Low Voltage Silicon Dioxide Reliability

by

Klaus Florian Schuegraf

B.S. (Rensselaer Polytechnic Institute) 1987
M.S. (Stanford University) 1988

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Engineering- Electrical Engineering
and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Chenming Hu, Chair
Professor Andrew Neureuther
Professor Leo Falicov

1994

The dissertation of Klaus Florian Schuegraf is approved:

April 28, 1994

Chair Date

L M. FALICOV APR 28 1994

Date

Date

University of California at Berkeley

1994

Low Voltage Silicon Dioxide Reliability

by

Klaus Florian Schuegraf

Abstract

Low Voltage Silicon Dioxide Reliability

by

Klaus Florian Schuegraf

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Chenming Hu, Chair

This study investigates the low voltage breakdown and conduction properties of thin thermal silicon dioxides with thickness ranging from 25 Å to 130 Å using silicon metal-oxide-semiconductor (MOS) capacitors and transistors.

Investigation of oxide breakdown shows that anode hole injection is the likely mechanism responsible for silicon dioxide wearout. A quantitative model for oxide breakdown based on anode hole injection is proposed. This model not only agrees with the predictions of an empirical inverse oxide field model ("1 over E model"), it offers a methodology to extrapolate high field oxide breakdown data into the low voltage operating regime. Investigation of low voltage silicon dioxide tunneling current shows that leakage current increases dramatically above that predicted by the Fowler-Nordheim theory for oxide voltages less than 3.15 Volts, possibly setting a scaling limit on thickness at 40 Å.

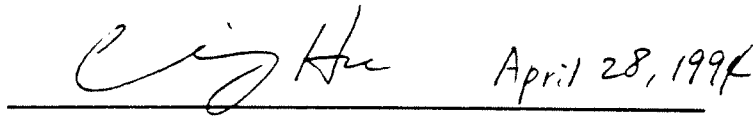
The model is also used to compare the breakdown of p^+ and n^+ polysilicon gate, showing no change in damage initiation mechanism. The temperature dependence of breakdown is also investigated, showing that anode hole injection models the breakdown characteristics accurately for temperatures below 150° C. The temperature

acceleration of breakdown is attributed to the oxide's reduced hole immunity at higher temperatures.

The anode hole injection model is thereby shown to model the breakdown characteristics of defect-free, "intrinsic" oxide very accurately. A defect model combining the breakdown mechanism of anode hole injection with "effective thinning" is able to characterize defect breakdown distributions by attributing the breakdown to an effective thinning of the oxide at some localized point.

Substrate current measurements during electrical breakdown stress establish the basis for the anode hole injection model. Investigations of substrate current in oxides thinner than 55 Å show that the anode hole injection current becomes dominated by the tunneling of valence-band electrons by means of either trap-assisted-tunneling through stress-induced traps in oxides between 45 and 55 Å or direct tunneling in oxides thinner than 45 Å. Nevertheless, the anode hole injection model derived for thicker oxides holds for the thinner oxides as well, despite the impediment to direct verification.

The dissertation of Klaus Florian Schuegraf is approved:



Professor Chenming Hu Date

Committe Chairman

To my family

Table of Contents

Acknowledgments

This thesis begets my most sincere gratitude for Professor Chenming Hu who served as my research advisor. Without Chenming's dedicated support, guidance, encouragement, and, above all, belief in my ability, I doubt very much that this thesis or my doctoral program would have reached fruition. I am very grateful for the "second chance" granted me and hope this thesis is but the first opportunity to display the education that Chenming has invested so much of his time and energy in. His enthusiasm, insight, standards of rigor and excellence will be sorely missed. His modesty, friendliness and approachability set the standards we should all aspire to.

I would also like to acknowledge the late Professor Shyh Wang, my previous research advisor, for his commitment to the belief that an education in the natural sciences requires schooling in fundamental math and physics. His patience and provision of the resources necessary to pursue graduate coursework in the physics department attests to this vision. He wished his students would pursue research of long-term and fundamental value to industry; I hope that my attempts satisfied this hope.

I am grateful to Professors Ping Ko, Andy Neureuther and Leo Falicov for serving on my qualifying examination committee and their subsequent role in providing the feedback leading to this manuscript. I would also like to thank the Microfabrication Laboratory staff, especially Katalin Voros, Robert Hamilton and Tariq Haniff, for their help in providing a friendly and functioning environment for device fabrication.

I would also like to express my utmost gratitude to the Joint Services Electronics Program (JSEP) for its financial support in the form of a most generous fellowship grant. It is my deepest hope that this work satisfies the JSEP Technical Coordinating

Committee's mission of fertilizing technological research vital to our national security interests. I reiterate my thanks for the trust that James Mink, Kenneth Davis and Horst Wittman placed in my abilities, not to mention the privilege of being supported by a program rooted in the legacy of Vannevar Bush.

It has been a privilege to interact with the students belonging to the solid-state device and quantum electronics research groups. They have brought insight, intelligence, entertainment and friendship to my tenure at Berkeley. Although by no means exhaustive, I would like to thank Fariborz Assaderaghi, Olga Blum, Mansun Chan, Jian Chen, Kai Chen, Gary Fedder, Mark Hadley, Patrick Harshman, Jian-Hui Huang, Kelvin Hui, Mike and Jack Judy, Julie Kenrow, Joe King, John Krick, Scott Kuusinen, Chester Li, Zhi-Hong Liu, Dave Loconto, James Ma, Eric Minami, Reza Moazzami, Donggun Park, Steve Parke, Khandker Quader, Bob Ried, Elyse Rosenbaum, Hyungcheol "Huck" Shin, Yaron Simler, Jay Tu, Robert Tu, Ashish Verma, Clement Wann, Kirt Williams and George Zhang for the interactive and fertile environment they shaped.

I would also like thank a few key friends and mentors who provided the support and encouragement for perseverance when it was most sorely needed. Firstly, I would like to thank Professor Robert M. Gray at Stanford University for suggesting the transfer to the University of California, Berkeley. I could not have surmised at the time how fortuitous I would discover this advice to be. I would like to thank Raymond and Vara Kraft, Rana "Ellen" Khayata, John Kibarian, Eric Boskin, Fariborz Assaderaghi, Patrick Harshman and Scott Silverman for believing in me, listening to me and helping me weather the storms; they have proved instrumental and I wonder who I would be without the time and energy they have given so selflessly.

Chapter 1

Introduction

1.1. Introduction

This chapter introduces the problems pertaining to the reliability of thin silicon dioxide in its use as an insulating material in silicon based integrated circuit technology. Silicon dioxide, a natural oxide of silicon, is an amorphous insulator displaying excellent interfacial properties, i.e. high silicon-silicon dioxide barrier height, excellent reproducibility and stability, etc., which enable the realization of commercially viable, high density integrated circuits of bipolar and field effect transistors (FET) on silicon substrate. Since the successful commercialization of silicon gate FET circuits in 1972, scaling of all device features, i.e. channel length, gate oxide thickness, junction depths, etc., has led to a more than 16000-fold increase in circuit density and a concommitant increase in circuit performance, enabling technological advance from centralized to personalized information processing and computing.

This scaling [1.1,1.2,1.3] has led to a reduction in FET gate length from 5 microns in 1976 to 0.5 microns in 1992, with 0.25 micron development efforts being vigorously pursued world-wide by leading-edge industrial concerns. The gate insulator thickness has also been scaled from more than 1000 Å, to below 100 Å. Due to adherence to industry standards, supply voltages have generally not been decreased; therefore, transistors belonging to subsequent generations of technology operate under ever-increasing electric fields. These high operating fields pose considerable risk to reliable circuit operation and have fertilized active research in the understanding of the

mechanisms which lead to the high-field degradation and wear-out observed in silicon MOSFETs.

A major recent industry trend is a move toward decreased power supply voltages, motivated by the desire to deliver integrated circuits with lower power consumption. Not only does this trend have important reliability ramifications, it also marks the arrival of "environmentally correct" applications. For instance, Digital Equipment Corporation has repeatedly advertised its "green" personal computer in the New York Times this spring[1.4]. Federal government procurement regulations also promise to accelerate the advance of low-power, energy-efficient computing. In addition to promising more energy efficient circuits and systems, this trend will enable an accelerated introduction of portable electronic systems, since reduced power consumption translates directly into extended battery life, not to mention enhanced consumer mobility and applications flexibility.

Therefore, understanding transistor reliability issues in the low voltage regime is a crucial milestone in the advance of silicon integrated circuit technology. This thesis focuses on one particular aspect of integrated circuit reliability, the breakdown of thin silicon dioxide at low voltages. The remainder of this first chapter introduces the important breakdown issue by reviewing the previous state of the art.

1.2. Time-dependent dielectric breakdown

The scaling of device dimensions, resulting in the thinning of oxide dielectrics in VLSI (Very Large Scale Integration) technologies, is driven by the desire to deliver more MOSFET current and, hence, circuit speed at low voltages for digital logic applications, to provide the same capacitance within a smaller area for DRAMs (Dynamic Random Access Memories), and to offer lower programming voltages for

non-volatile memories. Due to the high fields present in the silicon oxide mandated by this aggressive dielectric scaling and the presence of oxide defects, catastrophic failure of the dielectric has always been the predominant oxide reliability concern, thereby limiting scaling. Figure 1.1 illustrates the oxide fields at which specific reliability concerns are raised [1.1,1.5].

Since practical considerations limit the oxide integrity assessment on actual circuits, the oxide integrity question is usually abstracted to the breakdown study of capacitors whose area is comparable to or larger than the area of active transistor area in the circuit. We need to arrive at an accelerated testing paradigm which allows the engineer to extrapolate the lifetime of good product, based on a minimal number of short time, high field accelerated tests. This research focused on the development of techniques which allow an engineer to determine whether or not his or her oxide technology is sufficiently robust to guarantee the lifetime specification of her product at its intended operating voltage, including today's operating voltages at or below 3.3 Volts.

1.3. Intrinsic Breakdown

The term intrinsic breakdown refers to the breakdown of an oxide that is free of defects, thereby reflecting the practical upper limit of oxide reliability. Intrinsic lifetime can be determined either from the 100% cumulative failure point in the failure distribution or by measuring breakdown lifetime of small area test structures. Figure 1.2 illustrates the use of widely used E-method introduced by Berman [1.6,1.7] for projecting the lifetime from high-field stress data. This extrapolation method assumes an empirical linear relationship between the logarithm of the breakdown lifetime and applied stress field E_{ox}, such that

$$t_{bd} = \tau \exp\left(\frac{-\beta V_{ox}}{X_{ox}}\right) \tag{1.1}$$

where the electric field acceleration factor β reported in the literature has varied widely, from 7 to 1.5 decade/(MV/cm) as shown in Figure 1.3 [1.8]. Since a small variation in β can lead to orders of magnitude error in the extrapolated lifetime, the wide variation of β cast considerable doubt on this technique as an acceptable accelerated testing paradigm. Using Eq. (1.1), the statistical t_{bd} data under a high stress voltage can be used to predict the statistical distribution of oxide lifetime at a low operating voltage.

1.4. Hole-induced breakdown model

Unlike the aforementioned E-model, this model rests on a physical model for breakdown. As a result of the passage of electrons at high field, holes are created in the oxide or at the interfaces. Furthermore, experiments [1.9,1.10,1.11] have shown that an oxide's lifetime is the time required for the hole fluence, $Q_p(t)$, to reach some critical value Q_p, independent of the stress voltage. The hole fluence can be modeled as

$$Q_p(t) \propto J_n \alpha t \tag{1.2}$$

where $J_n \propto e^{-B/E_{ox}}$ is the Fowler-Nordheim tunneling current density with $B \approx 270$ MV / cm (the slightly larger value of B than the usual 250 MV/cm) accounts for the omission of the E_{ox}^2 term in the J_n expression) and $\alpha \propto e^{-H/E_{ox}}$ is the hole generation coefficient [1.12] with $H \approx 80$ MV / cm. Therefore the hole fluence follows

$$Q_p(t) \propto e^{-G/E_{ox}} \cdot t \tag{1.3}$$

where G = B + H ≈ 350 MV / cm. The breakdown time, t_{bd}, corresponds to the point at which $Q_p(t)$ reaches its critical value; therefore, t_{bd} has the following field dependence:

$$t_{bd} = \tau_o \cdot e^{G/E_{ox}} \tag{1.4}$$

τ_o was empirically found [1.13] to be 10^{-11} seconds. This theory, corroborated by the data in Figure 1.4, indicates that linear extrapolation of lifetime with respect to inverse field, $1/E_{ox}$, is more accurate than the E_{ox} extrapolation. Moreover, the "electric field acceleration factor," β, which is defined as the slope of log (t_{BD}) versus E_{ox} curve, can be written as

$$\beta \equiv \frac{-\,d\left(\log\left(t_{bd}\right)\right)}{d\,E_{ox}} = \frac{G}{2.3\,E_{ox}^2} = \frac{140}{E_{ox}^2} \quad \left[\frac{\text{decade}}{\text{MV / cm}}\right] \tag{1.5}$$

such that β cannot be constant, but is instead proportional to $\frac{1}{E_{ox}^2}$. Figure 1.3 demonstrates that this model agrees well with the published results for β thereby supporting the notion that lifetime is better extrapolated with respect to $1/E_{ox}$ than E_{ox}.

The above model works well for thin oxides, thinner than 120 Å. For thicker oxides, B in the Fowler-Nordheim expression is known to become large at large electric fields (> 8 MV/cm) due to hole accumulation in the oxide [1.8,1.14]. As a result, G is often larger than 350 MV/cm for accelerated test data of thick oxides. It is advised that 350 MV/cm rather than the measured G be used for thicker oxides in order to obtain a conservative projection of low field lifetime. (This is equivalent to extrapolating Q_{bd} to low field, using Fowler-Nordheim theory rather than measured large B to extrapolate the low field current density, and predicting low field lifetime by dividing Q_{bd} by J_n). This well documented complication makes it inadvisable to use the straightness of the log(t_{bd}) versus E curve of a thicker oxide to verify or disprove the

hole-induced breakdown model. Moreover, this complication is fortuitously absent when evaluating state-of-the-art and future thin oxides.

Also, note the good agreement between model and data in Figure 1.3, even when the oxide is quite thick. The thick oxide tests summarized in Figure 1.3 were concerned with non-intrinsic, defect induced early failures. Apparently, defects in thick oxides tend to behave like thinner oxides.

The rest of this thesis discusses the continued development of this hole-induced breakdown toward a comprehensive quantitative thin oxide breakdown model. This model enables low voltage silicon dioxide reliability characterization. Chapter 2 discusses anode hole injection, supplying theoretical underpinning to the empirical hole-induced breakdown model. The anode hole injection breakdown model enables the prediction of oxide charge- and time-to-breakdown, Q_{bd} and t_{bd}, respectively, at voltages below 5 Volts. Chapter 2 also discusses the low voltage, direct tunneling phenomenon. Chapter 3 examines the dielectric breakdown of the p^+ polysilicon gate, a technology receiving considerable attention in order to realize "surface" p-channel MOSFETs. Chapter 4 examines the role of temperature on accelerating oxide lifetime tests. Chapter 5 examines the role of defects in limiting the lifetime achievable from "intrinsic" quality oxide. Chapter 6 scrutinizes the correlations between substrate hole current and oxide breakdown, analyzing the contributions to substrate current in the thin oxide MOS system, establishing the validity of the anode hole injection model for oxides thinner than 55 Å. Chapter 7 discusses the expansive literature surrounding oxide breakdown and develops a model for oxide damage and breakdown that integrates the reported diverse observations and correlations into a coherent view of oxide breakdown. Chapter 7 also discusses innovations in silicon dioxide processing, trends in silicon oxide technology, and future needs for alternative dielectrics.

Maximum Acceptable Oxide Field

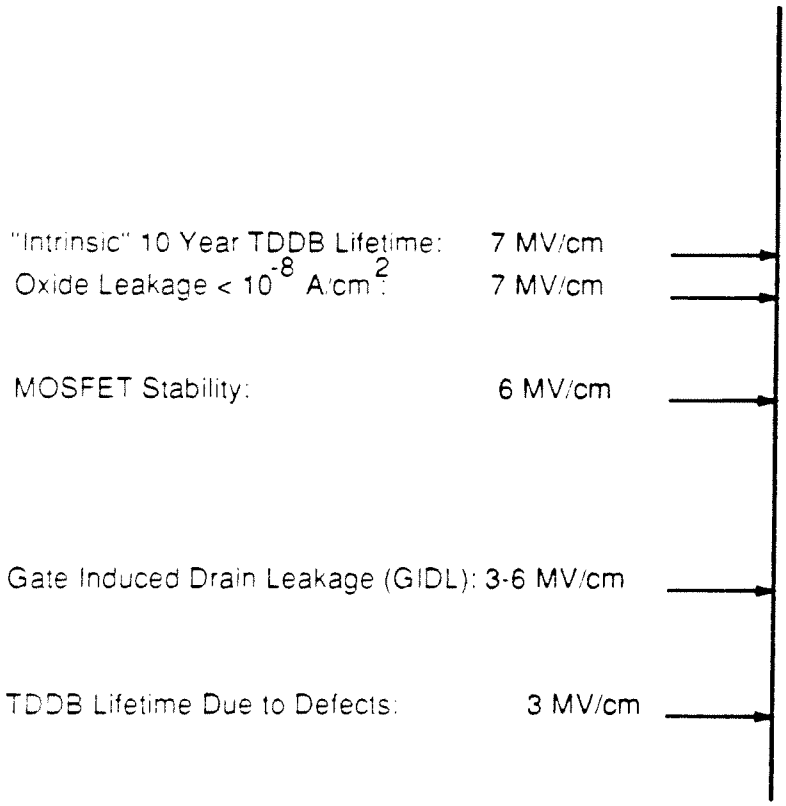


Fig. 1.1 Illustration of oxide fields at which specific device reliability concerns are raised.

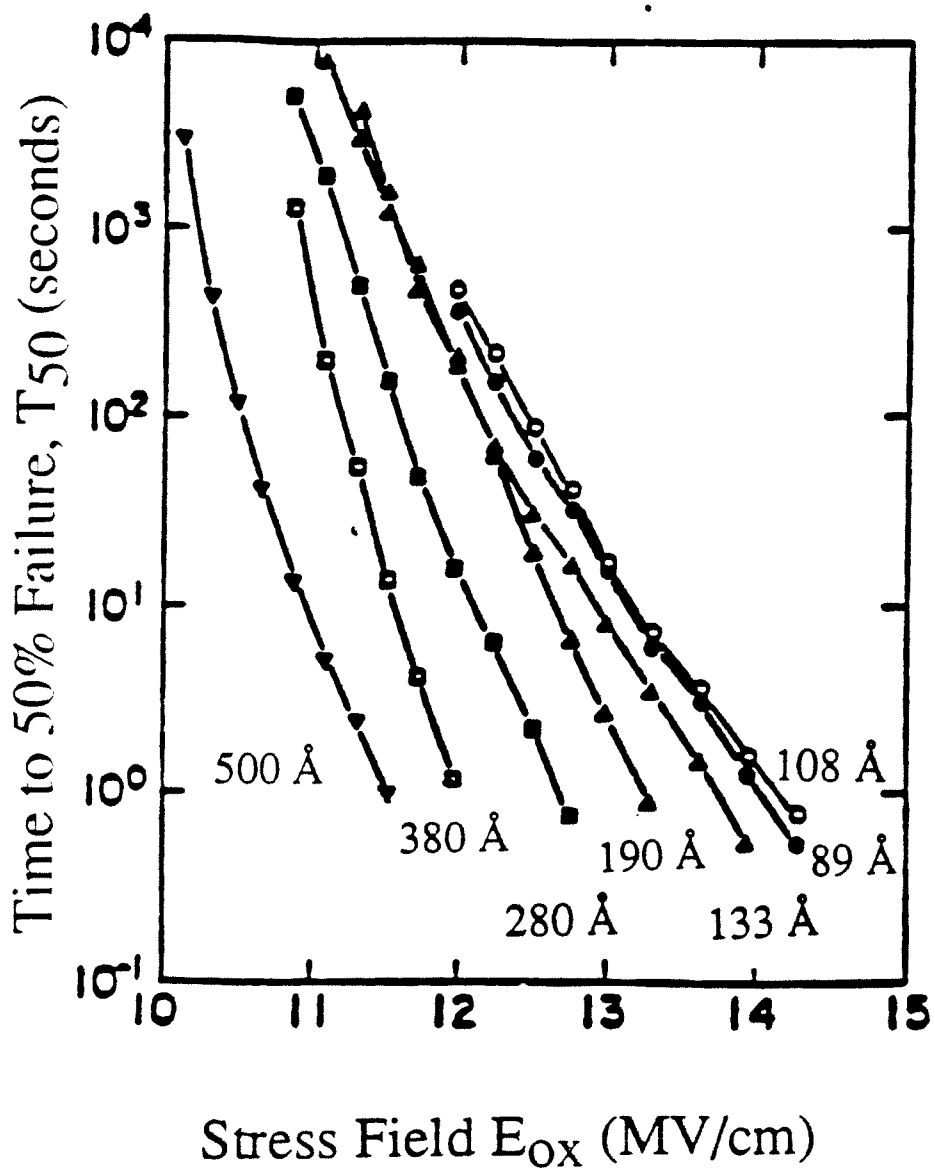


Fig. 1.2 Empirical linear extrapolation of breakdown lifetime with respect to oxide field[1.7].

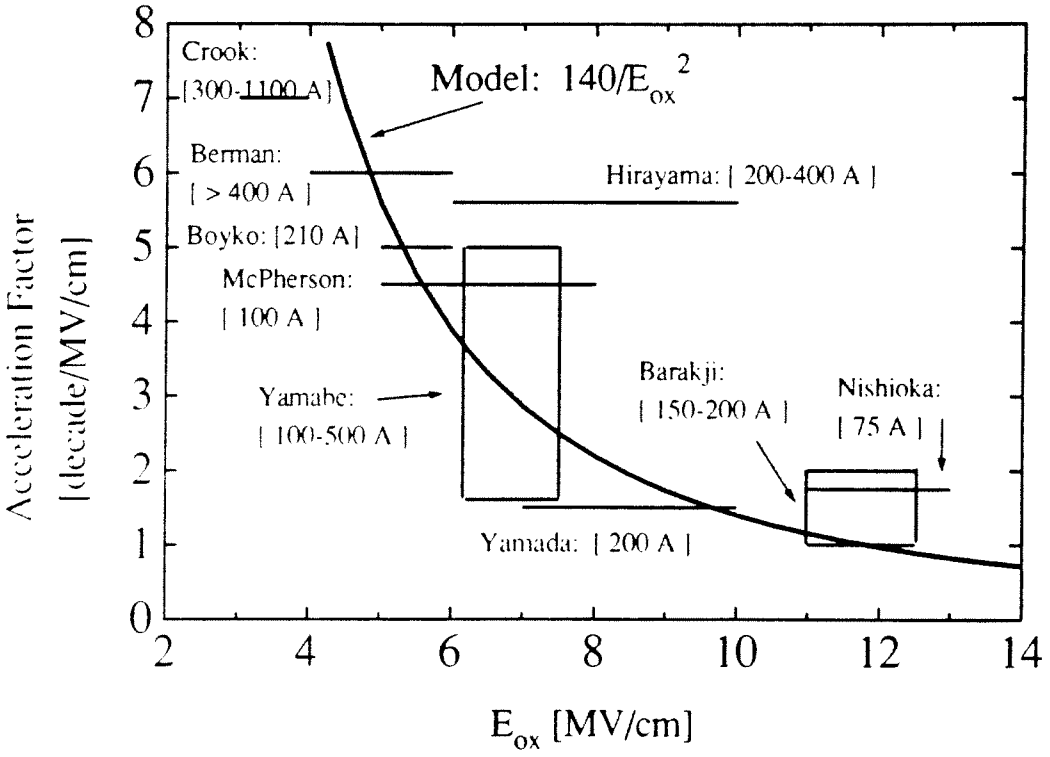


Fig. 1.3 Theoretical field acceleration β compared with reported results [1.8].

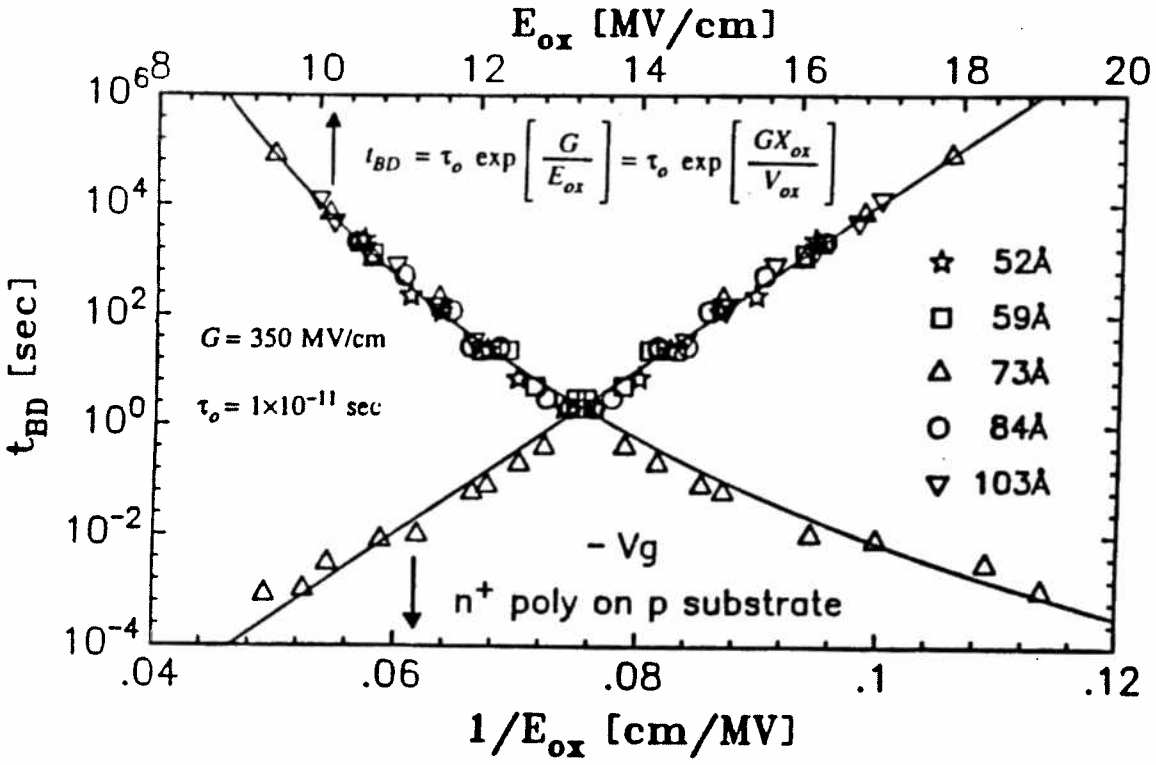


Fig. 1.4 Extrapolation of breakdown lifetime with respect to inverse oxide field consistent with 1/E hole-induced breakdown model [1.13].

1.5. References

[1.1] C. Hu, "Future CMOS Scaling and Reliability," *Proceedings of the IEEE*, vol. 81, no. 5, p. 682, 1993.

[1.3] P.K. Ko, "Approaches to Scaling," in VLSI Electronics Microstructure Science: Advanced MOS Device Physics, edited by N. G. Einspruch and G. Gildenblat, vol. 18, p. 3, 1989.

[1.3] R.H. Dennard, F.H. Gaensslen, H.N. Yu, V.L. Rideout, E. Bassous and A.R. LeBlanc, "Design of Ion-Implanted MOSFETs with Very Small Physcial Dimensions," *Proceedings of the IEEE*, vol. 81, no. 5, p. 682, 1993.

[1.4] *New York Times Business Day*, January 18, 1994.

[1.5] K.F. Schuegraf and C. Hu, "Reliability in Thin SiO_2," to appear in *Semiconductor Science and Technology*, May, 1994.

[1.6] A. Berman, "Time-Zero Dielectric Reliability Test by a Ramp Method," *International Reliability Physics Symposium*, p. 204, 1981.

[1.7] Y. Ozawa and K. Yamabe, "Film Thickness Dependence of TDDB Characteristics for Thermal Oxide," *Extended Abstracts of the 1991 International Conference on Solid State Devices and Materials*, p.240, 1991.

[1.8] I.-C. Chen and C. Hu, "Accelerated Testing of Time-Dependent Breakdown of SiO_2," *IEEE Electron Device Letters*, vol. 8, no. 4, p.140, 1987.

[1.9] I.-C. Chen, S. E. Holland and C. Hu, "Oxide Breakdown Dependence on Thickness and Hole Current - Enhanced Reliability of Ultra-Thin Oxides," *International Electron Devices Meeting*, p. 660, 1986.

[1.10] I.-C. Chen, J. Y. Choi, T. Y. Chan, and C. Hu, "The Effect of Channel Hot Carrier Stressing on Gate-Oxide Integrity in MOSFETs," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2253, 1988.

[1.11] K. R. Mistry, D. B. Krakauer, and B. S. Doyle, "Impact of Snapback-Induced Hole Injection on Gate Oxide Reliability of N-MOSFETs," *IEEE Electron Device Letters*, vol. 11, no. 10, p. 460, 1990.

[1.12] M. Knoll, D. Braunig, and W. R. Fahrner, "Comparative Studies of Tunnel Injection and Irradiation on Metal Oxide Semiconductor Structure," *Journal of Applied Physics*, vol. 53, no. 10, p. 6946, 1982.

[1.13] R. Moazzami, J. Lee, I.-C. Chen, and C. Hu, "Projecting the Minimum Acceptable Oxide Thickness for Time-Dependent Breakdown," *International Electron Devices Meeting*, p. 710, 1988.

[1.14] Y. Nissan-Cohen, J. Shappir, and D. Frohman-Bentchkowsky, "Measurement of Fowler-Nordheim Tunneling Current in MOS Structures Under Charge Trapping Conditions," *Solid State Electronics*, vol. 28, no. 7, p. 717, 1985.

Chapter 2

Intrinsic Low Voltage Breakdown Model

2.1. Introduction

This chapter focuses on the intrinsic SiO_2 low-voltage breakdown characteristics, namely, the field and voltage dependence of charge to breakdown, Q_{bd}, and breakdown lifetime, t_{bd}, thereby refining the empirical hole-induced 1/E breakdown model [2.1] by providing improved theoretical basis, that is, anode hole injection. Once again, the term intrinsic breakdown refers to the breakdown of an oxide that is free of defects, thereby reflecting the practical upper limit of oxide reliability. In practice, intrinsic breakdown is studied by examining oxides having small enough areas such that the breakdown reliability is independent of the area. The intrinsic breakdown reported here corresponds to less than 400 μm^2 capacitor and transistor test structures. Previous research [2.2,2.3,2.4,2.5] has demonstrated the interaction of holes with the oxide in initiating the damage which leads to catastrophic oxide failure. This chapter proposes a quantitative model for the physical process responsible for hole generation and injection into silicon dioxide during high-field electrical stress. The strength of this model lies in its ability to predict oxide breakdown at the very low operating voltages which are being introduced in state-of-the-art VLSI circuits. This chapter also discusses a model for the enhanced low-voltage leakage current due to direct tunneling in ultra-thin oxides.

