data set shown in Figure 3.8, Figure 3.9 is plotted in log-lin space, where E_{ox} is the x-axis, and $\left(\frac{n \cdot I_d}{I_{sub}}\right)$ is the y-axis. As one can see, Equation 3.4 achieves a much better fit with the data than both the empirical model in Figure 3.6 and Equation 3.2 in Figure 3.8. The Constant B in Equation 3.4 can be extracted from the slope in Figure 3.9.

One may notice that at the extreme values of E_{ox}, where $E_{ox} < 0.5$MV/cm and $E_{ox} > 4$MV/cm, the data deviates away from the model. Recall at these extreme values, either V_{gs} is very low, close to V_t, ($E_{ox} > 4$MV/cm) or V_{gs} is very high, close to V_{ds} ($E_{ox} < 0.5$MV/cm). At these bias conditions, other degradation mechanisms, such as hole-trapping and electron-trapping, start to play more important roles [3.6-7]. Thus, modeling the degradation rate in these bias regimes as due to the interface-trap generation may not achieve the accuracy observed in the mid-gate bias region.

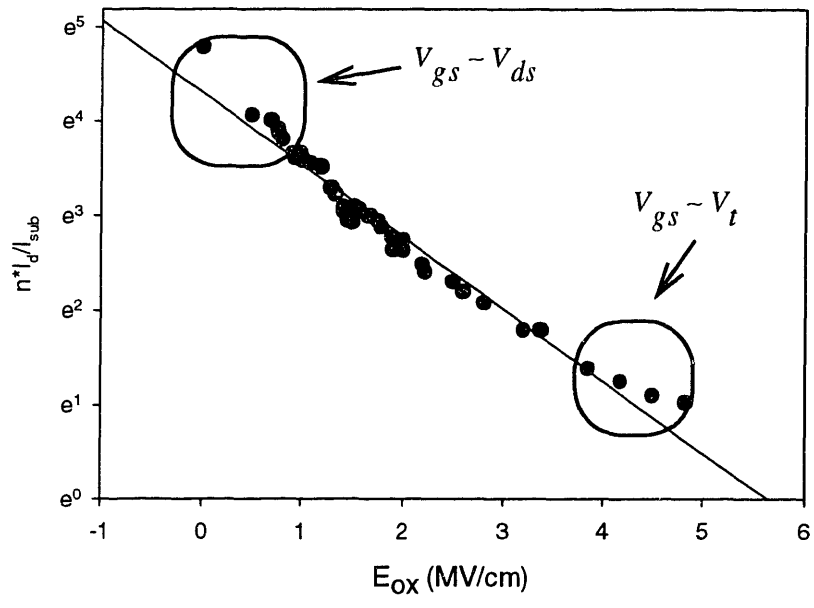

Figure 3.9: Physical model of **n** as a function of I_{sub} and E_{ox}.

3.3 Oxide-Field Dependence of the m and H Coefficients

3.3.1 Lifetime Correlation Plot

The **m** and **H** parameter extraction procedure was explained in Section 2.2.2. To briefly recapitulate the procedure, the parameters **m** and **H** were extracted from the slope and the y-intercept (at y = 1) of the regression line through the data of the **lifetime correlation plot**, in which $\frac{\tau I_d}{W}$ was plotted against $\frac{I_{sub}}{I_d}$ as in Figure 2.2. As was the case with the parameter **n**, stress experiments were carried out to check whether the parameters **m** and **H** would have any dependence on the oxide-field, and the result is shown in Figure 3.10. For each oxide-field shown, the stress conditions V_{ds} and V_{gs} were chosen such that the E_{ox} remained constant at that value. Since the E_{ox} is directly proportional to the difference between V_{gs} and V_{ds}, the difference $|V_{gs} - V_{ds}|$ was kept constant for each oxide-field.

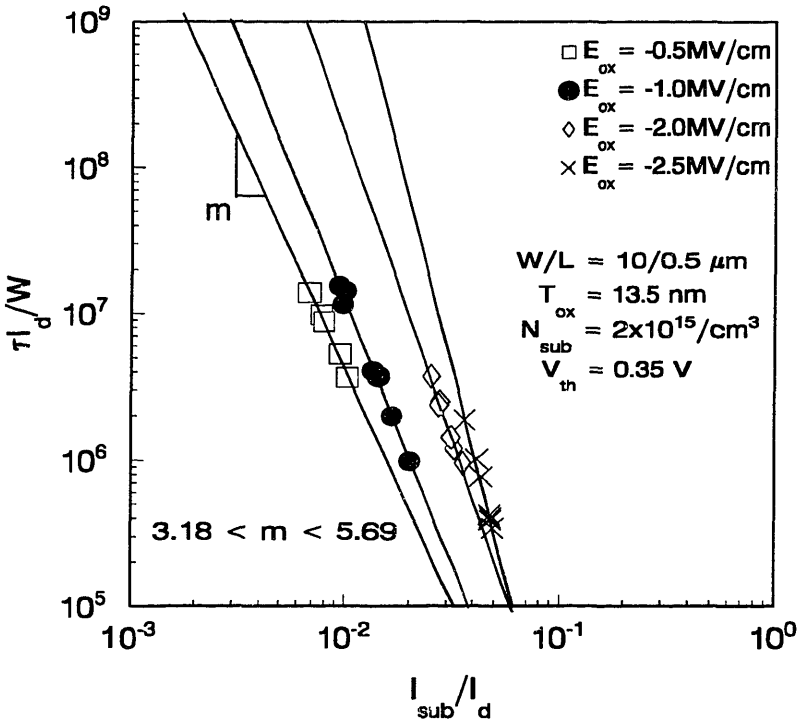

Figure 3.10: Extraction of the parameter **m** and **H** for several E_{ox} values.

As can be clearly seen in Figure 3.10, the parameters **m** and **H** are also oxide-field dependent. Recall that in deriving the degradation model Equation 2.6, the parameter **m** was defined as the ratio $\frac{\varphi_{it}}{\varphi_i}$ where ϕ_{it} is the critical energy the electrons must have in order to create interface-traps, and ϕ_i is the impact ionization critical energy. Since the impact ionization energy ϕ_i is constant at 1.1 eV for Si, independent of the electric field, we deduce that ϕ_{it}, the critical energy required for interface-trap generation, is a function of E_{ox}.

In order to further illustrate the oxide-field dependence of the voltage-acceleration factor **m**, Figure 3.11 shows a plot **m** vs. E_{ox}. We notice that the parameter **m** has a monotonic dependence on E_{ox}. In other words, the higher the E_{ox} is, the higher the value of **m** is. Given its monotonic dependence on E_{ox}, **m** can be empirically modeled as a second order polynomial function of E_{ox} as shown in Figure 3.11.

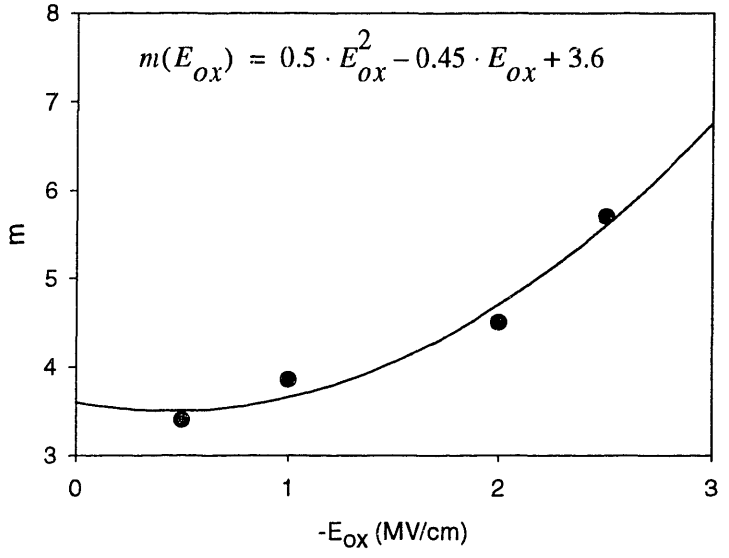

Figure 3.11: The empirical model of **m** as a function of E_{ox}.

3.3.2 Physical Theory of Oxide-Field Dependent m

Based on the above argument that it is ϕ_{it} that varies with the oxide-field while ϕ_i is constant at 1.1eV for Si, independent of the electric field, we can calculate that ϕ_{it} varies from 3.2eV to 5.2eV in Figure 3.11. Unlike the parameter **n**, the oxide-field dependence of **m** has been observed in a previous study [3.17]. This study also reports a similar trend (i.e. increasing ϕ_{it} with increasing E_{ox}) with a similar set of values. However, this past study presents only a qualitative physical explanation for the observed oxide-field dependence of ϕ_{it}. As E_{ox} becomes larger, $|V_{gs} - V_{ds}|$ increases, and thus, the energy band bends more in the Si. Thus, electrons must possess more energy to create the interface-trap; in other words, ϕ_{it} increases.

We can explore this theory further by identifying the different energy components that contribute to the total ϕ_{it}. Shown in Figure 3.12 are the energy band diagrams for the

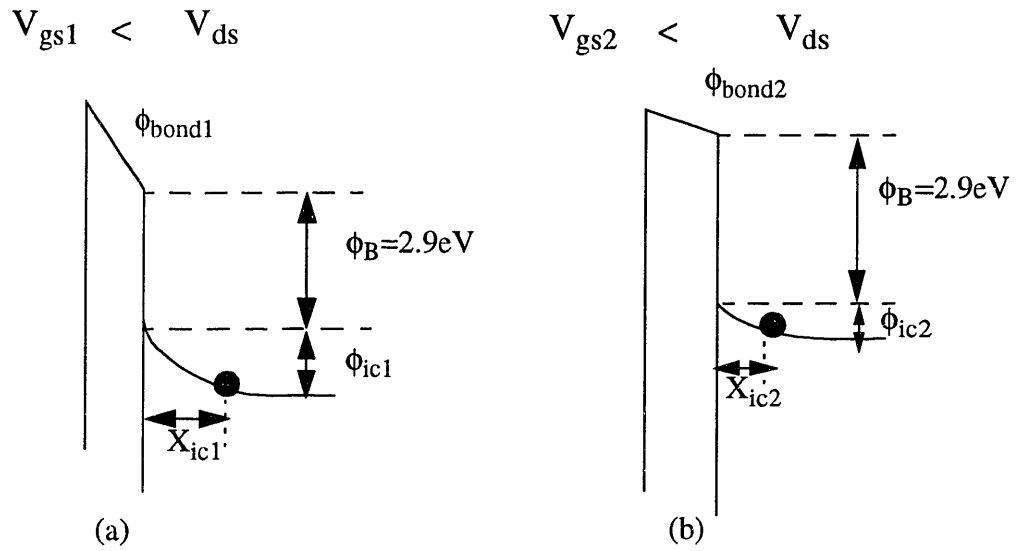

Figure 3.12: Energy band diagram that shows the E_{ox} dependence of ϕ_{it}.

saturation regime in which $V_{ds} > V_{gs}$. Assume for a fixed V_{ds}, V_{gs1} in (a) is less than V_{gs2} in (b). Thus, E_{ox1}, given by $\left|\frac{V_{gs1}-V_{ds}}{T_{ox}}\right|$, is greater than E_{ox2}, given by $\left|\frac{V_{gs2}-V_{ds}}{T_{ox}}\right|$. From the band diagram, we can identify three energy components that constitute the critical energy required to generate interface-traps. Let us write them as follows:

$$\Phi_{it} = \phi_{bond} + \phi_B + \phi_{ic} \tag{3.5}$$

The first term, ϕ_{bond}, is the energy required to break Si-H bond at the Si/SiO_2 interface. Recall from Chapter 1 that the interface-states are created by breaking the Si-H bonds at the interface, and that the interface-traps get filled when the subsequent hot-electron gets trapped in place of the H. Thus, in order to create the interface-traps, the Si-H bond must be first broken, and ϕ_{bond} represents the energy required to break the bond. Typically, ϕ_{bond} is reported to be ~ 0.3 eV [3.16].

The second term, ϕ_B, in Equation 3.5, is the energy barrier height that exists between Si and SiO_2. Although there have been some past analyses on this barrier height

lowering as a function of bias conditions [3.18-19], this effect is more pronounced when the energy band bends the other way, (i.e. $V_{gs} > V_{ds}$). Thus, in the saturation regime where hot-electrons are injected into the Si/SiO_2 interface (i.e. when $V_{gs} < V_{ds}$), ϕ_B can be assumed to be constant. As one can see from the band diagram in Figure 3.12, ϕ_B is determined by the work function difference and the energy band gap of Si and SiO_2, and thus, can be assumed to be constant at 2.9 eV for the bias conditions of interest [3.16].

The third term, ϕ_{ic}, in Equation 3.5, is the energy that hot-electron must possess to travel the distance X_{ic} in order to approach the Si/SiO_2 interface. From the shown band diagram in Figure 3.12, the energy band is bent more in (a) when E_{ox} is greater so that the electrons' current path is deeper into Si away from the Si/SiO_2 interface. Hence, in Figure 3.12, X_{ic1} is greater than X_{ic2}, requiring more energy for electrons to get injected over the Si/SiO_2 interface for the case (a) where E_{ox} is larger. Thus, because of more band bending that occurs in (a) where E_{ox} is larger, ϕ_{ic1} is greater than ϕ_{ic2}. Another component that is implicitly assumed in ϕ_{ic} is the energy loss as hot-electrons collide and become redirected toward the gate for injection. When the current path is deeper into the Si for the case (a), the energy loss component is also greater, and thus, contributing to the greater ϕ_{ic}. These three components quantitatively explain the monotonic behavior of ϕ_{it} as a function of E_{ox}.

Equation 3.5 quantifies ϕ_{it} as a sum of its constituent energies, and clearly explains why ϕ_{it} increases as E_{ox} increases. It is not feasible, however, to derive an equation that can accurately model each energy component because of the difficulty to measure experimentally each energy component as a function of E_{ox}. Furthermore, as will be shown in Chapter 5, it is not necessary to develop an accurate physical model for accurate hot-carrier AC-lifetime prediction. An empirical model shown in Figure 3.11 is sufficient for accurate lifetime prediction procedure. This issue will be further discussed in Chapter 5.

3.4 The Degradation Model Validity

The MOSFET lifetime can be predicted as a function of the power supply voltages from the lifetime correlation plot as in Figure 3.10. Each I_{sub}/I_d corresponds to a power supply voltage in Figure 3.10, and the y-axis gives the device lifetime for that particular supply voltage. As the supply voltage increases, the I_{sub}/I_d ratio increases, and hot-carrier degradation becomes more severe, hence shortening the device lifetime as can be seen in Figure 3.10. The device lifetime under the operating voltage, such as 3.3V for 0.35 μm technology, is calculated by extrapolating from the lifetime under the accelerated stress voltages through the regression line as in Figure 3.10.

This model tacitly assumes that more degradation always occurs at higher I_{sub}/I_d. This is generally true as in Figure 3.10 since the higher I_{sub}/I_d generally corresponds to higher power supply voltage. However, there are some bias conditions under which this relationship does not necessarily hold. An example was pointed out in Figure 3.3, in which the degradation was larger at lower I_{sub}/I_d. This occurred at a very low V_{gs}, which was barely above the threshold voltage V_t. In order to investigate this issue of model validity, numerous devices were stressed at low V_{gs}, and the lifetime correlation is plotted in Figure 3.13. As we can see, this set of data erroneously yields a positive slope, suggesting longer lifetime at higher stress voltages, which is not physically correct.

When using the degradation model Equation 2.6 to calculate the device lifetime based on the experimental data, one should remember that the model is not valid for the data taken from very low V_{gs}. Thus, when planning stress experiments, one should take into account that a sufficient number of devices must be stressed at V_{gs} values which are much greater than V_t.

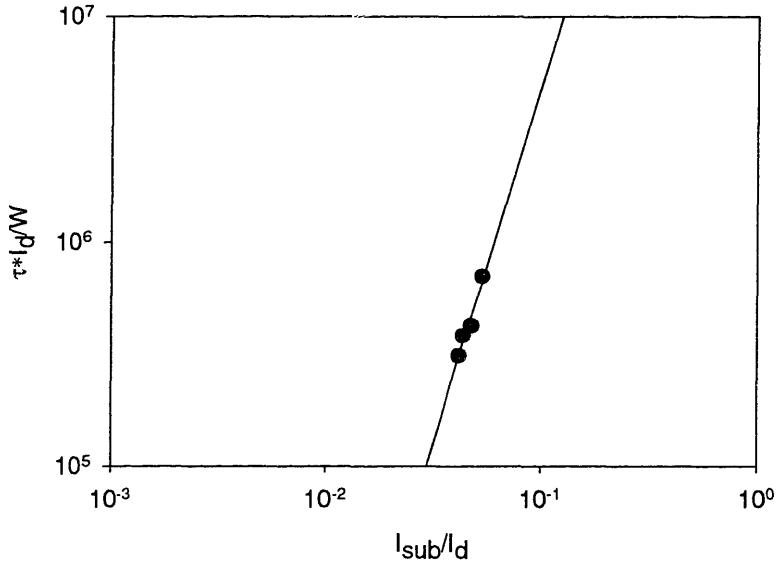

Figure 3.13: Invalidity of lifetime model at low V_{gs}; a positive rather than negative slope is observed.

References

[3.1]. 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, p111, 1983.

[3.2]. T. Tsuchiya, T. Kobayashi, and S. Nakajima, "Hot-Carrier Degradation Mechanism in Si nMOSFETs," Conference on Solid State Devices and Materials, p21, 1985.

[3.3]. P. Heremans, Rudi Bellens, Guido Groeseneken, and Herman Maes, "Consistent Model for the Hot-Carrier Degradation in n-Channel and p-Channel MOSFET's," IEEE Transactions on Electron Devices, Vol. 35, p2194, 1988.

[3.4]. S. Kim, B. Menberu, and J. Chung, "Oxide-Field Dependence of ; NMOS Hot-Carrier Degradation Rate and Its Impact on AC-Lifetime Prediction," International Electron Devices Meeting, p37, 1995.

[3.5]. V. Chan, S. Kim, and J. E. Chung, "Parameter Extraction Guidelines for Hot-Electron Reliability Simulation," International Reliabilty Physics Symposium, p32, 1993.

[3.6]. B. Doyle, M. Bourcerie, J. Marchetaux, and A. Boudou, "Interface State Creation and Charge Trapping in the Medium-to-High Gate Voltage Range ($V_d/2 > V_g > V_d$) During Hot-Carrier Stressing of n-MOS Transistors," IEEE Transactions on Electron Devices, Vol. 37, p244, 1990.

[3.7]. B. Doyle, M. Bourcerie, C. Bergonzoni, K. Mistry, and A. Boudou, "The Generation and Characterization of Electron and Hole Traps Created by Hole Injection During Low Gate Voltage Hot-Carrier Stressing of n-MOS Transistors," IEEE

Transactions on Electron Devices, Vol. 37, p1869, 1990.

[3.8]. J. Winnerl, A. Lill, D. Schmitt-Landsiedel, M. Orlowski, and F. Neppl, " Influence of Transistor Degradation on CMOS Performance and Impact on Lifetime Criterion," International Electron Devices Meeting, p204, 1988.

[3.9]. R. Bellens, P. Heremans, G. Groeseneken, and H. Maes, "A New Procedure for Lifetime Prediction of N-Channel MOS-Transistors Using the Charge Pumping Technique," International Reliability Physics Symposium, p8, 1988.

[3.10]. P. Heremans, R Bellens, G. Groeseneken, and H. Maes, "Consistent Model for the Hot-Carrier Degradation in n-Channel and p-Channel MOSFET's," IEEE Transactions on Electron Devices, Vol. 35, p2194, 1988.

[3.11]. M. Ancona, N. Saks, and D. McCarthy, " Lateral Distribution of Hot-Carrier-Induced Interface Traps in MOSFETs," IEEE Transactions on Electron Devices, Vol.35, p2221, 1988.

[3.12]. W. Hansch, "Modeling Hot-Carrier Reliability of MOSFET: What is necessary and what is possible?," IEEE International Electron Devices Meetings, p717, 1992.

[3.13]. S. Sun, M. Orlowski, and K. Fu, "Parameter Correlation and Modeling of the Power-Law Relationship in MOSFET Hot-Carrier Degradation," IEEE Electron Device Letters, Vol. 11, p297, 1990.

[3.14]. R. Bellens, P. Heremans, G. Groesseneken, and H. Maes, "A New Procedure for Lifetime Prediction of N-Channel MOS Transistors Using the Charge Pumping Technique," IEEE International Reliability Physics Symposium Proceeding, p8, 1988.

[3.15]. S. Mohamedi, V. Chan, J. Park, F. Nouri, B. Scharf, and J. Chung, "Hot-Electron-Induced Input Offset Voltage Degradation in CMOS Differential Amplifiers," IEEE International Reliability Physics Symposium, p 34, 1992.

[3.16]. C. Hu, S. Tam, F. Hsu, P.Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation - Model, Monitor, and Improvement," IEEE Transactions on Electron Devices, Vol. 32, p375, 1985.

[3.17]. 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 Letters, p1188, 1987.

[3.18]. Y. Toyoshima, H. Iwai, F. Matsuoka, H. Hayashida, K. Maeguchi, and K. Kanzaki, "Analysis on Gate-Oxide Thickness Dependence of Hot-Carrier-Induced Degradation in Thin-Gate Oxide nMOSFETs", IEEE Transactions on Electron Devices, Vol. 37, p1496, 1990.

[3.19]. K. Hoffman, C. Werner, and W. Weber, "Hot-Electron and Hole-Emission Effects in Short n-Channel MOSFETs", IEEE Transactions on Electron Devices, Vol. 32, p691, 1985

Chapter 4

Impact of the Oxide-Field Dependent Degradation Model Parameters on the MOSFET AC-Lifetime Prediction

The oxide-field dependence of the degradation parameters **n**, **m**, and **H** gives insight into physical issues of hot-carrier degradation, such as the decreasing degradation rate with increasing E_{ox} due to its opposing force, and the increasing critical energy for creating interface-traps with increasing E_{ox}. However, this oxide field dependence also adds complexity in terms of MOSFET AC-lifetime prediction [4.1-5]. MOSFETs in a circuit experience time-varying bias conditions, thus time-varying oxide-fields. In other words, the parameter values of **n**, **m**, and **H** are constantly changing as a function of time during AC circuit operation, and since the parameter values are different for different oxide-fields, there is no single set of values that can be used in Equation 2.6 in order to calculate the device lifetime in AC circuit operation.

4.1 Recursive Solution for AC Hot-Carrier Degradation

Before we derive a new equation which allows us to calculate the AC device lifetime under the varying bias conditions, let us first review how we calculated the DC device lifetime using Equation 2.6. Recall that once we define a lifetime criterion, such as $\Delta I_d/I_{d0}$ = 10%, then substitution of the lifetime criterion into Equation 2.6 yields

$$\left.\frac{\Delta I_d}{I_{d0}}\right|_{criterion} = \left(\frac{I_d}{WH}\left(\frac{I_{sub}}{I_d}\right)^m \tau\right)^n \tag{4.1}$$

Rearranging Equation 4.1 yields

$$\tau = \left(\frac{\Delta I_d}{I_{d0}}_{criterion}\right)^{\frac{1}{n}} WH(I_{sub})^{-m}(I_d)^{1-m} \tag{4.2}$$

After we extract the degradation parameters **n**, **m**, and **H** by DC stressing (i.e., applying constant V_{gs}, V_{ds}, and V_{bs}) as described in Section 2.2, and measuring I_{sub} and I_d at the corresponding bias conditions, we can calculate the lifetime of the device at its operating voltage by substituting appropriate quantities into Equation 4.2. Notice that this is the DC-lifetime of the device.

However, MOS devices in integrated circuits experience AC waveforms. In other words, the bias conditions for the devices, V_{gs}, V_{ds}, and V_{bs}, are constantly changing. Thus, direct substitution of the parameter values **n**, **m**, and **H**, and the currents I_{sub} and I_d at a single bias condition, into Equation 4.2 is not correct. These quantities change as a function of time in AC circuit operation. The challenge in calculating the AC-lifetime is to figure out how to combine the different **n**, **m**, and **H**, and I_{sub} and I_d values at different bias conditions, or more generally at different E_{ox}.

One way to calculate the degradation or the lifetime in AC waveforms is to break down the AC waveform into small time-steps, in each of which the oxide-field is approximately constant. Then, we can calculate the degradation in each time-step, using Equation 4.1 by substituting corresponding **n**, **m**, & **H**, and I_{sub} & I_d in each time-step since these quantities are constant during each time-step, in which the oxide-field is constant. Finally, we should combine all the degradations in each time-step to calculate the total degradation. This is illustrated in Figure 4.1. For each time-step from t_0 to t_1, from t_1 to t_2, etc., the bias condition, thus the oxide-field, is approximately constant. Hence, the degradation

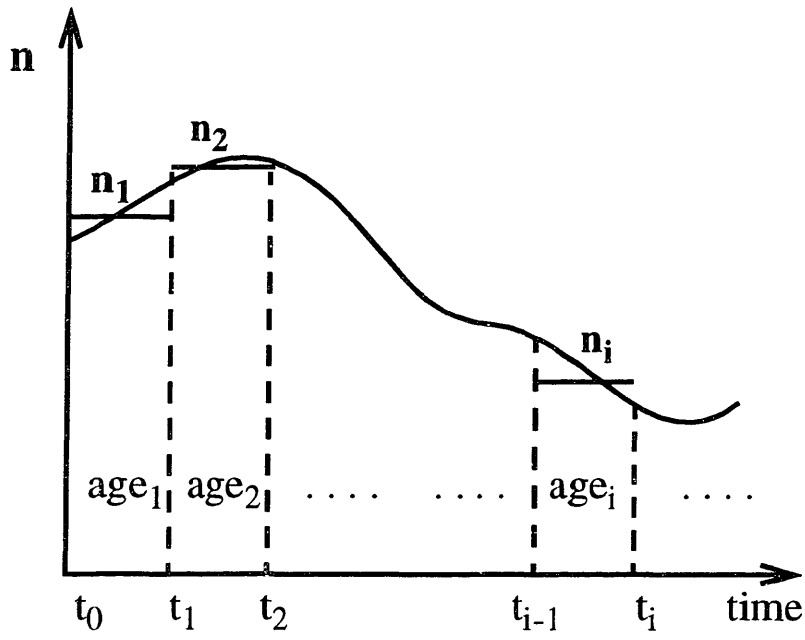

Figure 4.1: Quasi-DC approach to calculate AC-lifetime.

can be easily calculated by using Equation 4.1 in each time-step. In Figure 4.1, among the time-varying quantities such as **n**, **m**, and **H**, and I_{sub} and I_d, the parameter **n** is plotted as an example on the y-axis to illustrate how we can calculate the total degradation, using this piece-wise constant approximation.

In order to facilitate the illustration of how we calculate the total AC degradation, let us define a quantity called **age** as follows in Figure 4.1 [4.4]:

$$age_i = \frac{1}{W}\left(\frac{I_{d,i}}{H_i}\right)\left(\frac{I_{sub,i}}{I_{d,i}}\right)^{m_i}(t_i - t_{i-1}) \tag{4.3}$$

Now, combining Equation 2.5 and Equation 4.3, the degradation that occurred from t_{i-1} to t_i, $\Delta N_{it,\,i}$, can be written as

$$\Delta N_{it,\,i} = (age_i)^{n_i} \tag{4.4}$$

Our eventual task is to calculate the total degradation that occurred from t_0 to t_i, $\Delta N_{it,\,i_{total}}$.

