Nanoscale characterization of advanced high- κ gate dielectric stacks via scanning tunneling microscopy

Yew, Kwang Sing

2014

Yew, K. S. (2014). Nanoscale characterization of advanced high- κ gate dielectric stacks via scanning tunneling microscopy. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61038>

<https://doi.org/10.32657/10356/61038>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**NANOSCALE CHARACTERIZATION OF ADVANCED
HIGH- κ GATE DIELECTRIC STACKS VIA SCANNING
TUNNELING MICROSCOPY**

YEW KWANG SING

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2014

**NANOSCALE CHARACTERIZATION OF ADVANCED
HIGH- κ GATE DIELECTRIC STACKS VIA SCANNING
TUNNELING MICROSCOPY**

Yew Kwang Sing

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2014

ACKNOWLEDGEMENT

My sincere gratitude to my supervisor, *Prof. Ang Diing Shenp* from Nanyang Technological University (NTU), for his sound technical guidance, inspiring ideas, constructive discussions and constant encouragement throughout the course of my Ph.D research works. His immense knowledge in device physics, sharp analytical skills, effective writing skills, high research integrity and great passion in research have sparked my growing enthusiasm in research field. He is a kind and committed person, who is willing to spend time and patiently works things out with us whenever we faced difficulties in our works. *Prof. D. S. Ang* is more than a research supervisor to me; he has always provided me with invaluable advices, be it in research, career or life. I am very grateful and lucky to have the opportunity to work under *Prof. D. S. Ang*'s supervision.

I would like to express my sincere gratitude to *Prof. Pey Kin Leong* (during his tenure with NTU) for granting me access to the ultra-high vacuum scanning tunneling microscopy (UHV-STM) system and *Dr. Ong Yi Ching* (currently with TSMC Taiwan) for patiently training me in handling the UHV-STM system. She has been very kind to share her technical experience in the high- κ metal oxide characterization using STM. My appreciation to *Dr. Lakshmi Kanta Bera* for his technical advices on device fabrication process. This project is supported in part by the Singapore Ministry of Education under a research grant MOE2009-T2-1-050 and in part by the NTU.

My thanks to our collaborators (i) *Dr. Gennadi Bersuker* and his team members (*Dr. Pat Lysaght* and *Dr. Heh Dawei*) from SEMATECH (Austin, USA) for providing us samples and useful technical discussions, (ii) *Asst. Prof. Yu Hong Yu* and his

student, *Mr. Wu Ling*, for sample fabrication support, (iii) *Mr. Tang Lei Jun* from Institute of Microelectronics (IME), A*STAR, for electron energy loss spectroscopy (EELS) and transmission electron microscopy (TEM) support, and (iv) *Dr. Pan Jisheng* from Institute of Materials Research and Engineering (IMRE), A*STAR, for x-ray photoelectron spectroscopy (XPS) support. The collaborations have been very fruitful, making our joint publications possible.

I would like to acknowledge (i) NTU clean room staff for providing me equipments trainings and technical support throughout my sample fabrication work, (ii) *Mr. Chow Kam Wah* from Semiconductor Characterization Lab 2 for his technical support on the prober station, (iii) *Mr. Kenneth Kollin* and *Mr. Wang Zhouhang* from RHK Technology Inc. for their technical advices and support on the UHV-STM system, and (iv) NTU undergraduate students (*Mr. Tan Cher Hong* and *Mr. Cui Keer*) for their efforts in part of the electrical measurement works.

Many thanks to my other group members (*Mr. Gao Yuan*, *Miss. Boo Ann Ann*, *Mr. Gu Chenjie*, *Miss. Duan Tianli*, *Mr. Tham Wai Hoe*, *Mr. Zhang Haizhong*, *Mr. Zhou Yu*, and *Mr. Tung Zhi Yan*) and my other close friends (*Mr. Liu Pan*, *Mr. Cao Jin*, and *Mr. Chen Shuo*) for their constant moral support, technical sharing and friendships. My sincere appreciation to all the parties involved for the support throughout my exciting Ph.D journey.

I dedicate this dissertation to my parents (*Mr. Yew Kim Hok* and *Ms. Lock He Chun*), brothers (*Mr. Yew Kwang Hooi*, *Mr. Yew Kwang Meng*, *Mr. Yew Kwang Liang*, and *Mr. Yew Kwang Yong*) and sisters (*Ms. Lee Yee Ling*, *Ms. Ni Jun* and *Ms. Ho Min Kiat*). Their endless love has been a pillar of support to keep me moving forward in my life.

TABLE OF CONTENTS

ACKNOWLEDGEMENT.....	I
TABLE OF CONTENTS	III
ABSTRACT.....	VII
LIST OF FIGURES	XI
LIST OF TABLES	XXVII
LIST OF ABBREVIATIONS	XXVIII
LIST OF SYMBOLS	XXX
<i>CHAPTER 1 INTRODUCTION.....</i>	1
1.1 Background.....	1
1.1.1 Transistor's scaling towards nanometer regime	1
1.1.2 High- κ metal oxide as replacement gate dielectric.....	4
1.2 Motivation	6
1.3 Objectives	8
1.4 Organization of thesis.....	9
1.5 Original contributions.....	10
<i>CHAPTER 2 LITERATURE REVIEW.....</i>	13
2.1 Introduction	13
2.2 High- κ processing and its structural evolution.....	14
2.2.1 Atomic layer deposition.....	14
2.2.2 Crystallization in high- κ film.....	16
2.2.3 Interface issues.....	18
2.3 Electrical characteristics of metal/high- κ gate stack	20
2.3.1 Time dependent dielectric breakdown	20
2.3.2 Stress-induced leakage current	26
2.3.3 Threshold voltage instability.....	31

2.4	Physical failure mechanism	33
2.5	Post deposition annealing	35
2.5.1	Oxygen annealing	36
2.5.2	UV ozone annealing.....	37
2.5.3	Multi-deposition and annealing	39
2.6	Characterization by scanning tunneling microscopy	41
2.6.1	STM/CAFM study on ultrathin high- κ dielectric	45
2.7	Summary.....	48
CHAPTER 3 EXPERIMENTS.....		50
3.1	Introduction	50
3.2	Sample fabrication and preparation	51
3.2.1	High- κ deposited blanket sample.....	51
3.2.2	MOS capacitor fabrication.....	54
3.3	Scanning Tunneling Microscopy	55
3.3.1	Instrumentation	56
3.3.1.1	Mechanical structure	57
3.3.1.2	Control electronics	58
3.3.1.3	Scanning unit.....	59
3.3.1.4	Beetle-type scan head.....	61
3.3.2	Tip preparation.....	62
3.3.3	Characterization techniques	63
3.3.3.1	Constant current imaging (CCI).....	63
3.3.3.2	Conductance measurement – CITS and RVS	65
3.3.3.3	Localized constant voltage stressing.....	67
3.4	Prober system	68
3.5	Summary.....	70
CHAPTER 4 THE ROLE OF HIGH-K GRAIN BOUNDARY IN GATE STACK BREAKDOWN.....		71
4.1	Introduction	71
4.2	The impact of high- κ film quality on IL degradation of a bi-layer gate stack .	72
4.2.1	Experiments	73

4.2.2	IL degradation at pre-existing high- κ leakage path	74
4.2.3	IL degradation without pre-existing high- κ leakage path	78
4.2.4	Local electronic properties at high- κ grain and grain boundary	80
4.3	Two-step deposited and annealed HfSiO _x gate stack	86
4.3.1	Experiments	87
4.3.2	Electronic properties at grain and grain boundary	90
4.3.3	Misaligned GBs in the HfSiO _x film	101
4.3.4	Electrical characteristics of metal/HfSiO _x gate stack	103
4.4	Summary	106
 CHAPTER 5 RELIABILITY OF THE MULTISTEP DEPOSITED AND LOW TEMPERATURE ANNEALED HIGH-K GATE STACK.....		108
5.1	Introduction	108
5.2	Multistep deposited and UV ozone annealed HfZrO gate stack	110
5.2.1	Experiments	111
5.2.2	Topography and localized gate leakage density	113
5.2.3	Electrical uniformity and susceptibility against degradation.....	117
5.2.4	Traps profile by pulse <i>C-V</i> measurement.....	120
5.3	Enhanced high- κ reliability via two-steps (UV ozone cum RTA) anneal	125
5.3.1	Experiments	126
5.3.2	Electrical characterization of the HKMG gate stack	128
5.3.3	TDDDB and stress-induced degradation of the HKMG gate stack.....	131
5.3.4	Gate leakage current density and electrical uniformity by STM	132
5.3.5	Chemical analysis by EELS and XPS analysis.....	135
5.3.5	Reliability of ultrathin sub-1 nm EOT HKMG gate stack.....	141
5.4	Summary	147
 CHAPTER 6 BIMODAL WEIBULL DISTRIBUTION OF HIGH-K/METAL GATE STACK – INSIGHTS BY STM.....		149
6.1	Introduction	149
6.2	Experiments	150
6.3	Area-scaling Weibull distribution of high- κ /metal gate stack.....	151
6.4	TDDDB characteristic of grain and grain boundary	152

6.5	Bimodal Weibull distribution	157
6.5.1	Correlation of the GB-related defect density and gate area size.....	157
6.5.2	The role of GB in bimodal Weibull distribution.....	159
6.6	Summary.....	162
<i>CHAPTER 7 CONCLUSION & RECOMMENDATION.....</i>		164
7.1	Conclusion.....	164
7.2	Recommendations	166
LIST OF PUBLICATIONS		169
BIBLIOGRAPHIES		171

ABSTRACT

Gate leakage current is the major driving force for the integration of high- κ metal oxide into the silicon-based CMOS technology at sub-45 nm node. By virtue of the higher dielectric constant value, high- κ metal oxide provides a physically thicker insulator in suppressing the gate leakage current while achieving the gate capacitance identical to that of an ultrathin SiO₂ insulator, therefore enabling the further scaling of equivalent oxide thickness (EOT). However, the performance and reliability of high- κ metal oxide are severely compromised by an inherently high density of as-grown defects and the material phase change induced by the high-thermal-budget front-end processes. The reliability studies of the high- κ degradation mechanism give rise to inconsistent interpretations due to unresolved microscopic information, such as (i) the electronic properties of high- κ grain and grain boundary (GB), (ii) the role of GB-related defects in the interfacial layer (IL) breakdown (BD), (iii) the effect of GB-related defects on the gate stack time-dependent dielectric breakdown (TDDB) distribution, and (iv) the correlation of oxygen-vacancy defects (V_o 's) and gate leakage current. We address these issues by analyzing the microscopic data of the local defects extracted via spatially-resolved ultra-high vacuum scanning tunneling microscopy (UHV-STM).

The reliability study of the polycrystalline HfO₂/SiO_x gate stack at nanoscale level via STM shows directly that the degradation and BD of the SiO_x IL depends ultimately on the state of the thicker high- κ . A high degree of non-uniformity in the trap density distribution is observed for a given high- κ area, with the leakage sites tend to concentrate near GB. Under repeated stress, a significant degradation is observed at

the IL (as evidenced by a substantial increase of tunneling spectra under substrate injection bias) beneath a high- κ GB compared to its counterpart beneath a grain, making this location a preferential BD spot of the entire gate stack. These results provide useful insights into (i) GB-assisted BD in the MOS device, and (ii) the correlation of the IL-controlled BD and the high- κ film quality.

We demonstrated that the *multistep deposition cum rapid thermal annealing (RTA)* method could significantly reduce the effect of GB-related defects (resulting from high temperature 1000 °C anneal) on the high- κ film reliability. Localized electrical stress via STM probe reveals (i) a significantly narrower spread in the leakage current values, (ii) strong retardation of the gate stack BD process, even by a localized stress over the GB regions compared to the conventional single-step deposition process. The results are consistent with the MOS capacitor measurements. The improvements are ascribed to the formation of nanocrystalline structure (rather than large grains) in the thin films constituting the multistep dielectric, therefore resulting in misalignment in the GB-related low-resistance paths propagating through the dielectric.

From the STM measurement, we reveal that the substantial improvement seen in the multistep-deposited cum UVO-annealed TiN/HfZrO/SiO_x gate stack, relative to the rapid thermal-annealed (RTA) sample, is predominantly contributed by the suppressed crystallization of the high- κ layer (therefore a significant reduction in the density of GB-related defects as evidenced in the current map) rather than the UVO annealing effect. Compared with the as-dep sample, a marginal improvement is attributed to the reduction of V_o 's resulting from multistep UVO annealing. We show that by (i) extending the UVO annealing time after each deposition step, or (ii) increasing the number of deposition steps (i.e. by thinning down the high- κ sub-layer in each deposition step), a further improvement can be achieved.

A new method to enhance the reliability of gate-last ultrathin ~ 0.75 nm EOT HfO_2/TiN gate stack is demonstrated. A ~ 25 -Å HfO_2 film, deposited directly onto the HF-last Si surface via *multistep deposition cum two-step annealing (comprising a room-temperature UVO anneal and a subsequent 420 °C RTA)*, is shown to exhibit (i) approximately two orders of magnitude reduction in gate current density, (ii) approximately an order of magnitude longer time-to-breakdown (T_{BD}), and (iii) lower positive oxide trapped charge generation, as compared with the as-deposited, UVO-only and RTA-only gate stacks. EELS analysis indicates an increase in the amount of local oxygen atoms in the high- κ film. This inference is also borne out in XPS core level O 1s spectra, which show an increase in the Hf-O bonds coming from HfO_2 . This suggests an efficient incorporation of oxygen atoms into the high- κ bulk by the *two-step* annealing method.

We explain, via nanoscale TDDDB testing, the origin of the bimodal Weibull failure distribution obtained from the area-scaling of the high- κ /metal gate device BD statistics. The T_{BD} statistics at GB and grain in a polycrystalline high- κ stack are obtained individually from localized constant voltage stressing via STM probe. We observe comparable Weibull slopes for the T_{BD} distribution collected at GB and grain irrespective of initial state of pre-existing trap density, which can be explained with the percolation BD model. We show that the bimodal Weibull distribution is a combined effect: (i) the steep Weibull slope of the lower percentile, arising from large-area devices, is related to BD at GBs, and (ii) the shallow Weibull slope of the upper percentile, arising from small-area devices, is mostly related to grain BDs.

The novelty of this study lies in the application of STM in distinguishing GB and grain features in a polycrystalline high- κ film, thereby allowing us to probe the local electronic properties and BD strength at both locations individually. We develop the

multistep deposition cum two-step annealing method to improve the reliability of sub-1 nm EOT high- κ /metal gate for high- κ last integration scheme.

LIST OF FIGURES

Figure		Page
1.1	The number of transistors, using SiO ₂ /SiON gate dielectric, integrated in the different generation of Intel's microprocessors grows exponentially for the last three decades, as projected by Moore's law [1, 3].	2
1.2	Simulated I_g-V_g characteristic under inversion conditions for the various SiO ₂ thicknesses. The gate oxide can be scaled down to 20 Å before exceeding the 1 Acm ⁻² (at 1 V) limit for the high performance CMOS logic [6].	3
1.3	Gate current density and standby power consumption plotted as a function of gate voltage for a chip with a 15 Å SiO ₂ gate dielectric and the high-κ gate dielectric with the same EOT. The high-κ dielectric clearly suppresses the gate current density and thereby achieving low standby power consumption. Assumes at total gate area of 0.1 cm ² [10, 17].	4
2.1	Illustration of atomic layer deposition of Al ₂ O ₃ using Trimethylaluminum (TMA) precursor as chemical reactant and H ₂ O as oxidant. (a) TMA precursor is pulsed into the chamber and reacts with the OH groups on silicon surface, (b) a monolayer is formed and passivates the silicon surface, (d) unreacted TMA molecules and by-products are removed by N ₂ purge, (e,f) H ₂ O is pulsed into the chamber and remove CH ₃ groups, passivating surface with Al-OH, (g) unreacted H ₂ O molecules and by-products are removed by N ₂ purge. This cycle is repeated until the target thickness is achieved [30].	16
2.2	Cross-sectional transmission electron microscopy (TEM) of Si(100)/SiO ₂ /HfSiO gate stacks. The 20 nm HfSiO (30% SiO ₂) films were prepared by using metal-organic chemical vapor deposition (MOCVD) technique. (a) as-deposited film; the high-κ layer is in amorphous structure, (b) film grown at 650 °C and followed by 1000 °C annealing in N ₂ for 10 s, the high-κ layer has clearly crystallized [24].	17
2.3	Atomic force microscopy (AFM) images of 5 nm HfO ₂ films deposited at 200 °C on Si substrate. (a) as-deposited film; the surface roughness is 0.17 nm (rms), (b) film annealed at 900 °C in N ₂ ; the surface roughness is 0.27 nm (rms), (c) XRD scan of 5 nm HfO ₂	18

- films annealed at various temperatures. Monoclinic peaks are clearly observed after subjected to 800 °C and 900 °C anneal, indicating the presence of crystals in high- κ film [32].
- 2.4 Cross-sectional transmission electron microscopy (TEM) of atomic-layer-deposited 10 nm HfO₂ film on Si substrate. Prior to the HfO₂ deposition, the native oxide on Si surface was removed using diluted HF, subsequently a ~1 nm SiO₂ IL was grown. (a) as-deposited film; and (b) film subjected to 900 °C anneal in 1 bar O₂ for 10s. The IL thickness increases drastically and an obvious grain orientation is observed in the HfO₂ film, indicating crystallization [33]. 19
- 2.5 X-ray photoelectron spectroscopy (XPS) study of HfO₂/SiO₂ (~5 nm)/Si substrate annealed at 600 °C in N₂ for 30s. The thicknesses of atomic-layer-deposited HfO₂ in sample S1, S2, and S3 are 0 nm, 0.5 nm, and 4 nm respectively. The presence of Hf-O-Si peaks in sample S2 and S3 in (a) O 1s, and (b) Si 2p core-level spectra indicates the formation of hafnium silicate at the HK-IL interface due to an interaction between the HfO₂ and SiO₂ IL [35]. 20
- 2.6 (a) Measured Weibull slopes versus the physical thickness for SiO₂, Al₂O₃, and HfO₂ samples. At a given thickness, both Al₂O₃ and HfO₂ samples show a significantly lower Weibull slopes compared to SiO₂ sample [40]. (b) Weibull slope of Al₂O₃/SiO₂ gate stack failure distribution as a function of Al₂O₃ thickness. The Weibull slope shows a weak dependence on the high- κ thickness and annealing condition for the gate injection stress. For substrate injection stress, however, the value increases with the high- κ thickness but the effect pales after subjected to 700 °C/800 °C/900 °C annealing in N₂ [41]. 22
- 2.7 (a) Failure distribution of the Hf-based high- κ /metal gate devices based on vertical Poisson area scaling to a reference area of 1 μm^2 . The TDDB tests were carried out in inversion mode at 125 °C. The devices area ranges from 0.006 μm^2 to 50 μm^2 with a sample size of 150 devices per distribution. Two Weibull slopes are observed in the distribution. The upper percentile ($\beta \sim 0.8$) primarily consists of the T_{BD} data acquired from the small area devices whereas the lower percentile ($\beta \sim 1.5$) consists of the T_{BD} data acquired from the larger area devices [28]. (b) Area-scaled TDDB distribution of Al₂O₃(~9 nm)/metal gate devices subjected to constant voltage stressing at -6.8 V. Only a single Weibull slope is observed for the relatively large area devices [42]. 23
- 2.8 (a) Kinetic Monte Carlo simulated TDDB distributions of a high- κ /IL gate stack with various high- κ defect generation rates. As the defect generation rate is uniform for both high- κ and IL, a standard Weibull distribution with a single Weibull slope $\beta \sim 1.1$ is obtained. However, two Weibull slopes are observed in the TDDB distribution 25

- when the high- κ defect generation rate is increased to 3000 times compared with that in the IL [47]. (b) Simulation of Weibull plots based on a modified analytical cell-based percolation model, which shows the separate effect of grain boundary (GB), grain (G), and the interaction between G and GB (G-GB) in a $N \times N$ cells. The failure at pure G sites shows a straight line, while the GB sites shows a convex trend. The G-GB, on the other hand, shows a bimodal Weibull trend with the lower percentile follows the GB failure whereas the higher percentile follows the G failure [48].
- 2.9 (a) Defect density in HfO_2 “bulk” as a function of injected electron fluence at positive and negative stress polarity. The polarity-dependence of trap generation suggests that SILC conduction mechanism is related to the HfO_2 traps. (b) At the low trap density region, SILC shows a proportional one-to-one relationship with the HfO_2 “bulk” trap density. However, it increases rapidly at the high trap density region [50]. 26
- 2.10 (a) Quasi-static capacitance-voltage characteristic of $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stack as a function of gate voltage stress for 1000 s under gate injection mode. An obvious build up of interface states is observed, (b) under substrate injection mode, no interface states generation is observed. An obvious flatband voltage, V_{FB} , shift indicates electron trapping in Al_2O_3 layer [45]. 28
- 2.11 Schematic band diagram of $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stack during (a) gate injection stress. Holes mainly trapped near the SiO_2/Si interface and IL is believed to be the layer controlling the gate stack BD, (b) substrate injection stress. Traps creation occurs in the Al_2O_3 layer. Therefore, high- κ controls the gate stack reliability [45]. 28
- 2.12 Correlation of the low voltage SILC (measured at $V_g = 0.5$ V) and the charge pumping-measured (at low frequency of 1 kHz) N_{it} subjected to various gate stress voltages for SiO_2 (1.1 nm)/ HfO_2 (3 nm)/TiN nFET. A good correlation of the data indicates that the traps within IL control the SILC [44]. 29
- 2.13 A comparison of absolute gate leakage current (at accumulation bias, $V_g = -1$ V) versus EOT for the HfO_2 and SiO_2 dielectrics. SiO_2 leakage data is adapted from [55]. The thin HfO_2 dielectric exhibits a steeper slope in the curve compared to that of the SiO_2 , indicating a rapid increase in the gate leakage current and may even exceed SiO_2 at a low EOT (<1 nm). The inset shows an absolute leakage current evolution under constant voltage stressing [52]. 30
- 2.14 Evolution of threshold voltage shift $|\Delta V_t|$ of the (a) SiON [59] and (b) HfO_2 [60] p-MOSFETs under repeated stress/relax cycling (each stress/relax cycling is 1×10^3 s) under oxide fields (~ 7 MV/cm) at 32

- 100 °C. Compared with SiON, HfO₂ sample exhibits an obvious increase in the $|\Delta V_t|^{\text{eos}}$ and $|\Delta V_t|^{\text{eor}}$. The inset in (a) shows the definition of $|\Delta V_t|^{\text{eos}}$ and $|\Delta V_t|^{\text{eor}}$ at the end of the stress and recovery cycle, respectively.
- 2.15 Trapezoidal V_g pulse measurement ($t_r = t_f = 1 \mu\text{s}$; $\Delta t = 2 \mu\text{s}$) of P- and N-channel devices employing $\sim 14 \text{ \AA}$ HfO₂ gate dielectric. A pronounced transient I_d - V_g hysteresis is observed for both HfO₂ NMOS and PMOS as compared with the nitrided SiO₂ reference device at a comparable EOT [61]. 32
- 2.16 (a) Schematic illustration of trap (represented by the sphere) generation in the SiO₂ layer under applied electric field. As the number of generated traps reaches a critical value and forms a bridge connecting the gate and Si substrate, electrons/holes carriers may tunnel through the layer and induce intrinsic oxide BD. The BD path is depicted by the shaded spheres [39]. (b) The corrected oxygen deficiency in the percolation path based on the as-measured O K edge intensities near the BD spot. The local O deficiency in the center of the percolation path is relatively high. The diameter of the percolation (defective oxide area laterally) is estimated to be 30 nm for the post-BD leakage current of 2 μA [62]. 34
- 2.17 Schematic illustration of the proposed IL BD model for a bi-layer high- κ /IL gate stack. (a) GB facilitates the formation of oxygen-deficient filament in the IL region adjacent to the overlaying high- κ film. Shaded area indicates a high density of pre-existing V_o 's in IL. (b) Energy band diagram of the shaded area in IL, the oxygen deficient filament propagates toward the Si substrate accompanied by generation of additional V_o states (dotted lines). The waving line indicates trap relaxation due to an electron capture [65]. 34
- 2.18 (a) Equivalent oxide thickness (EOT) and gate current density, J_g , as a function of RTA time for annealing in O₂ or N₂ ambient at 700 °C. A 4.5 nm HfO₂ film was deposited on p -Si (100) substrate to form MOS capacitors. For the sample annealed in O₂ ambient, the J_g is reduced at the penalty of EOT increase [70] (b) Time-of-flight–secondary-ion-mass spectroscopy (TOF–SIMS) analysis of ¹⁶O₂ and ¹⁸O₂ signals for the as-deposited and 10 s annealed (at 900 °C in 0.1 bar ¹⁸O₂ atmosphere) HfO₂ films. A large fraction of ¹⁶O₂ is exchanged with ¹⁸O₂, suggesting that the oxygen diffusivity in HfO₂ is extremely high [33]. 37
- 2.19 (a) Gate current density versus gate voltage characteristics of Al/HfO₂ ($\sim 13 \text{ nm}$)/Si MIS structures subjected to O₂ treatment (O₂ ambient at 150 °C under 1000 psi for 2 hours) and UVO treatment (exactly similar condition as O₂ treatment, but under UV light illumination). The standard (i.e. as-deposited) sample receives no 39

	treatment. (b) 1 MHz high frequency $C-V$ characteristics. The O_2 and UVO treated samples exhibit smaller $C-V$ hysteresis (~ 0.7 V and ~ 0.2 V respectively) as compared with the standard sample (~ 1 V). In addition, a positive V_{FB} shift to indicates a lesser positive charge trapping in the UVO-treated sample [75].	
2.20	(a) Carbon concentration of $HfAlO_x$ films grown by LL-D&A technique in various ambiances. A significant reduction in the carbon content is observed for LL-D&A(NH_3) and LL-D&A(O_2) samples. (b) XPS core level Si 2p spectra of $HfAlO_x$ films grown by LL-D&A technique. LL-D&A(O_2) and LL-D&A(vac) samples exhibit an increase in the IL thickness, which is attributed to the diffusion of oxygen [78].	40
2.21	In quantum mechanics, an electron has a non-zero probability of tunneling through a potential barrier V even if the electron energy E is lesser (denoted by region II). The energy of the tunneled particle is identical except for a decrease in the amplitude.	41
2.22	One-dimensional metal-vacuum-metal tunneling junction model.	43
2.23	(a) 350×350 nm ² STM image of a thin SiO_2 dielectric. The area stressed with a high bias voltage of +10 V and 6 nA is indicated by the dotted line. (b) and (c) are the zoom-in high-resolution images of the corresponding area before and after degradation, respectively. The surface corrugation in (c) was clearly increased after modified by the hot-electron injection [85].	45
2.24	(a) STM measurement of Sc_2O_3 / La_2O_3 / SiO_x /Si gate stack at setpoint of -4 V, 20 pA. (I) depicts the topographical image of the gate stack whereas (II), (III), and (IV) depict the corresponding current maps acquired at -2.3 V, $+2.3$ V, and $+4$ V respectively. (b) STM measurement of the same area at setpoint of $+3.5$ V, 20 pA [104].	46
2.25	The $I-V$ characteristics extracted at various leakage sites (denoted as A, B, C and D in the current maps) [104].	46
2.26	STM topographical images of bi-layer La_2O_3 (4 nm)/ SiO_x (1 nm) gate stack scanned at different bias polarity. Bias voltage, V_s , is applied to the Si substrate. (a) at $V_s = +3.5$ V, the polycrystalline La_2O_3 layer is imaged. (b) at $V_s = -4$ V, the amorphous SiO_x IL is imaged [99].	47
3.1	Schematic view of STM setup. The tip-sample gap spacing is less than 10 Å. As a bias voltage is applied to the substrate of the sample, tunneling current in $\sim$ pA range is detected at the tunneling junction. Owing to an exponential relationship between the	56

	tunneling current and tip-sample gap distance, STM exhibits an excellent sensitivity to Z height.	
3.2	Mechanical structure (in plain view) of UHV STM system (UHV 3500 series) supplied by RHK Technologies Inc. The system consists of three main chambers: the fast entry lock-lock (FEL) chamber, preparation chamber and STM chamber. Three external air legs are mounted to the system to isolate the sample stage in the STM chamber from the surrounding vibrations and noises	57
3.3	Schematic block diagram of UHV STM feedback control electronics. At a given bias, the I_t sensed at the tunneling junction is compared with the preset value in order to generate an error z-signal, which will be processed by the feedback system for instantaneous rectification of the tip-sample vacuum gap spacing during imaging.	59
3.4	(a) Schematic diagram of piezoelectric hollow tube consisting of four longitudinal segments on the outer surface and the attached electrodes which control the three dimensional perpendicular motions in x , y , and z directions. (b) Application of opposite polarity voltages to a pair of diagonally opposite electrodes attached to the outer surface of the tube will produce lateral motion for the tip. (c) Extension and contraction Δl of the tube due to a transverse piezoelectric effect, which produces vertical motion for the tip..	60
3.5	STM scan head based on “beetle” design is mechanically and thermally stable. (a) top and (b) bottom views of the scan head. (c) The motion of the scan head outer legs ramping down the slope of a sample holder when approaching the tip to the sample surface	62
3.6	Apparatus used for the “loop” etching technique. High-purity tungsten wire is placed through the potassium hydroxide (KOH) thin film formed in the platinum ring. Both tungsten wire and platinum ring act as the potential electrodes. Under applied potential between the two electrodes, the place where tungsten wire in contact with KOH thin film would produce a sharp tip due to electrolysis reaction.	63
3.7	In constant current imaging (CCI) mode, the tip height is varied by the z-piezoelectric tube based on the output of the closed-loop feedback system in order to keep the I_t at constant when the tip is scanned across a sample surface. The tip retracts at physical protrusion (i.e. surface roughness; denoted at location A and B) and leakage site (i.e. electronic traps; denoted at location C) over a sample surface.	64
3.8	Spectra view of the CITS current map. The blue dots (as clearly seen in the zoom-in view) denote a spaced grid of spectroscopy pixels defined over a scan area. During CITS mode, the tip will probe at each spectroscopy pixel to acquire the local tunneling spectrum.	66
3.9	Schematic diagram representing the experimental set-up and the bias	66

	voltage behavior during ramp voltage stressing. Upon achieving the bias setpoint, the tip-sample height is locked by disabling the closed-loop feedback system. Subsequently, the bias voltage is ramped to measure the local tunneling spectra of an oxide film.	
3.10	Schematic diagram representing the experimental set-up and the bias voltage setting during constant voltage stressing. Upon achieving the bias setpoint, the tip-sample vacuum gap spacing is frozen by disabling the closed-loop feedback system. Subsequently, the bias voltage is increased to a higher value and kept constant to accelerate localized BD.	68
3.11	Schematic diagram representing the experimental set-up and the bias voltage behavior during constant voltage stressing. Upon achieving the bias setpoint, the tip-sample vacuum gap spacing is locked by disabling the closed-loop feedback system. Subsequently, the bias voltage is increased to a higher value to induce localized BD.	68
3.12	(a) Cascade Microtech 8-inch wafer probe station, a platform for holding device under test (DUT). (b) Keithley 4200-SCS Semiconductor Characterization System, a test environment for controlling the various electrical measurements.	69
4.1	Constant current STM images (CCI) of $\text{HfO}_2/\text{SiO}_2$ gate stack for a series of 6 consecutive scans at bias setpoint: $V_s = +5$ V, $I = 90$ pA. The color contrast in the CCI is attributed to the local variation of the tip height (i.e. a manifestation of the gate stack thickness and/or conductance) over a 500×500 nm ² high- κ surface. The bright spots in the images are observed to have shifted slightly for every subsequent scan due to the thermal drift. Two distinct observations: (i) Pre-existing bright spots, labeled as A and B, which exist at the 1 st scan. (ii) Electrical stress-induced bright spots, labeled as C, D, E, F and G, which appear gradually at a relatively uniform contrast area at the 3 rd scan (5 th scan) for C, D and E (F and G).	76
4.2	Tunneling spectra of bright spot A (pre-existing), extracted from the CITS current map, corresponding to the CCI from the 1 st scan to the 6 th scan in Fig. 4.1. A consecutive increase in the negative- I_t is observed for every successive scan, consistent with the corresponding increase in the brightness and dilation of bright spot A in Fig. 4.1. However, from the 3 rd scan onwards, the negative- I_t shows only marginal increase due to the tip retraction. Reference tunneling spectrum measured at a more uniform contrast area is shown.	77
4.3	Topography height profile of bright spot A extracted from the 1 st scan to the 6 th scan in the CCI in Fig. 4.1. The height and diameter of the bright spot increase significantly after the 2 nd scan, signifying local oxide degradation. The edge of bright spot A is taken as the height reference.	77

