12-2008

Reliability Analysis of Hafnium Oxide Dielectric Based Nanoelectronics

Rui Wan

University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Rui Wan entitled "Reliability Analysis of Hafnium Oxide Dielectric Based Nanoelectronics." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

Way Kuo, Yue Kuo, Major Professor

We have read this dissertation and recommend its acceptance:

Alberto Garcia, Myong. K. Jeong, Frank Guess

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Rui Wan entitled "Reliability Analysis of Hafnium Oxide Dielectric Based Nanoelectronics." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

Way Kuo
Major Professor

Yue Kuo
Co-chair of Committee

We have read this dissertation and
recommend its acceptance.

Alberto Garcia

Myong. K. Jeong

Frank Guess

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the
Graduate School

(Original signatures are on file with official student records)

RELIABILITY ANALYSIS OF HAFNIUM OXIDE DIELECTRIC BASED NANOELECTRONICS

A Dissertation
Presented for the
Doctoral of Philosophy
Degree
The University of Tennessee, Knoxville

Rui Wan
December 2008

To my dear family

ACKNOWLEDGMENTS

First and foremost I would like to express my appreciation to my advisor, Dr. Way Kuo, for his essential insight, guidance, and support throughout my Ph.D program. I feel fortunate under his guidance, and it was his encouragement and patience that helped me go through all the difficulties in the past four years.

Sincere appreciation goes to my co-chair, Dr. Yue Kuo, who has given me the great opportunity to study in the Thin Film Nano & Microelectronics Research Lab. I thank Dr. Kuo for the valuable knowledge and experience I have gained in the field of semiconductor devices under his guidance and help.

I would like to thank Dr. Alberto Garcia, Dr. M. K. Jeong, and Dr. Frank Guess for serving on the committee and for giving valuable suggestions in my research.

Dr. Wen Luo deserves special appreciation for her help, encouragements and recommendations when I felt frustrated in my study. Thanks also go to Dr. Jiang Lu, Jiong Yan, and Chien-Han Lin for the sample fabrication and for the wonderful time and discussions we shared.

Lastly, I want to thank my family for their continuous support throughout my life.

ABSRACT

With the physical dimensions ever scaling down, the increasing level of sophistication in nano-electronics requires a comprehensive and multidisciplinary reliability investigation. A kind of nano-devices, HfO_2-based high-k dielectric films, are studied in the statistical aspect of reliability as well as electrical and physical aspects of reliability characterization, including charge trapping and degradation mechanisms, breakdown modes and bathtub failure rate estimation.

This research characterizes charge trapping and investigates degradation mechanisms in high-k dielectrics. Positive charges trapped in both bulk and interface contribute to the interface state generation and flat band voltage shift when electrons are injected from the gate under a negative gate bias condition.A negligible number of defects are generated until the stress voltage increases to a certain level. As results of hot electrons and positive charges trapped in the interface region, the difference in the breakdown sequence is attributed to the physical thickness of the bulk high-k layer and the structure of the interface layer.

Time-to-breakdown data collected in the accelerated life tests are modeled with a bathtub failure rate curve by a 3-step Bayesian approach. Rather than individually considering each stress level in accelerating life tests (ALT), this approach derives the change point and the priors for Bayesian analysis from the time-to-failure data under neighborhood stresses, based on the relationship between the lifetime and stress voltage. This method can provide a fast and reliable estimation of failure rate for burn-in optimization when only a small sample of data is available.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHARTER 1
INTRODUCTION

As technology continues advancing, nano-scale semiconductor devices have entered all spheres of human life [1]. In the last three decades, the scaling of electronic devices has become the driving force for semiconductor industry to meet the market's demand for greater functionality and performance at a lower cost [2, 3]. By reducing the device dimensions of the integrated circuits, the manufacturing cost of each device can be reduced and the device performance can be also greatly improved in terms of switch speed [4]. One concern is that the tight reliability margins may limit miniaturization [5]. With the physical dimensions ever scaling down, the increasing level of sophistication in nano-electronics requires a comprehensive reliability investigation in all material, physical, electrical, and statistical aspects in early design stages. A brief description is presented next on some critical issues in the present semiconductor industry. These problems necessitate the research on the reliability of high k dielectric films.

1.1 MOSFET and MOS Capacitors

Metal-oxide-semiconductor field effect transistor (MOSFET) is the most used semiconductor device today. The schematic diagram of a MOSFET is shown in Figure 1. When a sufficient gate voltage is applied on the gate, a conducting channel is created between the source and the drain due to the modulation of charge concentration. For gate bias below the threshold value, the channel disappears and

This dissertation follows the style of the *IEEE Transactions on Reliability*.

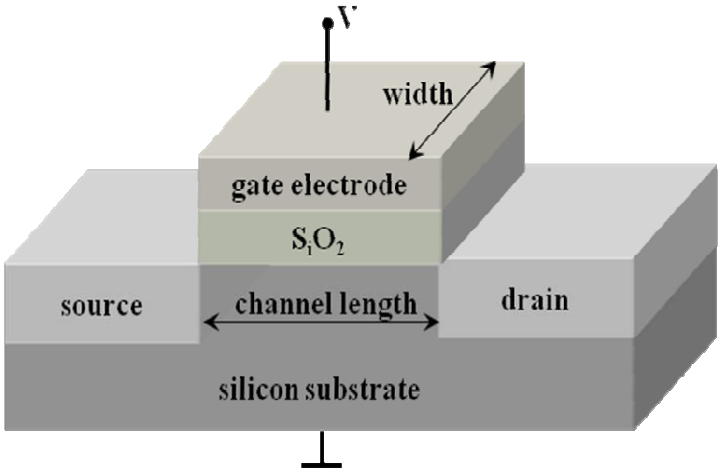


Figure 1. Schematic diagram of a MOSFET

only a very small subthreshold leakage current can flow between the source and the drain. The source is so named because it is the source of the charge carriers that flow through the channel; similarly, the drain is where the charge carriers leave the channel.

Working as the switch of a MOSFET, the performance of the dielectric S_iO_2 is crucial to the proper functioning of the transistor. A simple way to test gate oxide is to put it in a metal-oxide-semiconductor (MOS) capacitor [5], as shown in Figure 2. The metal and the semiconductor correspond to the gate electrode and silicon substrate of a transistor, respectively. The capacitance C is calculated as

$$C = \frac{\varepsilon_0 kA}{t} \tag{1}$$

where ε_0 is the permittivity of the vacuum ($8.85 \times 10^{-14} F/cm$), k is the dielectric constant of the gate dielectric material, A is the gate size, and t is the physical thickness of the gate dielectric.

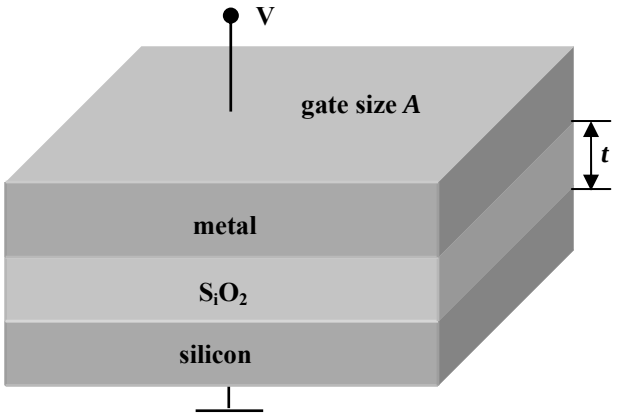


Figure 2. Schematic diagram of a MOS capacitor

1.2 Reliability Issues on Device Scaling

Figure 3 shows the advancement of the Intel processor [6]. It follows Moore's law that the number of transistors per integrated circuit increases exponentially and the integrated circuit density doubles every 18 months [4, 6]. Moore's law has been successfully maintained for 30 years and semiconductor industry expects that it will continue at least through the end of this decade [4].

The increased density of devices that results from reduced device dimensions has had a significant economic benefit. Because the semiconductor devices are fabricated by single-wafer or batch processes, scaling down the device dimensions and packing more devices onto each wafer have resulted in a dramatic and sustained reduction in the cost of each device [4]. The critical dimensions of CMOS devices decreased from 10-20μm to below the 0.1μm range over the last half-century of the microelectronic industry development [4].

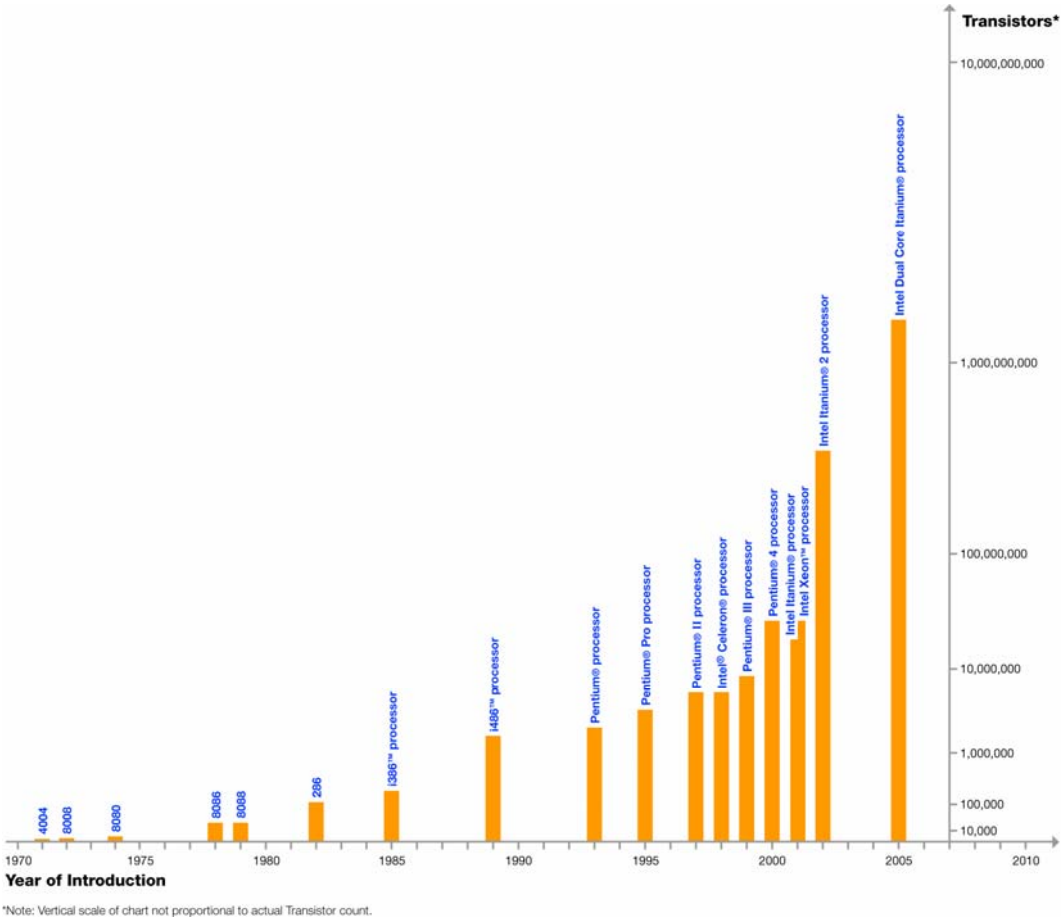


Figure 3. The advancement of the Intel processor [6]

The device performance and integration can be also improved by scaling the gate channel length of a MOSFET device. The drive current of a MOSFET available to switch its load devices increases linearly with the decrease of the physical channel length as shown in Eq. 2, which usually provides an improvement in the circuit speed [7].

$$I_D = \frac{W}{L}\mu C_{inv}(V_{GS} - V_T)V_{DS} \tag{2}$$

where I_D is the drive current between the source and drain, W is the effective channel width, L is the effective channel length, μ is the carrier mobility in the channel, C_{inv} is the gate capacitance density with channel in inversion, V_{GS} is the gate to source voltages, V_T is the threshold voltage, and V_{DS} is the drain to source voltage.

However, the gate area of a MOSFET device will also decrease with the reduction of the gate channel length, which will reduce the input capacitance and the drive current required by the load devices to achieve switching according to Eq. 1-2. Therefore, the physical thickness of gate dielectric film in Eq.1 should be approximately linearly scaled down together with the lateral channel dimension in order to maintain the same amount of gate control over the channel [4, 5].

Silicon dioxide (S_iO_2) has served as the gate dielectric for integrated circuit applications for more than 40 years, due to its excellent material and electrical properties [8-11]. However, as the oxide thickness scales down below 1.5nm[1], there exists a large direct tunneling gate-to-channel leakage current, resulting in excessively high power consumption. The scaling limit of S_iO_2 gate oxide in terms of leakage

[1] $1nm=10^{-9}m$

current should be 2.2-2.5nm for a low operating power application and 1.4-1.6nm for a high performance application [10, 12]. Scaling can also be limited by processing issues, such as high density of defects and poor uniformity of the gate oxide, the difficulty of S_iO_2 growth control, and the threshold voltage shift caused by boron diffusion from the polysilicon gate through the oxide. In order to have a physically thicker film but also maintain the capacitance, the use of dielectric materials with a large dielectric constant *k* is the only solution.

1.3 High-k Dielectric Materials

As discussed in previous section, the concerns of a high leakage current through the gate oxide and other reliability issues can be solved by using a new gate dielectric material with a high dielectric constant to replace the S_iO_2. The larger dielectric constant allows the use of physically thicker dielectric films without sacrificing film capacitance. With a thicker insulating layer, the concerns of a high tunneling current through the gate oxide, poor film uniformity and insufficient reliability can be solved efficiently. Figure 4 shows that using a 3.0 nm thick high-k gate dielectric layer may improve the capacitance by 60% and decrease the leakage current density by 2 orders of magnitude [13].

Metal oxides have been extensively studied as candidate materials because of their large dielectric constants, ranging from 20 to 80. Hafnium oxide (HfO_2) is selected as the most promising high-k gate dielectrics for the next generation of transistors. When HfO_2 was doped with an appropriate amount of zirconium (Zr), many of its physical and electrical properties, such as the crystallization temperature,

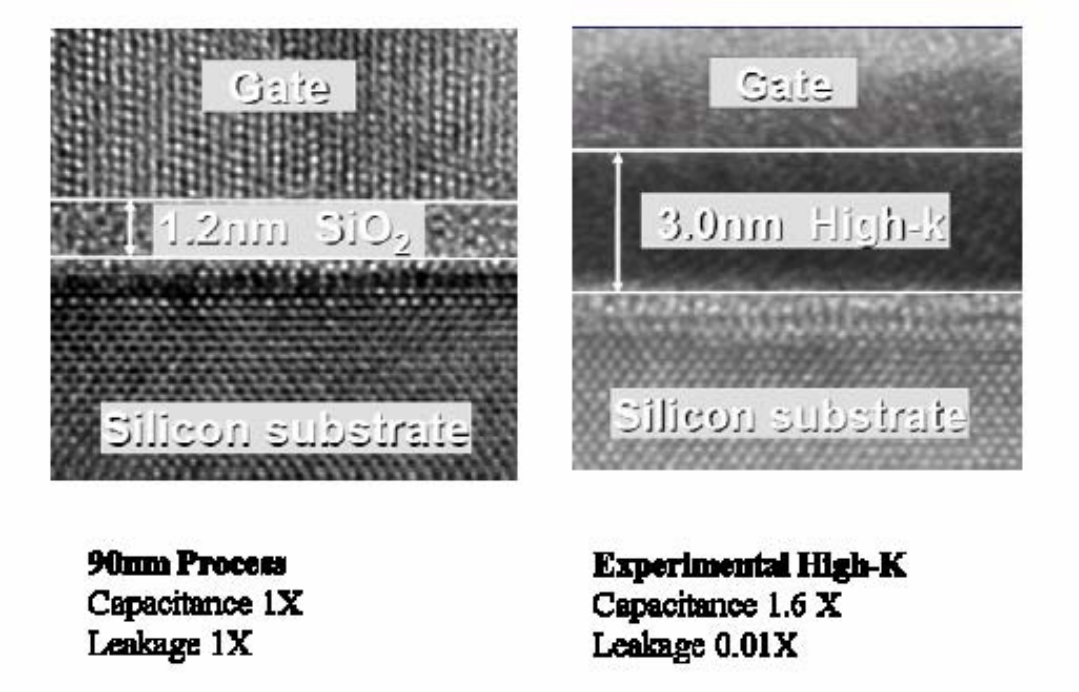


Figure 4. Benefit of using high-k gate dielectric material v.s. ultra-thin SiO_2 gate oxide [13].

effective *k* value and interface state density, could be improved [14, 15]. In spite of these promising characteristics, reliability characterization and evaluation are needed to fully understand the time-dependent performance of ZrHfOx high-k dielectrics.

1.4 Fabrication of HfOx-based films

Table 1 is the list of MOS capacitor structures in this study. Sample capacitors are fabricated and tested in the Thin Film Nano & Microelectronics Research Lab, Texas A & M University.

In sample preparation with the structure TiN/ZrHfO/SiON/Al, a 1nm-thick SiON was pre-deposited on a *p*-type silicon wafer. The Zr-doped HfO_2 ($ZrHfO_x$) film was

Table 1. High k dielectric films under test

Group	Gate Stack	EOT*
I	TiN/20s-deposited $ZrHfO_x$ / 1nm SiON/p-Si	1.8 nm
	TiN/45s-deposited $ZrHfO_x$ / 1nm SiON/p-Si	2.5 nm
II	TiN/ $ZrHfO_x$ /p-Si	1.79 nm
	TiN/ $RuZrHfO_x$ /p-Si	1.19 nm
	TiN/ $ZrHfO_x$ / RuO_x / $ZrHfO_x$ / p-Si	1.21 nm
III	Al/ $ZrHfO_x$ / p-Si	10 nm
	Al/ $ZrHfO_x$ / nc-Ru / $ZrHfO_x$ / p-Si	9 nm

* Equivalent Silicon Dioxide Thickness.

sputtered on the SiON surface using a composite Zr/Hf (12: 88 wt%) target in an Ar/O_2 (1:1) mixture at 5mTorr and 100W for 20 sec and 45 sec. The post deposition annealing (PDA) was done with rapid thermal annealing (RTA) 800^oC for 10s under the N_2/O_2 (20:1) atmosphere. The TiN film was then sputter deposited using a Ti target (100W) in N_2/Ar (1:50), followed by a 350^oC 10min forming gas (90% N_2/10% O_2) post metal annealing. The 120 nm thick TiN gate electrode has polycrystalline structure with a resistivity about 40μΩcm. The TiN layer was patterned with a lithography mask and wet etched with $NH_4OH:H_2O_2:H_2O$ (1:1:5) solution to form the gate pattern with an area of $7.85\times10^{-5}\,cm^2$. The backside of the wafer was deposited with an aluminum layer for ohmic contact formation. Finally, the capacitors were annealed at 300^oC in forming gas.

The EOT of 20s- and 45s-deposited $ZrHfO_x$/SiON samples were 1.8nm and 2.5nm, respectively, calculated from the capacitance-voltage (C-V) curves. The Zr/Hf

ratio in this film is 21/79 from X-ray photoemission spectroscopy results, which corresponding to the k value about 16, close to HfO_2.

The $ZrHfO_x$ and Ru-modified $ZrHfO_x$ samples were deposited by two different methods: by co-sputtering of separate Hf/Zr (100W) and Ru (100W) targets at room temperature for 20 sec (named $RuZrHfO_x$); by sequential deposition of the $ZrHfO_x$ (10 sec)/RuO_x (10 sec)/$ZrHfO_x$ (10 sec) tri-layer (named $ZrHfO_x/RuO_x/ZrHfO_x$) in one pumpdown [16]. All reactive sputtering processes were done in 1:1 Ar/O_2 at 5mTorr and the post-deposition annealing (PDA) was done at 1000^oC for 10 seconds in a 20:1 N_2/O_2 mixture by rapid thermal annealing (RTA) [16].

EOT was lowered due to the dipole enhancement mechanisms, i.e. the conductive RuOx could behave as induced dipoles or serve as quantum trapping centers where carrier to load via resonant tunneling [16]. XPS analysis show the interface layer of Ru-doped $ZrHfO_x$ ($RuZrHfO_x$) sample is close to SiO_2, but those of $ZrHfO_x/RuO_x/ZrHfO_x$ and $ZrHfO_x$ are close to Hf-silicate.

All three layer in the $ZrHfO_x$/nc-Ru/ZrHf O_x dielectric structure were deposited in one pump-down. Both top and bottom $ZrHfO_x$ layers were deposited by reactive sputtering using a Hf/Zr (88:12 wt%) composite target in an Ar/O_2 (1:1) mixture at 5 mTorr and room temperature. Sputtering powers of the $ZrHfO_x$ film and the embedded layer were 100W and 80W, respectively and the high-k film was rapid thermal annealed at 950^oC under N_2/O_2 (1:1) atmosphere [17]. A control sample without the nanocrystals embedded layer was prepared under the same sputtering and annealing condition for comparison.