Now, let us proceed to show how we can calculate the total degradation $\Delta N_{it,\,i_{total}}$. In Figure 4.1, up to the time-step t_1 from t_0, the total degradation is easily calculated by using Equation 4.1. During the time-step from t_0 to t_1, all the variables, such as **n**, **m**, and **H**, and I_{sub} and I_d, are approximately constant, and thus, age_1 is calculated by substituting appropriate quantities to Equation 4.3. The total degradation up to t_1, by using Equation 3.5, is

$$\Delta N_{it,\,1_{total}} = (age_1)^{n_1} \tag{4.5}$$

Similarly, the degradation from the time step t_1 to t_2 can be easily calculated by using Equation 4.4 as follows:

$$\Delta N_{it,\,2} = (age_2)^{n_2} \tag{4.6}$$

The challenging task, however, is to figure out how we can compute the total degradation from t_0 to t_2, $\Delta N_{it,\,2_{total}}$. The complexity arises because $n_1 \neq n_2$. If $n_1 = n_2$, then the total degradation up to t_2 is simply

$$\Delta N_{it,\,2_{total}} = (age_1 + age_2)^{n_2} \tag{4.7}$$

as in the DC case. However, when $n_1 \neq n_2$ as in the AC case, simple summation of ages is not valid.

In order to calculate correctly the total degradation up to t_2, $\Delta N_{it,\,2_{total}}$, the following procedure should be used.

1. Calculate the total cumulative effective **AGE**, $\mathbf{AGE_{eff}}$, that would have resulted at t_1 if the device were stressed from t_0 to t_1 at the degradation rate $\mathbf{n_2}$ rather than at $\mathbf{n_1}$.

   The $\mathbf{AGE_{eff}}$ can be solved by equating the total degradation at t_1 to $(AGE_{eff})^{n_2}$ as follows:

$$\Delta N_{it,\,1_{total}} = [age_1]^{n_1} = (AGE_{eff})^{n_2} \tag{4.8}$$

2. Solve Equation 4.8 for AGE_{eff}.

$$AGE_{eff} = [age_1]^{\frac{n_1}{n_2}} \tag{4.9}$$

3. Substitute Equation 4.9 into Equation 4.7, which can now be used since the degradation rates in the two time periods are now identical (i.e, $\mathbf{n} = \mathbf{n_2}$).

$$\Delta N_{it,\,2_{total}} = [AGE_{eff} + age_2]^{n_2} \tag{4.10}$$

4. Substituting Equation 4.9 for AGE_{eff} in Equation 4.10 yields

$$\Delta N_{it,\,2_{total}} = \left[age_1^{\frac{n_1}{n_2}} + age_2 \right]^{n_2} \tag{4.11}$$

where Equation 4.11 indicates the exact amount of AC degradation after t_2.

Notice, that in the case of a constant degradation rate (i.e., $n_1 = n_2 = n$) as in DC stressing, Equation 4.11 simplifies to

$$\Delta N_{it,\,2_{total}} = (age_1 + age_2)^{n} \tag{4.12}$$

which was derived in Equation 4.7 as a special case of constant **n** (i.e, $\mathbf{n_1} = \mathbf{n_2}$).

In order to calculate the exact amount of total AC degradation after t_i, Equation Equation 4.11 can be generalized to:

$$\Delta N_{it,\,i_{total}} = \left[\left(age_1^{\frac{n_1}{n_2}} + age_2 \right)^{\frac{n_2}{n_3}} + \ldots + age_i \right]^{n_i} \tag{4.13}$$

More concisely, Equation 4.13 can be written as:

$$\Delta N_{it,\,i_{total}} = \left(\left(\frac{\Delta I_d}{I_{d0}} \right)_{i-1,\,total}^{\frac{1}{n_i}} + age_i \right)^{n_i} \tag{4.14}$$

where the total degradation at $t = t_i$ (i.e., $\Delta N_{it,\,i_{total}}$) depends on the value of $\Delta N_{it,\,i-1_{total}}$ from the previous time step. Equation 4.14 states that the exact amount of total AC degradation at time-varying degradation rates can only be calculated using a

recursive solution. In other words, unless the degradation rate **n** is a constant, there is no simple, analytical formula into which a future time point of interest (such as 10 years) can be substituted and the exact amount of total AC degradation be calculated.

4.2 Two-Stage Example

For a typical integrated circuit manufactured, 10 years is a usually desired lifetime. If we assume a clock frequency of 100MHz (1 cycle = 10 ns), 10 years correspond to approximately 3.15×10^{16} cycles. Recall from above that the exact solution for AC-lifetime is not in a closed form, but rather in recursive form. This means that, if 1 cycle is broken down to 100 time steps, in each of which the oxide-field is approximately constant, one has to iterate more than 10^{18} times, using Equation 4.14, in order to calculate the exact amount of AC degradation at 10 years. Clearly, this is computationally infeasible.

The current practice of AC-lifetime prediction, called the $\mathbf{n_{eff}}$ algorithm, is to calculate the exact amount of AC degradation, using the recursive solution Equation 4.14, only for the first few cycles, and then extrapolate the AC degradation all the way out to 10^{16} cycles (10 years). This is schematically shown in Figure 4.2.

In order to illustrate the limitations of the existing $\mathbf{n_{eff}}$ extrapolation algorithm, we have simulated AC hot-carrier degradation, in which an NMOS device undergoes a simple AC waveform, which just consists of two, periodically alternating DC waveforms. Figure 4.3 shows the two alternating DC stressing conditions that comprise this simple AC waveform. Notice that, for this example, V_{ds} is fixed at 5.2V, whereas V_{gs} alternates between 1V and 2.5V.

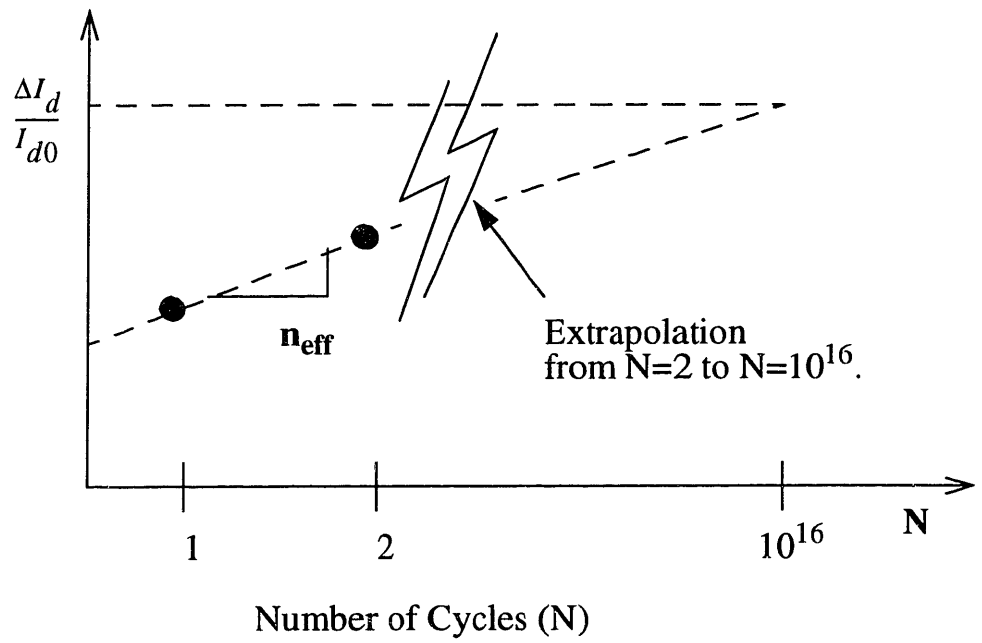

Figure 4.2: Existing n_{eff} algorithm to calculate AC degradation at 10 years.

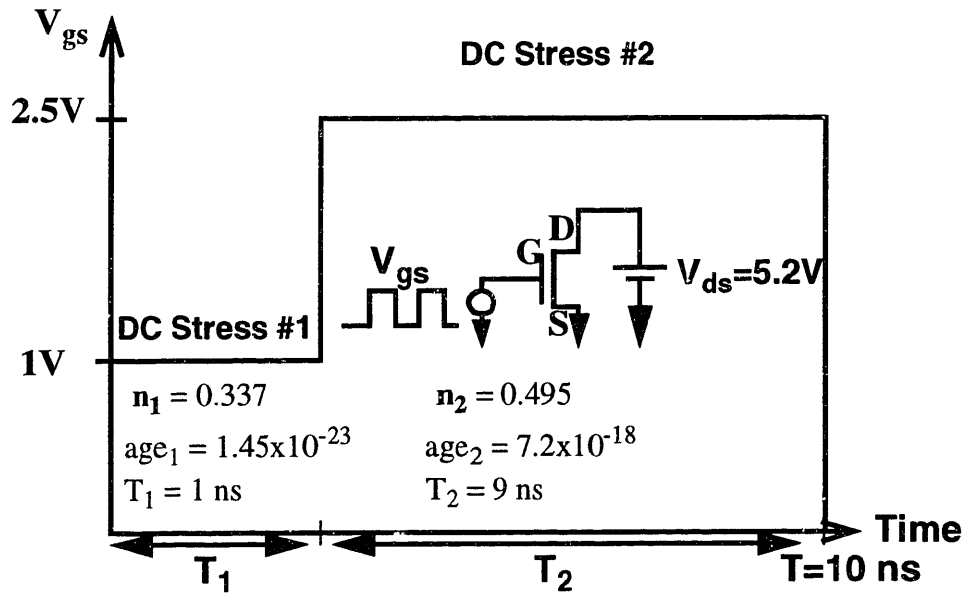

Figure 4.3: A simple AC waveform comprised of two DC stress conditions.

With the given information in Figure 4.3 for the two-stage example, such as age_1, age_2, $\mathbf{n_1}$, and $\mathbf{n_2}$, we have derived exact degradation equations, using similar steps as discussed in Section 4.1, and they are shown in Table 4.1. For comparison, the first column of the table shows the degradation equations for the DC case when the parameter **n** is constant, and the second column shows the degradation equations for this two-stage AC case when **n** alternates between $\mathbf{n_1} = 0.337$ and $\mathbf{n_2} = 0.495$. It is important to notice that even for this simple AC case, the degradation equation is quite complex and is in recursive form rather than in closed form as shown by Equation (2b) in Table 4.1.

	Constant **n**	Variable $\mathbf{n}(E_{ox}) = \{\mathbf{n_1}, \mathbf{n_2}\}$
1 Cycle (T)	(1a) $\frac{\Delta I_d}{I_{d0}}(T) = (AGE_T)^n$ $AGE_T = AGE_1 + AGE_2$	(1b) $\frac{\Delta I_d}{I_{d0}}(T_1) = (AGE_1)^{n1}$ $\frac{\Delta I_d}{I_{d0}}(T) = \left[AGE_1^{\frac{n_1}{n_2}} + AGE_2 \right]^{n_2}$
N Cycle (NT)	(2a) $\frac{\Delta I_d}{I_{d0}}(NT) = \left[\sum_{i=1}^{N} AGE_T \right]^n$	(2b) $\frac{\Delta I_d}{I_{d0}}(NT - T_2) = \left(\left(\frac{\Delta I_d}{I_{d0}}((N-1)T) \right)^{\frac{1}{n_1}} + AGE_1 \right)^{n_1}$ $\frac{\Delta I_d}{I_{d0}}(NT) = \left(\left(\frac{\Delta I_d}{I_{d0}}(NT - T_2) \right)^{\frac{1}{n_2}} + AGE_2 \right)^{n_2}$

Table 4.1 : Degradation equations for the two-stage example.

In order to check the validity of the existing $\mathbf{n_{eff}}$ algorithm for this two-stage example, we have applied the $\mathbf{n_{eff}}$ algorithm in order to calculate the AC-lifetime, and compared it with the exact AC-lifetime calculated by the recursive solution, Equation (2b) in Table 4.1. The result is shown in Figure 4.4. For this example, we have defined the lifetime criterion to be $\left(\frac{\Delta I_d}{I_{d0}}\right) = 1\%$. Notice in Figure 4.4 that the existing $\mathbf{n_{eff}}$ algorithm calculates the AC degradation, using Equation (2b) in Table 4.1, for the first two cycles (shown by x), and extrapolates the AC degradation all the way out to 10 years based on the first two cycles' degradation. The extrapolation is shown by solid line. On the same plot, we have also calculated the AC degradation, using Equation (2b) in Table 4.1, up to 10^8 cycles, (shown by filled diamond), and projected the AC degradation based on the asymptotic solution (shown by hollow diamond). In order to calculate the AC degradation up to 10^8 cycles recursively, using Equation (2b) in Table 4.1, it took about 14 days on a Sun

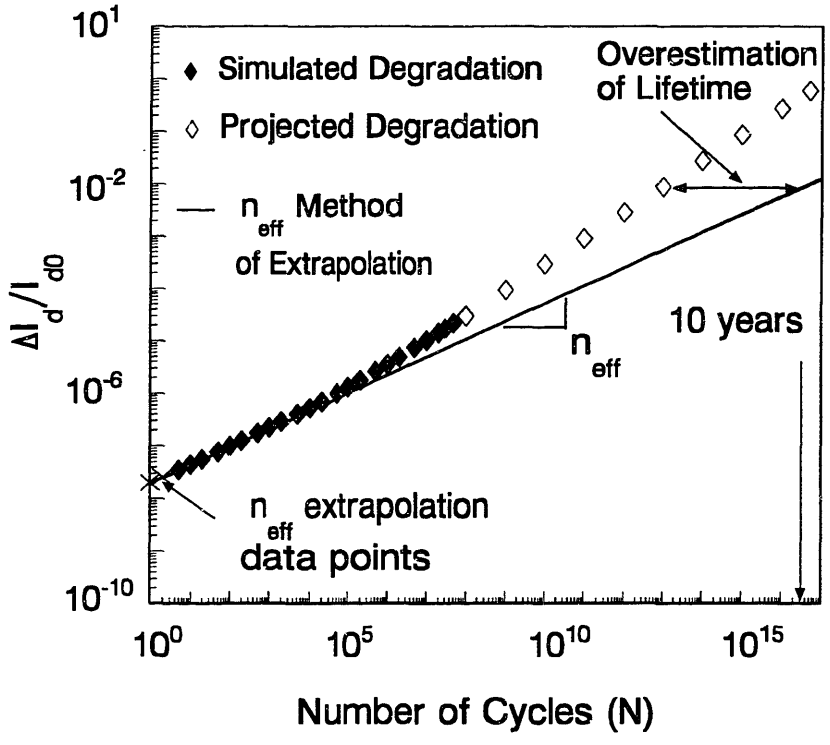

Figure 4.4: Overestimation of AC-Lifetime by the $\mathbf{n_{eff}}$ algorithm.

Sparc4 Workstation. Calculating the exact AC degradation up to 10^{16} cycles(= 10 years) recursively is practically impossible.

Notice that, when the lifetime criterion is $\left(\frac{\Delta I_d}{I_{d0}}\right)$ =1%, the existing $\mathbf{n_{eff}}$ algorithm overestimates the AC-lifetime by almost 3 orders of magnitude. An immense amount of error is introduced into AC-lifetime calculation by the $\mathbf{n_{eff}}$ algorithm. This example clearly demonstrates that, for AC degradation in which the MOSFET undergoes various degradation rates due to changing bias conditions, and thus, changing oxide-field, the traditional method of extrapolating all the way out to 10 years based on the first few cycles' degradation becomes invalid. This example serves as a motivation to develop a new, efficient algorithm for accurate AC-lifetime prediction, which is discussed in Chapter 5.

References

[4.1]. S. Kim, B. Menberu, and J. Chung, "Oxide-Field Dependence of the NMOS Hot-Carrier Degradation Rate and Its Impact on AC-Lifetime Prediction," International Electron Devices Meeting, p37, 1995.

[4.2]. V. Chan, S. Kim, and J. E. Chung, "Parameter Extraction Guidelines for Hot-Electron Reliability Simulation," International Reliability Physics Symposium, p32, 1993.

[4.3]. 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 Letters, p1188, 1987.

[4.4]. C. Hu, S. Tam, F. Hsu, P.Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation - Model, Monitor, and Improvement," IEEE Transactions on Electron Devices, Vol. 32, p375, 1985.

[4.5]. P. Ko, R. Muller, and C. Hu, "A Unified Model for Hot-Electron Currents in MOSFETs," IEEE International Electron Devices Meeting, p600, 1981.

Chapter 5

Dominant DC Degradation Asymptote (DDA) Algorithm for AC-Lifetime Prediction

5.1 Analysis of the n_{eff} algorithm

We have clearly demonstrated through the two-stage example in Chapter 4 that the $\mathbf{n_{eff}}$ algorithm can lead to significant amount of overestimation of AC-lifetime. Before we proceed to develop a new algorithm or method to predict the AC-lifetime more accurately, we need to summarize and analyze why the existing $\mathbf{n_{eff}}$ algorithm failed in the two-stage example.

First, we have shown that even for the simplest AC-case, in which only two DC periods alternate, the degradation equation is quite complex as shown in Table 4.1. The reason for this complexity comes directly from the nature of the hot-carrier degradation behavior, which has a non-linear, power-law time dependence. Since the degradation rate **n** appears in the exponential term in the degradation model, Equation 2.6, the total degradation cannot be computed by simply adding degradation in each time step, but must be computed by using the recursive solution as explained in Section 4.1.

Second, it is important to remember that the $\mathbf{n_{eff}}$ algorithm failed to predict the AC-lifetime even though it used the complex, recursive solution derived in Section 4.1. For the reader's convenience, the two-stage stress conditions and the $\mathbf{n_{eff}}$ AC-lifetime prediction procedure are replotted in Figure 5.1 and 5.2. Failure of the $\mathbf{n_{eff}}$ algorithm to predict the AC-lifetime accurately in Figure 5.2 is caused by that the extrapolation was based on the first few cycles' degradation, which was dominated by only one of the two stress conditions. More specifically, during the early stage of the AC degradation (up to 10^4

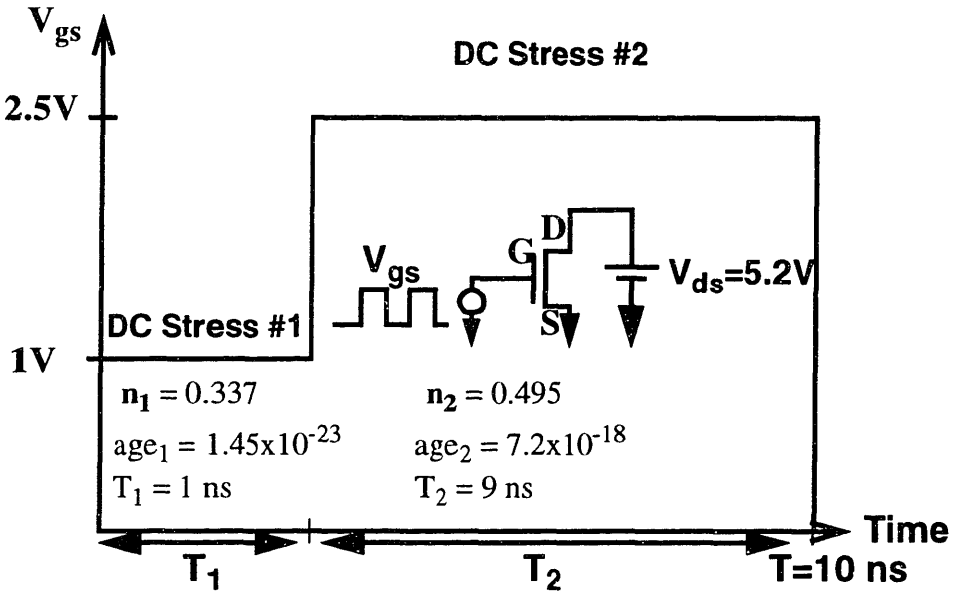

Figure 5.1: A simple AC waveform comprised of two DC stress conditions.

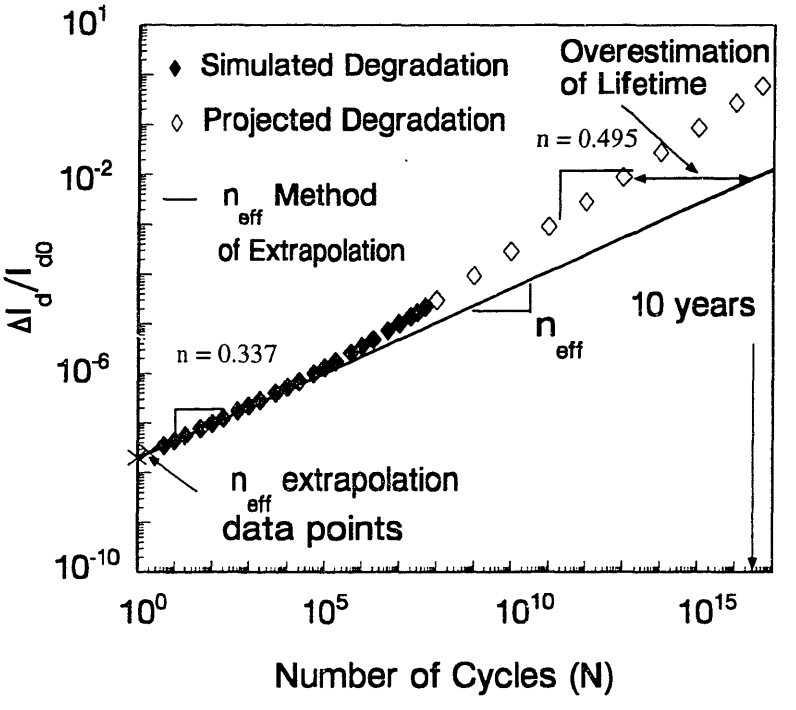

Figure 5.2: Overestimation of AC-Lifetime by the n_{eff} algorithm.

cycles), the degradation is dominated by DC #1 stress condition. As one can see in Figure 5.2, the AC degradation initially has a slope of n = 0.337, which is the slope of DC #1 stress condition. During the early stage of the AC degradation, the contribution of DC #2 stress condition is very minimal.

However, as one can see in Figure 5.2, at the time point of interest, which is 10 years that correspond to approximately 10^{16} cycles, DC #2 stress condition clearly dominates the overall AC degradation. The total AC degradation at the later stage (much beyond 10^4 cycles) has a slope of n = 0.495, which is the slope of DC #2 stress condition.

The reason why the existing $\mathbf{n_{eff}}$ extrapolation technique leads to significant overestimation of AC-lifetime in Figure 5.2 is that there exists a latent degradation component, namely DC #2 stress condition, which has smaller amount of degradation initially but has a higher degradation rate **n**. As stress time progresses, this component eventually dominates the overall AC degradation. Since the existing $\mathbf{n_{eff}}$ algorithm bases its extrapolation on the AC degradations of just the first few waveform cycles (which can be dominated by a degradation component that does not necessarily dominate at the future time point of interest), significant inaccuracy in the predicted AC-lifetime can be introduced [5.1-5].

5.2 Concept of the Dominant DC Degradation Asymptote (DDA)

As it was discussed above, different DC degradation components dominate the overall AC degradation at different times. Thus, the key to accurate AC-lifetime prediction is to identify the dominant DC degradation component at the future time point of interest and project the AC-lifetime based on such a dominant DC degradation component at the interested time point.

For the two-stage example demonstrated above, DC component #1 dominates the AC degradation during the early stage, and DC component #2 dominates the AC degradation during the more relevant time period (t = 10 years). Upon closer inspection of the AC degradation behavior, we observe that the AC degradation at any particular time point is dominated by one of the following equations as shown in Figure 5.3:

$$\frac{\Delta I_d}{I_{d0}} = (age_1 \times N)^{0.337} \tag{5.1}$$

$$\frac{\Delta I_d}{I_{d0}} = (age_2 \times N)^{0.495} \tag{5.2}$$

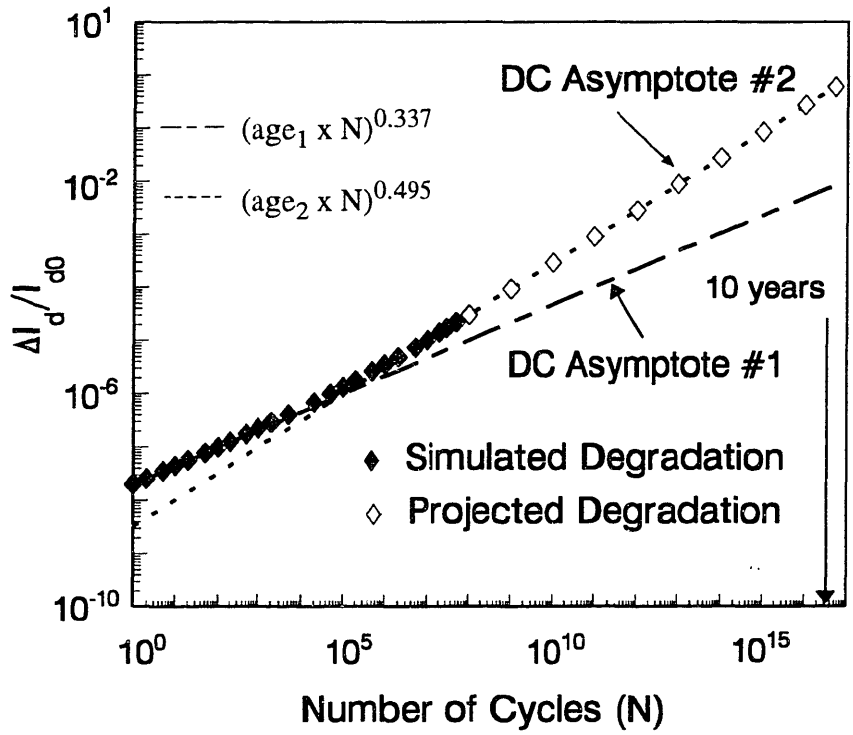

Figure 5.3: DC degradation asymptotes.

where age_1 and age_2 are given in Figure 5.1, and N is the number of waveform cycles that the device has experienced. Recall that age is a quantity that represents the amount of hot-carrier stress experienced by the MOSFET. Thus, age_1 represents the amount of hot-carrier stress experienced for one cycle through DC #1 stress condition in the two-stage example. Age_2 represents the corresponding quantity experienced through DC #2 stress condition.

As one can see in Figure 5.3, the AC degradation at any particular time point is dominated by either of the two equations. As a matter of fact, the AC degradation asymptotically approaches either of the two equations at all times. More specifically, during the early stage of degradation when DC #1 stress condition plays the dominant role, the AC degradation approaches DC #1 asymptote given by Equation 5.1. Correspondingly, during the later stage of degradation, the AC degradation approaches DC #2 asymptote given by Equation 5.2.

Thus, we define Equation 5.1 and 5.2 above to be DC asymptotes. In summary, there are several key observations and insights we draw from this two-stage example. First, the AC degradation, experienced by the MOSFET through an AC waveform, is always dominated by one of the DC components. Second, the AC degradation dominated by one of the DC component approaches a DC asymptote, which is in the form of Equation 5.1. The dominant DC component yields the dominant DC asymptote. Third, different DC components, and thus, different DC asymptotes dominate the AC degradation at different times. In other words, the dominant DC degradation asymptote (DDA) changes as a function of stress time. Therefore, for more general AC waveform, the key to predicting the AC-lifetime accurately is to identify the dominant DC component at the interested

time point, and project the AC-lifetime based on the dominant DC degradation asymptote (DDA).