-
4.4	Tunneling spectra of bright spot C (electrical stress-induced), extracted from the CITS current map, corresponding to the CCI from the 2 nd scan to the 6 th scan in Fig. 4.1. The negative- I_t shows a slight decrease initially at the 3 rd scan; note the subtle appearance of bright spot C in the CCI in Fig. 4.1(c). At the 4 th scan, the negative- I_t increases significantly; consistent with a significant increase in the size and brightness of bright spot C in Fig. 4.1(d). From the 5 th scan onwards, the negative- I_t shows a near-parallel shift to the less negative voltage regime. Reference tunneling spectrum measured at a more uniform contrast area is shown.	79
4.5	Topography height profile of bright spot C extracted from the 1 st scan to the 6 th scan in the CCI in Fig. 4.1. The height and diameter of the bright spot increase significantly after the 3 rd scan, signifying local oxide degradation. The edge of bright spot C is taken as the height reference.	79
4.6	(a) Zoom-in view ($150 \times 150 \text{ nm}^2$) of bright spot A in the topographical image in Fig. 4.1. The bias set point was $V_s = +5 \text{ V}$, $I = 90 \text{ pA}$ during image acquisition. Color contrast is adjusted to identify the higher region. Granular features are observed in the image. (b) Zoom-in view of the CITS current map corresponding to the topographical image in (a); bias voltage $V_s = -5 \text{ V}$. (c) and (d) show the cross-sectional height and current profiles of a granular feature, along dashed line 1-2 in (a) and (b) respectively.	81
4.7	Tunneling spectra of grain (denoted by solid (open) circle for the 1 st scan_Grain (6 th scan_Grain)) and GB (denoted by solid (open) triangle for the 1 st scan_GB (6 th scan_GB)) in bright spot A extracted from the CITS current map. The negative- I_t of GB (grain) increases by $\sim 36 \%$ ($\sim 4 \%$) after a series of successive ramp voltage stress, implying a higher degradation rate in the IL beneath a GB as compared with its counterpart beneath a grain. The tunneling spectra in the positive V_s regime all coincide because they are normalized to the same bias set-point ($+5 \text{ V}$, 90 pA).	82
4.8	Zoom-in view ($150 \times 150 \text{ nm}^2$) of the current map (at $V_s = -5 \text{ V}$) of bright spot A (a) at the 2 nd scan, and (b) at the 6 th scan. Grains, the granular features denoted by the green dashed line, are clearly observed in both images. An increase in the intensity of bright shades (corresponds to negative- I_t) encompassing the grains at the 6 th scan indicates a severe degradation of GBs	82
4.9	Local electric field across the IL (E_{IL}) beneath a high- κ GB as a function of EOT_{HK} . For a $\text{HfO}_2(5 \text{ nm})/\text{SO}_2(1 \text{ nm})$ gate stack, the local electric field across the normal high- κ layer is almost identical to that at the IL, as indicated by a converged curves at $EOT_{HK} = 5 \text{ nm}$. However, as the EOT_{HK} decreases due to degradation of low resistive GB, the E_{IL} is observed to increase rapidly, signifying a larger voltage drop across the IL. Assume the V_{ox} is 1.2 V .	85

-
4.10	(a) The tendency of oxygen vacancies to segregate along high- κ GBs could introduce conduction sub-bands in the high- κ energy bandgap, which facilitate electron tunneling through the high- κ layer. The low-resistive high- κ GBs induce a high “voltage loading” on the underlying IL, accelerating trap generation at this location. Upon achieving the critical trap density, a percolation path is formed and the entire gate stack BD occurs. (b) Schematic energy band diagram of $\text{HfO}_2/\text{SiO}_2$ gate stack under substrate injection bias. Electrons tunnel through the high- κ layer facilitated by sub-conduction bands present in the high- κ energy gap.	85
4.11	High-resolution cross-sectional transmission electron micrographs of the (a) single-step deposited HfSiO_x gate stack with 2.3 nm HfSiO_x and 1.4 nm interfacial SiO_x layer (IL) thicknesses, and (b) two-step deposited HfSiO_x gate stack showing identical HfSiO_x and IL thicknesses.	89
4.12	Electron energy loss spectroscopy (EELS) profiles of the $\text{TiN}/\text{HfSiO}_x/\text{IL}$ gate stacks processed using (a) single-step and (b) two-step (D&A) method. Both samples exhibit almost similar chemical composition profiles [129].	89
4.13	Constant-current STM images of the HfSiO_x/IL gate stack: (a) single-step deposited HfSiO_x ; (b) two-step deposited HfSiO_x . The scanned area is $100 \times 100 \text{ nm}^2$ and the bias set-point is $V_s = 3 \text{ V}$, $I_t = 25 \text{ pA}$. The contrast in the image is caused by variations in the vacuum spacing between the probe and the high- κ surface, needed to maintain I_t at the preset value as the probe scans across a given area. Bright shades correspond to micro protrusions and/or local conductive regions due to the presence of electronic traps. (c) and (d) are the corresponding CITS current maps ($@V_s = -3.5 \text{ V}$) for the single- and two-step deposited HfSiO_x respectively. Bright shades denote regions of higher I_t .	92
4.14	High resolution cross-sectional transmission electron micrographs of the (a) single-step and (b) two-step deposited HfSiO_x gate stacks. Grain orientation lines (denoted by dashed circle) can be observed in the single-step deposited HfSiO_x gate stack, suggesting the formation of larger grains in the high- κ film. The two-step deposited HfSiO_x , by contrast, exhibits a nanocrystalline structure.	93
4.15	Tunneling current versus bias voltage (I_t - V_s) characteristics measured at grain boundaries (GBs) and grains in the single-step deposited HfSiO_x gate stack. The I_t - V_s curves measured at the bright and dark shades in the CITS current map (Fig. 4.13(d)) of the two-step deposited HfSiO_x are also shown for comparison. The bias voltage, V_s , was swept from 5 to -5 V . All the tunneling spectra	95

coincide at $V_s = 5$ V due to normalization. The “plateau” in the negative V_s regime is caused by deep-depletion, which limits the increase of I_t .

- 4.16 (a) Tunneling current I_t distributions of the single- and two-step deposited HfSiO_x . The area probed by STM is $100 \times 100 \text{ nm}^2$. I_t is normalized to the value at the 63rd percentile. The y-axis gives the cumulative fraction of pixels having the I_t equal or less than a given value. The steeper distribution curve of the two-step deposited HfSiO_x implies an improved electrical uniformity. (b) Correlation between the I_t under substrate injection ($V_s = -2$ V) and tip injection ($V_s = 2$ V) biasing for the single- and two-step deposited HfSiO_x . Grain and grain boundaries of single- (two-) step sample are denoted by solid circle (triangle) and open circle (triangle) respectively. 96
- 4.17 (a) Evolution of the tunneling current I_t during constant voltage stressing, applied via the STM probe, on the bright (corresponds to grain boundaries) and dark (corresponds to grains) shades in the CITS current map of the single-step deposited HfSiO_x (Fig. 4.13(c)). The grain boundaries and grains were stressed at 3 V and 5.5 V, respectively. (b) The corresponding I_t evolution for CVS applied at the bright and dark shades in the CITS current map of the two-step deposited HfSiO_x (Fig. 4.13(d)). The T_{BD} data measured at the grain boundaries of the two-step HfSiO_x are generally longer than that measured at the grain boundaries of the single-step HfSiO_x . 98
- 4.18 Weibull distributions of the time-to-breakdown T_{BD} of the local constant voltage stress of Fig. 4.17. T_{BD} is defined as the time at which the tunneling current I_t increases abruptly to the compliance setting (10 nA). The T_{BD} measured at the GB of the two-step HfSiO_x (open triangle) is about one order of magnitude longer than that measured at the GB of the single-step HfSiO_x (open box). 100
- 4.19 Statistical spread in the tunneling current I_t of the single- and two-step HfSiO_x gate stacks. The scanned area is $100 \times 100 \text{ nm}^2$. All the values are normalized to the average I_t obtained in the 1st scan of the single-step HfSiO_x . (a) I_t measured under tip injection biasing, $V_s = 2$ V (probes traps in the high- κ [104]), and (b) I_t measured under substrate injection biasing, $V_s = -2$ V (probes traps in the IL [104]). The single-step HfSiO_x shows a substantial increase in the I_t spread after four successive scans, implying an inhomogeneous trap generation in the high- κ and IL. The two-step HfSiO_x , by contrast, shows a lesser increase in the I_t spread and the average I_t . 100
- 4.20 Schematic diagrams of the cross-sectional (a) single-step HfSiO_x , in which the grain boundaries propagate through the entire high- κ film, and (b) two-step HfSiO_x , in which the grain boundaries in the two thin high- κ layers are, in general, misaligned. 102

4.21	Schematic energy band diagrams of the single-step (a) and two-step (b & c) HfSiO _x gate stacks under substrate injection biasing. (i) and (ii) denote the high- κ sub-layer with and without the presence of grain boundaries respectively. The presence of grain boundaries induce the formation of conduction sub-bands in the high- κ bandgap, thus facilitating electrons tunneling through the layer.	102
4.22	(a) The gate leakage current density J_g versus gate voltage V_g characteristics of the MOS capacitors ($50 \times 50 \mu\text{m}^2$) employing the single- and multi-step (D&A) high- κ gate dielectrics. (b) Weibull distributions obtained from the statistics of the V_{BD} , defined as V_g corresponding to an abrupt increase in J_g depicted in (a).	104
4.23	(a) Stress-induced increase in the gate leakage current density J_g , measured at gate voltage $V_g = V_{FB} - 1 \text{ V}$, where V_{FB} is flat band voltage. The MOS capacitors ($50 \times 50 \mu\text{m}^2$) were stressed at $V_g = -2.2 \text{ V}$. (b) The corresponding increase in the hysteresis of the capacitance-voltage ($C-V$) curve, for the constant voltage stress of (a). The $C-V$ measurement was made by sweeping the gate voltage from -1 to 1 V and back to -1 V .	104
4.24	Comparison of interface trap (N_{it}) and bulk trap (N_{ot}) measured by the charge pumping at 1 MHz and single pulse methods for the single- and two-step HfSiO _x gate stacks [129].	105
5.1	Schematic illustration of the multistep deposition cum UVO annealing process for the 4step_UVO1min sample. An ultrathin ($\sim 0.5 \text{ nm}$) HfZrO layer was grown on the SiO ₂ /Si substrate via ALD (at a temperature of 250°C and a pressure of 0.2 Torr) and the sample was then subjected to <i>ex-situ</i> room-temperature anneal in O ₂ ambience under UV light (254 nm wavelength) for 1 min . The ozone concentration produced was $\sim 500\text{-}1000 \text{ ppm}$. This process is repeated until the final thickness is achieved.	112
5.2	High-resolution cross-sectional transmission electron micrographs of HfZrO/IL gate stacks deposited via ALD and annealed under various conditions: (a) RTA; (b) as-dep; (c) 2step_UVO30s; (d) 2step_UVO2min; and (e) 4step_UVO1min. All the gate stacks examined show nearly identical physical thicknesses for the HfZrO film and interfacial SiO _x layer.	113
5.3	The XPS core-level spectra of (a) Si $2p$ and (b) Hf $4f$ measured from RTA, 2step_UVO2min, and 4step_UVO1min gate stacks. No significant shift is observed. All the peak positions were referenced to the C $1s$ peak binding energy (BE) at 285 eV . [154].	113
5.4	Row (a) depicts the constant-current STM topography images of the (i) RTA, (ii) as-dep, (iii) 2step_UVO30s, (iv) 2step_UVO2min, and	116

- (v) 4step_UVO1min HfZrO gate stacks acquired at $V_s = 3$ V and $I_t = 25$ pA. Bright shades denote the locations where the tip retracts due to physical protrusions or local conductivities on the high- κ surface. Row (b) depicts the corresponding current maps (at $V_s = -3$ V) of the samples. Bright shades denote the regions of higher I_t . The scanned area is 100×100 nm².
- 5.5 Corresponding average tunneling current spectra of the various 100×100 nm² scanned regions depicted in Fig. 5.4. The current “plateau” in the negative bias regime is due the p-Si substrate entering into deep-depletion, which limits the number of carriers tunneling across the gate stack. In the positive bias regime, all the curves artificially converge to the set point of 0.1 nA, 5 V used for the tunneling current spectroscopy. The RTA sample shows the largest negative- I_t , indicating a rapid degradation at the underlying IL. However, the negative- I_t is clearly reduced for the as-dep and UVO annealed samples. 116
- 5.6 (a) Tunneling current I_t distributions of the various HfZrO gate stacks probed by STM over an area of 100×100 nm² at $V_s = -3$ V. The y-axis denotes the cumulative percentage of pixels having the I_t equal or less than a given value. To facilitate the comparison, the I_t of the various samples are normalized to the value at the 63rd percentile. The degree of the I_t distribution stretch-out indicates the local electrical uniformity for a given scanned area. (b) Stress-induced increase in the tunneling current I_t measured over a 100×100 nm² area of the various HfZrO gate stacks probed by STM. The I_t was measured at $V_s = -3$ V. The samples were repeatedly subjected to constant-current topography scanning at $V_s = 5$ V. A rapid degradation rate is observed for the RTA sample, whereas the as-dep and UVO annealed samples exhibit no significant increase in the I_t after successive stress. 119
- 5.7 (a) Schematic illustration of the programmed gate voltage waveform applied during the pulse C - V measurement. The ramp-up/down rate is set at 100 kV/s. Arrows denote the transition edges at which the C - V curve was recorded. The capacitor was charged with a negative voltage pulse ($V_g = -2.6$ V, pulse width = 1 ms) and then discharged at $V_g = 0.5$ V, with the C - V curve monitored at selected intervals (e.g. t_{D1} , t_{D2} , etc.). (b) Evolution of the fresh, after negative pulse ($V_g = -2.6$ V) charged, and discharged C - V curves. A nearly full discharge of the trapped “bulk” charge is observed after ~ 100 s. 123
- 5.8 Discharging-induced V_{FB} shift (relative to the first charged C - V curve) versus discharging time for the various HfZrO gate stacks. The estimated lateral shift (~ 100 mV) between the charged and fresh C - V curves expected from a quasi-static measurement which takes ~ 10 s to complete is shown for the RTA sample. 124

5.9	Schematic illustration of the multistep deposition cum two-step (UVO-RTA) annealing technique. An ultrathin (~ 0.7 nm) HfO_2 layer is grown on p-Si substrate via ALD (at a temperature of 250°C and a pressure of 0.2 Torr) and subsequently subjected to room-temperature UVO annealing for 1 min followed by RTA at 420°C in N_2 for 30 s. The ozone concentration produced was ~ 500 - 1000 ppm. This process is repeated until the final thickness is achieved.	127
5.10	High-resolution cross-sectional transmission electron micrographs of the (a) as-dep, (b) RTA, (c) RTA(1000°C), (d) UVO, (e) UVO-RTA, and (f) RTA-UVO samples. Almost all the samples show identical HfO_2 and SiO_2 IL thicknesses, except for the RTA(1000°C) which exhibits a thicker IL due to a high temperature involved.	128
5.11	(a) Quasi-static high-frequency (1 MHz) capacitance-voltage (C - V_g) characteristics of the $\text{TiN}/\text{HfO}_2/\text{IL}$ MOS capacitors subjected to various annealing conditions. The inset shows an almost identical EOTs (extracted from the corresponding C - V_g curves) for all the samples except for the one annealed at 1000°C . The gate stack work function is estimated to be 4.7 eV, typical of the TiN/HfO_2 stack. The lesser accumulation capacitance ($\sim 1.4 \mu\text{F}/\text{cm}^2 @ -1.15$ V) is because of the finite capacitance of the lowly doped ($\sim 6 \times 10^{15} \text{ cm}^{-3}$) p-Si substrate (Debye length ~ 500 Å). (b) Average gate current density versus gate voltage (J_g - V_g) characteristics, at low V_g regime, of the various gate stacks measured under room temperature (300 K). Each curve is an average of 20 devices.	130
5.12	(a) Time-to-breakdown T_{BD} distributions of the various $\text{TiN}/\text{HfO}_2/\text{IL}$ MOS capacitors subjected to constant gate voltage (-4.8 V) stressing at 125°C . Capacitor area: $50 \times 50 \mu\text{m}^2$. The stress test was not performed on the gate stack annealed at 1000°C since its breakdown voltage is less than 4.8 V. (b) Flat band voltage shift ΔV_{FB} as a function of stress time under constant gate voltage (-4.4 V) stressing at 125°C .	132
5.13	Row (a) depicts the $100 \times 100 \text{ nm}^2$ constant-current STM topography images of the (i) RTA(1000°C), (ii) as-dep, (iii) UVO, (iv) RTA, (v) RTA-UVO, and (vi) UVO-RTA HfO_2 gate stacks acquired at $V_s = 2.5$ V and $I_t = 15$ pA. Bright regions correspond to the locations where the probe retracts from the high- κ surface, either due to micro-protrusions or higher local conductance [104]. Row (b) depicts the corresponding current maps (at $V_s = -3$ V) obtained by the CITS technique. Bright shades denote regions of higher local I_t .	133
5.14	(a) Corresponding average I_t - V_s curves extracted from the $100 \times 100 \text{ nm}^2$ scanned area depicted in Fig. 5.13(b). Bias voltage V_s was swept from 5 V to -5 V. All the I_t - V_s curves converge at $V_s \sim 5$ V	134

- due to an adjustment of the vacuum gap between the probe and the high- κ surface by a feedback circuit, in order to achieve the same I_t at a given V_s . (b) Corresponding statistical spread in the I_t of the current maps (at $V_s = -3$ V; traps in the IL was probed under substrate injection biasing [104]) depicted in Fig. 5.13(b). To facilitate the comparison, I_t is normalized to the average I_t of the RTA(1000 °C) sample.
- 5.15 The XPS core-level spectra of (a) Hf 4*f* and (b) O 1*s* obtained from the HfO₂/IL gate stacks annealed at various conditions. All the spectra were referenced to C 1*s* peak binding energy (BE) at 285 eV. 138
- 5.16 (a) $I_{\text{Hf-O}}/I_{\text{(Hf-O-Si, Si-O)}}$ intensity ratios of the respective gate stacks obtained from the O 1*s* core level spectra in Fig. 5.15(b). The intensity of the peaks centered at BE ~530 eV and that at the shoulder BE ~531-532 eV correspond to the contents of HfO₂ and Hf silicate respectively. (b) The XPS core-level spectra of Si 2*p* obtained from the HfO₂/IL gate stacks annealed at various conditions. 138
- 5.17 Background corrected O K edge EELS spectra of (a) 2.5 nm HfO₂ film, and (b) 1 nm SiO₂ IL for the various gate stacks. The peak intensity corresponds to the amount of local oxygen atoms (i.e. lower peak intensity signifies oxygen deficiency). Double peaks characteristic can be seen in the O K edge spectra of HfO₂, consistent with the literature [172]. A growing 537.8 eV peak intensity is observed in both HfO₂ film and SiO₂ IL for the samples annealed with UVO. 140
- 5.18 High resolution cross-sectional transmission electron micrographs of the (a) as-dep; (b) UVO-only; (c) UVO-RTA; (d) UVO-RTA500°C; (e) UVO-RTA600°C; (f) RTA-UVO gate stacks. All the samples show almost identical HfO₂ (~2.5 nm) and IL (~0.3 nm) thicknesses, except for the UVO-RTA600°C sample which exhibits a thicker IL (~0.5 nm). 142
- 5.19 High-frequency (1 MHz) gate capacitance versus gate voltage (C_g - V_g) curves of the various TiN/HfO₂ gate stacks. Each C_g - V_g curve is an average data of more than ten devices. The gate stack work function is estimated to be 4.7 eV, typical of the TiN/ HfO₂ stack. The average k -value of the gate stack is 15, based on the extracted EOT of 7.5 Å (which corresponds to 4.6 μF/cm² oxide capacitance). The lesser accumulation capacitance (~2.6 μF/cm² @ -1.1 V) is because of the non-negligible Si capacitance of the lowly doped (~6 × 10¹⁵ cm⁻³) p-substrate (Debye length ~500 Å).. 143
- 5.20 (a) Average gate current density versus gate voltage (J_g - V_g) characteristics of the various TiN/HfO₂ gate stacks measured under room temperature (300 K). More than ten devices were measured for 145

each split. A significant reduction in J_g is observed for the UVO-RTA sample ($\sim 8 \text{ mA/cm}^2$ @ $V_{FB} - 1 \text{ V}$), further reduction is achieved by increasing the RTA temperature to 500°C as shown by the UVO-RT500 $^\circ\text{C}$ sample ($\sim 3.5 \text{ mA/cm}^2$ @ $V_{FB} - 1 \text{ V}$). (b) Gate current density J_g versus EOT at 300 K in comparison with the poly-Si/SiO₂ benchmark and the J_g reported in literature.

- 5.21 (a) Time-to-breakdown T_{BD} distributions of the various TiN/HfO₂ gate stacks subjected to constant voltage (-2.2 V) stressing at 125°C . The MOS capacitor area used in the TDDDB test is $20 \times 20 \mu\text{m}^2$. The Weibull slopes β of the respective distributions are ~ 1.5 . The UVO-RTA sample exhibits the longest T_{BD} with ~ 2.5 , 6, and 9 times longer than that of the RTA-UVO, RTA-only, and as-dep samples respectively. (b) Flat band voltage shift ΔV_{FB} as a function of stress time at 125°C . The applied electric field across the gate stack is estimated to be $\sim 10 \text{ MV/cm}$. The slowest positive trap generation is observed for the UVO-RTA sample. 147
- 6.1 (a) $50 \times 50 \text{ nm}^2$ constant-current STM image of the HfO₂/SiO_x gate stack at a set point of $V_s = 3 \text{ V}$ and $I_t = 25 \text{ pA}$. Under positive bias voltage, the polycrystalline HfO₂ layer is imaged, as verified by the granular features (bright regions of diameter $\sim 20 \text{ nm}$) in the image. (b) Corresponding current image (under substrate injection bias at $V_s = -3 \text{ V}$) showing (bright shades) “leaky” grain boundaries encompassing the less conductive grains. 153
- 6.2 (a) Evolution of the tunneling current I_t under localized constant voltage stressing at 4 V , applied via the STM probe, on the grain boundary and grain respectively (Fig. 6.1). As expected, the gate stack breakdown occurs much earlier for the stressing applied at the grain boundary than that applied at the grain. (b) Weibull distributions of the time-to-breakdown T_{BD} for the stressing applied at the GBs ($\beta = 1.45$) and grains ($\beta = 1.36$). T_{BD} is defined as the time when the tunneling current I_t increases abruptly to the compliance setting (10 nA). 155
- 6.3 $200 \times 200 \text{ nm}^2$ STM current map (at $V_s = -3 \text{ V}$) of the HfO₂/SiO_x gate stack acquired via constant imaging tunneling spectroscopy (CITS). The image is divided into smaller regions to analyze the percentage of the area occupied by the GBs (bright shades in the current map) as a function of gate area size. 158
- 6.4 Number of “devices” (obtained by dividing the current map in Fig. 6.3 into smaller regions) having a certain percentage of the gate area occupied by the grain. As the gate area size approaches $20 \times 20 \text{ nm}^2$, the grain BD becomes increasingly important as the number of regions having almost the entire area occupied by the grain increases substantially. At the same time, it is observed that there are a small number of regions with $> 50 \%$ of the area occupied by the GBs, implying the co-existence of the GB and grain BD influence on the 159

T_{BD} statistics of the small area devices.

- 6.5 (a) An example of a T_{BD} distribution due to a mixture of the grain and GB BDs in the small area devices. The distribution is obtained from the Weibull distribution for grain and GB BDs in Fig. 6.2(b), by taking the percentage of the grain and GB BDs to be 75 % and 25 %, respectively (cf. Fig. 6.4). (b) A Combination of the two T_{BD} distributions in Fig. 6.2(b) yields a bimodal Weibull distribution, with the steep Weibull slope at the lower percentile of the distribution predominated by the shorter T_{BD} at the GBs whereas the shallow Weibull slope at the upper percentile of the distribution predominated by the longer T_{BD} at the grains. 160
- 6.6 Weibull distributions of the T_{BD} for the enhanced field (at $V_s = 4.5$ V) accelerated stressing applied at the GBs (filled triangle) and grains (open triangle). T_{BD} is defined as the time when the I_t increases abruptly to the compliance setting (10 nA). 160

LIST OF TABLES

3.1	Gate stack structures and physical thicknesses of the high- κ film and IL measured from the cross-sectional HRTEM for Set A, B and C samples. The high- κ films were deposited via ALD and annealed in various conditions. All the samples were grown on p-Si (100) substrate.	52
-----	--	----

LIST OF ABBREVIATIONS

AFM	atomic force microscopy
ALD	atomic layer deposition
BD	breakdown
CAFM	conductive atomic force microscopy
CBM	conduction band minima
CCI	constant current imaging
CITS	constant current tunnelling spectroscopy
CMOS	complementary metal-oxide semiconductor
CVS	constant voltage stress
EELS	electron-energy loss spectroscopy
EOT	electrical oxide thickness
FEL	fast entry lock
GF	GlobalFoundries
IC	integrated circuit
IL	interfacial layer
LDOS	local density of state
MOCVD	metal-organic chemical vapor deposition
MOS	metal-oxide-semiconductor
MOSFET	metal-oxide-semiconductor field-effect transistor
MOSCAP	MOS capacitor
PDA	post deposition annealing
PLD	pulsed laser deposition
PVD	physical vapor deposition
RTA	rapid thermal annealing
RTO	rapid thermal oxidation
RVS	ramp voltage stress
SILC	stress-induced leakage current
SiON	Silicon Oxynitride
SIS	substrate injection stress

SoC	systems-on-a-chip
SPM	scanning probe microscopy
STM	scanning tunnelling microscopy
STS	scanning tunnelling spectroscopy
TAT	trap-assisted tunnelling
TDDDB	time-dependent dielectric breakdown
TEM	transmission electron microscopy
TIS	tip injection stress
TSMC	Taiwan Semiconductor Manufacturing Company
TSP	titanium sublimation pump
TOF–SIMS	time-of-flight–secondary-ion-mass spectroscopy
UHV	ultra-high vacuum
UVO	ultraviolet-ozone
XPS	x-ray photoelectron spectroscopy
XRD	x-ray diffraction

LIST OF SYMBOLS

μ	micro
κ or ε	dielectric constant
q	electronic charge
β	Weibull slope
$\text{\AA}$	Angstrom
σ	capture cross section
φ	wave function
C - V	capacitance-voltage characteristic
D_{it}	density of interface charge
E_{f}	Fermi level
E_{t}	energy level of electronic trap
E_{relax}	relaxation energy
I - V	current-voltage characteristic
I_{t}	tunneling current
J - V	current density-voltage characteristic
N_{t}	volume density of traps
V_{o}	oxygen vacancy
V_{s}	sample bias

Chapter 1 INTRODUCTION

1.1 Background

1.1.1 Transistor's scaling towards nanometer regime

Metal-oxide-semiconductor field effect transistor (MOSFET) integrated circuit (IC) entered the production scene in the early 1970s and has since undergone rapid progress. The fact that MOSFET scales successfully according to Moore's law, has enabled a high density of transistors to be integrated on a chip, thus achieving continuous improvements in performance (speed), low off-state power consumption, and reduced production costs [1-2]. This gives a significant impetus to the market expansion of Si-based microelectronics industry for the past five decades. Moore's law, established by Gordon Moore, projects the number of transistors per IC doubles every 18 months with the shrinking of the transistor's critical feature size. Fig 1.1 shows that the number of transistors integrated in every generation of Intel's microprocessors follows the Moore's law. For instance, the Dual-core Itanium-2 microprocessor, introduced in 2006, integrates 1.7 billion transistors on 596 mm² chip area using 90 nm process technology as compared to that of the Intel 4004 microprocessor, which integrates only 2300 transistors on 12 mm² chip area using 10 μm process technology [3].

Fig. 1.1: The number of transistors, using SiO_2/SiON gate dielectric, integrated in the different generation of Intel's microprocessors grows exponentially for the last three decades, as projected by Moore's law [1, 3].

Apart from the advanced lithography technology, another important key element that contributes to the successful downscaling of MOSFET is the superior electrical properties of the SiO_2/SiON gate insulator. The thermally grown SiO_2/SiON dielectric is highly stable on the single crystalline silicon substrate, forming an excellent Si-SiO₂ interface with a low intrinsic interface states density of $\sim 10^{10} \text{ cm}^{-2} \text{ eV}^{-1}$ [4]. In addition, the SiO_2/SiON dielectric exhibits stable thermal/chemical properties under high-thermal-budget complementary metal-oxide-semiconductor (CMOS) processing, excellent electrical isolation between the polysilicon gate and channel due to a relatively large bandgap of $\sim 9 \text{ eV}$, low electron and hole trapping rate, and large breakdown (BD) electric field of $\sim 15 \text{ MV/cm}$. However, further MOSFET downscaling to sub-45 nm technology node, in order to meet the market demand for high performance IC, faced great reliability challenges. From equation (1-1), it is known that a reduction in the channel length, L , or an increase in the gate dielectric

capacitance, C_{inv} , is the key to achieve an increased $I_{d,sat}$. In order to increase the C_{inv} , a reduction in the dielectric physical thickness, t , is required as shown in equation (1-2). However, the present SiO_2/SiON gate dielectric has reached a physical thickness of less than 2 nm, which is vulnerable to quantum mechanical tunneling [5-6]. Further thickness reduction to 1.5 nm will result in a gate current density exceeding the limit of $\sim 1 \text{ Acm}^{-2}$ (at 1 V) and thereby leading to an unacceptable power dissipation [2, 7-8] since the tunneling probability (corresponds to the gate leakage current) increases exponentially with a decrease in the dielectric thickness [5, 9] (Fig. 1.2). It was reported that for the SiO_2/SiON thickness less than 7 Å, the full band gap is not formed [10-11], thus a large charge carriers flow through the dielectric and result in an accelerated trap generation which eventually triggers oxide BD [12-14]. Therefore, it is evident that the further downscaling of SiO_2/SiON thickness is predominantly limited by the high rising gate leakage current.

$$I_{D,sat} = \frac{W}{L} \mu C_{inv} \frac{(V_G - V_T)^2}{2}. \quad (1-1)$$

$$C = \frac{\kappa \epsilon_0 A}{t} \quad (1-2)$$

Fig. 1.2: Simulated I_g-V_g characteristic under inversion conditions for the various SiO_2 thicknesses. The gate oxide can be scaled down to 20 Å before exceeding the 1 Acm^{-2} (at 1 V) limit for the high performance CMOS logic [6].

1.1.2 High- κ metal oxide as replacement gate dielectric

Based on the Intentional Technology Roadmap for Semiconductors (IRTS), the SiO_2/SiON gate dielectric with thickness less than 1.5 nm is required in the next generation MOSFET for both high performance and low operating power logic applications. However, as the CMOS device downscaling reaches atomic limit, tremendous researches and reliability studies have been focused on the high dielectric constant materials for the replacement of SiO_2/SiON dielectric in advanced CMOS technology as it is deemed to be the most promising solution at lower research and production cost due to the matured Si-based CMOS processes [10, 15-16]. The high- κ reliability study is useful for the successful implementation of new generation transistor such as FinFET. Intel had unveiled 22 nm SoC using tri-gate implementation of FinFET technology and TSMC is accelerating the production using the technology. To ramp FinFET on an established node is a great challenge due to the process reliability though Intel has shown promise in the technology. As pointed out in Section 1.1.1, C_{inv} is needed to be increased in order to achieve a larger $I_{\text{d,sat}}$. However, a further thinning of the physical oxide thickness is not feasible due to an unacceptable increase in the tunneling current, therefore, increasing the dielectric constant value is the only possible solution to enhance C_{inv} . An insulator with a large dielectric constant

Fig. 1.3: Gate current density and standby power consumption plotted as a function of gate voltage for a chip employing 15 Å SiO_2 gate dielectric and high- κ gate dielectric with the same EOT. The high- κ dielectric clearly suppresses the gate leakage current, thereby achieving low standby power consumption. Assumes at total gate area of 0.1 cm^2 [10, 17].