Figure 5. Black box and equipment inside used for testing [5].

All films are tested at room temperature in a black box to avoid disturbance from light and noise as shown in Figure 5. Current versus voltage characteristics and data of failure times are collected with HP 4155C, the semiconductor parameter analyzer. Capacitance versus voltage characteristics are measured with HP 4140B, the pico-ampere meter and DC voltage source. The HP 4140B and HP 4155C are also connected to a desktop computer for automatic control of measurements and data acquisition. All tests are programmed in LabVIEW 7.0, virtual instrument software developed by National Instruments.

1.5 Outline of the Dissertation

Chapter 2 first reviews several commonly used testing methods to characterize the reliability of gate dielectric films. The primary reliability concerns of gate dielectric films are introduced in the second part of this chapter, including degradation and

breakdown mechanisms, lifetime prediction and bathtub failure rate estimation.

Chapter 3 focuses on the charge trapping and degradation mechanisms in high-k dielectrics. A test method is proposed to investigate defect generating conditions and stress-induced defect density. Compared with $ZrHfO_X$, dielectric film with structure $ZrHfO_X$ / RuO_X / $ZrHfO_X$ is considered more promising in terms of leakage current, relaxation current, defect generating conditions and defect creation rate.

As results of hot electrons and positive charges trapped in the interface region, breakdown mechanisms are investigated in ultra-thin ZrHfOx/1 nm SiON high-k stacks in Chapter 4. The two-step breakdown phenomenon is observed on the 1.8nm EOT sample and the one-step breakdown phenomenon was detected on the 2.5nm EOT sample. The difference in the breakdown sequence is attributed to the physical thickness of the bulk high-k layer and the structure of the interface layer. Compared to the 2.5 EOT sample under the same gate bias condition, the higher stress on the interface of the 1.8nm EOT sample accelerates the defect generation in the interface region, which leads to its early interface breakdown. The bulk high-*k* breaks subsequently and separately at a higher voltage.

Chapter 5 investigates the charge trapping of Ruthenium nanocrystal embedded high-k film. Memory window in capacitance-voltage characteristics was mainly contributed by the negative flatband voltage shift due to hole trapping. Since the total charge trapping density is larger than the nanocrystal density, hole trapping might also occur in both nc-RuOx and in the interface between the nc-RuOx and ZrHfOx dielectric due to Coulomb blockade effect. The loosely trapped charges identified by relaxation current could be more likely located in the interface than in nc-RuOx, but not all of charges trapped in the interface could be recognized as loosely trapped

charge.

Following the experimental study concerning reliability characterization in previous three chapters, Chapter 6 and 7 presents a 3-step Bayesian procedure to estimate device bathtub failure rates, when the change point is unknown and the sample size is not large. Rather than individually considering each stress level in accelerating life tests (ALT), this approach derives the change point and the priors for Bayesian analysis from the time-to-failure under neighborhood stresses, based on the relationship between the lifetime and stress voltage.

Chapter 8 summarizes all of the work in this study and draws the conclusions.

CHARTER 2

RELIABILITY CONCERNS

ON HAFNIUM OXIDE BASED HIGH K FILMS

Reliability is defined as the probability of a device performing its designed function adequately for the intended period of time under the operating conditions [18]. For gate dielectric films, the designated function is to act as a sufficient insulator; the operating conditions include mainly the electric field and the temperature that correspond to the normal operation of a transistor; the intended lifetime could be 10 to 20 years depending on the application [5]. In replacing S_iO_2 with high k dielectrics, although the differences in the material properties lead to changes in the electric and reliability characteristics, the understanding the physical mechanisms, as well as the statistical behaviors, of oxide degradation and breakdown provides an advantageous starting point for studying high k films.

2.1 Stress Tests

The reliability of dielectric films can be characterized by means of stress testing [5]. According to the stress mode applied to the devices, testing methods can be categorized as the constant voltage stress (CVS) test, the constant current stress (CCS) test, the ramped voltage stress (RVS) test, and the ramp-relax stress (RRS) test [19, 20].

2.1.1 Constant voltage stress test

In CVS tests, a constant voltage is applied to the test structure, while the current evolution is monitored through a current versus time curve. A CVS test is usually used to measure time-to-breakdown and to predict the time-to-breakdown under the operating conditions, because this stress is more representative of the operating conditions in a real circuit environment [5]. Breakdown is detected if the leakage current is larger than a threshold or according to a criterion of change in the leakage current.

2.1.2 Constant current stress test

A constant current is applied on the gate electrode while the voltage across it is plotted against time. Breakdown is detected if a sudden drop of voltage occurs between two consecutive stress steps. This test is usually used to measure the injected charge to breakdown and, in some cases, time-to-breakdown as well [5].

2.1.3 Ramp voltage stress test

A RVS test applies a staircase voltage on the gate electrode, depicted in Figure 6. Leakage current flowing through the dielectric film is plotted against voltage at each step. This type of stress test is used to measure the maximal electric field a film can sustain before breakdown [5].

2.1.4 Ramp relax stress test

Ramp relax stress test, as shown in Figure 7, is usually used to detect dielectric breakdown by monitoring the relaxation current at a zero gate bias after each ramp step [20]. Figure 8 shows I-V curves that are measured using the ramp-relax test. The

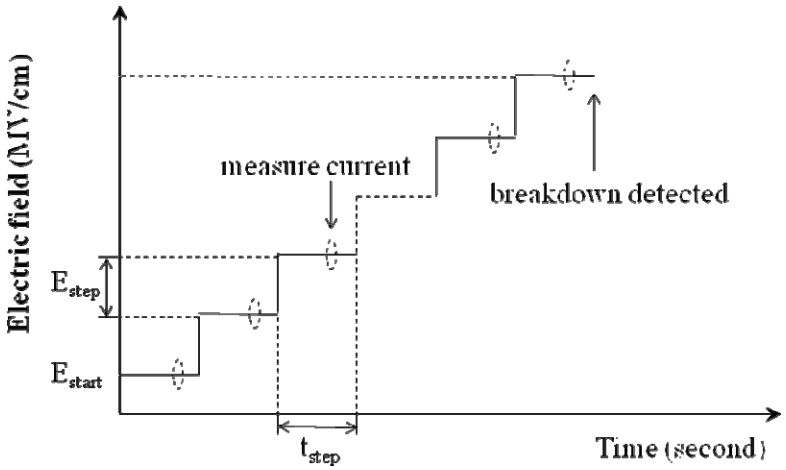


Figure 6. Stress voltage of a staircase of RVS [5]

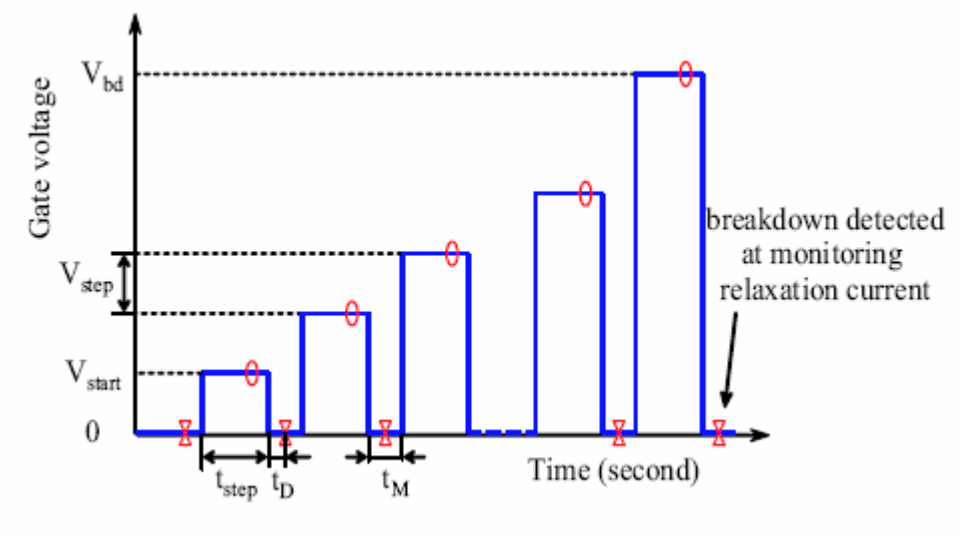


Figure 7. Ramp-relax of breakdown detection [20]

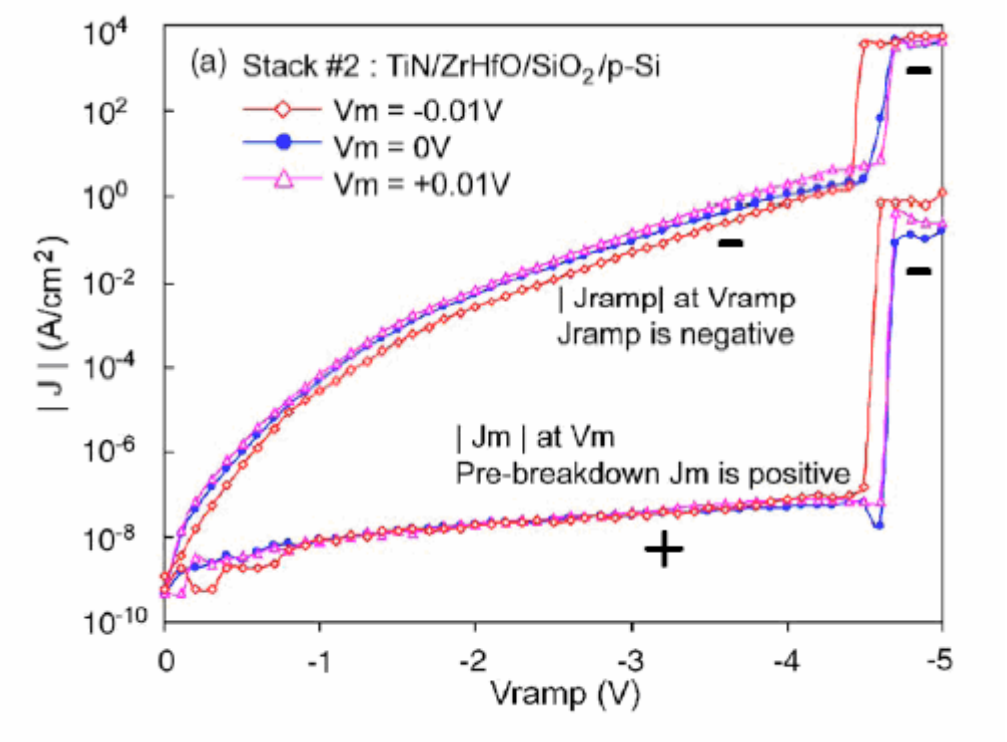


Figure 8. Ramp-relax test result on TiN/ZrHfO_x/2 nm SiO_2/p-Si [21].

leakage jump signals a breakdown, and the jump of the leakage current (Jramp) coincides with the disappearance of the relaxation current (Jm) implied by a sudden change of current direction [21]. Because the relaxation property is an essential element of high-k film integrity, the sudden disappearance of relaxation current can be used as the signal of high-k dielectric breakdown [20].

The ramp-relax test is very sensitive for breakdown detection when films are stressed in accumulation, in which case the relaxation and the leakage currents flow in opposite directions. This method avoids the difficulty of setting an empirical and quantitative criterion as is required by some common detection methods. It is simple and accurate even when the abrupt change in leakage current does not occur [20].

2.2 Charge Trapping in Connection with Degradation and Breakdown Mechanism

Degradation and breakdown mechanisms are fairly complex. The degradation is due to the continuous charge build-up that results from the charge injection when the dielectric is subjected to an electric field; the breakdown is initiated by the presence of a weak spot and manifested by the formation of a conductive path through the film [5]. Two models have emerged in recent years as the two most popular ways of describing the defect generation.

The anode hole injection model claims that breakdown is caused by anode hole injection. Hole-related defect generation usually occur at relatively high voltages [5]. As a consequence of the reaction of the released mobile hydrogen, a variety of defects, such as electron traps and interface states, gradually build up in the oxide and reach a point where the film breaks down [5]. The anode hole injection model is usually associated with the time-to-breakdown dependence on the electric field in the form of

$$t_{bd} \propto \exp\left(-\frac{a}{E}\right)$$

where E is the electric field across the oxide and a is a constant.

The thermo-chemical model considers the interaction of the applied electric field with dipole moments associated with oxygen vacancies [5]. The activation energy required for bond breakage is lowered by the dipolar energy, which leads to a quantitative prediction for the field dependence of the activation energy needed for breakdown and the dependence of the time-to-breakdown on the electric field in the form of [5]

$$t_{bd} \propto \exp(-E)$$

The thermo-chemical model is widely accepted, because the breakdown data appears to follow empirically an exponential dependence on the electric field [5]. However, the precise physical details of the trap generation mechanism remain speculative.

2.3 Bathtub Failure Rate Estimation

The failure rate of a semiconductor device usually has a bathtub shape, as shown in Figure 9, including a decreasing failure (DFR) due to infant mortality, a constant failure rate (CFR) representing the useful lifetime, and an increasing failure rate (IFR) reflecting the wear-out or aging effect [5]. The time when the failure rate turns from DFR into CFR is called the change point. The wear-out phase is not concerned under the reliability context due to the quick development and update of semiconductor electronic devices.

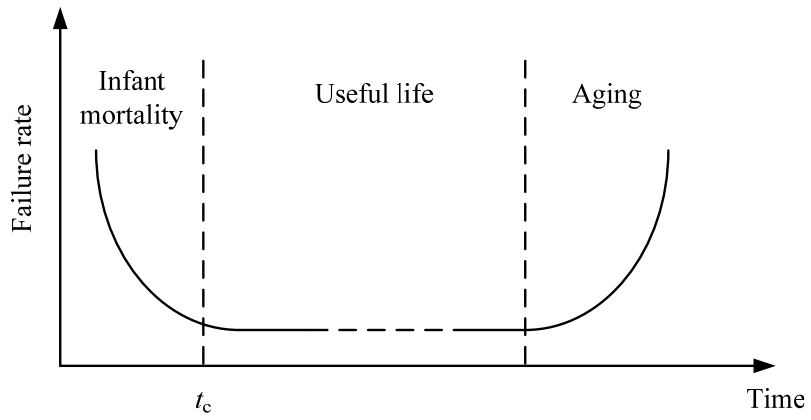


Figure 9. Bathtub failure curves

Since a large number of failures occur in the early stage of usage, manufactures usually employ a burn-in process to screen out the defective units and minimize the total cost of manufacturing, burn-in and warranty [5, 18]. Estimating the decreasing failure rate, constant failure rate and change point is the key to determining the optimal stress level and stress time of burn-in [5].

The breakdown of gate oxide is a weak-link problem. For instance, the failure of a microchip consisting of millions of transistors is defined by the first failure occurring on any one of these individual transistors. The Weibull distribution is known as the extreme value distribution that describes the smallest of many non-negative random variables. Its derivation suggests its application for a weakest-link type of problems [5]. Both experimental data and the physical examination of thin oxides provide the justification of using Weibull distribution to describe the decreasing failure rate [23]. The Weibull slope is a measure of critical defect density, which can also provide information of great value to fabrication engineers.

Several modified Weibull distributions are presented in Ref [24-27], which are mathematically capable of modeling bathtub failure rate. However, from an engineering point of view, it is difficult to apply these models in real life without further physical explanation on each parameter. Ref [28] and Ref [29] discussed the bathtub shaped failure rates with Weibull-Weibull and Weibull-exponentail models, respectively. The shape parameters of Weibull distributions are all greater than 1, which indicates only intrinsic defects are considered in the model. But this is not the case in real practice of nano-electronics. In Ref [30], the maximum-likelihood estimation (MLE) was used to estimate the change point in Weibull-exponential

mixture and it also showed that the optimal burn-in time should not exceed the change-point. However, the issue of simultaneously estimating the change-point and DFR still requires considerable research efforts and remains an area for research, especially when the sample is not large in early design stages.

2.4 Reliability Projection

Figure 10 illustrates the process of reliability projection [23]. A complete projection process usually includes:

1) estimating of the Weibull slope
2) projection of the characteristic life under the normal operating condition
3) scaling to a total gate
4) extrapolation to a low percentile life

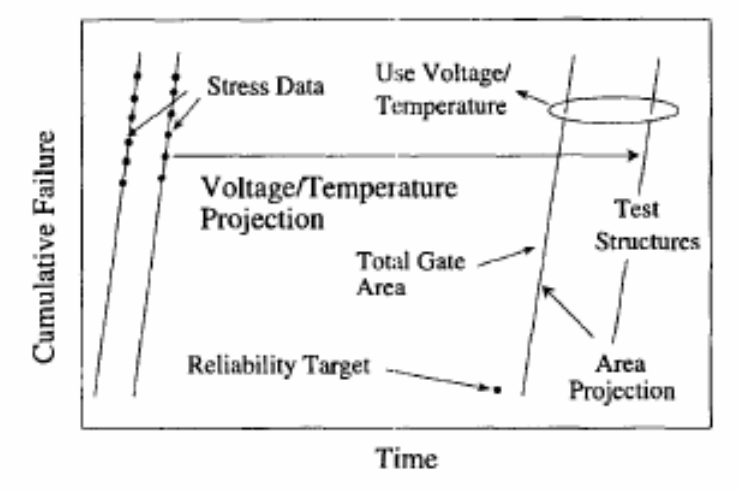


Figure 10. A schematic illustration of reliability projection [23]

In a typical dielectric life projection, the ALT data is first extrapolated from stress conditions to normal operating conditions. This step is legitimate under the assumption of identical failure mechanisms at all the stress levels involved. Particularly for dielectric thin films, time-to-breakdown data at an accelerated electric field are fitted with a Weibull distribution, for which the scale parameter of Weibull distribution can be estimated by means of least square regression or maximum likelihood estimation. Then the scale parameter at normal operating electric field can be extrapolated on the Weibull plot. Since the stresses are often performed on test structures with small area, a projection to the total area of a chip is required. Because a large number of breakdown events are unacceptable in practice, reliability specifications on gate oxide are usually defined in terms of a very low percentile lifetime. The last step of projection is to extrapolate the lifetime at higher failure percentiles to the lower cumulative failure percentiles using the Weibull breakdown distribution. A complete study of reliability projection on gate oxide has been reported in [23].

CHARTER 3

CHARGE TRAPPING AND DIELECTRIC DEGRADATION

Charge trapping in high-k dielectrics has been one of the most important concerns, which determine the application of these metal oxides in CMOS devices [31, 32]. The probability of bias-induced charge trapping in high-k gate stacks is extremely high due to the large densities of as-grown defects, which may function as electron traps and fixed charges [33, 34]. Charge trapping/detrapping can affect evaluation of mobility in transistors, cause threshold voltage instabilities, and greatly influence the reliability of the devices over the operation time [31-34]. Therefore, it is important to investigate the charge trapping mechanism in order to fully understand the channel mobility degradation and device reliability with high-k dielectrics.

3.1 Relaxation Current

Dielectric Relaxation is a bulk-related phenomenon which causes relaxation current following the direction of dV. Relaxation currents can be measured after the sudden removal of a constant voltage on the gate. The relaxation current decays with time following the Curie-von Schweidler law [35]:

---

*Reprinted with permission from "Dielectric breakdown and charge trapping of Ultrathin ZrHfO/SiON high-k Gate Stacks" by Rui Wan, Jiong Yan, Yue Kuo and Jiang Lu. Japanese Journal of Applied Physics, vol. 47, p. 1639, 2008.

*Reprinted with permission from "Charge trapping of Ultrathin ZrHfO/RuO/ZrHfO high-k Stacks" by Rui Wan, Chan-Han Lin, Yue Kuo and Way Kuo. The Proceedings of International Integrated Reliability Workshop, Fallen Leaf Lake, CA, 2007.

$$J / P = at^{-n}$$

where J is the relaxation current density(A/cm^2); P is the total polarization or surface charge density (V. nF/cm^2); t is the time in seconds; a is constant and n is real number close to 1.0. High-k films have relaxation currents one to two orders of magnitude larger than that of SiO_2 as shown in Figure 11 [20].