5.3 The Dominant DC Degradation Asymptote (DDA) Algorithm

Based on the concept of the dominant DC degradation asymptote as explained above, we have developed the following algorithm, which would allow to calculate AC-lifetime accurately. We explain below a step-by-step procedure that one can follow in order to calculate the AC-lifetime of a MOSFET undergoing a general AC waveform.

5.3.1 Degradation Model Parameter Characterization

The first step is to stress the MOS transistors at multiple bias conditions as shown in Section 2.2, and extract the degradation model parameters **n**, **m**, and **H**. Recall that these parameters are a function of the oxide field, E_{ox}, and thus, should be modeled as a function of E_{ox}. For a given technology, one should extract and characterize the degradation rate $\mathbf{n}(E_{ox})$ and the voltage-acceleration factor $\mathbf{m}(E_{ox})$ as shown in Figure 5.4.

5.3.2 Circuit Simulation

The second step is to simulate the desired circuit for one waveform cycle. Simulate the circuit with SPICE or its variants, and obtain V_{gs}, V_{ds}, E_{ox}, I_{ds}, and I_{sub} waveforms vs. time for the MOSFET device of interest. Conventional SPICE calculates V_{gs}, V_{ds}, and I_{ds} for the MOSFET. A substrate current model [5.6] must be implemented in order for SPICE to calculate I_{sub} as a function of V_{gs} and V_{ds}. E_{ox} can be calculated as $\frac{V_{gs} - V_{ds}}{T_{ox}}$.

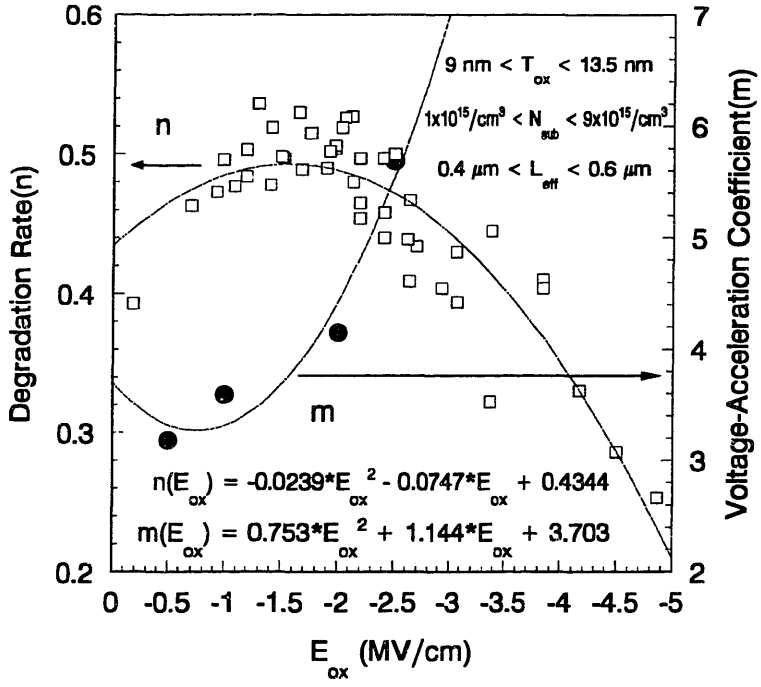

Figure 5.4: Characterization of the **n** and **m** as a function of E_{ox}.

5.3.3 Calculation of the Degradation Model Parameters

Based on the characterization of the degradation model parameters in Section 5.3.1 and the simulation of circuit waveforms vs. time in Section 5.3.2, one can calculate the degradation model parameter values for each time step. Calculate **n**, **m**, and **AGE** as a function of time for one waveform cycle. Also, compute the total cumulative AGE in one waveform cycle ($\mathbf{AGE_T}$).

Then, based on the calculated parameter values, plot **n** vs. time and **AGE** vs. time for one waveform cycle. From these plots, one can plot **n** vs. **AGE** by matching **n** with **AGE** at every time point. This is schematically illustrated in Figure 5.5.

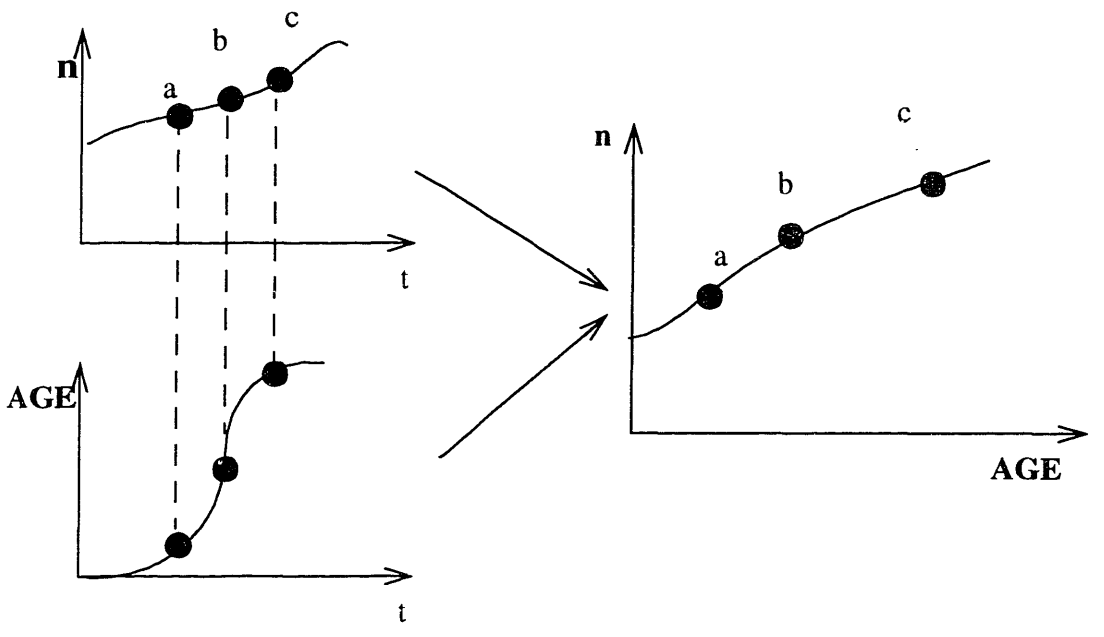

Figure 5.5: Plotting **n** vs. **AGE** for one waveform cycle.

5.3.4 Calculation of Weighted Average Degradation Rate: n_{ave}

In order to calculate the dominant degradation rate, n_{dom}, we begin by calculating the average degradation rate, n_{ave}, which is calculated by weighting the AGE. Recall that the AGE represents the amount of hot-carrier stress experienced by the MOSFET. The procedure of calculating the average degradation rate weighted by AGE is schematically shown in Figure 5.6. Once the plot of **n** vs. AGE for one waveform cycle is obtained as shown in Figure 5.6, we can compute the average degradation rate as follows:

$$n_{ave} = \frac{\int_0^b n(AGE)dAGE}{AGE_T} \tag{5.3}$$

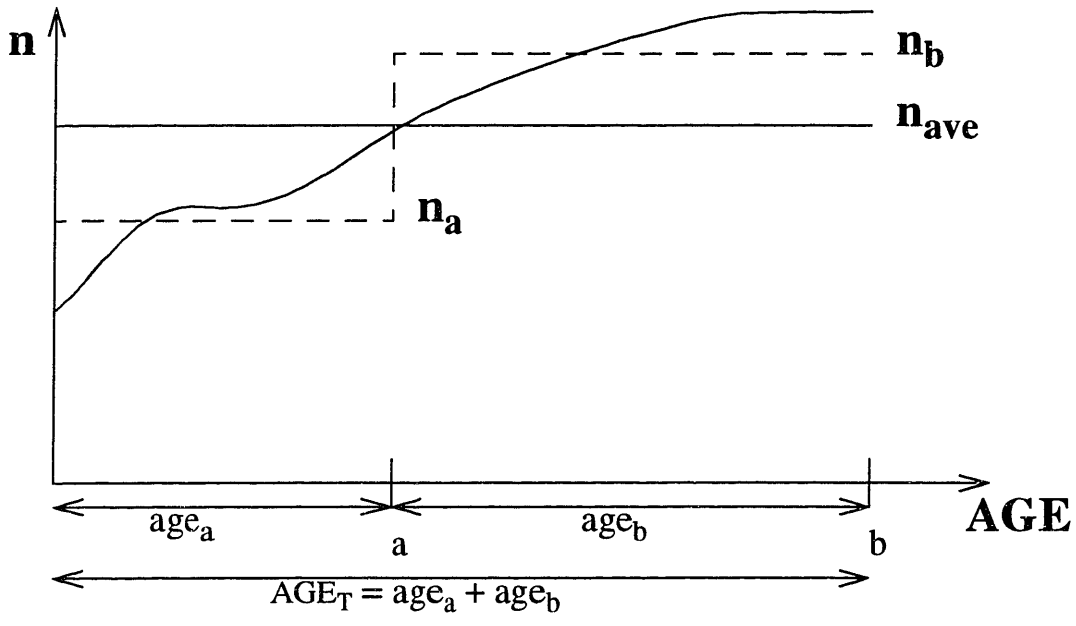

Figure 5.6: Calculation of weighted average **n** by AGE.

where AGE_T is a total cumulative AGE for one waveform cycle, and b is shown above in Figure 5.6.

After the $\mathbf{n_{ave}}$ is calculated, define two sub-averages $\mathbf{n_a}(\mathbf{n} < \mathbf{n_{ave}})$ and $\mathbf{n_b}(\mathbf{n} > \mathbf{n_{ave}})$, each weighted by the subset of corresponding ages as follows:

$$n_a = \frac{\int_0^a n(AGE)dAGE}{age_a} \tag{5.4}$$

$$n_b = \frac{\int_a^b n(AGE)dAGE}{age_b} \tag{5.5}$$

where a, b, age_a, and age_b are as shown above in Figure 5.6.

5.3.5 Projection of DC Asymptotes

After the weighted degradation rate $\mathbf{n_{ave}}$, $\mathbf{n_a}$, and $\mathbf{n_b}$ are calculated, the next step is to project the DC asymptotes on $\frac{\Delta I_d}{I_{d0}}$ vs. N (number of waveform cycles) plot. After having calculated the $\mathbf{n_{ave}}$, $\mathbf{n_a}$, and $\mathbf{n_b}$, we can lay down the following three DC asymptotes as shown in Figure 5.7.

$$\frac{\Delta I_d}{I_{d0}} = (AGE_T \times N)^{n_{ave}} \qquad \text{(Asymptote "ave")} \qquad (5.6)$$

$$\frac{\Delta I_d}{I_{d0}} = (age_a \times N)^{n_a} \qquad \text{(Asymptote "a")} \qquad (5.7)$$

$$\frac{\Delta I_d}{I_{d0}} = (age_b \times N)^{n_b} \qquad \text{(Asymptote "b")} \qquad (5.8)$$

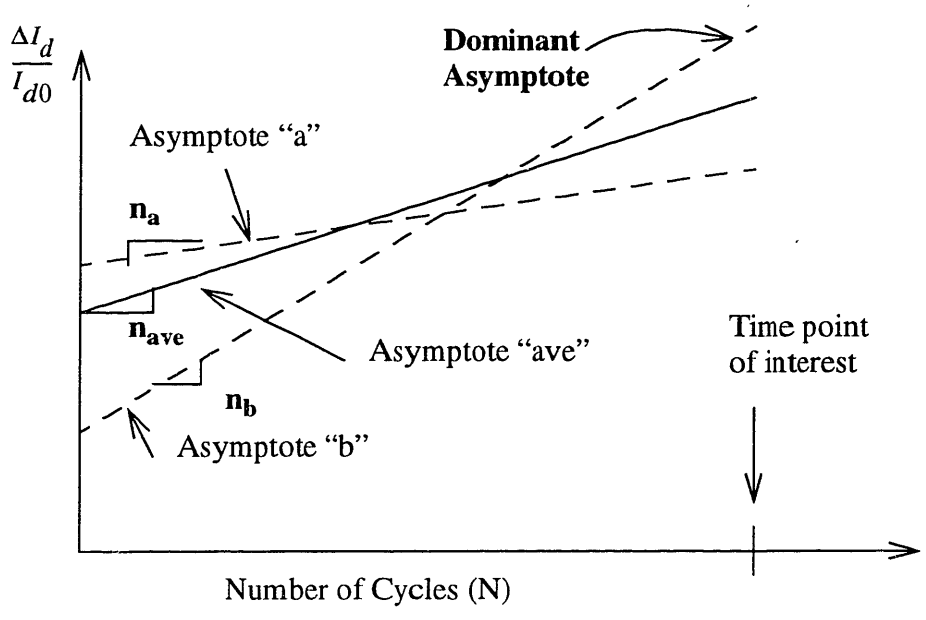

Figure 5.7: Projection of DC asymptotes.

5.3.6 Identification of the Dominant DC Asymptote (DDA)

We call Asymptote "a" and Asymptote "b" "twins" of the original Asymptote "ave." in Figure 5.7. The purpose of "twinning" the original Asymptote "ave" is to identify which age component and **n** value will dominate the AC degradation at the future time point of interest. After "twinning" the original Asymptote "ave," one should continue the "twinning" procedure with each asymptote, Asymptote "a," and Asymptote "b." The new resulting asymptotes are:

$$\frac{\Delta I_d}{I_{d0}} = (age_{aa} \times N)^{n_{aa}} \tag{5.9}$$

$$\frac{\Delta I_d}{I_{d0}} = (age_{ab} \times N)^{n_{ab}} \tag{5.10}$$

$$\frac{\Delta I_d}{I_{d0}} = (age_{ba} \times N)^{n_{ba}} \tag{5.11}$$

$$\frac{\Delta I_d}{I_{d0}} = (age_{bb} \times N)^{n_{bb}} \tag{5.12}$$

Equation 5.9 and 5.10 are "twins" of Asymptote "a", and Equation 5.11 and 5.12 are "twins" of Asymptote "b." This twining procedure is continued until the age_{xx} is negligibly small.

After the "twinning" procedure is completed, we can identify the dominant DC asymptote (DDA) by projecting each asymptote and observing which asymptote yields the maximum degradation at the future time point of interest. The asymptote which yields the most degradation at the interested time point becomes the DDA, based upon which the

AC-lifetime is calculated. For the example in Figure 5.7 where subsequent "twins" are not shown for simplicity and clarity of the plot, the Asymptote "b" is the dominant DC asymptote, and thus, we predict the AC-lifetime based on Asymptote "b."

Recall that the DDA does not give a complete path of the AC degradation over time. During the earlier stage of degradation, AC-degradation might follow some other DC asymptotes. However, by following the above procedure, we are assured that the AC-degradation will follow the DDA at the future interested time point because the DDA is computed by identifying the dominant degradation component at the future interested time point.

5.4 Demonstration of the DDA Algorithm

5.4.1 Verification of the Algorithm for the Two-Stage Example

Before we apply the DDA algorithm to a more general AC case, we have applied it to the two-stage example in order to verify the validity of the algorithm. We have followed the steps given from Section 5.3.1 to 5.3.6. Figure 5.8 shows the final step, in which the calculated AC degradation and the dominant DC degradation asymptote are shown. As can be seen, the DDA algorithm correctly predicts the AC degradation at 10 years.

5.4.2 Application of the DDA Algorithm to a CMOS Inverter

In order to demonstrate the usage of the DDA algorithm, we have applied the algorithm to an NMOS device in a CMOS inverter shown in Figure 5.9. The task is to calculate the AC-lifetime of the NMOS device as it experiences a time-varying AC waveform in the CMOS inverter. Primary design parameters for a CMOS inverter includes the input ramp

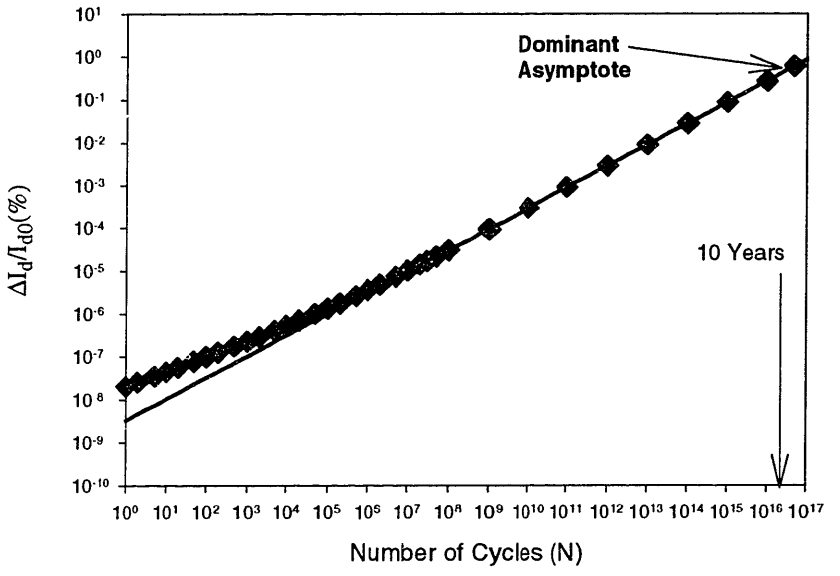

Figure 5.8: Application of the DDA algorithm to the two-stage example.

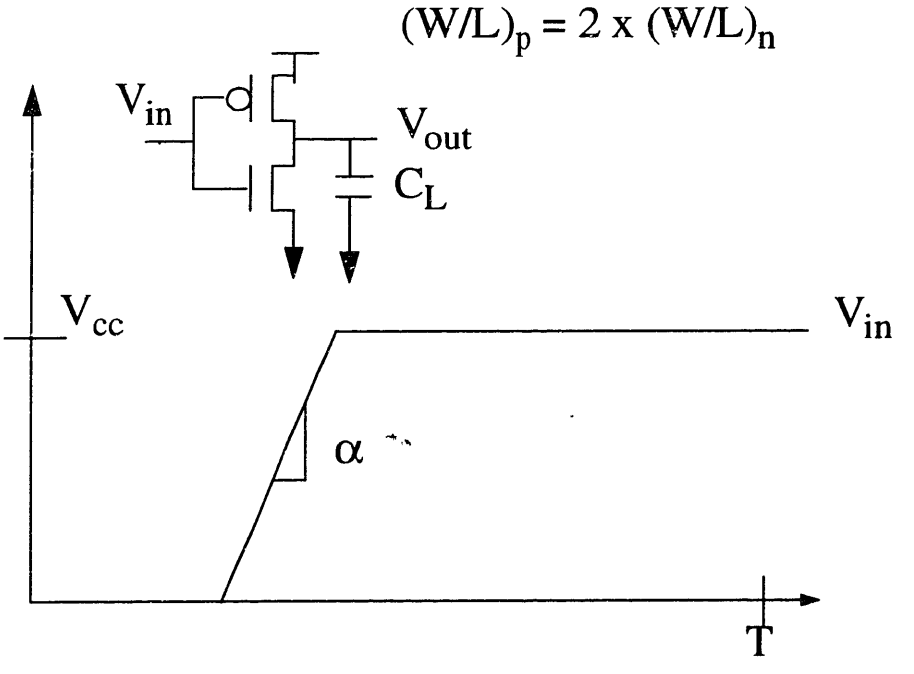

Figure 5.9: CMOS inverter along with definitions of the primary design parameters.

rate, shown as α, in Figure 5.9, and the load capacitance, C_L. We have applied the algorithm over a wide range of these design parameters in order to investigate the relationship between the AC-lifetime and the primary design parameters of an inverter. Notice that the PMOS device is scaled so that the output rise time and fall time are approximately equal. For this calculation, the lifetime criterion of $\frac{\Delta I_d}{I_{d0}}$ = 10% was used. T stands for one clock cycle in Figure 5.9.

The first step in the DDA algorithm is to characterize the degradation model parameters, **n**, **m**, and **H**, as a function of the oxide field, E_{ox}, for a given technology. As discussed in Chapter 3, this procedure involves DC stressing of multiple devices at numerous bias conditions. For this calculation, the parameters extracted in Figure 5.4 are used.

The second step is to simulate the circuit for one clock cycle. From the simulation, we should be able to obtain V_{gs}, V_{ds}, I_{ds}, and I_{sub} vs. time. As an example, Figure 5.10 is

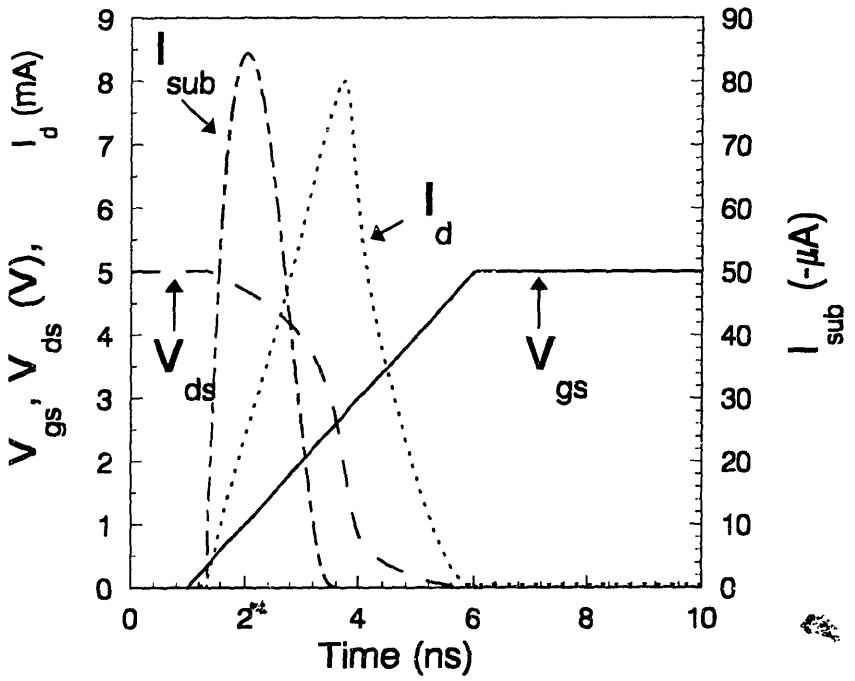

Figure 5.10: Circuit simulation for one cycle

shown for $\alpha = \frac{5V}{5ns}$ and $\frac{W}{C_L} = \frac{40\mu m}{0.5pF}$. Notice that E_{ox} for each simulation time step can be calculated from this plot. (Recall that $E_{ox} = \frac{V_{gs} - V_{ds}}{T_{ox}}$). Then, the degradation model parameters **n**, **m**, and **H** can be calculated for each simulation time step based on the extraction as in Figure 5.4. Recall that the parameters are modeled as a function of E_{ox}.

The next step is to calculate the weighted average degradation rate, $\mathbf{n_{ave}}$. In order to calculate the $\mathbf{n_{ave}}$, we should first plot **n**, **m**, and **AGE** vs. time for one clock cycle as shown below in Figure 5.11. From this plot, we can generate **n** vs. AGE plot by matching **n** with **AGE** at every simulation time step. Figure 5.12 shows such a plot for this example.

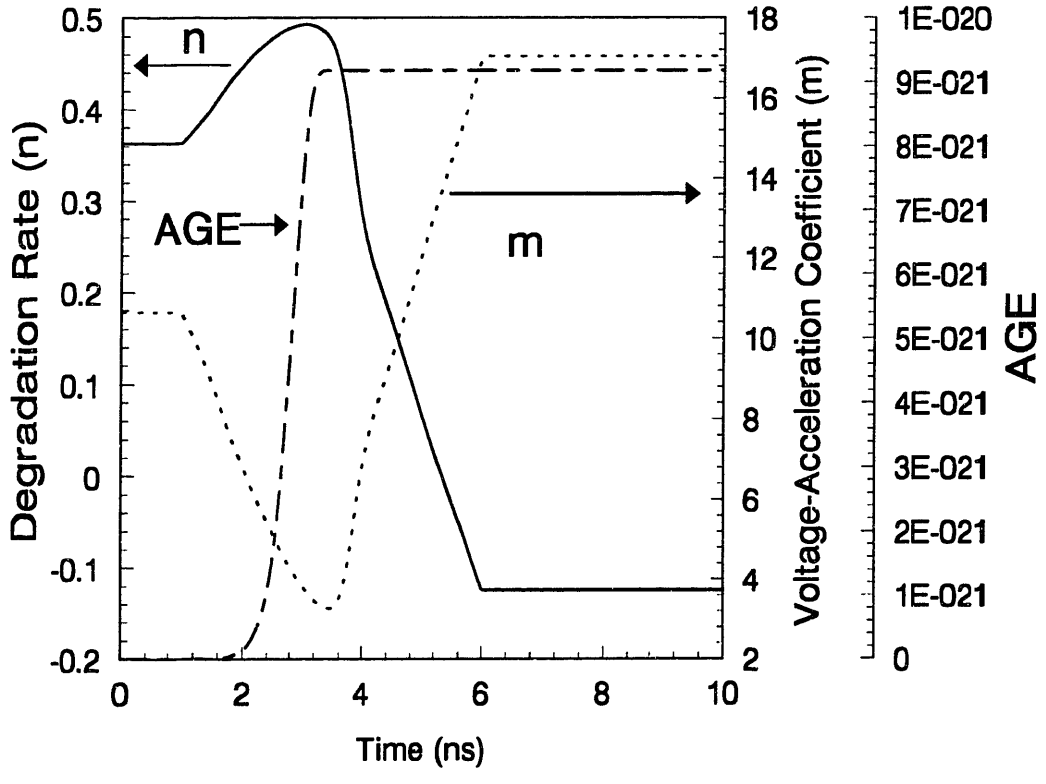

Figure 5.11: n, m, AGE vs. time for one clock cycle.

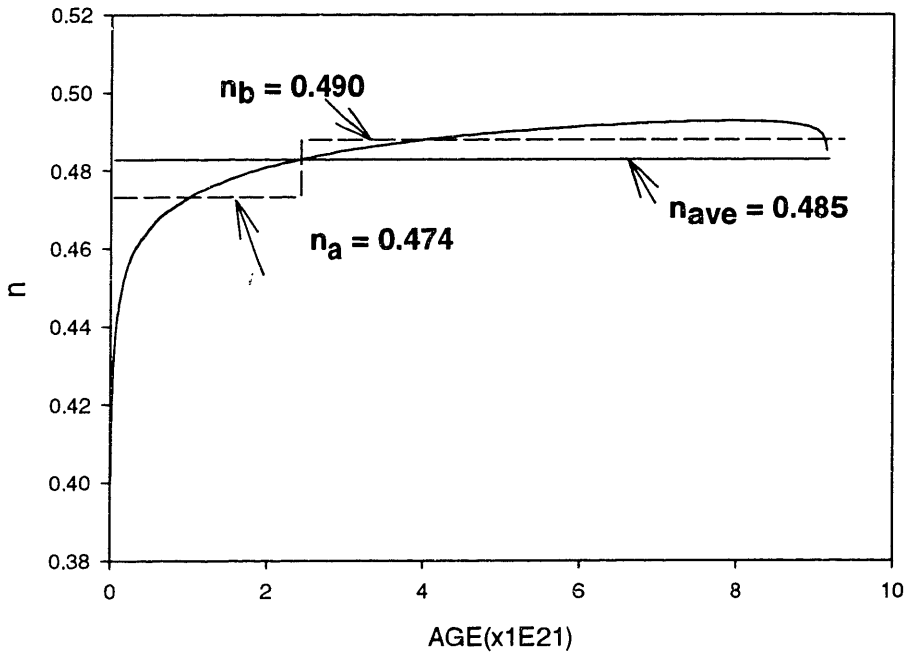

Figure 5.12: **n** vs. **AGE** to calculate $\mathbf{n_{ave}}$, $\mathbf{n_a}$, and $\mathbf{n_b}$.