2.2. Experiment

The devices used in this study were either capacitors or transistors with gate oxide thickness varying between 25 Å and 130 Å. Oxidation process was dry oxidation at 750-950°C. All devices have an in-situ phosphorus doped polysilicon gate deposited at 605°C. The capacitors were on n-type <100> 8-12 Ω-cm substrate. Constant-voltage stressing of capacitors was performed using an HP4140B picoammeter. Constant-voltage hole separation experiments[2.5,2.6] with transistors were performed using an HP4145B parameter analyzer. All the instruments were controlled by a PC using GPIB interface.

2.3. Anode Hole Injection Model

This model proposes that breakdown is a two-stage process [2.1], divided between the time where the oxide is slowly damaged under electrical stress and a much shorter, rapid runaway process, where a rapid final acceleration of damage due to electrical and/or thermal run-away leads to the formation of a permanent conductive path through the oxide. Since the first process dictates the breakdown time, our model addresses the period of accumulation of damage. Recent studies [2.2,2.3,2.4,2.5] clearly show that hole transport through the oxide precedes breakdown, indicating that holes cause damage to the oxide. This conclusion is not at all surprising when one realizes that holes have been found to generate bulk oxide electron traps [2.7,2.8,2.9,2.10] and interface traps [2.11]. Therefore all models of oxide breakdown variably emphasizing the roles of electron trapping [2.12], interface trapping [2.11], hole trapping [2.1] or resonant tunneling through electron traps [2.13] can all be reconciled by recognizing the role of hole injection. This model addresses the generation of holes for an oxide biased into the high-field tunneling regime as shown in

Figure 2.1, representing the conditions of accelerated test for determining oxide reliability. A fraction of the tunneling electrons reaching the anode are able to elastically transfer their entire energy to a deep valence-band electron [2.14,2.15,2.16,2.17]. Such an electron is promoted to the lowest available electron energy state, that is, the conduction band edge of the anode, thereby creating a "hot" hole, which tunnels back into the oxide. These injected holes act to increase the current density (at localized spots), probably through hole-induced trap generation [2.7,2.8,2.9,2.10] (see Chapter 7), until the final runaway process leads to catastrophic breakdown.

Mathematically, the hole tunneling current is given as $J_p = \alpha_p J_n \Theta_p$, where J_n is the incident electron tunneling current, $\alpha_p \Theta_p$ is the probability for a hole to be generated and to tunnel through the barrier. The quantum efficiency of the hole generation process is

$$\frac{J_p}{J_n} = \alpha_p \Theta_p = \alpha_p \exp\left(-\frac{\hat{B}}{E_{ox}} \left[\Phi_p(V_{ox}) \right]^{\frac{3}{2}} \right) \qquad (2.1)$$

where $\hat{B} = 8\pi\sqrt{2m_{p,ox}} / 3hq$, (h is Planck's constant), $m_{p,ox} = 0.2\ m_o$ and [2.6]

$$q\Phi_p = E_{g,SiO_2} - q\Phi_b - E_{gain} \qquad (2.2)$$

The energy gained from the oxide field before arrival at the anode in the Fowler-Nordheim tunneling regime, where $V_{ox} > \Phi_b$, calculated by means of a phenomenological energy relaxation model [2.18], is given as

$$E_{gain} = \Phi_b + (E_{ox}\lambda)\left[1 - \exp\left(\frac{-1}{\lambda}\left[X_{ox} - \frac{\Phi_b}{E_{ox}} \right] \right) \right] \qquad (2.3)$$

where $\lambda = 15$ Å is the mean free electron scattering length in the oxide conduction band [2.19]. In contrast, in direct tunneling, i.e. for $V_{ox} < \Phi_b$, electrons do not experience such scattering, thus the arrival energy of electrons at the anode is simply

$$E_{gain} = V_{ox} \text{ for } V_{ox} < \Phi_b . \tag{2.4}$$

Using **nMOS** transistors biased as in Figure 2.2, the tunneling electron current can be easily measured as gate current and the anode-injected hole current can be measured as the substrate current [2.5,2.6]. Figure 2.3 shows that the tunnel current density J_n decreases by over an order of magnitude during the oxide lifetime due to electron trapping. Moreover, as more clearly illustrated in Figure 2.4, the quantum hole generation efficiency remains constant for the duration of the oxide lifetime showing that the hole generation rate is strictly determined by the applied bias, consistent with Eq. (2.1) for $\alpha_p \cong 0.08$.

The injected hole quantity increases with time as $Q_p(t) = \int_0^t J_p(\hat{t})\, d\hat{t}$ until a critical hole fluence, Q_p, is reached, marking the breakdown event. The charge to breakdown, Q_{bd}, follows as

$$Q_{bd} = \frac{Q_p}{\alpha_p} \exp\left(\frac{\hat{B}}{E_{ox}} \left[\Phi_p(V_{ox}) \right]^{\frac{3}{2}} \right) \tag{2.5}$$

Figure 2.5 shows that the critical hole fluence at breakdown, Q_p, is independent of the stress voltage, while Q_{bd} decreases with increasing stress voltage according to Eq. (2.5). Figure 2.6 shows the excellent ability of Eq. (2.5) to predict intrinsic charge to breakdown for oxide thickness varying between 25 and 100 Å. The low voltage predictive ability of Eq. (2.5) attests to the hypothesis of anode hole injection since tunneling electrons cannot gain the energy necessary for impact ionization in the silicon dioxide at these low voltages. The rapidly rising Q_{bd} behavior in thinner oxides can be attributed to the fact that the hot hole energy, E_{gain}, becomes more sensitive to V_{ox} when scattering becomes weaker. The formalism leading to Eq. (2.5) may also be used in modeling thin oxide Q_{bd} derived from constant current stressing, by performing a

path integral between the initial and final stress voltages needed to maintain the constant current condition. Figure 2.7 compares values of Q_p measured using hole separation and those deduced fitting capacitor breakdown data to Eq. (2.5), showing that Q_p decreases with decreasing oxide thickness, reflecting a weakened hole immunity for thinner oxides.

2.4. Oxide Tunneling Current

This section discusses the calculation of oxide voltage from the applied voltage and introduces a model for tunnel current, J_n, which facilitates the modeling of the breakdown lifetime, t_{bd}. First, the necessity of accounting for polysilicon depletion in thin-oxide current conduction is demonstrated. Next, quantitative models for Fowler-Nordheim and direct tunneling are introduced.

2.4.1. Polysilicon Depletion

In order to model the leakage currents of thin oxides, it is necessary to relate the oxide voltage to the applied bias. Figure 2.8 compares the measured positive bias (substrate electron emission: SE) and negative bias (gate electron emission: GE) tunneling currents for a thin oxide capacitor. Whereas thicker oxides display parallel tunneling characteristics with an offset due to band-bending [2.20], these data show that the measured JV characteristics in very thin oxides are not parallel, which can be attributed to the depletion of the heavily in-situ doped n+ gate polysilicon [2.21,2.22,2.23]. Computation of the oxide voltage for each curve by subtracting for band bending in the poly-gate and the substrate reduces them to a single tunneling characteristic dependent only on V_{ox}, consistent with the notion that tunneling should be a unique function of the oxide voltage and thickness, independent of polarity. The polysilicon band

bending induces a voltage drop correcting the GE and SE curves under the transformations given in Table 2.1 [2.20,2.23] where $V_{poly} = \frac{\varepsilon_{ox}^2 E_{ox}^2}{2q\varepsilon_{Si}N_{poly}}$ until V_{poly} is pinned at 1.12 V due to the saturation of band bending in strong inversion. The maximum electrically active doping concentration appears to be bounded by approximately 5 x 10^{19} cm^{-3} for this in-situ doped polysilicon, with even lower dopant activation for implanted polysilicon [2.21,2.23]. The single tunneling characteristic dependent only on V_{ox} attests to the applicability of the depletion approximation [2.22]. Figure 2.9 shows that Eq. (2.5) is valid for predicting Q_{bd} for either positive or negative bias, providing the oxide voltage is calculated consistently with Table 2.1. However, Q_p for the gate emission case (negative gate voltage) was found to be 2 to 5 times lower than its substrate emission (positive voltage) counterpart. This indicates that the poly-SiO_2 interface is not as robust as the Si-SiO_2 interface.

2.4.2. Thin Oxide Conduction Model

This section introduces a model for quantitative modeling of the tunnel current, J_n, of Figure 2.8. Figure 2.10 illustrates the difference between direct and Fowler-Nordheim (FN) tunneling. The standard FN expression [2.24] is

$$J_n = A E_{ox}^2 e^{-\frac{B}{E_{ox}}}, \qquad (2.6)$$

where $B = \frac{8\pi\sqrt{2m_{ox}}\Phi_b^{3/2}}{3hq}$, and the electron effective mass, m_{ox}, is 0.5 m_o. Equation (2.6) represents tunneling through the triangular potential barrier of Figure 2.10(a), valid for $V_{ox} > \Phi_b$. However, for $V_{ox} < \Phi_b$, where the tunneling barrier is trapezoidal as in Figure 2.10(b), Eq. (2.6) is no longer valid but becomes [2.25],

$$J_n = A E_{ox}^2 \left(\frac{\Phi_b}{V_{ox}}\right)\left(\frac{2\Phi_b}{V_{ox}} - 1\right) \exp\left[\frac{-B\left(1-\left(1-\Phi_b^{3/2}\right)\right)}{E_{ox}}\right] \quad . (2.7)$$

This analytical formula does not approach zero as V_{ox} approaches zero. Therefore, it does not apply to the regime of $V_{ox} < 1$ Volt. An alternative approximation that overcomes this limitation is

$$J_n = A E_{ox}^2 \exp\left[\frac{-B\left(1-\left(1-\Phi_b^{3/2}\right)\right)}{E_{ox}}\right] . \quad (2.8)$$

Figure 2.11 demonstrates the utility of this model, highlighting the dramatic effect the changed barrier shape, from triangular to trapezoidal, exerts on increasing the leakage current for $V_{ox} < \Phi_b$. Observation of direct tunneling is generally limited to oxides thinner than 50 Å because the tunneling probability for thicker oxides is small. Thus, experimental constraints (i.e. current sensitivity of measurement equipment and capacitor area) limit observation of direct tunneling currents.

2.5. Breakdown Lifetime

Combining this closed form current model of Eq. (2.6) and (2.7) with the Q_{bd} model expression of Eq. (2.5) leads to a simple model to predict oxide lifetime, i.e. $t_{bd}=Q_{bd}/J_n$. Although carrier trapping can lead J_n to increase or decrease with stress time, this trapping is negligible for thin oxides so that $t_{bd} \approx Q_{bd} / J_n$. Figure 2.12 demonstrates exceptional agreement between t_{BD} theory and intrinsic breakdown data for thin oxide samples, highlighting the ability of the anode hole injection model to predict oxide lifetime at low operating voltages. Whereas the hole-induced 1/E extrapolation model predicts a constant extrapolation slope of 350 MV/cm [2.1,2.26], Figure 2.13 illustrates that $\log(t_{bd})$ may still be linearly extrapolated with respect to

1/E, albeit with increasing extrapolation slope for decreasing oxide thickness. Figure 2.14 shows this lifetime extrapolation slope follows

$$G = 400 + 190 \exp\left(-\frac{X_{ox} - 39}{\lambda}\right) \qquad \left[\frac{MV}{cm}\right] \tag{2.9}$$

for oxides thickness greater than 39Å, with λ = 15 Å. For thinner oxides, the extrapolation slope is about 650 MV/cm. Despite this increase in extrapolation slope, the maximum acceptable oxide field for 10 year lifetime varies by only about 10 % for a wide range of oxide thickness as shown in Figure 2.13.

2.6. Supply Voltage Limits

Figure 2.15 examines the role of two competing criteria, intrinsic TDDB (10 Year Lifetime at 25° C) and leakage current (0.1 pA/μm^2), in determining maximum acceptable operating voltage. Although these two criteria track each other in scaling from 130 Å to 60 Å, further scaling leads these criteria to diverge. The enhanced leakage current due to direct tunneling may require a rapid derating of supply voltage, posing an additional constraint on aggressively scaled oxide technologies below 40 Å. Different leakage requirements may relax the constraint on minimum oxide thickness; however, the rapid increase in direct tunneling current below 40 Å requires rapid supply voltage derating regardless of how much leakage current is permissible. While intrinsic breakdown sets the ultimate reliability limit, in practice, certain thickness margin is needed to cushion against defects whose breakdown is modeled by an effectively thinnest spot [2.27,2.28]. In practice, the ability to manufacture high quality, low defect ultra-thin oxides will limit the usable oxide thickness for a given operating voltage.

2.7. Summary and Conclusions

This chapter presented a new quantitative model for silicon dioxide breakdown, based on the concept of anode hole injection, valid for predicting insulator reliability performance for thicknesses between 25 and 130 Å. This model is suitable for predicting dielectric lifetime for reduced supply voltages and aggressively scaled oxide thicknesses. Additionally, a quantitative model for enhanced low voltage silicon dioxide leakage current due to the direct tunneling mechanism shows this current may pose additional oxide scaling constraints for oxides thinner than 40 Å. This criterion along with the manufacturability of defect-free ultra-thin silicon dioxide will ultimately determine the usable oxide thickness and supply voltage -- and hence the CMOS speed performance.

$\mid V_{ox} \mid = \mid V_g \mid$ - the following:		
	n+ poly on p sub	n+ poly on n sub
$+V_g$	V_{poly}	V_{poly}
$-V_g$	1.2	1.2

Table 2.1 J-V Experiments show how V_{ox} may be determined for different bias polarity and gate and substrate doping types. The 1.2 Volt offset for negative biases is due to band bending in the MOS system; it reflects an average across a variety of oxide thicknesses and technologies [2.22].

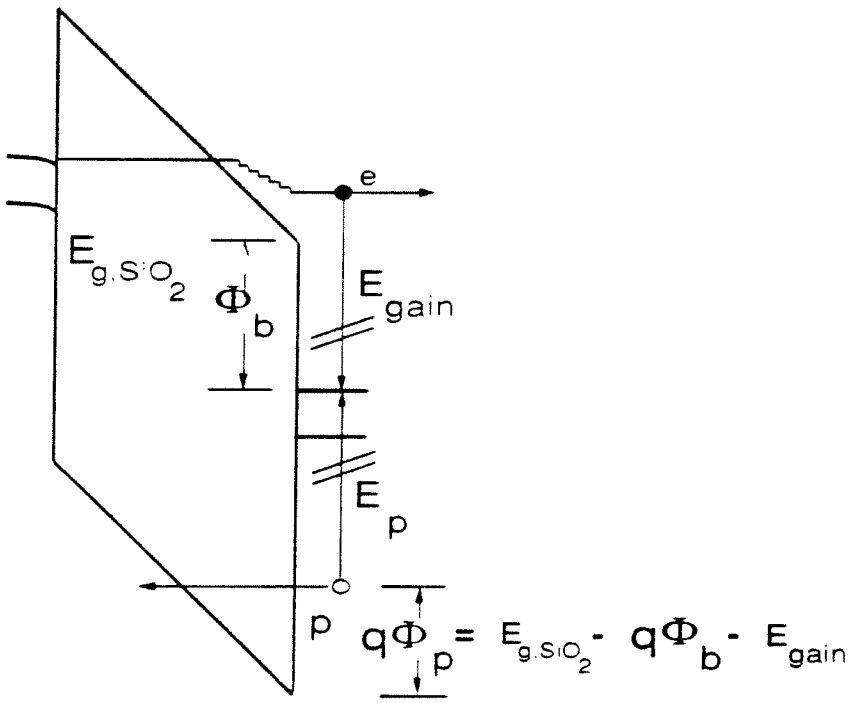


Fig. 2.1 Diagram of Anode Hole Injection Process. An incident tunneling electron arrives at the anode with energy, E_{gain}, to thermalize. This energy is transferred to a deep valence band electron, thereby exciting it to the lowest available energy state, the anode conduction band. This excitation creates a "hot" hole capable of tunneling back into the oxide.

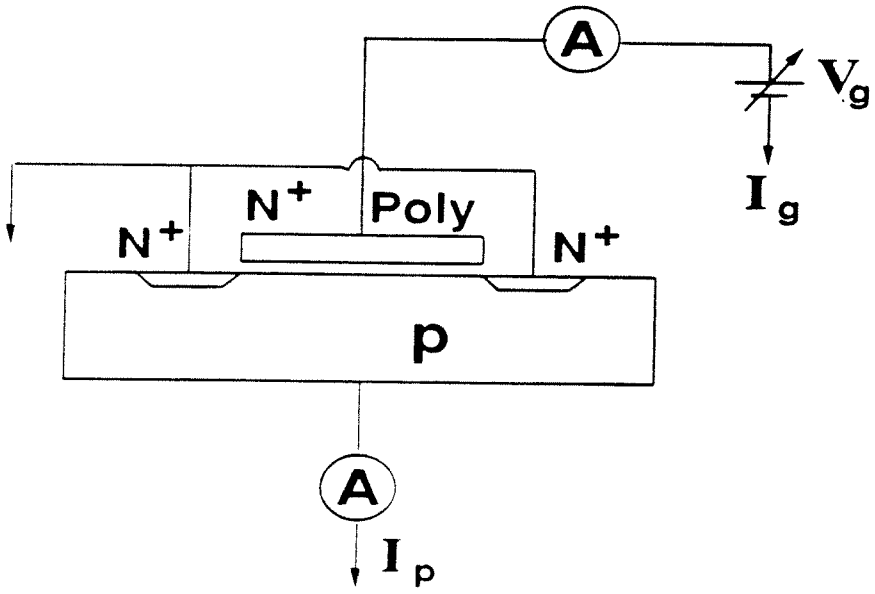


Fig. 2.2 Bias configuration of NMOSFET which enables separate measurement of tunneling electron and injected hole currents. Time integrated electron tunneling current is Q_{bd}, whereas time integrated hole current is Q_p.

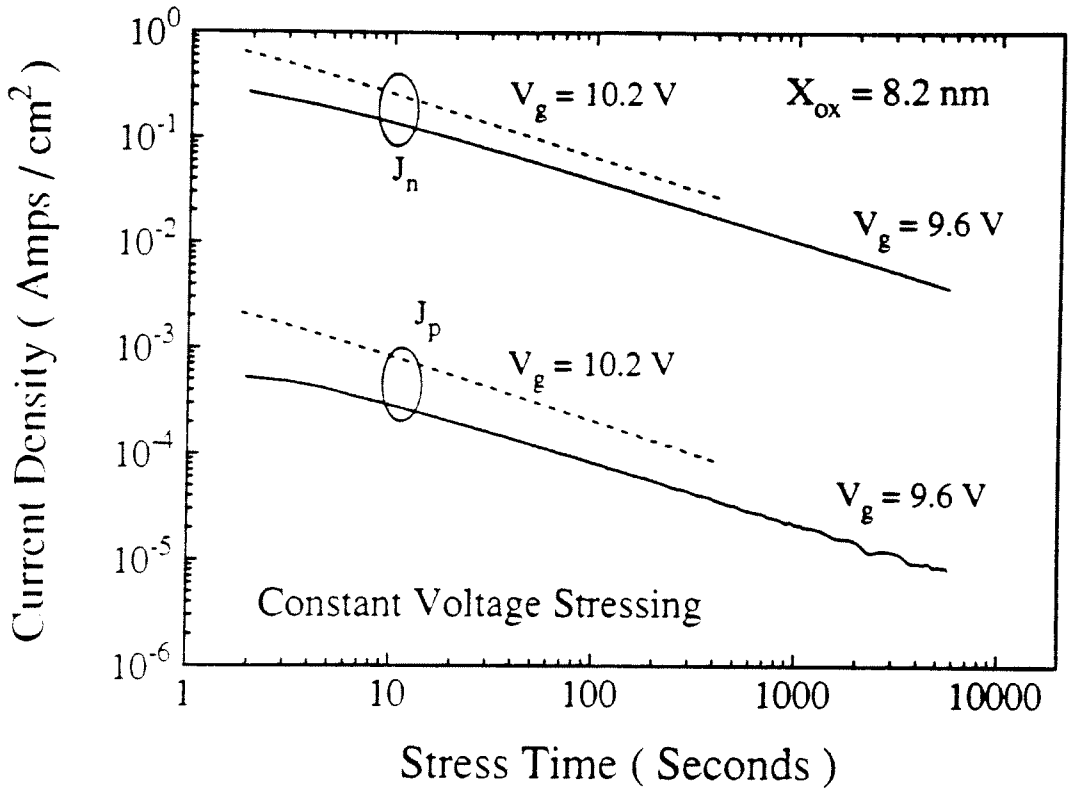


Fig. 2.3 Time evolution of gate current density, J_n, and injected hole density, J_p, during constant voltage stress of a 82 Å oxide, showing that electron trapping decreases the gate and injected hole tunneling currents while the ratio J_p/J_n is determined by the applied bias.

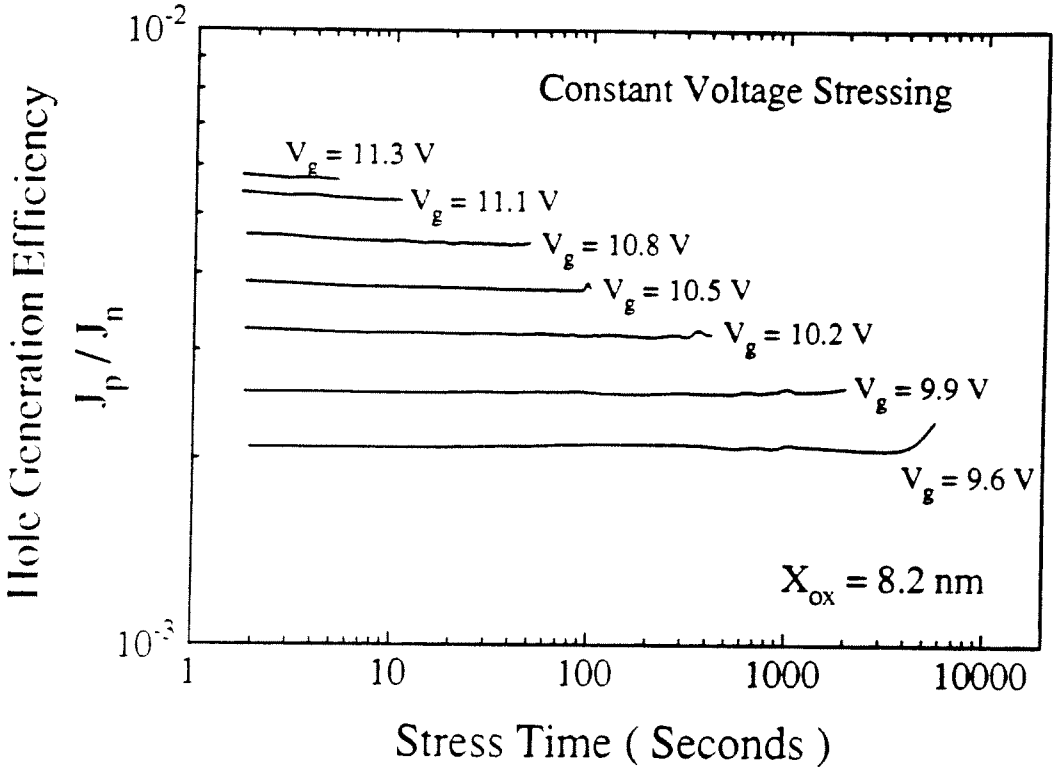


Fig. 2.4 Time evolution of the quantum hole generation efficiency, J_p/J_n, during constant voltage stress of a 82 Å oxide, clearly demonstrating that the applied bias determines this efficiency.

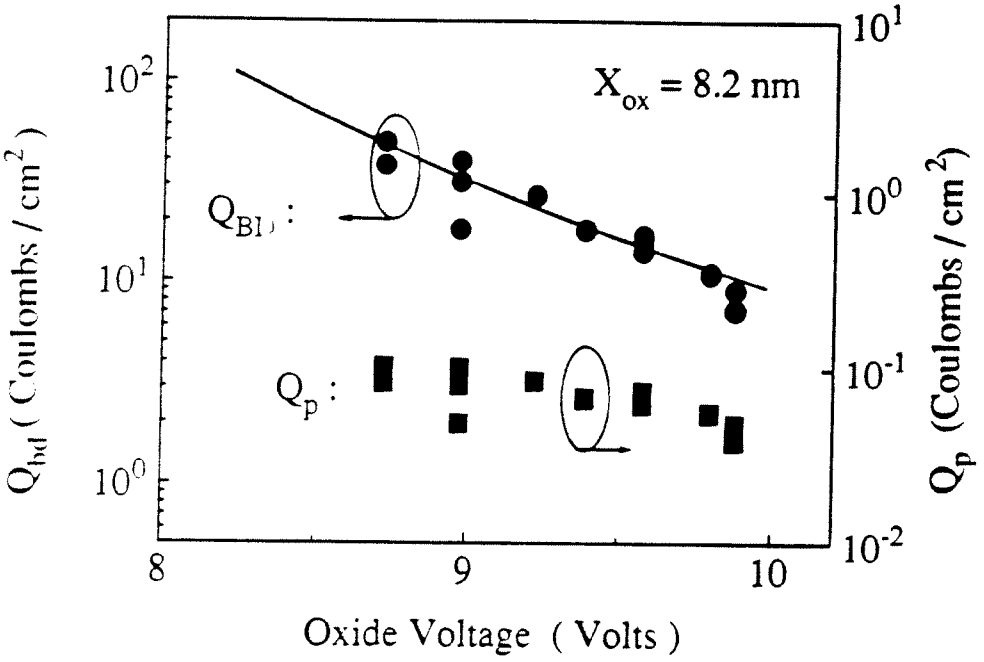


Fig. 2.5 Charge to breakdown, Q_{bd}, and hole fluence to breakdown, Q_p, for a 82 Å oxide, showing that Q_{bd} increases with decreasing stress voltage, while Q_p remains constant.

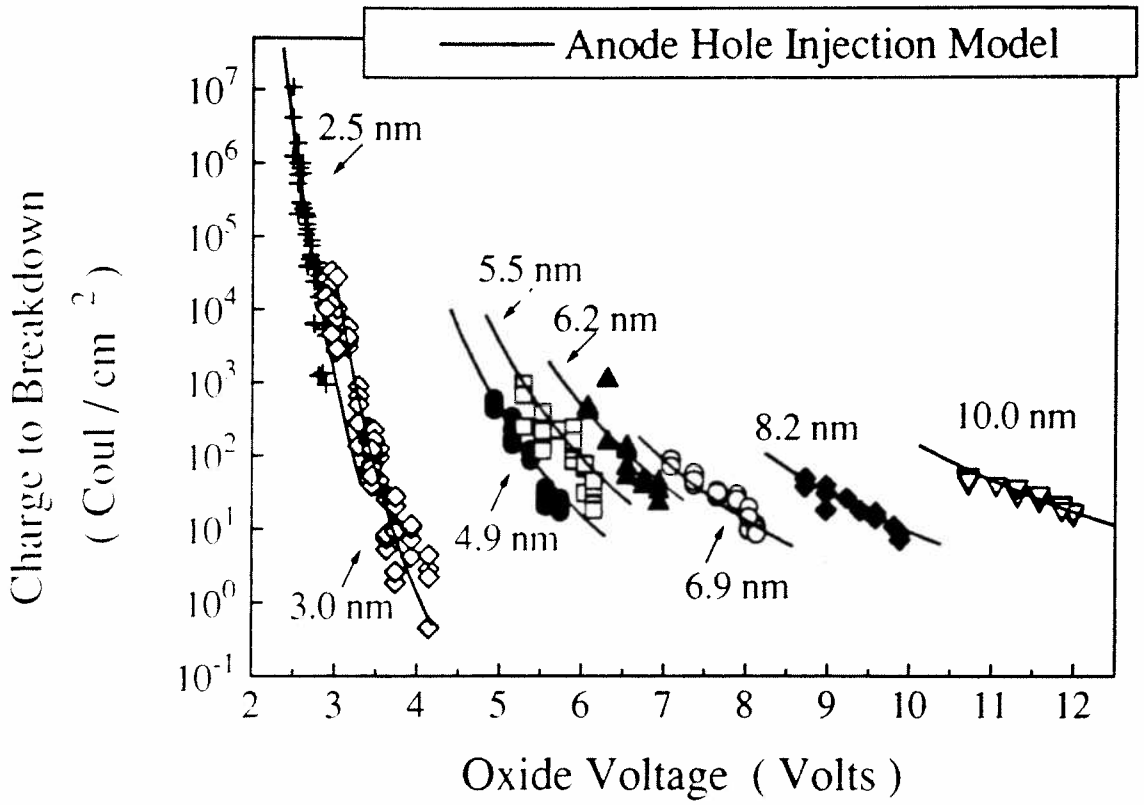


Fig. 2.6 Voltage dependence of thin oxide charge to breakdown, demonstrating ability to extrapolate to low voltages.

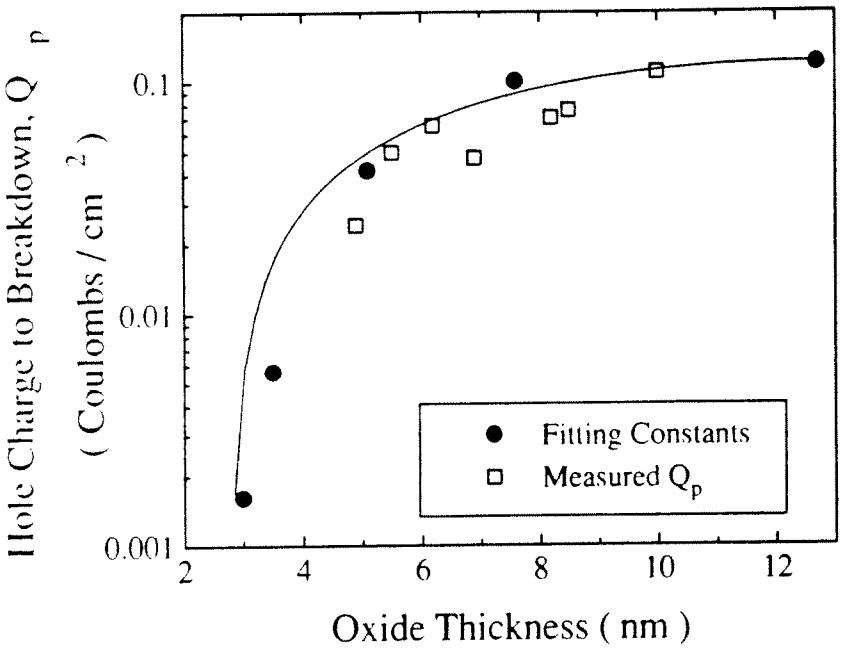


Fig. 2.7 Comparison of experimentally measured Q_p values using hole separation and those deduced from fitting capacitor breakdown data.