The relaxation behavior of metal-oxide high-k materials is contributed by charge trapping/detrapping and dielectric polarization/relaxation simultaneously. The mechanisms of charge trapping/detrapping and dielectric polarization/relaxation might not be physically distinguishable [36]. High-k gate stacks are know to be susceptible to charge trapping due to rich oxygen vacancies and oxygen interstitial atoms, imperfect boundary between the high-k layer and interface layer with different morphologies, and H atoms induced P_b-like centers during film

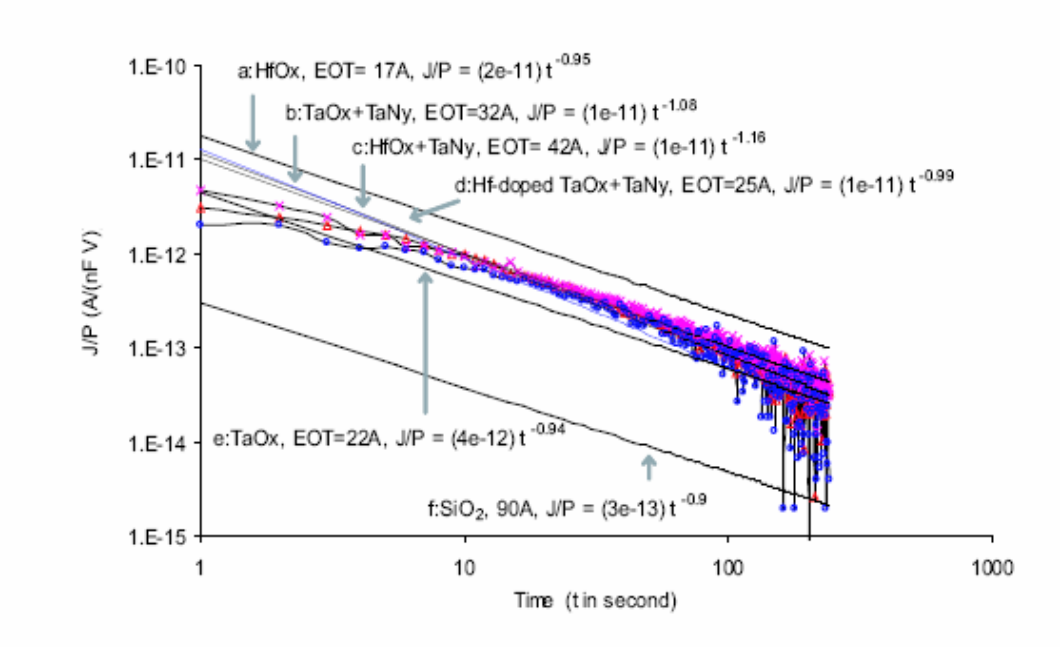


Figure 11. Comparison of relaxation current [20].

deposition [37-39]. Since the amplitude of the polarization is controlled by the electric field in the dielectric materials [40], the relaxation current can be used to compare the charge trapping characteristics in the high-k dielectric films under the same gate bias conditions.

3.2 Charge Trapping in $ZrHfO_X$ films

In order to understand the charge trapping of the high-*k* gate stack, a stress-relax method has been developed, as depicted in Figure 12. During the test, the decays of the relaxation currents are monitored and compared at each ramp step. The test interval should be long enough for the device to completely discharge.

Figure 13 shows the decays of I_{relax} with time at various cycles in Figure 12. When stressed with a lower voltage, i.e., -1.5V, -2V or -2.5V, the two relaxation currents after two consecutive stresses are the same, which indicates that a negligible number of new traps were created under the low voltage stress condition. However, when the stress voltage increased to -3V or -3.5V, the second relaxation current was higher than the first one. This corresponds to the new traps generated during second high-voltage stress. Under the larger gate bias condition, a large amount of electrons arrived at the anode with sufficient energy, which released hydrogen in the interface region. The liberated hydrogen transported into the gate dielectric and generated the hydrogen-induced positive charges [21, 41]. Stress-induced traps can also explain the large difference between the relaxation currents of a fresh device, e.g., I_1 and I_2, and those on the same capacitor after several high-voltage stresses, e.g., I_{11} and I_{12}. Since they were all measured after being stresses at the same -1.5V, this difference indicates

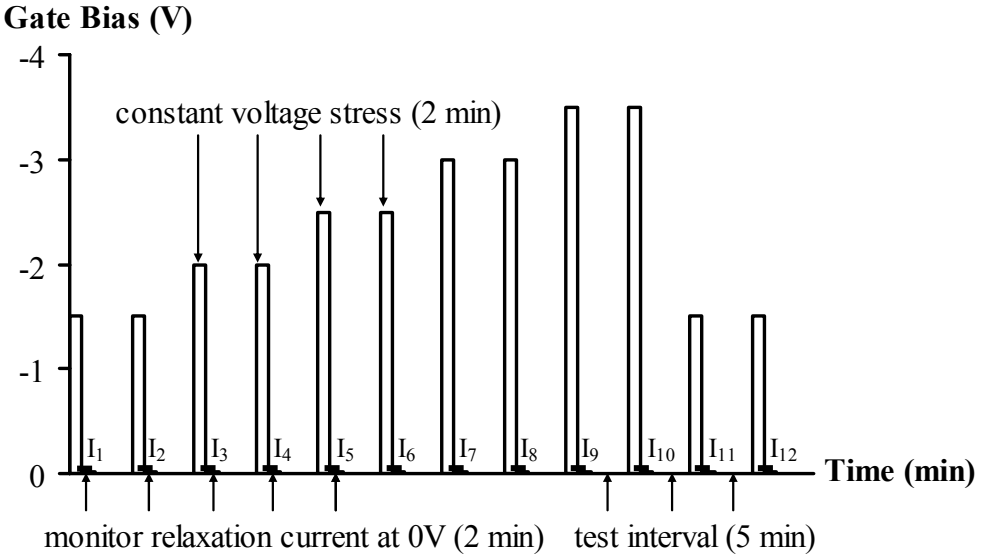


Figure 12. Pulse sequence for the proposed dielectric relaxation test

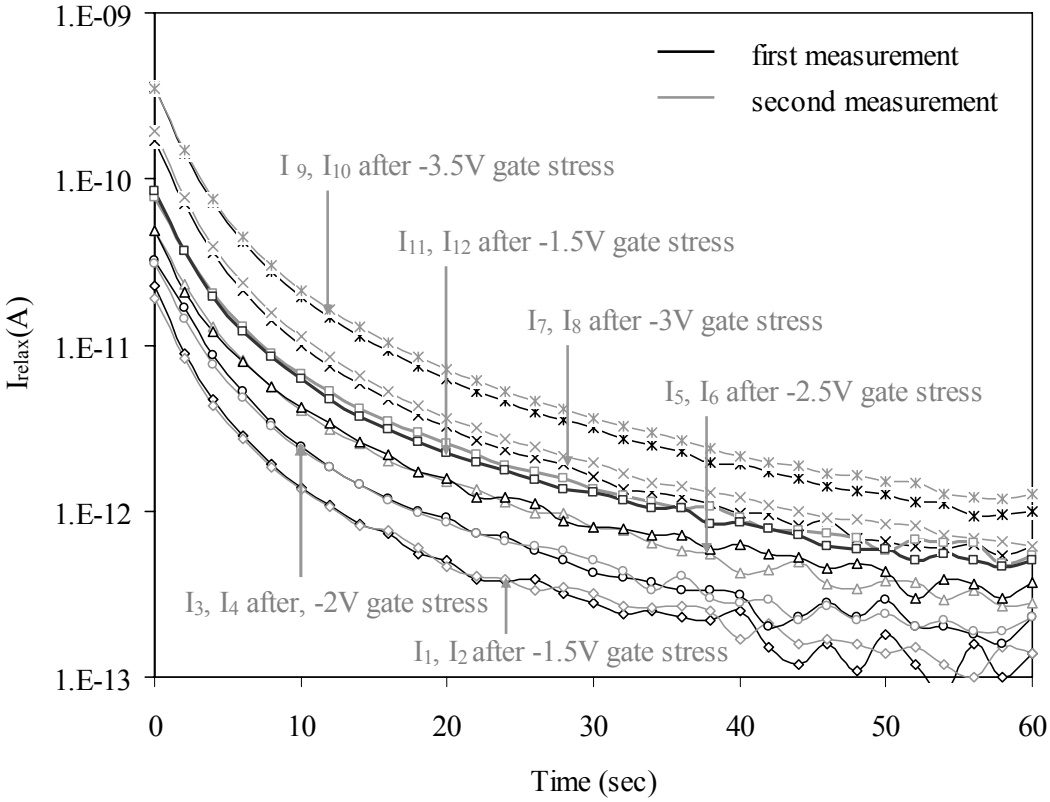


Figure13. Relaxation current decay with time of TiN/ 45s-deposited ZrHfO/1nm SiON/*p*-Si Stack. The stress and measurement methods are shown in Figure 12.

that the stress-induced trap generation is not reversible and the number of the newly created traps could be comparable to or even larger than that of the preexisted traps.

Figure 14 shows the I_{relax} as a function of the stress voltage V_{stress}. The measured relaxation current is shown as the black solid line. The two grey solid lines represent the linear regression fits to the two sets of data with gate bias ranges of (-1.5V to -2.5V) and (-2.5V to -3.5V). The relaxation current increases linearly with the magnitude of the stress voltage with a slope change at -2.5V. The slope becomes larger after -2.5V, which corresponds to new traps generated in the high-k stack during the high-voltage stress. The stress-induced trap creation in Fig. 12 can be calculated from the difference between the measured relaxation current under the gate bias of (-2.5V to -3V) and the prediction value extrapolated from the low-voltage linear model on (-1.5V to -2.5V), which is the black dash line in Fig. 14. The grey dash line represents a linear regression fit to the portion of relaxation current contributed by high-voltage stress induced trap generation. For this particular high-*k* stack, the trap creation rate is about $2.29 \times 10^{13} / \mathrm{cm}^2 \cdot \mathrm{V}$, which is comparable to the literature report [42].

3.3 Charge Trapping in $ZrHfO_X$ / RuO_X / $ZrHfO_X$ films

Figure 15 shows CV curves before and after constant voltage stresses (CVS). The CV measure provides the information on the effect of both bulk charges and interfacial charges near the substrate [39, 43]. There is a negative shift of V_{fb} after -3V CVS due to hydrogen-induced positive charges. Under the negative gate bias condition, a large amount of electrons arrived at the anode with sufficient energy, which released

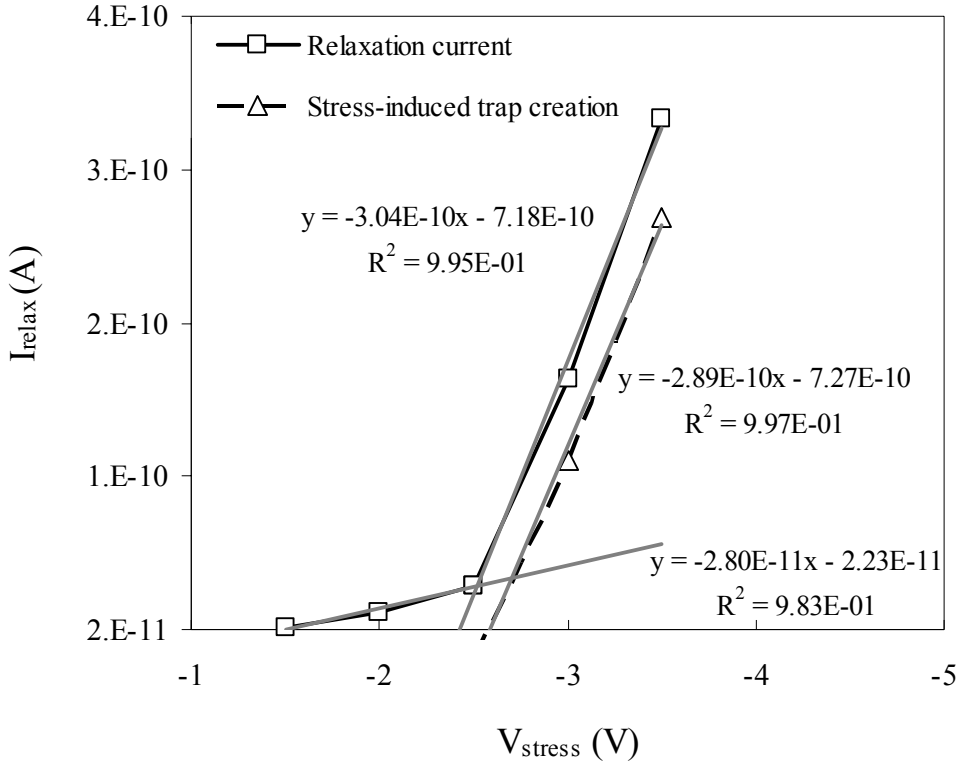
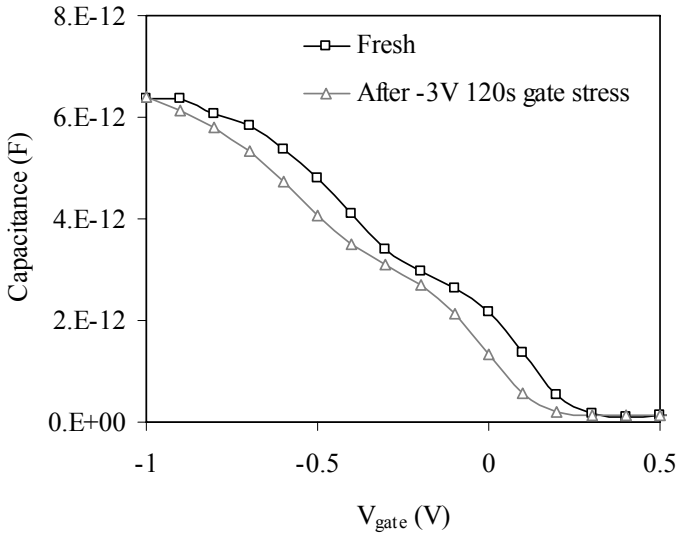
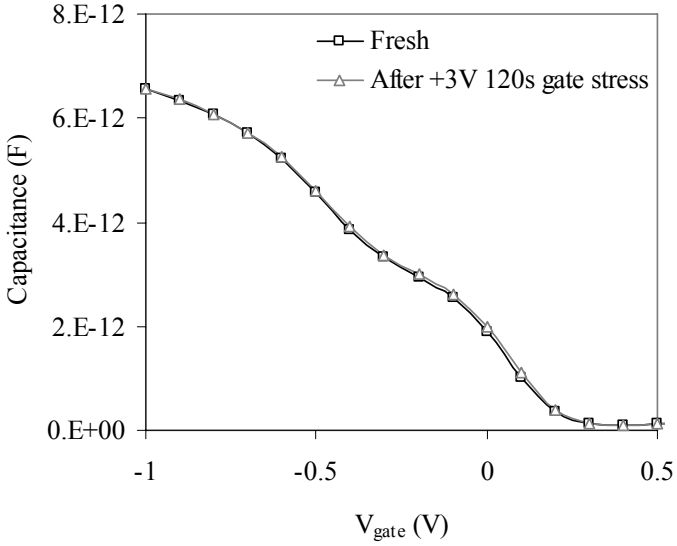


Figure 14. Relaxation current as a function of stress voltage, measured on TiN/ 45s-deposited ZrHfO/1nm SiON/*p*-Si stack with Figure 13 method.



(a)



(b)

Figure 15. CV characteristics of TiN / $ZrHfO_x$ / RuO_x / $ZrHfO_x$ / p-Si gate stacks before and after constant voltage stress tests

hydrogen in the interface region. The liberated hydrogen transported into the gate dielectric and generated the hydrogen-induced positive charges [21, 41]. Figure 15 (b) shows +3V CVS does not induce much trapping due to absence of both holes and hot electrons at the inverted substrate interface [21]. Therefore, hydrogen-induced positive charges trapped in both bulk and interface contribute to the interface state generation and V_{fb} shift when electrons are injected from the gate under a negative gate bias condition; meanwhile no noticeable electron or neutral traps created by anode hole injection degrade much the performance of TiN/$ZrHfO_x$/RuO_x /$ZrHfO_x$/p-Si.

Figure 16 shows CV characteristics of the TiN / $ZrHfO_x$ / RuO_x / $ZrHfO_x$ / p-Si capacitor measured at 100 kHz and 1 MHz. As the probe frequency decreases, the inversion capacitance increases, which was induced by slow states, e.g., positive charges that are located near the interface and able to communicate with the Si substrate over a wide range of time scales [21]. The increase of the capacitance ledge in C-V curve is a support for carrier tunneling phenomena caused by embedding RuO_x into $ZrHfO_x$ because the carrier tunneling effect is usually enhanced by decreasing the signal frequency [44]. No stress-induced leakage current is observed, as shown in Figure 17.

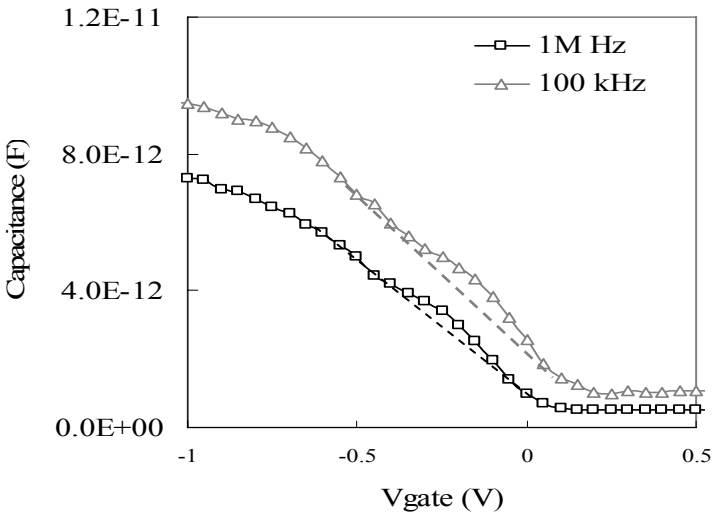
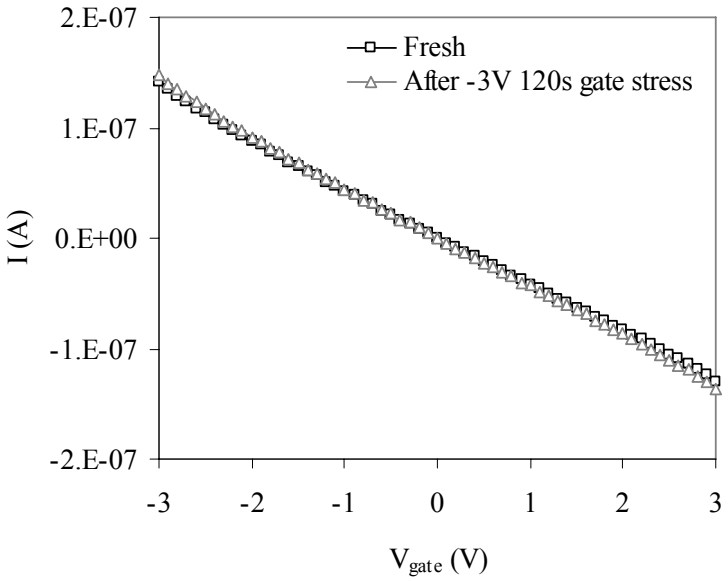
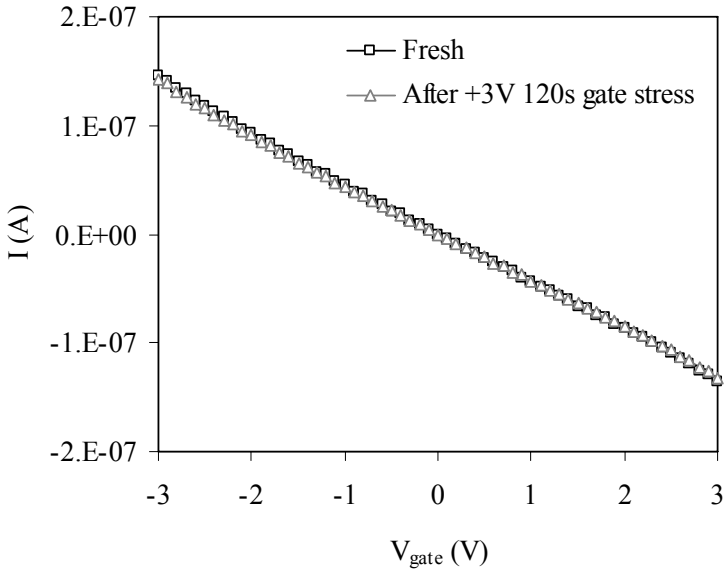


Figure 16. High frequency CV characteristics of $ZrHfO_X/RuO_X/ZrHfO_X$



(a)



(b)

Figure 17. IV characteristics of $ZrHfO_x$ / RuO_x / $ZrHfO_x$ before and after constant voltage stress tests

Figure 19 shows the decays of I_{relax} with time at various test cycles in Figure 18. As shown in Fig. 19 (a), two relaxation currents after consecutive stresses at -1.5 V, -2 V or -2.5 V are almost the same, which indicates a negligible number of new traps were created at the low voltage stress condition. When the stress voltage increased to -3.5 V -4 V or -4.5 V in Fig. 19 (b), the second relaxation current is larger because additional hydrogen-induced traps were generated during the second high-voltage stress [43, 45]. Measured under the same stress conditions, the large difference between I_1, I_2, and I_{21}, I_{22} in Fig. 19 (c) indicates that the stress-induced trap generation is not reversible and the number of the additionally created traps could be comparable to or even larger than that of the pre-existed traps. Although the relaxation current first increase with the stress voltage, it changes little after -4.5V, as shown in Fig. 19 (d). Transient conductivity has been considered as a reference of the charge trapping/detrapping and polarization/relaxation in the high-k layer [21]. There is always a limit in the amount of charges that can be trapped as well as the extent of polarization in a dielectric, which prevents the further increase in relaxation current and causes this saturation phenomenon.