(κ) value is known as “high- κ ” dielectric. The rare earth metal oxides have been extensively investigated for this purpose. Numerous studies of the various rare earth metal oxides have been reported, such as HfO_2 [17], CeO_2 [18], LaO_2 [19], ZrO_2 [20], etc. HfO_2 , for instance, has a relatively large κ value ranging from 15 to 25, as compared with the SiO_2 which has only 3.9. Therefore, based on equation (1-2), a much thicker high- κ dielectric can be used to suppress the tunneling current while achieving a small ($< 15 \text{ \AA}$) total equivalent oxide thickness (EOT) [21]. As shown in Fig. 1.3, it is clear that high- κ dielectric is required to meet the low-power application requirements. Compared with the 15- $\text{\AA}$ SiO_2/SiON , the high- κ dielectric with the same EOT clearly suppresses the gate current density and reduces the standby power consumption.

Due to the excellent electrical properties of the SiO_2/SiON dielectric, replacing the former in advanced CMOS technology is indeed a challenging task as the high- κ candidates need to meet many requirements as listed below [2, 10] ,

1. High dielectric constant value, such that the gate oxide can be employed for the next few generations of MOSFETs.
2. Forming a good oxide/Si interface. The high- κ material is required to be thermodynamically stable in contact with Si, such that no formation of low- κ interfacial layer (e.g. SiO_x or high- κ silicate).
3. Compatible with the conventional Si-based CMOS processing.
4. Providing an excellent electrical isolation between the metal gate and Si substrate. A large energy band gap is required to prevent the carrier injection into its bands at both oxide/Si and oxide/metal gate interfaces, thereby suppressing the gate leakage current.

5. Low electrically active defects density in the oxide “bulk” and at the oxide/Si interface to prevent charge trapping, which is associated with the threshold voltage (V_{th}) instability, low charge carriers mobility and BD.

Among the high- κ candidates investigated, Hafnium- and Zirconium-based dielectrics are the most promising candidates [2, 10] due to their good thermodynamic stability in direct contact with Si and a relatively large κ value. However, for a successful integration in the CMOS process, several key problems need to be resolved, such as (i) a high as-grown defects density at the oxide/Si interface and in the “bulk”; it was reported that the interface trap density at HfO_2/Si is in the order of $\sim 10^{11}$ - 10^{12} $cm^{-2}eV^{-1}$ [10, 22] as compared to that at the SiO_2/Si in the order of $\sim 10^{10}$ $cm^{-2}eV^{-1}$ [4], (ii) microstructural evolution in the high- κ film under high-thermal-budget front-end processes; for instance, HfO_2 dielectric tends to crystallize when subjected to ~ 600 - 700 °C annealing [23], (iii) formation of interfacial layer; the diffusion of oxygen species from the “bulk” or the surrounding environment reacts with the Si surface, thus forming a low- κ interfacial SiO_x layer (IL). In addition, an intermixing between the high- κ film and SiO_x IL forms a thin high- κ silicate layer, with the thickness increases under high temperature [24].

1.2 Motivation

The replacement of $SiO_2/SiON$ with high- κ dielectric has made the current reliability studies more complicated. Unlike the former, which retains its amorphous structure throughout the CMOS processing, the high- κ dielectric is kinetically less stable and undergoes phase change under high-thermal-budget front-end processes [23], thus giving rise to interface roughness and formation of structural defects,

particularly grain boundaries (GBs), as evidenced in the high resolution transmission electron micrograph (HRTEM) and x-ray diffraction (XRD) analysis. The broken bonds at the interfaces between randomly-orientated grains in a polycrystalline high- κ film present low resistive GBs that connect the metal gate and the Si substrate [15]. In addition, a lower (relative to the “bulk”) diffusion barrier at GBs tends to facilitate the diffusion of external oxygen species to the Si surface, thus forming a thin SiO_x IL at the high- κ /Si interface. These oxygen species may come from either the broken bonds in the oxide “bulk” or the process chamber wall [25-26]. The presence of SiO_x IL has been aggressively debated over its advantages and disadvantages on a device performance. On one hand, the IL is deemed important for the channel carrier mobility and the reliability margin since this layer determines the entire gate stack BD. On the other hand, the IL may restrain the EOT scaling due to a relatively low- κ value. Due to a more complicated high- κ gate stack structure (i.e. multilayer dielectrics with polycrystalline high- κ film) compared with a mere single amorphous SiO_2/SiON gate dielectric, there are more uncertainties arise in the interpretations of the high- κ devices electrical data. For instance, the area-scaled time-dependent dielectric breakdown (TDDB) plot obtained from the high- κ device BD statistics exhibit bimodal Weibull distribution, with the shallow Weibull slope of the higher percentile is mainly associated with the small area device BDs whereas the steep Weibull slope of the lower percentile is mainly associated with the large area device BDs [27-28]. In addition, non-uniformity in the distribution of the defect density throughout the high- κ dielectric contributes to the non-area scalability of the TDDB and I - V characteristics. The electronic property of GB is unknown and its role in the dielectric BD is unclear since the essential microscopic information is masked out due to an area-averaged effect in the standard electrical characterization.

When a transistor is scaled down to the nanometer regime, the electrical characteristics of a device are susceptible to small perturbations in properties such as the microstructural change, non-uniformity and oxide defects density. It is therefore crucial to have a metrology tool that can precisely measure the physical and electronic properties of the local defects that control the performance and reliability of a device. This information is essential to advance the understanding of the high- κ dielectric BD physics and thus facilitates the development of the advanced high- κ gate stack for future CMOS applications. The motivation of this research is to precisely measure, via superior lateral resolution ultrahigh vacuum scanning tunnelling microscopy (UHV-STM), the electrical characteristics of local defects for a high- κ stack prepared under various processes and annealing conditions. This characterization approach complements the standard electrical characterization method by providing microscopic insights, therefore improving the interpretation of the ambiguous experimental observations.

1.3 Objectives

The aim of this work is to study the electrical characteristics of the local defects in advanced high- κ stacks at nanoscale level and subsequently correlated to the macroscopic electrical data. The objectives are listed as below:-

- To separately study the electronic properties of grains and GBs in a polycrystalline high- κ film. The stress-induced degradation at GBs and their impact on the electrical state of the underlying IL, which controls the entire gate stack BD, will be examined.
- To study the effect of GBs on the reliability of an engineered high- κ gate stack prepared using the multistep deposition and RTA (600 °C, 30 s in N₂) technique.

- To study the room-temperature ultraviolet-ozone (RTUVO) annealing effect on the passivation of oxygen vacancies (V_o 's) in a multistep-deposited ultrathin high- κ film.
- To study the performance and reliability of the high- κ stack prepared using the multistep-deposition cum two-step annealing (comprising a RTUVO anneal and a subsequent 420 °C RTA) method.
- To reveal, via nanoscale TDDB testing, the origin of the bimodal Weibull failure distribution observed in the area-scaled TDDB plot obtained from the high- κ /metal gate device BD statistics.

High- κ deposited blanket samples were prepared for STM study, while the MOS capacitors employing an identical high- κ stack were prepared for the standard electrical measurements.

1.4 Organization of thesis

This thesis is divided into seven chapters, organized as follows:-

- *Chapter One* highlights the background, motivation, objectives and original contributions of the research.
- *Chapter Two* presents the literature review of the high- κ integration reliability issues, the challenges faced in the standard electrical characterization technique, and the microscopic data of high- κ oxide acquired via STM.
- *Chapter Three* presents the sample fabrication/preparation process, oxide characterization techniques and a brief description of the UHV-STM system. The electrical data of the high- κ sample is characterized by using prober while the microscopic data by STM.

- *Chapter Four* presents a study on the electrical characteristics of grains and GBs in a polycrystalline HfO₂ gate stack by using STM. The effect of the high- κ GBs on the degradation state of an ultrathin SiO_x IL is investigated. In addition, the reliability of the high- κ gate stack, prepared via the multistep deposition and RTA method, is discussed.
- *Chapter Five* investigates the effect of the room-temperature UVO annealing on the of V_o density in an ultrathin HfO₂ film. The performance and reliability of the multistep deposited cum two-step annealed (comprising a RTUVO anneal and a subsequent 420 °C RTA) HfO₂ film are examined.
- *Chapter Six* provides a comprehensive investigation, via STM, on the origin of the bimodal Weibull distribution observed in the area-scaled TDDB plot obtained from the high- κ /metal gate device BD statistics.
- *Chapter Seven* presents the conclusion of this report and the recommendations for the future research works

1.5 Original contributions

The major contributions of this research lie in the application of STM for the nanoscale characterization of local defects in high- κ stacks and subsequently correlated to the electrical measurement results obtained from MOS capacitors. The original contributions are listed as follows:-

- The topography images of grains and GBs in a 3 nm polycrystalline HfO₂ film are acquired in ultrahigh vacuum (~500 pTorr) by using STM. The electronic properties and stress-induced degradation rates of both features are demarcated via spatially-resolved STM probe. The *real-time* visualization of high- κ film wear-out and the trap generation at the underlying IL are demonstrated at nanoscale level. We show directly that the IL beneath a GB is highly susceptible to degradation

under stress compared to its counterpart beneath a “bulk”. This implies that an oxide BD preferentially occurs near GB.

- We perform nanoscale constant voltage stressing (CVS), via STM probe, over GBs and grains individually. The two sets of failure distributions plotted from the T_{BD} statistics of GBs and grains exhibit almost similar Weibull slopes, implying that a critical number of traps are required to induce BD regardless of the initial state of the pre-existing traps for a given dielectric thickness. Compared with grain, a lesser number of new traps are required to be generated at GB in order to meet the critical number, therefore resulting in a distinctively shorter T_{BD} .
- We show that the GBs effect in the HfO_2 gate stack can be suppressed via the multistep deposition and RTA (600 °C in N_2 for 30 s) technique. Unlike the standard single-step deposited HfO_2 stack, the two-step deposited sample of similar EOT retains a nanocrystalline structure (rather than transforms into a polycrystalline phase) following a 1000 °C anneal. This reduces the GB-related local leakage current density as a result of misaligned GBs in the multilayer high- κ dielectric. In addition, the averaged tunneling current for a given scan area is significantly reduced and a longer T_{BD} is achieved, consistent with the measurement results obtained at the device level.
- Our study shows that the V_o density localized in the conductive paths in HfZrO film can be suppressed via the room-temperature UVO annealing. Compared with the 600 °C RTA sample, we provide evidence, via STM, that the improvement seen in the multistep deposited and UVO-annealed HfZrO is a direct consequence of the suppressed high- κ crystallization by virtue of the low-thermal-budget. Further improvement due to the UVO annealing can be achieved by increasing the number of deposition steps or the UVO annealing time.

- A new method, via the multistep deposition cum two-step annealing (comprising a RTUVO anneal and a subsequent RTA at 420 °C), for enhancing the performance and reliability of gate-last sub-1 nm EOT HfO₂/TiN gate stack is developed. The gate stack exhibits approximately ~two orders of magnitude reduction in gate current density and ~one order of magnitude longer T_{BD} , compared with the as-deposited, UVO-only, and RTA-only gate stacks. The improvement is attributed to the thermal activation by the subsequent RTA that enhances the high- κ structural “repair” by the absorbed ozone. The sequence of the RTUVO annealing and the RTA steps is shown to play a crucial role. The improvement is diminished when either the RTA step is removed or the order of the two annealing steps reversed.
- We provide new insights, via nanoscale TDDDB testing, into the bimodal Weibull failure distributions obtained from the area-scaled T_{BD} plot of high- κ MOSFETs. The individual failure distributions of the GBs and grains in a polycrystalline high- κ gate stack are spatially resolved via localized CVS using STM probe. It is shown that (i) GB-related BD determines the steep Weibull slope at the lower percentile of the T_{BD} distribution (corresponds to the T_{BD} curve of the large area devices); (ii) the combined effect of grain- and GB-related BDs determines the shallow Weibull slope at the upper percentile of the T_{BD} distribution (corresponds to the T_{BD} curve of the small area devices).

Our research works have been published in established international conferences and peer reviewed journals, such as the *Solid States Devices and Materials (SSDM) symposium*, *IEEE Electron Device Letters (EDL)*, *IEEE Transaction Electron Device (TED)* and *Journal of Electrochemical Society (JES)*.

Chapter 2 LITERATURE REVIEW

2.1 Introduction

This chapter is divided into four main sections to systematically review the reliability issues arise from the employment of high- κ dielectric as gate dielectric and the extensive efforts to address the corresponding issues by various research groups. Section 2.2 reviews the integration issues resulting from the replacement of thermally grown SiO_2/SiON dielectric with atomic-layer-deposited high- κ dielectric in the conventional CMOS process. Unlike the former, the latter is thermodynamically unstable and therefore is more susceptible to microstructural change (i.e. phase segregation and uncontrollable IL re-growth) under high temperature process (e.g. source/drain activation spike anneal). Section 2.3 reviews the performance and reliability of high- κ devices (MOS transistor and capacitor). High- κ film is known to be inherently defective with oxide trapped charges and this combined with the presence of structural defects (i.e. GBs), have severely compromised a transistor's performance (e.g. transient threshold voltage V_{th} shift, mobility degradation, high gate current density, bimodal Weibull distribution, lower Weibull slope, etc). Section 2.4 reviews the failure mechanisms proposed in the literature. Section 2.5 reviews the various post deposition annealing (PDA) techniques (e.g. partial pressure oxygen

anneal, nitridation and multistep deposition cum RTA) experimented by various research groups in an effort to reduce the pre-existing defects in high- κ film.

Although the past published results have provided clear physical evidences (via HRTEM) of the transformation of an amorphous high- κ film into a polycrystalline structure following a high temperature anneal, reports concerning the effect of polycrystalline high- κ film on a device's performance and reliability is limited. This is due to the shortcomings of existing electrical characterization metrology, in which the microscopic details of the electronic properties of the crystallized high- κ film and the underlying IL are masked out. To overcome this complication, we employ advanced high-resolution UHV STM to study the high- κ film at nanoscale level. A detailed review of the application of STM in the high- κ film study is presented in Section 2.6.

2.2 High- κ processing and its structural evolution

2.2.1 Atomic layer deposition

Rare earth materials have shown great promise by virtue of the high dielectric constant value. However, the main challenge lies in the process technology in achieving a good quality film with low trap density. Unlike the SiO_2/SiON dielectric, which is “thermally” grown on the Si substrate by rapid thermal oxidation method and thus producing a high quality film, the high- κ dielectric is formed via external “deposition”, therefore the thin film deposition technology plays a crucial role for a successful high- κ integration in the CMOS process. The most common techniques (to date) for high- κ film deposition are physical vapor deposition (PVD), chemical vapor deposition (CVD) and atomic-layer-deposition (ALD). The PVD-based deposition techniques (e.g. sputtering deposition, pulse laser deposition, electron beam evaporation, etc) show several advantages such as high purity, simple deposition

process, low cost, excellent control of the IL growth and consistent composition between the target and the deposited film due to a relatively low process temperature in the vacuum chamber. However, as the technology node is aggressively scaled down, an ultrathin dielectric film (<10 nm) with good conformality is required. Thus, PVD is not expected to be a robust gate dielectric deposition technique due to its poor step-coverage. The CVD-based deposition technique is well established and has been in the production for years to deposit both the dielectric and metallic compounds. Metal-organic chemical vapor deposition (MOCVD), in particular, has been one of the high- κ deposition techniques focused by the industries due to its excellent composition control and film homogeneity. For some CVD processes, nucleation layer is not required to initiate the deposition; this made the grown high- κ film quality independent of the HF-last treated Si surface. The precursors used for deposition are generally volatile metal organics, which exist in liquid form at room temperature and can be delivered to the deposition chamber via a vaporizer. A major concern for the organic precursors is the carbon incorporation in the high- κ film. The ALD technique has received a tremendous interest from the industries in recent years. The Hf- and Zr-based dielectrics have been deposited predominantly by using ALD. Studies [29] have shown that ALD HfO₂ dielectric exhibits a higher capacitance and lesser flatband voltage, V_{FB} , shift following a CVS at 105 °C as compared to its counterpart deposited via MOCVD, indicating a more robust film. Fig. 2.1 illustrates the principle of ALD for Al₂O₃ growth. The chemical reactants in the ALD process are supplied into the deposition chamber via gas pulses under controllable timing. The film thickness depends on the number of reaction cycles. In addition, the surface chemical reaction is self-limiting, hence allowing a precise (up to atomic monolayer) thickness control, high density, low impurity level, excellent conformality and uniformity. This

deposition technique is not confined to only Hafnium- and Zirconium-based dielectrics, but also a wide range of high- κ materials such as Lanthanum, Cerium, etc. This makes the deposition of binary or ternary high- κ oxide achievable. The advantage of ALD offers a good prospective in the transistor downscaling and is becoming a major driving force for the high- κ integration.

Fig. 2.1: Illustration of atomic layer deposition of Al_2O_3 using Trimethylaluminum (TMA) precursor as chemical reactant and H_2O as oxidant. (a) TMA precursor is pulsed into the chamber and reacts with the OH groups on silicon surface, (b) a monolayer is formed and passivates the silicon surface, (d) unreacted TMA molecules and by-products are removed by N_2 purge, (e,f) H_2O is pulsed into the chamber and CH_3 groups are removed, passivating the surface with Al-OH , (g) unreacted H_2O molecules and by-products are removed by N_2 purge. This cycle is repeated until the target thickness is achieved [30].

2.2.2 Crystallization in high- κ film

Although ALD high- κ film shows a great potential as new replacement gate dielectric by virtue of its excellent conformity, the kinetic stability of the high- κ film

under high-thermal-budget front-end processes is a critical concern. Studies [24, 31] have shown that a thin HfSiO film crystallized after being subjected to source/drain dopant activation spike anneal at 1000 °C in N₂ for 10 s (Fig. 2.2). Compared with the as-deposited sample (Fig. 2.2(a)), the roughening of the top surface is clearly observed for the 1000 °C annealed HfSiO film (Fig. 2.2(b)). This observation is also borne out in the surface morphology of a thin 5 nm HfO₂ film (annealed at 900 °C) obtained by atomic force microscopy (AFM) as shown in Fig. 2.3(b). An obvious metastability is observed with the surface morphology exhibiting crack-like voiding. The high- κ film exhibits a surface roughness of 0.27 nm (rms) compared to that of the as-deposited film 0.17 nm (rms). This observation is consistent with the x-ray diffraction (XRD) analysis, which shows a clear appearance of monoclinic peaks after subjected to 800 °C and 900 °C anneal (Fig. 2.3(c)).

A polycrystalline high- κ film exhibits a great deal of non-uniformity in the distribution of trap density. The randomly oriented high- κ grains result in a severe atomic bonds breakage, thus presenting low resistive GBs that facilitate oxygen

Fig. 2.2: Cross-sectional transmission electron microscopy (TEM) of Si(100)/SiO₂/HfSiO gate stacks. The 20 nm HfSiO (30% SiO₂) films were prepared by using metal-organic chemical vapor deposition (MOCVD) technique. (a) as-deposited film; the high- κ layer is in amorphous structure, (b) film grown at 650 °C and followed by 1000 °C annealing in N₂ for 10 s, the high- κ layer has clearly crystallized [24].

diffusion through the layer. It is a great challenge to characterize the impact of GBs on the performance and reliability of a device by using standard characterization tools such as probe. The effect of GBs is typically masked out due to the area-averaged effect. Therefore, the application of STM in high- κ study provides an excellent alternative to complement the existing characterization techniques.

Fig. 2.3: Atomic force microscopy (AFM) images of 5 nm HfO_2 films deposited at 200 °C on Si substrate. (a) as-deposited film; the surface roughness is 0.17 nm (rms), (b) film annealed at 900 °C in N_2 ; the surface roughness is 0.27 nm (rms), (c) XRD scan of 5 nm HfO_2 films annealed at various temperatures. Monoclinic peaks are clearly observed after subjected to 800 °C and 900 °C anneal, indicating the presence of crystals in high- κ film [32].

2.2.3 Interface issues

Gate dielectric is liken to the “heart” of a transistor, the quality of this layer determines the overall performance of the device. The microstructural change in the high- κ “bulk” and an intermixing at the high- κ /Si interface under high-thermal-budget front-end processes are critical issues. As discussed in the previous section, high- κ film crystallization induces a high trap density distributed non-uniformly within the “bulk”. In this section, the high- κ /Si interface issue is reviewed as it plays a crucial role in determining a device’s performance, such as channel carrier mobility, TDDB,

V_{th} instability and EOT. The major interface issues identified are (i) large interface traps density near the Si substrate [22], (ii) inevitable IL re-growth [33-34], and (iii) inter-dielectric interaction at the high- κ /IL interface [34-35]. As discussed in Section 2.2.1, the high- κ dielectric is typically grown by external “deposition” (rather than “thermally” grown) such as PVD, CVD and ALD, this gives rise to a large interface trap density at the high- κ /Si interface. It was reported that the trap density at HfO_2 /Si interface is in the order of $\sim 10^{11}$ - $10^{12} \text{ cm}^{-2}\text{eV}^{-1}$ [10, 22], as compared with the SiO_2 /Si interface in the order of $\sim 10^{10} \text{ cm}^{-2}\text{eV}^{-1}$ [4]. Furthermore, the growth of thin sub-stoichiometric SiO_x IL at the high- κ /Si interface during deposition or PDA is often uncontrollable due to the diffusion (facilitated by GBs) of oxygen species, from high- κ “bulk” or the surrounding environment, interacting with the Si surface. Fig.2.4 depicts the TEM images of HfO_2 /SiO₂/Si gate stack subjected to 900 °C anneal in 1 bar O₂ for 10 s, which results in a substantial increase in the IL thickness from 1.2 nm to 3.6 nm. The presence of the thin IL limits the lowest achievable overall EOT scalability.

Fig. 2.4: Cross-sectional transmission electron microscopy (TEM) of atomic-layer-deposited 10 nm HfO_2 film on Si substrate. Prior to the HfO_2 deposition, the native oxide on Si surface was removed using diluted HF, subsequently a ~ 1 nm SiO_2 IL was grown. (a) as-deposited film; and (b) film subjected to 900 °C anneal in 1 bar O₂ for 10s. The IL thickness increases drastically and an obvious grain orientation is observed in the HfO_2 film, indicating crystallization [33].

Fig. 2.5: X-ray photoelectron spectroscopy (XPS) study of $\text{HfO}_2/\text{SiO}_2$ (~ 5 nm)/Si substrate annealed at 600 °C in N_2 for 30 s. The thicknesses of atomic-layer-deposited HfO_2 in sample S1, S2, and S3 are 0 nm, 0.5 nm, and 4 nm respectively. The presence of Hf-O-Si peaks in sample S2 and S3 in (a) O 1s, and (b) Si 2p core-level spectra indicates the formation of hafnium silicate at the HK-IL interface due to an interaction between the HfO_2 and SiO_2 IL [35].

Another critical issue is the intermixing of the high- κ film and SiO_x IL at the inter-dielectric high- κ /IL interface, giving rise to the formation of thin high- κ silicate layer as evidenced in the x-ray photoelectron spectroscopy (XPS) study [35]. Fig.2.5 shows the appearance of Hf-O-Si peaks in the XPS core-level spectra of the O 1s and Si 2p, indicating the presence of hafnium silicate at HfO_2 /IL interface. Both hafnium silicate and SiO_x IL have a relatively lower k value, thus resulting in a reduced effective k value of the entire gate stack. Moreover, the defects at this region may limit the carrier mobility/drivability of a device [36-38]. For a successful high- κ integration, interface engineering plays a crucial role to achieve a reliable small EOT device.

2.3 Electrical characteristics of metal/high- κ gate stack

2.3.1 Time dependent dielectric breakdown

Time-dependent dielectric breakdown (TDDB) testing, by subjecting n sample of devices to constant voltage stress (CVS) under an elevated temperature, is a standard

methodology to assess the reliability of a gate dielectric. The statistics of time-to-breakdown, T_{BD} , obtained from the TDDB test, form a Weibull failure distribution. T_{BD} is defined as the total time required for a sudden increase (by a few orders) in the gate current. The Weibull slope, β , of a failure distribution is an important factor in the reliability analysis for lifetime projection. According to the percolation BD model [39], β reflects the critical number of generated traps required to construct a conductive path through a dielectric layer in order to induce ultimate oxide BD. Percolation BD model is based on the percolation concept for modeling an oxide BD and is widely used to describe the intrinsic BD mechanism for SiO₂/SiON gate dielectric for the past decades. Under accelerated electric field, a continuous random generation of trapping centers in the “bulk” of an oxide induces gradual degradation and eventually triggers oxide BD when the accumulated number of trapping centers across an oxide, from the top gate electrode to the bottom substrate, reaches the critical level. Although the percolation BD model can consistently explain the TDDB results acquired from the SiO₂/SiON dielectric, it fails to resolve the TDDB results acquired from the high- κ dielectric, such as (i) a significantly lower Weibull slope compared with the SiO₂/SiON at a given physical thickness, (ii) Weibull slope shows bias polarity-dependence during stress, and (iii) bimodal Weibull slope in the area-scaled failure distribution.

Studies have reported a relatively low Weibull slope of the failure distribution for the devices employing Al₂O₃ [42] or HfO₂ [40] dielectric. Compared with the SiO₂ dielectric, the HfO₂ dielectric shows a weak thickness-dependence of Weibull slope. For instance, 5 nm and 9.5 nm SiO₂ yield Weibull slopes of ~ 2.2 and ~ 8.0 respectively, whereas the HfO₂ of similar physical thicknesses yield Weibull slopes of ~ 2.0 and ~ 4.0 as shown in Fig. 2.6(a). Several plausible explanations of the low Weibull slope

Fig. 2.6: (a) Measured Weibull slopes versus the physical thickness for SiO₂, Al₂O₃, and HfO₂ samples. At a given thickness, both Al₂O₃ and HfO₂ samples show a significantly lower Weibull slopes compared with SiO₂ sample [40]. (b) Weibull slope of Al₂O₃/SiO₂ gate stack failure distribution as a function of Al₂O₃ thickness. The Weibull slope shows a weak dependence on the high- κ thickness and annealing condition for the gate injection stress. For substrate injection stress, however, the value increases with the high- κ thickness but the effect pales after subjected to 700 °C/800 °C/900 °C annealing in N₂ [41].

have been proposed, i.e. (i) oxide BD depends on the defect density and size [43]. The critical number of generated traps, required to induce BD, is believed to be smaller for HfO₂ dielectric. In addition, the spacing between defects where tunneling of a trapped electron becomes probable is believed to be larger, (ii) ultrathin IL governs the gate stack BD [44], therefore the Weibull slope is small regardless of the high- κ thickness.

TDDDB study is further complicated by an unusual characteristic of the bias polarity-dependent Weibull slope. The failure distribution of Al₂O₃/SiO₂ gate stack with varied Al₂O₃ thickness is shown in Fig. 2.6(b) [41]. For gate injection mode, the Weibull slope is observed independent of the high- κ thickness even after subjected to different annealing conditions, suggesting that the IL controls the gate stack BD. For substrate injection mode, the Weibull slope of the as-deposited sample exhibits a strong dependence on the high- κ thickness (i.e. the Weibull slope increases with thickness), suggesting that the high- κ layer controls the gate stack BD. This effect,

Fig. 2.7: (a) Failure distribution of the Hf-based high- κ /metal gate devices based on vertical Poisson area scaling to a reference area of $1 \mu\text{m}^2$. The TDDDB tests were carried out in inversion mode at 125°C . The devices area ranges from $0.006 \mu\text{m}^2$ to $50 \mu\text{m}^2$ with a sample size of 150 devices per distribution. Two Weibull slopes are observed in the distribution. The upper percentile ($\beta \sim 0.8$) primarily consists of the T_{BD} data acquired from the small area devices whereas the lower percentile ($\beta \sim 1.5$) consists of the T_{BD} data acquired from the larger area devices [28]. (b) Area-scaled TDDDB distribution of Al_2O_3 ($\sim 9 \text{ nm}$)/metal gate devices subjected to constant voltage stressing at -6.8 V . Only a single Weibull slope is observed for the relatively large area devices [42].

however, is reduced after subjected to $700^\circ\text{C}/800^\circ\text{C}/900^\circ\text{C}$ annealing in N_2 . Degraeve *et al* [45] attributed the polarity-dependent Weibull slope to an asymmetry degradation of the high- κ gate stack.

Fig. 2.7(a) shows an area-scaled Weibull distribution obtained from the high- κ devices with various area sizes, ranging from $0.006 \mu\text{m}^2$ to $0.6 \mu\text{m}^2$ (A_x). The distribution is normalized to $1 \mu\text{m}^2$ (A_{norm}) by multiplying the values with $\ln(A_x/A_{\text{norm}})$. Two distinct Weibull slopes are observed; the lower percentile of the failure distribution exhibits a steep Weibull slope $\beta \sim 1.5$ whereas the upper percentile of the distribution exhibits a shallow Weibull slope $\beta \sim 0.8$. The regions at the upper and lower percentile of the distribution are primarily contributed by the T_{BD} data collected from the small and large area devices respectively. However, it is interesting to note that the normalized BD distributions for the devices with a relatively large area sizes,

ranging from $100 \mu\text{m}^2$ to $100000 \mu\text{m}^2$, merge to a single line and exhibit only a single Weibull slope [40, 42] as shown in Fig. 2.7(b). The bimodal Weibull characteristic has been reported by several research groups [27-28, 46-47] and yet no conclusive explanation to clarify the origin of this phenomenon. Nigam *et al* [47], via kinetic Monte Carlo (kMC) simulations based on a dual layer percolation model on a three-dimensional cubic grid, attributed the phenomenon to a different defect generation rate of the respective SiO_x IL and high- κ layer. As the high- κ defect generation rate is far higher (~ 3000 times) than that in the SiO_x IL, the TDDDB distribution shows bimodal characteristic (Fig. 2.8 (a)). These simulation results, however, were based on the assumptions that (i) the high- κ defect density for a small area device is very large and this layer degrades at a higher rate, thus the TDDDB measurement reflects only IL BD (thereby a small Weibull slope is observed), (ii) for a large area device with a low initial high- κ defect density ($< 0.3\%$), the Weibull slope corresponds to the entire gate stack thickness, hence a steep Weibull slope is obtained. Ribes *et al* [27] proposed that the T_{BD} of a high- κ gate stack is given by a convolution of the time to break the IL and the time to extend the BD path in the high- κ layer at the same location. The correlated high- κ BD is localized above the IL BD path; its localization does not correspond to the high- κ weakest point, thus it does not follow Poisson area scaling and gives rise to the bimodal characteristic. Raghavan *et al* [48] proposed a modified oxide BD percolation model, accounting the presence of GBs in a polycrystalline high- κ stack, to study the non-Weibull distribution. The author suggested that the failure distribution is related to the presence of GBs (Fig. 2.8 (b)) but does not clarify the mechanism involved.

Various explanations of the high- κ gate stack BD mechanism have been proposed based on the characteristics observed in the TDDDB distributions. However, there are

many unresolved ambiguities due to the following reasons; (i) bi-layer high- κ /IL structure makes the current reliability studies complicated; the high- κ degradation physics is not well understood and the impact of the high- κ quality on the IL degradation is unclear, (ii) high as-grown defects density in the high- κ “bulk” and interface result in inconsistent electrical measurements. (iii) limitations of the standard electrical characterization method. Nanoscale characterization, via STM, therefore is crucial since it offers a great advantage in resolving the microscopic information of the intrinsic BD mechanism. This will complement the standard electrical characterization method.

Fig. 2.8: (a) Kinetic Monte Carlo simulated TDDB distributions of a high- κ /IL gate stack with various high- κ defect generation rates. As the defect generation rate is uniform for both high- κ and IL, a standard Weibull distribution with a single Weibull slope $\beta \sim 1.1$ is obtained. However, two Weibull slopes are observed in the TDDB distribution when the high- κ defect generation rate is increased to 3000 times compared with that in the IL [47]. (b) Simulation of Weibull plots based on a modified analytical cell-based percolation model, which shows the separate effect of grain boundary (GB), grain (G), and the interaction between G and GB (G-GB) in a $N \times N$ cells. The failure at pure G sites shows a straight line, while the GB sites shows a convex trend. The G-GB, on the other hand, shows a bimodal Weibull trend with the lower percentile follows the GB failure whereas the higher percentile follows the G failure [48].