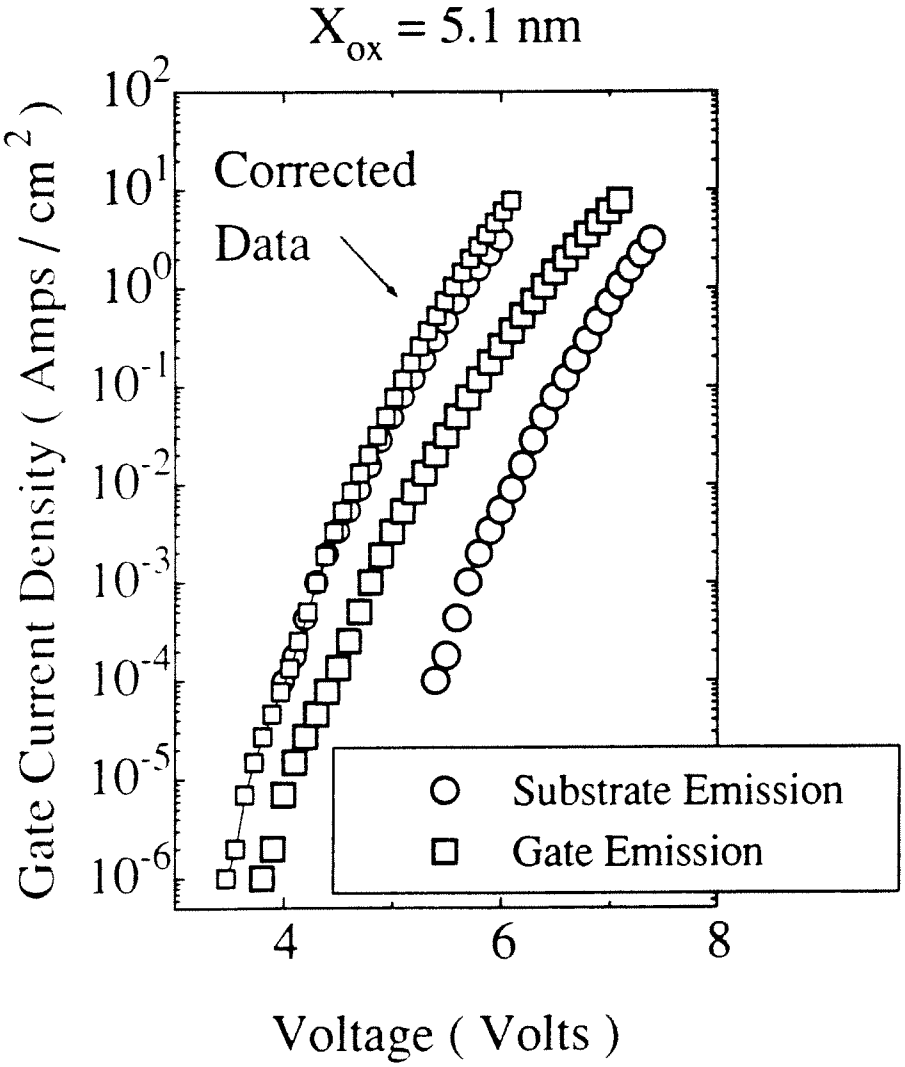


Fig. 2.8 J-V characteristics for 51 Å oxide capacitors with in-situ doped n^+ polysilicon gate. The corrected data for substrate emission and gate emission is computed according to the rules in **Table 2.1**.

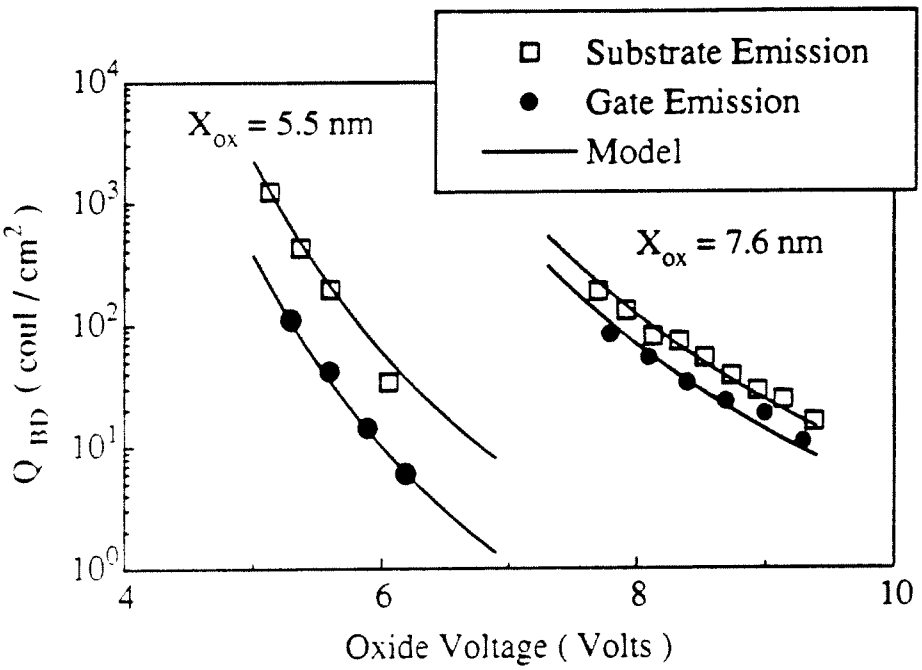


Fig. 2.9 Shows that Anode Hole Injection Model accurately predicts breakdown for positive and negative bias stresses.

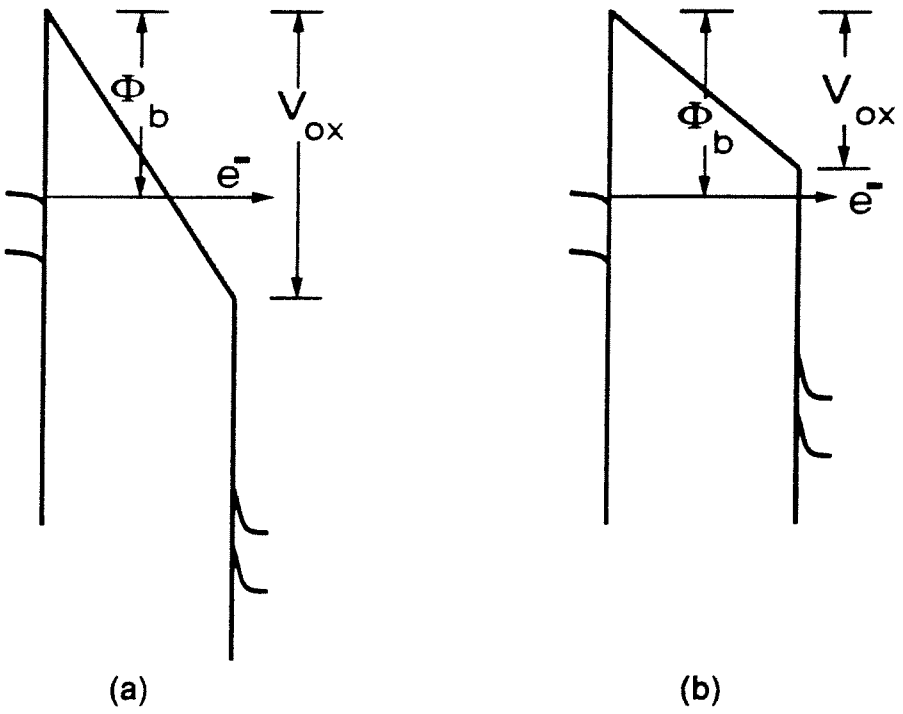


Fig. 2.10 Illustrates the physical difference between Fowler-Nordheim and Direct Tunneling. (a) Fowler-Nordheim Tunneling is associated with transversal of a triangular barrier. (b) Direct Tunneling is associated with transversal of a trapezoidal barrier, i.e. when $V_{ox} < \Phi_b$.

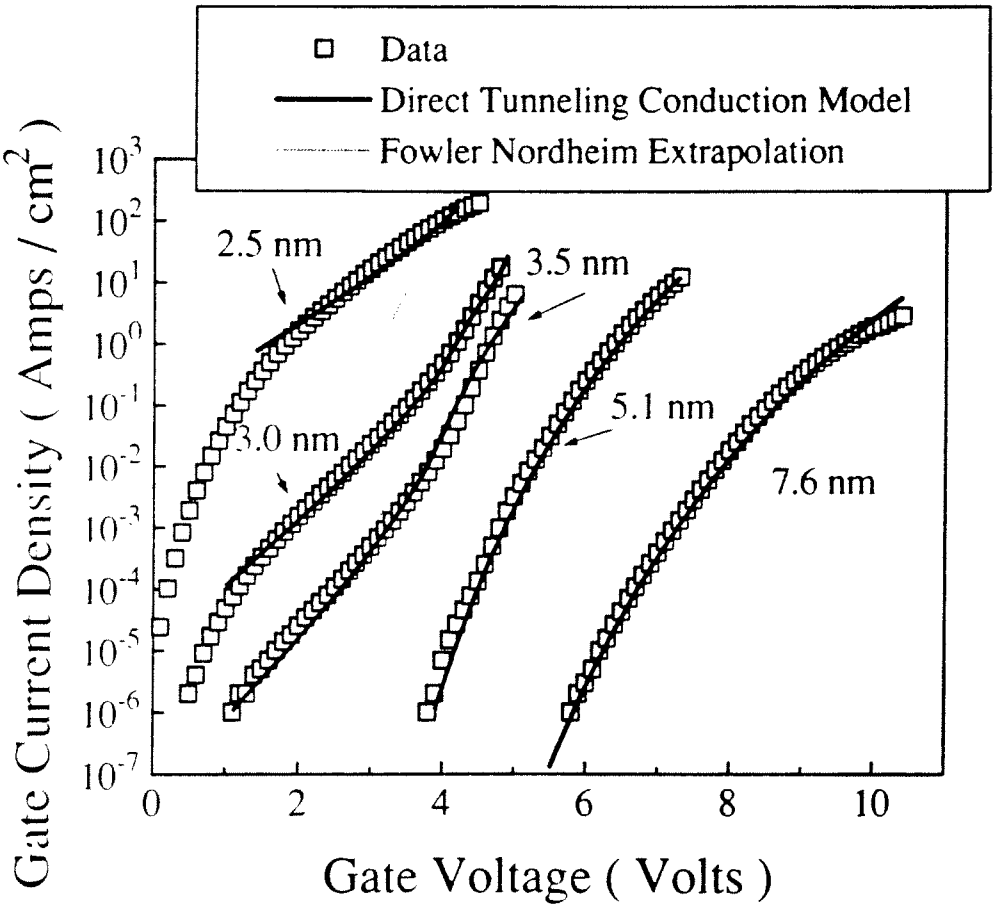


Fig. 2.11 Comparison between data and closed form oxide conduction current model which accounts for both Fowler-Nordheim and Direct Tunneling. The extension of Fowler-Nordheim theory for ultra-thin oxides illustrates the dramatic increase in leakage current caused by the subtle change in barrier shape, i.e. from triangular to trapezoidal, used to model the low voltage Direct Tunneling regime.

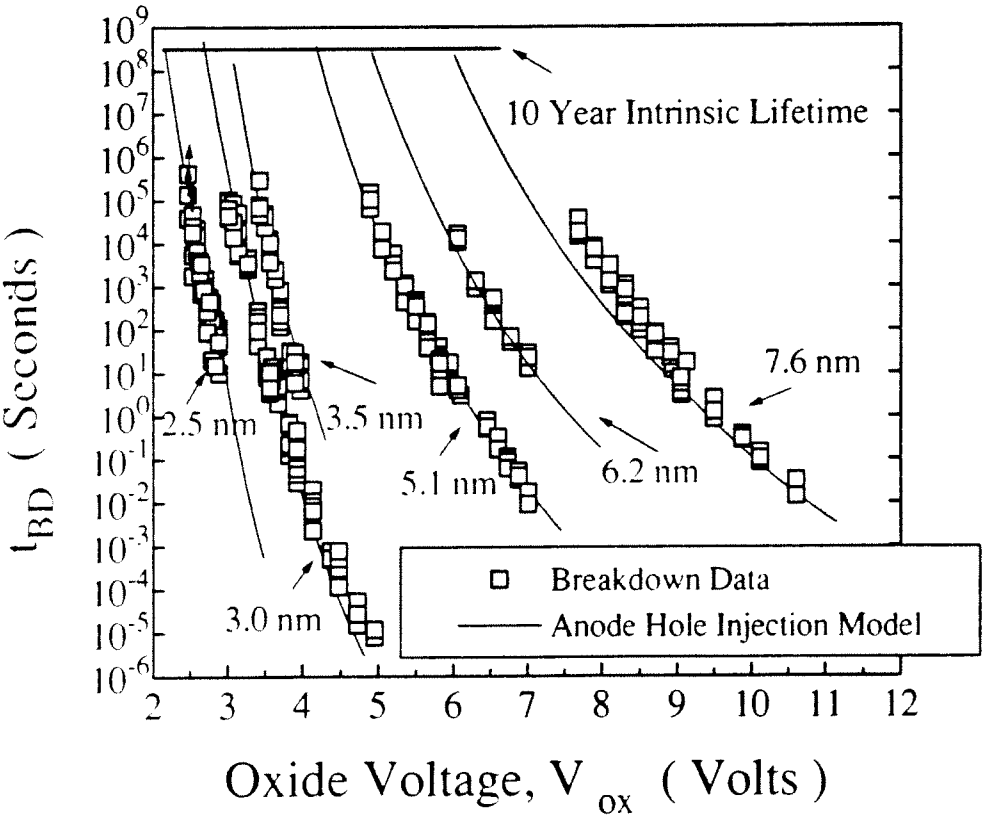


Fig. 2.12 Voltage dependence of breakdown lifetime for thin oxides predicted down to low voltages using Anode Hole Injection Model.

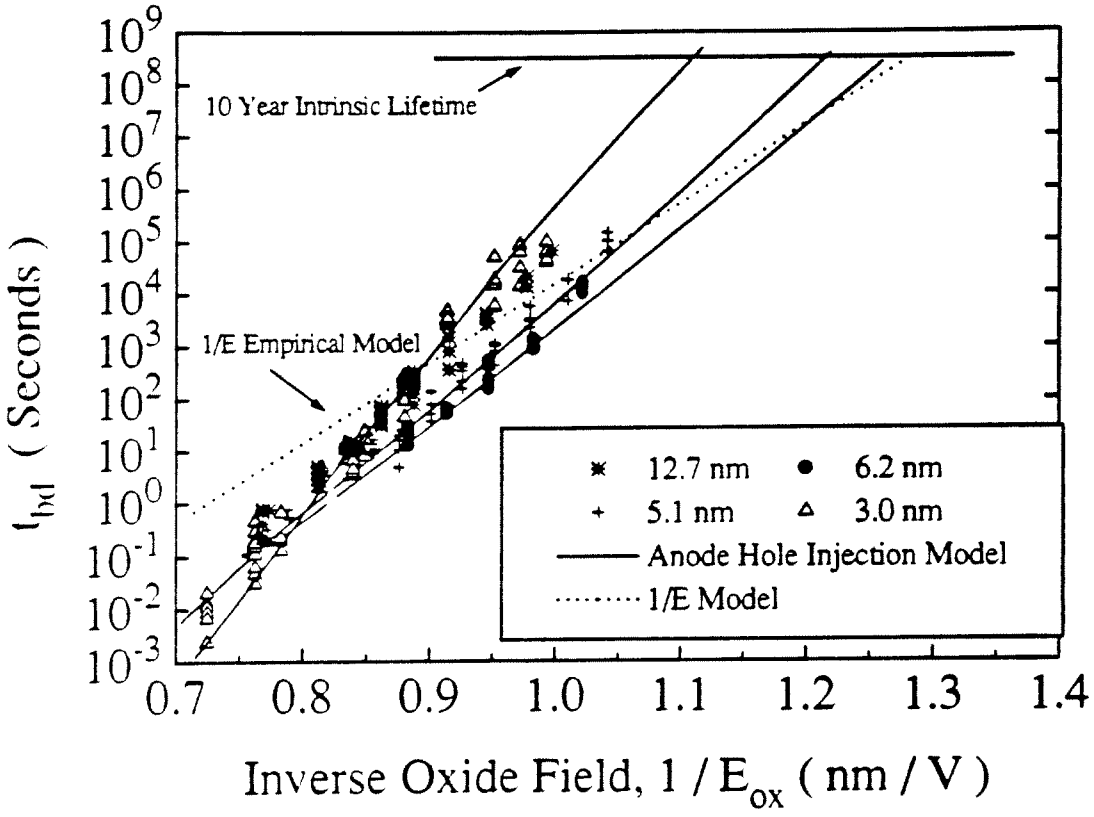


Fig. 2.13 Inverse field dependence of breakdown lifetime is linear with increasing extrapolation slope for thinner oxides.

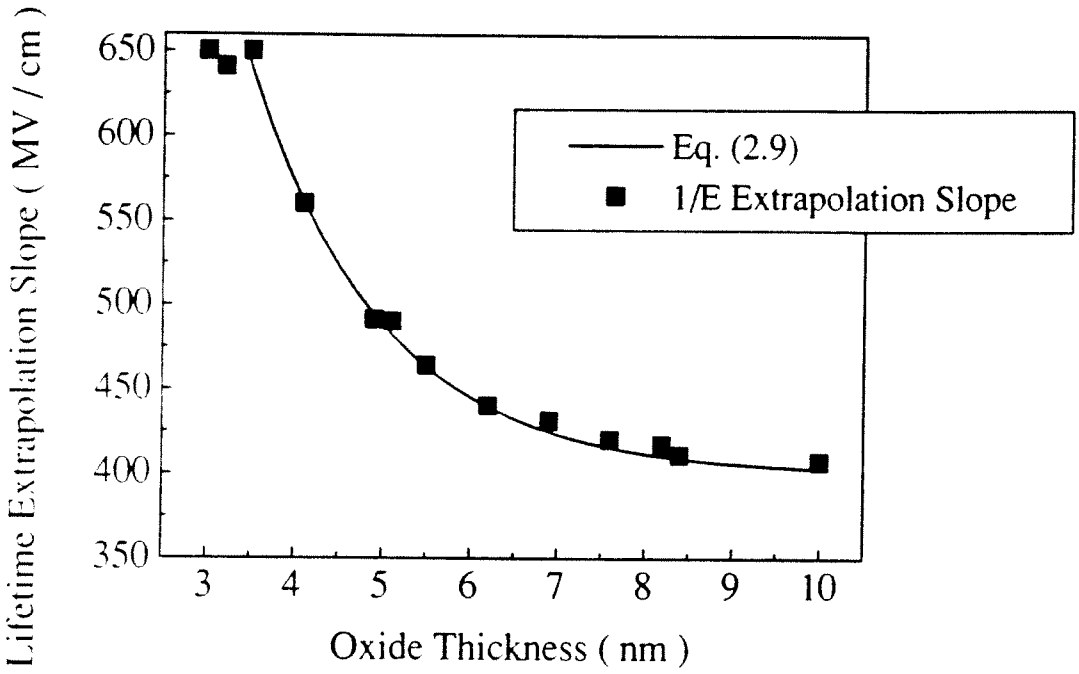


Fig. 2.14 Thickness dependence of inverse field oxide breakdown lifetime slope. Slope increases for thinner oxides due to reduced hole generation efficiency.

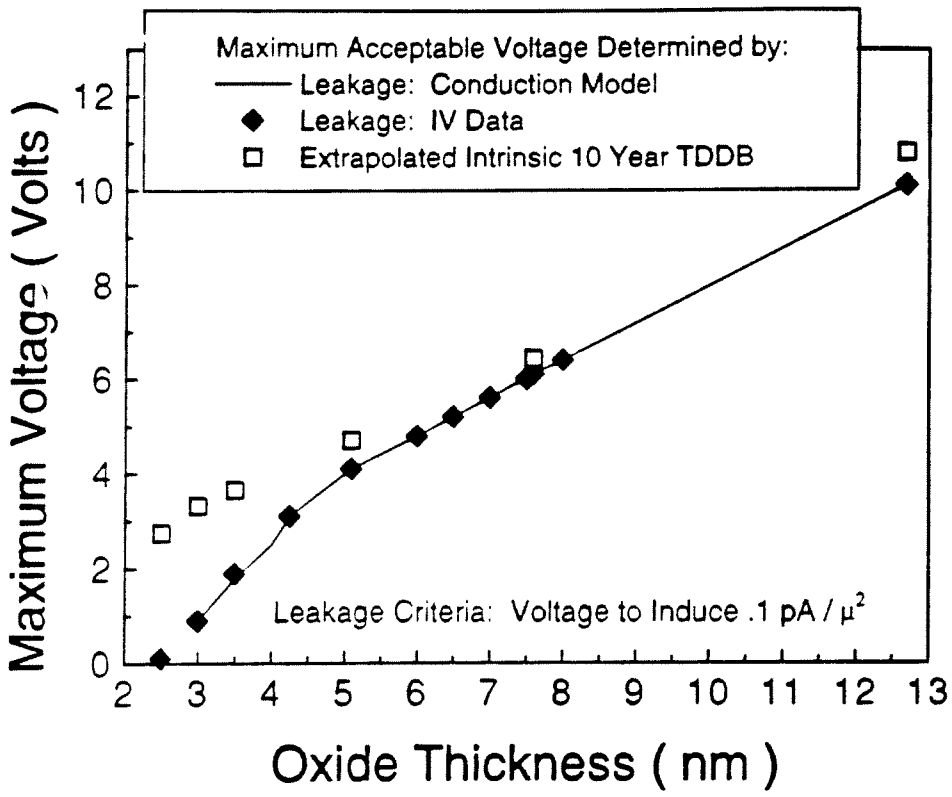


Fig. 2.15 Comparison of maximum acceptable operating voltages from viewpoint of 10 year **intrinsic** TDDB lifetime and Gate Leakage. These two criteria track each other in scaling from 130 Å to 60 Å, but further scaling leads them to diverge, as the enhance leakage current of direct tunneling mandates rapid voltage scaling.

2.8. References

[2.1] I.-C. Chen, S. E. Holland, and C. Hu, "Electrical Breakdown in Thin Gate and Tunneling Oxides," *IEEE Transactions on Electron Devices*, vol. ED-32, no. 2, p. 413, 1985.

[2.2] I.-C. Chen, J. Y. Choi, T. Y. Chan, and C. Hu, "The Effect of Channel Hot Carrier Stressing on Gate-Oxide Integrity in MOSFETs," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2253, 1988.

[2.3] K. R. Mistry, D. B. Krakauer, and B. S. Doyle, "Impact of Snapback-Induced Hole Injection on Gate Oxide Reliability of N-MOSFETs," *IEEE Electron Device Letters*, vol. 11, no. 10, p. 460, 1990.

[2.4] E. Rosenbaum, R. Rofan, and C. Hu, "Effect of Hot Carrier Injection on n- and p-MOSFET Gate Oxide Integrity," *IEEE Electron Device Letters*, vol. 12, no. 11, p. 599, 1991.

[2.5] I.-C. Chen, S. E. Holland and C. Hu, "Oxide Breakdown Dependence on Thickness and Hole Current - Enhanced Reliability of Ultra-Thin Oxides," *International Electron Devices Meeting*, p. 660, 1986.

[2.6] B. Eitan and A. Kolodny, "Two Components of Tunneling Current in Metal-Oxide-Semiconductor Structures," *Applied Physics Letters*, vol. 43, no. 1, p. 106, 1983.

[2.7] S. K. Lai, "Interface Trap Generation in Silicon Dioxide when Electrons are Captured by Trapped Holes," *Journal of Applied Physics*, vol. 54, no. 5, p. 2540, 1983.

[2.8] I.-C. Chen, S. E. Holland, and C. Hu, "Electron-trap Generation by Recombination of Electrons and Holes in SiO_2," *Journal of Applied Physics*, vol. 61, no. 9, p. 4544, 1987.

[2.9] H. Uchida and T. Ajioka, "Electron Trap Center Generation Due to Hole Trapping in SiO_2 Under Fowler-Nordheim Tunneling Stress," *Applied Physics Letters*, vol. 51, no. 6, p. 433, 1987.

[2.10] S. Ogawa, N. Shiono, and M. Shimaya, "Neutral Electron Trap Generation in SiO_2 by Hot Holes," *Applied Physics Letters*, vol. 56, no. 14, p. 1329, 1990.

[2.11] D. J. DiMaria, D. Arnold, and E. Cartier, "Degradation and Breakdown of Silicon Dioxide Films on Silicon," *Applied Physics Letters*, vol. 61, no. 19, p. 2329, 1992.

[2.12] E. Harari, "Dielectric Breakdown in Electrically Stressed Thin Films of Thermal SiO_2," *Journal of Applied Physics*, vol. 49, no. 4, p. 2478, 1978.

[2.13] B. Ricco, M. Ya Azbel, and M. H. Brodsky, "A Novel Mechanism for Tunneling and Breakdown in Thin SiO_2 Films," *Physics Review Letters*, vol. 51, p. 1795, 1983.

[2.14] Z. A. Weinberg, W. C. Johnson, and M. A. Lampert, "High-Field Transport in SiO_2 on Silicon Induced by Corona Charging of the Unmetallized Surface," *Journal of Applied Physics*, vol. 47, no. 1, p. 248, 1976.

[2.15] Z. A. Weinberg, M. V. Fischetti, and Y. Nissan-Cohen, "SiO_2-Induced Substrate Current and its Relation to Positive Charge in Field-Effect Transistors," *Journal of Applied Physics*, vol. 59, no. 3, p. 824, 1986.

[2.16] S. E. Holland, I.-C. Chen, and C. Hu, "Ultra-Thin Silicon-Dioxide Breakdown Characteristics of MOS Devices with n^+ and p^+ Polysilicon Gates," *IEEE Electron Device Letters*, vol. 8, no. 12, 1987.

[2.17] K. F. Schuegraf and C. Hu, "Hole Injection Oxide Breakdown Model for Very Low Voltage Lifetime Extrapolation," *International Reliability Physics Symposium*, p. 7, 1993.

[2.18] C. Chang, C. Hu, and R. W. Brodersen, "Quantum Yield of Electron Impact Ionization in Silicon," *Journal of Applied Physics*, vol. 57, no. 2, p. 302, 1985.

[2.19] G. Lewicki and J. Maserjian, "Tunneling in MOS Structures," *Journal of Applied Physics*, vol. 46, no. 7, p. 3032, 1975.

[2.20] E. Rosenbaum, R. Moazzami and C. Hu, "Implications of Waveform and Thickness Dependence of SiO_2 Breakdown on Accelerated Testing," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications*, p. 214, 1991.

[2.21] B. J. Fishbein and D. B. Jackson, "Performance Degradation of N-Channel MOS Transistors During DC and Pulsed Fowler-Nordheim Stress, *Proceedings of the International Reliability Physics Symposium*, p. 159, 1990.

[2.22] S. J. Wang, I.-C. Chen, and H. L. Tigelaar, "Effects of Poly Depletion on the Estimate of Thin Dielectric Lifetime", *IEEE Electron Device Letters*, vol. 12, no. 11, p. 617, 1991.

[2.23] K. F. Schuegraf, C. C. King and C. Hu, "Impact of Polysilicon Depletion in Thin Oxide MOS Technology," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications*, p. 86, 1993.

[2.24] M. Lenzlinger and E. H. Snow, "Fowler-Nordheim Tunneling into Thermally Grown SiO_2," *Journal of Applied Physics*, vol. 40, p. 278, 1969.

[2.25] K. F. Schuegraf, C. C. King and C. Hu, "Ultra-Thin Silicon Dioxide Leakage Current and Scaling Limit," *Symposium on VLSI Technology Digest of Technical Papers*, p. 18, 1992.

[2.26] R. Moazzami, J. Lee, I.-C. Chen, and C. Hu, "Projecting the Minimum Acceptable Oxide Thickness for Time-Dependent Breakdown," *International Electron Devices Meeting*, p. 710, 1988.

[2.27] J. C. Lee, I.-C. Chen and C. Hu, "Modeling and Characterization of Gate Oxide Reliability," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2268, 1988.

[2.28] K. F. Schuegraf and C. Hu, "Effects of Temperature and Defects on Breakdown Lifetime of Thin SiO_2 at Low Voltages," *International Reliability Physics Symposium*, p. 126, 1994.

Chapter 3

Gate Material Dependence of Anode Hole Injection

3.1 Introduction

This chapter presents the investigation of the breakdown mechanism of p^+ and n^+ polysilicon gate MOS structures. This comparison is of considerable technological relevance since realization of dual gate p+/n+ polysilicon CMOS technology [3.1] promises "surface" p-channel MOSFETs with superior off-state characteristics. We show that anode hole injection satisfactorily models this breakdown process, leading to accurate predictions of low voltage t_{bd} and Q_{bd}. The p^+ doped polysilicon exhibits poorer Q_{bd} than its n+ counterpart, not as the result of "enhanced hole generation efficiency," but rather as the result of reduced immunity to injected holes.

3.2. Experiment

The devices used in this study had oxide thickness varying between 30 Å and 75 Å. Oxidation process was dry oxidation at 750-950°C, depending on oxide thickness. The n^+ polysilicon devices have an in-situ phosphorus doped polysilicon gate deposited at 605°C. The p^+ polysilicon devices have boron implanted (dose: $3 \cdot 10^{15}$ cm^{-2}, energy: 20 keV) polysilicon gates, deposited undoped at 605°C. Capacitors were on n-type <100> epitaxial substrate. Constant-voltage stressing of capacitors was performed using an HP4140B picoammeter. Constant-voltage hole separation experiments using **n**MOSFETs [3.2,3.3] were performed using an HP4145B parameter analyzer. All the instruments were controlled by a PC using GPIB interface.

3.3. Results and Discussion

A basic understanding of the voltage dependence of intrinsic charge-to-breakdown, Q_{bd}, for n^+ and p^+ polysilicon gate determined by constant voltage stress requires a comparison at the same oxide, not gate voltage. Oxide voltage can be calculated from the gate voltage according to Table 3.1, deduced from J-V experiments as in [3.4,3.5] which account for band bending and polysilicon depletion [3.6,3.7,3.8]. Figure 3.1 plots Q_{bd} against the oxide voltage. The Q_{bd} dependence on J_{ox} is also shown for a second set of x-axes. Each J_{ox} corresponds to a certain V_{ox} (with electrons injected from the substrate) regardless of polysilicon doping type. Charge-to-breakdown, Q_{bd}, is lower for p^+ polysilicon gate, especially for oxide of smaller thickness.