From Figure 5.12, we can calculate the average degradation rate $\mathbf{n_{ave}}$, weighted by AGE, using the following equation:

$$n_{ave} = \frac{\int_0^{9.2x10^{-21}} n(AGE)dAGE}{9.2x10^{-21}} \tag{5.13}$$

We also calculate two sub-averages $\mathbf{n_a}$ and $\mathbf{n_b}$, by

$$n_a = \frac{\int_0^{2.6x10^{-21}} n(AGE)dAGE}{2.6x10^{-21}} \tag{5.14}$$

$$n_b = \frac{\int_{2.6x10^{-21}}^{9.2x10^{-21}} n(AGE)dAGE}{6.6x10^{-21}} \qquad (5.15)$$

These calculated values are graphically shown in Figure 5.12.

Once these values are calculated, the next step is to identify which one of the asymptotes with the above **n** values is the dominant one at the future time point of interest. In order to identify the dominant asymptote, we lay down these asymptotes on $\frac{\Delta I_d}{I_{d0}}$ vs. number of cycles plot as shown below in Figure 5.13. By choosing the asymptote which gives the most degradation at 10 years, we calculate the AC-degradation. For this particular example, the asymptote with $\mathbf{n_{ave}}$ is the dominant one at 10 years. None of the "twin"

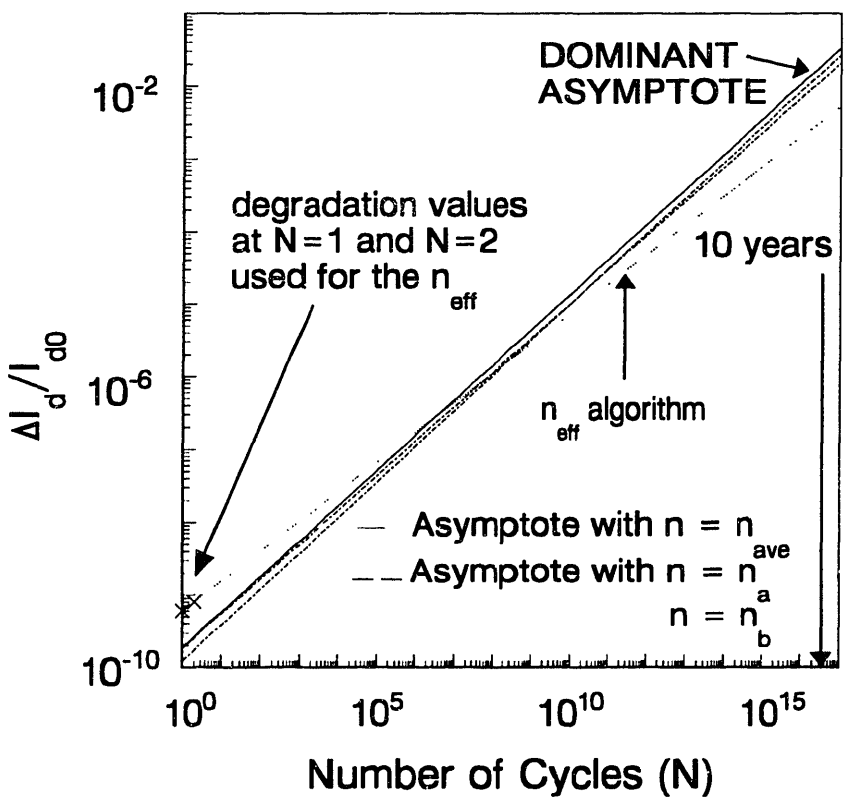

Figure 5.13: Identification of the dominant asymptote at 10 years.

asymptotes was more dominant at 10 years than the $\mathbf{n_{ave}}$ asymptote.

For comparison, the existing $\mathbf{n_{eff}}$ algorithm is also shown in Figure 5.13. As one can see, the $\mathbf{n_{eff}}$ algorithm overestimates the AC-lifetime by about 100 times for this example. This is because the initial dominant DC degradation component (upon which the $\mathbf{n_{eff}}$ extrapolation is based) is no longer dominant at 10 years. Significant amount of error is corrected by this new DDA algorithm.

In order to examine this issue further, the AC degradation was calculated for different CMOS inverters over a wide range of design parameters α and W/C_L, using both the new DDA algorithm and the existing $\mathbf{n_{eff}}$ algorithm. In Figure 5.14, the ratio of the AC-lifetime calculated by the existing $\mathbf{n_{eff}}$ algorithm(τ_{old}) over the AC-lifetime calculated

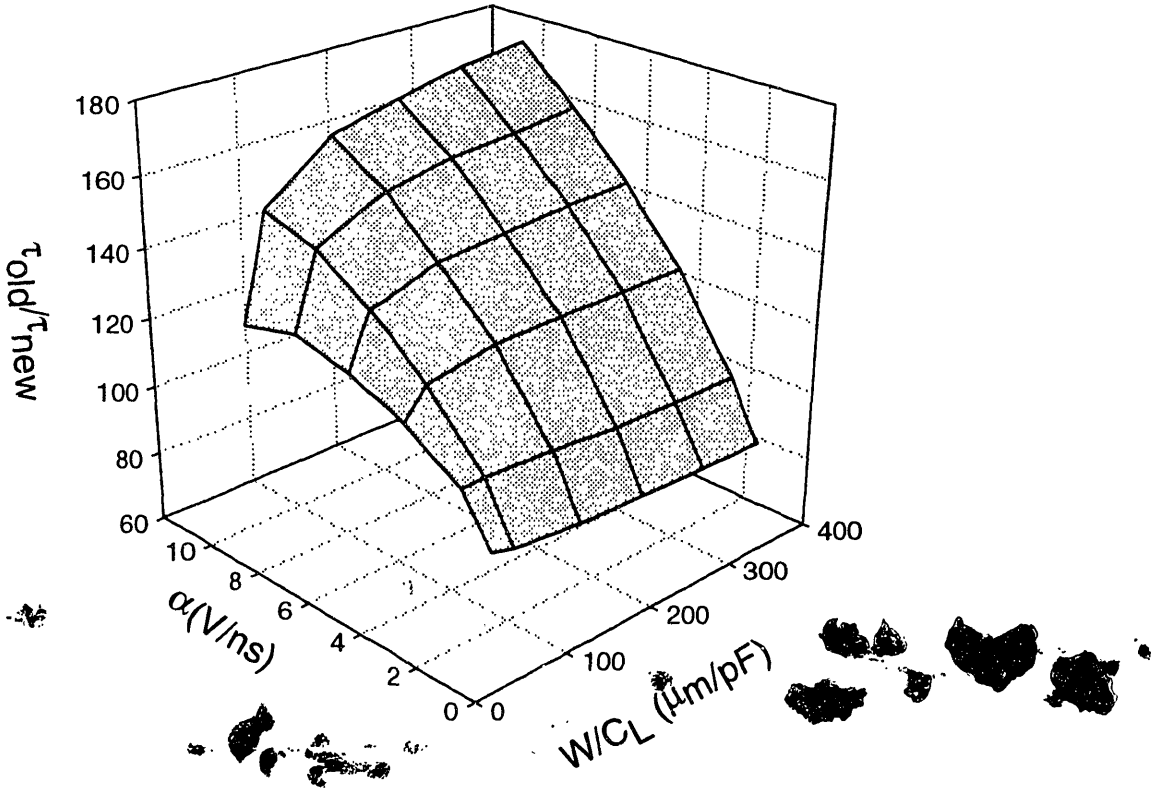

Figure 5.14: The ratio of AC-lifetime, τ_{old}/τ_{new}, between the $\mathbf{n_{eff}}$(old) and the DDA(new) algorithm

by the new DDA algorithm(τ_{new}) is shown. Over a wide range of design space, the existing $\mathbf{n_{eff}}$ algorithm is found to significantly overestimate the AC-lifetime.

This new DDA algorithm is computationally much more efficient than the exact recursive solution. For this inverter example, the DDA algorithm took less than 1 minute on Sun Sparc4 station for AC-lifetime calculation, whereas the exact recursive solution would require more than 10^{16} iterations, which would take many years.

5.5 Summary

Unlike the DC case, the AC waveform presents a complex solution to hot-carrier degradation. The complexity arises from the oxide-field dependent degradation model parameters in time-varying AC waveform. As a result, there exists no closed-form solution for AC hot-carrier degradation. An exact solution only exists in recursive form.

However, it is computationally infeasible to exactly calculate the AC-lifetime based on the recursive solution. It would require approximately 10^{18} iterations. Thus, the current practice, called the $\mathbf{n_{eff}}$ algorithm, calculates the AC-degradations based on the recursive solution for the first few cycles and then extrapolates the AC-degradation all the way out to 10 years, which corresponds to approximately 10^{16} cycles. Through the simple two-stage example of AC waveform, however, the $\mathbf{n_{eff}}$ algorithm is found to overestimate the AC-lifetime by a significant amount.

Based on the observation that a particular DC component dominates the overall AC degradation at a particular time, a new method, called the **DDA** algorithm, is developed. The new insight of this algorithm is to identify the dominant DC degradation component at the future time point of interest. This algorithm is shown to predict the AC-lifetime accurately. The DDA algorithm is also much more computationally efficient than the existing $\mathbf{n_{eff}}$ algorithm.

References

[5.1]. M. Kuo, K. Seki, P. Lee, J. Choi, P. Ko, and C. Hu, "Simulation of MOSFET Lifetime and AC Hot-Electron Stress," IEEE Transactions on Electron Devices, Vol. 35, p1004, 1988.

[5.2]. W. Sun, E. Rosenbaum, and S. Kang, "Fast Timing Simulation for Submicron Hot-Carrier Degradation," International Reliability Physics Symposium Proceeding, p65, 1995.

[5.3]. P. Li, C. Stamoulis, and I. Hajj, "iProbe-d: Hot-Carrier & Oxide Reliability Simulation," International Reliability Physics Symposium Proceeding, p275, 1994.

[5.4]. P. Lee, M. Kuo, K. Seki, P. Ko, and C. Hu, "Circuit Aging Simulator (CAS)," International Electron Devices Meeting, p134, 1988.

[5.5]. S. Aur, D. Hocevar,and P. Yang, "Circuit Hot-Electron Effect Simulation," International Electron Devices Meeting, p498, 1987.

[5.6]. P. Ko, R. Muller, and C. Hu, "A Unified Model for Hot-Electron Currents in MOSFETs," IEEE International Electron Devices Meeting, p600, 1981.

Chapter 6

Hot-Carrier Reliability at Cryogenic Temperature

6.1 Low Temperature Operation of MOSFET

For the past two decades, numerous studies have been published regarding the low temperature operation of MOSFETs [6.1-6.10]. These studies generally claim much improved performance for the MOSFET characteristics at cryogenic temperatures down to 77K. Traditionally, low temperature operation of MOSFET was studied for space application, such as satellite communication systems [6.4-6.6]. However, recently, low temperature operation is considered as an alternative VLSI scaling method as geometry scaling moves closer to its physical limits [6.7-6.10].

Although numerous low-temperature performance advantages have been identified [6.1-6.10], little work has been done in MOSFET reliability at cryogenic temperatures [6.11-6.13]. It is clear, however, that this reliability study is very important in order to maximize fully the performance advantages achievable at the low temperature.

There are two reasons why this dissertation investigated hot-carrier reliability of MOSFETs at cryogenic temperatures. First, without proper understanding of MOSFET reliability at cryogenic temperatures, it is possible to design a device to operate at the maximum performance, but only at the sacrifice of its device lifetime [6.14]. In order to optimize the performance vs. reliability trade-off according to circuit and system designers' need, the hot-carrier lifetime model needs to be reviewed and verified at cryogenic temperatures. A careful calibration of the model is necessary to predict hot-carrier lifetime at low temperatures [6.15].

Second, unlike various other reliability issues, such as electromigration, oxide-breakdown, and device latch-up, etc., hot-carrier reliability has been shown to be exacerbated significantly at low temperatures [6.11-13]. These past studies all found that the device lifetime is shortened almost by an order of a magnitude at 77K. Thus, it is possible that hot-carrier reliability will appear as a roadblock when the temperature scaling is adopted for future ULSI technologies [6.14].

The objective of this chapter is to address the aforementioned issues of hot-carrier reliability at cryogenic temperatures. First, we will discuss the modeling issues of hot-carrier reliability at low temperatures; how the room temperature model can be extended down to 77K, and how it needs to be calibrated for accurate lifetime prediction at cryogenic temperatures. Second, we will quantify how much hot-carrier lifetime is changed under various circumstances. From this discussion, we will attempt to investigate whether hot-carrier reliability actually improves or worsens at cryogenic temperatures. Third, based on the quantification of performance improvement and hot-carrier lifetime, a trade-off plot will be presented as a function of temperature. This plot will allow device and circuit designers to determine the optimal temperature based on the performance vs. reliability trade-off consideration.

6.1.1 Mobility Enhancement

The most notable improvement in the MOS device characteristics at cryogenic temperature is enhanced mobility of electrons. As the thermal energy decreases at cryogenic temperatures, the Si lattice vibrates less, and hence, electron-phonon scattering is reduced. As a result of the reduced phonon scattering, we observe enhanced electron mobility at low

temperature.

Figure 6.1 shows electron mobility extracted as a function of V_{gs}. The I-V data was taken from a large NMOS transistor (W/L=50/5μm) in the linear regime ($V_{ds} = 0.1V$). For this bias condition, μ_{eff} can be extracted from Equation 6.1 as follows [6.16]

$$\frac{\partial I_D}{\partial V_D} = \mu_{eff} \cdot C_{ox} \cdot (V_{gs} - V_T) \quad : \tag{6.1}$$

We observe that μ_{eff} is a strong function of V_{gs} because electrons in the channel get drawn closer to the Si/SiO_2 interface by V_{gs}, and hence, the surface roughness scattering increases. As a result, μ_{eff} decreases as V_{gs} increases. This surface normal E-field dependence of mobility has been observed in many studies [6.16-20]. Based on these studies, Sabnis and Clemens developed a universal mobility model in [6.21], which shows that μ_{eff} can be modeled as a universal function of effective normal E-field.

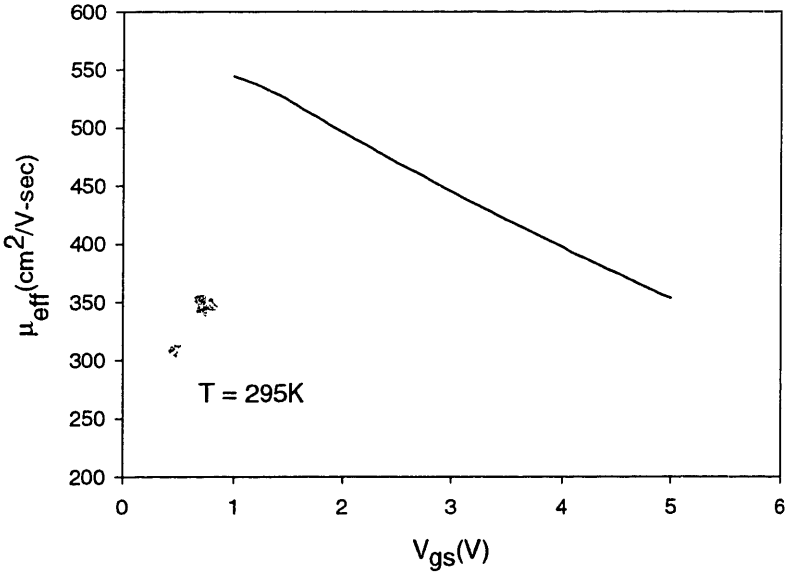

Figure 6.1: Effective mobility extraction at room temperature.

The concept is shown in Figure 6.2. A cross section of MOSFET along the normal direction is shown in Figure 6.2, in which X_i is the length of the inversion layer. The normal E-field experienced by an electron at the surface, x = 0, is due to the bulk charge, Q_B, and the inversion charge, Q_N, whereas the normal E-field experienced by an electron at $x = X_i$, is only due to the bulk charge, Q_B. In general, the E-field experienced by an electron at any point x is given by:

$$E(x) = \frac{Q_B}{\varepsilon_{si}} + \frac{q}{\varepsilon_{si}} \int_{x}^{x_i} n(y)dy \tag{6.2}$$

When the E-field in Equation 6.2 is averaged over the electron distribution in the inversion layer, an effective electric field, given by

$$E_{eff} = \frac{1}{\varepsilon_{si}}\left(Q_B + \frac{Q_N}{2}\right) \tag{6.3}$$

is obtained.

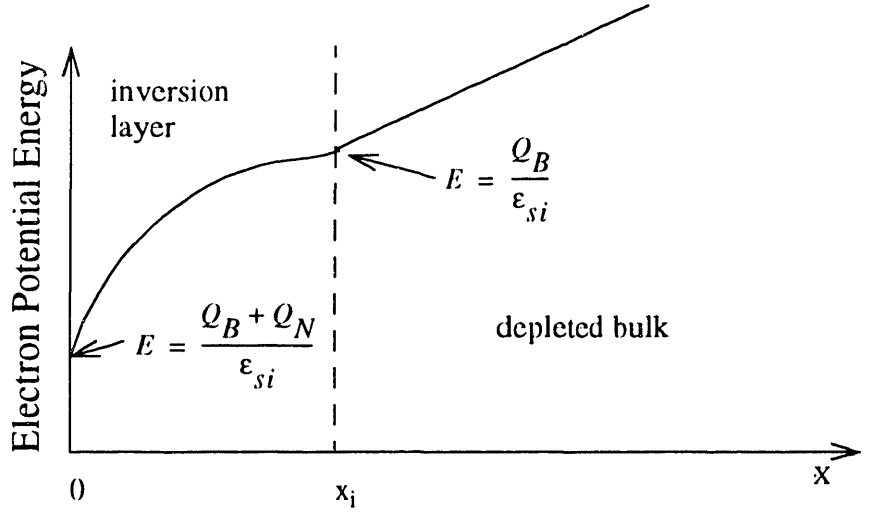

Figure 6.2: Concept of effective normal E-field [6.21].

When μ_{eff} was plotted against E_{eff} over a wide range of channel doping, a universal plot was obtained as shown in Figure 6.3 [6.21]. Other follow-up studies confirm that this relationship is valid over a range of other device parameters, such as T_{ox}, and in high temperature up to 413K [6.22-24].

In order to verify this universal relationship between μ_{eff} and E_{eff} at the cryogenic temperatures and to quantify how much the mobility is enhanced, μ_{eff} was extracted at a range of low temperatures, and plotted against E_{eff} as shown in Figure 6.4. There are several interesting observations to be made in Figure 6.4. First, the μ_{eff} vs. E_{eff} relationship is valid down to 90K. Second, at low E_{eff} (< 0.4 MV/cm), μ_{eff} is enhanced by approximately 300% at 90K compared to room temperature. Third, the μ_{eff} degradation due to high E_{eff} is more severe at low temperature. At T=90K, μ_{eff} decreases approximately 500 cm^2/V-sec from $E_{eff} \sim 0.4$ MV/cm to $E_{eff} \sim 1$ MV/cm, while at room temperature,

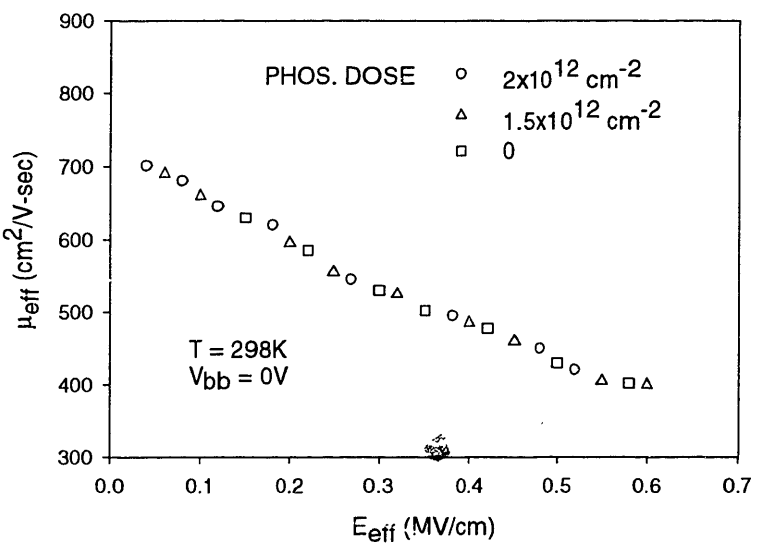

Figure 6.3: Universal mobility dependence on E_{eff} [6.21].

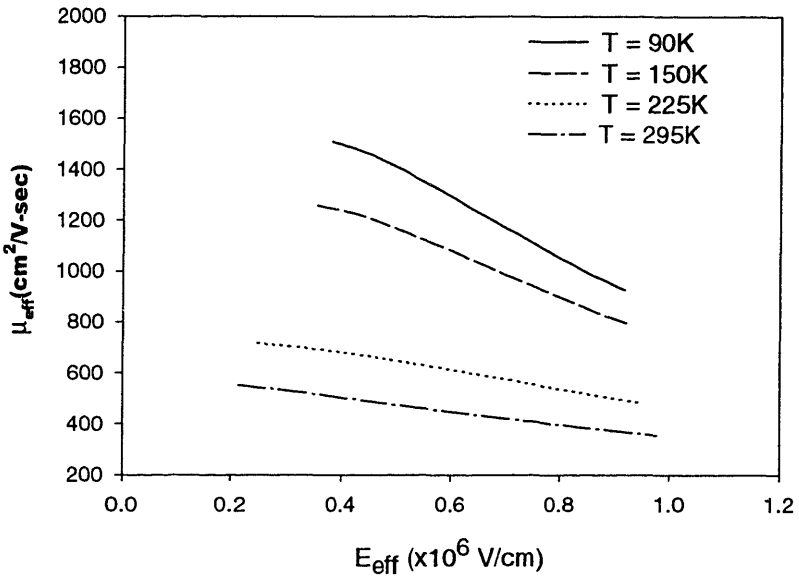

Figure 6.4: Mobility dependence on E_{eff} at cryogenic temperature.

μ_{eff} decreases about 150 cm^2/V-sec for the same range of E_{eff}. This relationship will be further explored in Section 6.3.

The enhanced mobility results in increased drain current at cryogenic temperature. For digital applications, I_{dsat} is an important quantity since it determines the speed of charging and discharging of capacitors, and thus, the dynamic timing behavior of digital circuits. Increased I_{dsat} results in faster switching speed, and thus, increases the throughput [6.25-27]. In order to quantify how much I_{dsat} increases at low temperature due to the enhanced mobility, measurements were taken at a fixed gate drive (V_{gs} - V_t), and the measured I_{dsat} was plotted against temperature in Figure 6.5 for the channel length of 0.3 μm and 1 μm. Comparing values at T=90K and T=295K, I_{dsat} increases approximately 80% for L_{eff} = 0.3 μm and 75% for L_{eff} = 1 μm at T=90K. Similar values have been reported in

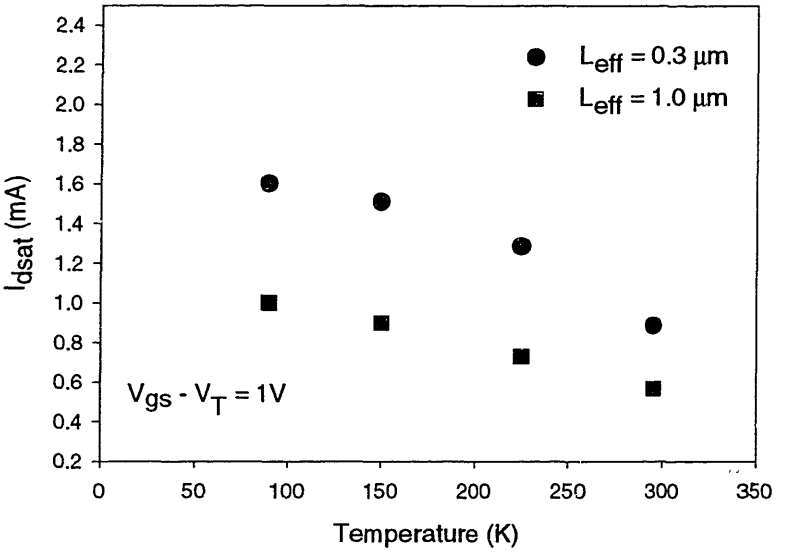

Figure 6.5: Improved performance of I_{dsat} at cryogenic temperature.

previous studies [6.28-30].

While I_{dsat} is an important parameter for digital circuit application, g_{msat} is another important parameter for analog circuit applications. For a common-source amplifier, which is used as a fundamental building block in many analog circuits, the output gain is shown to be proportional to $g_{msat} \cdot r_{out}$ [6.31]. Thus, in numerous studies [6.32-35], g_{msat} is used as the figure of merit for analog circuit applications. In order to quantify the g_{msat} improvement at cryogenic temperature, g_{msat} was extracted at a range of temperatures for the channel length of 0.3 μm and 1 μm, and the result is shown in Figure 6.6. As one can see, g_{msat} increases approximately 50% for both channel lengths at a fixed V_{ds} when it is compared between T=90K and T=295K. Similar values have been reported in previous low temperature device performance studies [6.32-35].

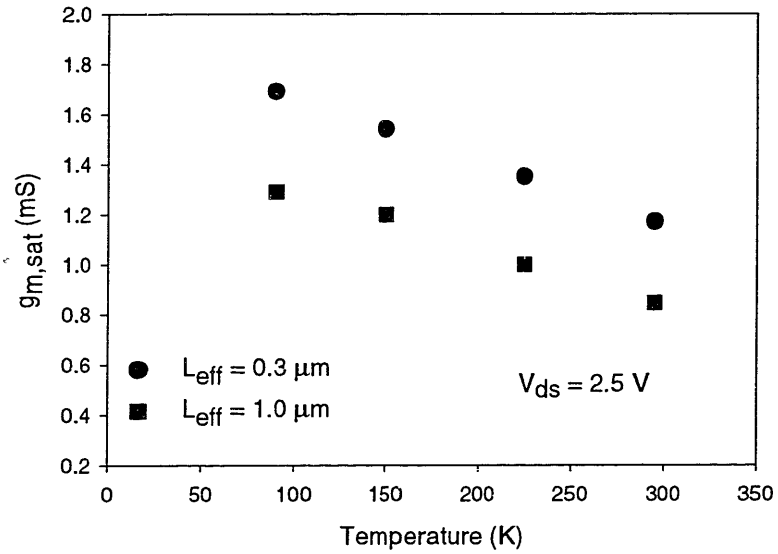

Figure 6.6: Improved performance of g_{msat} at cryogenic temperature.

6.1.2 Subthreshold Slope

Another noticeable advantage of MOSFET low temperature operation is its decreased subthreshold slope. As the subthreshold decreases, the device can switch on and off at a much faster rate, thus increasing the throughput of logic systems [6.36]. A decreased subthreshold slope also results in reduced leakage current at zero-gate voltage. This has been shown to be a critical design parameter, particularly for dynamic switching circuits for low power application [6.37].

The reduced slope directly comes from device physics. The subthreshold slope S is modeled as

$$S = \left(1 + \frac{C_{si}}{C_{ox}}\right) \cdot \frac{KT}{q} \cdot \ln 10 \tag{6.4}$$

in [6.38]. It is clear from Equation 6.4 that S decreases as T is lowered. In order to verify the theory with experimental data, the drain current was measured and plotted against V_{gs} in a log-lin plot in Figure 6.7. For the data shown in Figure 6.7, V_{ds} is biased at 0.1V with zero substrate bias V_{bs}. The channel length of the device is 0.45 μm. As can be seen, the subthreshold slope S linearly decreases from T=295K to T=90K. Similar values have been reported in previous studies [6.5, 6.8-10].