2.3.2 Stress-induced leakage current

Under high electric field stressing, traps are randomly generated in an oxide “bulk”. As the stress is removed, an increase (due to the generated traps that facilitate electrons tunneling) in the gate leakage current is observed at the low to moderate oxide electric field. This phenomenon is commonly known as stress-induced leakage current (SILC). The trap generation is related to hydrogen release by the hot electrons. The main driving force for the replacement of thin SiO_2/SiON with high- κ dielectric is the significant suppression of the gate leakage current [17, 50]. However, SILC may impose a reliability issue as it compromises the high- κ low leakage benefit. Similar to the TDDB results discussed in the previous section, reports have shown that SILC exhibits bias polarity-dependent characteristic [49]. However, the origin of the bias polarity-dependent characteristic of SILC remains unclear. The bi-layer high- κ /IL structure further complicates the reliability study as the correlation of SILC and the traps generated in the high- κ and SiO_x IL is ambiguous.

Fig. 2.9: (a) Defect density in HfO_2 “bulk” as a function of injected electron fluence at positive and negative stress polarity. The polarity-dependence of trap generation suggests that SILC conduction mechanism is related to the HfO_2 traps. (b) At the low trap density region, SILC shows a proportional one-to-one relationship with the HfO_2 “bulk” trap density. However, it increases rapidly at the high trap density region [49].

Crupi *et al* [49], via variable frequency charge pumping measurement on $\text{HfO}_2/\text{SiO}_2$ gate stack, suggested that the HfO_2 “bulk” trap density is directly related to SILC. In the experiment, the HfO_2 trap density as a function of injected electron fluence was carried out at both positive and negative stress polarity. At positive stress, the trap density increases rapidly, suggesting a direct correlation between SILC and HfO_2 “bulk” traps (Fig. 2.9(a)). At negative stress, on the other hand, the trap density exhibits a slower increase. Fig. 2.9(b) shows that as a gate stack is near to BD (i.e. corresponds to high trap density region), the SILC increases rapidly. Degraeve *et al* [45] conducted a charge trapping study on the $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stack by using quasi-static capacitance-voltage measurement. During gate injection stress, a rapid build up of interface states is observed in the CV curves depicted in Fig. 2.10(a), suggesting a fast degradation of SiO_2 IL. As the IL BD, a large voltage drop across the overlaying high- κ layer results in a rapid high- κ BD, therefore a low Weibull slope is obtained. During substrate injection stress, on the other hand, no obvious observation of interface state build-up (Fig. 2.10(b)). An asymmetry of the degradation is attributed to an asymmetry of the high- κ / SiO_2 structure as depicted in Fig. 2.11. A significant flatband voltage, V_{FB} , shift suggests a large electron trapping in Al_2O_3 . It is believed that the Al_2O_3 layer controls the gate stack reliability, as evidenced by the high- κ thickness-dependent of Weibull slope.

Study has shown that SILC exhibits a strong correlation to the evolution of the gate leakage current degradation features such as progressive BD and hard BD [51]. Therefore, for a bi-layer high- κ gate stack, the layer which induces the SILC growth is the major contributor to oxide degradation and determines the entire gate stack ultimate BD. Gennadi *et al* [44] suggested that IL is the “weak-link” which plays the crucial role for SILC growth in high- κ /IL gate stack. Fig. 2.12 shows an excellent

correlation between the stress time dependence of N_{it} and the stress time dependence of low voltage SILC measured under the same stress sequence. The interface trap density (ΔN_{it}) measured by the low frequency charge pumping corresponds to the traps within IL. Therefore, this correlation suggests that SILC is controlled by the trap generation in the IL, primarily near its interface with the high- κ .

Fig. 2.10: (a) Quasi-static capacitance-voltage characteristic of $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stack as a function of gate voltage stress for 1000 s under gate injection mode. An obvious build up of interface states is observed, (b) under substrate injection mode, no interface states generation is observed. An obvious flatband voltage, V_{FB} , shift indicates electron trapping in Al_2O_3 layer [45].

Fig. 2.11: Schematic band diagram of $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stack during (a) gate injection stress. Holes mainly trapped near the SiO_2/Si interface and IL is believed to be the layer controlling the gate stack BD, (b) substrate injection stress. Traps creation occurs in the Al_2O_3 layer. Therefore, high- κ controls the gate stack reliability [45].

Fig. 2.12: Correlation of the low voltage SILC (measured at $V_g = 0.5$ V) and the charge pumping-measured (at low frequency of 1 kHz) N_{it} subjected to various gate stress voltages for SiO_2 (1.1 nm)/ HfO_2 (3 nm)/TiN nFET. A good correlation of the data indicates that the traps within IL control the SILC [44].

A physically thicker high- κ film can effectively suppress the gate current density at least three orders magnitude lower than SiO_2 of similar EOT. However, as the high- κ thickness is further scaled down in order to meet the stringent requirement of IRTS roadmap, SILC becomes a critical concern. Fig. 2.13 shows a rapid increase (indicated by a steeper slope in the curve) in the gate leakage current for the HfO_2 film as compared with SiO_2 [52]. It is interesting to note that at the region less than 1 nm, an extrapolated line of the HfO_2 data is expected to intersect the SiO_2 data at a certain point, suggesting that the gate leakage current through the HfO_2 film is now larger than that through the SiO_2 . Therefore, the advantage of high- κ is increasingly less prominent when the dielectric thickness is in sub-1 nm [53-54]. In addition, an ultrathin HfO_2 film is reported to have a smaller effective dielectric constant ($K_{\text{eff}} \sim 10$ -12) compared to that in the “bulk” ($K_{\text{eff}} \sim 25$ -30). This is attributed to a lower film density (8.5 g/cm^3) and refractive index (~ 1.97) compared to that in the “bulk” (9.68 g/cm^3 , ~ 2.01) [52]. The reliability of ultrathin high- κ film is important and extensive

efforts to improve the film quality are necessary in order to achieve sub-1 nm EOT scaling.

Fig. 2.13: A comparison of absolute gate leakage current (at accumulation bias, $V_g = -1$ V) versus EOT for the HfO₂ and SiO₂ dielectrics. SiO₂ leakage data is adapted from [55]. The thin HfO₂ dielectric exhibits a steeper slope in the curve compared to that of the SiO₂, indicating a rapid increase in the gate leakage current and may even exceed SiO₂ at a low EOT (< 1 nm). The inset shows an absolute leakage current evolution under constant voltage stressing [52].

In summary, the present studies of the SILC characteristics observed in the high- κ gate stacks are reviewed. Unlike the standard metal-insulator-semiconductor (MIS) structure which contains a mere single SiO₂/SiON dielectric layer, the present high- κ gate stack typically consists of multilayer (due to the presence of IL) structure with an intermixing between high- κ and IL at the inter-dielectric interface, this complicates the reliability analysis as of which layer is the major contributor to SILC growth and how is the IL degradation correlates to the high- κ film quality. Therefore, an understanding on the physical mechanism is essential to clarify these issues.

2.3.3 Threshold voltage instability

Threshold voltage, V_{th} , instability is one of the critical reliability issues opposing the successful integration of high- κ dielectric in the CMOS process [2]. The V_{th} shift is closely related to an oxide charge trapping and detrapping. Three identified origins of V_{th} shift are: (i) the “bulk” and interface charge density, (ii) the fermi level pinning at the high- κ /polysilicon interface [56], and (iii) the dipole moment at the high- κ /IL interface [57]. It was shown that Hf-based gate oxide exhibits an efficient trapping due to a large charge trap density in the oxide “bulk” and at the interface [58]. For a bi-layer high- κ /IL gate stack, the V_{FB} can be derived from Poisson’s equation, in terms of the EOT value, t , as shown in equation (2-1) below [2]

$$V_{FB} = \Phi_{ms} + \frac{Q_{it}}{\epsilon_0 K_{SiO_2}} t + \frac{Q_b K_{hiK}}{2\epsilon_0 K_{SiO_2}} t^2, \quad (2-1)$$

where Φ_{ms} is the work function difference between the Si and the gate electrode, Q_{it} is the high- κ /SiO₂ interface fixed charge density, Q_b is the high- κ “bulk” fixed charge density, K_{SiO_2} and K_{hiK} are the dielectric constants of the SiO₂ and high- κ respectively. From equation (2-1), it is clearly shown that a large Q_b or Q_{it} will result in a severe V_{FB} instability for a device. Dynamic NBTI studies by Teo *et al* [59] and Gao *et al* [60] showed an evident progressive decrease in the recoverable component for the HfO₂ p-MOSFET as compared with SiON p-MOSFET (Fig. 2.14). The decrease is attributed to the conversion of partial recoverable component into a permanent form, hence giving rise to V_{th} instability.

Fig. 2.15 shows a comparison of the transient I_d - V_g hysteresis curves, measured using trapezoidal V_g pulses ($t_r = t_f = 1 \mu s$; $\Delta t = 2 \mu s$), for the HfO₂ and SiO₂ samples at a comparable EOT. Both HfO₂ NMOS and PMOS devices show very pronounced

hysteresis curves, with 0.15 V and 0.25 V shift respectively. In contrast, the nitrided SiO₂ reference device does not present any voltage shift. The authors attributed the V_{th} shift in the NMOS and PMOS devices to the presence of negative charges (due to electrons trappings) and positive charges (due to hole trappings) in the oxide respectively [61]. However, it is unclear whether the electron/hole trappings are

Fig. 2.14: Evolution of threshold voltage shift $|\Delta V_t|$ of the (a) SiON [59] and (b) HfO₂ [60] p-MOSFETs under repeated stress/relax cycling (each stress/relax cycling is 1×10^3 s) under oxide fields (~ 7 MV/cm) at 100 °C. Compared with SiON, HfO₂ sample exhibits an obvious increase in the $|\Delta V_t|^{eos}$ and $|\Delta V_t|^{eor}$. The inset in (a) shows the definition of $|\Delta V_t|^{eos}$ and $|\Delta V_t|^{eor}$ at the end of the stress and recovery cycle, respectively.

Fig. 2.15: Trapezoidal V_g pulse measurement ($t_r = t_f = 1 \mu s$; $\Delta t = 2 \mu s$) of P- and N-channel devices employing ~ 14 Å HfO₂ gate dielectric. A pronounced transient I_d - V_g hysteresis is observed for both HfO₂ NMOS and PMOS as compared with the nitrided SiO₂ reference device at a comparable EOT [61].

confined to one type of trap (i.e. similar to the P_{bo} center) or two different types of traps in an oxide. Improving the quality of an oxide film by reducing the “bulk” and interface traps density, therefore, is crucial to ensure a reliable performance of the high- κ device.

2.4 Physical failure mechanism

Degraeve *et al* [39] suggested that both anode hole injection model and electron trap density model are correlated to each other, therefore, the author proposed a consistent oxide BD model based on the percolation theory to clarify the intrinsic TDDDB mechanism in SiO_2 dielectric. According to the percolation BD model, traps are generated randomly in an oxide under electric field. As the density of the accumulated generated traps reaches a critical value such that the traps form a bridge connecting the top gate electrode and the bottom Si substrate as illustrated in Fig. 2.16(a), electron/hole carriers may tunnel through the dielectric via this conduction path, thereby giving rise to a large gate leakage current, which is known as “oxide BD”. Tremendous efforts to understand the “trap” generated in an oxide and its correlation to the percolation path have been made and most research groups attribute the “trap” to oxygen vacancy. Li *et al* [62], via TEM and electron energy loss spectroscopy (EELS) analysis on a BD spot in SiO_2 , showed that the center of a percolation path has the highest percentage of oxygen deficiency (Fig. 2.16 (b)) and the deficiency of oxygen spreads out radially from the center to its surrounding area. The Si–O composition changes from $SiO_{1.76}$ to $SiO_{0.7}$ as moving to the center of the percolation path. A large local current density at the center of a percolation path, therefore, is believed to be associated with the increasing oxygen vacancies, which serve as the traps responsible for the carrier trapping and tunneling through an oxide.

Most studies have attributed the high- κ degradation and BD to oxygen vacancies [63].

By using first principles calculations, McKenna *et al* [64] reported a strong oxygen

Fig. 2.16: (a) Schematic illustration of trap (represented by the sphere) generation in the SiO₂ layer under applied electric field. As the number of generated traps reaches a critical value and forms a bridge connecting the gate and Si substrate, electrons/holes carriers may tunnel through the layer and induce intrinsic oxide BD. The BD path is depicted by the shaded spheres [39]. (b) The corrected oxygen deficiency in the percolation path based on the as-measured O *K* edge intensities near the BD spot. The local O deficiency in the center of the percolation path is relatively high. The diameter of the percolation (defective oxide area laterally) is estimated to be 30 nm for the post-BD leakage current of 2 μ A [62].

Fig. 2.17: Schematic illustration of the proposed IL BD model for a bi-layer high- κ /IL gate stack. (a) GB facilitates the formation of oxygen-deficient filament in the IL region adjacent to the overlaying high- κ film. Shaded area indicates a high density of pre-existing V_o 's in IL. (b) Energy band diagram of the shaded area in IL, the oxygen deficient filament propagates toward the Si substrate accompanied by generation of additional V_o states (dotted lines). The waving line indicates trap relaxation due to an electron capture [65].

vacancy segregation and diffusion near a GB, which makes this location a preferential site for percolation path formation. The author found that the diffusion of oxygen vacancy along a GB is more rapid than in a “bulk” due to a high barrier for oxygen vacancy to climb out. Based on these results, Gennadi *et al* has proposed an IL BD model correlating the high- κ GB to the IL degradation for a bi-layer high- κ stack [65]. It is shown that the IL region adjacent to the high- κ film has a higher density of process-induced oxygen vacancies [66]. Under voltage stress, the low resistive high- κ GB results in a high electric field across the adjacent IL and hence inducing a rapid traps generation at this location (Fig. 2.17). An oxygen-deficient filament is formed at this region and subsequently initiates the formation of a BD path via propagation towards the substrate.

2.5 Post deposition annealing

The overall performance of a metal-oxide-semiconductor (MOS) transistor relies essentially on the performance and reliability of an ultrathin gate dielectric film. In the conventional CMOS processes, the high quality SiO₂ dielectric is grown by dry rapid thermal oxidation (RTO) at a relatively high temperature followed by RTA at temperature > 600 °C in N₂ ambient. This results in a reduction in the gate leakage current density, trapped charge density, interface state density, hot-carrier degradation and improved carrier mobility [4, 67]. Post deposition annealing (PDA) is an important step to improve the pre-existing defects in an oxide. Studies [68-69] have reported a reduction in the gate current density for a high- κ stack subjected to PDA. However, due to the thermodynamic instability of the high- κ dielectric in contact with Si, a high temperature PDA (> 600 °C) results in an uncontrollable increase in the IL thickness (thereby a larger EOT), formation of high- κ silicate layer at the inter-

dielectric interface, and phase change in the high- κ “bulk”. In order to improve the high- κ film quality while meeting the stringent requirement of the EOT scaling, extensive research efforts were carried out to explore a new treatment technique. This section will review the PDA techniques reported in the literature.

2.5.1 Oxygen annealing

The stability of high- κ stack against the PDA in O_2 -containing ambient is of high interest for a few reasons: (i) this step is performed intentionally to improve the oxygen vacancies in the high- κ “bulk” in order to suppress the gate current density [70] and oxide charge trapping; (ii) the high- κ stack is unintentionally exposed to the O_2 residual in the process chamber or furnace. Lee *et al* [70] investigated the O_2 annealing effect on the thermal stability and the leakage current of an ultrathin ~ 4.5 nm HfO_2 film. Although a short (less than 60 s) RTA in O_2 results in an evident reduction in the gate current density, the EOT is compromised by an increase in the IL thickness (Fig. 2.18(a)). Oxygen diffusivity in the high- κ film is the key mechanism that affects the IL re-growth with consequences on the control of the EOT [33]. Studies [33, 71] have reported a rapid IL re-growth, from 1.2 nm to 3.6 nm, for a thick 10 nm high- κ stack subjected to an annealing at 900 °C in 1 bar O_2 for 10 s. Ferrari *et al* [33] examined the oxygen diffusivity level in high- κ by subjecting the HfO_2 film to $^{18}O_2$ annealing.

Fig. 2.18(b) depicts the profiles of the $^{16}O_2$ and $^{18}O_2$ signals for the as-deposited and 10 s annealed (in $^{18}O_2$ atmosphere) HfO_2 films. Observation shows that a significant fraction (~ 50 %) of $^{16}O_2$ was exchanged with $^{18}O_2$ after annealing, pointing to a high oxygen diffusivity in high- κ film. In addition, the presence of GBs in a crystallized high- κ film further complicates the IL re-growth control. It is believed that

GBs facilitate the diffusion of oxygen through the high- κ and consequently inducing a rapid oxidation at the Si surface.

Fig. 2.18: (a) Equivalent oxide thickness (EOT) and gate current density, J_g , as a function of RTA time for annealing in O₂ or N₂ ambient at 700 °C. A 4.5 nm HfO₂ film was deposited on *p*-Si (100) substrate to form MOS capacitors. For the sample annealed in O₂ ambient, the J_g is reduced at the penalty of EOT increase [70] (b) Time-of-flight-secondary-ion-mass spectroscopy (TOF-SIMS) analysis of ¹⁶O₂ and ¹⁸O₂ signals for the as-deposited and 10 s annealed (at 900 °C in 0.1 bar ¹⁸O₂ atmosphere) HfO₂ films. A large fraction of ¹⁶O₂ is exchanged with ¹⁸O₂, suggesting that the oxygen diffusivity in HfO₂ is extremely high [33].

2.5.2 UV ozone annealing

Although PDA in O₂ or N₂ shows promising results in improving the high- κ dielectric performance, a relatively high annealing temperature imposes a minimum limitation for achievable EOT due to an inevitable increase in the IL thickness. Therefore, these techniques are not too favorable for an ultrathin (sub-1 nm) high- κ scaling. Moreover, the benefit of low leakage current in the high- κ dielectric is severely compromised due to crystallization. Overcoming the high-thermal-budget issue has received considerable concern, which leads to a switching from the gate-first integration scheme to the gate-last integration scheme in order for the high- κ deposition to be carried out after (or bypass) the source/drain dopant activation spike

anneal (~ 1000 °C) step. It is known that as-deposited high- κ film is usually loosely packed and contains a high density of as-grown defects such as oxygen vacancies, oxygen interstitials and carbon-related impurities [72], therefore a post deposition treatment at low temperature is necessary to improve the film quality. In recent years, some research groups [73-75] have demonstrated improved electrical performance for the high- κ device subjected to a low temperature annealing in O_2 under ultraviolet illumination (i.e. UVO annealing). O_2 , when exposed to the UV illumination, will produce active oxygen radicals and ozone [76]. It is believed that the oxygen radicals diffuse into the region near the high- κ surface and passivate the defects. Tang *et al* [73] reported an evident improvement in the electrical performance of an organic thin film transistors (OTFT) employing low temperature UVO-treated HfO_2 gate dielectric. Compared with the as-deposited (i.e. without UVO treatment) sample, a larger drain current, 10 times higher mobility, smaller sub-threshold slope and 3 times larger on/off ratio are observed. Zhao *et al* [74] attributed the suppression of leakage current and moisture absorption of a 4 nm La_2O_3 film (treated with UVO anneal at room temperature for 9 min) to the healing of oxygen vacancies in the film, since the suppression effect is similar to the O_2 ambient annealing. The surface roughness of the as-deposited La_2O_3 film, which typically increases with time in the air, is clearly suppressed for the UVO treated sample. Fig. 2.19 shows a comparison of the gate current density and $C-V$ characteristics between the standard (i.e. as-deposited), O_2 -treated and UVO-treated 13 nm HfO_2 films [75]. Compared with the standard sample, the UVO treated sample exhibits two orders of magnitude reduction in the gate current density and the smallest $C-V$ hysteresis (~ 0.2 V), indicating a lower trap density in the film (Fig. 2.19). In addition, the former shows the best EOT improvement (8.51 nm) as compared with the standard (12.01 nm) and the O_2 treated (10.16 nm) samples. A

Fig. 2.19: (a) Gate current density versus gate voltage characteristics of Al/HfO₂ (~13 nm)/Si MIS structures subjected to O₂ treatment (O₂ ambient at 150 °C under 1000 psi for 2 hours) and UVO treatment (exactly similar condition as O₂ treatment, but under UV light illumination). The standard (i.e. as-deposited) sample receives no treatment. (b) 1 MHz high frequency $C-V$ characteristics. The O₂ and UVO treated samples exhibit smaller $C-V$ hysteresis (~0.7 V and ~0.2 V respectively) as compared with the standard sample (~1 V). In addition, a positive V_{FB} shift to indicates a lesser positive charge trapping in the UVO-treated sample [75].

positive V_{FB} implies a significant suppression of positive charge trapping for the UVO-treated sample (Fig. 2.19(b)). Low temperature UVO annealing can only modify the quality near the film top surface region [77], therefore, efforts need to be made in order to remove the remaining defects located deep inside the “bulk”.

2.5.3 Multi-deposition and annealing

In order to reduce the high as-grown defects density inherent in the high- κ “bulk”, multi-deposition cum annealing technique have been proposed by some research groups [78-79]. Nabatame *et al* [78] demonstrated Layer-by-Layer Deposition and Annealing (LL-D&A) technique in reducing the residual impurities incorporated in the high- κ film. HfAlO_x/SiON(0.4 nm)/Si gate stacks with EOT ~1.38 nm were grown via LL-D&A at 650 °C in vacuum (0.1 mPa), O₂ ambient (130 Pa), and NH₃ ambient (130 Pa) respectively. It is shown that the residual carbon contents in the HfAlO_x film is

reduced by 50% for LL-D&A(O₂) and LL-D&A(NH₃) samples as compared with the standard single-step deposited sample, as shown in the secondary-ion-mass-spectroscopy (SIMS) profile (Fig. 2.20(a)). No obvious reduction in the gate leakage current is observed for both LL-D&A(O₂) and LL-D&A(NH₃) samples, but excellent electrical results such as small δV_{FB} (< 0.06 V), low subthreshold swing (~ 77 mV/dec), and improved peak mobility (~ 210 cm²/V_s) were reported. X-ray photoelectron spectroscopy (XPS) core level spectra of Si 2p (Fig. 2.20(b)) reveals an increase in the IL thickness for LL-D&A(O₂) sample due to a relatively high annealing temperature involved. Conley *et al* [79-80] showed that a film densification can be increased by removing the residual impurities (e.g. C, Cl, etc) in the “bulk” via a moderate annealing temperature (~ 420 °C in N₂ for 30 s) in multi deposition step. The film density can reach as high as ~ 10 g/cm³ compared with the standard sample (~ 7.7 g/cm³) which receives only single PDA.

Fig. 2.20: (a) Carbon concentration of HfAlO_x films grown by LL-D&A technique in various ambiances. A significant reduction in the carbon content is observed for LL-D&A(NH₃) and LL-D&A(O₂) samples. (b) XPS core level Si 2p spectra of HfAlO_x films grown by LL-D&A technique. LL-D&A(O₂) and LL-D&A(vac) samples exhibit an increase in the IL thickness, which is attributed to the diffusion of oxygen [78].

2.6 Characterization by scanning tunneling microscopy

Scanning probe microscopy (SPM), a technique in which a sharp tip is scanned across a sample to obtain the surface image at nanometer resolution, has evolved as a powerful way for surface characterization in many disciplines such as in the condensed-matter physics, chemistry, and biology. The earlier SPM instrument was scanning tunneling microscopy (STM), which was invented to image the surface topography at atomic resolution. By using a sharp tip, a lateral resolution in the order of 0.1 nm and vertical resolution in the order of 0.01 nm can be achieved. The basic principle of STM operation is based on the quantum mechanical electron tunneling in a given ambience (e.g. air, vacuum, etc). In the classical physics, an electron with energy E cannot penetrate into any region with potential barrier V larger than E . In the quantum physics, however, an electron has a non-zero probability of tunneling through a potential barrier even though $E < V$, as illustrated in Fig. 2.21.

Fig. 2.21: In quantum mechanics, an electron has a non-zero probability of tunneling through a potential barrier V even if the electron energy E is lesser than V (denoted by region II). The energy of the tunneled particle is identical except for a decrease in the amplitude.

In quantum physics, the state of an electron is described by a wave function $\psi(z)$. By solving Schrodinger's equation in equation (2-2), one would obtain equation (2-3)

and equation (2-4) for the classically allowed region ($E > V$) and forbidden region ($E < V$) respectively. The k in equation (2-3) is the wave vector whereas the k in equation (2-4) is the decaying constant.

$$-\frac{\hbar^2}{2m} \frac{d^2}{dz^2} \psi(z) + U(z)\psi(z) = E\psi(z). \quad (2-2)$$

$$\psi(z) = \psi(0)e^{\pm ikz}, \text{ in which } k = \frac{\sqrt{2m(E - U)}}{\hbar} \quad (2-3)$$

$$\psi(z) = \psi(0)e^{-\kappa z}, \text{ in which } \kappa = \frac{\sqrt{2m(U - E)}}{\hbar} \quad (2-4)$$

The probability of observing an electron near point z is proportional to $|\psi(0)|^2 e^{-2\kappa z}$, which has a non-zero value in the barrier region, thus a non-zero probability to penetrate a barrier. Considering the metal-vacuum-metal tunneling system in Fig. 2.22, the work function of a metal surface is defined as the minimum energy required to remove an electron from the “bulk” to the vacuum level. The fermi level is the upper limit of the occupied states in a metal. Taking the vacuum level as the reference point of energy, $E_F = -\phi$, and assuming that the work functions of the tip and the sample are equal, a sample state ψ_n with energy level E_n lying in between $E_F - eV$ and E_F has a chance to tunnel into the tip and vice versa. However, without a bias voltage, there is no net tunneling current. A bias voltage is required for tunneling to occur.

In STM experiment, when a tip is scanned across a sample surface, the condition of the tip usually does not vary, hence the electrons coming to the tip surface, $z = W$, have a constant velocity to flow into the tip. The tunneling current is directly proportional to the number of states on the sample surface within the energy interval

Fig. 2.22: One-dimensional metal-vacuum-metal tunneling junction model.

eV , which are responsible for the tunneling. By including all the sample states in the energy interval eV , the tunneling current can be written in terms of the local density of states (LDOS) as in equation (2-5). The typical value of work function is $\Phi \sim 4\text{eV}$, which gives a typical value of the decay constant $\kappa \sim 1 \text{ \AA}^{-1}$. From equation (2-5), the current decays about $e^2 \sim 7.4$ times per $\text{\AA}$. The dependence of the logarithm of the tunneling current with respect to the distance is a measure of the work function, or the tunneling barrier height. Therefore, it is clear that the tunneling current is extremely sensitive to the variance of the tip-sample distance, making STM a powerful technique for surface characterization.

$$\begin{aligned}
 I &\propto V \rho_S(0, E_F) e^{-2\kappa W} \\
 &\approx V \rho_S(0, E_F) e^{-1.025\sqrt{\phi} W}.
 \end{aligned}
 \tag{2-5}$$

In recent years, many researchers have started in exploring the application of SPM in the dielectric reliability study for CMOS application. As the dimensions of a transistor reach the nanometer regime, a variation in the electronic structure of a dielectric film affects the entire electrical characteristic of a device significantly. An

inhomogeneity in the high- κ electrical and physical properties give rise to inconsistent observations of a device performance [81-82], hence complicating the interpretation and analysis of electrical data acquired using the standard metrology tools. The powerful feature of STM appears to be an attractive technique which offers the advantage in resolving the local electronic structure of an ultrathin dielectric film at atomic scale. Moreover, the capability of STM to scan under various ambiances with virtually no damage or interference to the sample make it possible to observe the *real-time* oxide degradation process.

Studies of ultrathin SiO₂ reliability by utilizing SPM (e.g. STM, AFM, CAFM, etc) have been demonstrated by some research groups [83-89] and interesting results have been reported. Murrell *et al* [84] demonstrated the potential to spatially (with a lateral resolution better than 20 nm) probe the dielectric properties of SiO₂ using atomic force microscopy (AFM) with conducting cantilever (silicon cantilever coated with 100 nm titanium to ensure good electrical conductivity). The findings suggest that the BD strength of a MOS capacitor is not limited by the intrinsic dielectric strength alone, but by the imperfections or non-uniformities in the Si/SiO₂ structure. Studies of SiO₂ dielectric BD by using conductive atomic force microscopy (CAFM) were demonstrated by several other groups such as Yasue *et al* [90], Lanza *et al* [82], Fiorenza *et al* [91], Porti *et al* [92], etc. O'Shea *et al* [83] reported that the BD of SiO₂ over a small area ($\sim 500 \text{ nm}^2$) follows that observed in the conventional MOS device (with a maximum BD field of 13 MV/cm). Visualization of SiO₂ BD evolution from the pre-BD stage to the post-BD stage was demonstrated by Zhang *et al* [88] using CAFM in vacuum ($\sim 5 \text{ mPa}$). Watanabe *et al* [85] showed that the STM images can reveal the individual leakage sites and surface topography (e.g. atomic steps) of SiO₂ film (Fig. 2.23). Other research groups that apply STM for SiO₂ characterization

include Daniel *et al* [93], Fujita *et al* [86], Johnson *et al* [94], Wei *et al* [95], etc. Although STM is typically used to characterize the conductive samples, however, the fact that electrons could tunnel through an ultrathin (i.e. a few nanometer thickness) dielectric allows the probing of the electronic structure of local defects by STM.

Fig. 2.23: (a) $350 \times 350 \text{ nm}^2$ STM image of a thin SiO_2 dielectric. The area stressed with a high bias voltage of +10 V and 6 nA is indicated by the dotted line. (b) and (c) are the zoom-in high-resolution images of the corresponding area before and after degradation, respectively. The surface corrugation in (c) was clearly increased after modified by the hot-electron injection [85].

2.6.1 STM/CAFM study on ultrathin high- κ dielectric

The application of SPM in ultrathin oxide study is a significant breakthrough. However, the earlier reports of SiO_2 characterization results are more to a demonstration of the SPM capability in acquiring the local electronic structure at an atomic scale (rather than solving complications since SiO_2 characteristics are well understood at macroscopic level) until the transition of SiO_2 to high- κ dielectric. The non-area/thickness scalability in the electrical characteristics (e.g. I - V , C - V , TDDB, etc) of the high- κ film implies that an oxide degradation is not limited to the intrinsic wear-out alone, but also the non-uniformity and the existence of localized leakage sites [81]. Inconsistency in the interpretations of the high- κ devices electrical measurement data

is attributed to overlooked microscopic information in the high- κ oxide due to an area-averaged effect. Therefore, SPM characterization technique is increasingly important and attractive as more research groups are exploring this technique as an alternative way to access the microscopic information in order to resolve the ambiguities observed in the electrical measurements. Studies of the various high- κ metal oxides (e.g. HfO_2 [96-98], La_2O_3 [74, 99], ZrO_2 [100], CeO_2 [101], *etc*), via SPM, have been published in recent years.

Fig. 2.24: (a) STM measurement of $\text{Sc}_2\text{O}_3/\text{La}_2\text{O}_3/\text{SiO}_x/\text{Si}$ gate stack at setpoint of -4 V, 20 pA. (I) depicts the topographical image of the gate stack whereas (II), (III), and (IV) depict the corresponding current maps acquired at -2.3 V, $+2.3$ V, and $+4$ V respectively. (b) STM measurement of the same area at setpoint of $+3.5$ V, 20 pA [104].

Fig. 2.25: I - V characteristics extracted at various leakage sites (labeled as A, B, C, and D) across the $\text{Sc}_2\text{O}_3/\text{La}_2\text{O}_3/\text{SiO}_x/\text{Si}$ current maps in Fig. 2.24. [104].

Fig. 2.26: STM topographical images of bi-layer La_2O_3 (4 nm)/ SiO_x (1 nm) gate stack scanned at different bias polarity. Bias voltage, V_s , is applied to the Si substrate. (a) at $V_s = +3.5$ V, the polycrystalline La_2O_3 layer is imaged. (b) at $V_s = -4$ V, the amorphous SiO_x IL is imaged [99].