Figure 20 shows relaxation current versus stress voltage. The black solid line represents the measured relaxation current and the grey solid lines are the linear regression fits to the data under gate bias (-1.5V to -3V), (-3V to -4.5V) and (-4.5V to -6V). The relaxation current increases linearly with the magnitude of the stress voltage between -1.5 V and -4.5 V with a slope change at -3V. The larger slope against the gate bias after -3V indicates new trap creation occurs in the interface region or in the bulk high-k film during the high-voltage stresses. Using the method proposed in Ref [46], the trap creation rate can be calculated by the difference between the measured relaxation current under (-3V, -4.5V) and the prediction value

extrapolated from the low-bias linear model on (-1.5V, -3v). It is about $1.6\times10^{12}/\mathrm{cm}^2\cdot\mathrm{V}$ for this particular high-*k* stack.

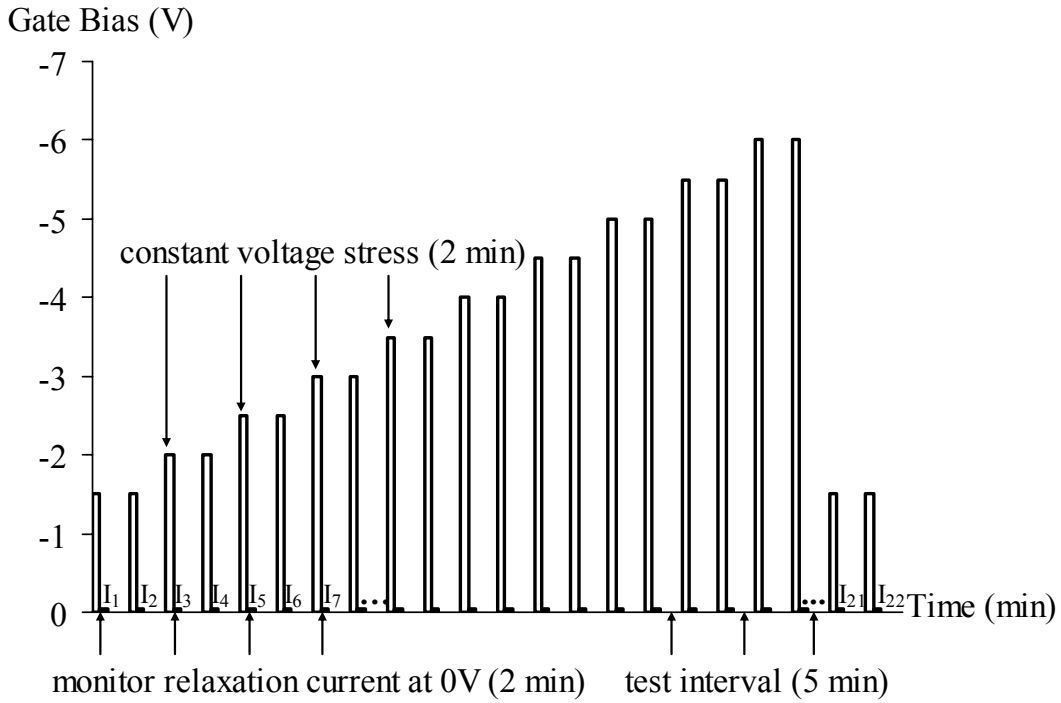
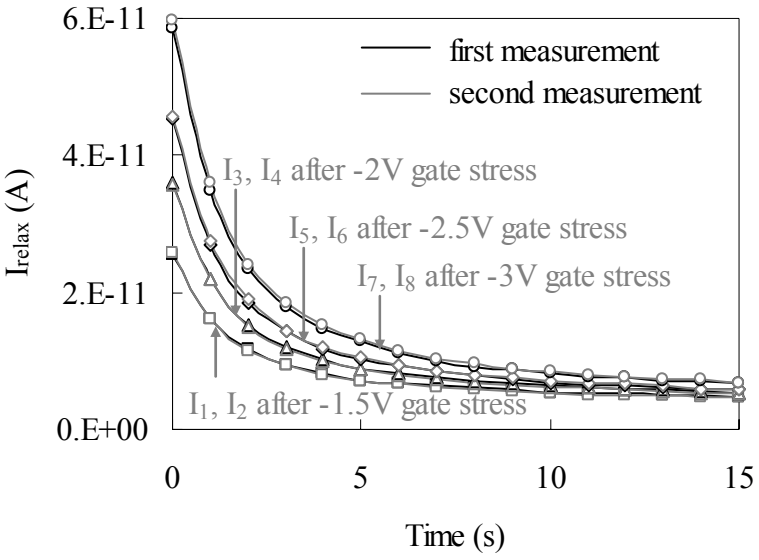
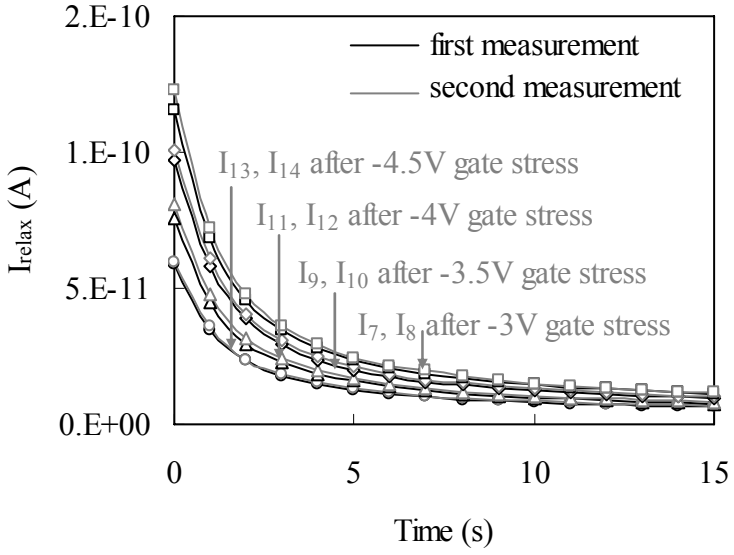


Figure 18. Pulse sequence for the proposed dielectric relaxation test

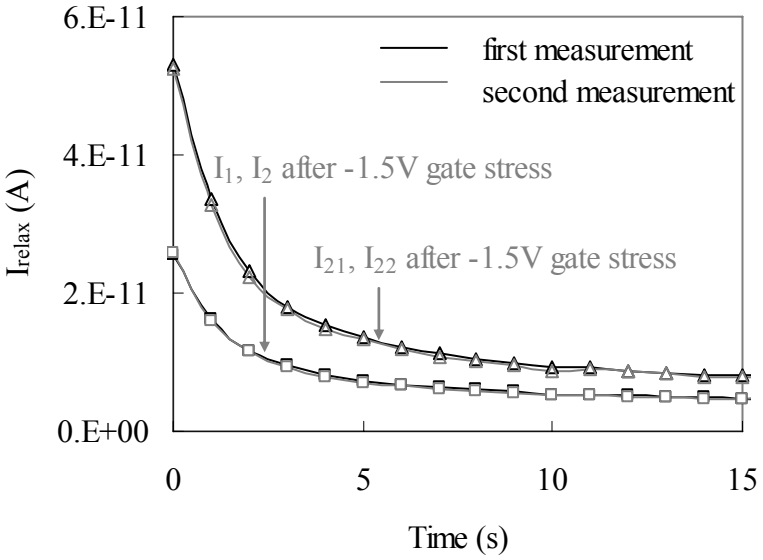


(a)

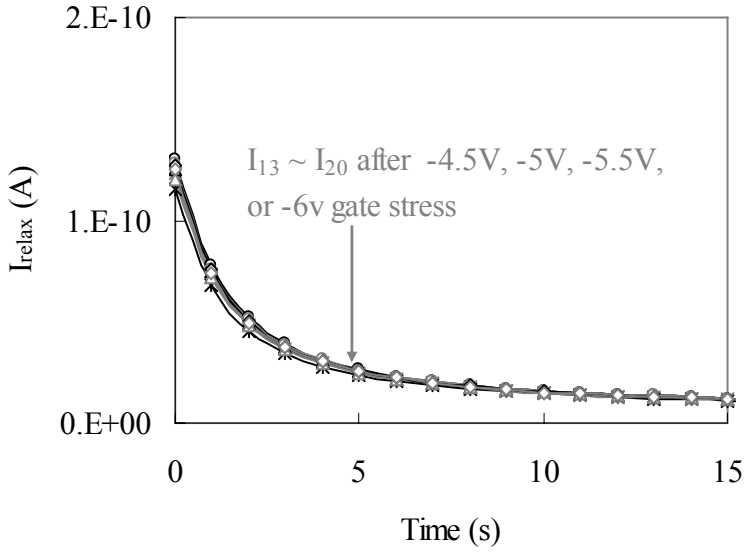


(b)

Figure 19. Relaxation current decay with time of TiN / $ZrHfO_X$ / RuO_X / $ZrHfO_X$ / p-Si stacks. The stress and measurement methods are shown in Fig. 18



(c)



(d)

Figure 19. Continued

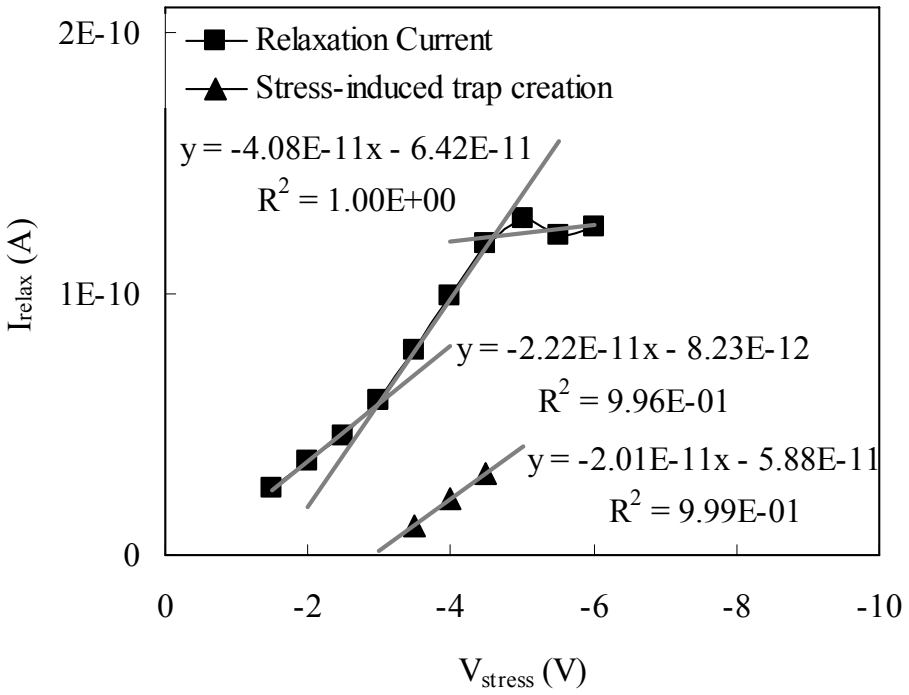


Figure 20. Relaxation current as a function of stress voltage

3.4 Comparisons between $ZrHfO_X$ / RuO_X / $ZrHfO_X$ and $ZrHfO_X$ films

Figure 21 compares the leakage currents and transient currents of the three high-*k* samples, using ramp-relax test. $ZrHfO_x/RuO_x/ZrHfO_x$ tri-layer has a lower EOT, the smallest leakage current and the smallest transient current than $ZrHfO_X$ and Ru $ZrHfO_X$ films. Therefore, $ZrHfO_x/RuO_x/ZrHfO_x$ may be a more promising gate dielectric material than $ZrHfO_x$ and $RuZrHfO_x$. During the stress,

1) New traps are created in $ZrHfO_X$ / RuO_X / $ZrHfO_X$ tri-layer after -3V, while defects are generated in $ZrHfO_X$ films when the stress voltage increased to -2.5V.
2) The defect generation rate of $ZrHfO_X$ / RuO_X / $ZrHfO_X$ film is one order of magnitude less than that of $ZrHfO_X$ films.
3) Saturation phenomenon of relaxation current was only observed on $ZrHfO_X$ / RuO_X / $ZrHfO_X$, while $ZrHfO_X$ has a small voltage-to-breakdown and fails before saturation.

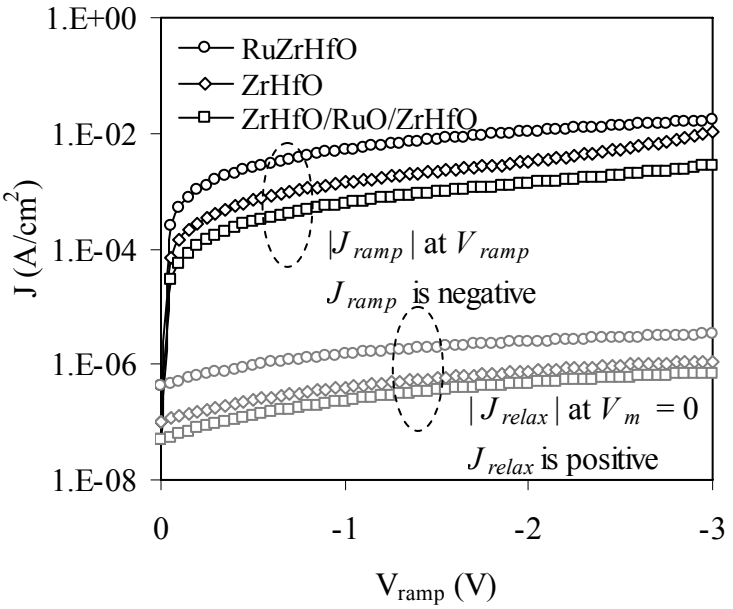


Figure 21. Ramp-relax test results of TiN/ZrHfO/RuO/ZrHfO/p-Si stack

CHARTER 4

DIELECTRIC BREAKDOWN MECHANISMS

For the high-*k* stacks with a 1nm SiON interfacial layer in group I, two types of breakdown phenomena were observed depending on whether the bulk and the interface layers failed simultaneously or separately [21]. Although the breakdown process is due to hot electrons and positive charges trapped in the interface region, the reason why different breakdown sequences occur is still unknown.

4.1 Dielectric Breakdown Detection

Breakdown detection is crucial for characterizing the reliability of gate dielectric films. An overestimation of reliability can result if initial breakdown event is not detected in soft breakdown, while underestimation occurs without consideration of stress induced leakage current. Due to the variety in film thicknesses, gate sizes, test structures, conduction mechanisms, fabrication processes and test equipments, the idea of a universal failure criterion is not practical [5]. Generally, the breakdown criterion is based on either leakage current or relaxation current.

4.1.1 Leakage current criterion

Gate dielectric films have low leakage current to ensure the proper functioning of the MOSFET. In general, this static conductivity is contributed by electron transport

*Reprinted with permission from "Dielectric breakdown and charge trapping of Ultrathin ZrHfO/SiON high-k Gate Stacks" by Rui Wan, Jiong Yan, Yue Kuo and Jiang Lu. Japanese Journal of Applied Physics, vol. 47, p. 1639, 2008.

across the film via a defect band in the bulk. The breakdown criterions based on leakage current include [5]:

1) Predefined level of leakage current
2) Increase of leakage current
3) Slope increase in I-V curve
4) Noise level of leakage current

A sudden increase of leakage current is widely used to detect breakdown in voltage stress tests. The relative change is defined as the ratio of two consecutive currents. This simple definition of breakdown might not be sufficient to identify the first breakdown event and might make it difficult to record both soft and hard breakdown [5]. Soft breakdown in CVS is characterized by a sudden increase in the measurement noise that accompanies the change of leakage current, though the current is well below the level of a hard breakdown [47].

4.1.2 Relaxation current criterion

The relaxation behavior and high-k film integrity are closely related. As introduced in Chapter 2, the sudden disappearance of relaxation current can be used as the signal of a high-k dielectric breakdown. This makes it most useful during RVS tests, in which a high k film fails at a relatively high field and the large leakage current in the background obscures the current jump caused by the breakdown [20]. While the leakage current monitors the breakdown of the whole stack including the bulk high-k layer and the interface layer, the relaxation current only detects the breakdown of bulk high-k layer.

4.2 Dielectric Breakdown Sequence and Mechanism

Figure 22 shows the XPS Si 2*p* core-level spectra of the 20s and 45s-deposited $ZrHfO_x$ samples after PDA. The samples were charge-corrected by aligning the Si substrate spectral feature to 99.3 eV [48, 49]. Both samples contain the 99.3 eV peak contributed by the Si wafer. Although originally both samples have the same SiON interface layer, their final interface compositions are different. For the 45s-deposited $ZrHfO_x$ sample, it contains the 102.4eV peak corresponding to SiON [48]. For the 20s-deposited $ZrHfO_x$ sample, however, the interface contains the 103.3 eV peak corresponding to the SiO_2 bond probably due to the easy diffusion of oxygen through the ultrathin high-k layer to oxidize the original SiON interface. The high-k layer of the 45s-deposited $ZrHfO_x$ sample may be thick enough to delay the oxygen diffusion into the interface and preserve the original SiON interface structure. Therefore, from the XPS data, the interface of the 45s-deposited high-k film has a higher dielectric constant than that of the 20s-deposited high-k film.

Ramp-relax test was used to monitor the relaxation behavior to identify the breakdown sequence of individual layers in the high-k/interface stacks [20, 21, 36]. Figure 23 shows the ramp-relax test results of the two high- k samples. The "+" and "-" stand for the polarity of the relaxation currents J_{relax}. After each ramp step, the bias voltage was removed for 0.5s and the J_{relax} was measured. From the J_{ramp}-V curves, it was estimated that the 20s-deposited sample has the Poole-Frenkel conduction mechanism, while the 45s-deposited sample has a conduction mechanism between the Poole-Frenkel and Schottky emission, which is similar to the report in [50]. The 45s-deposited $ZrHfO_x$ sample breaks down in one step in which J_{relax} and the leakage

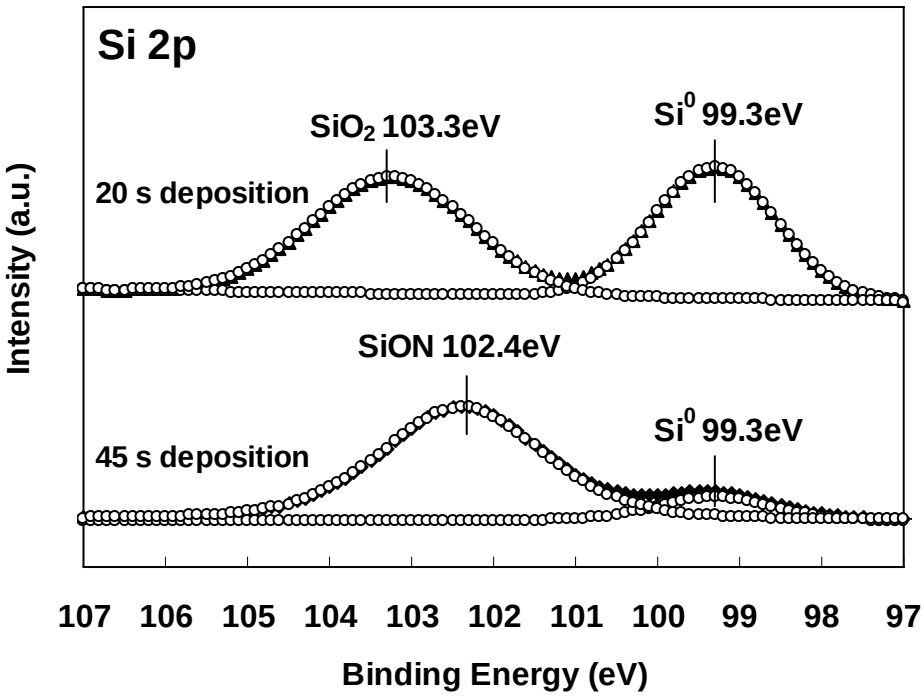


Figure 22. Si 2p Core Level Spectra of 20 s and 45 s-Deposited Zr-doped HfO_2 Films after Rapid Thermal Post-Deposition Annealing in N_2 and O_2 Atmosphere

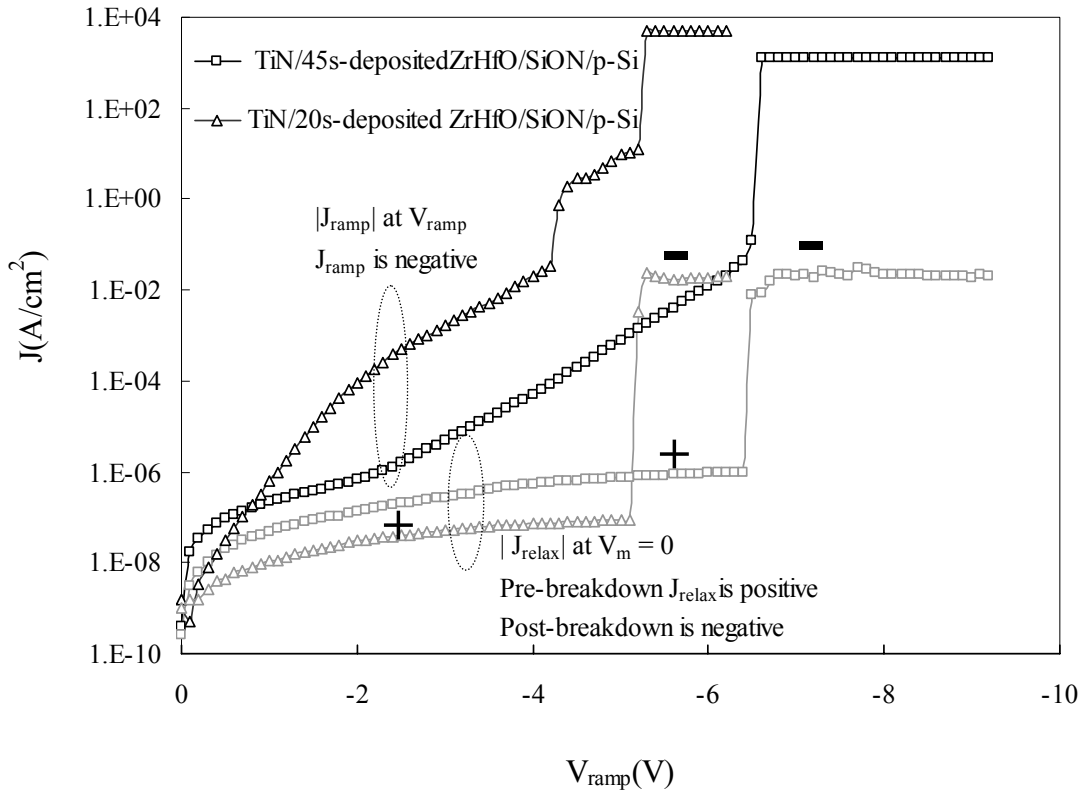


Figure 23. Breakdowns of TiN/20s- and 45s-deposited $ZrHfO_x$/1nm SiON/*p*-Si Stacks in Ramp-Relax Tests.

current J_{ramp} change abruptly at the same ramp voltage V_{ramp}. On the other hand, the 20s-deposited $ZrHfO_x$ sample shows two-step breakdown phenomenon where J_{ramp} jumps slightly at -4.3V first and sharply then at -5.2V.