To understand the difference better, we examine the data in light of an anode hole injection model which enables prediction of low voltage oxide breakdown parameters [3.9]. A fraction of the tunneling electrons reaching the anode are able to elastically transfer their entire energy to a deep valence-band electron [3.10,3.11,3.12]. Such an electron is promoted to the lowest empty electron energy state, i.e. the conduction band edge of the anode (the polysilicon gate), thereby creating a "hot" hole, which tunnels back into the oxide where it generates oxide traps enhancing electron tunneling. Breakdown occurs when a critical hole fluence, Q_p, has been reached. This critical hole fluence is thickness dependent, but independent of bias for a given thickness [3.2,3.9]. Charge-to-breakdown, Q_{bd}, may be lower for p^+ polysilicon gate because either Q_p is lower, i.e. oxides under p^+ polysilicon gate are more susceptible to hole injection, or more holes are injected into the oxide for each coul/cm^2 of electrons tunneling through the oxide. The latter explanation was thought possible [3.12] because "hotter" holes might be created in a p^+ polysilicon anode as the lowest empty electron states might be found at the top of the valence band rather than the bottom of the conduction band.

This line of reasoning suggests that hole injection from p^+ poly gate might be significantly enhanced, possibly explaining the reduced Q_{bd} for the p^+ gate. We performed hole separation measurements of p^+ and n^+ gate **n**MOSFETs. In this technique, the tunneling electron current is easily measured as gate current and the anode-injected hole current can be measured as substrate current [3.2,3.3]. Figure 3.2 shows no enhanced hole generation for the p^+ gate. This suggests that the density of available states at the valence band edge for accepting electrons promoted from deep in the valence band is not large enough to appreciably contribute to the anode hole injection current.

On the other hand, Figure 3.3 demonstrates that the measured hole fluence at breakdown, Q_p, is lower for p^+ than for n^+ poly-Si. This explains the reduced Q_{bd} for p^+ gate in Figure 3.1, where using the measured Q_p, the anode hole injection model [3.9] accurately predicts both n^+ and p^+ breakdown characteristics. A larger difference in Q_{bd} for thinner oxides is a direct result of the larger difference in Q_p for thinner oxides. Figure 3.4 shows the measured breakdown lifetime. It also shows the predicted lifetime, $t_{bd} = Q_{bd} / J_n$ (both Q_{bd} and J_n can be accurately predicted), down to low operating voltages.

3.4. Conclusions

This chapter has shown that the anode hole injection model is able to predict the breakdown characteristics of both p^+ and n^+ polysilicon MOS structures. Hole charge to breakdown is lower for p^+ structures reflecting a decreased immunity to hole injection, suggesting an interaction between injected holes and boron which is present in the oxide under the p^+ polysilicon. This process is further complicated due to the possible intermediary role that hydrogen [3.93.13,3.14], released by holes flowing through the

oxide, may play in reducing p^+ polysilicon Q_{bd}. This study suggests that "enhanced hole injection" [3.15] cannot explain bias temperature instability of p+ **p**MOSFET device parameters. Instead, oxide under p^+ polysilicon is more vulnerable to injected holes. Finally, although Q_{bd} is lower for p^+ structures, the maximum allowable voltage which satisfies 10 year intrinsic breakdown lifetime requirements is higher for p^+ structures due to band bending considerations.

$\|V_{ox}\| = \|V_g\|$ - the following:		
	$+V_g$	$-V_g$
n+ poly on n substrate	V_{poly}	1.3
n+ poly on p substrate	V_{poly}	1.3
p+ poly on n substrate	1.3	V_{poly}
p+ poly on p substrate	1.3	V_{poly}

Table 3.1 J-V Experiments show how V_{ox} may be determined for different bias polarity and gate and substrate doping types. The numerical voltage offsets reflect band bending in the MOS system. V_{poly} is calculated as $V_{poly} = \frac{\varepsilon_{ox}^2 E_{ox}^2}{2q\varepsilon_{Si}N_{poly}}$ until V_{poly} is pinned at 1.12 V due to the saturation of band bending in strong inversion. Another ~0.5 V is to be taken off the "effective V_{ox}" in calculating the tunneling current when the cathode is p^+ polysilicon on account of low surface electron concentration. With these "corrections," tunneling characteristics are reduced to a unique curve predicted by Fowler-Nordheim and direct tunneling conduction theory [3.4,3.5].

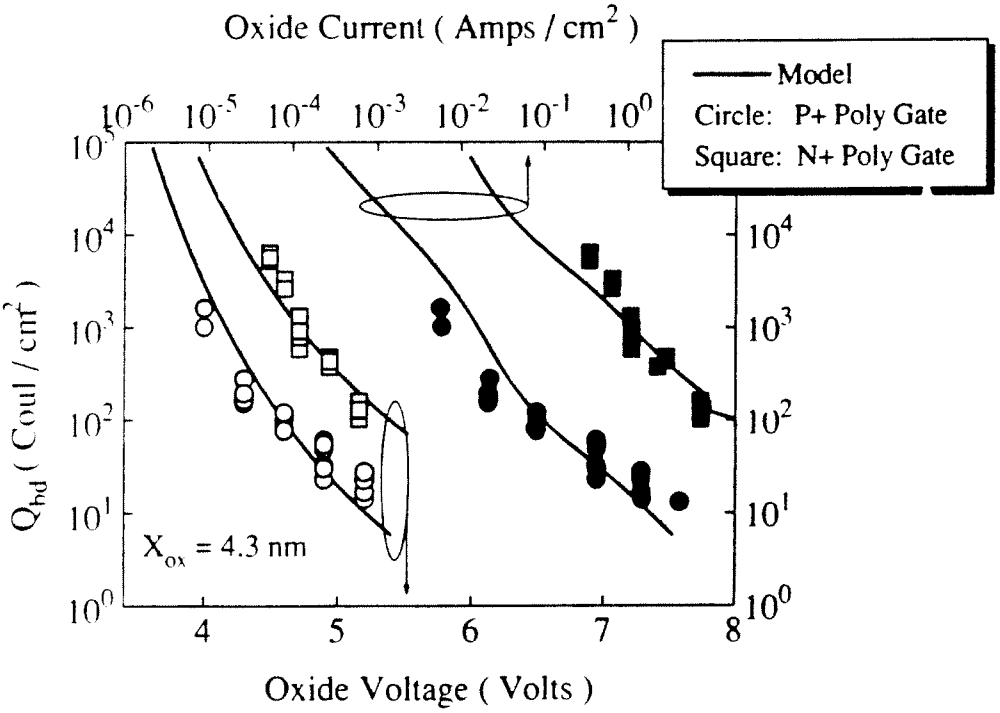


Fig. 3.1 Voltage and Current Dependence of $\mathbf{Q_{bd}}$ for oxide with n^+ and p^+ polysilicon gate. The bias dependence follows the prediction of the anode hole injection model [3.9]. $\mathbf{Q_{bd}}$ is lower for p^+ polysilicon gate.

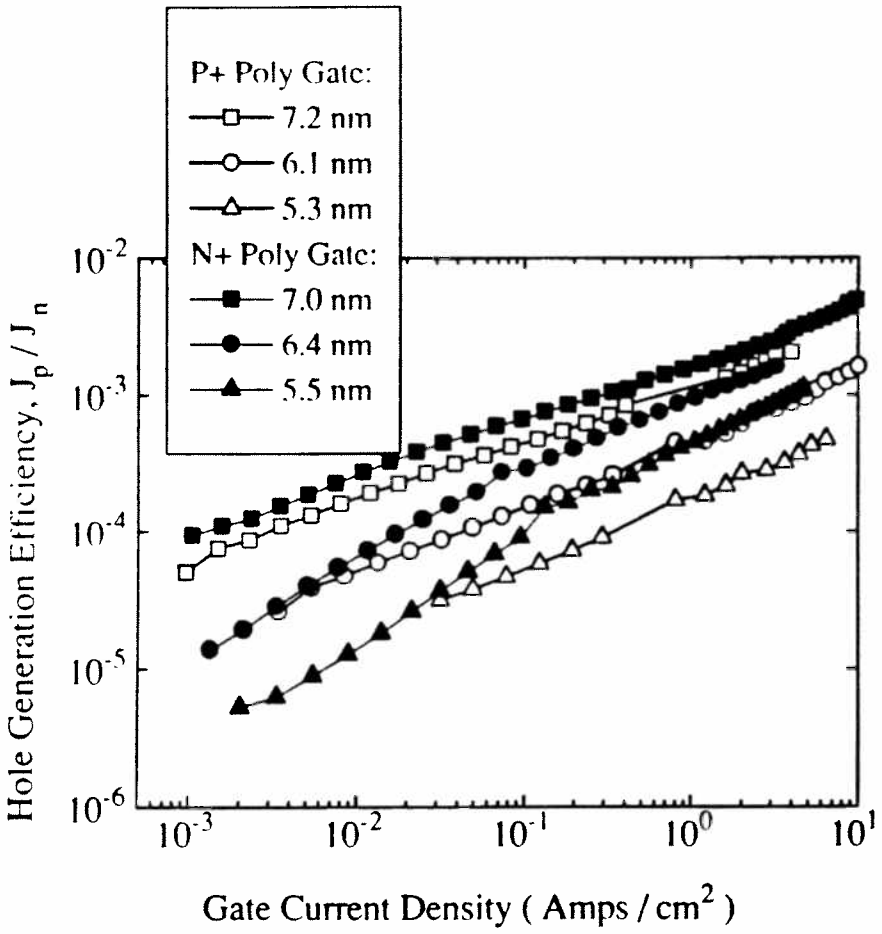


Fig. 3.2 Quantum Efficiency of Hole Generation Process, $\mathbf{J_p}$ / $\mathbf{J_n}$, inversely proportional to $\mathbf{Q_{bd}}$ and $\mathbf{t_{bd}}$, measured by the hole separation technique [3.2,3.3], is similar for both n^+ and p^+ polysilicon, showing no enhanced hole generation for the p^+ gate. Bias configuration for hole separation measurement is shown in **Fig. 2.2**. Tunneling electrons are supplied by the source and drain regions with current density $\mathbf{J_n}$. Holes generated in the anode tunnel towards the substrate where they are collected as substrate current, $\mathbf{J_p}$. Time integrated electron tunneling current is $\mathbf{Q_{bd}}$, whereas time integrated hole current is $\mathbf{Q_p}$.

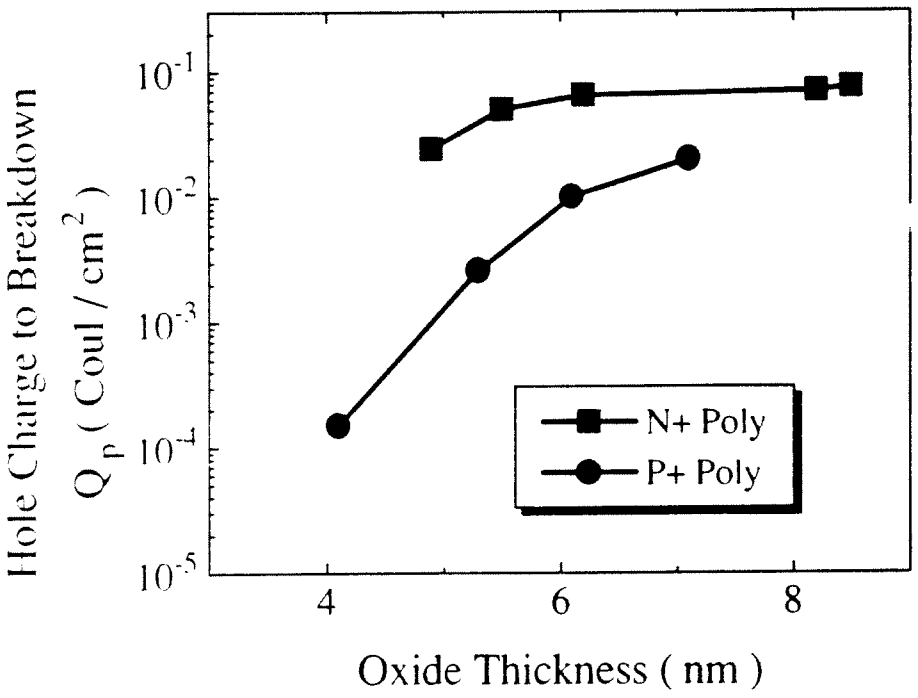


Fig. 3.3 Comparison of $\mathbf{Q_p}$ values measured using hole separation for n^+ and p^+ polysilicon showing that oxide under p^+ gate can sustain less hole injection at breakdown.

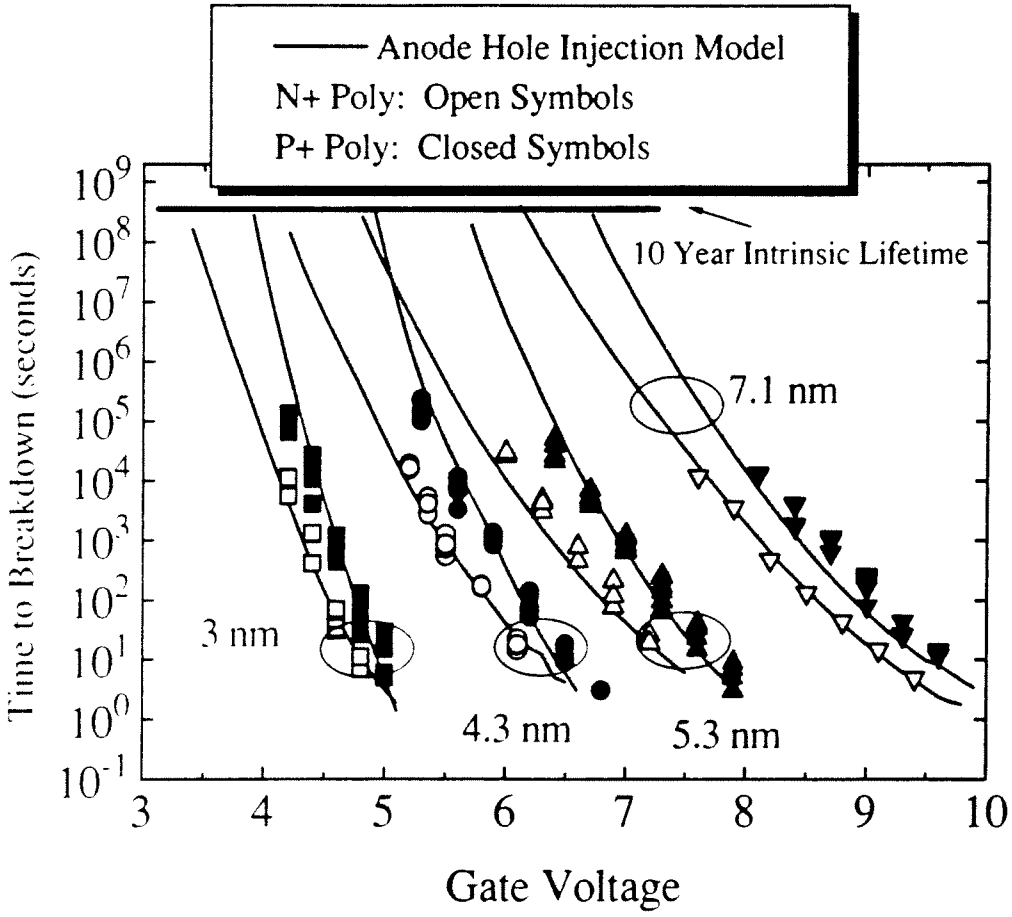


Fig. 3.4 Voltage dependence of breakdown lifetime for thin oxides showing that the Anode Hole Injection Model can accurately predict breakdown lifetime at low voltages for oxide under both n^+ and p^+ polysilicon gate.

3.5. References

[3.1] L. C. Parrillo, S. J. Hillenias, R. L. Field, E. L. Hu, W. Fichtner and M.-L. Chen, "A Fine-line CMOS that Uses P^+ Polysilicon Silicide Gates for NMOS and PMOS Devices," *International Electron Devices Meeting*, p. 418, 1984.

[3.2] I.-C. Chen, S. E. Holland and C. Hu, "Oxide Breakdown Dependence on Thickness and Hole Current - Enhanced Reliability of Ultra-Thin Oxides," *International Electron Devices Meeting*, p. 660, 1986.

[3.3] B. Eitan and A. Kolodny, "Two Components of Tunneling Current in Metal-Oxide-Semiconductor Structures," *Applied Physics Letters*, vol. 43, no. 1, p. 106, 1983.

[3.4] E. Rosenbaum, R. Moazzami and C. Hu, "Implications of Waveform and Thickness Dependence of SiO_2 Breakdown on Accelerated Testing," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications*, p. 214, 1991.

[3.5] K. F. Schuegraf, C. C. King and C. Hu, "Impact of Polysilicon Depletion in Thin Oxide MOS Technology," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications*, p. 86, 1993.

[3.6] B. J. Fishbein and D. B. Jackson, "Performance Degradation of N-Channel MOS Transistors During DC and Pulsed Fowler-Nordheim Stress, *Proceedings of the International Reliability Physics Symposium*, p. 159, 1990.

[3.7] S. J. Wang, I.-C. Chen, and H. L. Tigelaar, "Effects of Poly Depletion on the Estimate of Thin Dielectric Lifetime", *IEEE Electron Device Letters*, vol. 12, no. 11, p. 617, 1991.

[3.8] C.-L. Huang and N. D. Arora, "Measurements and Modeling of MOSFET I-V Characteristics with Polysilicon Depletion Effect," *IEEE Transactions on Electron Devices*, vol. ED-40, no. 12, p. 2330, 1993.

[3.9] K. F. Schuegraf and C. Hu, "Hole Injection SiO_2 Breakdown Model for Very Low Voltage Lifetime Extrapolation," to appear in *IEEE Trans. on Elec. Dev.*

[3.10] Z. A. Weinberg, W. C. Johnson, and M. A. Lampert, "High-Field Transport in SiO_2 on Silicon Induced by Corona Charging of the Unmetallized Surface," *Journal of Applied Physics*, vol. 47, no. 1, p. 248, 1976.

[3.11] Z. A. Weinberg, M. V. Fischetti, and Y. Nissan-Cohen, "SiO_2-Induced Substrate Current and its Relation to Positive Charge in Field-Effect Transistors," *Journal of Applied Physics*, vol. 59, no. 3, p. 824, 1986.

[3.12] S. E. Holland, I.-C. Chen, and C. Hu, "Ultra-Thin Silicon-Dioxide Breakdown Characteristics of MOS Devices with n^+ and p^+ Polysilicon Gates," *IEEE Electron Device Letters*, vol. 8, no. 12, 1987.

[3.13] N. S. Saks and D. B. Brown, "Observation of H^+ Motion During Interface Trap Formation," *IEEE Transactions on Nuclear Science*, vol. 37, no. 6, p. 1624, 1990.

[3.14] D. J. DiMaria, E. Cartier and D. Arnold, "Impact Ionization, Trap Creation, Degradation, and Breakdown in Silicon Dioxide Films on Silicon," *Journal of Applied Physics*, vol. 73, no. 7, p. 3367, 1993.

[3.15] Y. Hiruta, F. Matsuoka, K. Hama, H. Iwai, K. Maeguchi and K. Kanzaki, "+BT Instability in P^+ Poly Gate MOS Structure," *International Electron Devices Meeting*, p. 578, 1987.

Chapter 4

Temperature Acceleration of Breakdown

4.1. Introduction

This chapter presents a study of the temperature acceleration of oxide breakdown within the framework of the anode hole injection model. This study shows that the shape of the oxide breakdown characteristic, that is that of Q_{bd} and t_{bd}, is virtually independent of temperature for temperatures less than 150 °C. However, the oxide's immunity to hole injection, as measured by Q_p, decreases with increasing temperature with an activation energy of about 250 meV. This increased susceptibility to hole injection indicates that holes are more efficient mediators of the oxide damage leading to breakdown at higher temperatures.

4.2. Temperature Acceleration

It is generally accepted that higher temperatures degrade the oxide breakdown characteristics. This section discusses the temperature acceleration of oxide breakdown within the framework of the low-voltage anode hole injection model, showing that the degraded oxide integrity at higher temperatures is attributable to the oxide's reduced immunity to injected holes at higher temperatures.

4.2.1. Charge to Breakdown

The field dependent ratio of hole current to electron tunneling current, J_p / J_n, i.e. the quantum efficiency that a tunneling electron induces the injection of a hole into the oxide from the anode can be measured using the hole separation technique [4.1,4.2] of

an **n**MOSFET biased as described in Chapter 2. Figure 4.1 illustrates the temperature independence of the quantum efficiency of the hot-hole injection between 25 °C and 125 °C. Therefore the voltage dependence of Q_{bd} should follow the prediction of Eq. (2.5) as shown in Figure 4.2. The reduction in Q_{bd} in Figure 4.2 with increasing temperature is modeled solely by a reduction in Q_p, as shown in Figure 4.3. This reduction in Q_p shows that an oxide is more vulnerable to hole transport at higher temperatures. That is, holes are more effective agents of damage to the oxide at higher temperatures. The activation energy of Q_p was found to be 0.212 eV for 70 Å and 76 Å oxides and 0.327 eV for oxides thinner than 55 Å.

The number of holes flowing through an oxide to reach breakdown, Q_p, has been empirically observed to be constant independent of applied bias [4.1,4.3]. Moreover, hole injection has been shown to lead to bulk oxide trap generation [4.4,4.5,4.6,4.7] as well as interface trap generation [4.8]. This establishes a correlation between hole fluence and oxide damage. The temperature activation of Q_p therefore reflects an acceleration of the hole-mediated oxide damage mechanism.

The appearance of a temperature dependence of the injected hole quantum efficiency between 125 and 175 °C may explain the reported increase in activation energy at temperatures above 150 °C [4.9,4.10,4.11]. Not only is the oxide immunity to injected holes diminished above 150 °C, but the tunneling electrons become more efficient in generating and injecting holes into the SiO_2. Thus, the hole separation measurement offers insight into the physical mechanism responsible for increased activation energies of oxide breakdown above 150 °C. That is, the activation energy above 150 °C is the sum of two components: the first represents the decrease in Q_p at higher temperatures, the second the significantly increased efficiency of the anode hole injection process above 150 °C.

4.2.2. Time to Breakdown

Since breakdown time, t_{bd}, is related to Q_{bd} as

$$t_{bd} = Q_{bd}/J_n \qquad (4.1),$$

understanding the voltage and temperature dependence of the electron tunneling current, J_n, is essential. For applied voltages greater than 3.15 Volts, electrons tunnel through a triangular barrier and the tunneling current follows the Fowler-Nordheim relationship. For lower voltages, electrons tunnel through a trapezoidal barrier; this direct tunneling oxide leakage current is significantly enhanced over that predicted by the Fowler-Nordheim relationship[4.12]. Figure 4.4 shows that the increase of Fowler-Nordheim and direct tunneling current due to increased temperature is secondary compared with its exponential increase with increasing voltage. Moreover, the sensitivity on temperature decreases with decreasing thickness. The reduction in t_{bd} for increased temperature is accurately predicted using Eqs. (2.5) and (4.1) as shown in Figure 4.5, establishing that the anode hole injection model is able to predict very low voltage intrinsic breakdown lifetime.

Figure 4.6 shows that the logarithms of Q_{bd} and t_{bd} are linear functions of inverse field with a slope that is insensitive to temperature. Both Q_{bd} and t_{bd} are seen to decrease by a factor of approximately 20 between 25 °C and 125 °C. This acceleration factor holds true for oxides between 30 Å and 80 Å.

Figure 4.7 compares the t_{bd} activation energies for the thin oxides in this study with those reported in other studies [4.9,4.13,4.14,4.15]. While the activation energy appears to be independent of field for oxides thinner than 80 Å, it has been shown to decrease with increasing field for oxides thicker than 100 Å [4.11]. Perhaps the electron energy gained inside the oxide, significant for thick oxides, is temperature dependent. This is consistent with a temperature dependent scattering length [4.16].

4.3. Summary and Conclusions

This chapter completes the development of a comprehensive quantitative model for "intrinsic" silicon dioxide breakdown, based on the concept of anode hole injection, accounting for the effects of field and temperature acceleration. This model enables prediction of insulator reliability at operating voltages, 3.3 Volts and below for oxides below 100 Å. The next chapter considers the role of defects in limiting the realization of the high intrinsic oxide quality in large area structures, i.e. integrated circuits or canonical test structures that resemble them, developing a means of quantifying an oxide technology's defectivity using the physically-based intrinsic breakdown models.

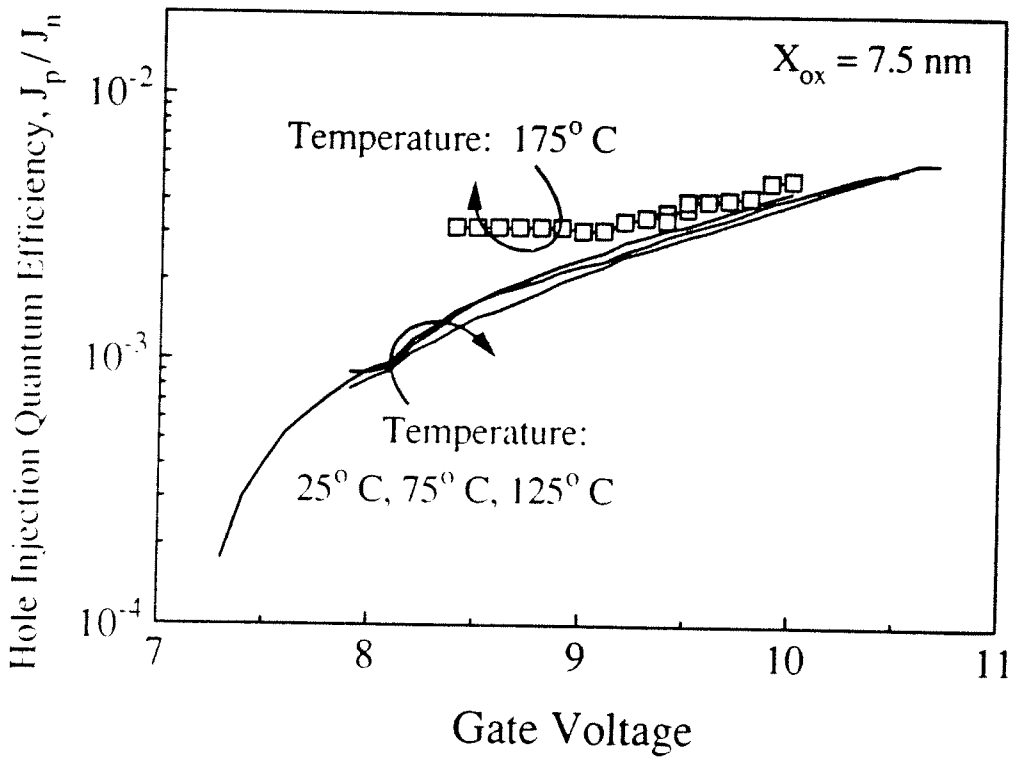


Fig. 4.1 Quantum Efficiency of Hole Generation Process, J_p / J_n, inversely proportional to $\mathbf{Q_{bd}}$ and $\mathbf{t_{bd}}$, is independent of temperature between 25° C and 125° C, measured by the hole separation technique [4.1]. Bias configuration for hole separation measurement is shown in **Fig 2.2**. Tunneling electrons are supplied by the source and drain regions with current density J_n. Holes generated in the anode tunnel towards the substrate where they are collected as substrate current, J_p.

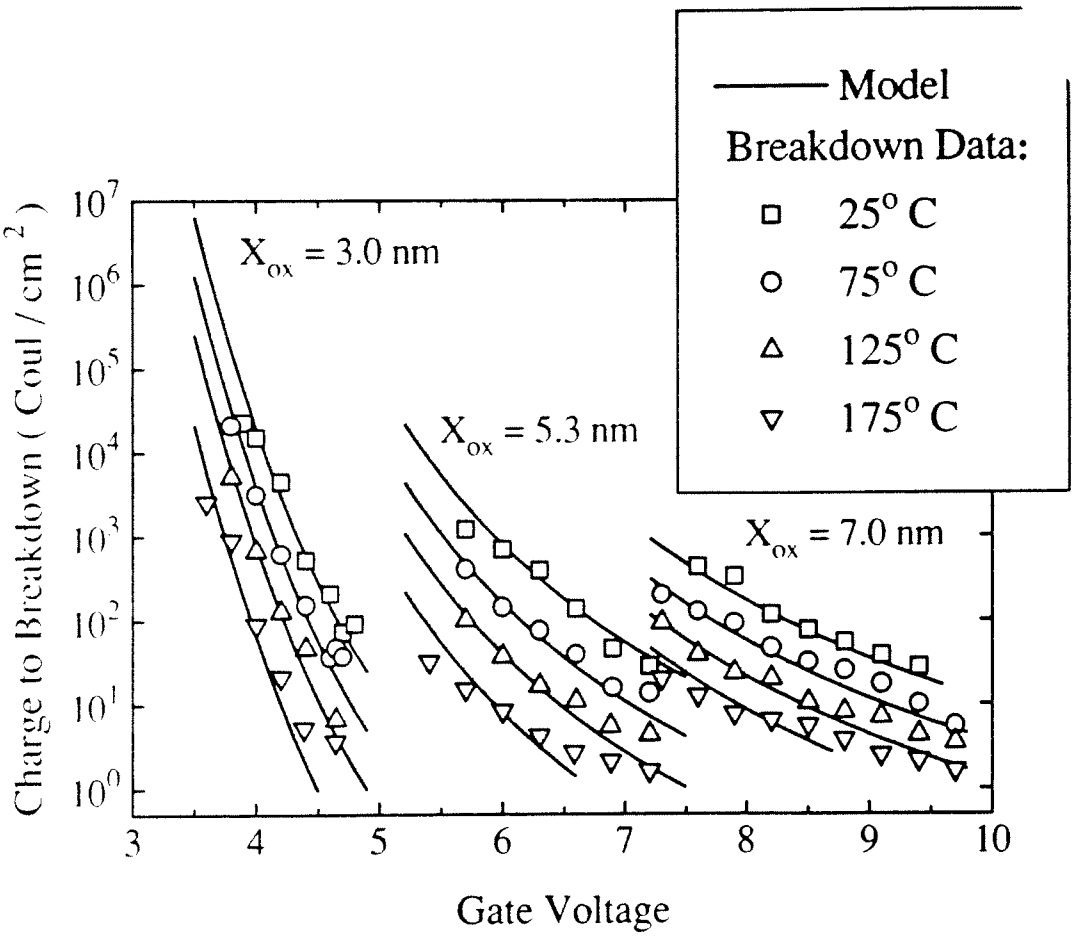


Fig. 4.2 Voltage dependence of $\mathbf{Q_{bd}}$ for several different temperatures. Reduction in $\mathbf{Q_{bd}}$ with increasing temperature is due to oxide's lower hole immunity at higher temperatures.