Although the high- κ phase change is clearly evidenced in the cross-sectional HRTEM (Fig.2.2) and AFM (Fig.2.3), the correlation of the GB-related defects and the high- κ degradation gives rise to contradicting inferences. Some researchers believed that the high- κ gate stack BD is related to GB-related defects [31, 102], whereas the others attributed to the “bulk” defects since a greater trap density is produced compared to that generated by the GBs, hence the GB-related defects are negligible [23, 103]. The electronic properties of GB and grain are unclear due to the limitations of TEM (i.e. no electrical information extracted) and standard characterization tools (i.e. area-averaged effect). Ong *et al* revealed that the leakage sites across a Sc_2O_3 / La_2O_3 / SiO_x /Si gate stack exhibit different I - V characteristics (as shown in Fig. 2.24 and Fig. 2.25) [104]. Two types of localized leakage sites were identified, type A and D are due to the lowering of the conduction band minimum of the high- κ and SiO_x IL respectively, whereas type B and C are due to the trap-assisted tunneling through the high- κ and SiO_x IL respectively. Tunneling current is most

sensitive to the barrier height variation at the emitting electrode [105], therefore, the capability of STM measurement at both polarities allows the electronic information at the tip/high- κ and the SiO_x/Si interfaces to be probed separately, as demonstrated in Fig. 2.26 [99].

2.7 Summary

The replacement of SiO_2/SiON with high- κ dielectric (rather than MOSFET's structure change) is the most promising solution for a transistor's downscaling due to the low research/production cost and the long established device physics. Although Hf- and Zr-based dielectrics are relatively compatible to the conventional CMOS processes, the performance and reliability of high- κ dielectrics are compromised by the as-grown defects and the microstructural evolutions (i.e. phase change and IL growth). Unlike the thermally grown SiO_2/SiON , high- κ dielectric is typically deposited via MOCVD, ALD, PVD, etc. Therefore, the interface trap density (D_{it}) and "bulk" defects are far higher compared to that in the SiO_2/SiON . ALD is the most promising high- κ deposition technique due to its capability in producing a high- κ film with atomic thickness control, high densification, low impurity level, excellent conformality and uniformity. Nevertheless, ALD is typically performed at a relatively low temperature (i.e. between room temperature to 300 °C), thus the high- κ film is usually loosely packed and contains a high defect density such as unreacted precursor ligands, oxygen interstitials, oxygen vacancies, and carbon-related impurities. PDA step is, therefore, important to (i) drive off the impurities, (ii) repair the pre-existing defects, and (iii) densify the film. However, the conventional PDA with a relatively high temperature (> 600 °C in N_2) results in a crystallization in the high- κ "bulk", IL

re-growth at the high- κ /Si interface, and reaction at the high- κ /metal interface. These microstructural evolutions cause a severe inhomogeneity in the defects distribution across the high- κ film. Consequently, reliability issues such as high gate current density, transient V_{th} instability, low Weibull slope, bimodal Weibull distribution, polarity-dependent Weibull slope, etc are observed. An inconsistency in the device performances, measured using the standard characterization tools, indicates that the microscopic information and the local leakage sites of an oxide film are overlooked in the data analysis due to area-averaged effect for MOSFET/MOSCAP device. Metrology tools, such as STM, that can extract the microscopic information of an oxide, thus is increasingly important and some researchers have started in exploring this technique to characterize the local defects of an ultrathin oxide. The recent published reports of the high- κ studies by using SPM have shown promising results, such as the revelation of different local leakage types, probing of the electronic properties of IL and high- κ layer respectively, etc. This provides useful insights into the ambiguous interpretations of the electrical data, such as area-scaled bimodal Weibull distribution, the correlation of GBs and the underlying IL degradation, UVO annealing effect on the high- κ film quality, etc. In this thesis, we utilize STM to address some of these issues. The nanoscale characterization results are compared with the electrical data acquired from the MOS capacitors to advance the understanding of high- κ gate stack degradation physics.

Chapter 3 EXPERIMENTS

3.1 Introduction

In the previous chapter, the current reliability issues faced in the high- κ integration are reviewed. The issues can be categorized into two major aspects, (i) process-induced microstructural evolutions, i.e. high- κ phase change and low- κ SiO_x IL re-growth at the high- κ /Si interface, (ii) electrical characterization, i.e. bimodal Weibull distribution, V_{th} instability due to a severe charge trapping, and high gate leakage current. It is known that the standard characterization methods faced limitations in resolving the high- κ BD physics due to the (i) device's nanometer dimension, thus the electrical characteristics are sensitive to the small perturbations in properties such as interface roughness and electronic/structural defects, (ii) high defects density inherent in high- κ film, (iii) multilayer structure resulting from uncontrollable IL re-growth. It is shown that STM is capable in providing atomic resolution imaging and direct probing of the defects localized in the high- κ and IL respectively, thus this technique offers useful insights into the high- κ BD mechanism.

This chapter is divided into three main sections; Section 3.2 presents the details of the samples prepared for both STM and device characterizations. High- κ deposited blanket samples and MOS capacitors employing the identical high- κ films were prepared for STM and device measurements respectively. Section 3.3 discusses the

UHV STM system's main components, operation modes, electronics control system and tip fabrication. Section 3.4 introduces the Cascade Microtech wafer probe station along with Keithley 4200-SCS parameter analyzer, which are used to perform electrical measurements such as I - V , C - V , and TDDB for MOS capacitor.

3.2 Sample fabrication and preparation

3.2.1 High- κ deposited blanket sample

Three sets of high- κ deposited blanket samples were prepared for our STM study as summarized in Table 3.1. Set A consists of two types of HfSiO_x gate stacks fabricated in SEMATECH. The HfSiO_x film contains 10% of SiO_x and was deposited via ALD on p-type Si (100) substrate. A ~ 1 nm SiO_2 IL was intentionally formed on the Si substrate prior to the HfSiO_x growth to improve the high- κ /Si interface. The total thickness of HfSiO_x film was 2 nm and was achieved in two deposition-and-annealing (D&A) steps. A ~ 1 nm HfSiO_x was first deposited and then subjected to an *in-situ* PDA at less than 700 °C in N_2 to avoid additional IL growth. Lee *et al* [70] reported no increase in the IL thickness for a 4.5 nm HfO_2 subjected to 700 °C annealing in N_2 for as long as 60 s. Upon the completion of the first PDA, another 1 nm layer of HfSiO_x was deposited, followed by a second PDA under the same condition (labeled as $\text{HfSiO}(2\text{-step})$). Due to the densification of the film during the first PDA [80, 106], additional ALD cycles were introduced in the second D&A step in order to achieve the identical total physical thickness as the control sample, which was prepared in a single D&A step (labeled as $\text{HfSiO}(1\text{-step})$). In order to ensure that the STM-measured results are directly relevant to an actual fully-processed gate stack, TiN (corresponds to the gate electrode) was sputter-deposited on the HfSiO_x film and the

wafer was subsequently subjected to a 1000 °C spike anneal in N₂ (corresponds to the source/drain dopant activation). The TiN was removed via wet etching prior to the STM study.

		High-k (nm)	SiO _x IL (nm)	
Set A	HfSiO _x			
	1-step	2	1	
	2-step	2	1	
Set B	HfZrO			
	as-dep	2	1	
	RTA(600°C)	2	1	
	2step_UVO30s	2	1	
	2step_UVO2min	2	1	
	4step_UVO1min	2	1	
Set C	HfO ₂		SET 1	SET 2
	as-dep	2.5	<i>none</i>	1.2
	RTA(420°C)	2.5	<i>none</i>	1.2
	RTA(420°C)-UVO	2.5	<i>none</i>	1.2
	UVO	2.5	<i>none</i>	1.2
	UVO-RTA(420°C)	2.5	<i>none</i>	1.2
	UVO-RTA(500°C)	2.5	<i>none</i>	1.2

Table. 3.1: Gate stack structures and physical thicknesses of the high- κ film and IL measured from the cross-sectional HRTEM for Set A, B and C samples. The high- κ films were deposited via ALD and annealed in various conditions. All the samples were grown on p-Si (100) substrate.

For Set B samples, a 150 mm prime grade p-Si (100) wafer with substrate resistivity < 10 ohms/cm was cleaned with standard RCA consisting of SC-1 (NH₄OH: H₂O₂: H₂O = 5: 1: 1 at 80 °C for 10 mins) and SC-2 (HCl: H₂O₂: H₂O = 5: 1: 1 at 80 °C for 10 mins) to remove the organic and ionic/metallic contaminations respectively. Immediately prior to the high- κ deposition, the substrate was dipped into a diluted hydrofluoric acid, HF (HF: H₂O = 1: 20), to remove the native oxide and to achieve H-terminated surface. A 1 nm thermal oxide was intentionally formed via rapid thermal oxidation on Si substrate followed with a 2 nm HfZrO film growth using Savannah-

S300 ALD system supplied by Cambridge Nanotech Inc. The ALD process was performed at 250 °C under 0.2 Torr using Tetrakis (dimethylamino) Hafnium and Tetrakis (dimethylamido) Zirconium as high- κ reactant precursors and H₂O as oxidizing agent. The ratio of as-deposited Hf/Zr is 3 : 1 as determined by XPS analysis. The HfZrO film was deposited in either two or four steps. After each step, HfZrO film was annealed *ex situ* at room temperature in O₂ under UV illumination (254 nm wavelength) generated by a low-pressure mercury vapor grid – UVO annealing. An interaction between the UV illumination and O₂ will produce oxygen radicals and ozone at a concentration of ~500-1000 ppm. The UVO annealing lasted 30 s or 2 minutes for the two-step case while it lasted 1 minute for the four-step case. The respective gate stacks are denoted as 2step_UVO30s, 2step_UVO2min and 4step_UVO1min. No high temperature annealing was performed so as to mimic the gate-last process. For comparison, as-dep HfZrO (single deposition step without annealing) and RTA HfZrO (single deposition step followed by 600 °C in N₂ for 30 s) were also fabricated. The total ALD cycles are kept constant to ensure that all the samples are with a similar thickness.

For Set C samples, a 150 mm prime grade p-Si (100) wafer was cleaned in exactly the same method as in Set B samples. A 2.5 nm HfO₂ film was grown via ALD on (~1 nm) SiO₂/Si substrate at 250 °C under 0.2 Torr. The high- κ reactant precursor used was Tetrakis (dimethylamino) Hafnium and the oxidant used was H₂O. The ALD was divided into four steps. In each step, a 0.7 nm HfO₂ was deposited and then subjected to two-step annealing (comprising an *ex situ* room temperature UVO anneal for one minute and a subsequent RTA at 420 °C in N₂ for 30 s). The ozone concentration was ~500-1000 ppm. For comparison, control samples with (i) as-deposited HfO₂ (i.e. without receiving any annealing), (ii) subjected to either UVO anneal or RTA, and (iii)

subjected to two-step annealing in reversed order (i.e. RTA followed by UVO anneal) were prepared. No high temperature annealing was performed so as to mimic the gate-last process. In order to examine the relative impact of the two-step annealing on the high- κ and SiO₂ IL respectively, an identical set of gate stacks were reproduced on HF-last (i.e. no IL growth) Si substrate.

Both Set B and C samples were produced in Nanyang Technological University (NTU) cleanroom. The ALD high- κ film thicknesses were measured using spectroscopy ellipsometer and confirmed by cross-sectional HRTEM. Prior to STM study, the samples were subjected to 300 °C anneal in UHV (~500 pTorr) chamber for 72 hours for degassing and removal of hydrocarbon-related elements present on the high- κ surface. The characterization of high- κ deposited blanket sample by using STM is fundamentally similar to the characterization of nano-sized MOS capacitor since the conductive probe may act as a gate electrode. In our work, we will correlate the STM-characterized results with the macroscopic electrical data acquired from the MOS capacitors to address unresolved reliability issues.

3.2.2 MOS capacitor fabrication

In order to examine the consistency and relevancy of the characterization results acquired at the nanoscale (i.e. STM measurement) and device levels, MOS capacitors employing the identical high- κ stacks described in section 3.2.1 were fabricated. 150 mm prime grade p-Si (100) wafer with substrate resistivity less than 10-ohms/cm was cleaned with standard RCA to remove the organic and ionic/metallic contaminations. A thick ~550 nm field oxide was grown on the Si surface via wet oxidation by subjecting the wafer to 1000 °C in O₂/H₂ for 4 hours in furnace. Subsequently, photolithography and buffered-oxide-etching (BOE) were performed to pattern the

device active area. The remaining field oxide on the Si substrate serves to protect and isolate the devices. The wafer was dipped with diluted HF to remove the native oxide and to achieve H-terminated Si surface prior to SiO₂ IL growth or a direct high- κ deposition under the various conditions described in the previous section. A ~100 nm TiN film was sputter-deposited and patterned to form the gate electrode. Finally, the MOS capacitors were subjected to forming gas anneal (FGA) at 425 °C for 30 minutes.

3.3 Scanning Tunneling Microscopy

The invention of STM based on the quantum-mechanical of electron tunneling for imaging conducting surface at atomic scale is a significant breakthrough in surface science. In recent years, researchers have been extensively exploring this powerful technique for thin oxide characterization. Due to an ultrathin (in a few nanometers) oxide thickness, current can be detected at the tunneling junction at a given bias, therefore STM can be employed to investigate the electronic properties and topography of an ultrathin high- κ film. A powerful feature of STM is the *real-space* visualization of an oxide degradation state at atomic scale. During STM scanning, a variation of the tip-sample gap spacing is converted into a topography image. Since tunneling current is exponentially dependent on the tip-sample gap spacing, electrons transfer at the tip apex involves only one or a few atoms at any given location of the tip over the sample surface. This allows the possibility to visualize sub-angstrom surface structure and detect the various atomic-scale defects present in an oxide film [107]. Fig. 3.1 shows a schematic view of STM setup. As a bias voltage is applied to the substrate of the sample, tunneling current in ~pA range is detected at the tunneling junction. An increase in the gap spacing will lead to an exponential decay in the

tunneling current. Therefore, STM is extremely sensitive to a small change in a surface contours.

Fig. 3.1: Schematic view of STM setup. The tip-sample gap spacing is less than 10 Å. As a bias voltage is applied to the sample's substrate, tunneling current in ~pA range is detected at the tunneling junction. Owing to an exponential relationship between the tunneling current and the tip-sample gap distance, STM exhibits an excellent sensitivity to Z height.

3.3.1 Instrumentation

UHV STM system, supplied by RHK Technologies Inc., was used in our research for the nanoscale characterization of ultrathin high- κ film. The system can be categorized into two major parts, i.e. mechanical components and control electronics. The mechanical components consist of vacuum pump (to achieve UHV condition in the STM chamber), pressure gauge (for pressure reading), piezo scanning unit (to control the lateral and vertical motions of probe tip), vibration isolation stage (to isolate the STM chamber from the surrounding vibrations and noises), and probe tip. The control electronics consist of computer interface (for probe tip motions control setting and spectroscopy data acquisition), pre-amplifier electronics (to detect and amplify the tunneling current in order to limit noise and interference), and feedback

control electronics (to establish a feedback loop for the tunneling current). The mechanical components and control electronics of the UHV STM system will be briefly discussed in this section.

3.3.1.1 Mechanical structure

Fig. 3.2 shows the main mechanical structure of UHV STM system consisting of three main chambers (i.e. fast entry lock-lock (FEL) chamber, preparation chamber, and STM chamber) separated by two gate valves. The FEL chamber is first vented to the ambient pressure (~ 760 Torr) by using N_2 gas before the sample is loaded onto a built-in magnetic transfer arm. Subsequently, the pressure is pumped down to $\sim 10^{-8}$ Torr by using both mechanical pump (which can achieve as low as $\sim 10^{-4}$ Torr) and turbo molecular pump (which can achieve as low as $\sim 10^{-8}$ Torr). The gate valves are opened and the sample is transferred from the FEL chamber into the STM chamber,

Fig. 3.2: Mechanical structure (in plain view) of UHV STM system (UHV 3500 series) supplied by RHK Technologies Inc. The system consists of three main chambers: the fast entry lock-lock (FEL) chamber, preparation chamber and STM chamber. Three external air legs are mounted to the system to isolate the sample stage in the STM chamber from the surrounding vibrations and noises.

without breaking the vacuum, using the magnetic transfer arm which serves as a transporter. The UHV ($\sim 10^{-10}$ Torr) condition in the STM chamber is maintained by a titanium sublimation pump (TSP) and a sputter ion pump. The sample stage in the STM chamber is integrated with heating feature, thus allowing the sample to be degassed in UHV environment. The storage elevator in the STM chamber provides temporary sample\|tip storage, and the wobble stick facilitates *in-situ* sample\|tip transfer between the storage elevator, sample stage, and magnetic transfer arm.

3.3.1.2 Control electronics

Fig. 3.3 shows the schematic block diagram of UHV STM closed-loop feedback control system. During approaching (in z -direction) the probe tip towards a sample surface, the current detected at the tunneling junction will be converted into a voltage by a current-voltage pre-amplifier (IVP). Since tunneling current, I_t , is exponentially dependent on the vacuum gap spacing between the tip and sample surface, as indicated in equation (2-5) in Chapter 2, a logarithmic amplifier is used to process the output signal of the IVP to ensure the signal shows a linear response with respect to the tunneling gap. Subsequently, the output of the logarithmic amplifier is fed to a comparator to generate an error signal through a comparison with a reference current (i.e. pre-determined feedback setpoint). The error signal is passed to the feedback electronics, which consists of proportional and integral control functions. Both the signals are summed to generate a feedback output and then applied to the piezoelectric tube through a high-voltage amplifier, which gives an effective sensitivity of $0.07 \text{ \AA/V}$ [108]. The discrepancy between the setpoint I_t and the detected I_t is eliminated by adjusting the tip-sample vacuum gap spacing (i.e. z -height). Alternatively, the tip-sample gap spacing can be frozen through an integrated sample-and-hold circuit for (i) sample conductance measurement. The local spectrum (i.e. I - V characteristic) at the

tunneling junction is acquired at a fixed tip-sample gap spacing with the feedback system temporarily disabled, and (ii) constant voltage stressing to induce local oxide degradation or BD. A careful setting and fine-tuning of the feedback parameter is crucial for acquiring a reliable data. The response time of the feedback system shall not be too fast or too slow. A low time constant and a high gain value will result in a high z-signal oscillation issue and the piezo creep effect due to the response time is too fast. On the other hand, if the response time is too slow, the tip may easily crash with the surface roughness and hence damaging the tip and losing the measurement accuracy.

Fig. 3.3: Schematic block diagram of UHV STM feedback control electronics. At a given bias, the I_t sensed at the tunneling junction is compared with the preset I_t value in order to generate an error z-signal, which will be processed by the feedback system for instantaneous rectification of the tip-sample vacuum gap spacing during imaging.

3.3.1.3 Scanning unit

The capability of STM to achieve atomic-scale imaging is attributed to an excellent property of the actuator made of piezoelectric materials. With the piezo-mechanical

coefficient in the range of 15-20 Å/V [107], a piezoelectric actuator may change its dimensions and shape under the influence of an applied electric field due to its anisotropic crystal structure. This property makes piezoelectric actuator an important motion driver for STM since it can control the movement of the probe tip with an amazing accuracy of 0.01 Å [109]. The commonly used motion driver is hollow tube scanner with the outer surface divided into four longitudinal segments equal in size, as shown in Fig. 3.4(a), and each segment is attached to an electrode for electrical signal propagation from the electronics control system. Whenever two equal voltages, with opposite polarity, are applied to a pair of electrodes diagonally opposite to each other, a tube bending effect is produced due to the contraction and expansion on different sides of piezoelectric (as illustrated in Fig. 3.4(b)). This yields a lateral motion (i.e. x - y direction scanning) for the tip mounted to the tube.

Fig. 3.4: (a) Schematic diagram of piezoelectric hollow tube consisting of four longitudinal segments on the outer surface and the attached electrodes which control the three dimensional perpendicular motions in x , y , and z directions. (b) Application of opposite polarity voltages to a pair of diagonally opposite electrodes attached to the outer surface of the tube will produce lateral motion for the tip. (c) Extension and contraction Δl of the tube due to a transverse piezoelectric effect, which produces vertical motion for the tip.

$$\Delta x = \frac{2 \cdot \sqrt{2} \cdot d \cdot l^2 \cdot V_x}{\pi \cdot D \cdot h} \quad (3-1)$$

$$\Delta l = \frac{d \cdot l \cdot V_z}{h} \quad (3-2)$$

The relationship between the tube lateral displacement Δx and the applied electric field is shown in equation (3-1) with d is the transverse piezoelectric coefficient (0.262 nm/V for PZT (lead zirconium titanate) material) of the tube, D is the average diameter of the tube, h is the wall thickness of the tube, and l is the length of the tube [108]. In addition, an electrode is attached to the inner wall surface of the tube. As a voltage is applied to the inner and the outer electrodes, the tube extends or shrinks due to the transverse piezoelectric effect, which produces a vertical motion for the tip (Fig. 3.4(c)). The relationship between the tube vertical displacement Δl and the applied electric field is shown in equation (3-2).

3.3.1.4 Beetle-type scan head

Scan head is one of the most important and delicate components of the system located in the UHV STM chamber. The “beetle-type” scan head consists of four piezoelectric scan tubes. One of the scan tubes is located at the center and underneath the scan head (as shown in Fig. 3.5(b)). This scan tube, mounted with STM probe, controls the two dimensional lateral scanning in orthogonal directions (x and y) for topography imaging or spectroscopy data acquisition. The other three piezoelectric scan tubes, with sapphire balls attached at the bottom (Fig. 3.5(b)), form the outer legs of the scan head. The outer legs, which stand on the cylindric ramp of a sample holder (Fig. 3.5(c)), render two functions, (i) approaching the probe tip to the sample surface under control by the closed-loop feedback system; an applied voltage induces the outer

legs “walking” down\up the ramp step by step until the z-height (i.e. the vacuum gap spacing between the probe tip and sample surface), required to achieve the pre-set current, is obtained [110], (ii) lateral motion of the scan head. Since the outer legs of the scan head are standing firmly on the sample holder, vibration noises (such as those transfer through the copper braids attached to the scan head) will not be coupled into the tunneling junction. Moreover, the stiff structure and the high resonant frequency of the scan head eliminate external disturbances, allowing a high resolution STM image is obtained.

Fig. 3.5: STM scan head based on “beetle” design is mechanically and thermally stable. (a) top and (b) bottom views of the scan head. (c) The motion of the scan head outer legs ramping down the slope of a sample holder when approaching the tip to the sample surface.

3.3.2 Tip preparation

A good preparation of STM tip is essential as the tip geometry has a strong influence on the topographical images and spectroscopic measurements. In order to achieve an excellent atomic scale image, a sharp and high-aspect ratio tip with the tunneling current confined to a single atom at the tip apex is necessary. In general, the tungsten (W) or platinum-iridium (Pt\Ir) wire is used to fabricate STM probe tip,

depends on the operation condition either in ambient or UHV environment. Tungsten wire (with 99% high purity and 0.25 mm diameter) can be etched to form a sharp tip by electrochemical etching method [111]. The tungsten tip faced no oxidation issue whenever used in UHV condition. The setup of a “loop” electrochemical etching for W tip fabrication is shown in Fig. 3.6.

Fig. 3.6: Apparatus used for the “loop” etching technique. High-purity tungsten wire is placed through the potassium hydroxide (KOH) thin film formed in the platinum ring. Both tungsten wire and platinum ring act as the potential electrodes. Under applied potential, the place where tungsten wire in contact with KOH thin film would produce a sharp tip due to electrolysis reaction.

3.3.3 Characterization techniques

In this section, the STM operation modes and the characterization techniques used to acquire the microscopic electronic properties of a high- κ stack are discussed. These include the constant current imaging (CCI), current-imaging-tunneling spectroscopy (CITS), and constant voltage stressing (CVS).

3.3.3.1 Constant current imaging (CCI)

In constant current imaging (CCI) mode, the instantaneous value of the I_t sensed at the tip, when it is scanned across a sample surface, is compared with the preset I_t by a closed-

loop feedback system. The output signal of the feedback system is amplified and subsequently sent to the z -piezoelectric element in order to vary the tip-sample vacuum gap spacing in such a way that the I_t is always stay constant, while the lateral motion of the tip, which is being controlled by the xy -piezoelectric elements, proceeds independently. From equation (2-5), it is shown that I_t is exponentially dependent on the barrier width (i.e. corresponds to the tip-sample gap spacing). Since I_t is very sensitive to the vacuum gap spacing (thereby surface corrugations), this phenomenon allows STM to acquire topography image at atomic resolution. For a typical work function of 4 eV, the I_t will increase by 10 times when the tip-sample gap spacing is decreased by 1 Å and vice versa [112]. During STM scanning of an ultrathin oxide surface, the tip will retract at locations with physical protrusion and/or leakage site. However, it will extend closer to the surface at locations with physical valley and/or non-conductive area (as illustrated in Fig. 3.7). The tip height variations will be recorded and manifested as topography image, which contains the convoluted information of the surface morphology and the local electronic properties of a sample.

Fig. 3.7: In constant current imaging (CCI) mode, the tip height is varied by the z -piezoelectric tube based on the output of the closed-loop feedback system in order to keep the I_t at constant when the tip is scanned across a sample surface. The tip retracts at physical protrusion (i.e. surface roughness; denoted at location A and B) and leakage site (i.e. electronic traps; denoted at location C) over a sample surface.

3.3.3.2 Conductance measurement – CITS and RVS

Although scanning tunneling spectroscopy (STS) provides local spectroscopy data of a thin oxide with an extraordinary spatial resolution, STM is constrained by the piezo-creep and the thermal drift effects. Piezo-creep occurs when the piezoelectric material continues to expand for some time or does not contract to the initial strain level immediately upon charging or discharging respectively. This effect results in a distortion of the features, either stretching or compressed, at the beginning of an image scanning. Thermal drift, on the other hand, is due to the surrounding temperature variations which results in a drift in the probe position during scanning, thereby causing a constant shift in the scan window. Both piezo-creep and thermal drift effects can cause inconsistency or artifacts in a scanned image, therefore complicating the data interpretation and analysis. Typically, the piezo-creep effect can be eliminated by letting the scanner piezoelectric legs “warm-up” for some time before a real image scanning. In addition, a proper control parameter setting will prevent the creep effect during spectroscopy data acquisition. Thermal drift, on the other hand, is a more commonly faced issue since an atomic scale image is highly susceptible to distortion when the tip motion is affected by thermal variations. Without a proper control of the surrounding temperature, it is extremely challenging to position the tip over a specific feature for spectroscopy data acquisition. However, current-imaging-tunneling spectroscopy (CITS) technique offers an effective solution in acquiring the local spectroscopy data without the thermal drift effect.

CITS technique allows both CCI scanning and STS measurement performed simultaneously. The tunneling spectra are extracted at all the positions on a sample surface determined by a regular grid of pixels defined over the scan area, as shown in Fig. 3.8. During CITS scan, the tip is scanned across an oxide surface at CCI mode

before it stops at a spectroscopy pixel and the feedback loop is temporarily disabled to perform a voltage ramp (Fig. 3.9). Subsequently, the feedback loop is re-enabled to allow CCI scanning until the next spectroscopy pixel is reached. This process is repeated until all the locations defined by the spectroscopy pixels are completed. The tunneling spectrum recorded at each pixel is mapped to form a current map at a given bias voltage. As the measurements are performed on a clean sample surface in UHV condition, the spectroscopic studies are free from the effect of the surrounding particle deposit.

Fig. 3.8: Spectra view of the CITS current map. The blue dots (as clearly seen in the zoom-in view) denote a spaced grid of spectroscopy pixels defined over a scan area. During CITS mode, the tip will probe at each spectroscopy pixel to acquire the local tunneling spectrum.

Fig. 3.9: Schematic diagram representing the experimental set-up and the bias voltage setting during ramp voltage stressing. Upon achieving the bias setpoint, the tip-sample height is frozen by disabling the closed-loop feedback system. Subsequently, the bias voltage is ramped to measure the local tunneling spectra of an oxide film.

3.3.3.3 Localized constant voltage stressing

The capability of STM in producing high resolution image enables us to specifically probe the local feature of interest (e.g. leakage sites) in an ultrathin high- κ film. By positioning the tip over a local leakage site and disabling the closed-loop feedback system, the bias voltage can be increased and kept constant in order to accelerate the localized oxide degradation and BD (Fig. 3.10). The switching-off of the feedback system is crucial in order to ensure that the tip-sample distance is being frozen even after an increase in the tunneling current resulting from local degradation. The pre-amplifier used in the system is IVP-300 with a low current noise (i.e. ~ 0.3 pA (rms) at 1.5 kHz bandwidth) and current compliance of 10 nA. The evolution of the stress time-dependent tunneling current is monitored by an oscilloscope attached to the current output. It is known that a fluctuation of the surrounding temperature results in an undesired drift in the probe tip, compromising the reliability of an atomic scale measurement. Therefore, a good control of the surrounding temperature (apart from a mechanically and thermally stable “beetle” scan head) is essential in minimizing the thermal drift effect. Fig. 3.11 demonstrates an example of a localized constant voltage stressing, via STM probe, over a HfSiO(3 nm)/IL(1 nm) gate stack at 4.5 V under controlled temperature. The feedback system was disabled upon achieving the bias setpoint. It is observed that the local tunneling current exhibits an abrupt increase and saturates (due to compliance) at 10 nA after 230 s (Fig. 3.11(a)); note that the z-signal (corresponds to the tip-sample gap spacing) shows insignificant drift of only 25 pm after 1000 s stress (or 0.015 Å/minute) (Fig. 3.11(b)). This allows us to perform nanoscale TDDB tests on a particular leakage site.

Fig. 3.10: Schematic diagram representing the experimental set-up and the bias voltage setting during constant voltage stressing. Upon achieving the bias setpoint, the tip-sample vacuum gap spacing is frozen by disabling the closed-loop feedback system. Subsequently, the bias voltage is increased to a higher value and kept constant to accelerate localized BD.

Fig. 3.11: Constant voltage stressing, via STM probe, over a HfSiO(3 nm)/IL(1 nm) gate stack at 4.5 V for 1000 s under controlled temperature. (a) evolution of the local tunneling current. The maximum measurable tunneling current is 10 nA due to compliance, (b) corresponding tip drift during stressing.

3.4 Prober system

Electrical characteristics, such as gate current density versus gate voltage (J_g - V_g), gate capacitance versus gate voltage (C_g - V_g), TDDB, and interface trap density (D_{it}) of a device are measured using 8-inch Cascade Microtech wafer probe station (Fig. 3.12(a)) and Keithley 4200 Semiconductor Characterization System with high

precision (i.e. sub-femto amp) resolution (Figs. 3.12(b)). The wafer probe station primarily consists of (i) a thermal chuck; which serves as a stage for holding 8-inch or smaller wafer; the temperature of the chuck is controlled by ATT System heater and can be set in the range from 25 °C to 300 °C, (ii) air-leg stabilizer; to isolate the thermal chuck from external vibrations, (iii) four probe heads; which are connected to the source-monitor units (SMUs) of Keithley 4200-SCS via 1.5 m low leakage tri-axial cables. Each probe head holds a sharp and conductive tungsten probe with radius 7 μm . Two remote pre-amps are installed at the rear panel of Keithley 4200-SCS to extend the system's measurement resolution from 100 fA to 0.1 fA by effectively adding five current ranges to SMU (10 nA, 1 nA, 100 pA, 10 pA, and 1 pA) for accurate low leakage measurements. Keithley Interactive Test Environment (KITE) characterization program provides user an interface for advanced test definition, parameter analysis, and graphing required for all the related electrical measurements.

Fig. 3.12: (a) Cascade Microtech 8 inch wafer probe station, a platform for holding device under test (DUT). (b) Keithley 4200 Semiconductor Characterization System, a test environment for controlling the various electrical measurements.

3.5 Summary

In summary, the sample fabrication, STM instrumentation and characterization techniques, and prober system are discussed. High- κ deposited blanket samples and MOS capacitors employing identical high- κ stacks were prepared for the STM and the standard electrical characterizations respectively. The STM results provide microscopic information for the high- κ dielectric degradation observed in the standard electrical measurements. The major components constituting the UHV STM system can be categorized into two parts, i.e. mechanical components (which include vacuum pumps, pressure gauge, piezo scanning unit, vibration isolation stage and probe tip) and control electronics (which include computer interface, pre-amplifier electronics, and feedback control electronics). The flexibility of the control electronics allows us to perform various characterizations techniques, such as TDDB testing and conductance measurement. By virtue of a sharp probe tip, the local electronic properties of a particular high- κ defect can be extracted individually for analysis, thus providing useful insights into the ambiguities arising from the macroscopic electrical results.