However, the J_{relax} changes its polarity only once at -5.2V. This indicates that the first jump of J_{ramp} is due to the breakdown of the SiO_2-like interface layer, which has a negligible relaxation current compared with that of the metal oxide high-k film [51]. J_{relax} maintains its magnitude and polarity after the first jump of J_{ramp}, because the $ZrHfO_x$ film still keeps its integrity. After the second jump of J_{ramp}, J_{relax} is in the same polarity as J_{ramp}, because the complete high-k stack was broken and lost its relaxation ability [21]. Few breakdown characteristics with several small jumps were also observed, which were taken as the soft breakdowns. At these points, there were no sharp drastic changes in the J_{ramp}-V and J_{relax}-V curves.

The difference in breakdown sequences of the 20s- and 45s-deposited samples is attributed to the physical thickness difference in the bulk high-k layers and the composition difference in the interface layers. For the double-layer dielectric stack, the voltage drops across the interface layer V_{IL} and the high-k bulk layer $V_{high\text{-}k}$ can be calculated with the following equations [40]

$$V_{IL} = \frac{V_{stack}}{\left(\frac{\varepsilon_{IL} t_{high-k}}{\varepsilon_{Ihigh-k} t_{IL}} + 1\right)} \tag{3}$$

$$V_{high-k} = \frac{V_{stack}}{\left(\frac{\varepsilon_{Ihigh-k} t_{IL}}{\varepsilon_{IL} t_{high-k}} + 1\right)} \quad (4)$$

where ε_{IL} and t_{IL} (ε_{high-k} and t_{high-k}) are dielectric constant and thickness of the interface layer (bulk high-k layer), respectively. Compared with the 45s-deposited sample, the 20s-deposited $ZrHfO_x$ sample has a smaller ε_{IL} and t_{high-k}; therefore, its interface layer is subjected to a higher voltage stress under the same gate bias condition. A negative shift of flat band voltage is observed after a negative gate voltage stress, which indicates positive charge trapping. During the stress, hydrogen-related holes were emitted by hot electron entering the anodic region. Positive charges were generated when these holes reacted with the weak spots, such as oxygen vacancies in the dielectric. The imperfect interfacial layers, e.g., SiON or SiO_2, as well as its interface with $ZrHfO_x$, contain a large number of such defects, which is the possible root cause of breakdown [21, 43]. The higher negative stress on the interface layer accelerates defect generation and leads to its early breakdown. The bulk high-*k* layer remains an insulator until it eventually breaks down at a higher voltage. Although the breakdown of the 45s-deposited $ZrHfO_x$ sample was initiated at its interface layer [52], the voltage drop across the interface layer was smaller due to the relatively thick bulk high-*k* film and the SiON-like interface structure. A larger voltage is required to break the interface layer. Therefore, the bulk high-*k* film and the interface layer break down almost simultaneously and it is difficult to differentiate the break down sequence of these two layers.

4.3 Summary

In summary, the breakdown sequence of the two ultra-thin ZrHfO/1 nm SiON high-k stacks has been investigated. The two-step breakdown phenomenon was observed on the 20s-deposited $ZrHfO_x$ sample and the one-step breakdown phenomenon was detected on the 45s-deposited $ZrHfO_x$ sample. The two-step breakdown follows the same mechanism as the one-step breakdown. When the stack is subjected to gate injection, a large amount of interface states are generated near the substrate due to hot electrons and positive charges trapped in the interface region. These defects lead to the early failure of the interface layer. The difference in the breakdown sequence is attributed to the physical thickness of the bulk high-k layer and the structure of the interface layer. Compared to the 45s-deposited $ZrHfO_x$ sample under the same gate bias condition, the higher stress on the interface of the 20s-deposited $ZrHfO_x$ sample accelerated the defect generation in the interface region, which leads to its early interface breakdown. The bulk high-*k* broke subsequently and separately at a higher voltage.

CHARTER 5

CHARGE TRAPPING OF RUTHENIUM NANOCRYSTAL EMBEDDED HIGH-K FILMS

The nonvolatile memory devices have been continuously scaled down in order to increase writing speed and lower opertating voltages. Ruthenium naoocrystal embedded high-k film is investigated as a replaccement for conventional flash nonvolatile memories in this research: Ru nanocrystal layer improve the data retention characteristic due to its larger work function and hole-based memory mechanism; Zirconium doped hafnium oxide layer imroves the program/retention performance because of its larger physical thickness and the lower electron barrier height than the conventional SiO_2 barrier.

5.1 Capacitance-Voltage Characteristics

The capacitance-voltage (C-V) characteristics of the nc-Ru embedded samples are shown in Figure 24 [68]. The C-V curves were measured from the accumulation region to the inversion region and back to the accumulation region. Ruthenium nanocrystals embedded sample has obvious hysteresis, which can be explained by the reversible storage of either positive (holes) or negative (electrons) charges in the nanocrystals, in the interface between the nanocrystals and high-k dielectric, and/or in the bulk high-k layer [60]. Under the same fabrication condition, the control sample without the embedded nc-RuOx layer has negligible hysteresis. The possibility of contamination with mobile ions or charge trapping in the bulk high-k dielectric layer

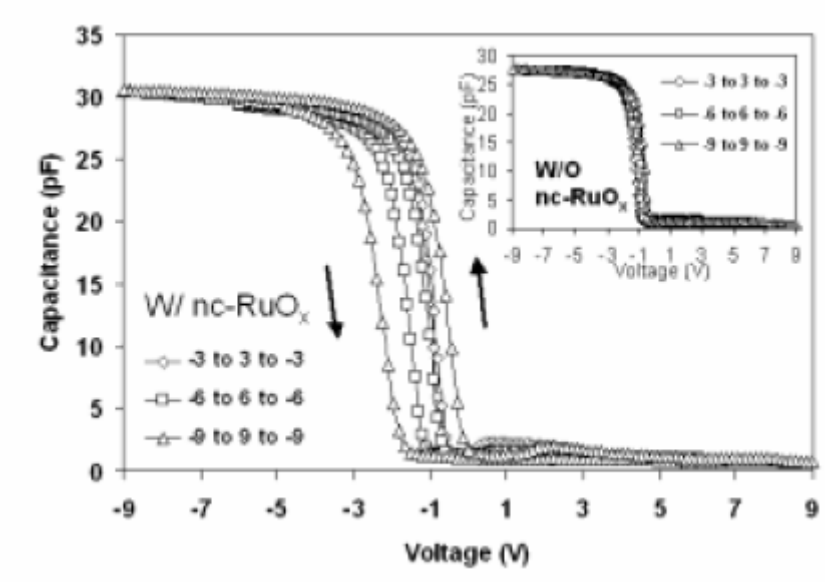
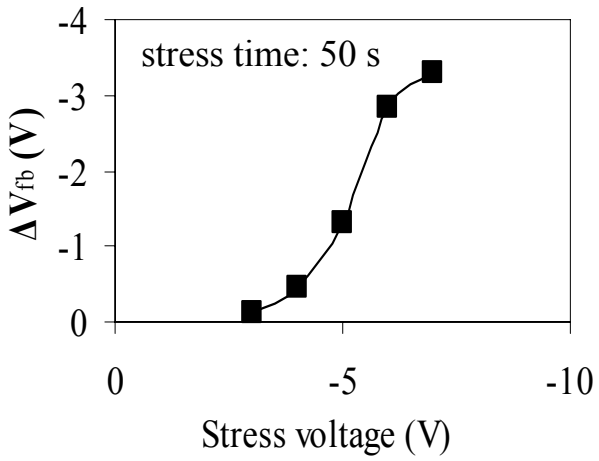


Figure 24. C-V curves for nc-RuOx embedded ZrHfOx films [68]

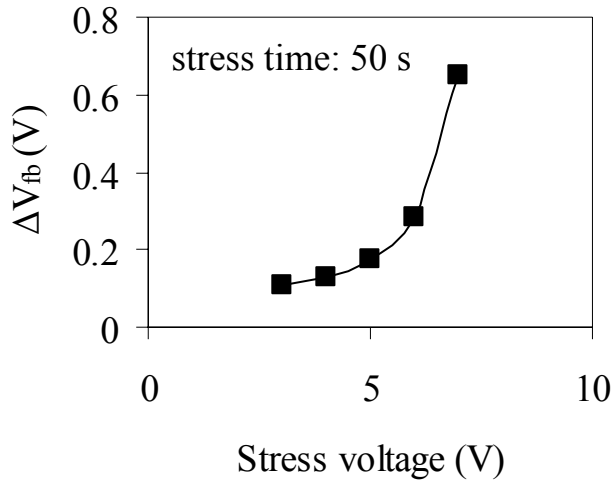
in the sample can be ruled out. It can be concluded that the hysteresis results from the presence of the nanocrystals. The memory window increases with the increasing of gate bias, showing the amount of charges injected increasing, the amount of stocked charges will also be increasing [60].

5.2 Electron and Hole Trapping Mechanisms

Figure 25 shows the flatband voltage shift ΔV_{fb} after 50-second voltage stresses. Different voltages and polarities were applied to study the opening of the voltage window for two different charge states. Obviously larger flatband voltage shift is observed under negative gate bias, which indicates positive charges were easily trapped in the nc-RuOx and the memory window was mainly contributed by the negative flatband voltage shift due to net hole trapping. Figure 25 (b) shows ΔV_{fb}



(a)



(b)

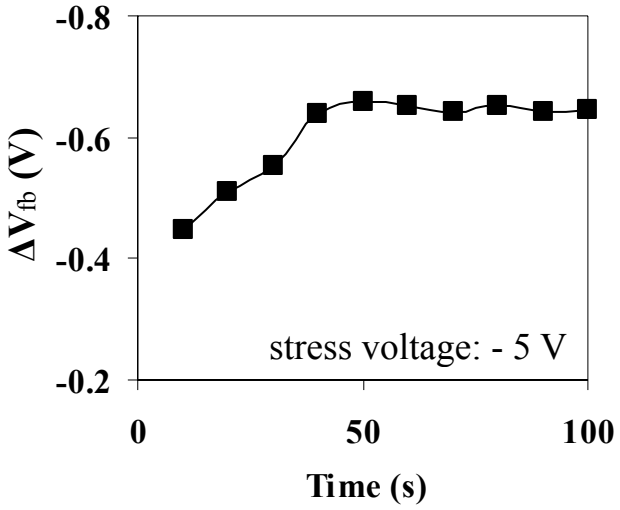
Figure 25. Flat band voltage shift after 50-second gate voltage stresses

might be influenced by electron injection in large positive gate voltage range. Since the low operation voltage is desirable in nonvolatile flash memory application, the nc-RuOx embedded capacitors can be used to trap and detrap holes with a low bias voltage [69].

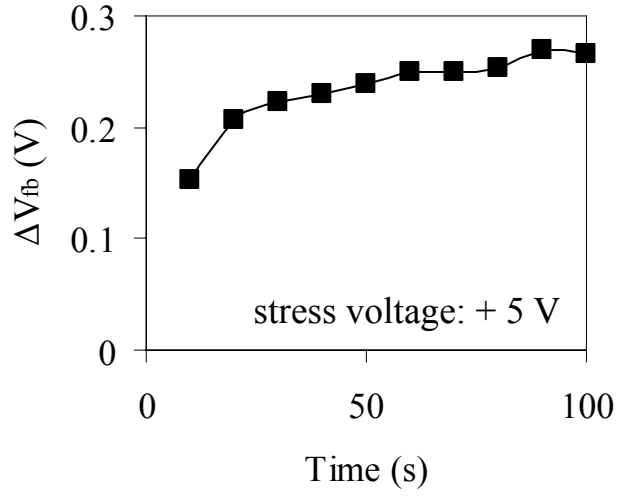
Charge trapping characteristics of nc-RuOx embedded ZrHfOx capacitors were examined by constant gate voltage stresses for different periods of time. Flatband voltage shifts under negative gate bias are consistently larger than those under positive gate bias for the same stress time, which indicates hole trapping dominates over electron trapping in nc-RuOx. When a -5V gate stress was applied, Figure 26 (a) shows the flatband voltage shift increases with the gate stress time, which implies the hole trapping is time dependent. The flatband voltage shift or hole trapping saturates after 50 seconds, suggesting that a limited amount of charges is allowed in each nc-Ru (Coulomb blockade effect) [60].

5.3 Charge Trapping Characteristics

Figure 27 and Figure 28 shows the decay of relaxation current with time after negative and positive gate bias stresses, respectively. The relaxation current was measured immediately after the removal of the stress voltage. Figure 29 summaries the relationship between the relaxation current and the stress voltage. The following conclusions might be summarized from Fig. 27-29.

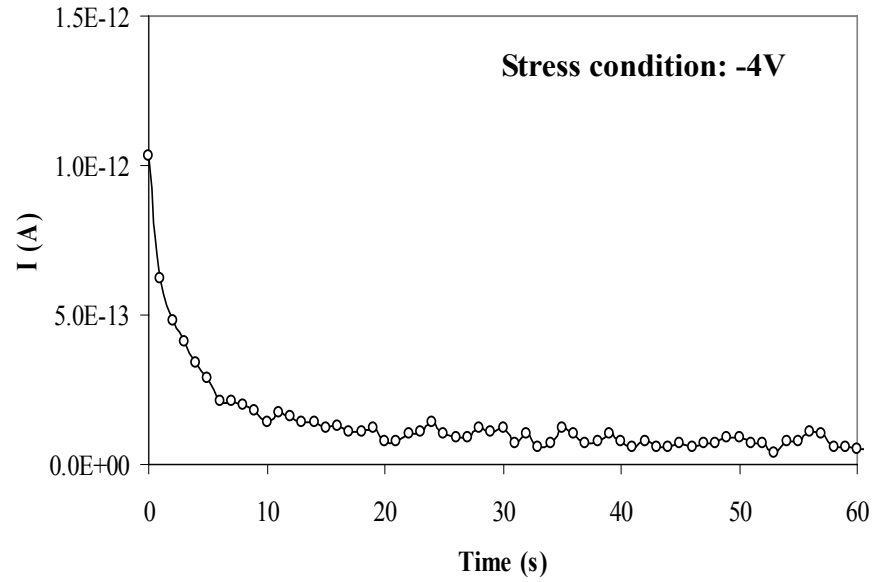


(a)

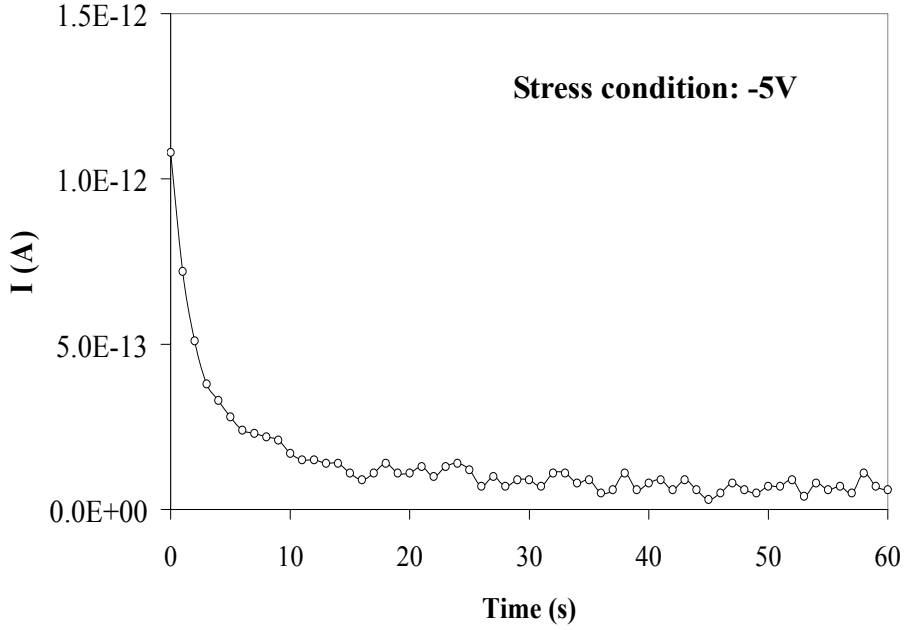


(b)

Figure 26. Flat band voltage shift as a function of gate stress time

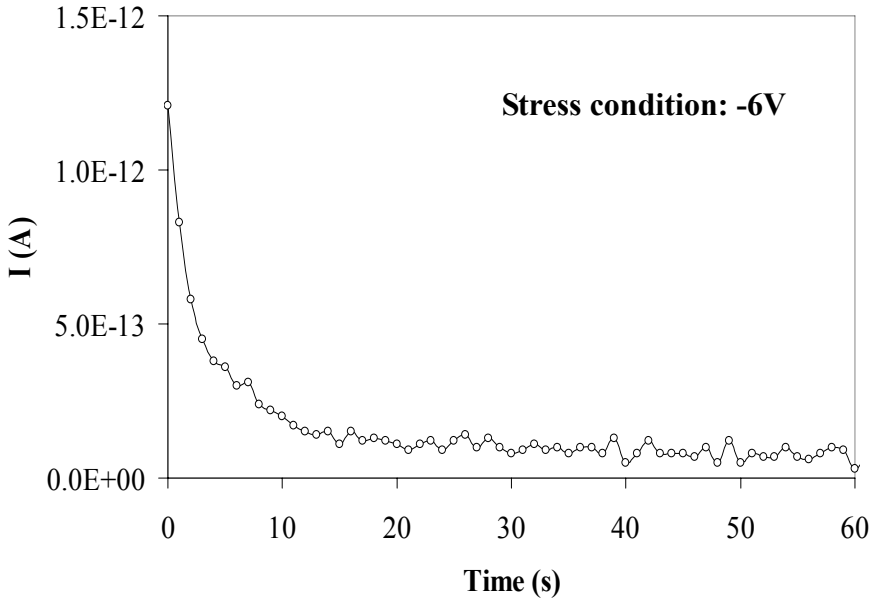


(a)