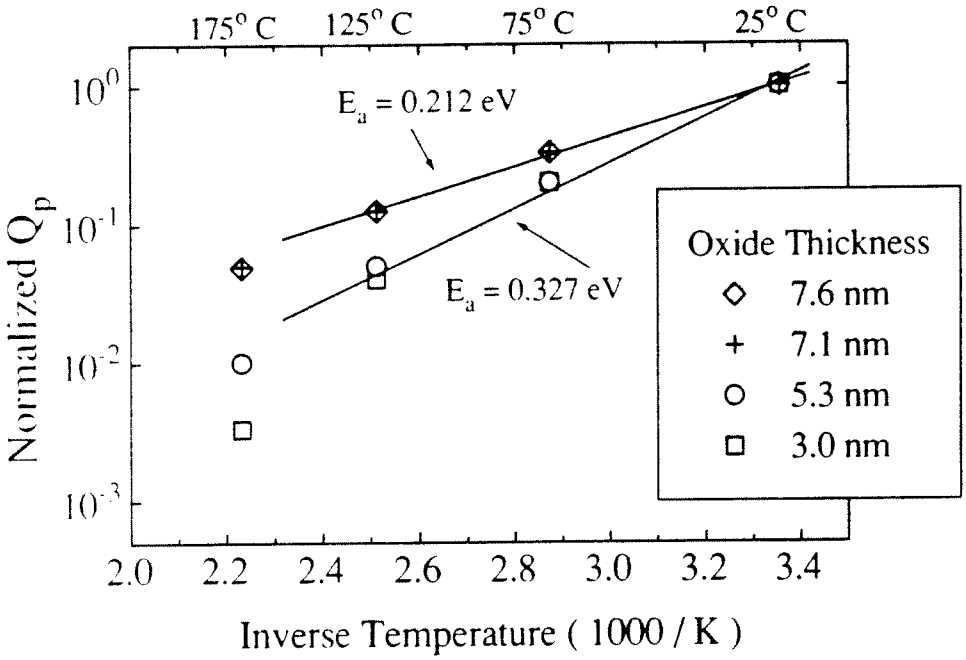


Fig. 4.3 Temperature dependence of $\mathbf{Q_p}$, normalized as $\mathbf{Q_p(T)/Q_p(T=25°C)}$, for several different oxide thicknesses. Reduction of $\mathbf{Q_p}$ reflects oxide's lower hole immunity at higher temperatures. Only $\mathbf{Q_p}$ is adjusted to model the breakdown data in **Fig. 4.2**.

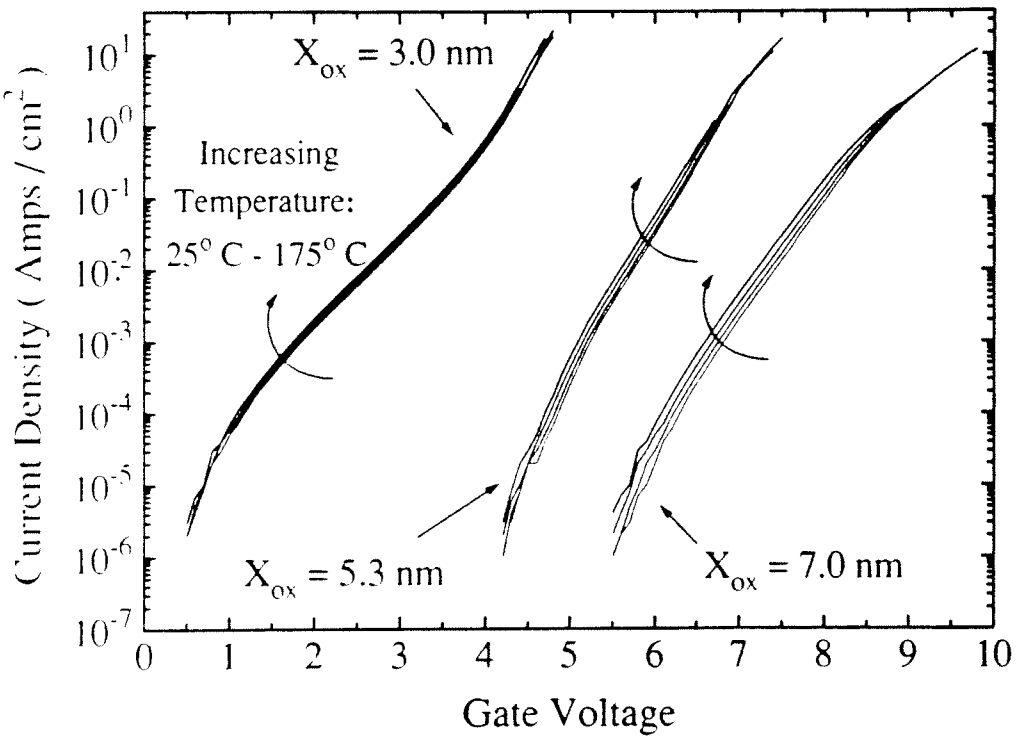


Fig. 4.4 Temperature dependence of electron tunneling current is negligible and decreases for thinner oxides.

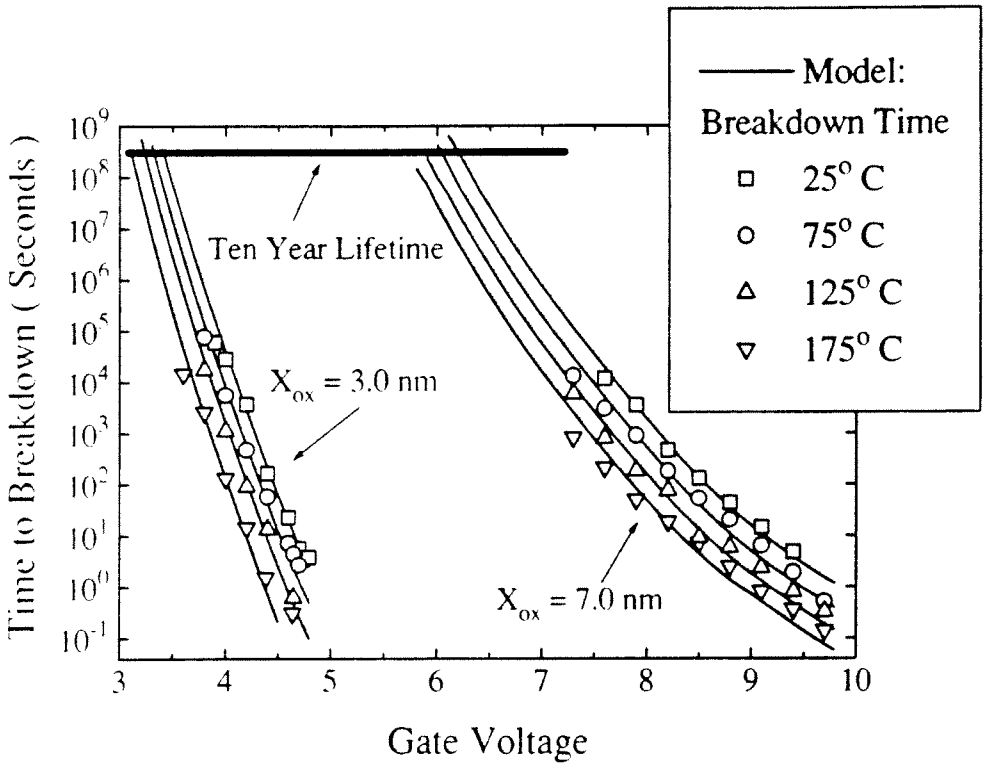


Fig. 4.5 Breakdown lifetime, $\mathbf{t_{bd}}$, decreases with increasing temperature, but Anode Hole Injection Model maintains its ability to predict the breakdown characteristics.

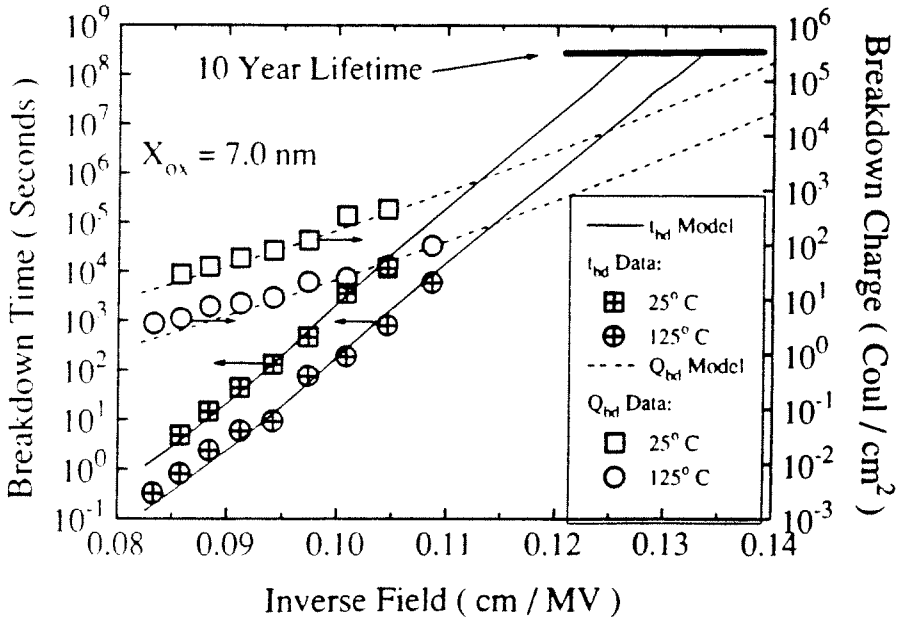


Fig. 4.6 Inverse field dependence of $\mathbf{t_{bd}}$ and $\mathbf{Q_{bd}}$ for 70 Å oxide shows nearly linear slope that is independent of temperature.

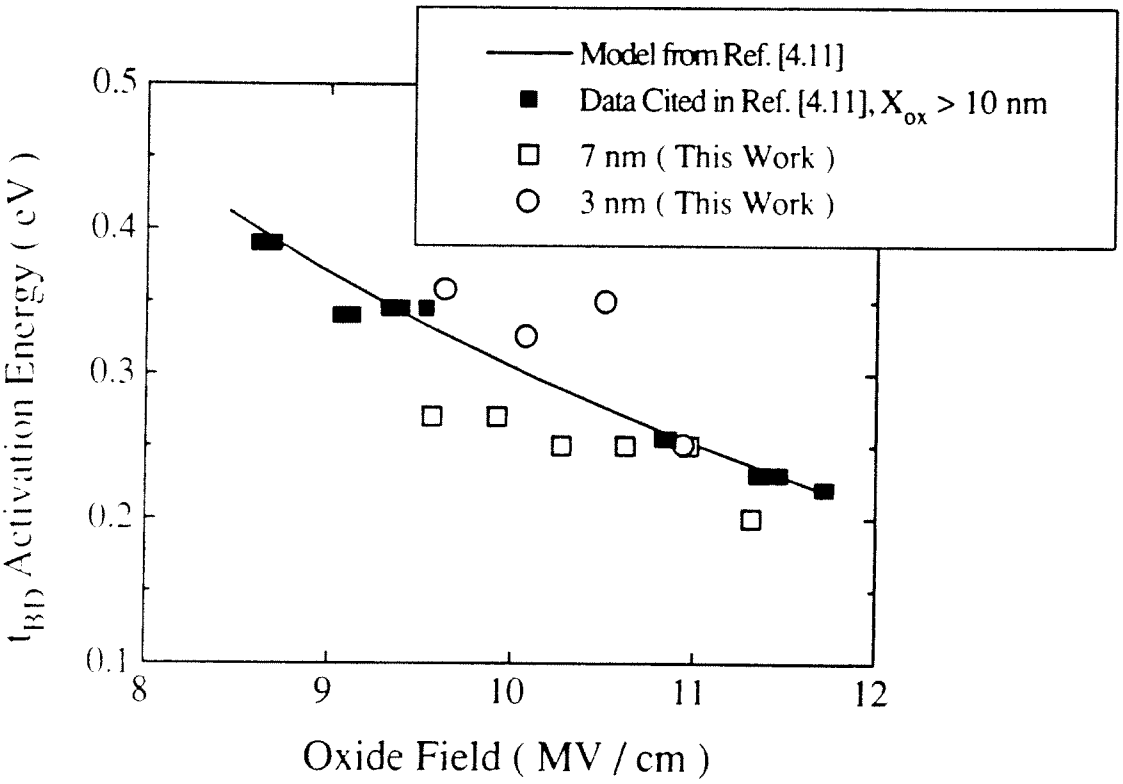


Fig. 4.7 Survey of temperature activation energies for breakdown lifetime reported in the literature. This work concludes that activation energy is independent of field for oxides thinner than 80 Å.

4.4. References

[4.1] I.-C. Chen, S. E. Holland and C. Hu, "Oxide Breakdown Dependence on Thickness and Hole Current - Enhanced Reliability of Ultra-Thin Oxides," *International Electron Devices Meeting*, p. 660, 1986.

[4.2] K. F. Schuegraf and C. Hu, "Hole Injection SiO_2 Breakdown Model for Very Low Voltage Lifetime Extrapolation," *IEEE Trans. on Elec. Dev.*, May, 1994.

[4.3] K. R. Mistry, D. B. Krakauer, and B. S. Doyle, "Impact of Snapback-Induced Hole Injection on Gate Oxide Reliability of N-MOSFETs," *IEEE Electron Device Letters*, vol. 11, no. 10, p. 460, 1990.

[4.4] S. K. Lai, "Interface Trap Generation in Silicon Dioxide when Electrons are Captured by Trapped Holes," *Journal of Applied Physics*, vol. 54, no. 5, p. 2540, 1983.

[4.5] I.-C. Chen, S. E. Holland, and C. Hu, "Electron-trap Generation by Recombination of Electrons and Holes in SiO_2," *Journal of Applied Physics*, vol. 61, no. 9, p. 4544, 1987.

[4.6] H. Uchida and T. Ajioka, "Electron Trap Center Generation Due to Hole Trapping in SiO_2 Under Fowler-Nordheim Tunneling Stress," *Applied Physics Letters*, vol. 51, no. 6, p. 433, 1987.

[4.7] S. Ogawa, N. Shiono, and M. Shimaya, "Neutral Electron Trap Generation in SiO_2 by Hot Holes," *Applied Physics Letters*, vol. 56, no. 14, p. 1329, 1990.

[4.8] D. J. DiMaria, D. Arnold, and E. Cartier, "Degradation and Breakdown of Silicon Dioxide Films on Silicon," *Applied Physics Letters*, vol. 61, no. 19, p. 2329, 1992.

[4.9] C. F. Chen, C. Y. Wu, M. K. Lee, and C. N. Chen, "Dielectric Reliability of Intrinsic Thin SiO_2 Films Thermally Grown on a Heavily Doped Si Substrate-Characterization and Modeling," *IEEE Transactions on Electron Devices*, vol. ED-34, no. 7, p. 1540, 1988.

[4.10] Y. Hokari, T. Baba and N. Kawamura, "Reliability of 6-10 nm Thermal SiO_2 Films Showing Intrinsic Dielectric Reliability," *IEEE Transactions on Electron Devices*, vol. ED-32, no. 11, p. 2485, 1985.

[4.11] R. Moazzami, J. C. Lee, and C. Hu, "Temperature Acceleration of Time-Dependent Dielectric Breakdown," *IEEE Transactions on Electron Devices*, vol. ED-36, no. 11, p. 2462, 1989.

[4.12] K. F. Schuegraf, C. C. King and C. Hu, "Ultra-Thin Silicon Dioxide Leakage Current and Scaling Limit," *Symposium on VLSI Technology Digest of Technical Papers*, p. 18, 1992.

[4.13] D. L. Crook, "Methods of Determining Reliability Screens for Time Dependent Dielectric Breakdown," *International Reliability Physics Symposium*, p. 1, 1979.

[4.14] K. Yamabe and K. Taniguchi, "Time-dependent Dielectric Breakdown of Thin Thermally Grown SiO_2 Films," *IEEE Transactions on Electron Devices*, vol. ED-32, no. 2, p. 423, 1985.

[4.15] D. A. Baglee, "Characteristics and Reliability of 100 Å Oxides," *International Reliability Physics Symposium*, p. 152, 1984.
[4.16] D. J. DiMaria, "Trap Creation in Silicon Dioxide Produced by Hot Electrons," *Journal of Applied Physics*, vol. 65, no. 6, p. 2342, 1989.

Chapter 5

Effect of Defects

5.1. Introduction

This chapter presents the results leading to the development of simple methodology for characterizing the defectivity of a thin oxide technology. Whereas the preceding discussion focused on understanding the voltage and temperature dependence of intrinsic breakdown, i.e. the breakdown of small area oxide samples, this chapter examines the role of defect-induced breakdown observable in large area oxide samples, i.e. those with a higher probability of containing gross oxide defects. Precisely these gross oxide defects limit the oxide yield and reliability of VLSI circuits. The following defect breakdown model offers a framework for determining whether an oxide technology offers ample lifetime at the circuit's operating voltage based on the physical anode hole injection model, by combining the intrinsic anode hole injection model with an "effective" thinning model [5.1].

5.2. Defect-Related Breakdown Model

Figure 5.1 summarizes the intrinsic 25 °C breakdown lifetime as a function of oxide thickness for several oxide voltages. This breakdown lifetime can be approximated as

$$t_{bd} = \tau_o e^{\frac{GX_{ox}}{V_{ox}}} \tag{5.1}$$

where G = 390 MV/cm and $\tau_O \approx 5 \cdot 10^{-13}$ sec. Despite increasing extrapolation slope, G, and decreasing pre-exponential constant as oxide thickness decreases (see Figure 2.14), Eq. (5.1) serves as an excellent approximation, thereby minimizing computational complexity of the effective thinning model. An intrinsic 400 year

lifetime (equivalent to 20 years at 125° C) is achievable with a 40 Å oxide for V_{ox} = 3.3 V; however, in practice, some thickness margin needs to be provided for defects in order to guarantee adequate dielectric reliability. This section uses the concept of effective thinning to model oxide defects in order to determine the density of defects that would cause premature breakdown.

Figure 5.2 shows the time-dependent dielectric breakdown data in cumulative percentage versus time for large area (1 mm^2) capacitors under positive bias. For each of the 5.6 V, 5.2 V, and 4.9 V bias conditions, there is a broad lifetime distribution spreading several orders of magnitude only approaching the intrinsic breakdown lifetimes. This distribution of breakdown lifetimes is due to defects of varying severity. The concept of "effective thinning" [5.1], illustrated in Figure 5.3, enables quantitative analysis of these statistical distributions of high field breakdown lifetime to accurately predict low field breakdown distributions. All oxide defects are modeled by an effective oxide thickness less than the nominal thickness determined by the breakdown lifetime as follows: each lifetime data point in Figure 5.2 is translated into an effective thickness according to the intrinsic breakdown model predictions of Figure 5.1. Alteratively, interpreting X_{ox} as X_{eff} in Eq. (5.1), the effective thickness follows from the breakdown lifetime as

$$X_{eff} = \frac{V_{ox}}{G} \cdot \ln\left[\frac{t_{BD}}{\tau_o}\right] \qquad (5.2)$$

where $X_{eff} < X_{ox}$ defines the effective thickness for a sample with nominal oxide thickness X_{ox}. The X_{eff} term may reflect "real" oxide thinning, surface asperities which enhance E_{ox} (mathematically equivalent to a reduction in X_{eff}, see Eq. (5.2)), or a locally reduced tunneling barrier height (reducing G in Eq. (5.2), again,

mathematically equivalent to a reduction in X_{eff}). All these effects can decrease t_{bd}. The device's "thinnest" spot, i.e. smallest X_{eff}, determines t_{bd}.

This approach of characterizing defect distributions is quite general and believed to be applicable for arbitrary sizes (areas) of oxide samples since its basis is the physical anode hole injection model. For every oxide technology, there exists an areal density of defects, $D(X_{eff})$, which uniquely defines the defect distribution independent of the origins of the defect [5.1]. Larger sample areas have a higher probability of incorporating a severe defect, causing the distribution to shift towards thinner effective thickness with a longer distribution tail. Knowing the probability of incorporating a defect of any given X_{eff}, one can calculate the probability of an oxide failing at a certain time for a specified supply voltage.

For ramp rates between 2 V/s and 20 V/s, X_{eff} may be deduced from the ramp voltage breakdown test (to first order) as

$$X_{eff}\ [nm] \approx \frac{V_{bd}\ [V]}{1.5} \qquad (5.3).$$

Figure 5.4 shows a refinement to Eq. (5.3) demonstrating that higher ramp rate leads to higher V_{bd} for the same X_{eff}. Equations (5.2,5.3), Figures. 5.1 and 5.4 enable the transformation of t_{bd} and V_{bd} statistics into X_{eff} statistics. The coincidence of the statistical distribution of effective thickness determined by this transformation as shown in Figure 5.5 supports the hypothesis that the distribution of defect-induced oxide breakdown lifetime can indeed be described by an "effectively thinnest" spot. The ability to determine the effective thickness distribution from the ramp voltage breakdown test establishes a particularly convenient way to evaluate an oxide technology's defectivity.

Once X_{eff} for a given percentile point in the distribution is known, the lifetime of that percentile can be extrapolated down to low operating voltages using the intrinsic breakdown model as in Figure 2.6. Figure 5.6 shows that this technique is able to not only predict the reliability of isolated points in the lifetime distribution (i.e. t_{50}), but that the high-field breakdown test can be used to predict the entire TDDB distribution at different voltages and temperatures. Figure 5.7 shows that distributions of Q_{bd} can also be predicted from the X_{eff} distribution and Eq. (2.5).

5.3. Supply Voltage Limits

Figure 5.8 summarizes the minimum effective thickness necessary to guarantee 30 year lifetime as a function of operating voltage for room temperature and 125° C operation. A 5.5 Volt (5 Volt plus 10 %) technology needs 76 Å, a 3.6 Volt (3.3 V plus 10 %) technology needs an effective thickness of at least 44 Å at 125°C, and a 2.8 Volt technology (2.5 Volt plus 10 %) needs an effective thickness larger than 34 Å. It is likely that a thickness margin of about 50 % (like 120 Å for 5.5 V) will be practical for an advanced sub-micron process. Moreover, Figure 5.8 shows the minimum acceptable ramp breakdown voltages needed in order to guarantee acceptable 30 year oxide reliability. They are 11.5 V for 5.5 V operation, 7.4 V for 3.6 V operation, and 6.0 V for 2.8 V operation. The population of oxides not meeting these V_{bd} requirements contain defects that may cause breakdown in less than 30 years. When interpreting the V_{bd} data, care must be taken to consider the gate work function and possible polysilicon gate depletion [5.2,5.3,5.4,5.5].

5.4. Summary and Conclusions

This chapter presents an extension to the intrinsic hole injection breakdown model to account for the role of defects in limiting achievable oxide breakdown lifetime. This simple model accounts for the role of defects in reducing oxide lifetime and enables the determination of whether an oxide technology provides sufficient margin for defects to guarantee adequate oxide failure rate. For 3.3 and 2.5 Volt operation, the minimum effective thickness is 44 and 34 Å, respectively. However, it is the ability to manufacture such high-quality low-defect ultra-thin silicon dioxide will ultimately determine the usable oxide thickness and supply voltage.

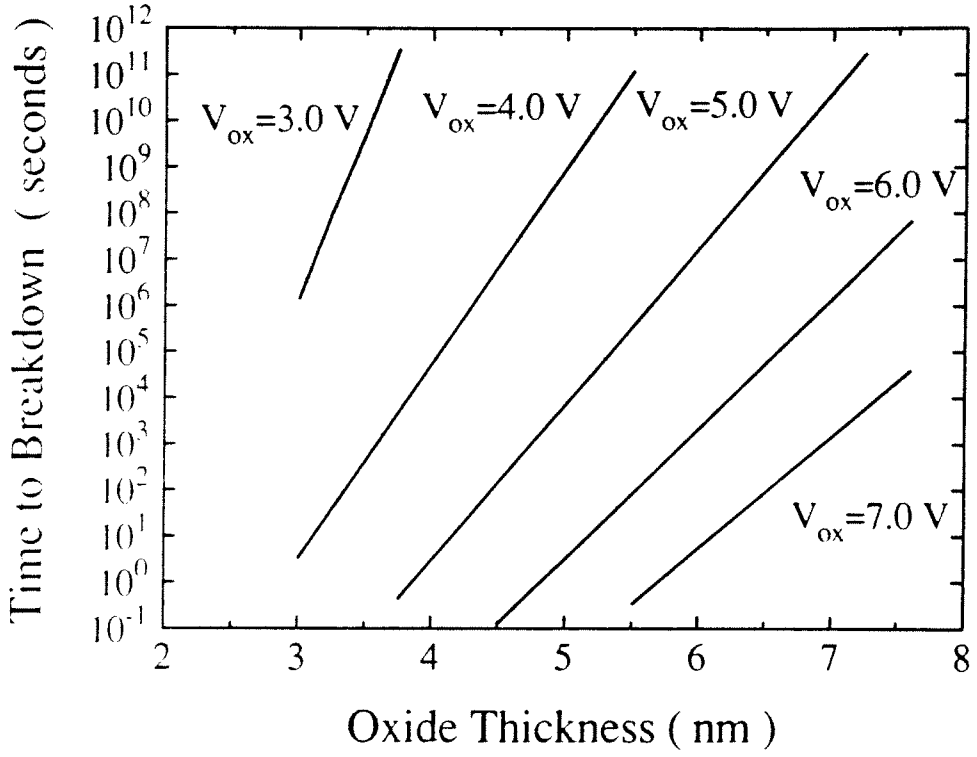


Fig. 5.1 Model predictions for oxide lifetime for several operating voltages as a function of oxide thickness.

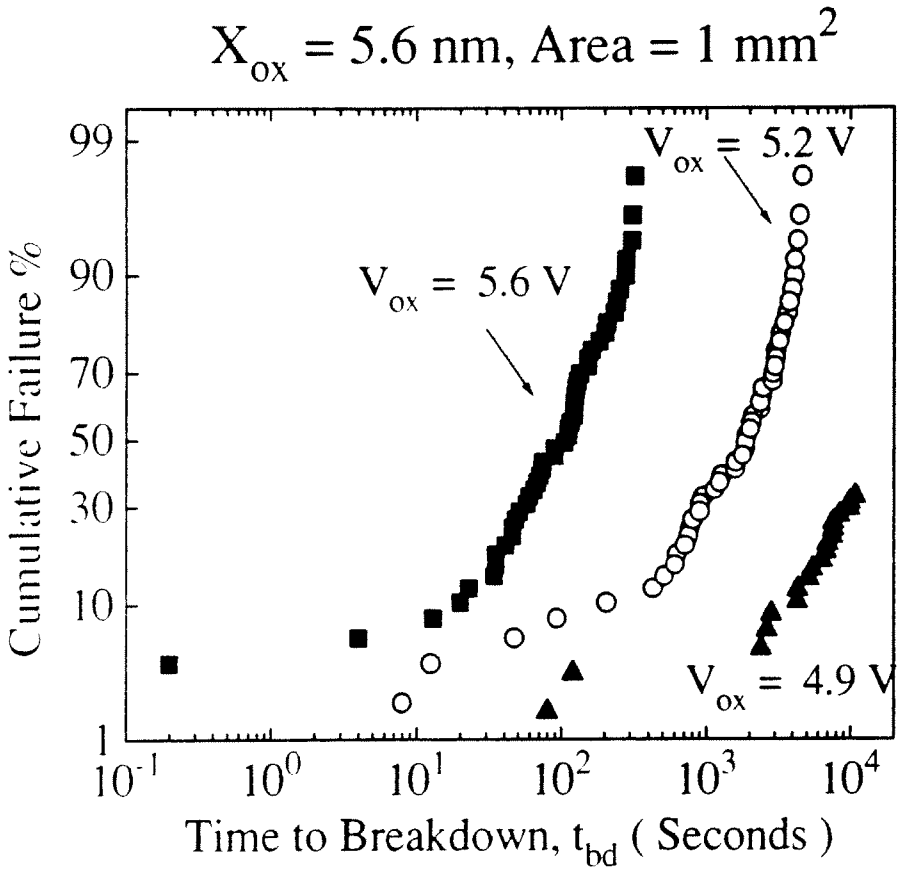


Fig. 5.2 Cumulative Fail Distribution of Defect Breakdown Time for several constant voltage stress levels. Oxide voltage is calculated by accounting for the effects of polysilicon depletion, band-bending and substrate resistance [5.2,5.3,5.4,5.5].

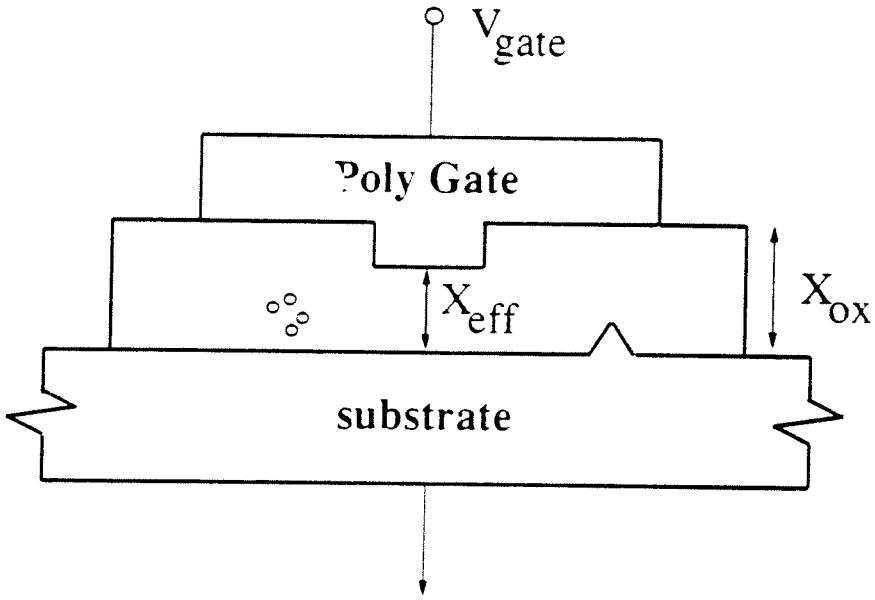


Fig. 5.3 Oxide defects are modeled by an effective oxide thickness which determines the nominal oxide's breakdown time. The abscissa in **Fig. 5.1** is interpreted as effective oxide thickness, $\mathbf{X_{eff}}$.