Chapter 4 THE ROLE OF HIGH-K GRAIN BOUNDARY IN GATE STACK BREAKDOWN

4.1 Introduction

High- κ integration is necessary in order to mitigate the exponential increase in the gate leakage current while at the same time increasing the gate capacitance [7]. However, the understanding, based on the electrical data acquired via conventional characterization method, of the high- κ breakdown (BD) mechanism remains unclear owing to unresolved microscopic details of the gate oxide. Electrical studies of the high- κ device (MOS capacitor and transistor) have resulted in several contradicting inferences [113]. For instance, stress-induced trap generation has been reported to precede in the high- κ layer (implied by an abrupt BD of the gate stack) or in the IL (implied by a progressive BD behavior of the gate stack) by different research groups. In addition, contradicting inferences on the high- κ grain boundaries (GBs) effect were also reported, with some authors [31, 102] attributed the gate stack BD to the presence of GBs whereas the others [23, 103] deduced that the GB effect is negligible since the “bulk” defects contributed a greater trap density than that produced by the former. The difference in the interpretation of the experimental observations has been attributed to the device area-averaged effect. While the trap generation is a localized phenomenon, the macroscopic

characterization technique may have averaged out the important information and therefore resulting in ambiguous observations [114]. The mechanism that triggers a gate stack BD needs to be clarified as this information is important for the research and development of the future electronics devices.

In this chapter, we overcome the area-averaged effect by capitalizing on the powerful nanoscopic characterization technique – scanning tunneling microscopy (STM) - to study the pre-existing and stress-induced defects in high- κ gate stack. The correlation of the high- κ film quality and the underlying IL degradation is examined via polarity-dependent stressing. We also present a systematic STM study of the high- κ gate stack formed via multistep deposition and annealing (D&A) method. The engineered high- κ stack exhibits improved electrical uniformity and higher resistance against constant gate voltage stress-induced breakdown, which is attributed to the reduced GB effect. The STM-measured results are compared with the electrical data measured from the MOS capacitors.

4.2 The impact of high- κ film quality on IL degradation of a bi-layer gate stack

Conventional TDDB test has provided valuable insights into the intrinsic breakdown mechanism of SiO_2 gate dielectric. However, by adopting the same approach in the reliability study of high- κ gate stack yields conflicting outcomes due to the fact that the latter has a more complicated structure (i.e. a multilayer structure with a polycrystalline high- κ film); (i) Some authors [44, 115] inferred that the entire gate stack BD was controlled by the ultrathin interfacial SiO_x layer (IL), on the basis that the BD characteristics were similar to those of a single ultrathin SiO_2 dielectric

(e.g. a low Weibull slope and a progressive BD characteristic); (ii) Others deduced that the entire gate stack BD was controlled by the thicker high- κ layer as an abrupt BD typical of that in a thick SiO₂ was observed [116]. TDDB test has shown that only a single instantaneous BD (i.e. a catastrophic jump in the leakage current) is observed, rather than a two-stage BD (i.e. either high- κ or IL BD first, followed by the BD of the other layer), for a bi-layer high- κ gate stack [28, 117]. A plausible explanation is that, whenever one of the layers fails, a huge applied voltage will drop across the remaining layer [115], resulting in an instantaneous BD due to a much higher electric field that exceeds the BD field of the latter. A major shortcoming of the standard characterization technique is its lack of ability to resolve the microscopic changes in the high- κ gate stack, therefore giving rise to an inconsistent device performance.

In this section, process- and stress-induced conduction paths in a HfO₂/SiO₂ gate stack were examined at nanoscale level by using STM. We show that the degradation of IL is inter-linked to the pre-existing traps in the thicker HfO₂ film. We observe a substantial increase in the tunneling current through the IL region adjacent to the high- κ leakage paths (irrespective of pre-exist or stress-induced), suggesting that the IL-controlled BD is a consequence of the high- κ “pre-breakdown”.

4.2.1 Experiments

A bi-layer gate stack comprising 5 nm ALD HfO₂ and 1 nm interfacial SiO₂ layer (IL) was formed on p-Si (100) substrate. The gate stack was subjected to PDA at 700 °C in N₂ for 30 s. STM study was carried out in UHV ($\sim 10^{-10}$ Torr) under controlled temperature. Bias voltage, V_s , was applied to the sample substrate and probe tip was grounded. STM constant-current imaging (CCI) was carried out at $V_s = +5$ V and tunneling current, $I_t = 90$ pA. The preset I_t was controlled by a built-in electric

feedback circuit which adjusts the tip and high- κ gap spacing instantaneously to keep the I_t at constant throughout the scan. The tip height variations were recorded and manifested as a topography image. The convoluted information of the surface morphology and electrical data can be separated by CITS technique. Upon achieving the preset I_t , the feedback system was temporarily disabled to hold the tip at a fixed position. The V_s was ramped from +5 V to -5 V to acquire the local tunneling spectrum. This process was repeated at all the positions on the high- κ surface. The tunneling spectra (at a chosen V_s) were mapped to construct a current image.

Tunneling current is highly sensitive to the barrier height at the electron emitting electrode [118], therefore STM has the capability to delineate the electronic and structural properties of high- κ and IL under different bias polarity. For instance, the image acquired under positive V_s is a manifestation of the electronic/structural properties of the high- κ layer (since the high- κ layer is adjacent to an electron emitting tip) [119]. Whenever the tip moves over a conductive area (i.e. due to the presence of electronic traps) or a physical protrusion, the tip will retract further away from the high- κ surface in order to preserve the tunneling barrier so as to keep the I_t at constant. The tip retraction will be manifested as bright shades in the topography images whereas the dark shades correspond to the local physical valley and/or the less conductive area [120].

4.2.2 IL degradation at pre-existing high- κ leakage path

Fig. 4 1 shows a series of 6 consecutive topographical images of $\text{HfO}_2/\text{SiO}_2$ gate stack under stress. The $500 \times 500 \text{ nm}^2$ topographical images were acquired at positive V_s , therefore the features in the images are the manifestations of the electronic/structural properties of the high- κ layer [99]. The thermal drift effect results in a slight shift in the features (i.e. bright spots A to G) for every successive scan. Nevertheless, the positions of

the features can be traced. Physical observation indicates that two types of leakage paths present in the $\text{HfO}_2/\text{SiO}_2$ gate stack; (i) pre-existing leakage paths (labeled as bright spots A and B in Fig. 4.1(a)), and (ii) electrical stress-induced leakage paths (labeled as bright spots C, D, E, F and G in Figs. 4.1(c) and (e)). Bright spots A and B are observed at the 1st scan (Fig. 4.1(a)). Under successive stress, the dilations in bright spots A and B are clearly seen; note that the bright spot diameter increases significantly at the 6th scan (Fig. 4.1(f)) as compared with that in the 1st scan (Fig. 4.1(a)). The EELS analysis of SiO_2/SiON BD spot reveals that the center of a percolation path has the highest percentage of oxygen deficiency and it spreads out radially from the center to its surrounding area [62]. This indicates that the dilations of the pre-existing conductive paths are induced by the generation of oxygen vacancies spreading out radially [121]. On the other hand, bright spots C, D, and E appear gradually at the 3rd scan (Fig. 4.1(c)) at the more uniform contrast area. Since the experiment was conducted in UHV (~ 500 pTorr) condition at a relatively low bias voltage (within -5 V and +5 V), the new bright spots appear at the 3rd scan are unlikely to be the deposits from the tip or surrounding, they are attributed to the stress-induced leakage paths. This is confirmed by the dilations of the bright spots at consistent positions.

The evolution of the electronic property of bright spot A was examined by CITS technique. Tunneling current is most sensitive to an electron emitting electrode [105], therefore the tunneling current in the negative voltage regime (henceforth denoted as negative- I_t) in the I - V characteristic reflects the electronic property of the underlying SiO_2 IL. We scrutinize the bright spot A negative- I_t evolution (corresponds to the IL degradation) and its corresponding topographical change (corresponds to the high- κ degradation) in Fig. 4.1 in order to understand the relative impact of high- κ degradation on the underlying IL reliability in a bi-layer gate stack. The evolutions of negative- I_t and the topography of

bright spot A are summarized in Fig. 4.2 and Fig. 4.3 respectively. The result shows that the negative- I_t increases consecutively for every successive scan, suggesting that in the presence of a pre-existing high- κ leakage path, trap generation occurs readily in the underlying SiO₂ IL. However, it is noted that the negative- I_t shows a lesser increase after the 2nd scan. An examination on the corresponding topographical height profiles in Fig. 4.3 reveals that the lesser negative- I_t increase coincides with a significant increase (corresponds to the tip retraction) in the maximum topographical height at the 3rd scan. The tip retraction is to compensate an increase in the tunneling current through the local pre-existing leakage path resulting from the localized oxide degradation. A similar trend is also observed for bright spot B.

Fig. 4.1: Constant-current STM images (CCI) of HfO₂/SiO₂ gate stack for a series of 6 consecutive scans at bias setpoint: $V_s = +5$ V, $I_t = 90$ pA. The color contrast in the CCI is attributed to the local variation of the tip and high- κ gap spacing (i.e. a manifestation of the gate stack thickness and/or conductance) over a 500×500 nm² high- κ surface. The bright spots in the images are observed to have shifted slightly for every subsequent scan due to the thermal drift. Two distinct observations: (i) Pre-existing bright spots, labeled as A and B, which exist at the 1st scan. (ii) Electrical stress-induced bright spots, labeled as C, D, E, F and G, which appear gradually at a relatively uniform contrast area at the 3rd scan (5th scan) for C, D and E (F and G).

Fig. 4.2: Tunneling spectra of bright spot A (pre-existing), extracted from the CITS current map, corresponding to the CCI from the 1st scan to the 6th scan in Fig. 4.1. A consecutive increase in the negative- I_t is observed for every successive scan, consistent with the corresponding increase in the brightness and dilation of bright spot A in Fig. 4.1. However, after the 2nd scan, the negative- I_t shows lesser increase due to the tip retraction. Reference tunneling spectrum measured at a more uniform contrast area is shown.

Fig. 4.3: Topographical height profiles of bright spot A extracted from the 1st scan to the 6th scan in Fig. 4.1. The height and diameter of the bright spot increase significantly after the 2nd scan, signifying a local oxide degradation. The edge of bright spot A is taken as the height reference.

4.2.3 IL degradation without pre-existing high- κ leakage path

Bright spots C, D, E, F and G appear in the topographical image only after a series of successive electrical stress, therefore the observed micro-protrusions on the high- κ topographical image are not related to the local physical surface roughness. Instead, they are due to the localized evolutions of the high- κ electronic properties. The trap generation in the SiO₂ IL at the locations where the stress-induced high- κ leakage paths (i.e. bright spots C, D, E, F and G) appear are examined. Only the topography and I - V evolutions of bright spot C are discussed since the other bright spots exhibit consistent observations. The electronic property of bright spot C was extracted only from the 2nd scan (Fig. 4.1(b)) onwards since the location where this bright spot appears only visible in the 2nd scan window. The subtle appearance of bright spot C at the 3rd scan (Fig. 4.1(c)) implies the onset of high- κ layer degradation; note that the negative- I_t (at the 3rd scan in Fig. 4.4) remains unchanged. This observation suggests that the SiO₂ IL beneath bright spot C remains robustly intact when the high- κ layer degrades. After the formation (as denoted by the dilation of bright spot in Fig. 4.1(d)) of bright spot C, however, the negative- I_t for the subsequent scans (from the 4th scan onwards) show a near-parallel shift towards the less negative voltage regime, suggesting that a stable high- κ leakage path is formed. It is interesting to note that bright spot C shows a similar I - V evolution as in bright spot A, i.e. the negative- I_t increases substantially only after the high- κ layer degrades. Therefore, it is believed that the same mechanism governs the SiO₂ IL degradation irrespective of the pre-existing or stress-induced high- κ bright spots. Electrical studies suggested that IL determines the entire high- κ /IL gate stack BD [44]. Our STM characterization result, however, reveals that the high- κ layer quality has a non-negligible effect on the IL-controlled BD.

Fig. 4.4: Tunneling spectra of bright spot C (electrical stress-induced), extracted from the CITS current map, corresponding to the CCI from the 2nd scan to the 6th scan in Fig. 4.1. At the 3rd scan, a subtle appearance of bright spot C is observed in the CCI (Fig. 4.1(c)); note that the negative- I_t remains unchanged, indicating that the IL quality remains intact. At the 4th scan, a significant increase in the brightness and dilation of bright spot C is observed in the CCI (Fig. 4.1(d)); note the huge increase in the negative- I_t , signifying an onset of IL degradation. From the 4th scan onwards, the negative- I_t shows a near-parallel shift to the less negative voltage regime. Reference tunneling spectrum measured at a more uniform contrast area is shown.

Fig. 4.5: Topographical height profiles of bright spot C extracted from the 2nd scan to the 6th scan in Fig. 4.1. The height and diameter of the bright spot increase significantly at the 4th scan, signifying a local oxide degradation. The edge of bright spot C is taken as the height reference.

4.2.4 Local electronic properties at high- κ grain and grain boundary

A closer examination over bright spot A reveals the existence of a cluster of granular features within this region. Based on the height and current profiles (extracted from the local topographical image and its corresponding current map respectively (Fig. 4.6)), these granular features (generally in 20 ~ 50 nm diameter) exhibit the same characteristics, i.e. the granular features appear as micro-protrusions on the high- κ surface (denoted by bright shades in the topographical image in Fig 4.6(a)) and non-conductive (denoted by dark shades in the current map in Fig 4.6(b)). In addition, the current map shows that the boundaries (bright circumferences) along granular regions have a higher tunneling current. This observation is consistent with that reported in the previous STM studies, which have been shown to be the grains of a polycrystalline high- κ film [102, 114]. The capability to image the high- κ polycrystalline structure allows us to perform localized stress separately on grain and grain boundary (GB) in order to study their corresponding electronic properties, which provides useful insights for GB-assisted oxide BD [63].

In order to investigate the stress-induced evolution of the electronic properties of grain and GB, the local tunneling spectra of grain and GB at the 1st scan and the 6th scan are extracted for comparison. In Fig. 4.7, the negative- I_t of grain shows a marginal increase even after five successive ramp voltage stress (RVS); note that at $V_s = -5$ V, the negative- I_t at the 6th scan (denoted by 6thscan_Grain) increases by only ~ 4 % compared to that at the 1st scan (denoted by 1stscan_Grain). In contrast, under the same stress condition, GB exhibits a substantial increase in the negative- I_t ; note that at $V_s = -5$ V, the negative- I_t at the 6th scan (denoted by 6thscan_GB) increases by ~36 % compared to that at the 1st scan (denoted by 1stscan_GB). The trap generation rate is almost ~ 9 times higher as compared to that at grain. Based on the V_s of 5 V, vacuum gap of ~5 Å, HfO₂ dielectric constant of 24, and EOT of ~1.8 nm, the voltage drop across the gate stack is estimated to be ~0.8 V.

The tunneling current, at -5 V, measured at grain and GB show ~ 25 pA and ~ 130 pA respectively (Fig. 4.7). The resistance at grain and GB, thereby, can be estimated to be ~ 25 G ohm and ~ 5 G ohm respectively. Literature reports the resistance of 12 nm HfO_2 film to be $\sim 10^{11}$ ohm [122]. After successive ramp voltage stress, the GB resistance drops to ~ 4 G ohm while the grain shows no significant degradation. Complex impedance plot can be used to estimate the grain and GB resistance, however, it is not in the scope of the thesis.

This result is consistent with the color contrast in the zoom-in current map of bright spot A, in which the intensity of the bright circumference (corresponds to negative- I_t) encompassing the granular features at the 6th scan shows (relative to the 2nd scan) an obvious increase (Fig. 4.8), indicating a severe GB degradation. A huge negative- I_t increase implies that the ultrathin IL beneath a GB is highly susceptible to trap generation as

Fig. 4.6: (a) Zoom-in view ($150 \times 150 \text{ nm}^2$) of bright spot A in the topographical image in Fig. 4.1. The bias set point was $V_s = +5$ V, $I_t = 90$ pA during image acquisition. Color contrast is adjusted to identify the higher region. Granular features are observed in the image. (b) Zoom-in view of the CITS current map corresponding to the topographical image in (a); bias voltage $V_s = -5$ V. (c) and (d) show the cross-sectional height and current profiles of a granular feature, along dashed line 1-2 in (a) and (b) respectively.

Fig. 4.7: Tunneling spectra of grain (denoted by filled (open) circle for the 1st scan_Grain (6th scan_Grain)) and GB (denoted by filled (open) triangle for the 1st scan_GB (6th scan_GB)) in bright spot A extracted from the CITS current map. The negative- I_t of GB (grain) increases by $\sim 36\%$ ($\sim 4\%$) after a series of successive ramp voltage stress, implying a higher degradation rate in the IL beneath a GB as compared with its counterpart beneath a grain. The tunneling spectra in the positive V_s regime all coincide because they are normalized to the same bias set-point (+5 V, 90 pA).

Fig. 4.8: Zoom-in view ($150 \times 150 \text{ nm}^2$) of the current map (at $V_s = -5 \text{ V}$) of bright spot A (a) at the 2nd scan, and (b) at the 6th scan. Grains, the granular features denoted by the green dashed line, are clearly observed in both images. An increase in the intensity of bright shades (corresponds to negative- I_t) encompassing the grains at the 6th scan indicates a severe degradation of GBs.

compared to its counterpart beneath a grain under stress. Due to the phase segregation in high- κ under high-thermal-budget front-end processes, the randomly-oriented high- κ

grains result in the formation of highly defective structural defects (i.e. low resistive GBs) at the interface between grains, therefore causing a large voltage drop across the underlying IL. The presence of large trap density at the GB facilitates electrons tunneling through the thicker high- κ layer at a lower energy state [123-124], therefore giving rise to a larger negative- I_t as compared with grain.

The local electric field across the IL (E_{IL}) beneath a high- κ leakage path (or GB) can be estimated based on the relationship between the bias voltage applied across the entire bi-layer gate stack (V_{ox}) and the voltage drop and capacitance in the respective dielectric layers, as shown in equation (4-1), (4-2) and (4-3) below [125]. V_{HK} and C_{HK} are the local voltage drop and capacitance of the high- κ layer respectively, whereas V_{IL} and C_{IL} are the local voltage drop and capacitance of the underlying IL respectively. V_{ox} is the total summation of V_{HK} and V_{IL} . For our HfO_2/SiO_2 gate stack, since the EOT_{HK} of the HfO_2 layer (~ 0.8 nm) is closed to the EOT_{IL} of the underlying SiO_2 IL (~ 1 nm), the local voltage drop across the IL and normal high- κ are expected to be almost identical (i.e. $V_{IL} = V_{HK} = V_{ox}/2$). However, the high- κ GB presents low resistance path that enables electrons tunneling through the layer. Considering that EOT_{HK} is proportional to V_{HK} from equation (4-2), one can estimate the local electric field (at a given V_{ox}) across the IL

$$V_{ox} = V_{IL} + V_{HK} \quad (4-1)$$

$$\frac{V_{IL}}{V_{HK}} = \frac{1/C_{IL}}{1/C_{HK}} = \frac{E_{IL} \cdot T_{IL}}{E_{HK} \cdot T_{HK}} = \frac{EOT_{IL}}{EOT_{HK}} \quad (4-2)$$

$$E_{IL} \cdot T_{IL} = \frac{V_{ox}}{\left(1 + \frac{EOT_{HK}}{EOT_{IL}}\right)} \quad (4-3)$$

as a function of EOT_{HK} based on equation (4-3). Fig. 4.9 shows the E_{IL} beneath a high- κ GB as a function of EOT_{HK} . It can be clearly seen that as the EOT_{HK} decreases due to degradation of low resistive GB under stress, the E_{IL} increases rapidly, giving rise to enhanced-field assisted trap generation at this location [113, 125]. Our STM characterization results show that the leakage current through a thin polycrystalline high- κ gate stack preferentially flows via the GBs, underlining the important role played by GBs in the charge transport and degradation of high- κ gate stack. *Ab initio* calculations demonstrated that the positively charged oxygen vacancies tend to segregate and diffuse near the GBs [64], with the oxygen vacancy defect densities at GBs estimated to be $\sim 2.1 \times 10^{21} \text{ cm}^{-3}$ as compared to grain which shows $\sim 3 \times 10^{19} \text{ cm}^{-3}$ [126]. The large defect densities form localized conductive sub-bands that facilitate electrons tunneling through the high- κ layer, as illustrated in Fig. 4.10(a). This indicates a strong correlation between oxygen vacancy defect and GB-assisted gate stack breakdown.

In summary, the correlation of the high- κ film quality and the underlying IL degradation in a bi-layer gate stack, comprising 5 nm HfO_2 and 1 nm SiO_2 IL, is studied at nanoscale level via STM. This study provides evidences that the IL-controlled BD depends ultimately on the “pre-breakdown” of the thicker high- κ film; the stress-induced dilation of the bright spots in the topographical image (corresponds to high- κ layer degradation) consistently precedes an increase in the negative- I_t (corresponds to underlying IL degradation). Result shows that the presence of low resistive high- κ GBs could induce a large voltage drop across the underlying IL, therefore resulting in a rapid degradation rate at this location (as evidenced by a large increase in the negative- I_t of the IL adjacent to GB (+36 %) as compared to grain (+4 %)). This location is a weak spot that is highly susceptible to the formation of percolation path connecting the gate electrode and the Si

substrate. The study of the correlation of the high- κ film quality and the IL-controlled BD is essential to advance the understanding of the high- κ gate stack BD physics.

Fig. 4.9: Local electric field across the IL (E_{IL}) beneath a high- κ GB as a function of EOT_{HK} for a HfO_2 (5 nm)/ SiO_2 (1 nm) gate stack. As the EOT_{HK} decreases due to the degradation of low resistive GB, the E_{IL} is observed to increase rapidly, signifying a larger voltage drop across the IL. Assume the V_{ox} is 1.2 V.

Fig. 4.10: (a) The tendency of oxygen vacancies to segregate along high- κ GBs could introduce conduction sub-bands in the high- κ energy bandgap, which facilitate electron tunneling through the high- κ layer. The low-resistive high- κ GBs induce a high “voltage loading” on the underlying IL, thus accelerating trap generation at this location. Upon achieving the critical trap density, a percolation path is formed and the entire gate stack BD occurs. (b) Schematic energy band diagram of HfO_2/SiO_2 gate stack under substrate injection bias. Electrons tunnel through the high- κ layer facilitated by sub-conduction bands present in the high- κ energy gap.

4.3 Two-step deposited and annealed HfSiO_x gate stack

It is imperative for the advanced metal-oxide-semiconductor (MOS) field-effect transistor to employ a thicker high-permittivity (high- κ) gate dielectric, replacing the ultrathin SiO₂/SiON, in order to keep the gate leakage current at check [7]. Hafnium-based high- κ gate dielectrics have shown great promise, however, their ability to maintain a low leakage current under an aggressive EOT scaling remains one of the primary concerns. The structural defects in a high- κ dielectric can severely compromise the performance and reliability of a gate stack [24]. In particular, the GBs in a crystallized high- κ dielectric have been shown to present localized conductance paths which could induce enhanced-field assisted trap generation at the underlying interfacial SiO_x layer (IL) [65, 127]. Improving the high- κ film quality is important since it has a direct influence on the performance and reliability of the underlying IL (a layer which is believed to control the gate stack BD) as discussed in the previous section. Extensive efforts to improve the high- κ film quality have been reported. For instance, the multistep deposition-cum-annealing (D&A) gate dielectric formation process [128] was shown to be effective for reducing gate leakage current without negatively affecting its electrical characteristic. Conley *et al* [80] and Yeo *et al* [106] reported that a high- κ dielectric deposited and annealed in multiple steps was denser than that achieved via a single step. The improvement was attributed to a more efficient removal of impurity (the most common being carbon originated from the precursor used for the atomic layer deposition (ALD) process) from a thinner high- κ film, as was verified by Nabatame *et al* [78]. However, the film density was not related to its electrical properties. Bersuker *et al* [129] suggested that the observed leakage current reduction in the multilayer high- κ dielectric stacks is caused by the misalignment of the GBs in each of the thin dielectric layer. Indeed, recent work [82, 104, 130-131]

based on scanning probe microscopy (SPM) reported the presence of localized leakage-current paths in the high- κ gate stack after subjected to high temperature anneal, with the density of these leakage paths increases with temperature [82]. The plan-view transmission electron microscopy (TEM) reveals the formation of polycrystalline structure in the high- κ dielectric [99]. The STM [127] current map shows that the localized leakage paths are mainly due to the presence of GBs. When a high- κ thickness is comparable to the average grain size, the GB contributes a significant leakage current through the dielectric.

In this section, we present a detailed and systematic STM study of the reliability characteristics of the high- κ hafnium silicate (HfSiO_x) gate dielectric formed via a multistep deposition-cum-annealing method. We show that STM, by virtue of its excellent spatial resolution (~ 2 Å), is capable of resolving the microscopic details (of feature size < 10 nm) in a thin high- κ dielectric.

4.3.1 Experiments

The HfSiO_x (containing 10% SiO_x) film was deposited via ALD on p-Si (100) substrates. The total thickness of HfSiO_x was ~ 2 nm and was achieved in two deposition-cum-annealing (D&A) steps. Prior to the growth of the HfSiO_x , a ~ 1 nm interfacial SiO_x layer (IL) was intentionally formed to improve the interface between the high- κ film and the Si-substrates. A ~ 1 nm HfSiO_x was first deposited and then subjected to an *in-situ* post-deposition anneal (PDA) at a temperature less than 700 °C in N_2 to avoid additional growth of the IL. Lee *et al* [70] reported no IL thickness increase in the (1 nm) SiO_2 /(4.5 nm) HfO_2 stack subjected to 700 °C annealing in N_2 for as long as 60 s. Upon the completion of the first PDA, an additional 1 nm layer of HfSiO_x was deposited, followed by a second PDA under the same conditions. Because

of the densification of the film during the first PDA [80, 106], several extra ALD cycles were added to the second deposition step in order to match the final physical thickness of the control sample, which was prepared in a single D&A step.

The STM study was performed on blanket-deposited samples. To ensure that the studied gate stacks are identical to those used in the full-processed devices, the entire device stack fabrication process was reproduced. The TiN film (corresponds to the metal gate) was sputter-deposited onto the HfSiO_x film and the sample was subjected to a 1000 °C spike anneal in N_2 . The TiN film was then removed by wet etching process prior to the STM study. From the high-resolution cross-sectional TEM in Fig. 4.11, both single- and two-step HfSiO_x/IL gate stacks are confirmed to have almost identical high- κ and IL thicknesses. In addition, the electron energy loss spectroscopy (EELS) analysis reveals almost similar chemical composition profiles for both gate stacks as shown in Fig. 4.12 [129]. Electrical measurements were performed on the MOS capacitors fabricated using the same gate stack process. The STM characterization results are compared with the electrical measurements in order to examine the consistency of the data obtained at the microscopic and device levels.

STM was performed in UHV ($\sim 10^{-10}$ Torr) condition under controlled temperature. Therefore, experimental artifacts such as electrodeposition and/or anodic oxidation were eliminated [119] and thermal drift effect was minimized. The V_s was applied to the Si substrate and the STM probe (Pt/Ir) was grounded. The vacuum gap between the probe and the high- κ surface was varied via an electromechanical feedback loop to keep the preset I_t constant during topography imaging. Local variations in the conductivity of the high- κ stack were examined by the CITS; during topography scanning at a given position, the feedback loop was temporarily disabled to allow the local current-voltage (I_t - V_s) characteristic of the dielectric directly underneath the probe to be measured by

Fig. 4.11: High-resolution cross-sectional transmission electron micrographs of the (a) single-step HfSiO_x gate stack with 2.3 nm HfSiO_x and 1.4 nm interfacial SiO_x layer (IL) thicknesses, and (b) two-step HfSiO_x gate stack showing identical HfSiO_x and IL thicknesses.

Fig. 4.12: Electron energy loss spectroscopy (EELS) profiles of the $\text{TiN}/\text{HfSiO}_x/\text{IL}$ gate stacks processed using (a) single-step and (b) two-step (D&A) method. Both samples exhibit almost similar chemical composition profiles [129].

applying a voltage-ramp. A set of the I_t values (at a given V_s) obtained at different points (or pixels) on the high- κ dielectric forms a current map, which depicts local conductivity variations over the scanned area.

4.3.2 Electronic properties at grain and grain boundary

Figs. 4.13(a) and (b) show the $100 \times 100 \text{ nm}^2$ constant-current topography images of the single-step and two-step HfSiO_x/IL gate stacks, respectively. During the acquisition of the topography image, a positive bias voltage ($V_s = 3 \text{ V}$) was applied to the gate stack substrate and the tunneling current ($I_t = 25 \text{ pA}$) was held constant by a electromechanical feedback circuit. Electrons were injected from the STM probe and the injection probability, which in turn determined I_t , was predominantly controlled by the vacuum gap that separated the probe from the dielectric surface [104, 132]. Microscopic changes in the vacuum gap, needed to achieve a constant I_t as the probe was scanned across the high- κ surface, gave rise to the contrast in the topography image. Such changes correspond to variations in the high- κ topography or local conductivity induced by the presence of electronic traps. Bright shades correspond to locations where the probe retracts from the high- κ surface (in order to maintain the preset I_t) due to micro protrusions or local increase in the I_t due to the presence of electronic traps [104]. The topography and electronic effects can be distinguished via reference to the corresponding CITS current maps depicted in Figs. 4.13(c) and (d).

In Fig. 4.13(a), a cluster of coarse granular structures (denoted by bright shades), with an average size of $\sim 20\text{-}30 \text{ nm}$, are clearly seen in the topography image of the single-step HfSiO_x . This observation is consistent with the result shown in Fig. 4.6, suggesting the formation of polycrystalline structure in the single-step HfSiO_x . The average size of the granule agrees well with the average grain size observed in the plan

view TEM [119]. The topography image of the two-step HfSiO_x (Fig. 4.13(b)), in contrast, does not show any distinct granular features. Instead, the bright shades, with an average size less than ~ 10 nm are dispersed randomly throughout the high- κ surface, suggesting that the high- κ is in a nanocrystalline phase [133]. Further evidence supporting the above inference is found in the corresponding CITS current maps of the single- and two-step HfSiO_x depicted in Figs. 4.13(c) and (d), respectively. The color contrasts in the current maps denote variations of conductivity across the high- κ gate stacks, with bright shades represent high-leakage sites. In forming the current maps, a negative bias voltage ($V_s = -3.5$ V) is selected, which corresponds to electron injection from the substrate. Although the negative- I_t in this case is controlled by the effective tunneling barrier at the SiO_x/Si interface, the correspondence between the CITS current map and the topography image, especially for the single-step HfSiO_x gate stack, shows that the quality of the overlaying high- κ film also has a direct influence on the electronic property of the underlying IL. A localized conductive path through the high- κ film would translate into a higher voltage drop across the underlying SiO_x IL and hence a higher I_t at a given V_s [127]. The inverse correlation between the dark regions in the CITS current map (Fig. 4.13(c)) and the bright shades in the topography image (Fig. 4.13(a)) of the single- step HfSiO_x gate stack implies that the bright shades in the latter do not arise from a local I_t increase due to the presence of electronic traps in the high- κ film, but rather resulting from the variations in the morphology of the polycrystalline high- κ film. This suggestion is in agreement with another salient feature – the bright shades, which surround the dark regions in the CITS current map (Fig. 4.13(c)). We attribute these bright shades to “leaky” GBs encompassing the less conductive grains (denoted by dark shades).

A physical analysis from high-resolution cross-sectional TEM supports the above inference. In the micrograph of the single-step HfSiO_x (Fig. 4.14(a)), grain orientation lines can be clearly observed in the high- κ layer (see for example the dashed circle). However, these lines are rarely seen in the micrograph of the two-step HfSiO_x (Fig. 4.14(b)). This observation points to the presence of a longer range order in the single-step HfSiO_x as compared to the two-step sample. It corroborates the STM results that the single-step HfSiO_x has transformed into a polycrystalline phase while the two-step HfSiO_x contains mostly randomly oriented nanocrystals.

Fig. 4.13: Constant-current STM images of the HfSiO_x/IL gate stack: (a) single-step HfSiO_x ; (b) two-step HfSiO_x . The scanned area is $100 \times 100 \text{ nm}^2$ and the bias set-point is $V_s = 3 \text{ V}$, $I_t = 25 \text{ pA}$. The contrast in the image is caused by variations in the vacuum spacing between the probe and the high- κ surface, needed to maintain I_t at the preset value as the probe scans across a given area. Bright shades correspond to micro protrusions and/or local conductive regions due to the presence of electronic traps. (c) and (d) are the corresponding CITS current maps ($@V_s = -3.5 \text{ V}$) for the single- and two-step HfSiO_x respectively. Bright shades denote regions of higher I_t .