As briefly mentioned above, low temperature operation also reduces the zero-volt gate leakage current, which minimizes static power dissipation. Viewed in another way, short-channel effects are reduced at cryogenic temperatures. In order to verify this, the same I_{ds}-V_{gs} plot in log-lin space is shown in Figure 6.8 for a channel length of 0.3 μm. As can be seen in the Figure, not only does the slope S decrease, but also the leakage current at V_{gs} = 0V is significantly reduced at T = 90K compared to T = 295K. This results in a tremendous amount of power savings for low power application [6.37].

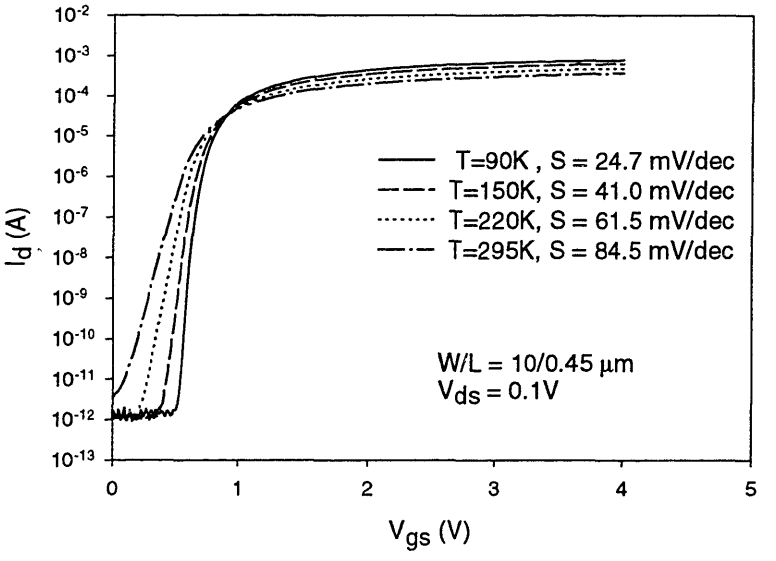

Figure 6.7: Subthreshold slope comparison as a function of temperature.

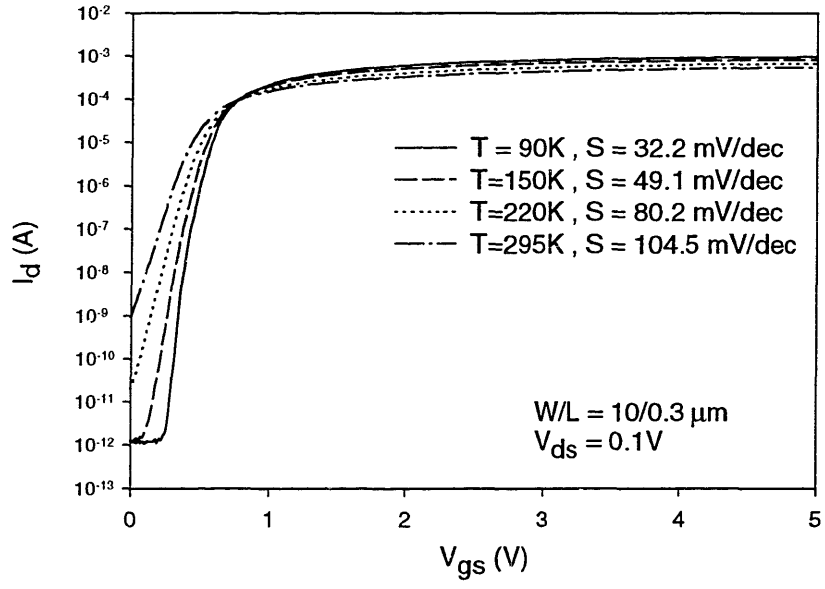

Figure 6.8: Reduced leakage current at low temperature.

6.1.3 Threshold Voltage (V_t) Characterization

For a wide width and long channel device, the threshold voltage V_t is modeled as in [6.38]

$$V_t = V_{FB} + 2\phi_B + \frac{\sqrt{2qN_s\varepsilon_{si}(2\phi_B + V_{bs})}}{C_{ox}} \tag{6.5}$$

where V_{FB} is a flat-band voltage, given by the work function difference between gate and the Si bulk, and ϕ_B is the bulk Fermi potential, given by

$$\phi_B = \left(\frac{KT}{q}\right)\ln\frac{N_s}{n_i} \tag{6.6}$$

Since both V_{FB} and ϕ_B in Equation 6.5 are temperature-dependent, V_t is also temperature dependent. In the case of n^+ polysilicon gate, which is a dominant technology in modern

CMOS microelectronics, V_t is determined primarily by the temperature dependence of the band gap E_g and ϕ_B [6.39]. Overall, the increase of ϕ_B at low temperatures, is due to the reduction of n_i in Equation 6.6, which increases V_t at cryogenic temperatures. In order to verify this increase in V_t at low temperatures, Figure 6.9 shows I_{ds}-V_{gs} measurements in the linear regime at different temperatures. As can be seen in the circled region in the Figure, V_t is approximately 200mV greater at T=90K compared to that at T=295K. This is in accordance with several previous studies [6.4, 6.7-9]. Figure 6.9 also shows increased drain current at the low temperatures in the linear regime due to the enhanced mobility as discussed above.

In Section 6.1.2, it was shown that the leakage current is significantly reduced at the cryogenic temperature. In other words, the short-channel effect is significantly reduced. One such a phenomenon is reduced V_t roll-off as shown in Figure 6.10. The threshold voltage V_t, defined by constant current method at $I_d = \frac{W}{L} \cdot 0.1\mu A$, generally

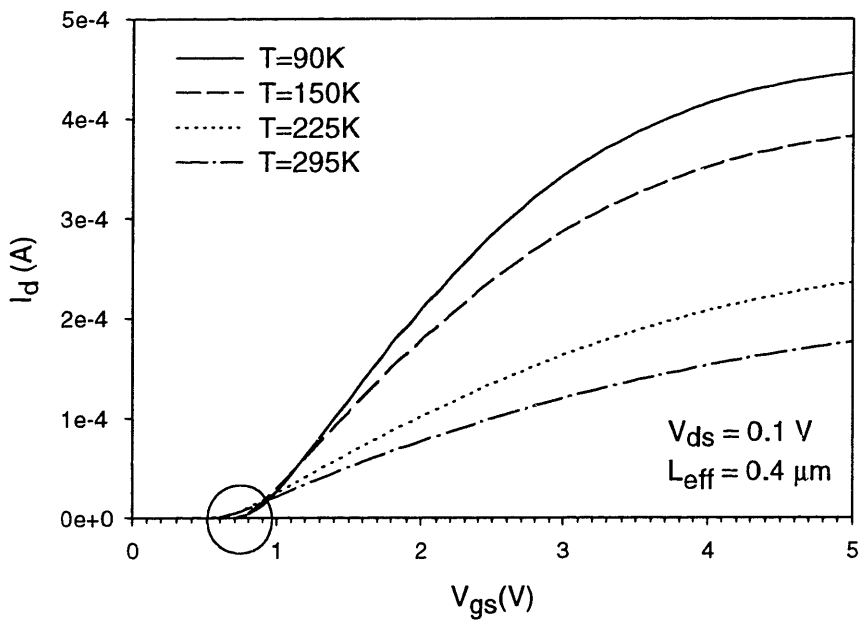

Figure 6.9: Increased V_t at low temperatures.

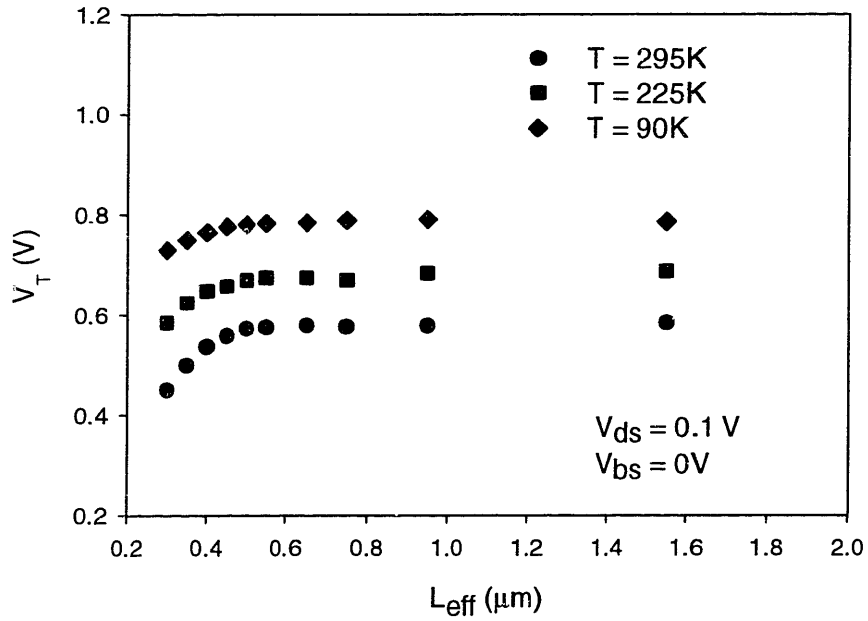

Figure 6.10: Reduced V_t roll-off at low temperatures.

decreases as the channel length decreases. For example, in Figure 6.10 above, V_t decreases approximately 150mV from a long-channel (L_{eff} > 1μm) device to a short-channel (L_{eff} = 0.3μm) device at room temperature. This is evident from comparing Figure 6.7 and Figure 6.8. The increased leakage current for a short-channel device results in reduced V_t, and this phenomenon is called the V_t roll-off.

An interesting observation in Figure 6.10 is that the V_t roll-off is significantly reduced at T=90K. Because of the reduced leakage current for short-channel devices at low temperatures, as can be seen in Figure 6.7 and Figure 6.8, the V_t reduction as L_{eff} decreases is not as pronounced at T=90K as it is at T=295K. This reduced V_t roll-off at the cryogenic temperature yields a significant impact on device and circuit performance modeling. For many circuit simulation device models, such as BSIM3 SPICE model, a large number of computational iterations is necessary to achieve an accurate fit between the

measured V_t data and the simulated V_t data. Reduced V_t roll-off results in reduced number of iterations, thus reducing the simulation time [6.40]. Also, circuit designers often prefer uniform V_t across the channel length [6.41]. Uniform V_t allows them to design circuits without getting into too many details of device modeling.

6.1.4 Summary of Enhanced Performance at Low Temperature

Section 6.1.1 through 6.1.3 discuss the performance improvement of MOS device at the cryogenic temperature down to T=90K. The purpose of these sections was to review some fundamental device characteristics at low temperatures, and verify that these data are typical compared with numerous previous studies on improved device performance at low temperature. These data will be used again in Section 6.5 where the performance vs. reliability trade-off is quantified as a function of temperature.

To summarize, the most notable figure of merit at the cryogenic temperature is enhanced mobility. The enhanced mobility results in increased current drivability and transconductance, which are critical design parameters for digital and analog circuits. Next, the subthreshold slope S was extracted over a range of temperatures. As expected from fundamental device physics, S decreases at low temperature, allowing the device to switch much faster. Reduced leakage current for a short-channel device was also observed. Finally, we discussed that V_t increases at low temperature due to the increased ϕ_B. Also, we noticed that the V_t roll-off was reduced at T=90K, allowing improved efficiency in circuit simulation.

6.2 Hot-Carrier Degradation Behavior at Cryogenic Temperature

Section 6.1 discusses the performance advantages obtained by operating the MOSFET at low temperatures down to 77K. The most distinct feature at cryogenic temperature is the enhanced mobility, which yields higher current drive. The higher mobility at low temperature is a direct result of increased energy of electron due to reduced phonon-scattering.

Unfortunately, the same reason that gives a performance boost at low temperature, exacerbates hot-carrier degradation at low temperatures. Since electrons are more energetic at cryogenic temperatures, they have a higher probability to overcome the energy barrier between Si and SiO_2 interface, and create interface-traps, and electron-traps in SiO_2. Hence, without proper understanding of hot-carrier reliability at cryogenic temperatures, successful low temperature operation of MOSFET cannot be optimized. Thus, the remainder of Chapter 6 is devoted to understanding and modeling hot-carrier reliability at cryogenic temperatures.

Section 6.2 discusses how the room temperature model can be extended to low temperatures, Section 6.3 discusses the dominant degradation mechanisms at a range of temperatures down to 77K, and Section 6.4 compares device lifetimes obtained at various bias conditions at different temperatures. Based on these results, Section 6.5 presents a trade-off analysis of performance vs. reliability for circuit and system designers at different temperatures, and Chapter 6 ends with its summary in Section 6.6.

6.2.1 Extension of Room Temperature Model

Before we begin the study of hot-carrier reliability at cryogenic temperature, the first step is to verify whether the room temperature model can be extended down to liquid nitrogen

temperature 77K. The room temperature model derived in Chapter 2 is written again below for reader's convenience as Equation 6.7:

$$\frac{\Delta I_d}{I_{d0}} = \left(\frac{I_d}{WH} \left(\frac{I_{sub}}{I_d} \right)^m t_{stress} \right)^n . \tag{6.7}$$

Recall that Equation 6.7 is a semi-empirical model where $\frac{\Delta I_d}{I_{d0}} = A \cdot t_{st.ess}{}^n$ is derived from measured data. Other quantities, such as I_{sub} and I_d in constant A, are derived from physical understanding of how hot-carriers are generated.

In order to verify whether t^n dependence is still valid at the low temperature, accelerated stress voltages are applied, and $\frac{\Delta I_d}{I_{d0}}$ is plotted against the stress time in log-log space as shown in Figure 6.11. As can be seen, the degradation still follows t^n dependence

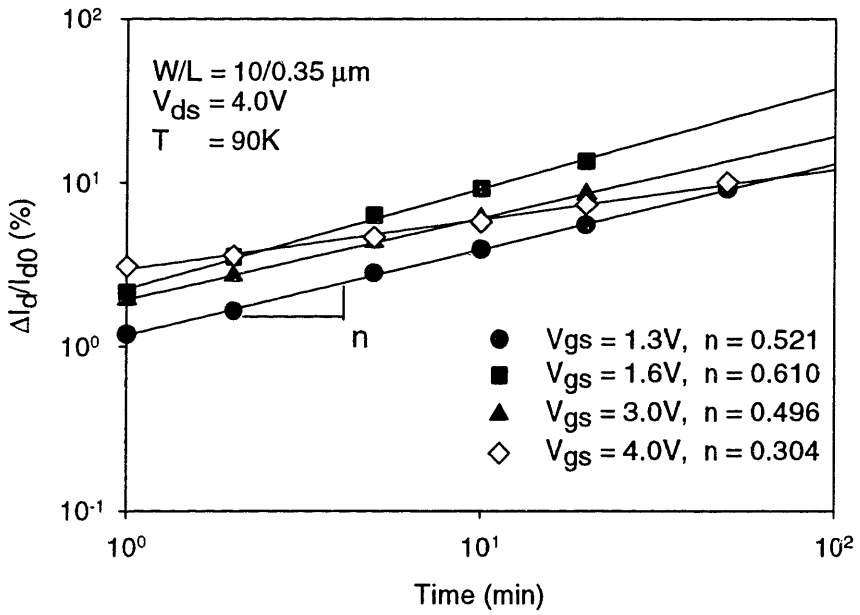

Figure 6.11: Degradation time dependence at cryogenic temperature.

at cryogenic temperature. It is also interesting to observe that the degradation rate **n** has similar values to the ones extracted at room temperature. In order to further verify the bias dependence of degradation and its rate at T=90K, Figure 6.12 plots $\Delta I_d/I_{d0}$, **n**, I_{sub}, and I_{sub}/I_d against V_{gs} for a fixed bias of V_{ds}. One can notice that although the absolute value of $\Delta I_d/I_{d0}$ is different at T=90K, the general trend of degradation is quite similar to that at room temperature. For example, the degradation seems to follow the shape of I_{sub} at T=90K. Also, the peak **n** condition is different from the peak $\Delta I_d/I_{d0}$ condition as it was of room temperature. As V_{gs} increases, the normalized quantity I_{sub}/I_d decreases. Figure 6.11 and 6.12 suggest that the degradation behavior at liquid nitrogen temperature is quite similar to that at room temperature, and hence make it feasible to use the same parameter extraction procedure to predict the device lifetime at cryogenic temperature.

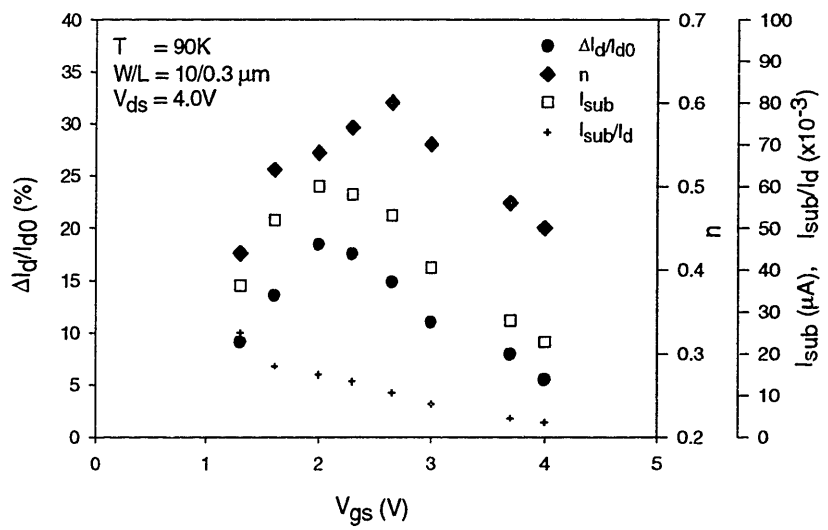

Figure 6.12: Bias dependence of degradation and its parameters.

6.2.2 Bias Dependence of the Model Parameters

As discussed in Chapter 4 and 5, an accurate extraction of the degradation model parameters is critical for the accurate MOSFET lifetime prediction. Thus, this section will focus on the parameter extraction procedure, from which we not only extract the model parameters to calculate the device lifetime at cryogenic temperatures, but also gain much better insight on physical understanding of hot-carrier degradation at low temperatures.

Figure 6.12 shows that the degradation rate **n** is bias-dependent. Recall also that in Chapter 3, the degradation rate **n** was modeled as follows:

$$n = \frac{I_{sub}}{I_d} \cdot \exp(-B \cdot E_{ox}) \tag{6.8}$$

where E_{ox} is given by $\left|\frac{V_{gs} - V_{ds}}{T_{ox}}\right|$. In order to check whether this model is still applicable

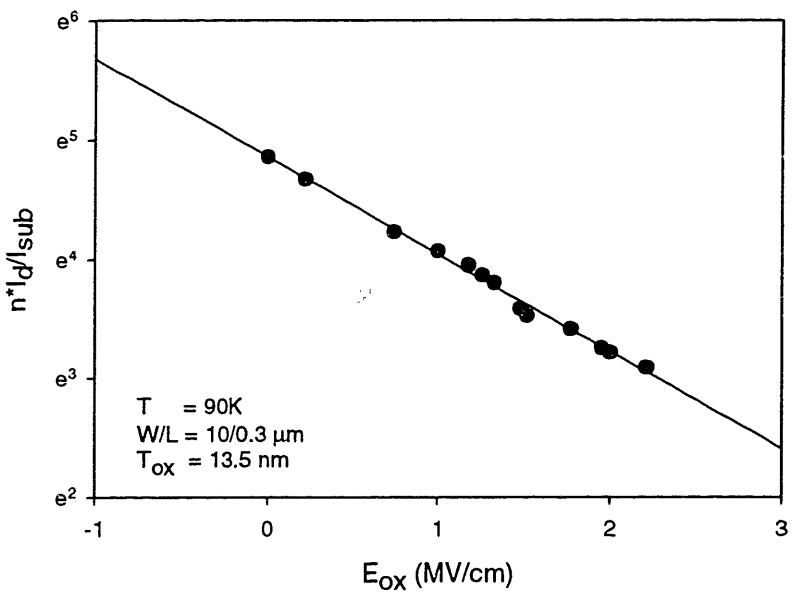

Figure 6.13: Model of **n** at low temperature.

at low temperature, $\frac{n \cdot I_d}{I_{sub}}$ is plotted against E_{ox} in log-lin space in Figure 6.13. As was the case at room temperature, Equation 6.8 still holds valid at T = 90K. In other words, the degradation rate **n** still has the same dependence on the vertical oxide-field, E_{ox}, and the peak lateral E-field, E_p. More specifically, we observe that **n** is proportional to the normalized quantity I_{sub}/I_d (which represents E_p), and inversely proportional to E_{ox} with exponential dependence.

Having verified the validity of **n** model at T = 90K, we are now ready to make the lifetime correlation plot from which we can extract the parameters **m** and **H**. Recall that they were modeled as a function of E_{ox} at room temperature. Figure 6.14 shows the corresponding lifetime correlation plot at T = 90K. As it can be seen, **m** and **H** still have a strong monotonic dependence on E_{ox}.

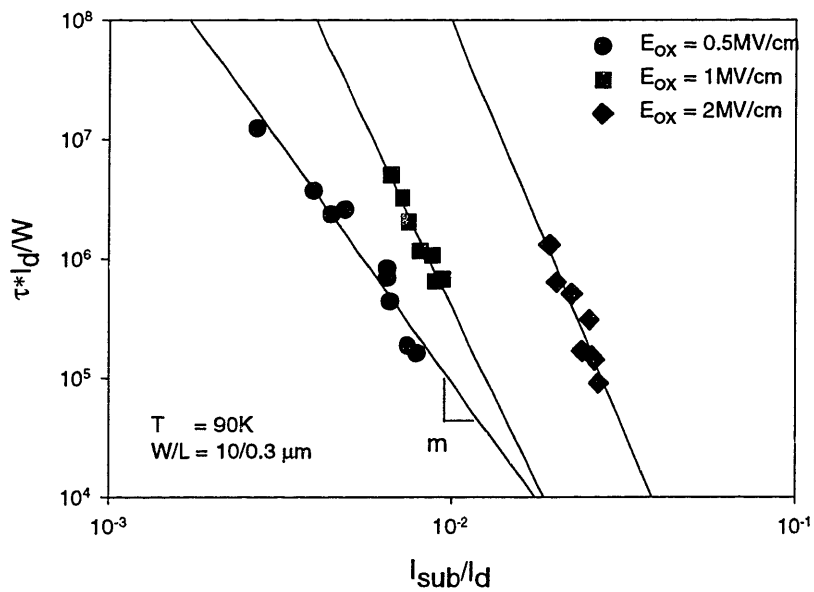

Figure 6.14: m and **H**'s dependence on E_{ox} at cryogenic temperature.

In deriving the degradation model, **m** was defined as the ratio, $\frac{\phi_{it}}{\phi_i}$, where ϕ_{it} is the critical energy required to create the interface-trap. Section 3.3 physically explained why ϕ_{it} has a dependence on E_{ox}. In order to compare the critical energy, ϕ_{it}, at different temperatures, the lifetime correlation plot is shown in Figure 6.15, which compares the slope **m** between room temperature and T = 90K for a fixed E_{ox} = 1MV/cm. As can be seen, **m** is greater at low temperature. Similar values are extracted at other E_{ox}'s at T = 90K and T = 300K, and they are shown in Figure 6.16.

Based on Figure 6.15 and 6.16, one might jump to conclusion that ϕ_{it} is greater at low temperatures. In other words, hot-electrons must possess greater amount of energy to create interface-traps at cryogenic temperatures. However, there is no clear physical explanation for this phenomenon. The barrier height between Si and SiO_2 does not change

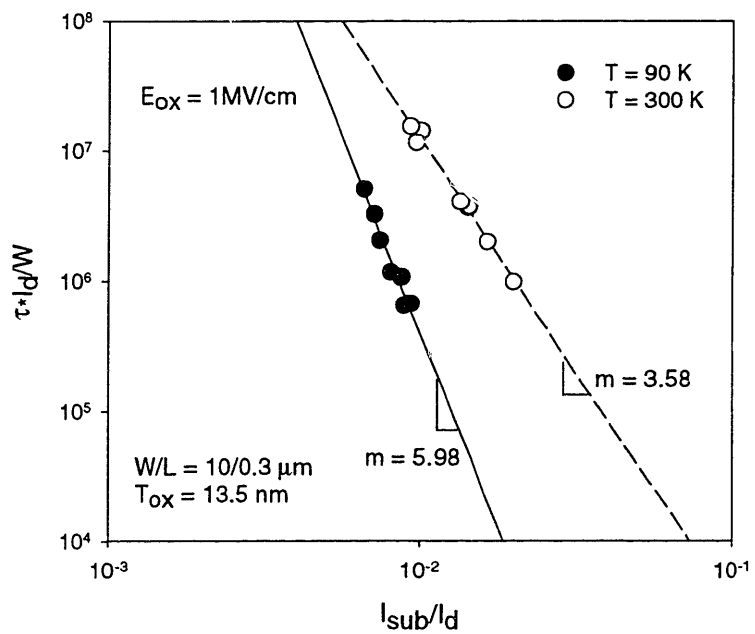

Figure 6.15: Comparison of **m** at different temperatures.

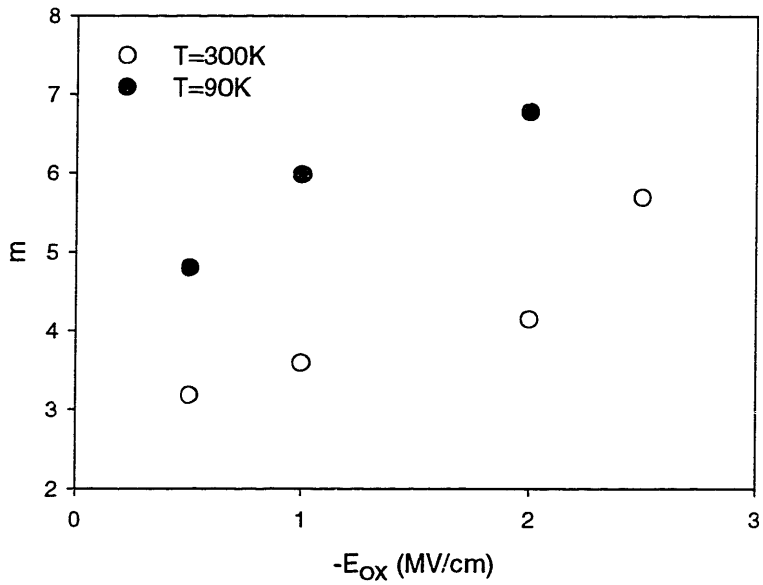

Figure 6.16: Comparison of **m** for a range of E_{ox}.

significantly as the temperature decreases from T = 300K to T = 90K [6.38-39]. There may be some other mechanisms that are responsible for the observed phenomenon, but, at this point, it is not clear whether ϕ_{it} indeed increases as the temperature decreases. A physical discussion will be carried out in Section 6.4 to further investigate this issue.

Based on Section 6.2, which shows the bias-dependence of **n**, and the oxide-field dependence of **m** and **H** at T = 90K, it seems feasible that the room temperature model can be extended down to T = 90K with appropriately extracted parameters. After having observed $\Delta I_d/I_{d0}$'s dependence on I_{sub}, and how the rate **n** varies with E_p and E_{ox}, it is further assured that the room temperature model derived in Chapter 2 is still applicable down to liquid nitrogen temperature with different parameter values.