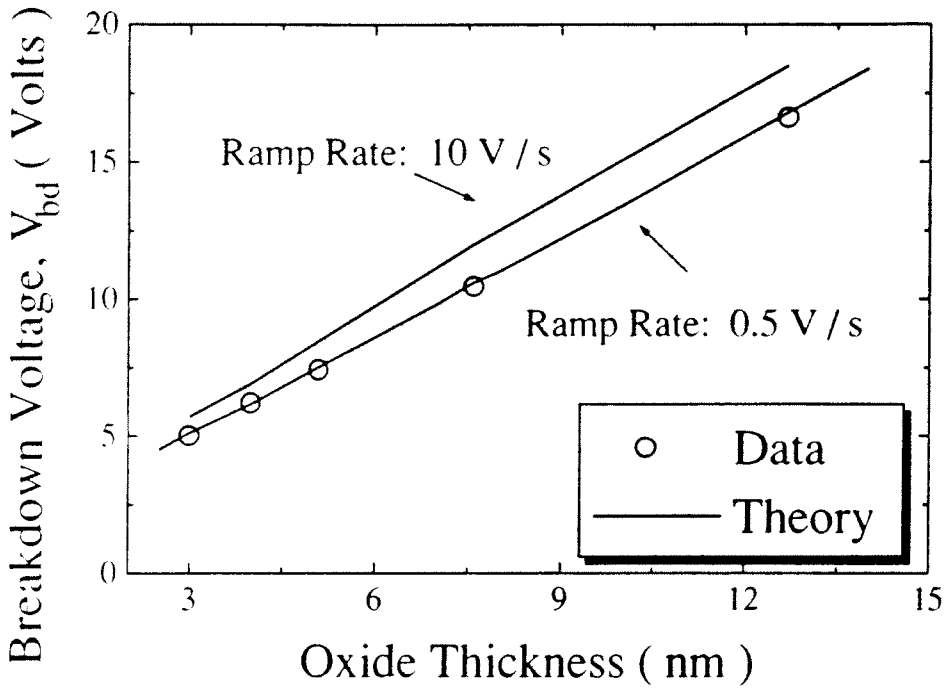


Fig. 5.4 Ramp breakdown voltage, $\mathbf{V_{bd}}$, as function of oxide thickness for different ramp rates. The theoretical lines are predicted from the anode hole injection model. The abscissa may again be interpreted as effective thickness allowing a rapid characterization of oxide defects.

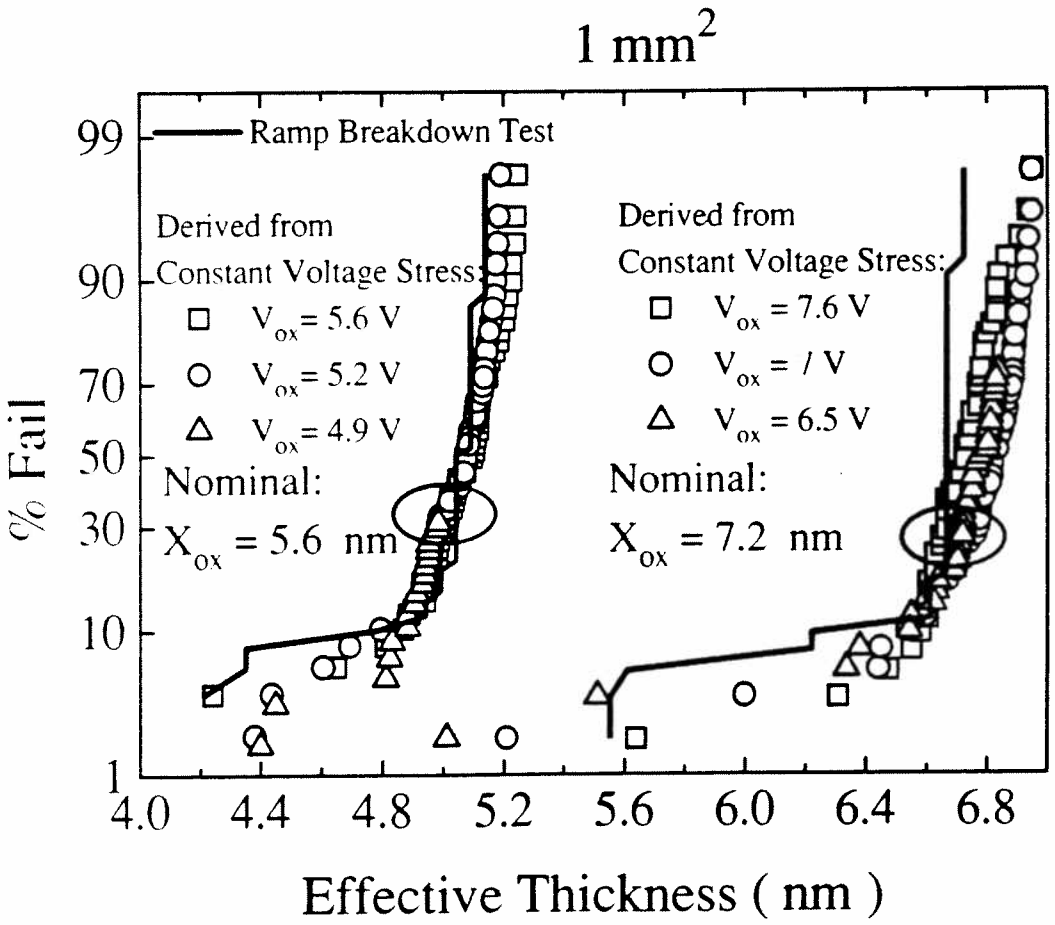


Fig. 5.5 Statistical distribution of effective thickness, $\mathbf{X_{eff}}$, supports the hypothesis that defects in even aggressively scaled oxides can be modeled by effective thinning. This distribution can be used to predict low voltage dielectric reliability parameters.

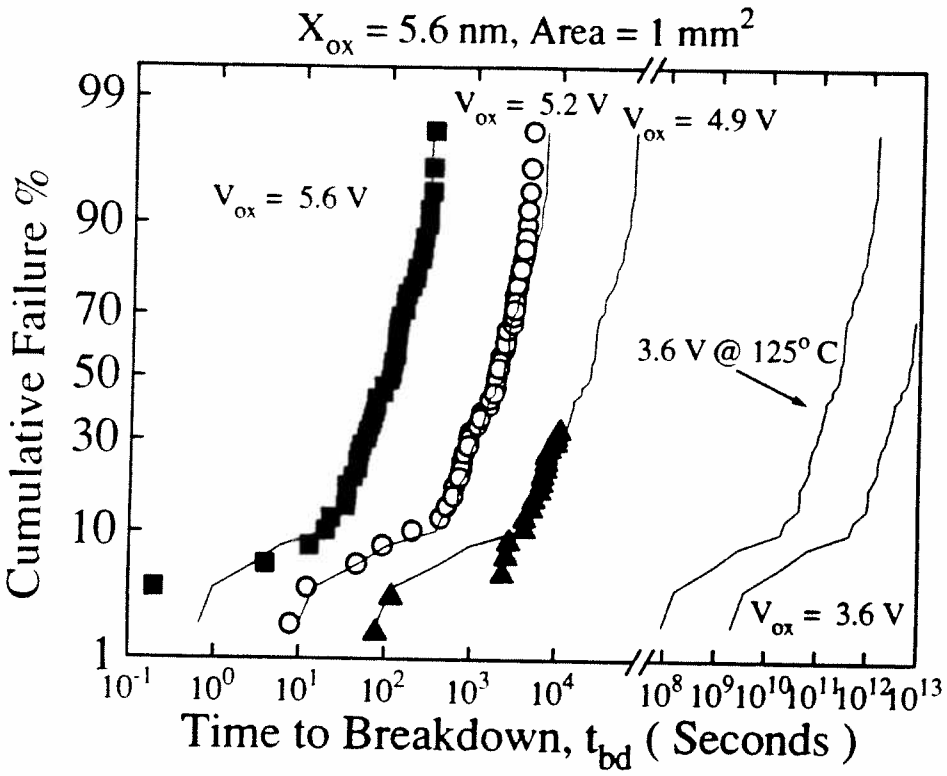


Fig. 5.6 Shows that the averaged high field statistical defect lifetime distribution can be used to predict the lifetime distribution for low voltages and different temperatures. Temperature acceleration is determined from the activation energy of $\mathbf{Q_p}$ in **Fig. 4.3**.

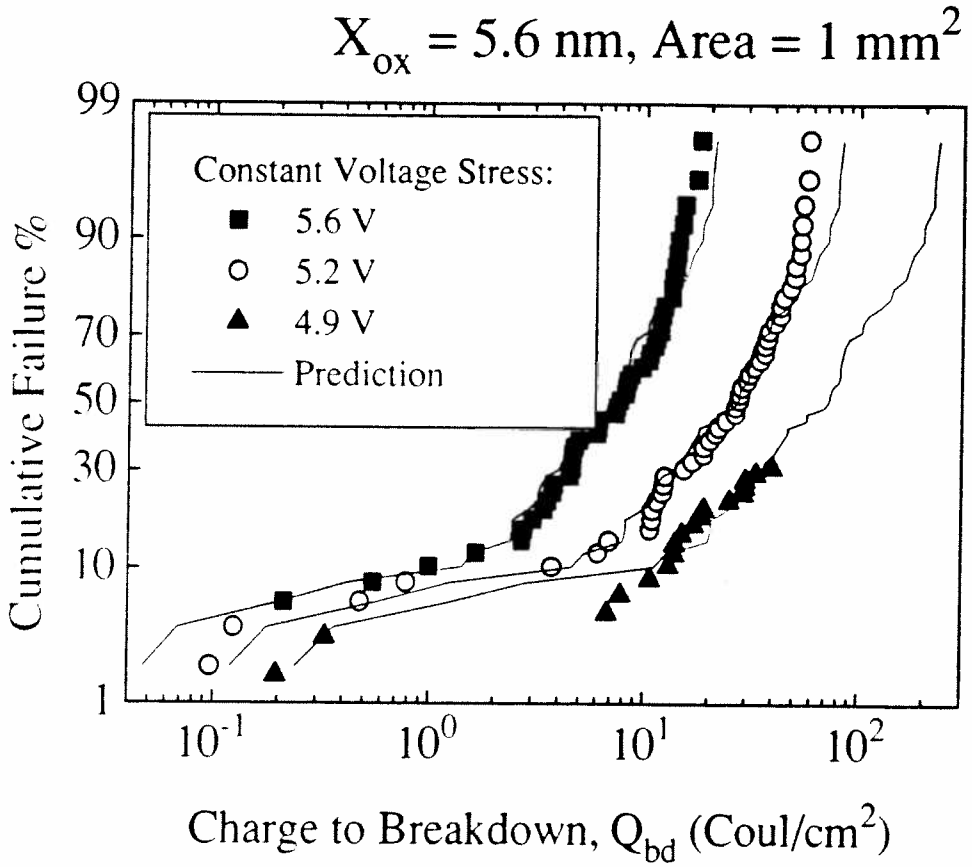


Fig. 5.7 Illustrates that the $\mathbf{X_{eff}}$ defect distribution (**Fig. 5.5**) can also predict the statistical $\mathbf{Q_{bd}}$ distribution.

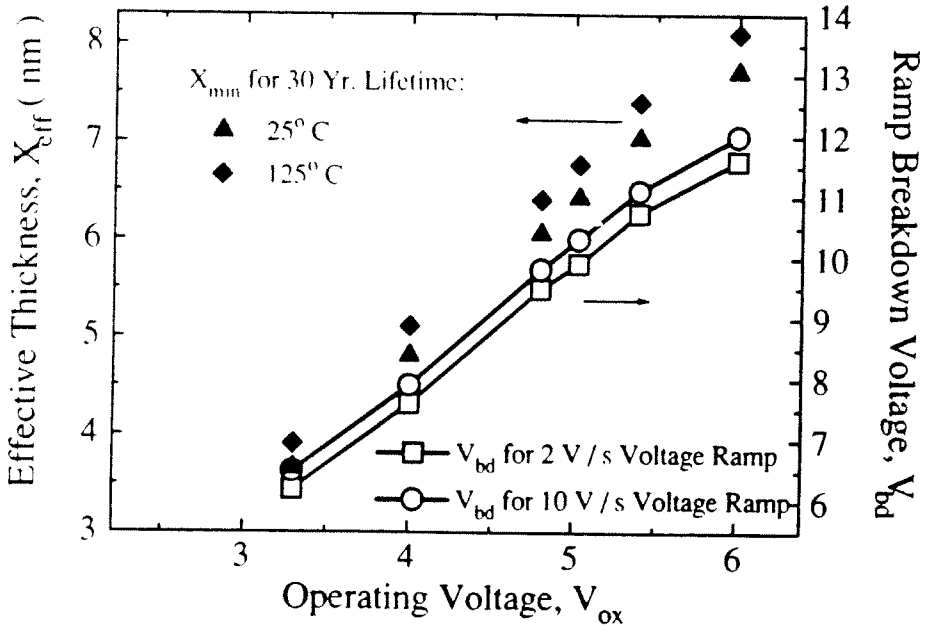


Fig. 5.8 Allows for determination of minimum effective thickness to guarantee 30 year dielectric lifetime as a function of operating voltage. Indicates minimum ramp breakdown voltages at 25° C necessary to guarantee 30 year lifetime at 125° C.

5.5. References

[5.1] J. C. Lee, I.-C. Chen and C. Hu, "Modeling and Characterization of Gate Oxide Reliability," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2268, 1988.

[5.2] E. Rosenbaum, R. Moazzami and C. Hu, "Implications of Waveform and Thickness Dependence of SiO_2 Breakdown on Accelerated Testing," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications*, p. 214, 1991.

[5.3] B. J. Fishbein and D. B. Jackson, "Performance Degradation of N-Channel MOS Transistors During DC and Pulsed Fowler-Nordheim Stress, *Proceedings of the International Reliability Physics Symposium*, p. 159, 1990.

[5.4] S. J. Wang, I.-C. Chen, and H. L. Tigelaar, "Effects of Poly Depletion on the Estimate of Thin Dielectric Lifetime", *IEEE Electron Device Letters*, vol. 12, no. 11, p. 617, 1991.

[5.5] K. F. Schuegraf and C. Hu, "Hole Injection SiO_2 Breakdown Model for Very Low Voltage Lifetime Extrapolation," *IEEE Transactions on Electron Dev*ices, May, 1994.

Chapter 6

Limit of Substrate Current in Predicting SiO_2 Breakdown

6.1. Introduction

The discussion in Chapters 2, 3, and 4, exploited the correlation between substrate current and charge-to-breakdown, Q_{bd}, characteristic to establish the anode hole injection model. This chapter explores the origin of the substrate current of a MOSFET when the gate oxide undergoes Fowler-Nordheim stress. While the measured substrate current is entirely due to anode hole injection for oxides thicker than 55 Å, tunneling by valence band electrons contributes to the substrate current in thinner oxides. Valence band electron tunneling current is shown to increase with oxide stressing similar to low voltage gate oxide leakage; apparently, both are enhanced by trap-assisted tunneling. This chapter examines these additional contributions to substrate current and shows that prediction of substrate hole current still accurately predicts SiO_2 breakdown despite the role that valence-band electron tunneling plays in limiting the correlation between measurable substrate current and breakdown.

First, the anode hole injection model is briefly reviewed and shown to predict both the oxide breakdown characteristic and the substrate current for oxides thicker than 55 Å. Next, valence-band electron tunneling is shown to be the source of substrate current for oxides thinner than 45 Å, thereby decoupling the link between substrate current and oxide breakdown. Third, in the transition region between 45 and 55 Å, where the anode hole injection and valence band electron tunneling currents are approximately of the same magnitude, trap creation during electrical stress can be observed in the substrate current as valence-band-electron trap-assisted-tunneling. The stress created

traps greatly enhance the tunneling probability of valence band electrons and begins to exceed the anode hole injection current. This is unimportant for X_{ox} thinner than 45 Å as the fresh valence band electron tunneling is large to begin with. However, despite these complications, the anode hole injection model can still predict the thickness and voltage dependence of the oxide charge-to-breakdown characteristic, Q_{bd}.

6.2. Experiment

The devices used in this study were either capacitors or transistors with gate oxide thickness varying between 25 Å and 130 Å. Oxidation process was dry oxidation at 750-950°C. All devices have an in-situ phosphorus doped polysilicon gate deposited at 605°C. The capacitors were on n-type <100> 8-12 Ω-cm substrate. Constant-voltage stressing of capacitors was performed using an HP4140B picoammeter. Constant-voltage hole separation experiments [6.1,6.2] with transistors were performed using an HP4145B parameter analyzer. All the instruments were controlled by a PC using GPIB interface.

6.3. Substrate Current: Anode Hole Injection

The mathematical model development for hole tunneling current was developed in Chapter 2. The quantum efficiency of the hole generation process, given by Eq. (2.1), is inversely proportional to Q_{bd}. Figure 6.1 repeats Figure 2.4, clearly indicating that the hole-generation quantum efficiency remains constant for the duration of the oxide lifetime showing that the hole generation rate is strictly determined by the applied bias. The injected hole fluence increases with time as $Q_{ps}(t) = \int_0^t J_p(t)\,dt$ until a critical hole

fluence, Q_p, is reached, marking the breakdown event. Charge to breakdown, Q_{bd}, is given by Eq. (2.5) as

$$Q_{bd} = \frac{Q_p}{\alpha_p} \exp\left(\frac{\hat{B}}{E_{ox}}\left[\Phi_p(V_{ox})\right]^{\frac{3}{2}}\right) \tag{6.1}$$

Figure 6.2 not only demonstrates the proportionality between Q_{bd} and the inverse of the hole injection quantum efficiency, i.e. $1/\left[J_p/J_n\right]$, but also demonstrates the effectiveness of the anode hole injection theory postulated by Eqs. (2.1) and (6.1). Figure 6.3 shows that the critical hole fluence at breakdown, $Q_p \approx 0.1$ Coul/cm^2, is independent of the stress voltage, while Q_{bd} decreases with increasing stress voltage according to Eq. (6.1) for oxides with thickness between 55 and 100 Å. For thinner oxides, the appearance of valence-band-electron tunneling currents or valence-band-electron trap-assisted-tunneling currents must be condsidered as is done in the following section; however, subsequent sections show that the equations of the anode hole injection model retain their ability to predict thin oxide reliability parameters.

6.4. Substrate Current: Valence-Band-Electron Tunneling

For oxides thinner than 45 Å, the substrate current measured by the hole separation technique results from tunneling by valence band electrons as shown in Figure 6.4 [6.3,6.4]. Figure 6.4 illustrates a situation where electrons may tunnel both from the cathode conduction and valence bands. In the hole separation bias configuration, conduction band electrons are supplied by the source and drain. Tunneling valence band electrons leave behind holes which are driven from the surface by the depletion-region field and collected as substrate current. Valence band electron tunneling is only detectable in such very thin oxides because the tunneling probability of valence-band-

electrons with barrier height of 4.25 eV in thin oxides is sufficiently high to dominate the anode hole injection current as shown in Figure 6.5. Thus, for these thin oxides, the substrate current J_p no longer reflects the flow rate of damaging holes through the oxide nor serves as an useful tool for studying oxide breakdown. Moreover, the time-integrated substrate current $Q_S(t) = \int_0^{t_{bd}} J_S(t)\,dt$ no longer measures the quantity of holes injected into the oxide and is therefore no longer constant (independent of V_{ox}) at breakdown, as shown in Figure 6.6. At $V_{ox} \approx 4.0$ Volts, corresponding to a V_g of 5.1 Volts, where anode hole injection is expected to dominate (see Figure 6.5), Q_S finally reaches the 0.01 Coul/cm^2 critical hole fluence shown in Figure 6.3. In Figure 6.6, Q_{bd} still follows Eq. (6.1) while Q_p is taken to be constant at 0.01 Coul/cm^2. Therefore, although it is difficult or impossible to determine Q_p by integrating the substrate current as for thicker oxides, the anode hole injection model is still able to predict the very thin oxide breakdown characteristics provided we choose Q_p as a fitting parameter to match $Q_{bd}(V_{ox})$ with Eq. (6.1). Figure 6.7 compares values of Q_p measured using hole separation (where meaningful) and those deduced fitting capacitor breakdown data to Eq. (6.1), showing that Q_p decreases with decreasing oxide thickness, reflecting a weakened hole immunity for thinner oxides.

6.5. Substrate Current: Valence-Band-Electron Trap-Assisted-Tunneling

For transistors with oxide thickness between 45 Å and 55 Å, the substrate current measurement allows direct observation of trap creation during high-field stress. Trap-assisted-tunneling of valence band electrons arises, dominating the anode-hole-injection substrate current (which was shown to remain constant with time in Figure 6.1) measured from a fresh device. Figure 6.8 displays the time evolution of the quantum

hole generation efficiency as in Figure 6.1 for a 49 Å device. For higher bias, the ratio J_p / J_n is time invariant, following the anode hole injection model. For lower fields, the substrate current increases during stress due to an additional current contribution, the tunneling of valence band electrons through stress induced traps. Figure 6.9 compares electron and substrate currents of a device before and after Fowler-Nordheim stress where a stress-induced leakage component is seen not only in the electron tunneling current [6.5], but also in the substrate current. The stress-induced leakage "tail" in the electron tunneling characteristic results from trap-assisted-tunneling [6.6,6.7] where electrons direct tunneling into a band of traps [6.8] near the cathode observe a Fowler-Nordheim type current relationship with an effectively reduced barrier height [6.5,6.9]. The trap-assisted-tunneling of valence band electrons explains the increase of hole injection quantum efficiency in Figure 6.8. Only the lower bias levels exhibit this increase since the anode hole injection component dominates the stress induced leakage component at higher biases as shown in Figure 6.9.

Again, Eq. (6.1) accurately predicts the Q_{bd} characteristic as shown in Figure 6.10. Since the time-integrated substrate current, Q_S, measures the substrate current contributions of both the anode hole injection and the valence band electron trap-assisted-tunneling processes, it is not constant for all biases. However, using the high field Q_p of Figure 6.10 (0.01 Coul/cm^2) in Eq. (6.1) enables modeling the Q_{bd} characteristic in Figure 6.10.

6.6. Oxide Reliability Prediction: Anode Hole Injection Model

Figure 6.11 summarizes the power of the anode hole injection model to predict the bias dependence of thin oxide Q_{bd} for oxides between 25 and 100 Å. The rapidly rising Q_{bd} behavior in thinner oxides (Q_{bd} exceeds 10^7 Coul/cm^2 for 25 Å oxide at 2.4 V)

can be attributed to the fact that the hot hole energy, E_{gain}, becomes quite small and more sensitive to V_{ox} when scattering becomes weaker. A small hot hole energy means many electrons may flow through the oxide (large Q_{bd}) before the critical hole fluence of anode-injected holes, Q_p, is reached.

6.7. Summary and Conclusions

This chapter has examined the role of substrate current arising during Fowler-Nordheim stress in predicting oxide reliability. The correlation between this substrate current and oxide breakdown for oxide thickness greater 55 Å leads to the development of the anode hole injection model. This chapter shows that additional mechanisms involving the tunneling of valence band electrons obscure the correlation between measured substrate current and oxide breakdown for oxides thinner than 55 Å. Nevertheless, the anode hole injection model is still able to predict oxide breakdown for oxides thinner than 55 Å, establishing the low voltage oxide breakdown model.

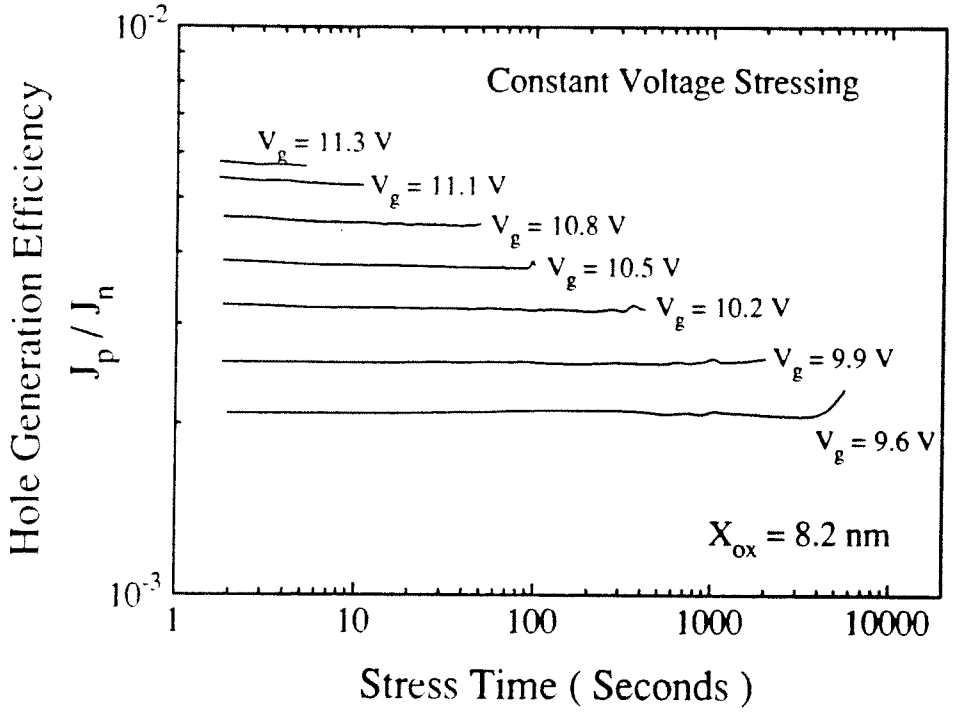


Fig. 6.1 Time evolution of the quantum hole generation efficiency, J_p/J_n, during constant voltage stress of a 82 Å oxide, clearly demonstrating that the applied bias determines this efficiency.

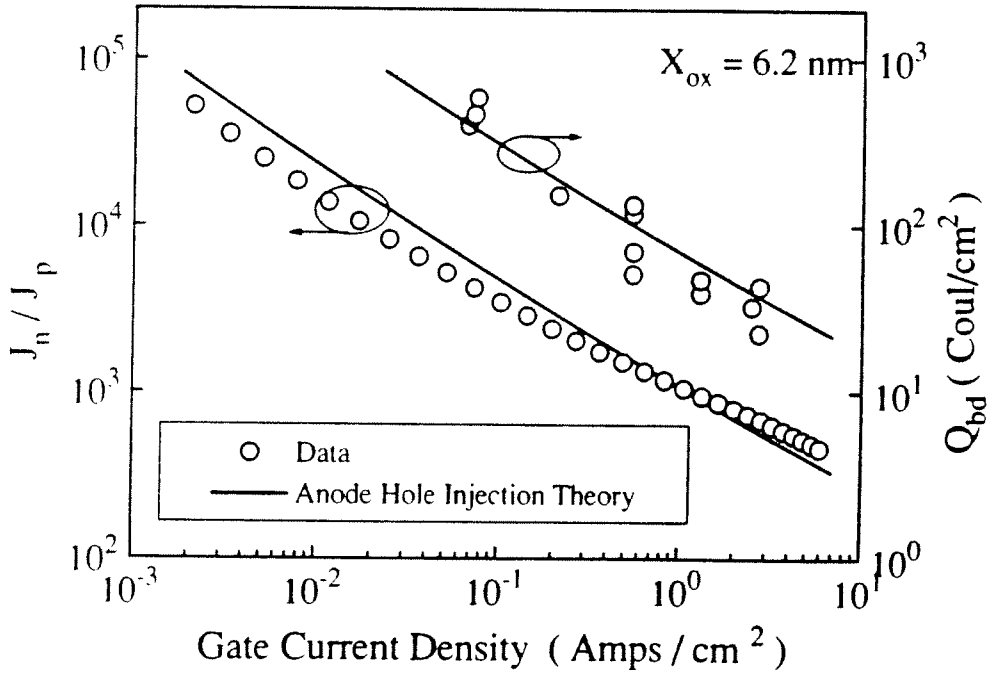


Fig. 6.2 Shows that both Q_{bd} and the inverse of the hole injection quantum efficiency, $1/\left[J_p/J_n\right]$, follow the same bias dependence.

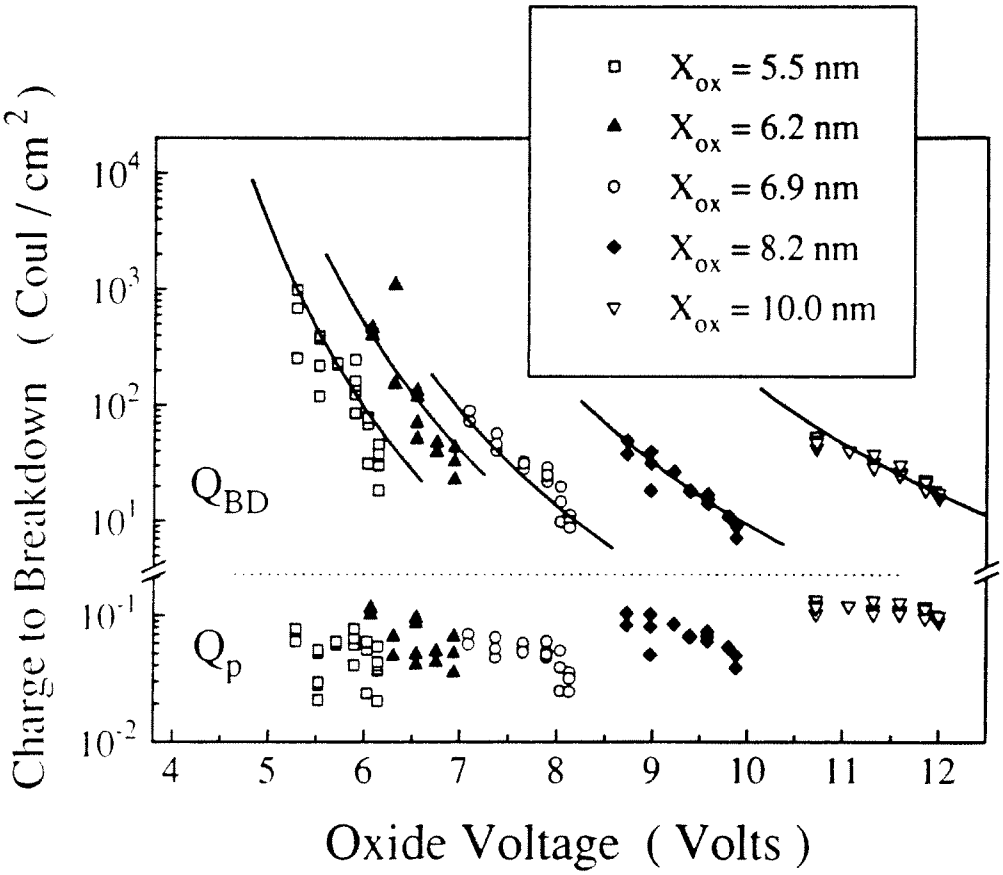


Fig. 6.3 Shows that while the critical hole fluence at breakdown, Q_p, is almost independent of the stress voltage, the charge-to-breakdown, Q_{bd}, decreases with increasing stress voltage according to the anode hole injection model.