Fig. 4.14: High resolution cross-sectional transmission electron micrographs of the (a) single-step and (b) two-step HfSiO_x gate stacks. Grain orientation lines (denoted by dashed circle) can be observed in the single-step HfSiO_x gate stack, suggesting the formation of larger grains in the high- κ film. The two-step HfSiO_x , by contrast, exhibits a nanocrystalline structure.

Both STM and TEM results clearly indicate that crystallization of the HfSiO_x is suppressed when it is deposited in thin layer increments. It has been reported that thinner HfSiO_x film was less prone to crystallization [131, 134]. It is believed that the effective diffusion length to overcome a critical nucleation size for two-dimensional lateral growth is increased when the film thickness is reduced. While this might mean that crystallization of the initial 1 nm HfSiO_x was suppressed during the first PDA, it could not explain the reduced crystallization of the final two-step HfSiO_x during the second PDA and the 1000 °C spike anneal, since the total physical thickness was the same as that of the single-step HfSiO_x (cf. Fig. 4.11). We reckon that suppressed crystallization of the final two-step HfSiO_x may be caused by an effect similar to that of Zener pinning [135]. Densification of the initial 1 nm HfSiO_x during the first PDA might have increased the barrier for GB movement during subsequent anneals of a thicker film, thereby inhibiting the crystallization of HfSiO_x as a whole. Extended x-

ray absorption fine-structure (EXAFS) study of thin HfO₂ film attributes the tendency to form small grains in a thin film to the surface energy effects [133]. In addition, the ratio of (tetragonal) *t*-HfO₂ to (monoclinic) *m*-HfO₂ is shown to increase as the film is scaled thinner, which may be seen as an advantage since tetragonal phase has a higher permittivity ($\kappa = 28\text{--}29$) compared to that of monoclinic phase ($\kappa = 16\text{--}18$) [133].

The local tunneling spectra at all the locations on the high- κ gate stack were measured via CITS technique. The I_t - V_s characteristics averaged over the scanned area in Figs. 4.13(c) and (d) of the single- and two-step HfSiO_x gate stacks, are compared in Fig. 4.15. The “ I_t plateau” in the negative V_s regime is a result of the p-Si substrate being biased into deep depletion. Two effects restrict the rise in I_t in the negative V_s regime: (i) the oxide field does not increase as quickly as in the case of a positive bias since the deep-depleted p-Si absorbs most of the increase in the applied $|V_s|$; (ii) the slow minority-carrier generation rate limits the tunneling of electrons from the conduction band of the p-Si. It can be observed that the I_t measured over the dark shade regions in the CITS current map of the single- and two-step HfSiO_x gate stacks (cf. Figs. 4.13(c) and (d)) are similar (denoted by the filled circle and triangle respectively in Fig. 4.15). However, a comparison of the I_t measured over the bright shade regions between the two samples produce an interesting result, with the single-step HfSiO_x exhibits a distinctly higher I_t . It should be mentioned that prior to the I_t - V_s measurement, i.e. during the topography imaging step, the vacuum gap between the STM probe and the high- κ surface was already adjusted by the electromechanical feedback system to achieve the same I_t (10 pA) at a given V_s (5 V) [104]. This explains the convergence of all the I_t - V_s curves at $V_s = 5$ V. The difference between the open circle and triangle hence does not reflect the actual difference in the I_t measured at the bright shades of the two HfSiO_x samples. In fact, one would expect the I_t difference to

Fig. 4.15: Tunneling current versus bias voltage (I_t - V_s) characteristics measured at grain boundaries (GBs) and grains in the current map (Fig. 4.13(c)) of the single-step HfSiO_x gate stack. The I_t - V_s curves measured at the bright and dark shades in the current map (Fig. 4.13(d)) of the two-step HfSiO_x are also shown for comparison. The bias voltage, V_s , was swept from 5 to -5 V. All the tunneling spectra coincide at $V_s = 5$ V due to normalization. The “plateau” in the negative V_s regime is caused by deep-depletion, which limits the increase of I_t .

be larger than that depicted, since the vacuum gap between the probe and the high- κ surface of the single-step HfSiO_x must have necessarily been increased in order to compensate for the initially larger I_t at $V_s = 5$ V before the feedback loop was disabled for I_t - V_s measurement. The result therefore clearly shows (i) the presence of localized leakage paths having much lower resistance in the single-step HfSiO_x; (ii) the formation of these paths has been suppressed in the two-step HfSiO_x.

Since the CITS current map and the I_t - V_s measurement show that the two-step D&A method clearly suppress the localized low resistance leakage paths in high- κ film, we further investigate the impact of this method on the electrical uniformity of the two-step HfSiO_x. The statistical distributions of the I_t of the single- and two-step HfSiO_x are compared in Fig. 4.16(a). The y-axis gives the cumulative distribution of pixels

(normalized to the total number) in the CITS current map having the I_t equal or less than a given value. To allow a better comparison of the spread in the two distributions, the I_t plotted on the x -axis is normalized to the value at the 63rd percentile. The much steeper transition in the characteristic of the two-step HfSiO_x indicates a tighter I_t distribution, i.e. a better electrical uniformity as compared to that of the single-step HfSiO_x. In accordance to Fig. 4.15, this observation shows that the leakage paths which give rise to the “lower tail” of the single-step HfSiO_x characteristic are largely absent in the two-step HfSiO_x. Fig. 4.16(b) depicts the correlation between the I_t extracted under negative and positive V_s , which correspond to the electron injection from the substrate and tip, respectively. When the I_t is measured over the GBs in the tip-injection mode, the corresponding I_t in the substrate-injection mode is higher than the normal I_t (i.e. whenever no high- κ GBs present). This confirms the impact of high- κ film quality on

Fig. 4.16: (a) Tunneling current I_t distributions of the single- and two-step HfSiO_x gate stacks. The area probed by STM is $100 \times 100 \text{ nm}^2$. I_t is normalized to the value at the 63rd percentile. The y-axis gives the cumulative fraction of pixels having the I_t equal or less than a given value. The steeper distribution curve of the two-step HfSiO_x implies an improved electrical uniformity. (b) Correlation between the I_t under substrate injection ($V_s = -2 \text{ V}$) and tip injection ($V_s = 2 \text{ V}$) biasing for the single- and two-step HfSiO_x. Grain and GBs of the single- (two-) step sample are denoted by filled and open circles (triangles) respectively.

the underlying IL reliability [127]. It is also apparent that the I_t taken over the GBs in the two-step HfSiO stack is much lower, indicating a larger effective tunneling barrier. On the other hand, the I_t 's taken over grains are similar for both gate stacks.

As discussed in the earlier sections (cf. Fig. 4.6 and Fig. 4.13), the GBs represent preferential leakage paths, therefore one would expect the BD spot occurs near the GBs. Based on this fact, a single-step HfSiO_x might be expected to exhibit a poorer BD characteristic than that of the two-step HfSiO_x, since GBs, which extend the entire dielectric film (as observed in Fig. 4.14(a)), are potential precursors for the BD paths. Constant voltage stress (CVS) was applied locally over the high- κ via the STM probe. As CVS would lead to an increase in the I_t due to the localized generation of electronic traps, the feedback system was deactivated to keep the probe at a constant height above the high- κ surface (which in turn kept the voltage across the high- κ gate stack at constant). Using the current map in Fig. 4.13(c) as a reference, CVS was applied at the grains and GBs of the single-step HfSiO_x. The time at which the I_t increases abruptly to the compliance setting (10 nA), signifying the BD of the gate stack, is observed to be much shorter for CVS applied on the GBs compared to that on the grains as shown in Fig. 4.17(a). This is despite the fact that a higher stress voltage was applied to the grains (the higher voltage was needed to keep the time-to-dielectric breakdown, T_{BD} , below a certain value to minimize the thermal drift effect). This observation shows that the GBs severely reduce the T_{BD} of the high- κ gate stack. As a comparison, CVS was also applied on the dark and bright shades in the current map (Fig. 4.13(d)) of the two-step HfSiO_x. The T_{BD} for CVS applied on the dark shades is similar to the T_{BD} for CVS applied on the grains of the single-step HfSiO_x (Fig. 4.17(b)), in agreement with the local I_t - V_s spectra depicted in Fig. 4.15. More importantly, the T_{BD} for CVS applied on the bright shades is much longer than the T_{BD} for CVS applied on the GBs of the

Fig. 4.17: (a) Evolution of the tunneling current I_t during constant voltage stressing, applied via the STM probe, on the bright (corresponds to GBs) and dark (corresponds to grains) shades in the CITS current map of the single-step HfSiO_x (Fig. 4.13(c)). The GBs and grains were stressed at 3 V and 5.5 V, respectively. (b) The corresponding I_t evolution for CVS applied on the bright and dark shades in the CITS current map of the two-step HfSiO_x (Fig. 4.13(d)). The T_{BD} data measured on the bright shades are generally longer than that measured on the GBs of the single-step HfSiO_x .

single-step HfSiO_x . Consistency of the above observations is confirmed by the Weibull distributions depicted in Fig. 4.18. The results show clearly that the time-dependent dielectric breakdown (TDDB) of the single-step HfSiO_x is compromised by the more extensive crystallization of the high- κ film. The T_{BD} measured on the bright shades of the two-step HfSiO_x is about one order of magnitude longer than that measured on the GBs of the single-step HfSiO_x . This implies a lower trap density at the GB of the two-step HfSiO_x , therefore a longer stress time is required in order to generate a sufficient number of traps to form a conductive path propagating through the gate dielectric.

The degradation of the gate stack is also studied by examining the evolution of the I_t distribution in a given area under repeated STM scanning. Electrical stress was applied during local I_t - V_s measurement by extending the range of the voltage-ramp. Electronic trap generation in the high- κ and IL are separately probed based on the I_t extracted at positive and negative V_s respectively. When V_s is positive, electrons are injected from the STM probe and the resultant I_t is more strongly influenced by the electronic trap generation in the high- κ , especially near the surface of the high- κ [104]. Conversely, the electronic trap generation in the IL is more effectively monitored by the I_t obtained under negative V_s (i.e. electrons are injected from the Si substrate). Figs. 4.19(a) and (b) compare the I_t distributions, extracted under positive and negative V_s respectively, of the 1st scan and the 4th scan for the single- and two-step HfSiO_x gate stacks. To facilitate the comparison, I_t is normalized to the average I_t obtained from the 1st scan of the single-step HfSiO_x. Relative to the single-step HfSiO_x, a slower increase in the average positive I_t as well as its spread is evident for the two-step HfSiO_x (Fig. 4.19(a)). This indicates a slower degradation rate and a more uniform trap distribution in the latter. Similar observation applies to the evolution of the negative I_t , which is sensitive to the trap generation in the underlying IL (Fig. 4.19(b)). This confirms the impact of the high- κ film quality on the stress-induced degradation of the IL [104, 127].

Fig. 4.18: Weibull distributions of the time-to-breakdown T_{BD} of the local constant voltage stress of Fig. 4.17. T_{BD} is defined as the time at which the tunneling current I_t increases abruptly to the compliance setting (10 nA). The T_{BD} measured on the bright shades of the two-step HfSiO_x (open triangle) is about one order of magnitude longer than that measured on the GB of the single-step HfSiO_x (open box).

Fig. 4.19: Statistical spread in the tunneling current I_t of the single- and two-step HfSiO_x gate stacks. The scanned area is $100 \times 100 \text{ nm}^2$. All the values are normalized to the average I_t obtained in the 1st scan of the single-step HfSiO_x. (a) I_t measured under tip injection biasing, $V_s = 2 \text{ V}$ (probes traps in the high- κ [104]), and (b) I_t measured under substrate injection biasing, $V_s = -2 \text{ V}$ (probes traps in the IL [104]). The single-step HfSiO_x shows a substantial increase in the I_t spread after four successive scans, implying an inhomogeneous trap generation in the high- κ and IL. The two-step HfSiO_x, by contrast, shows a lesser increase in the average I_t and its spread.

4.3.3 Misaligned GBs in the HfSiO_x film

High-resolution cross-sectional TEM (Fig. 4.11) confirms that the physical thicknesses of the HfSiO_x film and the underlying IL are identical for both gate stacks. The quality of the IL for both gate stacks are similar since this layer was initially grown on a bare Si wafer before it was split into smaller pieces for HfSiO_x deposition using single- or two-step D&A method. This is confirmed by the EELS results (Fig. 4.12), which show almost similar chemical composition profiles for both samples. Therefore, the observed improvements in the electrical characteristics of the two-step HfSiO_x must be necessarily related to a difference in the structure of the high- κ itself. Based on the STM and TEM results, which both show a higher degree of crystallinity in the single-step HfSiO_x, we provide a consolidated explanation for the difference in the electrical characteristics of the two gate stacks. Given that the average size of the granular features is ~20-30 nm (Fig. 4.13(a)), we reckon that the likelihood of having GBs which extend through the entire 2 nm thick HfSiO_x (as illustrated schematically in Fig. 4.20(a)) is considerably higher in the single-step sample (cf. Fig. 4.14(a) showing grain orientation lines extending throughout the entire single-step HfSiO_x). Oxygen vacancies V_{os} , which tend to segregate near a GB [136-137] and form conduction sub-bands in the energy band gap [138], is believed to be the main defects contributing the formation of localized leakage path. The fact that the bright shades encompassing the regions of lower conductivity self-consistently shows that they arise from the boundaries between different grains (Fig. 4.13(c)). On the other hand, the probability for the formation of such boundaries is reduced in the two-step sample due to suppressed crystallization of the entire high- κ film. Even if crystallization of the individual layers is to occur, the GBs are likely to be misaligned (as illustrated in Fig.

4.20(b)), thus producing an additional tunneling barrier which reduces the leakage current through the gate stack (Fig. 4.21).

Fig. 4.20: Schematic diagrams of the cross-sectional (a) single-step HfSiO_x , in which the GBs propagate through the entire crystallized high- κ film, and (b) two-step HfSiO_x , in which the GBs in the two thin high- κ sub-layers are, in general, misaligned.

Fig. 4.21: Schematic energy band diagrams of the single-step (a) and two-step (b & c) HfSiO_x gate stacks under substrate injection biasing. (i) and (ii) denote the high- κ sub-layer with and without the presence of GBs (percolation path) respectively. The presence of GBs induces the formation of conduction sub-bands in the high- κ bandgap, thus reducing the high- κ tunneling barrier, as represented schematically by a reduction of the bandgap. Dotted lines show the conduction and valence band edges in the absence of the GB.

4.3.4 Electrical characteristics of metal/HfSiO_x gate stack

The above improvements in the electrical and reliability characteristics of the two-step HfSiO_x, as observed in the STM, are also manifested in the device-level measurement data. For a more comprehensive study, we extend our work by preparing MOS capacitors with the gate stack processed using four D&A steps (each individual layer of HfSiO_x was further thinned down to ~0.5 nm). Similar to the case in the two-step HfSiO_x, due to the film densification after PDA, several extra ALD cycles were added in the 4th D&A step in order to match the total physical thicknesses of the single- and two-step HfSiO_x gate stacks. As shown in Fig. 4.22(a), the MOS capacitors employing the multi-step (D&A) HfSiO_x exhibit a lower gate leakage current density and a higher BD voltage (V_{BD}) as compared to those of the single-step HfSiO_x. Further improvement in the BD strength (though marginal) is achieved by increasing the number of D&A steps to four (Fig. 4.22(b)). In terms of stress-induced degradation, a greater number of D&A step appears to bring a more substantial improvement. With respect to the single-step sample, the four-step sample shows a greater relative reduction in the stress-induced increase of the gate leakage (Fig. 4.23(a)) and hysteresis (Fig. 4.23(b)) than the two-step sample. These results clearly show that the four-step sample has a better immunity to stress-induced electronic trap generation. Since both single- and two-step HfSiO_x devices exhibit comparable interface (N_{it}) and bulk trap (N_{ot}) densities (Fig. 4.24), a reduction in the gate leakage current observed in the two-step sample is believed not related to the interface quality and the bulk trap density [129]. The different extents to which the four-step D&A fabrication affects the degradation and BD improvements of the gate stack may be explained in terms of the different mechanisms involved. BD is a localized process and therefore is determined by the weakest spot in the gate stack. Once the formation of GBs (which extend

Fig. 4.22: (a) Gate leakage current density J_g versus gate voltage V_g characteristics of the MOS capacitors ($50 \times 50 \mu\text{m}^2$) employing the single- and multi-step (D&A) high- κ gate dielectrics. (b) Weibull distributions obtained from the statistics of the V_{BD} , defined as V_g corresponding to an abrupt increase in J_g depicted in (a).

Fig. 4.23: (a) Stress-induced increase in the gate leakage current density J_g , measured at gate voltage $V_g = V_{FB} - 1$ V, where V_{FB} is flat band voltage. The MOS capacitors ($50 \times 50 \mu\text{m}^2$) were stressed at $V_g = -2.2$ V. (b) The corresponding increase in the hysteresis of the capacitance-voltage ($C-V$) curve, for the constant voltage stress of (a). The $C-V$ measurement was made by sweeping the gate voltage from -1 to 1 V and back to -1 V.

Fig. 4.24: Comparison of interface trap (N_{it}) and bulk trap (N_{ot}) measured by the charge pumping at 1 MHz and single pulse methods for the single- and two-step HfSiO_x gate stacks [129].

throughout the entire high- κ) is suppressed, an increase in the number of D&A steps may not yield further significant improvement. However, an increase in the gate leakage and hysteresis depends on the aggregate number of traps generated in a given area. Report [79] showed that RTA in N_2 can effectively drive off the impurities in a high- κ film as the thickness is thinned down. We reckon that densification and impurity removal from the high- κ would continue to improve by thinning down the individual layers via a larger number of D&A steps.

In summary, the topography and electrical characteristics of the HfSiO_x/IL gate stack fabricated by a single-step and a two-step D&A method are compared in detail via STM. High resolution TEM and EELS analyses confirm that both gate stacks exhibit similar physical thicknesses and chemical composition profiles. After the same high temperature ($\sim 1000^\circ\text{C}$) spike anneal, the two-step HfSiO_x is found to retain a nanocrystalline phase while the single-step HfSiO_x has transformed into a polycrystalline phase. The result shows that a multi-step D&A method could

significantly improve the thermal stability of the high- κ film with respect to the grain formation process. This translates into improvements in the electrical uniformity and BD voltage of the two-step HfSiO_x . We attribute the enhanced electrical characteristics to the suppressed formation of GBs that extend throughout the entire high- κ film. The improvements in the electrical characteristics observed via STM are found to correspond to those observed via device-level measurements. This study shows that the multi-step D&A method is a promising means for improving the thermal stability and reliability of high- κ gate stacks.

4.4 Summary

Process- and stress-induced localized conduction paths in the $\text{HfO}_2/\text{SiO}_2$ gate stack were examined by using scanning tunneling microscopy (STM). We show that the degradation of the interfacial SiO_2 layer (IL) is inter-linked to the leakage path (i.e. GB) in the thicker HfO_2 film. In the presence of high- κ GBs, the underlying IL is highly susceptible to degradation (as evidenced by a substantial increase in the tunneling spectra under substrate injection biasing) compared to its counterpart beneath a grain. A rapid trap generation rate at the IL region adjacent to a high- κ GB forms a “bridge” connecting the high- κ leakage path to the Si substrate, thereby triggering the entire gate stack BD. This unambiguous observation provides useful insights into the (i) grain-boundary assisted BD in the MOS device, and (ii) correlation of the IL-controlled BD and the quality of the overlying high- κ film.

We also present a detailed and systematic STM study of the high- κ hafnium silicate (HfSiO_x) gate dielectric grown via a single-step and a multi-step D&A method. Unlike the single-step HfSiO_x , which exhibits coarse granular features indicating a

transformation to a polycrystalline phase, the multi-step HfSiO_x of comparable EOT is observed to retain the nanocrystalline phase following a high temperature (~1000 °C) spike anneal. A localized electrical stress, applied via the STM probe, shows that the multi-step HfSiO_x has a better leakage current uniformity and a higher BD voltage compared to the single-step HfSiO_x. The STM result is consistent with the device-level measurements using the MOS capacitors employing the similar gate stacks. The improvements are ascribed to a partially suppressed formation of the GBs in the thin films constituting the multi-step dielectric, thus reducing the GB-related low-resistance paths. The results show that the multi-step (D&A) process can improve the thermal stability of the high- κ gate dielectrics.

Chapter 5 RELIABILITY OF THE MULTISTEP DEPOSITED AND LOW TEMPERATURE ANNEALED HIGH-K GATE STACK

5.1 Introduction

Although high- κ dielectric presents promising solution in suppressing gate leakage current (therefore reducing off-state leakage power) for transistor scaling, other reliability issues that compromise the performance of a transistor arise, such as threshold voltage pinning, mobility degradation due to soft optical phonons, etc [139-140]. The replacement of polysilicon gate with metal gate has been shown not only succeeded in suppressing the poly depletion effect and screening the soft optical phonons for improved mobility, but also in work function tuning [141]. The replacement metal gate (RMG) method [142], by virtue of the limited thermal budget after metal gate deposition, is an attractive means for gate stack work function tuning in sub-45 nm high- κ /metal gate (HKMG) technologies [142-144]. However, the high- κ gate oxide is typically formed in the early stage of a transistor processing in the conventional RMG approach [142], therefore, the high- κ gate oxide reliability may be compromised due to phase change caused by subsequent high thermal-budget front-end processes such as source/drain dopant activation step (which is typically carried

out at 1000 °C spike anneal) [145-146]. In Chapter 4, the trap generation rates at GBs and grains were spatially resolved by using STM. We show directly that the IL beneath GB exhibits a substantial increase in the tunneling current compared to that at the grain under stress, indicating that GB is an ideal BD location. First principle simulations indicated that the concentration of oxygen vacancies segregated at GBs can be as high as $2.5 \times 10^{14} \text{ cm}^{-2}$, which is sufficient to form percolation paths for electron tunneling [64]. Therefore, the presence of GBs may severely compromise the reliability of a high- κ stack. Depositing the high- κ oxide in the final stage (as in the high- κ last integration scheme) [143-144] could circumvent this problem but further optimization of the gate stack is constrained since as-grown defects in the high- κ oxide could not be optimally removed by the low-thermal-budget of back-end-of-line processes. Studies have reported improved reliability of the high- κ film by post deposition UV ozone (UVO) annealing at room temperature [74, 147]. The improvement is ascribed to the passivation of oxygen vacancies. However, the density of oxygen vacancies in the bulk may not be effectively removed since the UVO anneal affects only the surface of the high- κ oxide [77].

In this chapter, we present our study on the electrical characteristic of the high- κ gate stack formed via multistep deposition cum UVO annealing. The thin high- κ sub-layer deposited in each step is annealed at room-temperature in O_2 under UV illumination. The reliability of the multistep deposited and UVO annealed high- κ stack was examined at nanoscale and device level. Subsequently, we extend our work by investigating the thermal effect on the oxygen vacancies “repair” by UVO annealing. This was carried out by subjecting the thin high- κ sub-layer deposited in each step to two-step annealing (comprising a room-temperature UVO anneal and a subsequent RTA at various

temperatures). The chemical compositions, physical and electrical characteristics of the engineered high- κ gate stacks were systematically studied.

5.2 Multistep deposited and UV ozone annealed HfZrO gate stack

Although the hafnium-based dielectric shows promising potential for replacing ultrathin SiO₂/SiON [148], a high density of oxygen vacancies (V_O 's) inherent in the high- κ dielectric gives rise to gap states leading to transient charge trapping (consequently causing threshold voltage instability in CMOS transistors) and increased gate leakage current [149-151]. Overcoming the pre-existing traps in the high- κ dielectric is crucial to ensure a reliable performance of transistors. Attempts to reduce the V_O concentration via post-deposition nitridation [151], annealing in O₂ [78, 152] and directly oxidizing Hf using ozone [153] have been reported. Although improvements could be achieved, the relatively high annealing temperature involved results in an undesired microstructural evolutions, such as (i) an increase in the SiO_x IL physical thickness due to the oxygen diffusion from the oxide “bulk” or surrounding environment, and (ii) inter-dielectric mixing at the high- κ /IL interface, forming a thin high- κ silicate layer. The presence of IL and high- κ silicate may compromise the EOT of a gate stack since these layers have a relatively low dielectric constant. It is, therefore, essential to develop a means to reduce the V_O density in the low thermal-budget processes (as in the gate-last scheme) to meet the stringent EOT scaling requirement.

Study has shown that UVO annealing at room temperature could lower the gate leakage current of high- κ stack by reducing the V_O density [74]. This offers a

promising approach to control the V_O density at low temperature. However, UVO anneal affects only the surface region of the high- κ [77]. A recent work showed that the multistep deposition cum UVO annealing method could improve the high- κ film reliability [154]. However, the understanding was incomplete since the comparison was made in relation to a sample annealed at a relatively high temperature (600 °C). The relative importance of the UVO annealing and the thermal-budget reduction is unclear.

In this section, we examine the multistep deposited cum UVO annealed HfZrO gate stack via STM and pulse capacitance-voltage (C - V) measurement. The STM and pulse C - V results are compared to those of the rapid thermal annealed (RTA) and as-deposited samples. The microscopic data on the tunneling current obtained via STM, in conjunction with the trap density profiles derived using the pulse C - V method, provide insights into the effect of thermal-budget reduction and UVO annealing on the quality of the HfZrO gate stack.

5.2.1 Experiments

An approximately 1.5 nm thick thermal oxide was intentionally formed via rapid thermal oxidation on p-Si (100) substrate prior to the growth of a 2 nm HfZrO film via ALD at a temperature of 250 °C under 0.2 Torr. The ~1.5 nm interfacial SiO_x layer (IL) was grown to ensure a good interface quality with the Si substrate. The HfZrO film was deposited in either two or four deposition-cum-annealing (D&A) steps. In each step, after a thin HfZrO sub-layer was deposited, the sample was subjected to *ex situ* annealing at room temperature in O₂ under ultraviolet illumination (UVO), as illustrated in Fig. 5.1. For the case of HfZrO gate stack prepared in two-step D&A, the UVO annealing lasted 30 s or 2 minutes in each step; while for the case of HfZrO gate

stack prepared in four-step D&A, the UVO annealing lasted 1 minute in each step. The respective gate stacks are denoted as 2step_UVO30s, 2step_UVO2min and 4step_UVO1min. No high temperature annealing (e.g. source/drain dopant activation spike anneal at ~ 1000 °C) was performed so as to mimic the gate-last process. For comparison, as-dep HfZrO and RTA HfZrO were also fabricated. These two samples were achieved under exactly the same ALD conditions as the multistep deposited sample except that the UVO annealing step was skipped. The RTA sample received *ex situ* PDA at 600 °C for 30 s in N₂ after completion of the ALD while the as-dep sample received no annealing. The high-resolution transmission electron micrographs (TEM) show almost identical HfZrO and IL thicknesses for all the samples (Fig. 5.2). XPS analysis shows no significant difference in the chemical composition (Fig. 5.3). The STM study was performed on the blanket-deposited samples in UHV ($\sim 10^{-10}$ Torr) condition under controlled temperature. The bias voltage, V_s , was applied to the p-Si substrate and the STM probe (Pt/Ir) was grounded. On some samples, TiN film (as metal gate electrode) was sputter-deposited and patterned to form MOS capacitors for pulse C - V measurement. All the TiN/HfZrO/SiO_x capacitors were subjected to forming gas anneal at 425 °C for 30 minutes.

Fig. 5.1: Schematic illustration of the multistep deposition cum UVO annealing process for the 4step_UVO1min sample. An ultrathin (~ 0.5 nm) HfZrO layer was grown on the SiO₂/Si substrate via ALD (at a temperature of 250 °C and a pressure of 0.2 Torr) and the sample was then subjected to *ex situ* room-temperature anneal in O₂ under UV illumination (254 nm wavelength) for 1 min. The ozone concentration produced was ~ 500 -1000 ppm. This process was repeated until the final thickness was achieved.

Fig. 5.2: High-resolution cross-sectional transmission electron micrographs of HfZrO/IL gate stacks deposited via ALD and annealed under various conditions: (a) RTA; (b) as-dep; (c) 2step_UVO30s; (d) 2step_UVO2min; and (e) 4step_UVO1min. All the gate stacks examined show nearly identical physical thicknesses for the HfZrO film and interfacial SiO_x layer.

Fig. 5.3: The XPS core-level spectra of (a) Si 2p and (b) Hf 4f measured from RTA, 2step_UVO2min, and 4step_UVO1min gate stacks. No significant shift is observed. All the peak positions were referenced to the C 1s peak binding energy (BE) at 285 eV [154].

5.2.2 Topography and localized gate leakage density

Fig. 5.4(a) shows the *constant-current* STM topography images of the various HfZrO gate stacks. These images were all acquired under the same fixed positive $V_s = 3$ V and tunneling current $I_t = 25$ pA, which was kept constant by an electromechanical feedback circuit. The probability of an electron tunneling from the STM probe was

controlled by the tunneling barrier at the probe/high- κ interface, thus the contrast in the topography image in Fig. 5.4(a) is a manifestation of the high- κ structure variations [119, 155]. An inverse correlation between the bright granular features ($\sim 10\text{-}30$ nm in diameter) in the topography image (Fig. 5.4(a)(i)) and the dark regions in the corresponding current map (Fig. 5.4(b)(i)) for the RTA sample indicates that the HfZrO film has crystallized after subjected to 600°C anneal in N_2 for 30 s. On the other hand, no such features are apparent in the as-dep (Fig. 5.4(a)(ii)) as well as the UVO-annealed (Figs. 5.4(a)(iii), (iv) and (v)) samples. This means that the high- κ film of the as-dep sample remains amorphous and the UVO-annealed samples retain this structure by virtue of the room temperature under which the annealing was performed.

Spatial variation in I_t due to electronic traps was examined via CITS current map acquired simultaneously during the constant current imaging mode [145, 155]. This technique allows an accurate probing of the local electronic properties of a particular feature on the high- κ surface without the thermal drift effect. Following the adjustment of the vacuum gap to maintain a constant I_t at a given position, the feedback circuit was disabled momentarily and the V_s was swept to obtain the I_t - V_s curve. This was repeated for all the positions on the high- κ surface. Combining the I_t (at a given V_s) for all the positions on the high- κ surface forms a current map (i.e. Fig. 5.4(b)) which depicts the I_t spatial variation across the sample. Bright shades denote the locations of larger I_t due to the presence of electronic traps.

From the current maps (Fig. 5.4(b)) examined, the RTA sample shows the most number of bright spots (Fig. 5.4(b)(i)), indicating the highest defects density in the gate stack. The bright spots can be seen encompassing the regions of lower conductivity grains in the topography image in Fig. 5.4(a)(i). This implies that the leakage paths are mainly situated near the GBs [104, 155]. On the other hand, the

bright spots density (which are randomly dispersed) is evidently reduced in the current map of the as-dep sample (as shown in Fig. 5.4(b)(ii)). It is observed that the density of bright spots in the current map of the 2step_UVO30s sample (Fig. 5.4(b)(iii)) is comparable to that of the as-dep sample. A further reduction in the density of bright spots is achieved with either an increase in the UVO annealing time after each deposition step (i.e. the 2step_UVO2min sample shows reduced bright spots relative to the 2step_UVO30s sample) or an increase in the number of deposition steps (i.e. the 4step_UVO1min sample shows reduced bright spots relative to the 2step_UVO2min sample). The total UVO annealing time is the same for the 2step_UVO2min and 4step_UVO1min samples. The reduction in the number of leakage sites depicted by Figs. 5.4(b)(i) to (v) is in agreement with the corresponding lowering of the average I_t of the scanned area as shown in Fig. 5.4.

It should be emphasized that the difference in the bright spot density of the as-dep and the multistep deposited cum UVO annealed samples pales in comparison to that between the as-dep and the RTA samples. In conjunction with the topography images of Fig. 5.4(a), the results show that a large part of the improvement observed in the multistep deposited cum UVO annealed samples in fact stems from the suppressed crystallization of the HfZrO film. This is contrary to the claim of [154], which compared the UVO annealed samples directly to the RTA sample and attributed the “improvement” in the former to the passivation of oxygen vacancies present at the GBs by UVO annealing. The authors also ascribed part of the improvement to the misalignment of the GBs in the individual high- κ sub-layers in the case of the multistep deposited sample [154]. It is believed that the multistep deposition help suppress the formation of GBs, as opposed to a single step deposition. However, a comparable quality of the as-dep (formed in a single step deposition and received no

Fig. 5.4: Row (a) depicts the constant-current STM topography images of the (i) RTA, (ii) as-dep, (iii) 2step_UVO30s, (iv) 2step_UVO2min, and (v) 4step_UVO1min HfZrO gate stacks acquired at $V_s = 3$ V and $I_t = 25$ pA. Bright shades denote the locations where the tip retracts due to physical protrusion and/or local conductivity on the high- κ surface. Row (b) depicts the corresponding current maps (at $V_s = -3$ V) of the samples. Bright shades denote the regions of higher I_t . The scanned area is 100×100 nm².