6.3 Degradation Mechanism

Once we have verified the applicability of the room temperature model down to T = 90K, we are now in a position to study the dominant hot-carrier degradation mechanism at low temperature. The methodology of the degradation mechanism study is to first observe the degraded I-V characteristics of the MOSFET, and then to form a hypothesis based on the degraded I-V data. Charge-pump current measurement will further illuminate the degradation mechanisms, either proving or disproving the hypothesis.

Figure 6.17 and 6.18 show the degraded I_d vs. V_d characteristics at T = 300K and T = 90K, respectively. For both plots, MOSFETs with identical dimensions W/L=10/0.3μm were stressed at identical stress conditions, V_{ds} = 4.2V, and V_{gs} = 2.4V for 30 minutes. V_{ds} was chosen such that a sufficient amount of degradation occurs at room temperature, and yet such that, not an excessive amount of degradation occurs at cryogenic temperatures. V_{gs} was determined near the maximum I_{sub} condition for the chosen V_{ds}.

It can be seen that at T = 90K, the drain current degradation is more severe for both forward and inverted (with drain and source switched) measurement. Also, at both temperatures, the drain current degradation is more severe in the linear regime than in the saturation regime. The current reduction may be attributed to the presence of stress-induced negative charge *either at Si/SiO_2 interface or in the SiO_2* [6.42]. This negative charge increases V_t and decreases μ_{eff}, which together reduces the drain current. Now, the task is to find out whether it is the *fixed negative charge in SiO_2* or *the negative charge built at the Si/SiO_2 interface* that reduces the current.

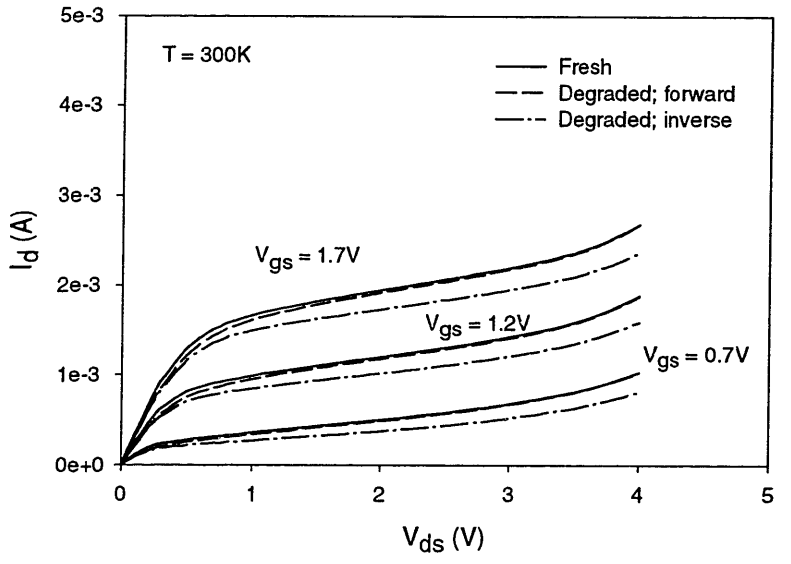

Figure 6.17: Degraded I-V characteristics after stressing at T = 300K.

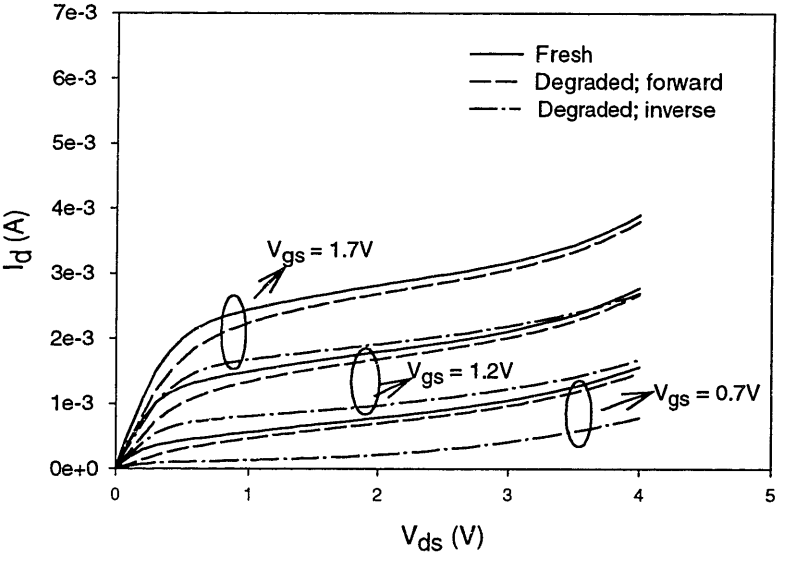

Figure 6.18: Degraded I-V characteristics after stressing at T = 90K.

In order to make further observation on the degraded I-V characteristics, Figure 6.19 shows both fresh and degraded I_d vs. V_{gs} in the linear regime where V_{ds} = 0.1V. To observe the drain current both in the subthreshold and inversion regime, I_d is plotted against V_{gs} in log-lin and lin-lin space. One can clearly observe from Figure 6.19 that I_d in the subthreshold regime only slightly degrades, whereas it degrades quite noticeably in the inversion regime. This strongly suggests that the negative charge induced by hot-carrier is not a fixed charge in SiO_2, but rather a Si/SiO_2 interface charge. If it is a fixed charge in SiO_2, a parallel shift of I-V curve should occur. In order to further explore this speculation, the MOSFET was stressed with strong V_{bs} and V_{gs} with both V_{ds} and V_{gs} grounded. For this bias condition, it is known that the *substrate hot-electrons* rather than the *channel hot-electrons* are injected into SiO_2 over the energy barrier, and negative fixed charge becomes trapped in SiO_2. The experimental setup is shown in Figure 6.20.

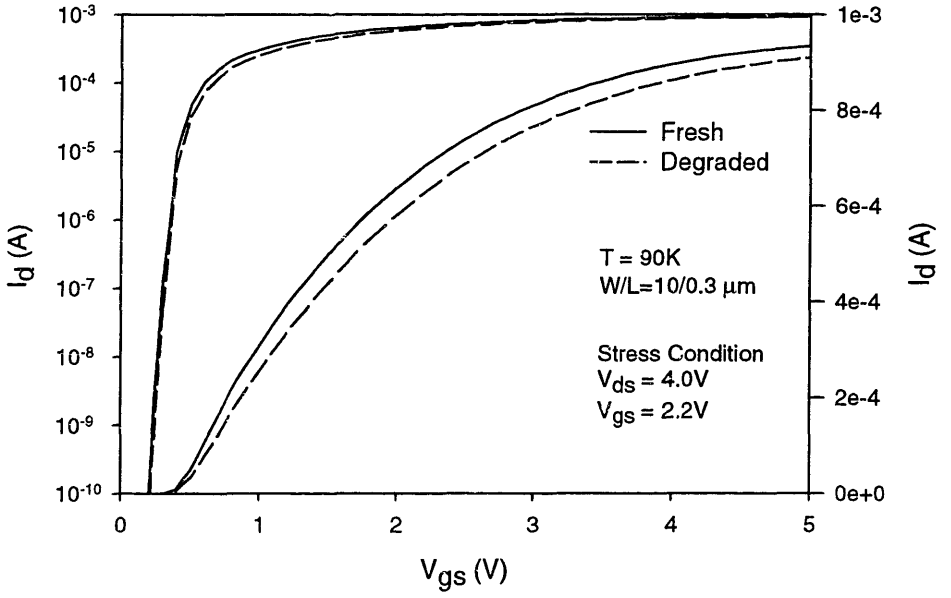

Figure 6.19: Fresh and Degraded I_d vs V_{gs} at T = 90K.

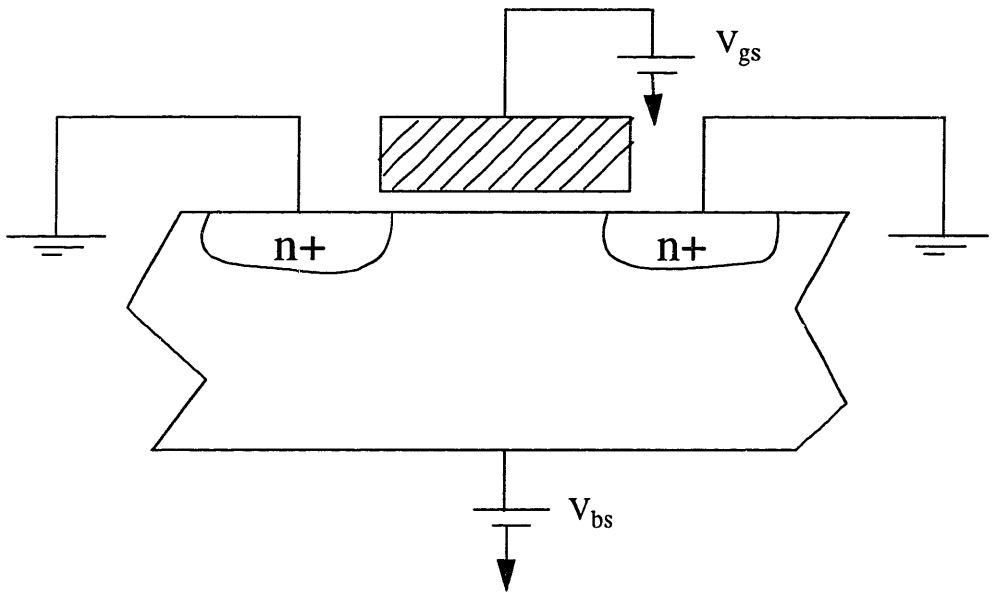

Figure 6.20: Experimental setup for the substrate hot-electron injection.

After the MOSFET was stressed with V_{gs} =3V and V_{bs} = -10V for 30 minutes, I-V degradation was observed as shown in Figure 6.21. For this substrate hot-electron injection, it has been reported that fixed electron-trap is a dominant degradation mechanism [6.43]. As one can see, for the fixed electron-trap in SiO_2, a parallel shift of I-V characteristics occurs. This is logically sound because the fixed charge should not vary as we change the bias conditions. However, for the channel hot-electron injection shown in Figure 6.19, the drain current degrades significantly in the inversion regime, but does not degrade as noticeably in the subthreshold regime. This can be attributed to the interface-trap because the interface-trap is charged negatively only when it is filled with electrons. In the subthreshold regime, the interface-trap is empty, whereas in the inversion regime, it gets filled with electrons, hence showing the degraded I-V characteristics as in Figure 6.19. Based on this argument, interface-trapping seems to be the dominant degradation

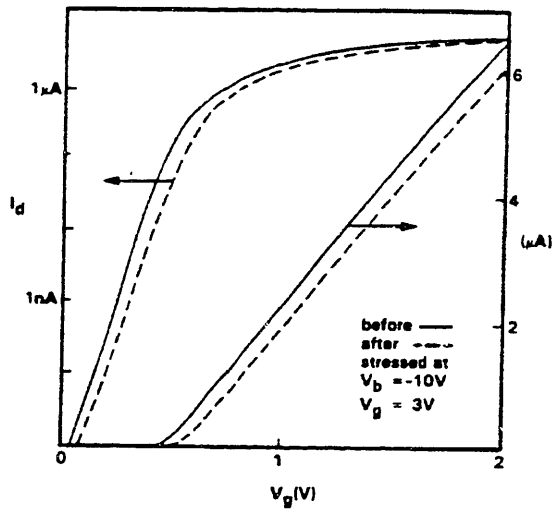

Figure 6.21: Degraded I-V characteristics after substrate hot-electron injection. T_{ox} = 20nm, W=L=100μm, [6.43].

mechanism at cryogenic temperatures.

The presence of interface-traps at the drain end is further supported in Figure 6.22, in which I_d is plotted against V_{ds} in log-lin space for two different V_{gs}'s - one below the threshold, and the other above the threshold. Even though the interface-states are created during the stress, they are empty when measured below the threshold, hence showing little degradation for both forward and inverse mode. However, when the device is measured above the threshold, we observe a markedly different characteristics. For the forward mode, the drain current decreases significantly in the linear regime and gradually approaches its original value in the saturation regime. This is because the interface-traps that are created during the stress are mostly filled with electrons in the linear regime but mostly empty in the saturation regime. For the inverse mode where the drain and source are interchanged, however, we observe that the drain current saturates at a value well below its original value. This is because the created interface-traps at the source end

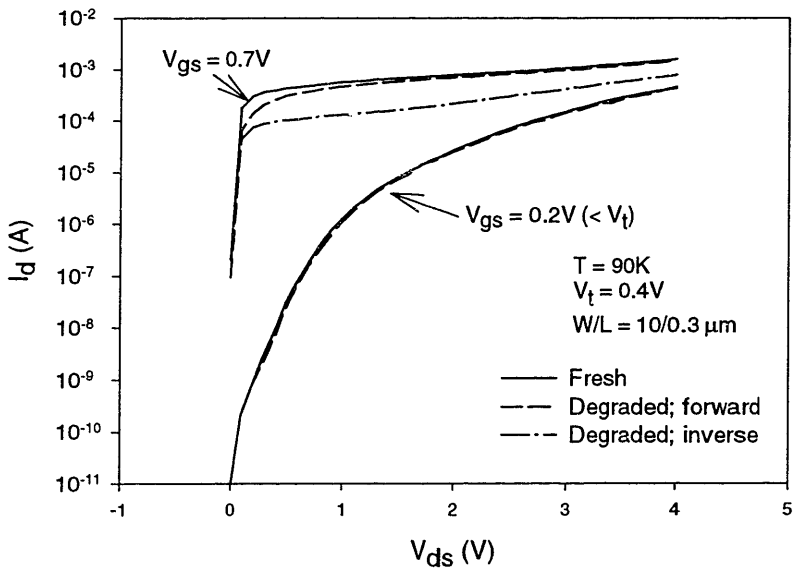

Figure 6.22: Degraded I_d vs V_{ds} for above and below threshold.

become mostly filled regardless of V_{ds} bias conditions as long as the device is in inversion.

In order to verify the hypothesis of interface-trap generation, charge-pump current was measured after the devices were stressed at cryogenic temperature. Recall from Chapter 1 that as interface-traps are generated, the charge-pump current increases because the number of captured electrons at the interface increases. Figures 6.23 to 6.25 show how the I_{cp} curve changes after the device was stressed at low temperature. In Figure 6.23, I_{cp} is plotted against V_{base} after the device was stressed near the maximum I_{sub} condition. The upper curve shows I_{cp} after the device was stressed, the lower curve shows the original I_{cp} results, and the middle curve shows how much I_{cp} has increased by subtracting the lower curve from the upper curve.

As one can observe, I_{cp} has increased by a substantial amount, indicating a signifi-

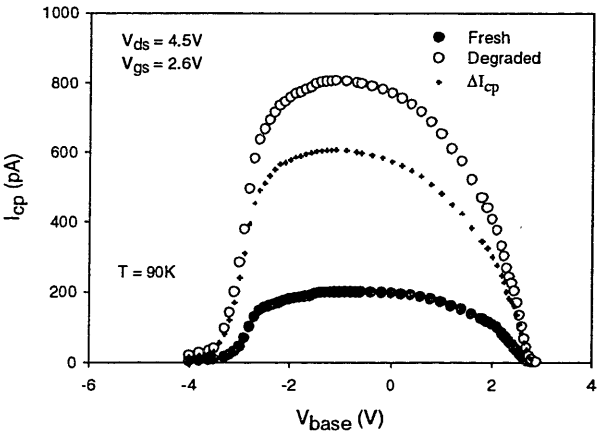

Figure 6.23: I_{cp} vs V_{base} after stressed at peak I_{sub}.

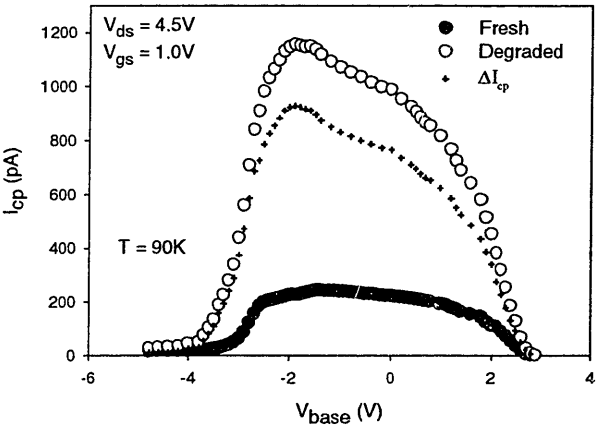

Figure 6.24: I_{cp} vs V_{base} after stressed at low V_{gs}.

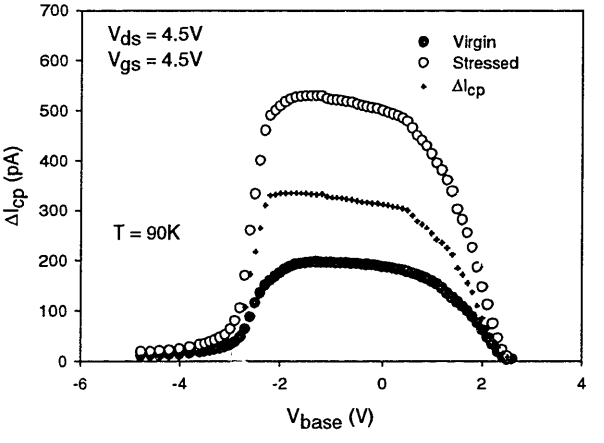

Figure 6.25: I_{cp} vs V_{base} after stressed at high V_{gs}.

cant increase in the number of interface-traps after the device was stressed. The fact that the left edge of the curve in Figure 6.23 shifted to the left may suggest a possible positive-charge (hole) traps in the oxide, but it is evident that ΔI_{cp}, which is directly proportional to the increase of interface-traps, clearly dominates the overall degradation mechanism.

Figures 6.24 and 6.25 show similar I_{cp} characteristics after the devices were stressed at different bias conditions - low V_{gs} ($V_{gs} \sim V_t$) and high V_{gs} ($V_{gs} = V_{ds}$). For both conditions, I_{cp} significantly increases even though the edge of the curve shifts a bit more at low V_{gs} condition. Figure 6.23 through 6.25 verify our hypothesis of interface-trap generation at cryogenic temperature, which was initially based on the interpretation of I-V characteristics of MOSFETs.

In order to compare the degraded I_{cp} behavior of low temperature to that of room temperature, ΔI_{cp} is plotted against stress time in Figures 6.26 and 6.27. In Figure 6.26,

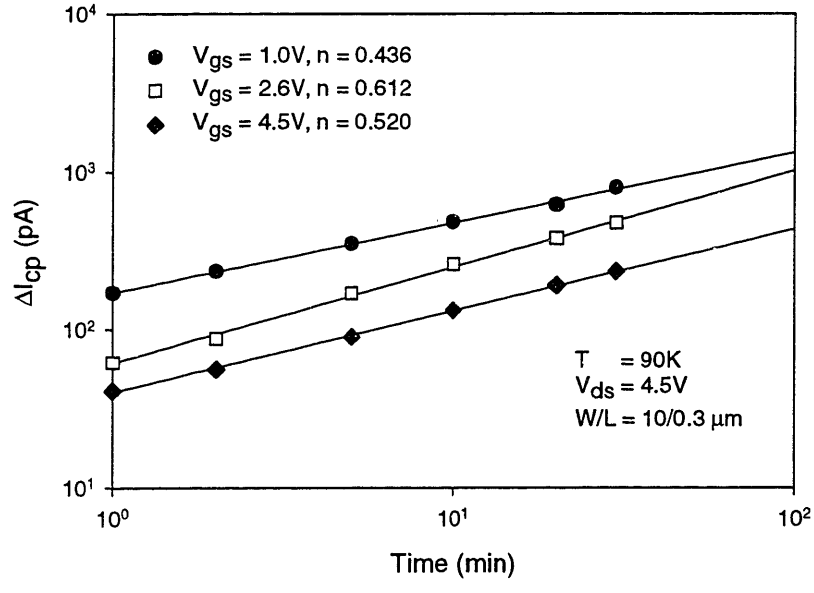

Figure 6.26: Bias-dependence of degradation rate at T = 90K.

ΔI_{cp} at V_{base} = 0V in Figures 6.23 through 6.25 is plotted against stress time. The degradation rate **n** extracted from ΔI_{cp} clearly has a bias-dependence as Equation 6.8 predicts. Figure 6.27 shows a similar plot after the room temperature stressing. Even though the absolute value of degradation rate is different, the trend is the same - i.e., **n** increases, reaches its maximum, and decreases as V_{gs} increases. This is consistent with Equation 6.8 since this model was shown to be valid at both room and cryogenic temperatures with appropriately extracted I_{sub}/I_d and B.

Figures 6.26 and 6.27 together verify that the interface-trap is the dominant degradation mechanism at both room and cryogenic temperatures, and the rate of interface-trap generation can be modeled with number of hot-electrons *generated* (I_{sub}/I_d) and *injected* into Si/SiO_2 interface (E_{ox}). From this, it can be concluded that we should focus the device design on improving the quality of Si/SiO_2 interface in order to minimize the hot-electron

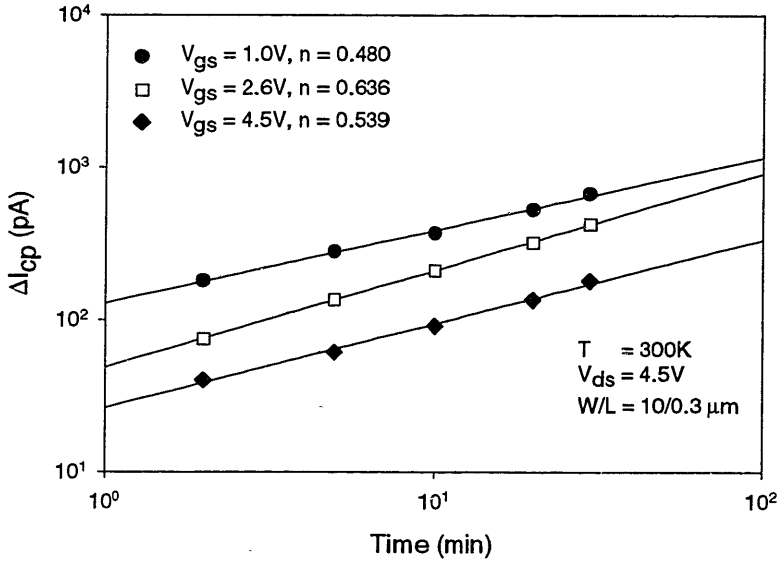

Figure 6.27: Bias-dependence of degradation rate at T = 300K.

degradation at both room and cryogenic temperatures.

6.4 Hot-Electron Device Lifetime at Cryogenic Temperatures

We have verified the validity of the room temperature model extension down to liquid nitrogen temperature, and have also found that the interface-trap is the dominant degradation mechanism. Now, we are ready to proceed to evaluate hot-carrier reliability at cryogenic temperature - i.e., whether the reliability improves or worsens at low temperatures. In order to evaluate the reliability, we will study how the device lifetime changes as we move the device operation temperature from T=300K to T=90K.

For digital circuit designers, an important device performance metric is the drain current, I_d. The higher the current, the faster the device can switch. For analog circuit designers, an important parameter is the transconductance g_m. Typically, gain of an amplifier is directly proportional to g_m. Thus, these two parameters' degradation, $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, will be used as lifetime criteria to calculate the device lifetime over a wide range of temperatures.

6.4.1 Lifetime Based on Device Degradation Parameters

First, we use the lifetime criterion of $\Delta I_d/I_{d0} = 10\%$ to calculate the device lifetime. This definition has been typically used in numerous studies that investigated the impact of hot-electron degradation on digital circuits [6.44-46]. Figure 6.28 shows a lifetime correlation plot with this lifetime definition over a wide range of temperatures. The devices used were 10/0.3 μm. Since it is known that the device lifetime is a function of E_{ox} from Chapter 3, E_{ox} was fixed at 1MV/cm for all the stressings performed for Figure 6.28. As one can.

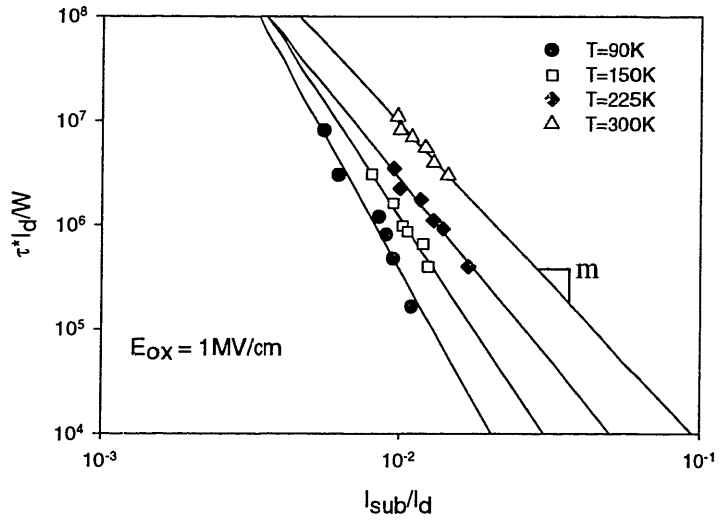

Figure 6.28: Device lifetime comparison between different temperatures with the lifetime criterion of $\Delta I_d/I_{d0} = 10\%$.

clearly see, the device lifetime is shortened at low temperatures for the same I_{sub}/I_d. Recall that I_{sub}/I_d represents a normalized peak lateral E-field in the channel, E_p. In other words, for the same strength of E_p, the device degradation in terms of $\Delta I_d/I_{d0}$ is more severe as the temperature is lowered from T=300K to T=90K. Figure 6.28 is also consistent with Figure 6.15, which showed that the extracted slope **m** is greater at lower temperature.

In order to further verify the lifetime correlation between temperatures, we use a different lifetime definition, $\Delta g_m/g_{m0} = 20\%$, and its lifetime correlation plot is shown in Figure 6.29. As discussed above, $\Delta g_m/g_{m0}$ is an important parameter for analog circuit designers, and hence, important to monitor this quantity for analog circuits, such as differential amplifiers and current mirrors, etc.

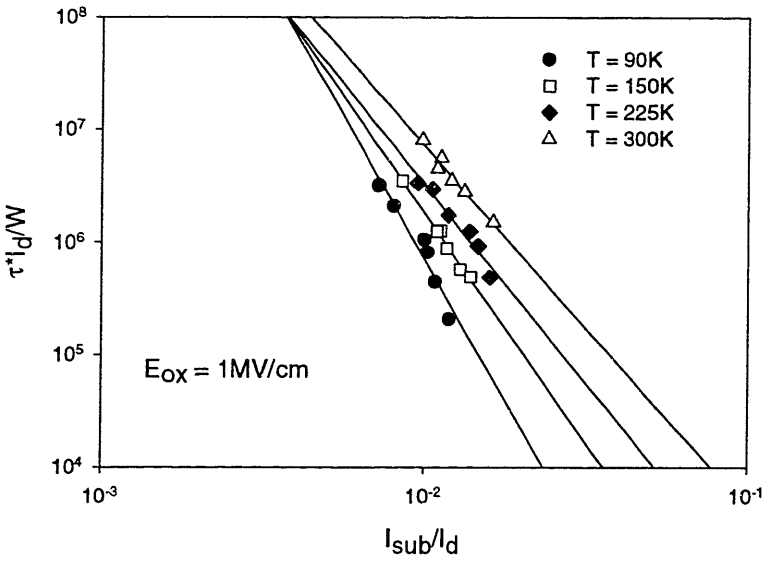

Figure 6.29: Device lifetime comparison between different temperatures with the lifetime criterion of $\Delta g_m/g_{m0} = 20\%$.