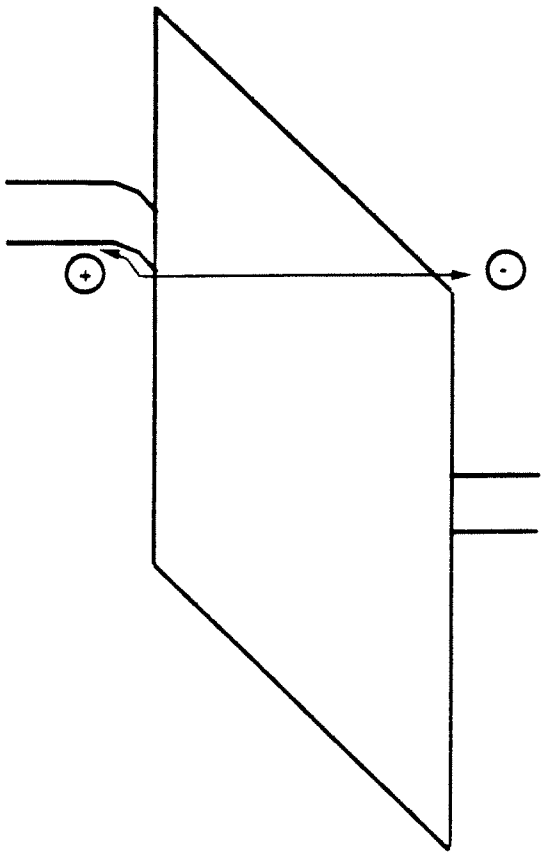


Fig. 6.4 Band diagram illustrating the situation where electrons may tunnel from both the cathode conduction and valence bands. The valence band component is directly measurable as **n**MOSFET substrate current for oxides thinner than 45 Å.

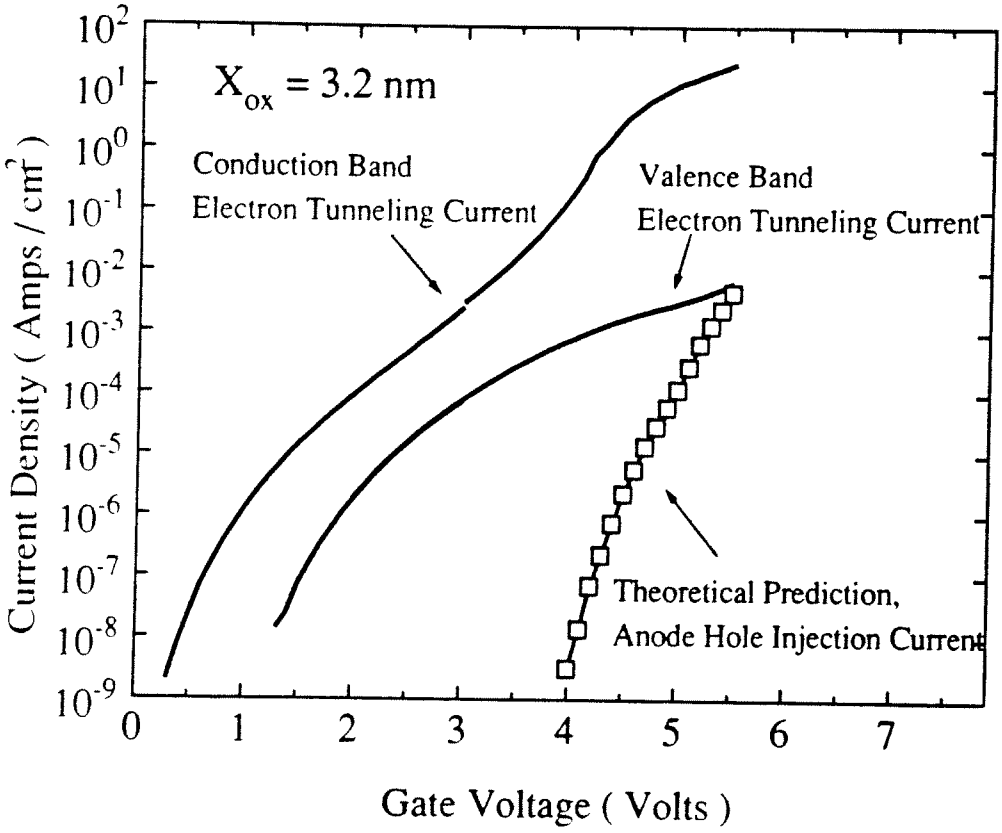


Fig. 6.5 Valence band electron tunneling is shown to be much larger than the amount of anode hole injection current predicted from theory in a 32 Å sample.

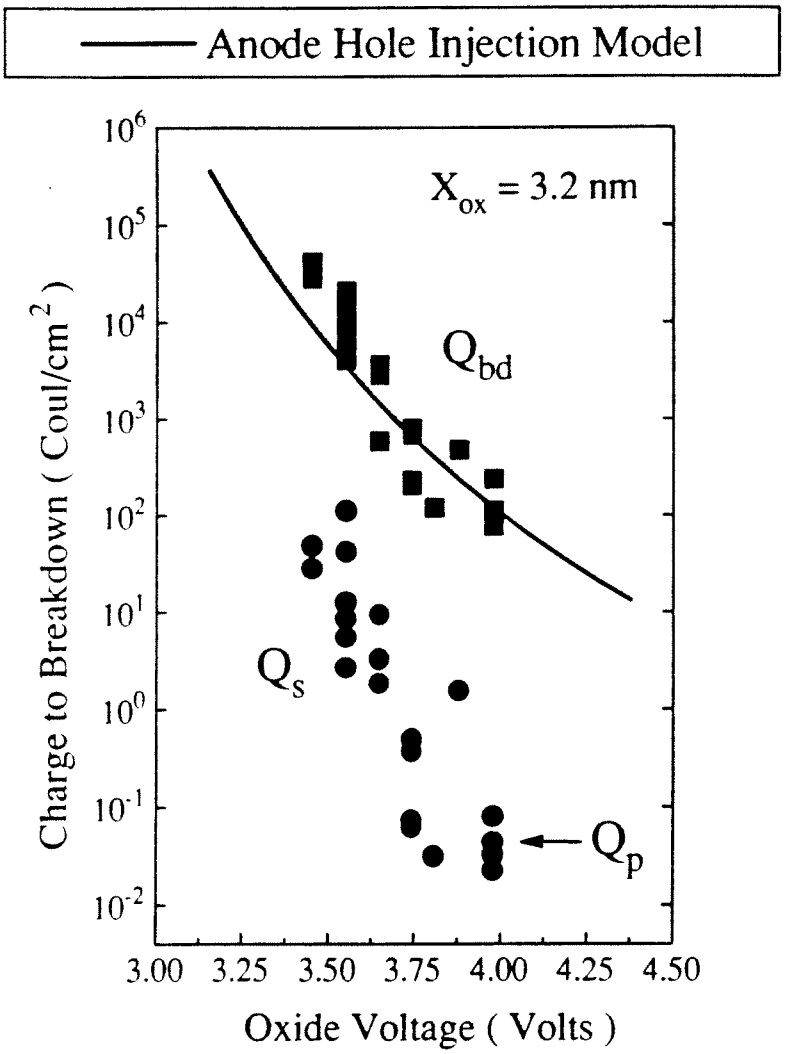


Fig. 6.6 Charge-to-breakdown, Q_{bd}, follows the prediction of the anode hole injection model although the time-integrated substrate current at breakdown, Q_S, is no longer constant because the substrate current reflects valence-band-electron tunneling rather than anode hole injection. At higher stress voltage, Q_S approaches the Q_p value needed as a fitting parameter for the Q_{bd} characteristic because the valence-band-electron tunneling and anode hole injection currents become equal.

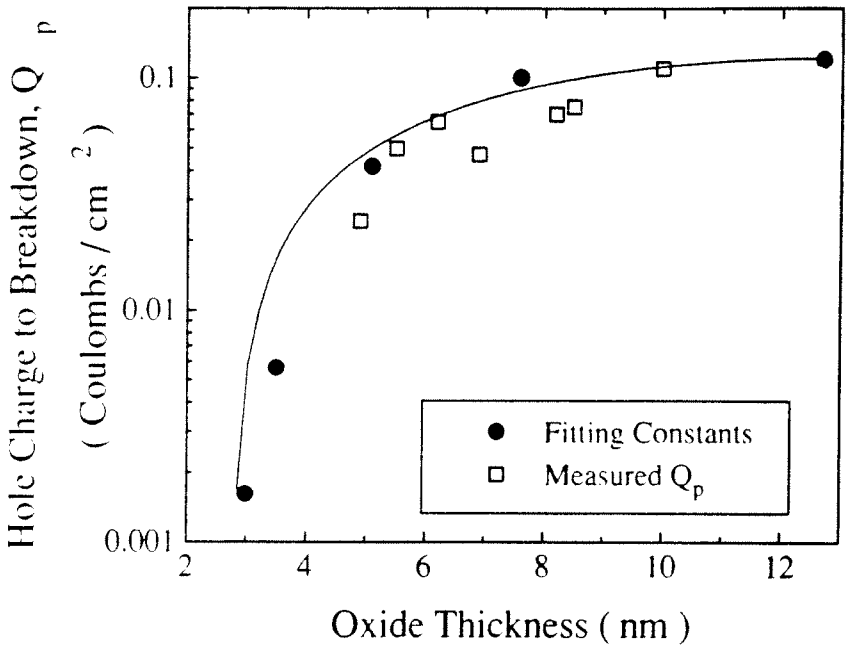


Fig. 6.7 Comparison of experimentally measured Q_p values using hole separation and those deduced from fitting capacitor breakdown data.

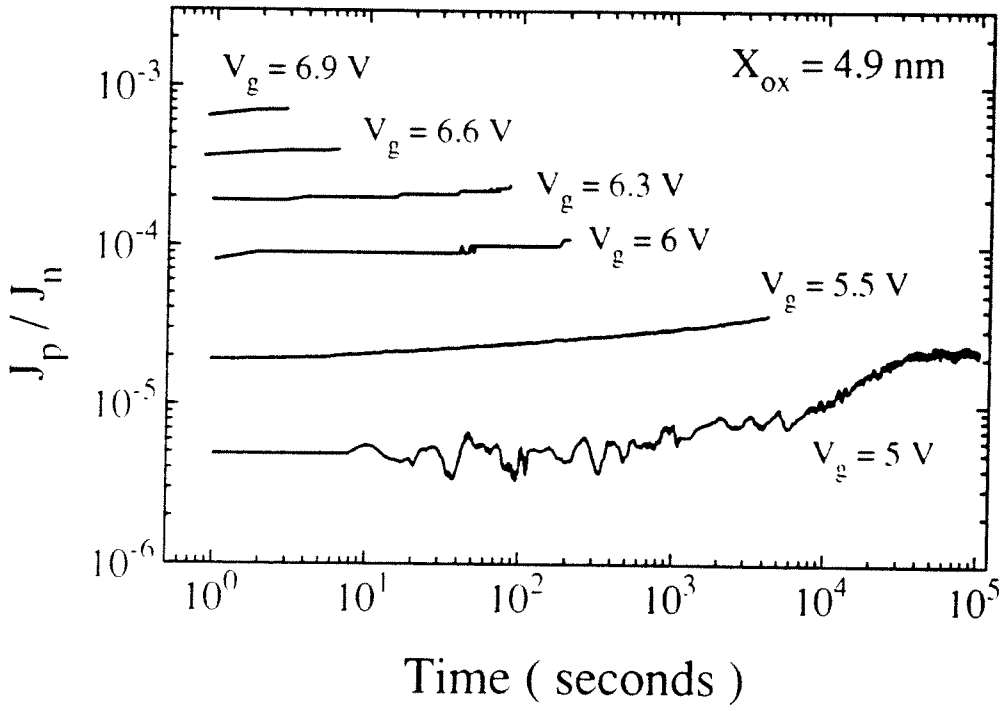


Fig. 6.8 Illustrates the time evolution of J_p / J_n for a 49 Å oxide. At higher bias voltages, this ratio remains time invariant as in thicker oxides. At lower voltages, the J_p increases due to the additional component of valence band tunneling through stress induced traps.

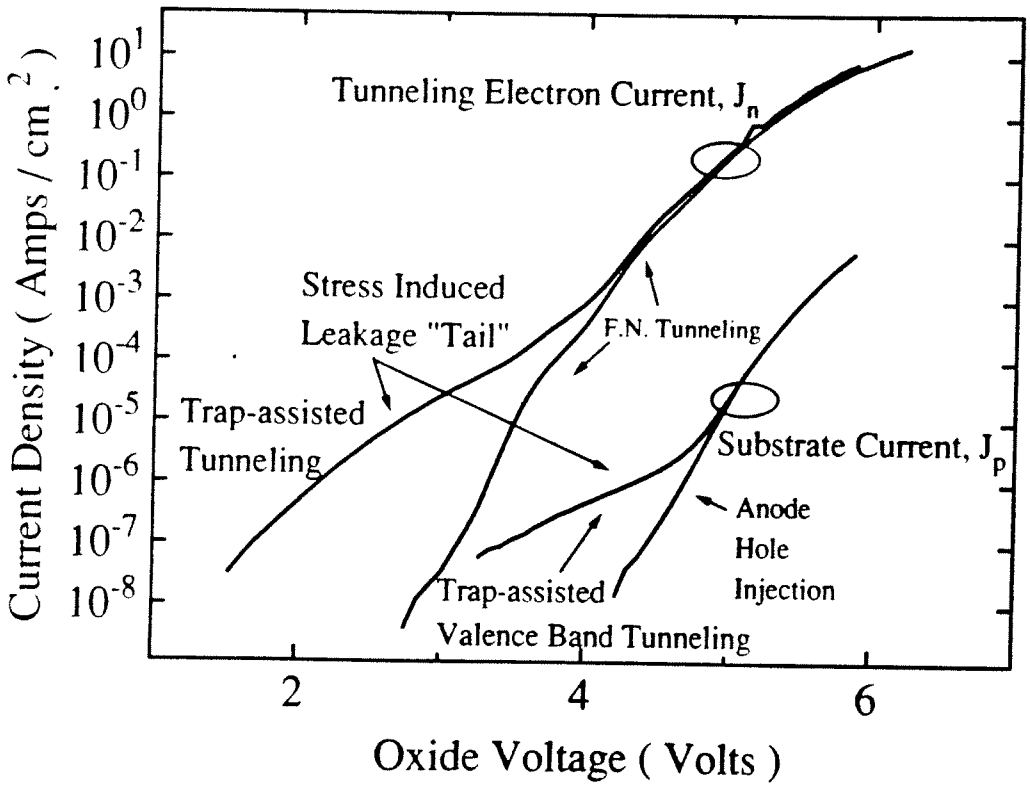


Fig. 6.9 Illustrates the appearance of stress-induced leakage "tails" in both the tunneling electron and substrate current characteristics. The tail in the conduction-band tunneling electron characteristic is due to trap-assisted tunneling. The tail in the substrate current characteristic is due to trap-assisted valence-band electron tunneling.

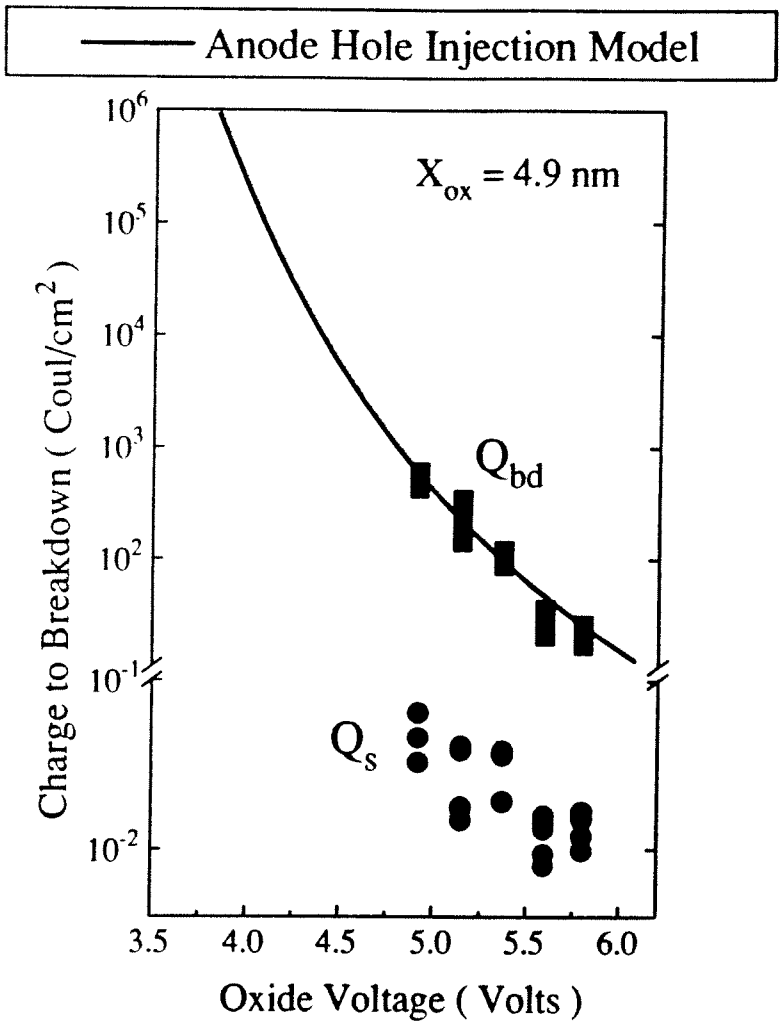


Fig. 6.10 Charge-to-breakdown, Q_{bd}, follows the prediction of the anode hole injection model although the time-integrated substrate current at breakdown, Q_S, increases at lower oxide voltages due to the additional contribution by the valence-band electron trap-assisted tunneling process to the substrate current. At higher stress voltage, Q_S equals the hole charge, Q_p, needed to induce breakdown because the anode hole injection current exceeds this additional contribution.

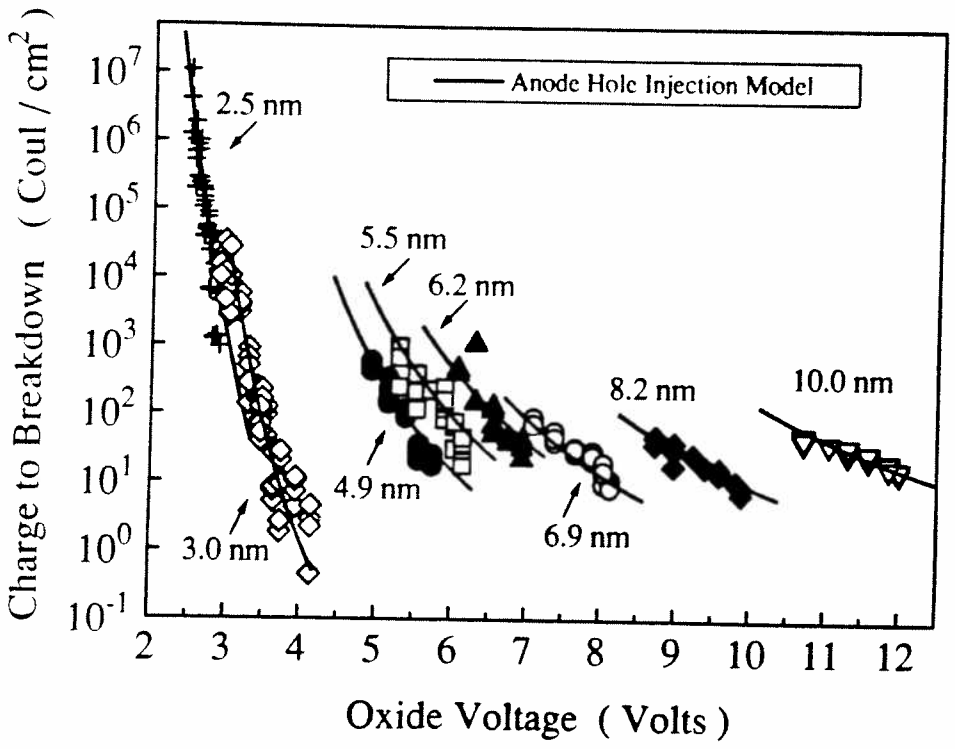


Fig. 6.11 Voltage dependence of thin oxide charge to breakdown, demonstrating ability to extrapolate to low voltages.

6.8. References

[6.1] I.-C. Chen, S. E. Holland and C. Hu, "Oxide Breakdown Dependence on Thickness and Hole Current - Enhanced Reliability of Ultra-Thin Oxides," *International Electron Devices Meeting*, p. 660, 1986.

[6.2] B. Eitan and A. Kolodny, "Two Components of Tunneling Current in Metal-Oxide-Semiconductor Structures," *Applied Physics Letters*, vol. 43, no. 1, p. 106, 1983.

[6.3] C. Chang, C. Hu, and R. W. Brodersen, "Quantum Yield of Electron Impact Ionization in Silicon," *Journal of Applied Physics*, vol. 57, no. 2, p. 302, 1985.

[6.4] Z. A. Weinberg and M. V. Fischetti, "Investigation of the SiO_2-Induced Substrate Current in Silicon Field-Effect Transistors," *Journal of Applied Physics*, vol. 57, no. 2, p. 443, 1985.

[6.5] P. Olivo, T. N. Nguyen, and B. Ricco, "High-Field-Induced Degradation in Ultra-Thin SiO_2 Films," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2259.

[6.6] R. Rofan and C. Hu, "Stress-Induced Oxide Leakage," *IEEE Electron Device Letters*, vol. 12, no. 11, p. 632, 1991.

[6.7] R. Moazzami and C. Hu, "Stress-Induced Current in thin Silicon Dioxide Films," *International Electron Devices Meeting*, p. 139, 1992.

[6.8] I. Lundstrom, S. Christensson, and C. Svensson, "Carrier Trapping Hysteresis in MOS Transistors," *physica status solidi (a)*, vol. 1, p. 395, 1970.

[6.9] B. Ricco and M. Ya. Azbel, "Physics of Resonant Tunneling. The One-Dimensional Double-Barrier Case," *Physical Review B*, vol. 29, no. 4, p. 1970, 1984.

Chapter 7

Summary, Conclusions, and Future Work

7.1. Introduction

During the course of electrical stress, oxide damage manifests itself as generated interface traps [7.1], generated bulk traps [7.2,7.3,7.4,7.5,7.6], low level stress induced leakage [7.7,7.8], and, finally, catastrophic breakdown. The first part of this chapter attempts to integrate these many diverse reported observations of oxide breakdown into a coherent picture of the wearout process leading to oxide breakdown. Next, the properties of "improvements" to thermal-oxide technology are discussed. This section shows that although these improvements may improve oxide quality locally, only the stacked gate oxide technology is able to consistently deliver improvements to oxide defect density. The final section briefly discusses the needs for future dielectrics which offer high capacitance, reliability, manufacturability and low leakage.

7.2. Breakdown Process

Many studies have correlated a particular observation of these manifestations to oxide breakdown and drew conclusions about the cause of oxide wearout. For instance, Harari [7.6] noted a fixed amount of electron trapping before breakdown and reasoned that electron trapping establishes high internal fields which lead to material rupture. This model fails to account for the role of holes in the breakdown process. Hot-hole injection by non-Fowler-Nordheim processes [7.9,7.10,7.11,7.12] such as junction avalanche show that hole fluence, not electron fluence, is correlated to oxide breakdown. Hole fluence is also much more deleterious to oxide integrity than electron

fluence in that the oxide can sustain a fluence of about 100 times more electrons than holes. (This agrees with Figure 2.4 that one hole is injected into the oxide for every several hundred electrons tunneling through the oxide.) Others [7.13,7.14] have suggested that current density can be enhanced by hole trapping and this process leads to oxide wearout. This scenario fails to account for the fact that trapped holes are converted into electron traps [7.2,7.15] and that hole trapping eventually leads to net negative charge trapping in the oxide.

Others have shown the correlation between interface trap generation rate and breakdown under static Fowler-Nordheim stress [7.1,7.16], that is, the higher the N_{it} generation rate, the lower the charge to breakdown. Yet, dynamic bipolar stress generates considerably more interface states than static stress, while also yielding a longer, not shorter, breakdown lifetime [7.17,7.18]. Moreover, Chen [7.9] showed that large interface state generation under channel hot-carrier stress did not degrade oxide breakdown integrity. These experiments serve to argue against a causal relationship between interface traps and breakdown.

Breakdown appears to be a localized process triggered by a local increase in stress current density. Although the small area of this locally large current density precludes direct measurement of current enhancement, the study of low-level stress-induced leakage [7.7,7.8,7.19] seems to indicate a conduction mechanism of trap assisted tunneling through neutral traps. Such neutral traps [7.2] created by a two step process of hole trapping and subsequent recombination with electrons [7.3,7.4,7.5] may serve to explain the localized increase in current density, be it through the mechanisms of charge assisted tunneling [7.13], trap-assisted tunneling [7.20] or a closely related resonant tunneling [7.21]. Since hole fluence leads to the generation of bulk and interface electron traps [7.2,7.3,7.4,7.5], it is not surprising that there is often, though

not always, a correlation between oxide breakdown and electron trapping or interface trap generation. In other words, holes in the oxide are the agent responsible for electron trap and interface trap generation as well as breakdown. There, a casual observation might conclude that breakdown is correlated with any of these other quantities. Recent models emphasizing the motion of hydrogen in trap generation and breakdown can also be linked to hole injection [7.1,7.22,7.23]. Another model proposes that bond breaking by the energetic tunneling electrons forms a conductive path from anode to cathode [7.24]. This may be describing the same hole-induced damage and leakage model espoused here.

Others believe that heating is the final cause of oxide breakdown. A post-breakdown failure analysis would reveal that breakdown is a point of localized melting [7.25], reflecting a large energy discharge at breakdown. However, since the discharge process is virtually instantaneous, compared with the damage accumulation time, the oxide breakdown lifetime is determined by the accumulation of the critical damaging hole fluence. Discharge is but the firework display marking the end of the oxide lifetime.

The anode hole injection model includes and integrates all these diverse and valid observations about the breakdown process. Breakdown is a two-step process. The first part is a long process, possibly spanning years, where the oxide is slowly damaged under electrical stress. The second step is a very short runaway event, on the order of perhaps a microsecond, where a rapid final acceleration of damage due to electrical and/or thermal run-away leads to the formation of a permanent conductive path through the oxide. The injection of holes into the oxide is a precursor to oxide breakdown. Hole injection and trapping lead to the generation of both bulk electron and interface traps and possibly the release of hydrogen [7.2,7.3,7.4,7.5,7.15,7.23].

Therefore, hole injection, electron trapping, interface trapping, and perhaps even hydrogen motion can all be correlated to oxide wearout unless special situations such as AC stress and hot hole injection are considered. The electron traps, so produced, can enhance current density through the oxide. The enhanced current density generates more damage. This positive feedback process causes destructive breakdown to occur at a weak spot.

7.3. Oxide Technology Developments

This section reviews research concerning potential oxide technology improvements to dry thermal oxide with respect to oxide charge trapping and oxide breakdown characteristics. First, the effects of oxide "hardening" by chlorination, fluorination, and nitridation on bulk and interface charge trapping are presented. It is shown that incorporation of these impurities can indeed "harden" an oxide to the effects of high-field stress intrinsically. However, these processing "improvements" have yet to display the reduction in oxide defectivity, as measured by either defect density or effective thickness, necessary to spearhead their introduction into semiconductor manufacturing. Next, discussion of a promising technology for improving oxide defectivity, that of the stacked gate oxide, shows that oxide defectivity can be improved dramatically by "filling" oxide defects with a thin CVD (chemical vapor deposition) oxide.

7.3.1. Bulk and Interface Charge Trapping: Incorporation of Cl, F, and N in SiO_2

This section reviews the development of "hardened" oxide technology, concentrating on the improvement of the gate oxide's robustness to bulk charge trapping and interface trap generation. Channel hot carrier stress results primarily in interface trap generation

in **n**MOSFETs[7.26] and bulk electron trapping in **p**MOSFETs[7.27]. Temperature bias stress results in the generation of interface traps [7.28]. High-field stress results in interface and bulk trapping [7.14]. Radiation stress results in interface and bulk trapping [7.23]. Process development in the form of impurity incorporation has recently been shown to alleviate the susceptibility of an oxide to these trapping mechanisms.

Impurity incorporation of fluorine and chlorine may improve oxide charge trapping. For example, Figure 7.1 shows that ion-implanted fluorine is able to improve oxide resistance to hot carrier and Fowler-Nordheim tunneling stress by 3 to 5 times [7.29]. Figure 7.2 shows that chlorine introduction by annealing in a TCA ambient with much lower concentration than usually used provides an optimal point for the suppression of interface state generation by hot carrier stress [7.30]. The hardness improvement for chlorine and fluorine incorporation is thought to result from strain relaxation in the SiO_2, the formation of Si-F or Si-Cl bonds in place of Si-H and strained Si-O bonds. However, excess incorporation of chlorine and fluorine results in the creation of non-bridging oxygen centers resulting in a reduction of oxide hardness [7.29]. Therefore, these competing mechanisms exhibit an optimal concentration of incorporated chlorine or fluorine for improving oxide hardness. Whether a 3 to 5 times improvement in trapping rate can justify this added process complexity is unclear. However, fluorine incorporation may be a side benefit of tungsten-CVD deposition in IC processing [7.31].

Nitridation also offers promise for hardening the silicon-silicon dioxide interface. Figure 7.3 shows the dramatic improvement in channel hot carrier immunity achievable with reoxidized NH_3 annealed MOSFETs. Unfortunately, hydrogen incorporation during the NH_3 anneal leads to increased charge trapping [7.32]. Nitridation in N_2O

[7.33,7.34] has created a lot of excitement because of the elimination of hydrogen in the anneal process and the promise of good process control and reproducibility. However, nitridation in N_2O has shown lesser improvements (similar to fluorine incorporation) in hot carrier immunity as shown in Figure 7.4 [7.35].

7.3.2. Stacked Thermal/CVD SiO_2

These alternative hardened gate oxide technologies have been shown to reduce the charge trapping and trap generation rates of the silicon dioxide. The small improvement in intrinsic oxide lifetime due to fluorine, chlorine, or nitrogen incorporation is welcome but insignificant when compared with the overwhelming impact of oxide defects on oxide lifetime. Only the composite stacked thermal/CVD gate oxide [7.36] has been able to consistently demonstrate a significant reduction in defect density desired to deliver an improvement in IC yield and reliability. Stacked thermal/CVD oxide films are fabricated by growing a thermal oxide film followed by CVD oxide deposition, densification, and, optionally, reoxidation. Table 7.1 [7.37] shows how the defect density of the stacked films excels over conventional thermal oxide, especially for thin oxides. Table 7.2 shows the wafermap of breakdown voltages for a 100 Å thermal oxide compared with those for a 200 Å stacked gate with 100 Å initial thermal oxide and 100 Å of CVD oxide, showing that the CVD layer not only boosts the breakdown voltages according to the additional 100 Å oxide layer, but actually repairs the defects in the original 100 Å oxide to achieve a very uniform breakdown map [7.38]. In other words, those defects with 1 Volt-breakdown voltage did not become 10 Volt-breakdown defects after the 100 Å CVD oxide is added, but became perfect 18 Volt-breakdown oxide. This supports a picture in which oxide

defects are very small in size so that a thin CVD oxide can "fill" and "planarize" the "pinholes." The origin of the defects may be micropores [7.36].