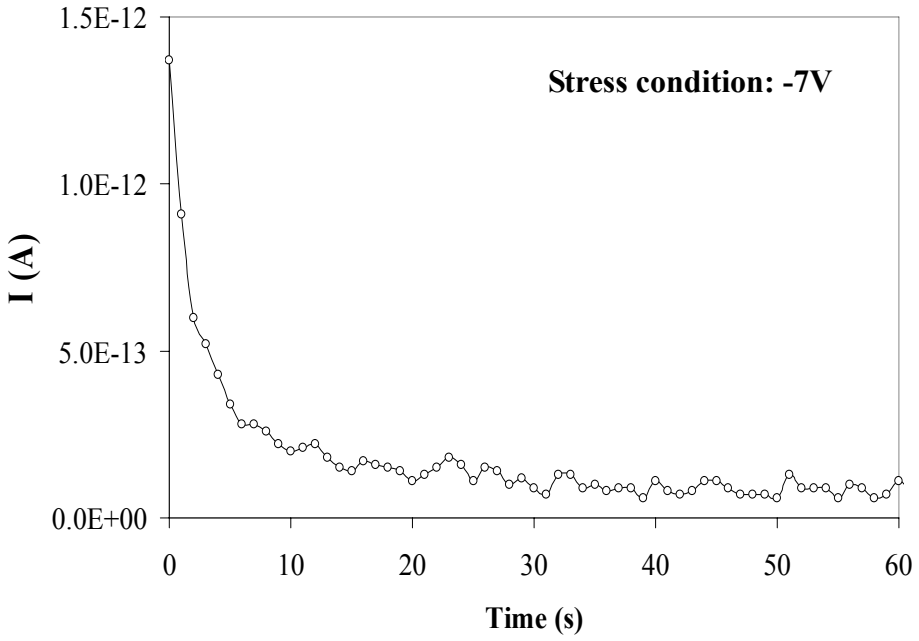


(b)

Figure 27. Relaxation current decay with the time of nc-RuOx embedded ZrHfOx film after 120-second negative gate bias stresses

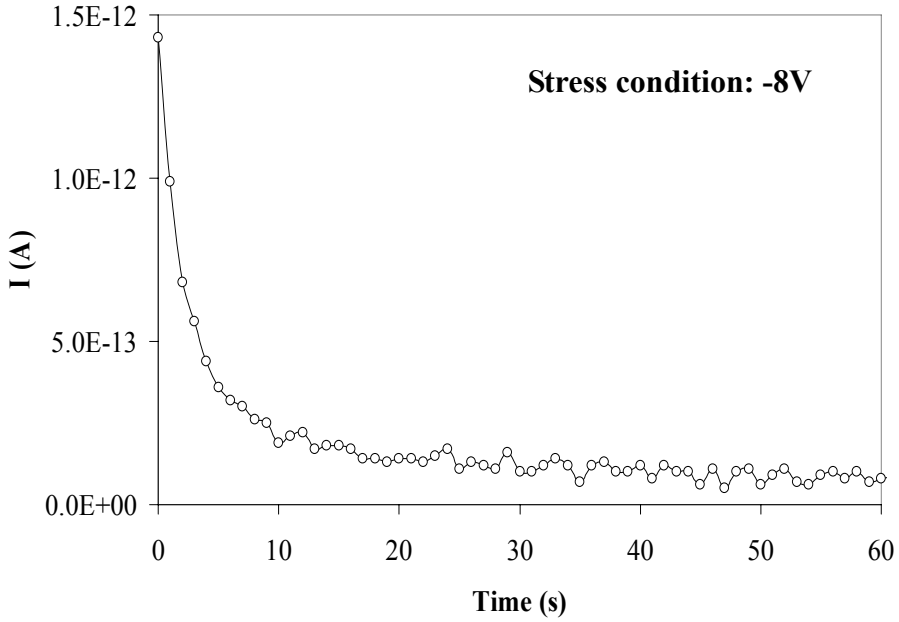


(c)



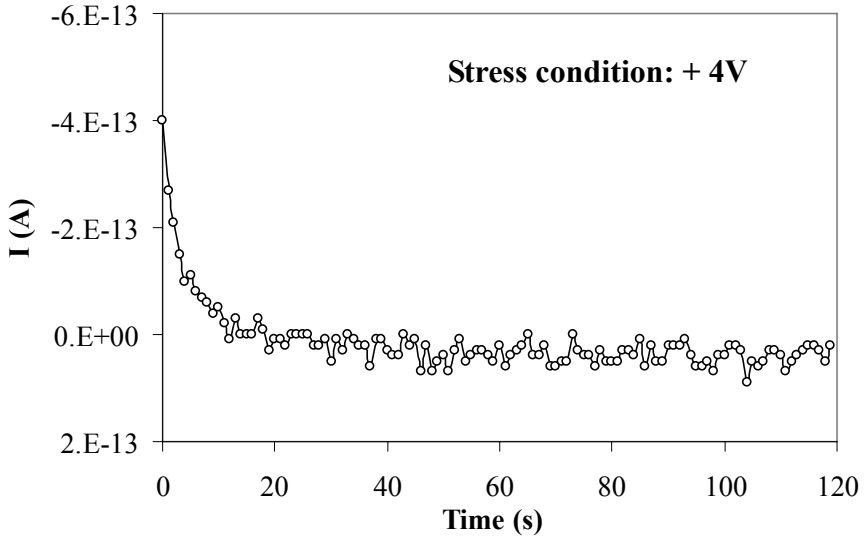
(d)

Figure 27. Continued



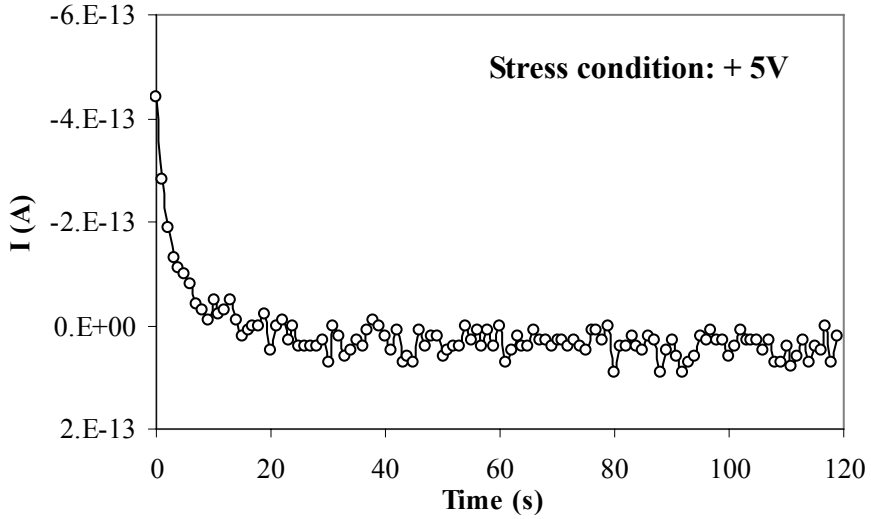
(e)

Figure 27. Continued

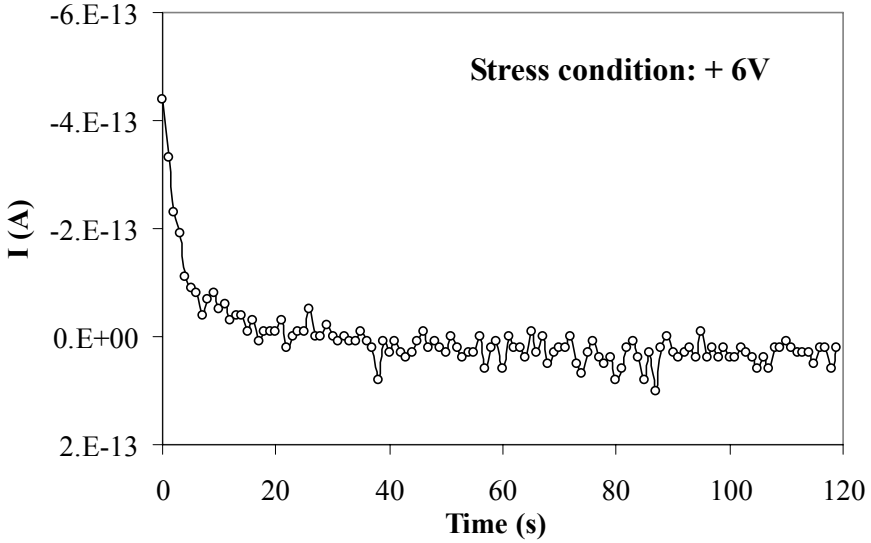


(a)

Figure 28. Relaxation current decay with the time of nc-RuOx embedded ZrHfOx film after 120-second positive gate bias stresses



(b)



(c)

Figure 28 Continued

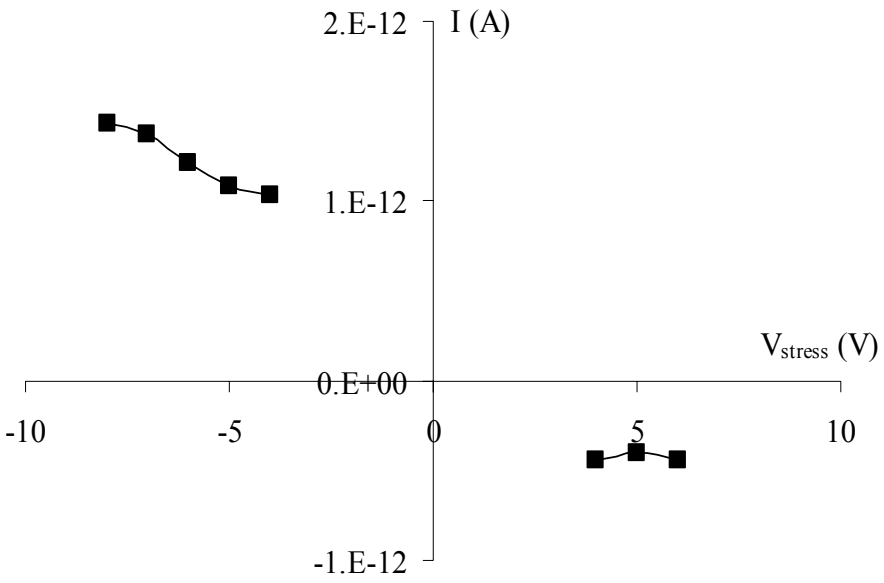


Figure 29. Relaxation current versus stress voltage on nc-RuOx embedded ZrHfOx films

1) Relaxation current increases with negative stress voltage
2) The relaxation current after a negative voltage stress is obviously larger than that after a positive stress with same voltage level
3) Relaxation current is independent of positive stress voltage in a low gate voltage range (-4V, -6V)

The relaxation current after a negative gate bias stress is attributed to the detrapping of loosely trapped holes during the stress [18]. The number of the loosely trapped charges can be calculated from the relaxation current. For instances, the densities of the released traps are $1.14\times10^{11}\,\mathrm{cm}^{-2}$ and 0.96×10^{11} after a 120-second -6V and -8V gate stresses, respectively.

The total charge trapping density Q in nc-RuOx embedded film can be estimated by [70]

$$Q = \frac{C_{fb}\Delta V_{fb}}{q}$$

where q is the electron charge ($1.6\times10^{19}C$). Q is about $1.4\times10^{12}\,\mathrm{cm}^{-2}$ in the sweep of range ± 9 V. The sheet of nc-RuOx was about $8\times10^{11}\,\mathrm{cm}^{-2}$ determined by TEM in Ref [68].

Since the total charge trapping density is obviously larger than the nanocrystal density, hole trapping might occur not only in the nc-RuOx layer but also in the interface between the nc-RuOx and ZrHfOx dielectric due to Coulomb blockade effect. The loosely trapped charges identified by relaxation current could be more likely located in the interface than in nc-RuOx. However, the difference between the total charge trapping density and the nanocrystal density is about $6\times10^{11}\,\mathrm{cm}^{-2}$, which

might not be completely explained by the loosely trapped charge density identified by relaxation current ($\sim 1\times 10^{11}\,cm^{-2}$). A possible speculation is that not all of charges trapped in the interface will be released after the removal of stress voltage. In other words, some of charges trapped in the interface between the high-k dielectric and nc-RuOx will remain in the device after the removal of the stress, and not all of them could be recognized as loosely trapped charge.

CHAPTER 6
BAYESIAN APPROACH TO BATHTUB FAILURE RATE ESTIMATION

6.1 Sample Time-to-Failure Data

Within a reasonable time and at an acceptable cost, it is unlikely to directly test whether a reliability specification for normal operation is satisfied. First of all, the lifetime under the operating condition could be in years; secondly, a test structure has a much smaller gate area to avoid process-induced defects; and finally, a large sample size is needed to observe failure times at a very low percentile [5].

The alternative is to use failure data collected in accelerated life tests (ALT) to project the reliability under the operating conditions. In the ALT, devices are subjected to higher than normal level of a stress to speed up the degradation process. ALT is usually designed and performed at several stress levels, in order to project the reliability under the normal operating condition and decide the optimal burn-in stress level under certain cost and time constraints. To collect time-to-breakdown data, constant voltage stress (CVS) tests were performed on $HfTaO_x$ films in this study. The ordered breakdown times, first presented in [5], are given in Table 2. The test capacitors were randomly allocated to a stress level. Each was tested individually and broke down independently. The electric stresses were chosen from 7.7 to 8.1MV/cm to ensure linear acceleration factor and identical failure mechanism at each stress level.

Table 2. Ordered time-to-breakdown of MOS devices [5]

Electric field	Ordered time-to-breakdown (s)
8.1 MV/cm	1 5 10 10 10 11 14 20 28 68 78 92 93 116 125 128 141 167 219 311 407 502 506 682 709 772 1209 1282 1485 1638 2154 2443
7.9 MV/cm	1 2 9 12 35 46 72 74 82 107 142 153 193 251 290 348 399 511 556 1104 1509 1535 1756 2376 2843 3140 3514 3616 3882 4583
7.7 MV/cm	9 18 20 25 29 66 124 127 175 221 249 341 362 552 630 760 782 794 906 932 968 1378 1386 1664 1728 2229 2249 2338 4058 4986 6312 6400 6847 8474

Although breakdown is a random event that cannot be exactly predicted in advance, there are deterministic acceleration models which relate the lifetime as a function of the applied electric stress. The most commonly used model for voltageacceleration is the inverse power law, as shown in Figure 30. Let t_i and t_j be the failure times at two accelerated electric field E_i and E_j, respectively. Then the inverse power law is

$$t_j = \left(\frac{E_j}{E_i} \right)^{\kappa} t_i \tag{5}$$

where κ is material-specific exponent [51].

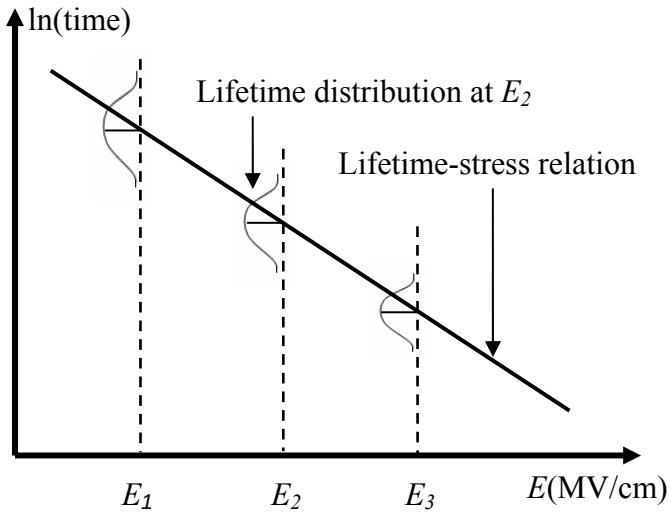


Figure 30. Relationship between the failure distributions in ALT.

6.2 3-Step Bayesian Approach to Bathtub Failure Rate Estimation

In the context of reliability, time-to-failure data are usually collected in the accelerated life tests. In this research, a group of voltage stress acceleration tests were conducted on the next generation high-*k*-based capacitors at three voltage stress levels. Different from separately analyzing the failure rate at each stress level in the literature, this research considered the relationship between the lifetime and voltage stress level. A three-step procedure is proposed to estimate the change point and DFR simultaneously when the sample size is not large. It includes:

1) change point estimation through exponentiality test on the time-to-failure data under the neighborhood stresses;
2) derivation of priors for DFR estimation from time-to-failure data under the

neighborhood stresses;

3) a Bayesian process to estimate the DFR.

Since the posterior likelihood is hardly sharply peaked for a small sample size, the prior distributions are critical in this case. With the consideration of the electrical relationship between the lifetime and voltage stress level, the priors for DFR estimation were mathematically derived from the time-to-failure data under neighborhood stresses in ALT. Thus, it can provide a fast and reliable estimation even when only a small sample of data is available at each stress level. In practice, this gives the manufacturer great advantages in terms of time and cost savings.

6.3 Exponentiality Test and Change Point Estimation

To estimate the change point at stress E^*, a goodness-of-fit test on the exponential distribution [5] is applied to the time-to-failure data under two neighborhood voltage stresses $E^* \pm \delta$, $\delta > 0$. The purpose is to determine the order j for each stress level at which the sample is partitioned into two parts, $N_1 = (t_{(1)},..., x_{(j-1)})$ and $N_2 = (t_{(j)},...,t_{(n)})$, where N_2 is tested for exponentiality against the alternatives of DFR and IFR. The idea of using the exponentiality test for sample partition can be illustrated with Figure 31 [5]. It should be noted the boundaries of the lifetime under stress E^* can be roughly estimated from the neighborhood data under stress $E^* \pm \delta$ based on the inverse power law (6). With an appropriate δ, the mean or standard deviation of time-to-failure before or after t_c at E^* is no less than its corresponding statistic at $E^* + \delta$, but no more than its corresponding statistic at $E^* - \delta$. Later, these boundaries can be also used to the priors for DFR estimations from a practical

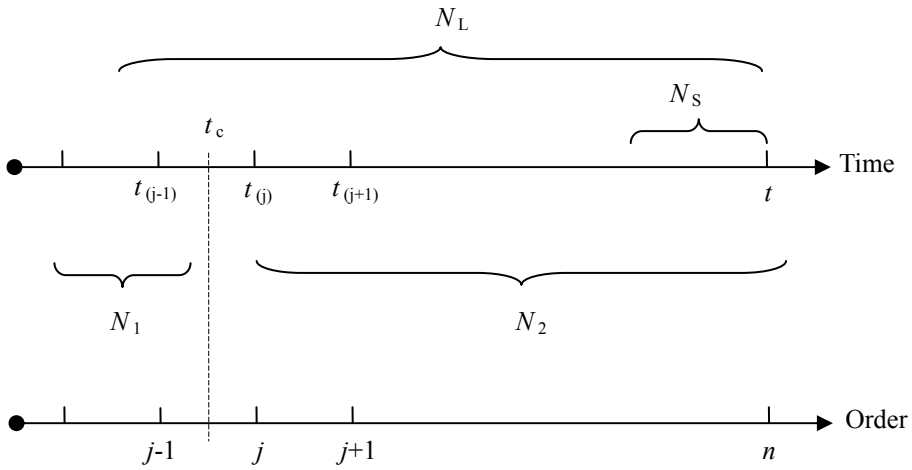


Figure 31. Illustration of the partition of an ordered data set [5]

method proposed in [62].

The ordered data set of failure time $(t_{(1)},...,t_{(n)})$ consists of two sections, the early N_1 of DFR and the latter N_2 of CFR. In case of N_L where the last several data points of DFR are wrongly taken as from CFR, N_L shows a DFR instead of the supposed CFR. N_S is the opposite case where a few data points of CFR are mistaken as from DFR, then N_S has a CFR or possibly IFR. Theoretically, a reasonable approximation of change point ($t_c = (t_{(j-1)} + t_{(j)})/2$) for a given data set should be the sample point $t_{(j)}$ that leads to the N_2 fitting well with an exponential distribution; if there are multiple attempted partition points that result in good fitting of N_2, then the point of a lower or the lowest order is recommended.

To test the goodness-of-fit of the t-set on Exp (t_c, λ), the $\mathbb{N}$ transformation and the $\mathbb{K}$ transformation are often made on the lifetime data. Suppose $t_{(i)}$, i = 1, … , n, is an ordered sample from Exp (t_c, λ). After $\mathbb{N}$ transformation, the normalized spacing

$$t'_{(i)} = (n+1-i)(t_{(i)} - t_{(i-1)}) \tag{7}$$

is independently and identically distributed as Exp (0, λ), where i = 1, … , n and $t_{(0)} = 0$. Through $\mathbb{K}$ transformation,

$$t''_{(i)} = \frac{\sum_{k=1}^{i} t'_{(k)}}{\sum_{k=1}^{n} t'_{(k)}}, \qquad i = 1, \ldots, n\text{-}1 \tag{8}$$

is an order sample of size n-1 from the uniform distribution U (0,1). With the $\mathbb{N}$ and $\mathbb{K}$ transformations, the exponentiality test on the t-set is equivalent to testing the uniformity of the t"-set.

The simplest test for the uniformity of the t"-set is based on $\overline{T}$, the mean of the t"-set. The distribution of $\overline{T}$ converges quickly to the normal. Because $\overline{T}$ tends to be large for an original IFR sample and to be small for a DFR sample, $\overline{T}$ can be used as a one-tail statistic to guard against the alternative of either IFR or DFR [5]. Table 3 summaries the results of sample partition and the occurrence interval at E = 7.9 MV/cm as an example.