As can be clearly seen, a very similar lifetime correlation is shown in Figure 6.29 when $\Delta g_m/g_{m0}$ is used in place of $\Delta I_d/I_{d0}$. Not only are the lifetimes reduced at lower temperatures for both cases, but the values of the extracted slope **m** are very similar to each other. In order to show the comparison of **m** ($= \frac{\phi_{it}}{\phi_i}$) between different temperatures based on different lifetime criteria, Figure 6.30 shows the extracted ϕ_{it} values as a function of temperature for a fixed E_{ox} = 1MV/cm. As can be seen, ϕ_{it} monotonically increases as the temperature decreases for both cases of extraction - one based on $\Delta I_d/I_{d0}$, and the other $\Delta g_m/g_{m0}$.

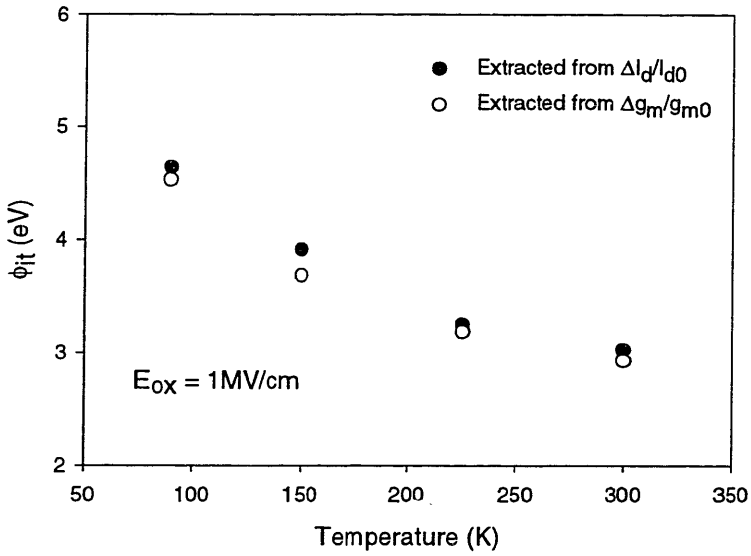

Figure 6.30: Extracted ϕ_{it}s over a range of temperatures.

6.4.2 Lifetime Based on Physical Damage

Recall that $m = \frac{\phi_{it}}{\phi_i}$ where ϕ_{it} is the critical energy for interface-trap generation and ϕ_i is the impact ionization energy. Since ϕ_i is a constant independent of temperature [6.38], it must be the case that ϕ_{it} increases as the temperature is lowered. However, there is no physical theory that would support this phenomenon. In Chapter 3, we identified three components that constitute ϕ_{it}: ϕ_{bond}, ϕ_B, and ϕ_{ic}, where ϕ_{bond} is the energy to break the Si-H bond at the Si/SiO_2 interface, ϕ_B is the barrier height between Si and SiO_2, and ϕ_{ic} is the amount of band bending. None of these quantities should be dependent upon temperature [6.15, 6.38, 6.47].

The question, then, is why the extracted slope **m** is steeper at lower temperature as shown in Figures 6.28 and 6.29. Before we answer this question, we first need to experimentally verify that ϕ_{it} is indeed constant over the temperature range as some physical the-

ories suggest [6.15, 6.38, 6.47]. In order to verify this, we have stressed the MOSFETs over the range of temperatures, and have monitored ΔI_{cp}, the increase of charge pumping current. Recall from Section 6.3 that we have discovered that the interface-trap is the dominant degradation mechanism over the temperature range of interest, and thus, ΔI_{cp} is a very useful quantity that can give a direct measurement of physical damage.

Figure 6.31 shows the lifetime correlation plot, using ΔI_{cp} = 500pA as the lifetime criterion. This definition was chosen because ΔI_{cp}= 500pA corresponds to approximately $\Delta I_d/I_{d0}$ = 10% at room temperature. The device dimension of W/L = 10/0.3 μm was used for all the stressings. We find a remarkable result in Figure 6.31. When ΔI_{cp} is used as the degradation monitor, the device lifetime is not a function of temperature any more. Furthermore, the extracted slope **m**, and thus, ϕ_{it} is constant from T=300K to T=90K. Assum-

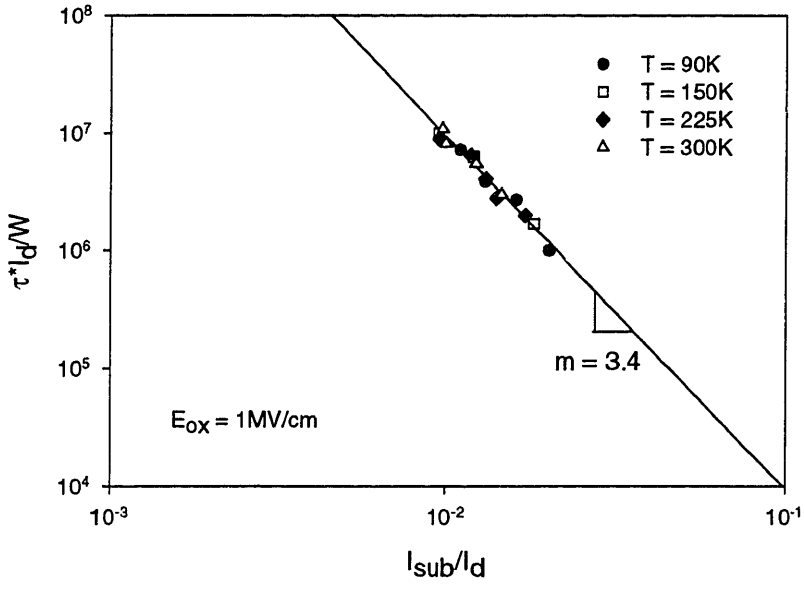

Figure 6.31: ϕ_{it} extraction using ΔI_{cp} as degradation monitor.

ing ϕ_i=1.1eV [6.38], we extract ϕ_{it}=3.1eV. This is the value that we extracted at T=300K, using device degradation monitors, $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, as shown in Figure 6.30.

This result suggests that the amount of physical damage, i.e. the increase in the number of interface-traps, is constant for the same normalized E-field, I_{sub}/I_d, from room temperature down to liquid nitrogen temperature. Also, ϕ_{it}, the critical energy for generating interface-traps, is constant over the temperature range of interest. Its value is surprisingly consistent with the early Berkeley model, which made numerous simplifications [6.48].

From the results of Figures 6.28 through 6.31, we conclude that the amount of physical hot-carrier damage is the same for the same E_p, but its *manifestation* in terms of device performance metrics, such as $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, are different at different temperatures. More specifically, for the same amount of interface-traps created, its impact on device performance metrics, such as $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, is greater at cryogenic temperature. Thus, extracting ϕ_{it} based on the device degradation parameters, $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, can mislead to conclusion as shown in Figure 6.30 that ϕ_{it} increases at lower temperatures, whereas it actually stays constant as shown in Figure 6.31. This result is also consistent with the conclusion that can be drawn by putting previous works together [6.15, 6.38, 6.47].

6.4.3 Physical Explanation of Lifetime Behavior

Having observed that the amount of physical hot-carrier damage is the same at the same E_p, but its *manifestation* is greater at lower temperature, we attempt to develop a physical explanation for this observed phenomenon in this section. When interface-traps are cre-

ated and filled with electrons after the hot-carrier stressing, negative charges at Si/SiO_2 interface are built up. There are two direct mechanisms that reduce the drain current, I_d, as a result of built-up negative charges at the interface. First, the threshold voltage increases, and thus, I_d decreases at the same gate and drain bias. Second, the mobility of channel electrons decreases due to increased scattering. Since I_d is directly proportional to channel electron mobility, it decreases when mobility decreases.

We observe from Figures 6.28 through 6.31 that I_d decreases more even for the same amount of physical damage at lower temperatures, and thus, we closely inspect both V_t and μ at low temperatures. Compared to room temperature, V_t seems to increase a bit more at lower temperatures, but its differential increment is not sufficient to explain the significantly higher $\Delta I_d/I_{d0}$ at lower temperature. Thus, we speculate that mobility degradation is more severe at lower temperatures. In order to verify this hypothesis, we have extracted effective mobility, μ_{eff}, from fresh and degraded devices at both T=300K and T=90K, and the result is shown in Figure 6.32. As can be observed, mobility degradation is significantly more severe at cryogenic temperature, hence, proving our hypothesis.

According to a previous study, mobility degradation due to interface-traps can be modeled as follows [6.49]:

$$\mu_{eff} = \frac{\mu_0}{1 + K \cdot N_{it}} \tag{6.9}$$

where μ_0 is the inversion-layer mobility with no degradation, **K** is the interface-trap scattering coefficient, and N_{it} is the number of interface-traps generated per unit area, cm^2. The parameter **K** models how much the mobility degrades due to increased scattering caused by interface-traps.

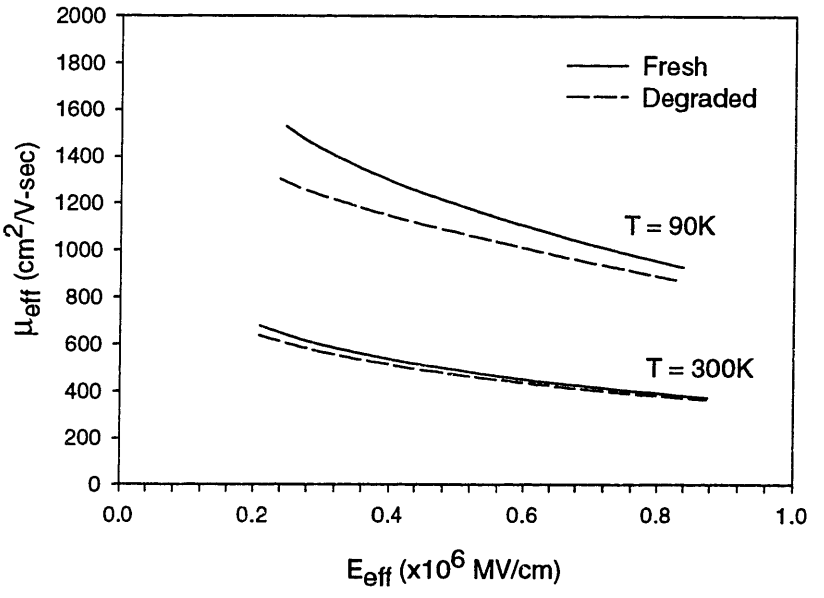

Figure 6.32: Mobility degradation after hot-carrier stressing at room and cryogenic temperature.

In order to extract the parameter **K**, we utilize the relationship between $\Delta I_d/I_{d0}$ and ΔI_{cp}. Under the interface-trap generation, the drain current reduction can be modeled as follows [6.50]:

$$\frac{\Delta I_D}{I_{D0}} = K \cdot \left(\frac{l}{L_{eff}}\right) \cdot N_{it} \tag{6.10}$$

where l is the length of degraded region near the drain end in the channel.

The increase in I_{cp} due to interface-traps can be modeled as in [6.51]:

$$\Delta I_{cp} = q \cdot f \cdot W_{eff} \cdot l \cdot N_{it} \tag{6.11}$$

where f is the signal frequency, and l is the length of the degraded region in the channel. By solving Equation 6.11 for N_{it} and substituting it into Equation 6.10, we derive the fol-

lowing relationship between $\Delta I_d/I_{d0}$ and ΔI_{cp}:

$$L_{eff} \cdot \frac{\Delta I_D}{I_{D0}} = K \cdot \left(\frac{1}{qfW_{eff}}\right) \cdot \Delta I_{cp} \tag{6.12}$$

Equation 6.12 states that we can extract the parameter **K** from the slope by plotting $L_{eff} \cdot \frac{\Delta I_D}{I_{D0}}$ against ΔI_{cp}. Recall, however, that μ_{eff} is a function of E_{eff} as shown in Figure 6.32. Hence, **K** is also a function of E_{eff}. This result is also consistent with the work in [6.50]. Figure 6.33 shows the correlation plot between $L_{eff} \cdot \frac{\Delta I_D}{I_{D0}}$ and ΔI_{cp} for a fixed E_{eff} = 0.5 MV/cm. We observe that the slope, and thus, **K** increases as the temperature is lowered. In other words, the sensitivity of mobility degradation to interface-traps becomes greater as the temperature is lowered.

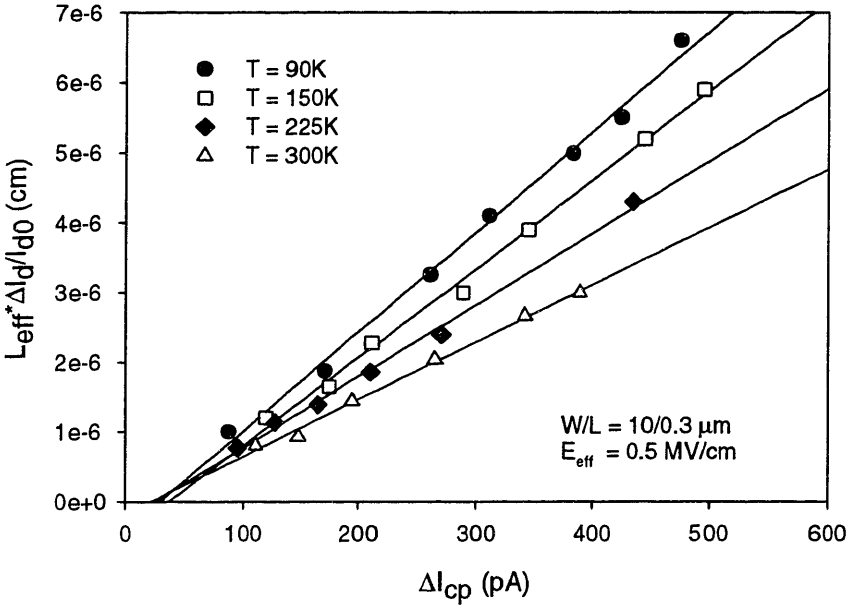

Figure 6.33: Extraction of the parameter **K** over the range of temperatures.

Table 6.1 shows the list of **K** values as a function of temperature for the fixed E_{eff} = 0.5 MV/cm. The increased **K** value accounts for larger mobility degradation at cryogenic temperature as Equation 6.9 implies. To recapitulate, even though the amount of physical hot-carrier damage (interface-trap) is the same for the same strength of E_{peak}, its *manifestation* in terms of device degradation, such as $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, is greater at low temperatures because the mobility degradation is greater at such temperatures. Greater mobility degradation is a direct consequence of greater sensitivity parameter **K**.

In order to show that this is a universal phenomenon over a wide range of measurement conditions, **K** was extracted, using the $\Delta I_d/I_{d0}$ and ΔI_{cp} relationship, at multiple E_{eff} values over the temperature range of interest, and the result is shown in Figure 6.34. As can be seen, **K** is a strong function of E_{eff} and temperature.

Temperature (K)	**K** (1×10^{-12} cm^2)
90	7.77
150	6.50
225	4.07
300	2.91

Table 6.1: Extraction of the **K** parameter as a function of temperature. E_{eff} = 0.5MV/cm.

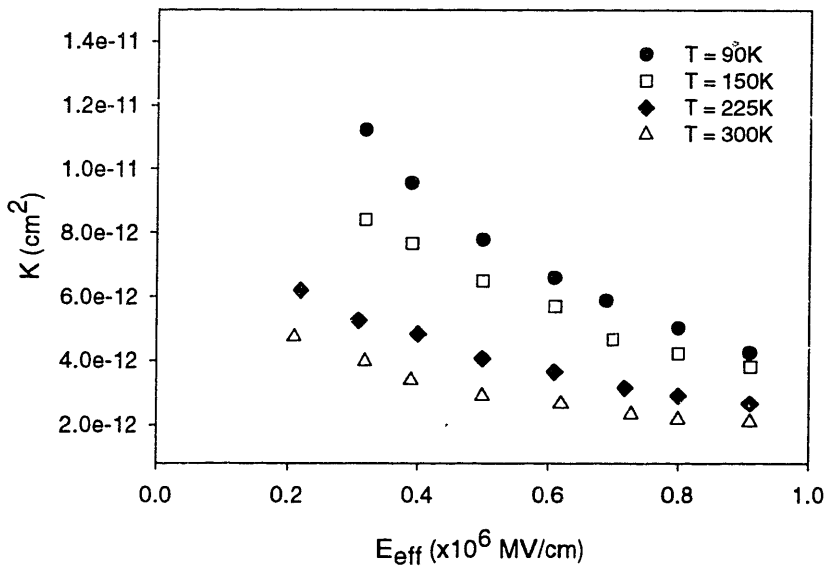

Figure 6.34: Extraction of **K** over a wide range of measurement conditions, E_{eff}, for different temperatures.

Having observed that the device degradation is greater due to greater mobility degradation at cryogenic temperature, the fundamental question to ask is why the mobility degradation is enhanced at low temperature even for the same amount of interface-traps generated. More specifically, the question can be stated as why the sensitivity parameter **K** increases at low temperature.

Recall that **K** is an interface-trap scattering coefficient. In other words, **K** relates how much mobility degrades due to increased scattering as the interface-traps are generated. There are several scattering mechanisms, such as phonon, coulomb, and surface roughness scattering. Phonon scattering is not dependent upon interface-traps generated [6.51]. However, coulomb and surface roughness scattering are directly influenced by the generated interface-traps [6.52]. This implies that the increased **K** is a result of increased coulomb and surface roughness scattering at cryogenic temperature. This implication is

consistent with previous studies that have reported greater influence of these scattering mechanisms on overall inversion layer electron mobility at low temperatures [6.53-6.54].

At cryogenic temperature, the charged scattering center is believed to be located at Si/SiO_2 interface. Since the coulomb scattering is primarily due to the interface charge rather than impurity ions in silicon, the increased interface-traps reduce the inversion layer mobility due to the increased coulomb scattering. Surface roughness scattering is a result of fluctuating potential caused by the imperfect interface. This mechanism is strongly process dependent, and hence, is independent of temperature [6.52]. As a result, surface roughness scattering becomes essential at low temperature where phonon scattering is significantly reduced.

To summarize why mobility degradation is enhanced for the same amount of interface-traps generated as the temperature is lowered, we believe that other scattering mechanisms, such as coulomb and surface roughness scattering, exert greater influence on overall inversion layer electron mobility at low temperatures. It is well known that at room temperature, μ is mainly determined by phonon scattering, and hence, the role of coulomb and surface roughness scattering can be disregarded for the purpose of extracting μ_{eff} for device modeling [6.53]. However, as the temperature is lowered, phonon scattering significantly decreases, and hence, other mechanisms play essential roles in determining the overall μ_{eff} [6.52, 6.54]. It is for this reason, we believe, that the mobility degradation, and thus, the drain current degradation becomes greater at low temperature for the same amount of physical damage - interface-traps - created.

6.5 Performance vs. Reliability Trade-off

We have investigated the major hot-carrier reliability issues at low temperature - applicability of room temperature model to cryogenic temperature, the dominant degradation mechanism, and the greater manifestation of the physical damage in terms of device performance parameters, such as $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$. Having understood that, we are now in a position to develop a trade-off plot between performance and reliability as the MOSFETs get cooled from room temperature down to liquid nitrogen temperature. In other words, we want to develop a quantitative metric of how much performance gets improved vs. how much reliability gets sacrificed as the temperature is reduced.

There are several standards to gauge device performance and reliability, and hence, the first step is to define a figure of merit that would serve as direct monitor for performance and reliability. The approach taken in this chapter is to define performance and reliability metrics from a circuit designer's perspective. As explained in Section 6.1, I_d is an important quantity for digital circuit designer, whereas g_m is an important one for analog circuit designer. Thus, I_d and g_m will be used as performance metrics. For reliability, device lifetime is of primary concern for designers. An analog designer is interested in how long the device would last, meeting the g_m specification. Similarly, a digital designer needs to know how long the device would last, meeting the I_d specification. Both designers need such information as the device gets cooled down to low temperature for a fixed supply voltage. As an example, for cryogenic circuit application, a digital designer would want to know how much I_d increases and how long the MOSFET would last with less than 10% I_d degradation, for a channel length of 0.3 μm operating at 3V, as the temperature is

lowered from room temperature to liquid nitrogen temperature. A similar definition can be applied to analog design space as well. There is an interesting point that should be noticed in this definition. Recall from Section 6.4 that the *amount of physical damage is the same for the same* E_p even as the device's temperature is lowered to 90K. However, from a circuit designer's perspective, the physical damage ΔN_{it} and E_p are not interesting figures of merit. What is important to the circuit designer is *how much the device performance parameters,* I_d *and* g_m*, change at a fixed power supply voltage as the temperature is lowered.* Hence, we adopt to use the *manifestation of the physical hot-carrier damage*, i.e., $\Delta I_d/I_{d0}$ and Δg_m, as the lifetime criteria rather than ΔN_{it}.

For performance metrics, I_d and g_m were characterized in Section 6.1 as a function of temperature and channel length. Figures 6.5 and 6.6 show such a characterization. For reliability metrics, device lifetime based on $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$ can be used. Measuring exact device lifetime at typical operational voltage is almost an impossible task. For example, it would take a long period of time to stress the 0.3 μm device at 3V and observe 10% drain current reduction. Hence, the devices have traditionally been stressed at much higher voltages, the device lifetimes have been extracted at such higher voltages, and the device lifetime at operational voltage has been extrapolated from the lifetimes extracted at such higher voltages. Although this extrapolation approach seems to be feasible and might be the only viable methodology for any generic reliability testing, there is an inherent risk of introducing error by such huge extrapolation. This is especially true in hot-carrier testing because the device lifetime is an exponential function of stressing voltage.

Thus, for the performance vs. reliability trade-off analysis in this chapter, we have decided to stress the device at the lowest possible stress voltage. In other words, the stress

voltage was chosen such that a noticeable amount of degradation was visible after one day of stressing, and the lifetime is extracted from the stressed data without extrapolating down to the operational voltage. Hence, the lifetime obtained from the stressed data may not accurately represent the lifetime at operational voltage, but, for the trade-off analysis, it should give a good understanding of how performance vs. reliability trade-off is guided.

Since we have characterized the figures of merit, I_d and g_m, for the performance analysis in Section 6.1, we now extract the device lifetime at the lowest possible stress voltage. Figure 6.35 shows an example. An NMOS device with 10/0.3 μm dimension was stressed for $3x10^4$ seconds. V_{ds} chosen was 3.5V to ensure that a clearly noticeable amount of degradation was visible at room temperature. The same V_{ds} was applied to the NMOS devices at different temperatures as shown in Figure 6.35, and the degradation, $\Delta I_d/I_{d0}$, was measured. For each temperature, V_{gs} was chosen to be a peak I_{sub} condition.

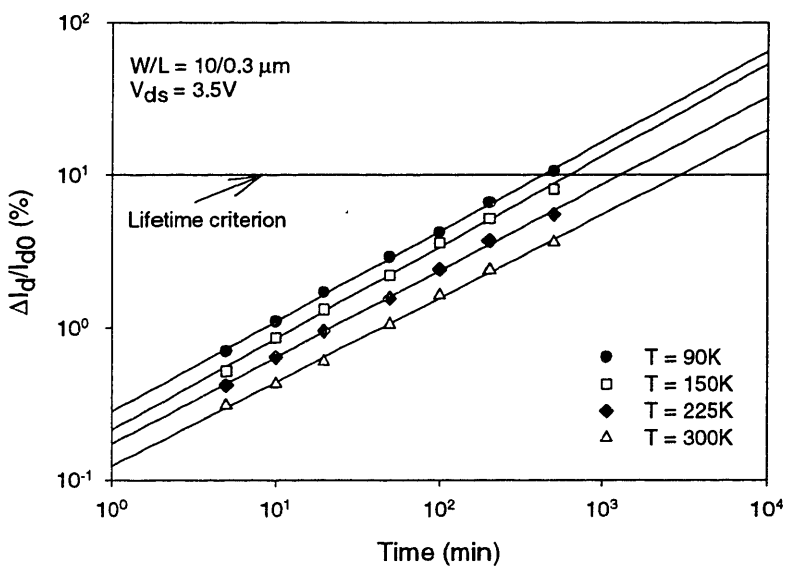

Figure 6.35: Lifetime extracted based on $\Delta I_d/I_{d0}$ = 10% at different temperatures.

Defining the lifetime criterion as $\Delta I_d/I_{d0}$=10%, we observe that the lifetime is shortened at cryogenic temperature. Recall, however, that this is the lifetime extracted when stressed at V_{ds} = 3.5V. A typical operational voltage for 0.3 μm technology is 3V. A similar plot is shown in Figure 6.36 when $\Delta g_m/g_{m0}$=20% is used as the lifetime criterion. Recall that g_m is an important parameter for analog designers. We observe a very similar dependence of the device lifetime on temperature in Figure 6.36.

In order to quantify how the device lifetime changes as a function of temperature, we plot the normalized device lifetime against temperature in Figure 6.37. As one can see in the figure, the lifetime obtained at room temperature is defined as 1. Figure 6.37 reveals that, when the lifetime is defined in terms of device performance parameters, $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, rather than the physical damage ΔN_{it}, the device lifetime is an exponential function of temperature. In other words, the lifetime gets exponentially shorter as the tempera-

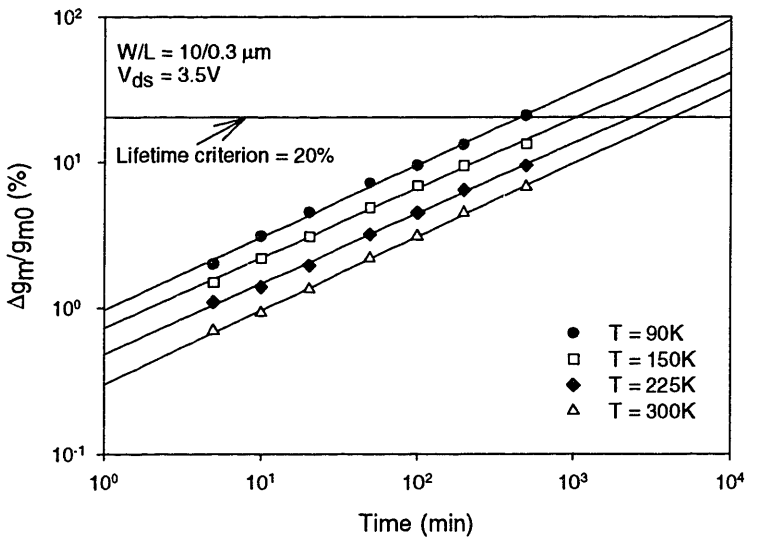

Figure 6.36: Lifetime extracted based on $\Delta g_m/g_{m0}$=20% at different temperatures.