The optimization of stacked gate composition depends on the defect densities of the thermal oxide and CVD oxide layers, that is, effective defect density reduction requires the defects of one layer to compensate those in the other [7.39]. Figure 7.5 shows the defect density of 140 Å (total) stack oxide versus the CVD oxide thickness. Notice the large decrease in defect density from zero CVD oxide to 25 Å CVD oxide thickness. This shows that even a 25 Å thin layer of CVD oxide suffices to fill thermal oxide micro-pores. Moreover, since a very thin layer is sufficient to fill pores, stacked oxide technology should be scalable merely by reducing the thickness of the thermal oxide layer. Stacked oxide may be practical even for 50 Å gate oxides. Uniformity of CVD film thickness is expected to be good. After all, deposited (stacked) dielectrics are widely used in DRAM's with only 50 Å of equivalent oxide thickness. Other advantageous features of this technology include superior hot carrier immunity, tighter control of MOSFET parameters, and resistance to process-induced damage [7.40]. Figure 7.6 shows that the stack oxide technology provides more robustness to process induced damage caused by severe via etch conditions.

7.4. Future Dielectric Needs for Silicon MOS Technology

The enormous success of the microelectronics industry rests on the ability to constantly scale both channel length as determined by the state-of-the-art in lithography tools and gate oxide thickness as determined by reductions in oxide defect density. The limits to oxide scaling presented in Chapter 2, as summarized by Figure 2.15, indicate that the silicon dioxide MOS system may reach its technology maturity around the year 2009 [7.41], as gate oxide thickness approaches 40 Å, with a characteristic feature size

(channel length) of 90 nm. Interestingly, inherent scaling limitations on lithography-limited channel length are not observable even for channel lengths as short as 40 nm[7.42]. The attainment of good device characteristics, for instance, low off-state leakage and high drive current, necessitates a gate-oxide dielectric thinner than 40 Å. Therefore, scaling appears to be fundamentally limited by the realization of a low-leakage gate dielectric offering a capacitance whose equivalent oxide thickness ($X_{equivalent} = X_{dielectric}\ \varepsilon_{oxide} / \varepsilon_{dielectric}$) is less than 40 Å. However, relaxing the constraint on acceptable leakage current through the gate oxide, may realize integrated circuit technologies with oxide thickness below 40 Å. In fact, this leakage constraint could be relaxed if device parameters such as threshold voltage, subthreshold swing, mobility, etc. prove to be stable in the presence of the large direct tunneling current densities seen in thin oxides at low voltages. This question promises controversy and fertile ground for further research. Ultimately, however, the scaling trend will exhaust the capabilities of the silicon MOS system.

Engineering progress in dynamic random access memory (DRAM) technology serves a glimpse into the future. A DRAM is a highly repetitive structure consisting of two essential device components: a MOSFET transistor with shortest channel length achievable by current lithographical capability and a capacitor which must provide about 30-40 femtocoulombs per memory cell. This capacitance requirement portends the evolution of the MOS technology. In order to realize this high capacitance within a very small area, DRAM manufacturers have followed a variety of different strategies to circumvent the 40 Å oxide thickness scaling limitation. Most current strategies provide the necessary capacitance using a dielectric thickness at the fundamental limit (about 50 Å equivalent thickness by means of an oxide-nitride-oxide sandwich layer) while increasing capacitor area through the use of three dimensional structures. Alternative

strategies have focused on the development of high dielectric constant materials such as tantalum pentoxide or barium strontium titanate [7.43]. Tantalum pentoxide has failed to find wide spread acceptance due to difficulties in its manufacturability and its relatively low dielectric constant of only 25. Although this dielectric constant is about six times larger than silicon dioxide, the success of three dimensional capacitor structures has been able to successfully thwart the considerable technological challenge of introducing a new material (and potential contaminant) into silicon manufacturing environments. However, the complexity of these three dimensional structures has been increasing dramatically in order to continue providing the required 30 femtocoulombs within an ever-smaller cell area. Therefore, materials of an even higher dielectric constant (approximately 1000), such as BST and PZT are subjects of active research. Successful implementation of such a high dielectric constant material may relieve the DRAM industry from its chronic worry concerning the realization of reliable, manufacturable and scalable capacitor structures. Moreover, successful implementation of a high dielectric constant material as a DRAM capacitor may lead to the development of transistor structures utilizing this material, thereby overcoming the inherent fundamental hurdle to device scaling posed by the limits of silicon dioxide as gate dielectric material.

7.5. Summary and Conclusions

This chapter has discussed an integrative model for the many diverse and seemingly contradictory observations concerning correlations between oxide wearout and its eventual breakdown. Additionally, technological improvements to the "thermal" oxide were reviewed, showing that oxides "hardened" to trap generation exhibit superior local characteristics, but do not necessarily provide lower defectivity, as does stacked

gate oxide technology. Finally, the conclusion concerning the scaling limit of silicon dioxide as a dielectric material is discussed, including an assessment of future technology directions for the thin dielectrics needed to realize successful advances in silicon integrated circuit technology.

	Defect Density, D_0 (cm^{-2})	
Oxide Thickness (Å)	Conventional Thermal SiO_2	Stacked Thermal/CVD SiO_2
250	0.35	0.07
210	0.35	0.10
150	0.70	0.17
125	1.00	0.20
100	1.28	0.28

Table 7.1 Demonstrates the dramatic reduction in dielectric defect density achievable with thermal/CVD stack technology. Data for 30 mm^2 structures. Adapted from [7.37].

A) 100 Å Thermal Oxide						B) 200 Å Composite Oxide with 100 Å Thermal Oxide					
11	3	3	1	11	1	18	18	18	18	18	18
11	2	11	11	3	3	18	18	18	18	18	18
11	1	1	1	11	4	18	18	18	18	18	19
11	11	1	3	11	11	18	18	18	18	18	19
1	1	11	11	10	4	18	18	18	18	18	19
3	1	1	11	1	8	18	18	18	18	18	18
3	11	1	2	4	11	18	18	18	18	18	19
2	2	1	1	11	11	18	18	18	18	18	19
1	2	1	2	11	11	18	18	18	18	18	19
11	7	10	11	11	11	18	18	18	18	18	19
2	3	1	2	2	1	18	18	18	18	18	18
5	1	6	1	11	11	18	18	18	18	18	19
11	11	1	3	11	11	19	18	18	18	18	14
11	5	1	1	1	11	19	18	18	18	18	19
4	8	2	1	11		19	19	18	18	19	

Table 7.2 Demonstrates the ability of thermal/CVD stack technology to "repair" or "fill" defects by comparing wafer map data for breakdown voltage for films before and after deposition of 100 Å CVD layer [7.38]. Reproduced with permission by Leo Yau, Intel, Inc.

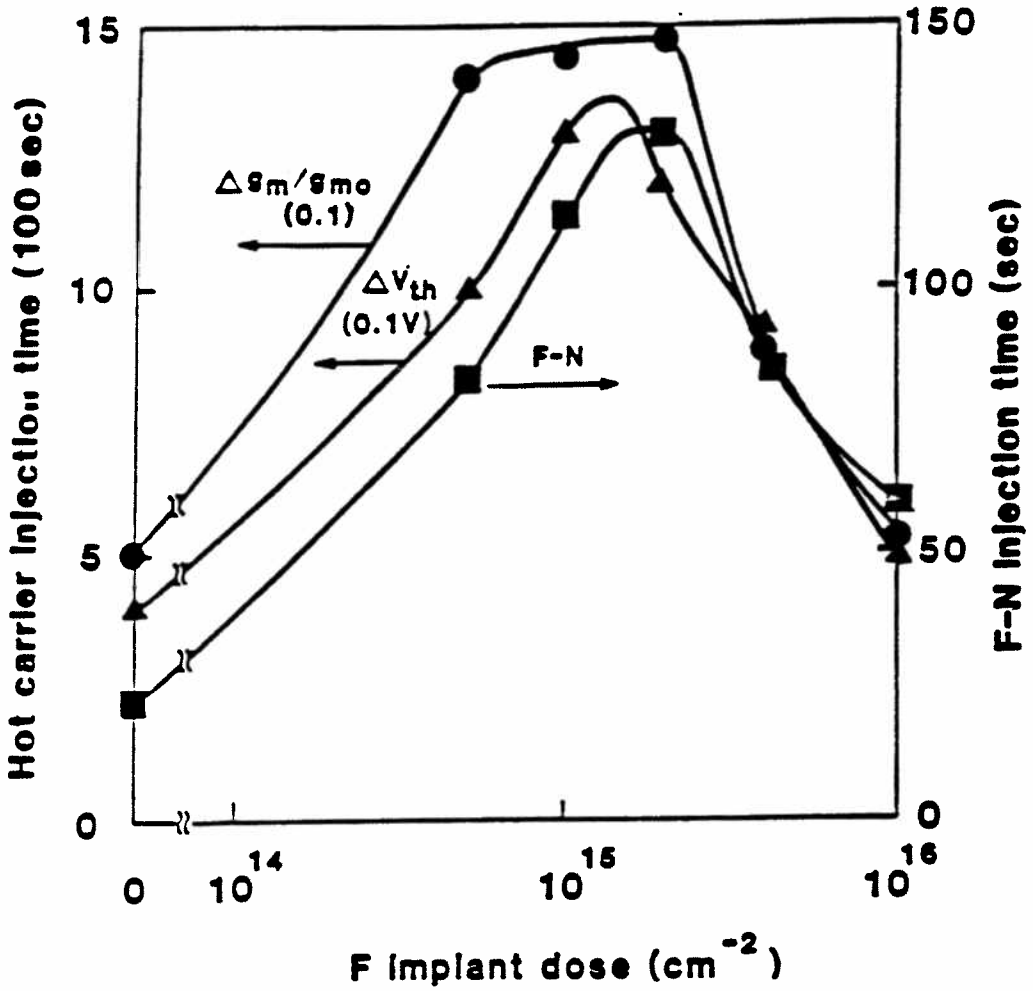


Fig. 7.1 Illustrates the oxide hardening to Fowler-Nordheim and channel hot carrier stress obtained by incorporating fluorine into SiO_2 [7.29].

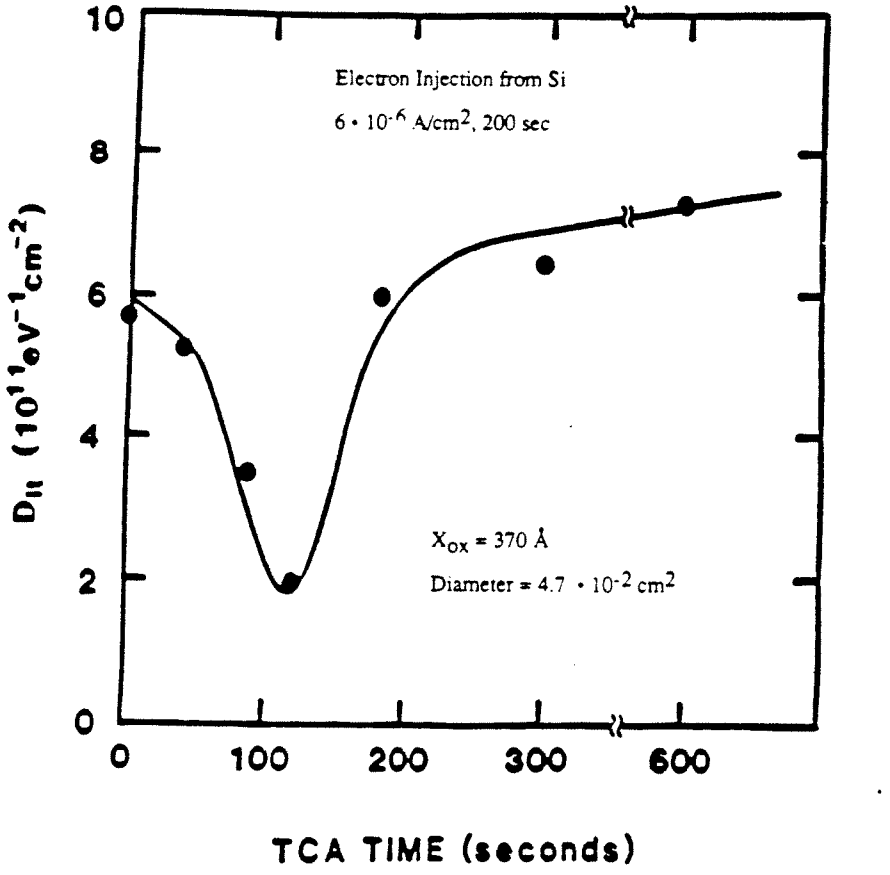


Fig. 7.2 Shows that chlorine introduction into SiO_2 by TCA anneal has an optimum point for suppressing the interface state generation of hot carrier stress [7.30].

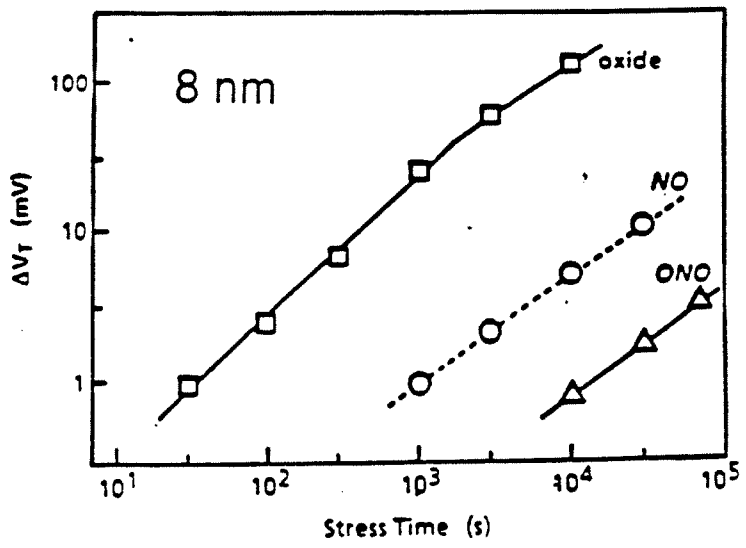


Fig. 7.3 Demonstrates that reoxidized, NH_3 annealed MOSFETs promise dramatic resistance to channel hot carrier stress [7.32].

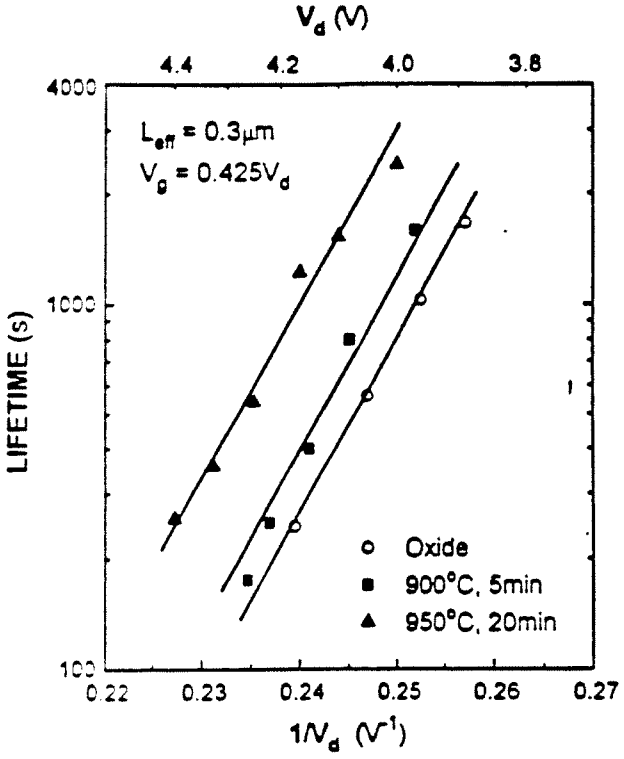


Fig. 7.4 Shows the improvement in MOSFET hot carrier lifetime available for N_2O annealed oxides [7.35].

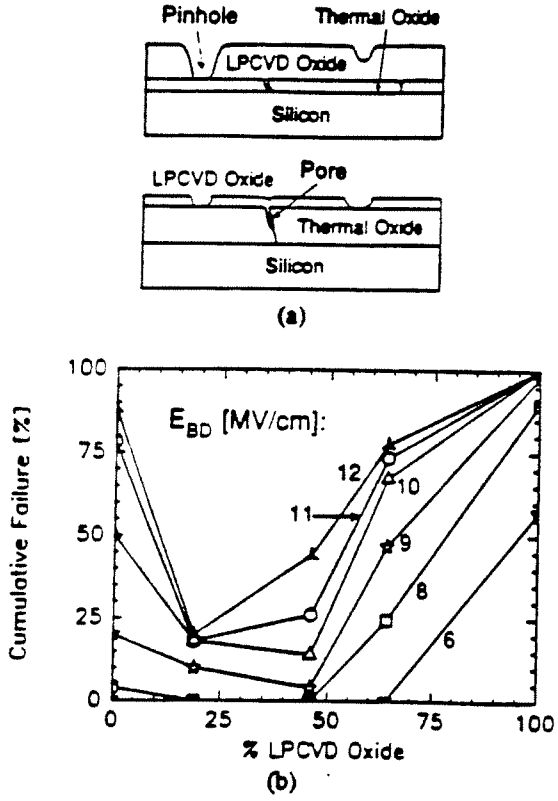


Fig. 7.5 (a) Illustrates that optimization of thermal/CVD stack technology composition results from a balance of the defect quality of the thermal and CVD process. (b) A 25 Å CVD oxide layer is sufficient to dramatically reduce the defect density of a 140 Å (total) stack oxide structure [7.39].

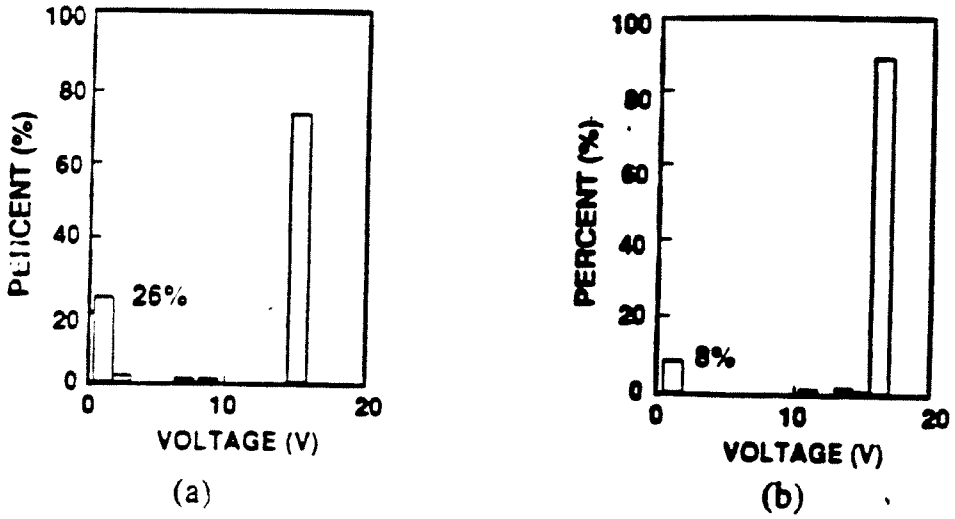


Fig. 7.6 Comparison of breakdown voltage histograms for thermal oxide and stacked oxide subject to severe process induced damage [7.40].

7.6. References

[7.1] D. J. DiMaria, D. Arnold, and E. Cartier, "Degradation and Breakdown of Silicon Dioxide Films on Silicon," *Applied Physics Letters*, vol. 61, no. 19, p. 2329, 1992.

[7.2] S. K. Lai, "Interface Trap Generation in Silicon Dioxide when Electrons are Captured by Trapped Holes," *Journal of Applied Physics*, vol. 54, no. 5, p. 2540, 1983.

[7.3] I.-C. Chen, S. E. Holland, and C. Hu, "Electron-trap Generation by Recombination of Electrons and Holes in SiO_2," *Journal of Applied Physics*, vol. 61, no. 9, p. 4544, 1987.

[7.4] H. Uchida and T. Ajioka, "Electron Trap Center Generation Due to Hole Trapping in SiO_2 Under Fowler-Nordheim Tunneling Stress," *Applied Physics Letters*, vol. 51, no. 6, p. 433, 1987.

[7.5] S. Ogawa, N. Shiono, and M. Shimaya, "Neutral Electron Trap Generation in SiO_2 by Hot Holes," *Applied Physics Letters*, vol. 56, no. 14, p. 1329, 1990.

[7.6] E. Harari, "Dielectric Breakdown in Electrically Stressed Thin Films of Thermal SiO_2," *Journal of Applied Physics*, vol. 49, no. 4, p. 2478, 1978.

[7.7] P. Olivo, T. N. Nguyen, and B. Ricco, "High-Field-Induced Degradation in Ultra-Thin SiO_2 Films," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2259.

[7.8] R. Moazzami and C. Hu, "Stress-Induced Current in thin Silicon Dioxide Films," *International Electron Devices Meeting*, p. 139, 1992.

[7.9] I.-C. Chen, J. Y. Choi, T. Y. Chan, and C. Hu, "The Effect of Channel Hot Carrier Stressing on Gate-Oxide Integrity in MOSFETs," *IEEE Transactions on Electron Devices*, vol. ED-35, no. 12, p. 2253, 1988.

[7.10] K. R. Mistry, D. B. Krakauer, and B. S. Doyle, "Impact of Snapback-Induced Hole Injection on Gate Oxide Reliability of N-MOSFETs," *IEEE Electron Device Letters*, vol. 11, no. 10, p. 460, 1990.

[7.11] E. Rosenbaum, R. Rofan, and C. Hu, "Effect of Hot Carrier Injection on n- and p- MOSFET Gate Oxide Integrity," *IEEE Electron Device Letters*, vol. 12, no. 11, p. 599, 1991.

[7.12] S. Haddad, C. Chang, B. Swaminathan, and J. Lien, "Degradations Due to Hole Trapping in Flash Memory Cells," *IEEE Electron Device Letters*, vol. 10, no. 3, p. 117, 1989.

[7.13] J. Maserjian and N. Zamani, "Observation of Positively Charged State Generation Near the Si/SiO_2 Interface During Fowler-Nordheim Tunneling," *Journal of Vacuum Science and Technology*, vol. 20, no. 3, p. 743, 1982.

[7.14] M.S. Liang, C. Chang, Y. T. Yeow, C. Hu, and R. W. Brodersen, "MOSFET Degradation Due to Stressing of Thin Oxide," *IEEE Transactions on Electron Devices*, vol. ED-31, no. 9, p. 1238, 1984.

[7.15] D. J. DiMaria, "Correlation of Trap Creation with Electron Heating in Silicon Dioxide," *Applied Physics Letters*, vol. 51, no. 9, p. 655, 1987.

[7.16] Y. Tang, C. Chang, S. Haddad, A. Wang, and J. Lien, "Differentiating Impacts of Hole Trapping vs. Interface States on TDDB Reduction in Thin Oxide Gated Diode Structures," *International Reliability Physics Symposium*, p. 262, 1993.

[7.17] E. Rosenbaum, Z. Liu, and C. Hu, "The Effects of Oxide Stress Waveform on MOSFET Performance," *International Electron Device Meeting*, p. 719, 1991.

[7.18] E. Rosenbaum, Z. Liu, and C. Hu, "Silicon Dioxide Breakdown Lifetime Enhancement under Bipolar Bias Conditions," to appear in *IEEE Transactions on Electron Devices*.

[7.19] R. Rofan and C. Hu, "Stress-Induced Oxide Leakage," *IEEE Electron Device Letters*, vol. 12, no. 11, p. 632, 1991.

[7.20] I. Lundstrom, S. Christensson, and C. Svensson, "Carrier Trapping Hysteresis in MOS Transistors," *Phys. Stat. Sol. (a)*, vol. 1, p. 395, 1970.

[7.21] B. Ricco, M. Ya Azbel, and M. H. Brodsky, "A Novel Mechanism for Tunneling and Breakdown in Thin SiO_2 Films," *Physics Review Letters*, vol. 51, p. 1795, 1983.

[7.22] D. J. DiMaria, E. Cartier and D. Arnold, "Impact Ionization, Trap Creation, Degradation, and Breakdown in Silicon Dioxide Films on Silicon," *Journal of Applied Physics*, vol. 73, no. 7, p. 3367, 1993.

[7.23] N. S. Saks and D. B. Brown, "Observation of H^+ Motion During Interface Trap Formation," *IEEE Transactions on Nuclear Science*, vol. 37, no. 6, p. 1624, 1990.

[7.24] D. R. Wolters and A. T. A. Zegers-Van Duynhoven, "Electronic Charge Transport in Thin SiO_2 Films," in *The Physics and Technology of Amorphous SiO_2*, A. B. Devine, Ed., Plenum Press, New York, p. 391, 1988.

[7.25] S. Chiang, R. Wang, T. Speers, J. McCollum, E. Hamdy, and C. Hu, "Conductive Channel in ONO Formed by Controlled Dielectric Breakdown," *Symposium on VLSI Technology Digest of Technical Papers*, p. 20, 1992.

[7.26] C. Hu, S. C. Tam, F. C. Hsu, P. K. Ko, T. Y. Chan, and K. W. Terrill, "Hot-electron-induced MOSFET Degradation - Model, Monitor and Improvement," *IEEE Transactions on Electron Devices*, vol. ED-32, no. 2, p. 375, 1985.

[7.27] M. Koyanagi, A. G. Lewis, J. Zhu, R. A. Martin, T. Y. Huang, and J. Y. Chen, "Investigation and Reduction of Hot Electron Induced Punchthrough (HEIP) Effect in Submicron PMOSFETs," *International Electron Devices Meeting*, p. 722, 1986.

[7.28] A. K. Sinha and T. E. Smith, "Kinetics of the Slow-Trapping Instability at the Si/SiO_2 Interface," *Journal of Electrochemical Society*, vol. 125, p. 743, May 1978.

[7.29] Y. Nishioka, K. Ohyu, Y. Ohji, N. Natuaki, K. Mukai, and T.-P. Ma, "Hot-Electron Hardened Si-gate MOSFET Utilizing F Implantation," *IEEE Electron Device Letters*, vol. 10, no. 4, p. 141, 1989.

[7.30] Y. Nishioka, E. F. Da Silva, Jr., Y. Wang, and T.-P. Ma, "Dramatic Improvement of Hot-Electron-Induced Interface Degradation in MOS Structures Containing F or Cl in SiO_2," *IEEE Electron Device Letters*, vol. 9, no. 1, p. 38, 1988.

[7.31] V. Jain, D. Pramanik, K. Y. Chang, and C. Hu, "Improved Sub-micron CMOS Device Performance Due to Fluorine in CVD Tungsten Silicide," *Symposium on VLSI Technology Digest of Technical Papers*, p. 91, 1991.

[7.32] T. Hori and H. Iwasaki, "Improved Hot-Carrier Immunity in Submicrometer MOSFETs with Reoxidized Nitrided Oxides Prepared by Rapid Thermal Processing," *IEEE Electron Device Letters*, vol. 10, no. 2, p. 64.

[7.33] H. S. Mamose, T. Morimoto, Y. Ozawa, M. Tsuchiaki, M. Ono, K. Yamabe, and H. Iwai, "Very Lightly Nitrided Oxide Gate MOSFETs for Deep-Submicron CMOS Devices," *International Electron Devices Meeting*, p. 359, 1991.

[7.34] H. Hwang, W. Ting, D. L. Kwong, and J. Lee, "Electrical and Reliability Characteristics of Ultrathin Oxynitride Gate Dielectric Prepared by Rapid Thermal Processing in N_2O," *International Electron Devices Meeting*, p. 421, 1990.

[7.35] Z. H. Liu, "The Effects of Furnace N_2O Annealing on MOSFETs," *International Electron Devices Meeting*, p. 625.

[7.36] P. K. Roy, R. H. Doklan, E. P. Martin, S. F. Shive, and A. K. Sinha, "Synthesis and Characterization of High Quality Ultra-Thin Gate Oxides for VLSI/ULSI Circuits," *International Electron Devices Meeting*, p. 714, 1988

[7.37] P. K. Roy and A. K. Sinha, "Synthesis of High Quality Ultra-Thin Gate Oxides for ULSI Applications," *AT&T Technical Journal*, p. 155, Nov./Dec. 1988.

[7.38] L. Yau, Intel, Inc., presented at SRC Workshop on Oxide Reliability, Jan. 1990.

[7.39] R. Moazzami and C. Hu, "A High-Quality Stacked Thermal/LPCVD Gate Oxide Technology for ULSI," *IEEE Electron Device Letters*, vol. 14, no. 2, p.72, 1993.

[7.40] H. H. Tseng, P. J. Tobin, J. D. Hayden, K. M. Chang, and J. W. Miller, "A Comparison of CVD Stacked Gate Oxide and Thermal Gate Oxide for 0.5 μm Transistors Subjected to Process Induced Damage," *IEEE Transactions on Electron Devices*, vol. ED-40, no. 3, p. 613, 1993.

[7.41] C. Hu, "Future CMOS Scaling and Reliability," *Proceedings of the IEEE*, vol. 81, no. 5, p. 682, 1993.

[7.42] M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro and H. Iwai, "Sub-50nm Gate Length nMOSFETs with 10 nm Phosphorus Source and Drain Junctions," *International Electron Devices Meeting*, p. 119, 1993.

[7.43] P. Singer, "Directions in Dielectrics in CMOS and DRAMs," *Semiconductor International*, p. 56, April, 1994.

Appendix I

List of Acronyms

BST	**Barium Strontium Titanate**, a dielectric with relative dielectric constant greater than 500.
CVD	**Chemical Vapor Deposition**
CMOS	**Complementary Metal Oxide Silicon**, a low power dissipation integrated circuit technology.
D_{it} (or N_{it})	**Interface Trap Density** at silicon-silicon dioxide interface.
DRAM	**Dynamic Random Access Memory**, a high-density, volatile memory.
FET	**Field Effect Transistor**, four terminal, majority carrier switch of primary importance in silicon integrated circuit technology.
FN	**Fowler-Nordheim**
GPIB	**General Purpose Interface Bus**, a communications protocol for controlling measurement instrumentation by means of a personal computer (PC).
IC	**Integrated Circuit**, circuit consisting of transistors and other components integrated onto a single underlying substrate of silicon.
MOSFET	**Metal-Oxide-Semiconductor Field Effect Transistor**, a FET using silicon dioxide as gate dielectric.
nMOS	**n**-channel Field Effect Transistor.
PC	**Personal Computer**

PZT	**Lead Zirconate Titanate, $PbZr_xTi_{1-x}O_3$,** a high dielectric constant ferroelectric material.
pMOS	**p**-channel Field Effect Transistor.
Q_{bd}	**Charge to Breakdown**, charge flowing through a dielectric to reach the breakdown event.
SiO_2	**Silicon Dioxide**, an amorphous natural oxide of silicon, displaying excellent insulative properties, with relative dielectric constant of 3.9.
t_{bd}	**Time to Breakdown**, time required to breakdown a dielectric at a specified voltage.
t_{50}	**50% quantile** of the Time to Breakdown data.
TDDB	**Time Dependent Dielectric Breakdown**
VLSI	**Very Large Scale Integration**, integration level with more than 100,000 transistors in a single silicon integrated circuit. Defined as the state-of-the-art in integration density. Predicted advances in integration density are referred to by some as **ULSI (Ultra Large Scale Integration)**.