Based on the inverse power law, the material-specific exponent κ and $t_c\big|_{E^*}$ could be estimated from (9)-(10).

$$\kappa = \log_{(E^*+\delta)/(E^*-\delta)}^{t_c\big|_{E^*+\delta}/t_c\big|_{E^*-\delta}} \tag{9}$$

$$t_c\big|_{E^*} = \frac{1}{2}\left(f(E^* + \delta) + f(E^* - \delta)\right)$$
$$= \frac{1}{2}\left(t_c\big|_{E^*+\delta} \Big/ \left(\frac{E^* + \delta}{E^*}\right)^{\kappa} + \left(\frac{E^*}{E^* - \delta}\right)^{\kappa} t_c\big|_{E^*-\delta}\right) \quad (10)$$

For illustration, Table 4 shows the procedure of change point estimation at E = 7.9MV/cm.

Table 3. Sample partitions at E = 8.1 and 7.7 MV/cm, and occurrence interval at E = 7.9 MV/cm

Sample Partition					
E (MV/cm)	j	N_1		N_2	
		Mean (s)	S.D. (s)	Mean (s)	S.D. (s)
8.1	14	33.846	35.066	789.263	716.513
7.7	13	117.000	109.383	2578.864	2438.318

Occurrence Interval				
E (MV/cm)	Mean (s)		S.D. (s)	
	$N_{Weibull}$ (L_M, U_M)	N_{exp}	$N_{Weubull}$ (L_S, U_S)	N_{Exp}
7.9	(33.846, 117.000)	(789.263, 2578.864)	(35.066, 109.383)	(716.513, 2438.318)

Table 4. Change point estimation at E = 7.9 MV/cm

E (MV/cm)	J	$t_c = \frac{(t_{(j)} + t_{(j-1)})}{2}$	κ	$f(E^* + \delta)$	$f(E^* - \delta)$	$t_c\big\|_{E^*}$
$E^* + \delta = 8.1$	14	104.5s	-23.95	190.18s	190.20s	190.19s
$E^* - \delta = 7.9$	13	351.5s				

CHAPTER 7

PRIORS AND POSTERIOR SIMULATION

7.1 Prior Derivation

Generally, posterior likelihood is less sharply peaked with a small sample size or a large number of parameters. In this case, the prior distributions are critical. Three sources of information can be considered in determining the priors in this study. They are

1) reports on SiO_2 films in literature,
2) experimental results on other high-*k* films in recent publications,
3) failure data under other accelerated stress conditions in this study [5].

From the engineering point of view, dielectric breakdown is sensitive to the structure and fabrication of film stacks. As a material, structure and deposition techniques vary among SiO_2 and high-*k* films prepared in different research labs, it is challenging work to combine information from various sources and utilize it in a systematic way. Hence, only the third source, from the capacitors fabricated under identical conditions but tested under different accelerated stresses, is the most appropriate and reliable to derive the priors in this study.

Weibull distribution Weibull (α, β) is assumed to describe the DFR

$$f(x) = \beta\alpha^{\beta} t^{\beta-1} e^{-(\alpha t)^{\beta}}, \alpha > 0, \beta > 0$$

where $1/\alpha$ is the 63.2^{th} percentile life, and the shape parameter β is a measure of the

critical defect density, which is unique for each failure mechanism [63, 64].

In the Bayesian context, α and β are treated as random variables which are given their corresponding prior distributions. Given the occurrence intervals for the mean [L_M, U_M] and standard deviation [L_S, U_S] in Chapter 6, the expected values of M and S could be estimated by the inverse power law:

$$E(M) = \frac{1}{2}\left(L_M \Big/ \left(\frac{E^* + \delta}{E^*} \right)^{\kappa_M} + \left(\frac{E^*}{E^* - \delta} \right)^{\kappa_M} U_M \right) \tag{11}$$

$$E(S) = \frac{1}{2}\left(L_S + U_S \right) \tag{12}$$

where $\kappa_M = \log_{(E^*+\delta)/(E^*-\delta)}^{L_M/U_M}$.

The mean and standard deviation can be written as follows.

$$E(M) = E(t|\alpha, \beta) == \frac{1}{\alpha}\Gamma(1 + \frac{1}{\beta}) \tag{13}$$

$$E^2(S) = V(t|\alpha, \beta) == \frac{1}{\alpha^2}\left(\Gamma(1 + \frac{2}{\beta}) - \Gamma^2(1 + \frac{1}{\beta}) \right) \tag{14}$$

The estimations of α and β could be obtained by solving (11) ~ (14).

$$\hat{\beta} = f_\beta\left(E(M), E(S) \right) = \left\{ x \left| \frac{\Gamma(1 + \frac{1}{\beta})}{\left(\Gamma(1 + \frac{2}{\beta}) - \Gamma^2(1 + \frac{1}{\beta}) \right)^{0.5}} - \frac{E(M)}{E(S)} = 0 \right. \right\} \tag{15}$$

$$\hat{\alpha} = f_{\alpha}\left(x, E(M), E(S)\right) = \frac{\Gamma\left(1 + \frac{1}{f_{\beta}(E(M), E(S))}\right)}{E(M)} \tag{16}$$

Alternative prior distributions are selected from truncated uniform, normal, and gamma densities for further sensitivity analysis of posterior inference. For simplicity, the same parameter notation is used for each random variable just to show the form of its density function.

Case 1: $\alpha \sim$ Normal (a, b) and $\beta \sim$ Normal (c, d)

Since α has more influence on the mean and the shape parameter β has more influence on the standard deviation, the boundaries of mean and standard deviation are used to roughly estimate the variance of α and β, respectively, based on the extension of Camp-Meidel to the inequality of Tchebychev [60].

$$V(\alpha) = \frac{\left(\left|\hat{a} - f_{\alpha}(L_M, E(S))\right|/3\right)^2 + \left(\left|\hat{a} - f_{\alpha}(U_M, E(S))\right|/3\right)^2}{2} \tag{17}$$

$$V(\beta) = \frac{\left(\left|\hat{\beta} - f_{\beta}(E(M), L_S)\right|/3\right)^2 + \left(\left|\hat{\beta} - f_{\beta}(E(M), U_S)\right|/3\right)^2}{2} \tag{18}$$

Then, the prior parameters can be estimated as follows:

$$a = \hat{\alpha}, \quad b = V(\alpha)$$

$$c = \hat{\beta}, \quad d = V(\beta)$$

Case 2: $\alpha \sim$ Normal (a, b) and $\beta \sim$ Gamma (c, d).

Similarly,

$$a = \hat{\alpha} = \frac{\Gamma\left(1 + \frac{1}{f_\beta(E(M), E(S))}\right)}{E(M)}$$

$$b = V(\alpha) = \frac{\left(\left|\hat{a} - f_\alpha(L_M, E(S))\right|/3\right)^2 + \left(\left|\hat{a} - f_\alpha(U_M, E(S))\right|/3\right)^2}{2}$$

For the other parameter β, $\beta \sim$ Gamma (c, d), we have

$$E(\beta) = \hat{\beta} = \frac{c}{d} \tag{19}$$

$$V(\beta) = \frac{\left(\left|\hat{\beta} - f_\beta(E(M), L_S)\right|/3\right)^2 + \left(\left|\hat{\beta} - f_\beta(E(M), U_S)\right|/3\right)^2}{2} = \frac{c}{d^2} \tag{20}$$

Solve (19)~(20), we obtain

$$c = \frac{\hat{\beta}^2}{V(\beta)}$$

$$d = \frac{\hat{\beta}}{V(\beta)}$$

Case 3: $\alpha \sim$ Uniform (a, b) and $\beta \sim$ Uniform (c, d).

Again, boundaries of mean and standard deviation are used to develop the prior parameters of α and β, respectively.

$$a = \frac{\Gamma\left(1 + \frac{1}{f_\beta(L_M, E(S))}\right)}{E_M}$$

$$b = \frac{\Gamma\left(1 + \frac{1}{f_\beta(U_M, E(S))}\right)}{E_M}$$

$$c = \min\left\{f_\beta(E(M), L_S), f_\beta(E(M), U_S)\right\}$$

$$d = \max\left\{f_\beta(E(M), L_S), f_\beta(E(M), U_S)\right\}$$

For illustration, the derived priors at E = 7.9MV/cm are summarized in Table 5.

Table 5. Prior distributions at E=7.9MV/cm

Case	*A*	*B*
1	Normal (0.0179, 8.4899×10^{-5})	Normal (0.8472, 0.0538)
2	Normal (0.0179, 8.4899×10^{-5})	Gamma (13.3358, 15.7451)
3	Uniform (0.0093, 0.0322)	Uniform (0.5898, 1.7968)

7.2 DFR Estimations

Since the marginal densities of Weibull distribution are analytically intractable, Markov Chain Monte Carlo (MCMC) simulation is the easiest way to get reliable results without evaluting integrals [65]. MCMC methods are to sample from specified target distributions by creating a Markov chain that has the target distribution as its equibrium distribution [61]. The state of the chain after a large number of steps is then used as a sample from the target distribution. The quality of the sample improves as a function of the number of steps.

A special MCMC algorithm, the Gibbs sampler, is particularly well-adapted to high-dimensional Bayesian problems [65]. Gibbs sampler simulates *n* ramdom variables sequentially from the *n* univariate conditionals rather than generating a single *n*-dimensional vector in a single pass using the full joint distribution. Each iteration of the Gibbs sampler cycles through the unknown parameters, drawing a sample of one parameter conditional on the latest value of all the others [5]. The sample draws on one parameter can be regarded as simulated observations from its marginal distribution after a large number of iterations. Hence, the marginal density is reconstructed by averaging over the conditional density of the sample draws on this parameter. This marginal distribution is the posterior of the parameter and can easily be used to make inferences on the functions of the model parameters [5].

WinBUGS, a Windows version of the Bayesian inference using Gibbs sampling, is used in this study for posterior inference. In Gibbs sampling, simulated draws are sampled from full conditional distributions, but WinBUGS does not require explicit evaluation of the integration constant of a full conditional density function [5]. The

full conditional densities of the Weibull parameters are derived for different priors as follows for reference.

7.2.1 Full conditional densities

The sampling density of t_i from N_1 follows Weibull distribution W (α, β).

$$f(t_i|\alpha,\beta) = \beta\alpha^{\beta}t^{\beta-1}e^{-(\alpha t)^{\beta}}$$

Let $n_1 = j$-1. The joint density of the sample data $\mathbf{T}=\{t_1, t_2,\ldots,t_{n1}\}$ and the model parameters α and β can be written as

$$\begin{aligned} f(\mathbf{T},\alpha,\beta) &= g(\alpha)g(\beta)\prod_{i=1}^{n_1} f(t_i|\alpha,\beta) \\ &= g(\alpha)g(\beta)\prod_{i=1}^{n_1} \beta\alpha^{\beta}t_i^{\beta-1}e^{-(\alpha t_i)^{\beta}} \\ &= g(\alpha)g(\beta)\beta^{n_1}\alpha^{\beta N_1}\prod_{i=1}^{n_1} t_i^{\beta-1}e^{-(\alpha t_i)^{\beta}} \end{aligned}$$

The full conditional density of each parameter is derived as follows.

$$\begin{aligned} f(\alpha|\mathbf{T},\beta) &\propto g(\alpha)\prod_{i=1}^{n_1} f(t_i|\alpha,\beta) \\ &\propto g(\alpha)\alpha^{\beta n_1}\prod_{i=1}^{n_1} t_i^{\beta-1}e^{-(\alpha t_i)^{\beta}} \end{aligned}$$

$$\begin{aligned} f(\beta|\mathbf{T},\alpha) &\propto g(\beta)\prod_{i=1}^{n_1} f(t_i|\alpha,\beta) \\ &\propto g(\beta)\beta^{n_1}\alpha^{\beta n_1}\prod_{i=1}^{n_1} t_i^{\beta-1}e^{-(\alpha t_i)^{\beta}} \end{aligned}$$

where $g(\alpha)$ and $g(\beta)$ denote the prior distributions of α and β, respectively.

Substitute $g(\alpha)$ and $g(\beta)$ with the density functions in Table 5 to see the actual form of the full conditionals.

Case 1: $\alpha \sim$ Normal (a, b) and $\beta \sim$ Normal (c, d)

$$f(\alpha|\mathbf{T},\beta) \propto \alpha^{\beta n_1} e^{-\frac{(\alpha-a)^2}{2b}} \prod_{i=1}^{n_1} t_i^{\beta-1} e^{-(\alpha t_i)^\beta}$$

$$f(\beta|\mathbf{T},\alpha) \propto \beta^{n_1} \alpha^{\beta n_1} e^{-\frac{(\beta-c)^2}{2d}} \prod_{i=1}^{n_1} t_i^{\beta-1} e^{-(\alpha t_i)^\beta}$$

Case 2: $\alpha \sim$ Normal (a, b) and $\beta \sim$ Gamma (c, d).

$$f(\alpha|\mathbf{T},\beta) \propto \alpha^{\beta n_1} e^{-\frac{(\alpha-a)^2}{2b}} \prod_{i=1}^{n_1} t_i^{\beta-1} e^{-(\alpha t_i)^\beta}$$

$$f(\beta|\mathbf{T},\alpha) \propto \beta^{n_1+c-1} \alpha^{\beta n_1} e^{-d\beta} \prod_{i=1}^{n_1} t_i^{\beta-1} e^{-(\alpha t_i)^\beta}$$

Case 3: $\alpha \sim$ Uniform (a, b) and $\beta \sim$ Uniform (c, d).

$$f(\alpha|\mathbf{T},\beta) \propto \alpha^{\beta n_1} \prod_{i=1}^{n_1} t_i^{\beta-1} e^{-(\alpha t_i)^\beta}$$

$$f(\beta|\mathbf{T},\alpha) \propto \beta^{n_1} \alpha^{\beta n_1} \prod_{i=1}^{n_1} t_i^{\beta-1} e^{-(\alpha t_i)^\beta}$$

7.2.2 Posterior simulation and sensitive analysis

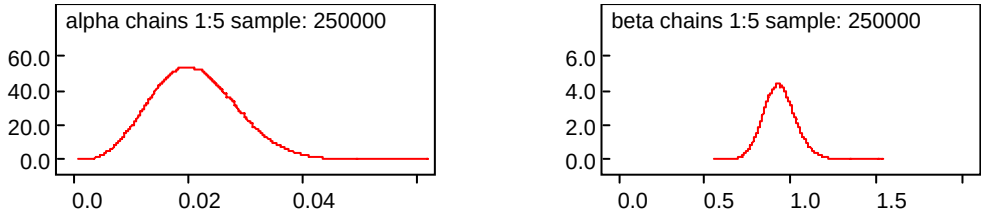
The Mtropolis algorithm in WinBUGS is based on a univariate normal proposal distribution at the current value. The standard deviation of the normal proposal is tuned over the first 4000 iterations in order to get an acceptance rate between 20% and 40%. All summary statistics for the model will ingore information from this adapting phase. The sampling method is slice sampling. This method can avoid the need to sample from non-standard distibutions [66]. It has an adaptive phase of 500 iterations which will be discard from all summary statistics.

As an example, the output statistics of the WinBUGS simulation are reported in Table 6, given the priors listed in Table 5. To prevent posterior dependence on the starting point of a simulation, five chains with over-dispersed starting points are used in one MCMC simulation. The Each chain continues for 100000 iterations, whereas the first 50000 iterations are discarded before any data analysis. Gelman-Rubin convergence statistic R is defined as the ratio of the width of the central 80% interval of the pooled chains to the average width of the 80% width of the individual chains [67]. The good convergence is confirmed by the fact that the statistic R is stable around 1. It indicates the variance between the different chains is no longer larger than the variance within each individual chain.

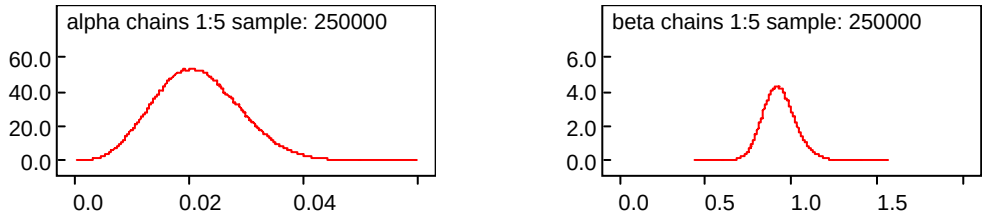
Irrespective of the prior case, the posterior of each model parameter has a peaked normal-like kernel density as shown in Figure 32. For both α and β, the difference in the posterior statistics is less than 15% among the three prior cases, except the posterior variance of β under the uniform prior is larger than the other two due to its large prior variances. However, this truncated uniform prior is still acceptable because

Table 6. Posterior statistic under the priors in Table 5.

Case	A			β		
	Mean	Median	Variance	Mean	Median	Variance
1	0.0213	0.0209	5.3626×10^{-5}	0.9372	0.9330	0.0090
2	0.0217	0.0213	5.4893×10^{-5}	0.9305	0.9255	0.0093
3	0.0186	0.0179	2.6925×10^{-5}	0.9233	0.8979	0.0395

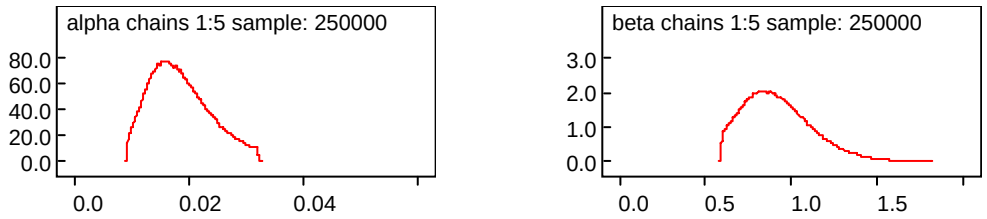


(a) Case 1 in Table. 7



(b) Case 2 in Table. 7

Figure 32. Posterior kernel densities of α and β under the derived priors



(c) Case 3 in Table. 7

Figure 32. Continued

of its peaked posterior likelihood for a very small sample (N =12). The posterior mean of β plus one standard deviation may greater than 1. It means the failure rate estimated from this approach might not be a strictly decreasing but an acceptable DFR-like estimation in this particular example. As shown in Table 5, the model parameters are generally insensitive to small changes in the priors, but the influence of priors is not negligible on the posterior. The sample data and the priors are both considered important in this case, which makes it a key point to develop reasonable prior.

Assuming that time-to-breakdown data in the useful life follows an exponential distribution, constant failure rate estimation using Bayesian approach has been demonstrated in Ref [5]. It can be also roughly estimated by the decreasing failure rate at the change point.

7.3 Summary

A 3-step Bayesian procedure is proposed to estimate the bathtub failure rate when the change point is unknown and the sample size is not large. Rather than individually considering each stress level in ALT, this work considers the relationship between the

lifetime and voltage stress level as determined by electrical and physical properties of electronic devices. Based on this inverse power law, the change point and the priors used to estimate DFR and CFR in Bayesian analysis can be derived from the time-to-failure data under neighborhood stresses. To estimate the change point, an exponentiality test including transformation is performed without estimating the unknown distribution parameters. The derived prior distributions enable this Bayesian procedure to provide a fast and reliable estimation from the sharply peaked posterior likelihood with a small sample size. In practice, this gives the manufacturer great advantages in terms of time and cost savings.

CHAPTER 8
CONCLUSIONS

8.1 Summary and Contributions

The topic of this dissertation is multidisciplinary, and the content is broad. It deals not only with the statistical aspects of reliability, but also with the electrical and physical aspects of reliability characterization. As an example of nano-devices, HfOx-based high-k dielectric thin films are studied in depth and analyzed with respect to charge trapping and degradation mechanisms, ch breakdown modes and bathtub failure rate estimation. The results of this work can be summarized in four parts.

1) Charge Trapping in connection with Dielectric Degradation

Charge trapping in high-k dielectrics has been one of the most important concerns, which determine the application of these metal oxides in CMOS devices. The probability of bias-induced charge trapping in high-k gate stacks is extremely high due to the large densities of as-grown defects, which may function as electron traps and fixed charges. Charge trapping/detrapping can affect evaluation of mobility in transistors, cause threshold voltage instabilities, and greatly influence the reliability of the devices over the operation time. Chapter 3 characterizes charge trapping and investigates degradation mechanisms in high-k dielectrics. Positive charges trapped in both bulk and interface contribute to the interface state generation and V_{fb} shift when electrons are injected from the gate under a negative gate bias condition; meanwhile no noticeable electron or neutral traps created by anode hole degrade much the performance of high-k dielectric films. A negligible number of defects are generated

until the gate bias stress increases to a certain level. The stress-induced trap generation is not reversible under high gate bias conditions and the number of these newly generated traps could be comparable to or even larger than that of the pre-existed ones. Compared with $ZrHfO_X$, dielectric film with structure $ZrHfO_X$ / RuO_X / $ZrHfO_X$ is considered more promising in terms of leakage current, relaxation current, defect generating conditions and defect creation rate.