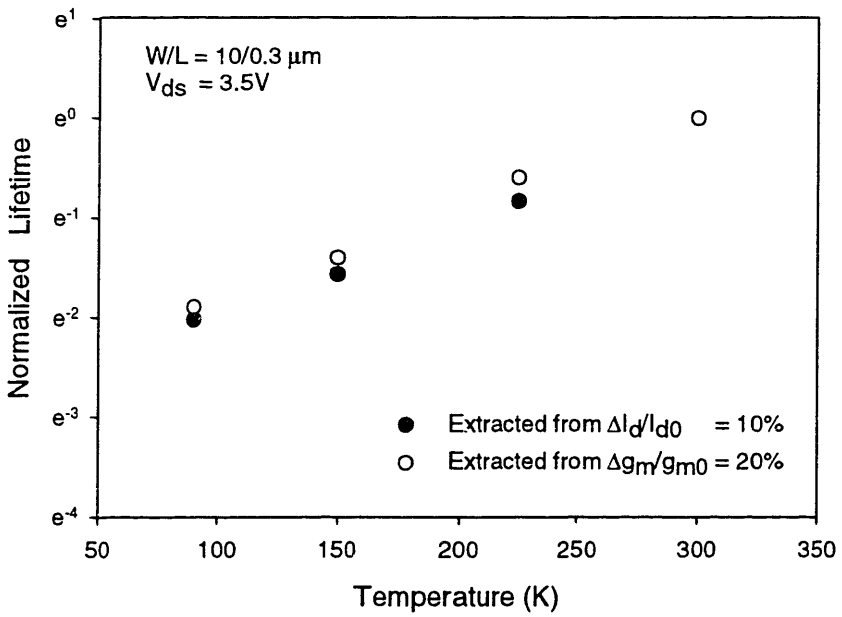

Figure 6.37: Exponential dependence of the device lifetime on temperature.

ture is lowered from T=300K to T=90K. This is a very important finding since it implies that hot-carrier reliability can be a bottleneck for the low temperature operation of MOSFETs.

There are two reasons that account for the reduced lifetime at cryogenic temperatures for the same fixed supply voltage. First, experimental measurement shows that the I_{sub}/I_d ratio, thus E_p, is greater at low temperatures even for the same supply voltage. This is due to increased I_{sub} because of the increased electron mean-free path. Hence, we can deduce that the amount of physical damage, ΔN_{it}, is also greater at low temperature even for the same stressing condition. Second, as explained in detail in Section 6.5, even for the same amount of physical damage, the device degradations, $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, are greater at low temperatures. These two mechanisms collectively account for the signifi-

cantly reduced device lifetime at cryogenic temperatures. Recall that this result is based on the stressing condition V_{ds} = 3.5V for the channel length of 0.3 μm device. At the operational voltage 3V, the actual lifetime values would be different, but we expect the same trend in terms of temperature dependence of the device lifetime.

In order to quantify performance vs. reliability trade-off, Figure 6.38 plots normalized performance against normalized reliability, using Figure 6.5 and Figure 6.37. The plot is shown for different temperatures for the channel length of 0.3 μm. Again, I_d at T=300K was used as value 1 to normalize the performance as a function of temperature. One interesting point to observe from Figure 6.38 is that although the lifetime extracted from $\Delta g_m / g_{m0}$ gets reduced as rapidly as that extracted from $\Delta I_d / I_{d0}$ as the temperature is lowered, the performance improvement is not as significant. This suggests that for analog circuits, the low temperature operation may marginally boost the performance, but the device lifetime gets reduced significantly. Even though the overall circuit lifetime at cryogenic temp-

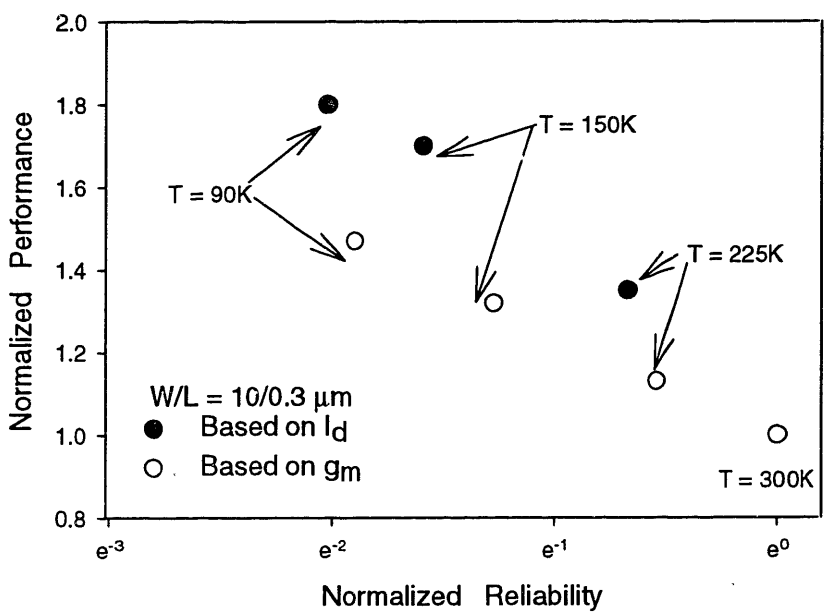

Figure 6.38: Performance vs. Reliability trade-off at different temperatures.

erature is entirely a different study, the analog designer should keep in mind that each device's lifetime in analog circuits can get reduced by a significant amount.

In order to observe how this trend changes at different channel lengths, this analysis has been repeated for a range of device channel lengths from 0.3 μm to 1 μm, and the result is shown in Figure 6.39. For Figure 6.39, $\Delta I_d/I_{d0}=10\%$ is used as the lifetime criterion. One can observe that the lifetime suffers more at low temperatures as the channel length is reduced. This suggests that the low temperature operation of short-channel devices requires careful attention to the device lifetime. A similar plot for different channel lengths at different temperatures is shown with $\Delta g_m/g_{m0}=20\%$ definition in Figure 6.40. The trend of reduced reliability for short-channel devices is even more visible in this plot. As the channel length is reduced, the normalized performance does not improve,

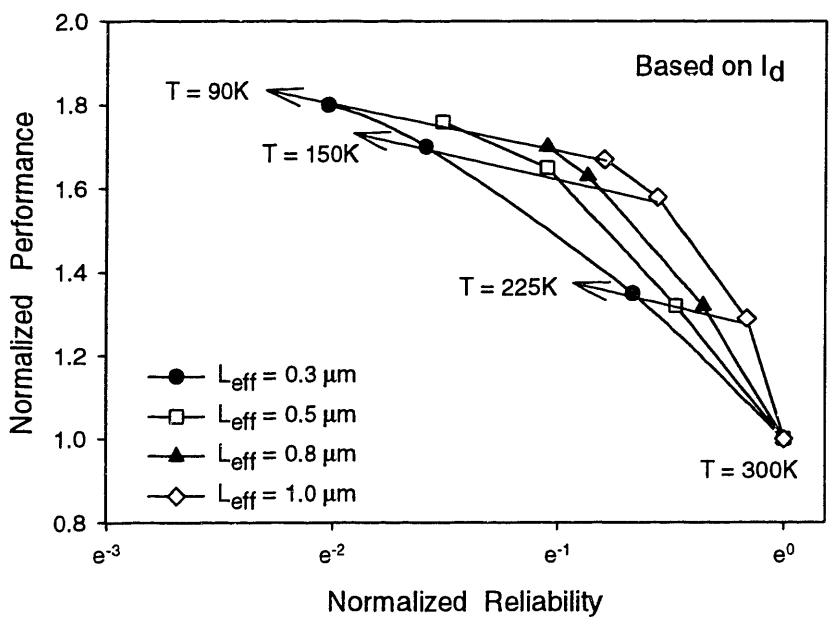

Figure 6.39: Performance vs. Reliability trade-off with I_d as parameter as a function of channel length and temperature.

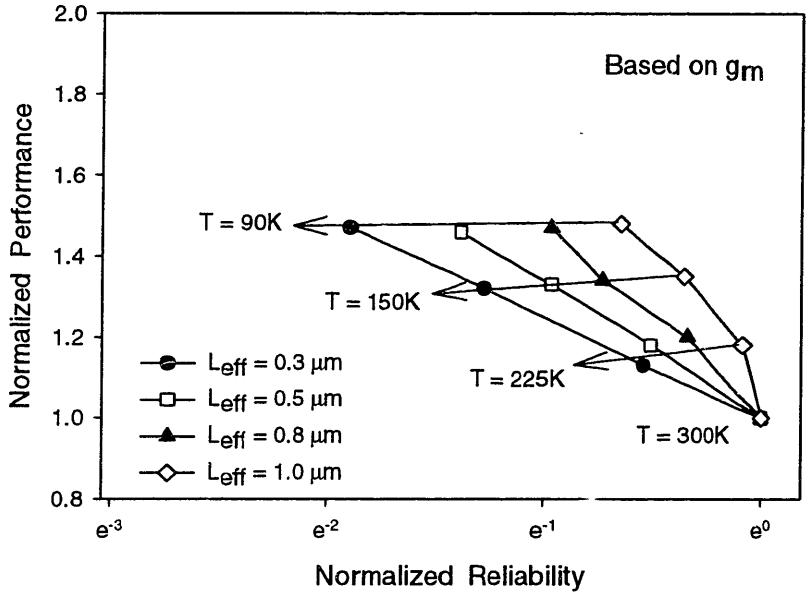

Figure 6.40: Performance vs. Reliability trade-off with g_m as parameter as a function of channel length and temperature.

whereas the normalized lifetime suffers much more significantly at low temperatures.

The trade-off plots from Figure 6.38 to Figure 6.40 show that the low temperature operation of MOSFETs requires careful attention to hot-carrier reliability. Even though the device performance boost is observed, the device lifetime is significantly reduced at cryogenic temperatures from the circuit designers' perspective. It is also interesting to notice that the performance is *linearly* dependent on temperature with negative correlation, whereas the lifetime is *exponentially* dependent on temperature with positive correlation. The reduced reliability is even more visible for short-channel devices for both digital and analog circuit applications.

6.6 Summary of Temperature Dependence

In this chapter, we have studied temperature dependence of device performance and hot-carrier reliability from room temperature down to liquid nitrogen temperature. First, we have characterized the improved device characteristics, such as enhanced mobility and steeper subthreshold slope, at low temperatures. In terms of device performance parameters, this translates to higher current drive, I_d, and higher transconductance, g_m, which are important parameters of interest for circuit designers.

For hot-carrier reliability, we have experimentally verified the applicability of the room temperature model at cryogenic temperatures with appropriately extracted model parameters. This verification is further supported by studying the dominant degradation mechanism at low temperatures. Over the range of temperatures from T=300K down to T=90K, we find that interface-trap generation is the dominant degradation mechanism as is the case in room temperature. Since the room temperature model calculates hot-carrier degradation based on ΔN_{it} mechanism, it is not surprising that this model can be extended down to liquid nitrogen temperature where ΔN_{it} is still the dominant mechanism.

An interesting observation that we made for low temperature reliability is that the amount of physical damage, i.e., increase in the number of interface-traps, is the same as in room temperature for the same I_{sub}/I_d ratio. This suggests that the physical damage is not exacerbated at cryogenic temperatures for the same strength of E-field. However, we do observe much greater degradation, even for the same I_{sub}/I_d ratio, in terms of device performance parameters, such as $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$. This is due to enhanced mobility degradation at low temperature after hot-carrier stressing. The enhanced mobility degrada-

tion is attributed to different scattering mechanisms, such as coulomb and surface roughness scattering, having greater influence on the overall inversion layer electron mobility.

After having understood hot-carrier reliability issues at low temperature, we carried out a trade-off analysis of performance vs. reliability as a function of temperature and device channel length. The metrics chosen for performance are I_d and g_m, and the metrics for reliability are defined to be the device lifetime at the same power supply voltage, given the lifetime criteria of $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$. The trade-off plot shows that the device lifetime, when defined in terms of $\Delta I_d/I_{d0}$ and $\Delta g_m/g_{m0}$, suffers at low temperature. The shortened lifetime is attributed to *increased I_{sub}/I_d ratio* even for the same power supply voltage, and *greater manifestation of hot-carrier damage* at cryogenic temperature. Performance parameters, such as, I_d and g_m, however, improve at low temperatures as reported from numerous previous studies. The trade-off analysis gives a good guideline to circuit designers in terms of their anticipation of performance improvement and reliability sacrifice as the circuits are designed for low temperature applications.

References

[6.1]. F. Gaensslen, V. Rideout, E. Walker, and J. Walker, "Very Small MOSFET's for Low-Temperature Operation," IEEE Transactions on Electron Devices, Vol. 24, p218, 1977.

[6.2]. G. Baccarani, M. Wordeman, R. Dennard, "Generalized Scaling Theory and Its Application to a 1/4 Micrometer MOSFET Design," IEEE Transactions on Electron Devices, Vol. 31, p452, 1984.

[6.3]. F. Balestra and G. Ghibaudo, "Brief Review of the MOS Device Physics for Low Temperature Electronics," Solid-State Electronics, Vol. 37, p1967, 1994.

[6.4]. G. Ghibaudo, F. Balestra, and A. Emrani, "A Survey of MOS Device Physics for Low Temperature Electronics," Microelectronic Engineering, p833, 1992.

[6.5]. S. K. Tewksbury, "Attojoule MOSFET Logic Device Using Low Voltage Swings and Low Temperature," Solid-State Electronics, p255, 1985.

[6.6]. Avid Kamgar, "Subthreshold Behavior of Silicon MOSFETs at 4.2K," Solid-State Electronics, p537, 1982.

[6.7]. Giorgia Baccarani, Matthew Wordeman, and Robert Dennard, "Generalized Scaling Theory and Its Application to a 1/4 Micrometer MOSFET Design," IEEE Transactions on Electron Devices, Vol. 31, p452, 1984.

[6.8]. Jack Sun, Yuan Taur, Robert Dennard, Stephen Klepner, "Submicrometer-Channel CMOS for Low-Temperature Operation," IEEE Transactions on Electron Devices, Vol. 34, p19, 1987.

[6.9]. Tong-Chern Ong, Ping Ko, and Chenming Hu, "50A Gate-Oxide MOSFET's at 77K," IEEE Transactions on Electron Devices, Vol. 34, p2129, 1987.

[6.10]. You-WEn Yi, Kazuya Masu, Kazuo subouchi, and Nobuo Mikoshiba, "Temperature-Scaling Theory for Low-Temperature-Operated MOSFET with Deep-Submicron Channel," Japanese Journal of Applied Physics, Vol. 27, p1958, 1988.

[6.11]. Albert Henning, Nelson Chan, and James Plummer, "Substrate Current in N-Channel and P-Channel MOSFETs Between 77K and 300K: Characterization and Simulation," International Electron Device Meeting, p573, 1983.

[6.12]. D. Ju, R. Reich, J. Schrankler, M. Hot, and G. Kirchner, "Transient Substrate Current Generation and Device Degradation in CMOS Circuits at 77K," International Electron Device Meeting, p569, 1983.

[6.13]. Alexandre Acovic, Michel Dutoit, and Marc Ilegems, "Characterization of Hot-Electron-Stressed MOSFETs by Low-Temperature Measurements of the Drain Tunnel Current," IEEE Transactions on Electron Devices, p1467, 1990.

[6.14]. Robert Fox, and Richard Jaeger, "MOSFET Behavior and Circuit Considerations for Analog Applications at 77K," IEEE Transactions on Electron Devices, Vol. 34, p114, 1987.

[6.15]. Chien Nguyen-Duc, Sorin Cristoloveanu, and Gilles Reimbold, "Effects of Localized Interface Defects Caused by Hot-Carrier Stress in n-Channel MOSFETs at Low Temperature," IEEE Electron Device Letters, Vol. 9, p479, 1988.

[6.16]. K. Fu, "Mobility Degradation due to the Gate Field in the Inversion Layer of MOSFETs," IEEE Electron Device Letters, p292, 1985.

[6.17]. Y. Cheng and E. Sullivan, "Effect of Coulomb Scattering on Silicon Surface Mobility," American Institute of Physics, p187, 1974.

[6.18]. C. Jacoboni, C. Canali, G. Ottaviani, and A. Quaranta, "A Review of Some Charge

Transport Properties of Silicon," Solid-State Electronics, p77, 1977.

[6.19]. R. Coen and R. Muller, "Velocity of Surface Carriers in Inversion Layers on Silicon," Solid-State Electronics, p35, 1980.

[6.20]. A. Hartstein, A. Fowler, and M. Albert, "Temperature Dependence of Scattering in the Inversion Layer," Surface Science, p181, 1979.

[6.21]. Anant Sabnis and James Clemens, "Characterization of the Electron Mobility in the Inverted <100> Si Surface," International Electron Device Meeting, p48, 1979.

[6.22]. Winfried Soppa and Hans-Gunther Wagemann, "Investigation and Modeling of the Surface Mobility of MOSFETs from -25 to 150C," IEEE Transactions on Electron Devices, Vol. 35, p970, 1988.

[6.23]. Shiuh-Wuu Lee, "Universality of Mobility-Gate Field Characteristics of Electrons in the Inversion Charge Layer and Its Application in MOSFET Modeling," IEEE Transactions on Computer Aided Design, Vol. 8, p724, 1989.

[6.24]. Mong-Song Liang, Jeong Yeol Choi, Ping Ko, and Chenming Hu, "Inversion-Layer Capacitance and Mobility of Very Thin Gate-Oxide MOSFETs," IEEE Transactions on Electron Devices, Vol. 33, p409, 1986.

[6.25]. John Wakerly, Digital Design Principles & Practices, Prentice Hall, 1997.

[6.26]. Jan Rabaey, Digital Integrated Circuits, Prentice Hall, 1996.

[6.27]. Neil Weste and Kamran Eshraghian, Principles of CMOS VLSI Designs, Addison-Wesley, 1992.

[6.28]. Siegfried Selberherr, "MOS Device Modeling at 77K," IEEE Transactions on Electron Devices, p1464, 1989.

[6.29]. C. Huang and G. Gildenblat, "An Accurate Engineering Model of an n-Channel

MOSFET for 60-300K Temperature Range," Solid-State Electronics, p1309, 1990.

[6.30]. Jacquelyn Hill, Renato Pires, Richard Anderson, "Scaling Silicon MOSFET's for 77K Operation," IEEE Transactions on Electron Devices, p2497, 1991.

[6.31]. Charlie Sodini, Microelectronics, Prentice Hall, 1997.

[6.32]. G. Gildenblat, L. Colonna-romano, D. Lau, and D. Nelson, "Investigation of Cryogenic CMOS Performance," International Electron Device Meeting, p268, 1985.

[6.33]. Armond Hairapetian, Daniel Gitlin, and C. Viswanathan, "Low-Temperature Mobility Measurements on CMOS Devices," IEEE Transactions on Electron Devices, p1448, 1989.

[6.34]. Ken'ichiro Sonoda, Kenji Taniguchi, and Chihiro Hamaguchi, "Analytical Device Model for Submicrometer MOSFETs," IEEE Transactions on Electron Devices, Vol. 38, p2662, 1991.

[6.35]. Michio okoyama, Tetsuya Hidaka, You-Wen Yi, Kazuya Masu, and Kazuo Tsubouchi, "Low Temperature MOSFET Operation by the Temperature Scaling Theory," International Conference on Solid State Devices and Materials, p499, 1992.

[6.36]. M. Deen, and Z. Yan, "DIBL in Short-Channel NMOS Devices at 77K," IEEE Transactions on Electron Devices, Vol. 39, p908, 1992.

[6.37]. V. Rideout, F. Gaensslen, and A. LeBlanc, "Design considerations for ion implanted n-channel MOSFETs," IBM J. Res. Develop., p50, 1975.

[6.38]. S.M. Sze, VLSI Technology, McGraw Hill, 1988.

[6.39]. G. Gildenblat, MOS Advanced Device Physics, Chapter 5, Academic Press, 1989.

[6.40]. Hong-June Park, Ping Ko, and Chenming Hu, "A Charge Sheet Capacitance Model of Short Channel MOSFETs for SPICE," IEEE Transactions on Computer-

Aided Design, Vol. 10, p376, 1991.

[6.41]. Zhi-Hong Liu, Chenming Hu, Jian-Hui Huang, Tung-Yi Chan, Min-Chie Jeng, Ping Ko, and Y. C. Cheng, "Threshold Voltage Model for Deep-Submicrometer MOSFETs," IEEE Transactions on Electron Devices, p86, 1993.

[6.42]. J.A. Bacchitta, T. L. Honan, and R. L. Anderson, "Hot-electron-induced degradation in MOSFET's at 77K," IEEE Transactions on Electron Devices, Vol. Ed-32, p1850-157, 1985.

[6.43]. F. Hsu, and S. Tam, "Relationship Between MOSFET Degradation and Hot-Electron-Induced Interface-State Generation," IEEE Electron Device Letters, Vol. 5, p50, 1984.

[6.44]. W. Jiang, H. Le, S. Dao, S.Kim, B. Stine, J. Chung and P. Marcoux, "Key Hot-Carrier Degradation Model Calibration and Verification Issue for Accurate AC Circuit-Level Reliability Simulation," International Reliability Physics Symposium, p300, 1997.

[6.45]. K. Quader, P. Ko, and Chenming Hu, "Simulation of CMOS Circuit Degradation due to Hot-Carrier Effects," International Reliability Physics Symposium, p16, 1992.

[6.46]. P. Lee, M. Kuo, K. Seki, Ping Ko, and Chenming Hu, "Circuit Aging Simulator (CAS)," International Electron Device Meeting, p. 134, 1988.

[6.47]. Chenming Hu, Advanced MOS Device Physics, Chapter 3, Academic Press, 1989.

[6.48]. C. Hu, S. Tam, F. Hsu, P.Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation - Model, Monitor, and Improvement," IEEE Transactions on Electron Devices, Vol. 32, p375, 1985.

[6.49]. S. Sun, and J. Plummer, "Electron Mobility in Inversion and Accumulation Layers on Thermally Oxidized Silicon Surfaces," IEEE Transactions on Electron Devices, Vol. 27, p1497, 1980.

[6.50]. J. Chung, P. Ko, and C. Hu, "A Model for Hot-Electron-Induced MOSFET Linear-Current Degradation Based on Mobility reduction Due to Interface-State Generation," IEEE Transactions on Electron Devices, Vol. 38, p1362, 1991.

[6.51]. Neil Ashcroft and N. Mermin, Solid-State Physics, Saunders College Publishing, 1976.

[6.52]. N. Arora, G. Gildenblat, "A Semi-Empirical Model of the MOSFET Inversion Layer Mobility for Low-Temperature Operation," IEEE Transactions on Electron Devices, Vol. 34, p28, 1987

[6.53]. T. Ando, A. Fowler, and F. Stern, "Electronic Properties of Two-Dimensional Systems," Rev. Mod. Phys., vol. 54, p437, 1982.

[6.54]. A. Hartstein, A. Fowler, and M. Albert, "Temperature Dependence of Scattering in the Inversion Layer," Surface Science, vol. 98, p181, 1980.

Chapter 7

Conclusion

7.1 Summary of Research

This dissertation has investigated major issues of hot-carrier reliability of MOSFETs. Starting with some background theories, it has investigated the existing room temperature model of NMOS hot-carrier degradation. The first major issue that was discussed was the degradation model parameter extraction. Chapter 2 discussed the procedure of the model parameter extraction; how they can be accurately extracted, and what the physical meaning of each parameter is.

Chapter 3 continued to characterize the model parameters, and discovered the universal dependence of the time-acceleration factor **n** and the voltage-acceleration factor **m** on the oxide field. It was also shown that these parameters are not constant, but a strong function of bias conditions of the MOSFET, such as V_{ds} and V_{gs}. The implication of bias-dependent parameters necessitates revisiting the traditional device-lifetime prediction procedure. Chapter 4 discussed how erroneous the traditional lifetime prediction method can be by not carefully taking into account the time-varying degradation model parameters. Based on close inspection of the degradation behavior, however, a new algorithm, which identifies the dominant degradation component at the future time point of interest, is presented in Chapter 5. This algorithm is shown to be much more accurate and efficient in terms of simulation time.

Chapter 6 discussed hot-carrier reliability at cryogenic temperatures. First, the motivation for low temperature operation of the MOSFET was discussed by showing the

improved device performance characteristics. In order to fully maximize the performance benefits at low temperature, however, the hot-carrier reliability issue must be clearly addressed at low temperature. It was found that even though the physical hot-carrier damage is not exacerbated, its *manifestation* in terms of device performance parameters, I_d and g_m, is shown to degrade significantly more at cryogenic temperature. This raises a concern for circuit designers whose primary interest in the device is its performance parameters, such as I_d and g_m, rather than the physical damage. Having formulated that, a trade-off analysis was carried out between device performance and reliability as a function of temperature and device channel length. The trade-off analysis shows that both device engineers and circuit designers need to be aware of the risk of the low temperature operation of MOSFETs since hot-carrier reliability can serve as a design bottleneck due to the reduced device lifetime at cryogenic temperature.

7.2 Future Research

This dissertation focused on hot-carrier reliability issues at the device level, i.e., how the device lifetime can be accurately predicted, and how it changes as we decrease the operation temperature. In the Microsystems Technology Laboratories (MTL) at MIT, circuit-level hot-carrier reliability studies have been also carried out [7.1-3]. These studies have focused on identifying important issues affecting hot-carrier reliability in both digital and analog sub-circuits, such as ring oscillator, adder, differential amplifier, and current mirror circuits. There are also some previous theses elsewhere, whose primary research was at the circuit level [7.4-5].

In the work presented here, however, we have not examined the results or implica-

tions at the system level, which encompasses a range of end products that are actually being used by consumers. For example, we still need deeper understanding of how microprocessor performance degrades over time due to hot-carrier degradation. As another example of digital circuit, memory circuits, such as DRAM, SRAM, and CAM, need to be analyzed in terms of how their performance, such as their access time, degrades over time due to hot-carrier degradation. In the analog domain, the variety of circuits is even greater. Operational amplifiers, frequency synthesizers, and A/D converters are prominent ones to name a few. More research needs to be carried out to understand how hot-carrier reliability affects these systems' performance. This dissertation and others have formed a good foundation for understanding hot-carrier reliability at the device and circuit level. Future research in this field would need to focus on higher-level systems to go beyond academic impact.

References

[7.1]. Vei-han Chan, Hot-Carrier Reliability Evaluation for CMOS Devices and Circuits, Ph.D dissertation, Massachusetts Institute of Technology, 1995.

[7.2]. Wenjie Jiang, Hot-Carrier Reliability Assessment in CMOS Digital Integrated Circuits, Ph.D dissertation, Massachusetts Institute of Technology, 1998.

[7.3]. Huy Le, On the Methodology of Assessing Hot-Carrier Reliability of Analog Circuits, Ph.D dissertation, Massachusetts Institute of Technology, 1999.

[7.4]. Khandker Quader, Hot-Carrier Reliability of Integrated Circuits, Ph.D dissertation University of California at Berkeley, 1993.

[7.5]. Peter Lee, Modeling and Simulation of Hot-Carrier Effects in MOS Devices and Circuits, Ph.D dissertation, University of California at Berkeley, 1989.

THESIS PROCESSING SLIP

FIXED FIELD: ill. _____ name _____

index _____ biblio _____

▶ COPIES: (Archives) Aero Dewey (Eng) Hum

Lindgren Music Rotch Science

TITLE VARIES: ▶ ☐ _____

_____

_____

NAME VARIES: ▶ ☐ _____

_____

IMPRINT: (COPYRIGHT) _____

_____

▶ COLLATION: 151 l. _____

_____

▶ ADD: DEGREE: _____ ▶ DEPT.: _____

SUPERVISORS: _____

_____

_____ _____

_____ _____

_____ _____

_____ _____

_____ _____

_____ _____

_____ _____

NOTES:

cat'r: date:

page: J181

▶ DEPT: E. E.

▶ YEAR: 1999 ▶ DEGREE: Ph.D.

▶ NAME: KIM, Seok Won Abraham

_____