2) Dielectric Breakdown Mechanisms

Breakdown mechanisms are fairly complex. General speaking, breakdown is a two-step process, with progressive dielectric degradation followed by final destruction of the film. Chapter 4 investigated the breakdown sequence of the two ultra-thin ZrHfO/1 nm SiON high-k stacks. When the stack is subjected to gate injection, a large amount of interface states are generated near the substrate due to hot electrons and positive charges trapped in the interface region. These defects lead to the early failure of the interface layer. However, for this kind of high-*k* stack, two types of breakdown phenomena were observed depending on whether the bulk and the interface layers failed simultaneously or separately. The difference in the breakdown sequence is attributed to the physical thickness of the bulk high-k layer and the structure of the interface layer. Compared to the 2.5 EOT sample under the same gate bias condition, the higher stress on the interface of the 1.8nm EOT sample accelerated the defect generation in the interface region, which leads to its early interface breakdown. The bulk high-*k* broke subsequently and separately at a higher voltage.

3) Charge Trapping Ruthenium Nanocrystal embedded high-k film

Charge trapping is also studied on Ruthenium nanocrystal embedded high-k film, a promising alternative to the conventional flash nonvolatile memories. Memory window in capacitance-voltage characteristics was mainly contributed by the negative flatband voltage shift due to hole trapping. Since the total charge trapping density is larger than the nanocrystal density, hole trapping might also occur in both nc-RuOx and in the interface between the nc-RuOx and ZrHfOx dielectric due to Coulomb blockade effect. The loosely trapped charges identified by relaxation current could be more likely located in the interface than in nc-RuOx, but n not all of charges trapped in the interface could be recognized as loosely trapped charge.

4) Bathtub Failure Rate Estimations

A 3-step Bayesian-based procedure is proposed to estimate the CFR and DFR, when the change point is unknown and the sample size is not large. Rather than individually considering each stress level in ALT, this research considers the relationship between the lifetime and voltage stress levels as determined by electrical and physical properties of electronic devices. Based on this inverse power law, the change point and the priors used to estimate DFR and CFR in Bayesian analysis can be derived from the time-to-failure under neighborhood stresses. To estimate the change point, an exponentiality test including transformation is performed without estimating the unknown distribution parameters. The derived prior distributions enable this Bayesian procedure to provide a fast and reliable estimation from the sharply peaked posterior likelihood with a small sample size. In practice, this gives the manufacturer great advantages in terms of time and cost savings. This paper used the real lifetime data on the next generation high-k-based capacitors as an

example, but the discussed methods can be applied to any voltage-stress accelerations.

8.2 Future Directions

The research covered in this dissertation may be extended in the following directions

1) Temperature is also a key stress factor that affects the reliability of dielectric thin films. Although temperature stress is not as crucial as the electric field, it is necessary to know how dielectric films behave under different temperature levels. In reality, the electric field and temperature are interactive. Hence, reliability should be examined and modeled when both of these stress factors are considered simultaneously.

2) Data mining classification techniques, such as linear classifier and support vector machine may be applicable to change point estimation. The SVM embedded Bayesian approach to failure rate estimation can be described in Figure 33.

Upon written request, the author would be glad to provide the LabView programs of stress tests, and the WinBUGS models for Markov chain Monto Carlo simulation used in the dissertation.

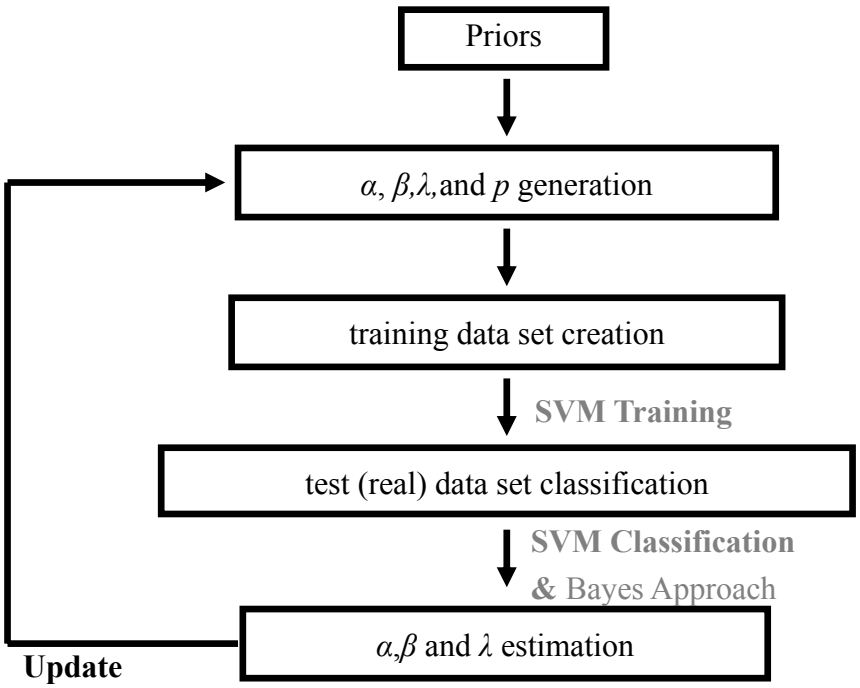


Figure 33. SVM Embedded Bayes Approach

REFERENCES

[1] National Research Council, *Implications of Emerging Micro and Nanotechnologies*, Washington, D. C.: the National Academics Press, 2002.

[2] S. Wolf, *Silicon Proessing for the VLSI Era: Deep-Submicron Process Technology*, vol.4, 1st ed., Latice Press, Sunset Beach, CA, 2002.

[3] R. Dennard, F. Gaenssle, V. Rideout, E. Bassous, and A. Leblanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol.9, no.5, p.256, 1974.

[4] J. Lu, *Hafnium-Doped Tantalum Oxide High-k Gate Dielectric Films for Future CMOS Technology*, Ph.D Dissertation, Texas A & M University, College Station, TX, 2004.

[5] W. Luo, *Reliability Characterization and Prediction of High-k Dielectric Thin Film*, Ph.D Dissertation, Texas A & M University, College Station, TX, 2004.

[6] http://download.intel.com/pressroom/images/events/moores_law_40th/Microprocessor_Chart.jpg (accessed on August 2008)

[7] H. C. Casey, Jr., *Devices for Integrated Circuits: Silicon and III-V Compound Semiconductors*, John Wiley & Sons, New York, 1999.

[8] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-k gate dielectrics: Current status and materials properties considerations," *Journal of Applied Physics*, vol.89, p. 5243, 2001.

[9] M. L. Green, E. P. Gusev, R. Degraeve, and E. L. Garfunked, "ultrathin (<4 nm) SiO_2 and Si–O–N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits," *Journal of Applied Physics*, vol.90, p. 2057, 2001.

[10] M. Houssa, *High-k Gate Dielectrics*, Institute of Physics Publishing, Bristol, UK, 2004

[11] T. Hori, *Gate Dielectrics and MOS ULSIs: Principles, Technologies, and Applications*, Springer, Berlin, Germany, 1997.

[12] *International Technology Roadmap for Semiconductors*, Semiconductor Industry Association, San Jose, CA, 2001.

[13] Intel website http://www.intel.com/technology/silicon/., (accessed on September, 2005).

[14] J. Yan, Y. Kuo, and J. Lu, "Zirconium-Doped Hafnium Oxide High-*k* Dielectrics with Subnanometer Equivalent Oxide Thickness by Reactive Sputtering," *Electrochemical and Solid-State Letter*, vol. 10, p. H199, 2007.

[15] D. H. Triyoso, R. I. Hegde, J. K. Schaeffer, D. Roan, P. J. Tobin, S. B. Samavedam and B. E. White, Jr., "Impact of Zr addition on properties of atomic layer deposited HfO_2," *Applied Physics Letter*, vol. 88, p. 222901, 2006.

[16] C. –H Lin, Y. Kuo and J. Lu, "Influence of Ru Dopant on the Dielectric Properties of Zr-doped HfO_2 High-k Thin Film," *Electrochem. Soc. Symp. Proc. Advanced Gate Stack, Source/ Drain, and Channel Engineering for Si-Based CMOS*, 2007.

[17] C.-H. Yang, Y. Kuo, C.-H Lin, R. Wan, and W. Kuo, "Relaxation behavior and breakdown mechanisms of nanocrystals embedded Zr-DOPED HfO_2 high-k thin films for nonvolatile memories," in *Proceedings of Materials Research Society*, 2008.

[18] W. Kuo, W. –T. Chien, and T. Kim, *Reliability, Yield, and Stress Burn-in*. Boston, Massachusetts: Kluwer Academic Publisher, 1998.

[19] G. Ghibaudo, G. Pananakakis, R. Kies, E. Vincent, and C. Papadas, "Accelerated dielectric breakdown and wear out standard testing methods and structures for reliability evaluation of thi oxides," *Microelectronics Reliability*, vol. 39, p. 597, 1999.

[20] W. Luo, Y. Kuo, and W. Kuo, "Dielectric Relaxation and Breakdown Detection of Doped Tantalum Oxide High-k Thin Films," *IEEE Transactions on Device and Materials Reliability*, vol. 4, no. 3, p. 488, 2004.

[21] W. Luo, T. Yuan, Y. Kuo, J. Lu, J. Yang and W. Kuo, "Breakdown phenomena of zirconium-doped hafnium oxide high-k stack with an inserted interface layer," *Applied Physics Letter*, vol. 89, p. 072901, 2006.

[22] J. H. Stathis, "Physical and predictive models of ultra thin oxide reliability in CMOS devices and circuits," in *Proceedings of the 39th Annual International Reliability Physics Symposium*, p. 132, 2001.

[23] E. Y. Wu, W. W. Abadeer, L. –K. Lai, S. –H. Lo, and G. R. Hueckel, "Challenges for accurate reliability projections in the ultra-thin oxide regime," in Proceedings of the 37th Annual International Reliability Physics Symposium, p. 57, 1999.

[24] G. S. Mudholkar and D. K. Srivastava, "Exponentiated Weibull family for analyzing bathtub failure-rate data," *IEEE Transactions on Reliability*, vol. 42, no. 2, pp. 299-302, 1993.

[25] Z. Chen, "A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function," Statistics & Probability Letters, vol. 49, pp. 155-161, 2000.

[26] M. Xing, Y. Tang and T. N. Goh, "A modified Weibull extension with bathtub-shaped failure rate function," *Reliability Engineering & System Safety*, vol. 76, pp. 279-285, 2002.

[27] C. D. Lai, M. Xie, and D. N. P. Murthy, "A modified Weibull distribution," *IEEE Transactions on Reliability*, vol. 52, no. 1, pp. 33-37, 2003.

[28] E. T. Wondmagegnehu, J. Navarro and P. J. Hernandez, "Bathtub shaped failure rates from mixtures: a practical point of view," *IEEE Transactions on*

Reliability, vol. 54, no. 2, pp. 270-275, 2005.

[29] B. Boukai, "Bayes sequential procedure for estimation and for determination of burn-in time in a hazard rate model with an unknown change-point," *Sequential Analysis*, vol. 6, no.1, pp. 37-53, 1987.

[30] K. Chou and K. Tang, "Burn-in time and estimation of change-point with Weibull-exponential mixture distribution," *Decision Sciences*, vol. 23, no.4, pp.973-986, 1992.

[31] S. Ramanathan, P. C. Mclntyre, S. Guha, and E. Gusev, "Charge trapping studies on ultrathin ZrO_2 and HfO_2 high-k dielectrics grown by room temperature ultraviolet ozone oxidation," *Applied Physics Letters*, vol. 84, p. 389, 2004.

[32] Y. Zhao, C. D. Young, M. Pendley, Kenneth Matthews, Byoung Hun Lee, and George A. Brown, "Effective Minimization of Charge Trapping in High-k Gate Dielectrics with an Ultra-short Pulse Technique," in *Proceedings of 2004 International Conference on Solid-State and Integrated Circuits Technology*, p. 407, 2004.

[33] G. Bersuker, B. H. Lee, H. R. Huff, J. Gavartin and A. Shluger, "Mechanism of Charge Trapping Reduction in Scaled High-k Gate Stacks," in *Defects in High-k Gate Dielectric Stacks*, vol. 220, p. 227, Springer Netherlands, 2006.

[34] Xing Zhou, *Charge Trapping Properties of Alternative High-k Dielectrics in MOS Devices*. Ph.D Dissertation, Vanderbilt University, 2006.

[35] A. K. Jonscher, "Dielectric Relaxation in Solids," Journal of Physics, D: Applied Physics, vol. 32, p. 57, 1999.

[36] W. Luo, T. Yuan, Y. Kuo, J. Lu, J. Yang and W. Kuo, "Charge trapping and dielectric relaxation in connection with breakdown of high-k gate dielectric stacks," *Applied Physics Letter*, vol. 88, p. 202904, 2006.

[37] D. M. Fleetwood, M. R. Shaneyfelt, W. L. Warren, J. R. Schwank, T. L. Meisenheimer and P. S. Winokur, "Border Traps: Issues for MOS Radiation Response and Long-term Reliability," *Microelectronics Reliability*, vol. 35, p. 403, 1995

[38] A. Kerber, E. Cartier, R. Degraeve, P. J. Roussel, L. Pantisano, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, "Charge Trapping and Dielectric Reliability of SiO/sub 2/-Al/sub 2/o/sub3/gate stacks with TiN electrons," *IEEE Transactions on Electron Devices*, vol. 50, p.1261, 2003.

[39] A. Kerber, E. Cartier, G. Groeseneken and H. E. Maes, "Stress Induced Charge Trapping Effects in SiO2/Al2O3 gate stacks with TiN electrodes," *Journal of Applied Physics*, vol. 94, p. 6627, 2003.

[40] Z. Xu, L. Pantisano, A. Kerber, R. Degraeve, E. Cartier, S. De Gendt, M. Heyns, and G. Groeseneken, "A study of relaxation current in high-/spl kappa/ dielectric stacks," *IEEE Transactions on Electron Device*, vol. 51, no. 3, p. 402, 2004.

[41] M. Houssa, V. Afanas'ev, A. Stesmans and M. M. Heyns, "Polarity dependence of defect generation in ultrathin SiO_2/ZrO_2 gate dielectric stacks, " *Applied Physics Letter*, vol. 79, p. 3134, 2001.

[42] S. Hall, O. Buiu and Y. Lu, "Direct Observation of Anomalous Positive Charge and Electron-Trapping Dynamics in High-k Films Using Pulsed-MOS-Capacitor Measurements," *IEEE Transactions on Electron Devices*, vol. 54, p. 272, 2007.

[43] D. J. DiMaria and J. W. Stasiak, "Trap creation in silicon dioxide produced by hot electrons," *Journal of Applied Physics*, vol.65, p. 2342, 1989.

[44] C. –H. Lin, Y. Kuo and J. Lu, "Influence of Ru Dopant on the Dielectric Properties of Zr-doped HfO_2 High-k Thin Film," in *Proceedings of Electrochemical Society Symposium on Advanced Gate Stack, Source/ Drain, and Channel*

Engineering for Si-Based CMOS, 2007.

[45] C. H. Lee, J. Meteer, V. Narayanan, and E. C. Kan, "Self-assembly of metal nanocrystals on ultrathin oxide for nonvolatile memory applications," *Journal of Electronic Material*, vol. 34, p.1, 2005.

[46] R. Wan, J. Yan, Y. Kuo and J. Lu, "Dielectric breakdown and charge trapping of ultrathin ZrHfO/SiON high-*k* gate stacks," *Japanese Journal of Applied Physics*, vol.47, p. 1639, 2008.

[47] J. S. Suehle, "Reliability Characterization of Ultra-thin Film Dielectrics," in *Characterization and Metrology for ULSI Technology: 1998 International Conference*, p. 115, 1998.

[48] H. Reisinger, G. Steinlesberger, S. Jakschik, M. Gutsche, T. Hecht, M. Leonhard, U. Schroder, H. Seidl, and D. Schumann, "A comparative study of dielectric relaxation losses in alternative dielectrics," *2001 IEDM Technical Digest*, p. 267, International Electron Device Meeting, 2001.

[49] P. D. Kirsch, C. S. Kang, J. Lozano, J. C. Lee and J. G. Ekerdt, "Electrical and spectroscopic comparison of HfO_2/Si interfaces on nitrided and un-nitrided Si(100)," *Journal of Applied Physics*, vol.91, p. 4353, 2002.

[50] M. Houssa, M. Tuominen, and M. Naili, "Trap-assisted tunneling in high permittivity gate dielectric stacks," *Journal of Applied Physics*, vol.87, p. 8615, 2000.

[51] J. R. Jameson, W. Harrison, P. B. Griffin, and J. D. Plummer, "Double-well model of dielectric relaxation current," *Applied Physics Letter*, vol. 84, p. 3489, 2004.

[52] Y. Kuo, J. Lu, S. Chatterjee, J. Yan, T. Yuan, H. C. Kim, J. Peterson, M. Gardner, S. Chatterjee, and W. Luo, "Sub 2 nm thick zirconium doped hafnium oxide high-k gate dielectrics," *ECS Transactions*, vol. 1, no. 5, p. 447, 2006.

[53] P. Cappelletti, C. Golla, P. Olivo, and E. Zanoni, Flash Memories, p.4, Kluwer Academic Publishers, Boston, 1999.

[54] S. Tiwari, F. Rona, K. Chan, L. Shi, and H. Hanafi, "A silicon nanocrystals based memory," *Applied Physics Letter*, vol. 38, p. 1377, 1996.

[55] J. De Blauwe, "Nanocrystal nonvolatile memory devices," *IEEE Transactions on Electron Device*, vol. 1, no. 1, p. 72, 2002.

[56] D. –W Kim, S. Hwang, T. Edgar, and S. Banerjee, "Characterization of SiGe quantum dots on SiO_2 and HfO_2 grown by rapid thermal chemical deposition for nanoelectronic devices," *Journal of the Electrochemical Society*, vol. 150, p. G240, 2003.

[57] Z. Liu, C. Lee, V. Narayanan, G.. Pei, and E. C. Kan, "Metal nanocrystal memories- part I. Device design and fabrication," vol. 49. p.1606, 2002.

[58] G. Wilk, R. Wallace, and J. Anthony, "Hafnium and zirconium silicates for advanced gate dielectrics," *Journal of Applied Physics*, vo87, p. 484, 2000.

[59] J. Lu, Y. Kuo, and J-Y Tewg, "Hafnium-doped tantalum oxide high-k gate dielectrics," *Journal of the Electrochemical Society*, vol. 153, p. G410, 2006.

[60] B. Garrido, S. Cheylan, O. Gonzalez-Varona, A. Perez-Rodriguez, and J. R. Morante, "The Effect of Additional Oxidation on the Memory Characteristics of Metal-oxide-semiconductor capacitors with Si Nanocrystals, " Applied Physics Letter, vol. 82, p. 4818, 2003.

[61] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, Bayesian Data Analysis, Boca Raton, Florida: Chapman & Hall/CRC, 2000

[62] H.G. -Pulido, V. A.-Torres and J. A. Christen, "A practical method for obtaining prior distributions in reliability," *IEEE Transactions on Reliability*, vol. 54, no. 2, p. 262, 2005.

[63] R. Degraeve, B. Kaczer and G. Groeseneken, “Degradation and breakdown in thin oxide layers: Mechanisms, models and reliability prediction,” *Microelectronics Reliability*, vol. 39, p. 1445, 1999.

[64] H. Jeffreys, *Theory of Probability*. Third edition, London: Oxford University Press, 1961.

[65] A. E. Gelfand and A. F. M. Smith, “Sampling-based approaches to calculating marginal densities,” *Journal of the American Statistical Association*, vol. 85, p. 398, 1990.

[66] R. M. Neal, “Slice sampling, ” *Technical Reprot*, no. 2005, Department of Statistics, University of Toronto, Canada, 2000.

[67] D. Spiegelhalter, A. Thomas, N. Best, and D. Lunn, WinBUGS User Manual, Version 1.4, http://www.mrc-bsu.cam.ac.uk/bugs, 2003.

[68] C. –H. Lin and Yue Kuo, “Embedding of Nanocrystalline Ruthenium Oxide in Zirconium-doped Hafnium Oxide High-k Film for Nanvolatile Memories,” ECS Transactions, vol. 13, p.465, 2008.

[69] J. Lu, C. –H. Lin and Yue Kuo, “Nonvolatile Memories Based on Nanocrystalline Zinc Oxide Embedded Zirconium-doped Hafnium Oxide Thin Film,” ECS Transactions, vol. 11, p.509, 2007.

[70] D. B. Farmer and R. G. Gordon, “High density Ru nanocrystal deposition for nonvolatile memory applications,” *Journal of Applied Physics*, vol. 101, p. 124503, 2007.

VITA

Rui Wan received her B. E. and M.S. degrees in electrical engineering from Tianjin University, Tianjin, China. She earned her Ph.D degree from the Department of Industrial Engineering, University of Tennessee, Knoxville, in 2008. Her research interests are in reliability testing and characterization, as well as statistical modeling and prediction of nano electronics.