Fig. 5.5 Corresponding average tunneling current spectra of the various 100×100 nm² scanned regions depicted in Fig. 5.4(b). The current “plateau” in the negative bias regime is due the p-Si substrate entering into deep-depletion, which limits the number of carriers tunneling across the gate stack. In the positive bias regime, all the curves artificially converge to the set point of 0.1 nA, 5 V used for the tunneling current spectroscopy. The RTA sample shows the largest negative- I_t , indicating a rapid degradation of the underlying IL. However, the negative- I_t is clearly reduced for the as-dep and UVO annealed samples.

annealing) and the 2step_UVO30s samples shows that the proposed mechanism does not play any major role in the samples studied. To begin with, the as-dep sample is already amorphous-like; multi-step deposition is thus not expected to offer further advantage in eliminating the GBs that may directly link the gate and substrate. The further improvements seen in the 2step_UVO2min and 4step_UVO1min samples may be ascribed, respectively, to a greater removal of oxygen vacancy defects under prolonged UVO annealing time and a more efficient removal of oxygen vacancy defects by UVO annealing from a thinner high- κ layer (since UVO anneal only affects the surface region of the high- κ [77]). The following mechanism is thought to be dominant in the photolysis of ozone [156].

The photolysis of ozone by an ultraviolet radiation produces excited oxygen atoms, O(${}^1\text{D}$), which diffuse into the high- κ layer and occupy the oxygen vacancy sites near the surface region. More extensive “repair” is expected (i) under prolonged UVO annealing time and (ii) in a thinner high- κ layer for a given annealing time.

5.2.3 Electrical uniformity and susceptibility against stress-induced degradation

It is believed that the oxygen vacancy defects play the major role in facilitating electron tunneling through an oxide. Removal of oxygen vacancies by UVO annealing, therefore, means a reduction in the gate leakage current density of a dielectric. To quantitatively examine the impact of the room-temperature UVO annealing on the electrical uniformity of the various HfZrO gate stacks, we analyze and compare the

statistical distribution of the I_t (at $V_s = -3$ V) based on the current maps in Fig. 5.4(b). Fig. 5.6(a) shows the I_t distributions of the various samples by plotting the cumulative number of pixels having the I_t equal or lesser than a specified value. Stretch-out of the I_t distribution curve is an indication of the I_t non-uniformity in the scanned area [157]. The I_t distribution curve for the RTA sample (denoted by solid square) shows the greatest stretch-out, indicating the highest degree of I_t non-uniformity. However, the stretch-out is evidently reduced in the case of the as-dep sample (denoted by solid line), pointing to a better electrical uniformity due to the suppressed crystallization of the HfZrO film. The 2step_UVO30s sample (denoted by solid circle) shows a nearly identical stretch-out as the as-dep sample, whereas a further reduction in the stretch-out is observed for the 2step_UVO2min and 4step_UVO1min samples (denoted by open triangle and open circle respectively), signifying a significant reduction of oxygen vacancy defects when the UVO annealing time is extended or the number of D&A step is increased. The result in Fig. 5.6(a) is consistent with the current maps of Fig. 5.4(b).

To assess the immunity of the UVO annealed samples against enhanced field accelerated trap generation, we investigate the stress-induced degradation of the HfZrO gate stacks by repeatedly scanning a given region at a relatively high V_s (~ 5 V). The generation of stress-induced leakage current (SILC) or $\Delta I_t/I_t^0$, which directly probes the generation of “bulk” traps, is monitored at $V_s = -3$ V. The SILC results of the various samples are compared in Fig. 5.6(b). As expected, the RTA sample exhibits the fastest rate of increase in the SILC. In contrast, the SILC generation rate is significantly suppressed in the as-dep sample. Since the as-dep sample was not subjected to UVO annealing, the observed improvement must be necessarily stems directly from the suppressed crystallization of the HfZrO film rather than the

passivation of oxygen vacancies at the GBs by UVO annealing. We attribute the relative ease of “bulk” trap generation in the RTA sample to a faster “wear-out” of the leaky GBs since the oxygen vacancy defects segregated along the GBs are readily acting as “stepping stones” for electrons tunneling under stress. It was reported that oxygen vacancies tend to segregate at the GBs due to a lower diffusion barrier relative to the bulk [64, 138]. In addition, the oxygen vacancy density at the IL region beneath oxygen-deficient GBs is found to be higher [113]. These oxygen vacancies give rise to electronic traps responsible for an increased I_t , leading to a higher trap generation rate. Our STM study [145] showed that the T_{BD} for constant voltage stressing on GB is more than one order of magnitude shorter than that on grain (as depicted in Figs. 4.18

Fig. 5.6: (a) Tunneling current I_t distributions of the various HfZrO gate stacks probed by STM over an area of $100 \times 100 \text{ nm}^2$ at $V_s = -3 \text{ V}$. The y-axis denotes the cumulative percentage of pixels having the I_t equal or less than a given value. To facilitate the comparison, the I_t of the various samples are normalized to the value at the 63rd percentile. The degree of the I_t distribution stretch-out indicates the local electrical uniformity for a given scanned area. (b) Stress-induced increase in the tunneling current I_t measured over a $100 \times 100 \text{ nm}^2$ area of the various HfZrO gate stacks probed by STM. The I_t was measured at $V_s = -3 \text{ V}$. The samples were repeatedly subjected to constant-current topography scanning at $V_s = 5 \text{ V}$. A rapid degradation rate is observed for the RTA sample, whereas the as-dep and the UVO annealed samples exhibit no significant increase in the I_t after successive stress.

in Chapter 4). Further reduction, relative to the as-dep sample, of SILC generation is observed in the UVO annealed samples, which is ascribed to an additional elimination of oxygen vacancies by the UVO anneal. However, it should be pointed out that the improvement pales in comparison to that observed between the RTA and as-dep samples.

5.2.4 Traps profile by pulse C - V measurement

Due to the limitation of frequency (typically 100 Hz, which gives a tunneling time less than 10 ms), charge pumping technique is incapable to measure the traps situated deep inside a thick oxide “bulk” [158]. The standard quasi-dc C - V measurement, on the other hand, takes at least seconds and is not suitable to probe the fast detrapping event near an oxide’s interface [159]. Zhang *et al* [160] demonstrated the capability of pulse C - V method in probing the traps distributed throughout the entire (6 nm) Al_2O_3 /(2 nm) SiO_2 gate stack by virtue of its unlimited discharging time. Since pulse C - V method does not rely on the pulse frequency, the traps present at the interface and “bulk” can be effectively probed. We apply this method to derive the trap density profiles of the various studied gate stacks via Agilent B1500A Waveform Generator/Fast Measurement Unit [160-161]. The programmed gate voltage (V_g) waveform applied on the sample is depicted in Fig. 5.7(a). The rising and falling edges of the pulse (corresponding to a transition from accumulation to deep depletion mode and vice versa) induced displacement current proportional to the capacitance, as described by equation (5-2) below,

$$I_c = C \frac{dV}{dt} \quad (5-2)$$

I_c is the current flowing into the capacitor, C is capacitance of the dielectric and dV/dt is the voltage ramp rate (V/s). The first ramp-up (arrow 0) measured the C - V curve of a fresh device (denoted by filled square in Fig. 5.7(b)). This was followed by pulsing the gate to -2.6 V for an interval of 1 ms to induce hole trapping in the “bulk” traps. The effect of the negative gate pulse on the C - V curve (denoted by open circle in Fig. 5.7(b)) was monitored by the second ramp-up (arrow C). Hole trapping is evident from the lateral shift of the C - V curve along the negative V_g axis. Hole trapping has also been observed in the ultra-thin SiON gate dielectrics under comparable negative oxide field (e.g. [162]). The device was then biased at $V_g = 0.5$ V and the positive bias induced detrapping effect on the C - V curve was measured after selected discharging times (t_{D1} , t_{D2} , etc.). It is noted that after ~ 100 s, the C - V curve was almost shifted back to that of the fresh device, implying that the trapped holes (induced by negative $V_g = -2.6$ V for 1 ms) had nearly all detrapped. It should be mentioned that the C - V curve of a fresh device derived from the pulse C - V method matched that obtained by the standard quasi-static dc C - V measurement, thus validating the pulse C - V reconstruction. For all the C - V measurements, the ramp-rate was fixed at 100 kV/s. This corresponded to a transition time of 31 μ s for a 3.1-V V_g swing. The short transition time minimized the charge trapping/detrapping effect during the C - V measurement [161, 163], therefore improving the accuracy of the quantitative measurement of the trap density. Moreover, the ability to measure the first C - V curve within a short instant (~ 1 ms) following the application of the “charging pulse” allows the effect of “bulk” traps with short emission time constants to be probed. This explains the relatively large lateral shift of the first “charged” C - V curve (relative to the fresh curve) as compared to that normally seen under the quasi-static C - V measurement. The effective positive trapped charge density for the gate stack derived

from the observed V_{FB} shift is in the order of 10^{12} cm^{-2} , consistent with that deduced from the pulse current-voltage measurement data reported in other studies (e.g. [164]), this confirms the quality of the gate dielectric used in this study.

The evolution of V_{FB} shift or ΔV_{FB} (relative to the first charged $C-V$ curve; denoted by red open circle in Fig. 5.7(b)), as a result of the detrapping effect induced by the positive bias $V_g = 0.5 \text{ V}$, is examined as a function of the cumulative discharge time for the various HfZrO gate stacks (Fig. 5.8). To facilitate the comparison, the data for the 2step_UVO30s HfZrO are not shown since they are nearly identical to those of the as-dep sample. The dashed-horizontal line, obtained from the difference between the charged and fresh V_{FB} , represents the total charge trapped in the gate stack for the given interval of 1 ms negative gate voltage ($V_g = -2.6 \text{ V}$) pulse. With an increase in the discharging time, the holes trapped further away from the IL/Si interface would detrapp, resulting in a steady increase of ΔV_{FB} . This shows that a slower measurement would yield a much smaller lateral V_{FB} shift between the charged and fresh $C-V$ curves. For instance, $\sim 10 \text{ s}$ after the charging pulse, the lateral shift is reduced to a mere 100 mV for the RTA sample. For a given difference between two discharge times, a larger incremental change of ΔV_{FB} implies a higher density of “bulk” traps at a particular location in the gate stack. The detrapping time for “bulk” traps situated at 1 nm from the Si/IL interface (corresponding to the IL/HfZrO interface denoted by the vertical dotted line) is estimated to be 40 ms [165].

The gradient of the ΔV_{FB} versus discharge time curve and the total ΔV_{FB} (denoted by dashed-horizontal line) of the RTA HfZrO are the largest (Fig. 5.8), which implies that this sample has the highest “bulk” trap density. As expected, the “bulk” trap density in the IL of the RTA sample is higher than that of the as-dep sample, in good agreement with the observed tunnelling current density in the corresponding STM

Fig. 5.7: (a) Schematic illustration of the programmed gate voltage waveform applied during the pulse C - V measurement. The ramp-up/down rate is fixed at 100 kV/s. Arrows denote the transition edges at which the C - V curve was recorded. The capacitor was charged with a negative voltage pulse ($V_g = -2.6$ V, pulse width = 1 ms) and then discharged at $V_g = 0.5$ V, with the C - V curve monitored at selected intervals (e.g. t_{D1} , t_{D2} , etc.). (b) Evolution of the fresh, negative pulse charged, and discharged C - V curves. A nearly full discharge of the trapped “bulk” charge is observed after ~ 100 s.

Fig. 5.8: Discharging-induced V_{FB} shift (relative to the first charged $C-V$ curve) versus discharging time for the various HfZrO gate stacks. The estimated lateral shift (~ 100 mV) between the charged and fresh $C-V$ curves expected from a quasi-static measurement which takes ~ 10 s to complete is shown for the RTA sample.

current images (Fig. 5.4(b)). In addition, this observation is consistent with the simulation work showing that the quality of the IL is compromised by the defective high- κ layer [113]. A much larger difference is, however, apparent in the high- κ layer, with the as-dep sample clearly having a much lower “bulk” trap density relative to the RTA sample. A further reduction in the “bulk” trap density is observed in the UVO annealed samples, although it is lesser than the difference between the RTA and the as-dep samples. The trap density profiles of the various gate stacks, derived by the pulse $C-V$ method, are consistent with the leakage current density in the current maps obtained from STM, indicating that a large part of the improved characteristics of the UVO annealed sample arises from the suppressed crystallization of the HfZrO film. Nonetheless, a further reduction of the “bulk” trap density could be achieved by increasing the UVO annealing time or the number of deposition steps (with a thinner high- κ layer in each step).

In Summary, we carried out a detailed study of the multistep deposited and UVO annealed HfZrO gate stack by using STM and pulse C - V measurement in an attempt to understand the improved reliability reported in a recent study. Physical evidence shows that a large part of the improvement seen in the multistep deposited and UVO annealed sample, relative to the RTA sample, is a direct consequence of the suppressed crystallization of the HfZrO film by virtue of the reduced thermal-budget in the former. Compared to the as-dep sample, further improvement due to UVO annealing could be achieved by increasing the number of deposition steps or extending the annealing duration. This may be ascribed to a greater efficiency of oxygen vacancies reduction by UVO annealing in a longer time or in a thinner high- κ sub-layer deposited in each step, since UVO annealing affects only the high- κ surface region. This shows that multistep deposition and UVO annealing is a viable means for further enhancing the robustness of the high- κ gate dielectric in a gate-last process.

5.3 Enhanced high- κ reliability via two-steps (UV ozone cum rapid thermal) anneal

High- κ last integration scheme allows the high- κ gate oxide to be integrated into MOSFET without undergoing phase change induced by the high-thermal-budget front-end processes, thereby improving the film reliability [143-144]. However, as-deposited high- κ film is usually loosely packed and contains an inherently high density of as-grown defects, such as oxygen vacancies (V_O 's), which give rise to the gap states leading to transient charge trapping (hence threshold voltage instability) and gate leakage current [149-151]. Attempts to reduce the V_O density via annealing in O_2 ambience [78], direct oxidation of Hf in ozone [153] or post-deposition nitridation

[151] were reported. However, the improvement is usually accompanied by an undesirable increase of the EOT due to the high temperature involved. Moreover, the formation of the GB-related weak spots arising from the crystallization of the high- κ dielectric further impacts the gate stack reliability [145, 155]. In the previous section, we shows that the electrical characteristics of the multistep deposited and UVO annealed HfZrO gate stack are better than the similar gate stack annealed in N₂ at a high temperature (600 °C). STM study shows, however, that a large part of the perceived improvement actually stems from the suppressed crystallization of the HfZrO film [146]. Compared to the as-deposited (as-dep) HfZrO gate stack, only marginal improvement was achieved through the UVO annealing [146]. In this section, we demonstrate a multistep deposition cum *two-step* annealing (comprising a room-temperature UVO anneal and a subsequent 420 °C RTA) method for enhancing the high- κ film quality in a gate-last process. In particular, UVO anneal followed by a low temperature (400 °C – 500 °C) RTA is found to yield a significant improvement in the electrical characteristics of the HfO₂ gate stack without any EOT penalty.

5.3.1 Experiments

An approximately 1 nm thermal oxide was intentionally formed via rapid thermal oxidation on 150 mm prime grade p-Si (100) substrate followed by 28 Å HfO₂ film growth via ALD at 250 °C and 0.2 Torr. Tetrakis (dimethylamino) hafnium was used as the high- κ Hf precursor and H₂O was used as the oxidizing agent. Since UVO anneal was shown to affect only the high- κ oxide surface region [77], the ALD was divided into four steps to thin down the high- κ sub-layers. In each ALD step, a thin ~7-Å HfO₂ film was deposited and the sample was then subjected to *ex situ* UVO annealing at room temperature for one minute followed by RTA at 420 °C (unless

stated otherwise) in N_2 for 30 s (this sample henceforth denoted as UVO-RTA gate stack), as illustrated in Fig. 5.9.

No high temperature annealing was performed so as to mimic the high- κ last process. During UVO annealing, ozone (500–1000 ppm) was produced via interaction between O_2 and UV illumination (254 nm wavelength) generated by a low-pressure Hg vapor grid. For comparison, control samples with 1) as-deposited (as-dep) HfO_2 (i.e. without any anneal); 2) either only UVO anneal [147] or RTA [78] was carried out in each deposition step; 3) the order of the two-step anneal reversed (i.e. high- κ layer was subjected to RTA 420 °C followed by UVO anneal) were also fabricated. High-resolution TEM confirms almost identical HfO_2 film and IL thicknesses for the various gate stacks (Fig. 5.10), except for the thicker IL in the RTA(1000 °C) sample due to a high temperature annealing involved. The STM study was performed on the blanket-deposited samples in UHV ($\sim 10^{-10}$ Torr) under controlled temperature. The bias voltage V_s was applied to the p-Si substrate and the STM probe (Pt/Ir) was grounded. Some samples were sputter-deposited with TiN film (as metal gate electrode) and patterned to form the MOS capacitors for electrical measurement. All the devices were subjected to a final forming gas anneal at 425 °C for 30 min.

Fig. 5.9: Schematic illustration of the multistep deposition cum two-step (UVO-RTA) annealing method. An ultrathin (~ 0.7 nm) HfO_2 layer is grown on p-Si substrate via ALD (at 250 °C and 0.2 Torr) and subsequently subjected to room-temperature UVO (500-1000 ppm) annealing for 1 min followed by RTA (420 °C, N_2 , 30 s). This process is repeated until the final thickness is achieved.

Fig. 5.10: High-resolution cross-sectional transmission electron micrographs of the (a) as-dep, (b) RTA-only, (c) RTA(1000 °C), (d) UVO-only, (e) UVO-RTA, and (f) RTA-UVO samples. Almost all the samples show identical HfO₂ and SiO₂ IL thicknesses except for the RTA (1000 °C), which exhibits a thicker IL due to a high temperature annealing involved.

5.3.2 Electrical characterization of the HKMG gate stack

Fig. 5.11(a) shows the quasi-static high-frequency capacitance versus gate voltage ($C-V_g$) characteristics (each $C-V_g$ curve is an average of 20 devices) of the TiN/HfO₂/IL gate stacks subjected to various annealing conditions. Relative to the as-dep sample (denoted by open circle), only the samples subjected to UVO annealing (i.e. UVO-only, RTA-UVO and UVO-RTA samples) show evidently more positive flat band voltage (V_{FB}). V_o 's typically exist as deep positively charged defect levels, V_o^{2+} , and are the dominant states in the high- κ oxide [166]). The positive V_{FB} shift, therefore, clearly indicates a reduced density of the positive oxide trapped charge in the high- κ resulting from UVO annealing [74], similar to that achieved in O₂ annealing [167]. On the other hand, only a marginal V_{FB} shift is observed for the gate stack annealed with only RTA. Compared to the UVO-only sample, an incremental positive V_{FB} shift for the RTA-UVO sample indicates that the preceding RTA has a subtle effect on the V_o density. In other words, a reduced V_o density in the sample is attributed mainly to the UVO annealing. Interestingly, however, an additional positive

V_{FB} shift, corresponding to a further reduction of the positive oxide trapped charge, is observed for the UVO-RTA sample as compared to the RTA-UVO sample. This implies a further removal of V_o 's by the 420 °C RTA carried out after the UVO anneal. The EOTs extracted from the $C-V_g$ data of the respective gate stacks are shown in the inset. The nearly identical EOTs (except for the one annealed at 1000 °C) indicates no IL re-growth after annealed at 420 °C in N_2 , consistent with that observed in [167]. Clearly, the multistep deposition and two-step anneal do not affect the overall gate stack thickness. The larger EOT of the gate stack annealed at 1000 °C is ascribed to an increase of the IL thickness [70].

The corresponding gate current density J_g versus V_g curves are shown in Fig. 5.11(b). The 1000 °C RTA gate stack exhibits the largest J_g due to a crystallization of the HfO_2 after a high temperature anneal [145, 155]. From the current map obtained by STM (see Fig. 5.13(b)(i)), it is observed that most of the leakage sites preferentially concentrate near the GBs of a crystallized HfO_2 film [145], thus making “leaky” GB an ideal BD location. Compared to the as-dep sample, the J_g of the 420 °C RTA sample (denoted by filled circle) is reduced by about one order of magnitude (i.e. $\sim 17 \mu A/cm^2$ at $V_{FB} - 1$). It was shown that 400-450 °C anneal could densify a high- κ film without inducing phase change [80]. The reduced J_g may be attributed to the removal of carbon-related impurities and un-reacted precursor ligands from the HfO_2 film by the low temperature anneal [80], which help retains the film amorphous structure. For the gate stack annealed with only UVO, similar improvement in the J_g was obtained (denoted by open triangle). However, a further reduction of J_g is achieved when the UVO anneal was preceded by 420 °C RTA, i.e. RTA-UVO sample. This implies that the removal of V_o 's from the HfO_2 film by UVO annealing could proceed more efficiently after the impurities were driven off by the 420 °C RTA. It is interesting to

note that the J_g could be further reduced (by one order of magnitude reduction relative to the sample annealed with either only UVO or RTA) when the sequence of the two anneals are reversed, i.e. high- κ sub-layer was subjected to room-temperature UVO annealing followed by 420 °C RTA. It is believed that the subsequent RTA not only drove off the impurities, but also provided the thermal activation which enhanced the “repair” rate of the structural defects via migration of the oxygen from the UVO anneal at the “surface” to fill the V_o ’s present at the “deeper” region of the film.

Fig. 5.11: (a) Quasi-static high-frequency (1 MHz) capacitance-voltage ($C-V_g$) characteristics of the TiN/HfO₂/IL MOS capacitors subjected to various annealing conditions. The inset shows almost identical EOTs (extracted from the corresponding $C-V_g$ curves) for all the samples except for the one annealed at 1000 °C. The gate stack work function is estimated to be 4.7 eV, typical of the TiN/HfO₂ stack. The lesser accumulation capacitance ($\sim 1.4 \mu\text{F}/\text{cm}^2$ @ -1.15 V) is because of the finite capacitance of the lowly doped ($\sim 6 \times 10^{15} \text{ cm}^{-3}$) p-Si substrate (Debye length ~ 500 Å). (b) Average gate current density versus gate voltage (J_g-V_g) characteristics of the various gate stacks measured under room temperature (300 K). Each curve is an average of 20 devices.

5.3.3 TDDB and stress-induced degradation of the HKMG stack

Apart from the substantial J_g reduction (Fig. 5.11(b)), the UVO-RTA anneal also yields a significant improvement in the gate stack immunity against stress as shown in Fig. 5.12. Under constant gate voltage (-4.8 V) stressing at 125 °C, the UVO-RTA sample exhibits the least susceptibility to trap generation, as evidenced by the longest T_{BD} compared to that of the other samples (Fig. 5.12(a)). The T_{BD} distributions of the UVO-only and the RTA-only samples are clustered together, in agreement with the observation on the very similar J_g of these samples in Fig. 5.11(b). A significantly longer T_{BD} is evident for the two-step annealed samples, with the UVO-RTA sample showing ~ 4 and ~ 12 times longer T_{BD} as compared with the RTA-UVO and the as-dep samples, respectively. The result clearly shows that the BD strength of the gate stack is significantly enhanced by the two-step anneal. A greater increase in the T_{BD} of the UVO-RTA sample consistently shows that when the RTA is preceded by the UVO anneal, a more efficient removal of oxygen vacancies is achieved. This inference is also borne out in Fig. 5.12(b), which compares the positive trapped charge generation rate induced by the negative constant V_g (-4.4 V) stressing at 125 °C for the various gate stacks. Relative to the as-dep sample, the UVO-RTA sample shows the least V_{FB} shift, followed by the RTA-UVO, RTA-only and UVO-only samples. Clearly, the two-step annealed samples are more resistant to the positive trapped charge generation than the single-step annealed samples. The smallest V_{FB} shift of the UVO-RTA sample indicates that this annealing sequence is most effective in reducing the bulk trap precursors in the gate stack, in agreement with the other observations.

Fig. 5.12: (a) Time-to-breakdown T_{BD} distributions of the various TiN/HfO₂/IL MOS capacitors subjected to constant gate voltage (-4.8 V) stressing at 125 °C. Capacitor area: $50 \times 50 \mu\text{m}^2$. The TDDB test was not performed on the gate stack annealed at 1000 °C since its BD voltage is less than 4.8 V. (b) Flat band voltage shift ΔV_{FB} as a function of stress time under constant gate voltage (-4.4 V) stressing at 125 °C.

5.3.4 Gate leakage current density and electrical uniformity by STM

Fig. 5.13(a) shows the constant-current topography images of the various HfO₂ gate stacks acquired under a fixed positive bias voltage $V_s \sim 2.5$ V and tunneling current $I_t \sim 15$ pA [99, 155]. The microscopic changes in the vacuum gap (due to the probe height variations controlled by a feedback circuit), needed to keep the I_t at constant as the probe was scanned across the high- κ surface, give rise to the contrast in the topography image, which corresponds to the convolution data of the high- κ surface morphology and the local conductivity induced by the presence of electronic traps [104]. The morphology and electronic properties of the high- κ stack can be demarcated via reference to the corresponding current maps (Fig. 5.13(b)) obtained by CITS [155]. The contrast in the current map denotes the variations of the I_t across the high- κ gate stack – bright shades represent locally higher leakage sites due to electronic traps.

Fig. 5.13: Row (a) depicts the $100 \times 100 \text{ nm}^2$ constant-current STM topography images of the (i) RTA(1000 °C), (ii) as-dep, (iii) UVO-only, (iv) RTA-only, (v) RTA-UVO, and (vi) UVO-RTA HfO_2 gate stacks acquired at $V_s = 2.5 \text{ V}$ and $I_t = 15 \text{ pA}$. Bright regions correspond to the locations where the probe retracts from the high- κ surface, either due to micro-protrusion or higher local conductivity [104]. Row (b) depicts the corresponding current maps (at $V_s = -3 \text{ V}$) obtained by CITS. Bright shades denote regions of higher local I_t .

The RTA (1000 °C) sample clearly shows a polycrystalline structure of HfO_2 film after subjected to 1000 °C anneal [104, 145, 155]. This is evidenced by the corresponding current map (Fig. 5.13(b)(i)), which indicates the leaky GBs encompassing the less conductive grains (i.e. denoted by bright granular regions in the topography in Fig. 5.13(a)(i)). These features, however, are clearly absent for the as-dep, UVO-only, RTA-only, RTA-UVO and UVO-RTA samples. The bright shades in the topography images of these samples are dispersed randomly throughout, suggesting that the high- κ films are in amorphous-like structure. Moreover, the density of the leakage sites in the corresponding current maps is clearly reduced, with the two-step annealed samples exhibit the lowest leakage density as shown in Figs. 5.13(b)(v) and (vi). This result is in good agreement with the corresponding lowering of the

Fig. 5.14: (a) Corresponding average I_t - V_s curves extracted from the 100×100 nm² scanned area depicted in Fig. 5.13(b). Bias voltage V_s was swept from 5 V to -5 V. All the I_t - V_s curves converge at $V_s \sim 5$ V due to an adjustment of the vacuum gap between the probe and the high- κ surface by a feedback circuit, in order to achieve the same I_t at a given V_s . (b) Corresponding statistical spread in the I_t of the current maps (at $V_s = -3$ V; traps in the IL was probed under substrate injection biasing [104]) depicted in Fig. 5.13(b). To facilitate the comparison, I_t is normalized to the average I_t of the RTA(1000 °C) sample.

average I_t measured from the scanned area (Fig. 5.14(a)). It should be mentioned that during CITS mode, the vacuum gap between the probe and the high- κ surface was already adjusted by the electromechanical feedback system to achieve the same I_t at a given V_s prior to the local I_t - V_s measurement, therefore all the I_t - V_s curves converge at $V_s \sim 5$ V. The difference between the STM-measured I_t - V_s and that obtained from the MOS devices (see Fig. 5.11(b)) is attributed to the presence of vacuum gap between the probe and the high- κ surface, which absorbs part of the V_s applied across the gate stacks. Therefore, the I_t measured by STM in Fig. 5.14(a) does not reflect the actual value and it is expected to be larger.

The stress-induced oxide degradation and the corresponding electrical uniformity in a given area of the respective gate stacks are examined by monitoring the evolution

of the I_t statistical distribution (at $V_s = -3$ V) under repeated STM scanning (Fig. 5.14(b)). The gate stack was stressed by extending the voltage-ramp during local I_t - V_s measurement in CITS mode. The electronic trap generation in IL is probed by the I_t extracted under negative V_s (i.e. electrons are injected from the Si substrate) whereas the local electrical uniformity is monitored by the spread of the I_t distribution [155]. To facilitate the comparison, I_t of the various gate stacks are normalized to the average I_t of the RTA(1000 °C) sample (arbitrarily selected). A slower increase in the average negative I_t as well as its spread is evident for the two-step annealed samples as compared with the single-step annealed counterparts, pointing to a better immunity against trap generation and a lower degree of I_t non-uniformity. Both two-step annealed samples (i.e. RTA-UVO and UVO-RTA samples) exhibit a nearly identical distribution spread. The result in Fig. 5.14(b) is consistent with the current density depicted in the current maps in Fig. 5.13(b). This clearly shows the impact of the high- κ film quality on the stress-induced degradation of the underlying IL.

5.3.5 Chemical analysis by EELS and XPS analysis

From the electrical measurements, it is clearly shown that the post-UVO RTA step substantially reduces the J_g , positive oxide trapped charge generation and prolongs the T_{BD} , as compared with the sample annealed with only UVO. However, the role of the post-UVO RTA step in achieving the improvements is unclear due to a lack of understanding on the chemical structural change in the HfO₂ film. The improvements cannot be explained by the densification of the high- κ film and the reduction of the precursor-related impurities (e.g. carbon, un-reacted precursor ligands, etc) in the high- κ bulk since a lesser J_g reduction was achieved for the case when the order of the two-step annealing steps was reversed (i.e. RTA-UVO sample) and for the case annealed

with only RTA. The chemical structures of the various HfO_2/IL gate stacks were examined by XPS measurements using monochromatic Al $K\alpha$ x-ray source ($h\nu = 1486.7$ eV). The XPS core level spectra of Hf $4f$ and O $1s$ were referenced to C $1s$ peak binding energy (BE) at 285 eV. The Hf $4f$ spectra of the samples are shown in Fig. 5.15(a). Relative to the as-dep sample, the Hf $4f_{7/2}$ peaks of the RTA and RTA-UVO samples shift to a higher BE 17.1 eV in the vicinity of Si (i.e. 17~17.5 eV), indicating the presence of Hf silicates (HfSi_xO_y) at the interface region [168]. A further incremental shift to BE 17.6 eV is observed for the RTA(1000 °C) sample, which is attributed to the formation of a thicker Hf silicate layer induced by the high temperature annealing involved. This result is consistent with the corresponding TEM in Fig. 5.10(c), which shows a thicker IL. For the sample annealed with only UVO, however, a trend similar to the as-dep sample is observed, indicating no obvious intermixing occurs. It is interesting to note that when UVO anneal precedes RTA in the two-step annealing method, a lesser Hf $4f_{7/2}$ peak BE shift is observed, as shown by the UVO-RTA sample. In addition, the intensity of the valley between Hf $4f_{7/2}$ and Hf $4f_{5/2}$ peaks is comparable to that of the as-dep and the UVO-only samples. This suggests no obvious Hf silicate formation, even if there is intermixing occurs, only a thin Hf silicate layer is formed at the interface region. The corresponding XPS core level spectra of O $1s$ are shown in Fig. 5.15(b). The major peak centered at BE ~530 eV is attributed to the Hf–O bonding in HfO_2 [169] whereas the peak centered at BE ~531 to 532 eV at the shoulder is associated with the Hf–O–Si bonding (in Hf silicate) and the Si–O bonding (in SiO_x) [168, 170]. The $I_{(\text{Hf-O})} / I_{(\text{Hf-O-Si, Si-O})}$ intensity ratios in Fig. 5.16(a) imply the relative contents of HfO_2 and Hf silicate/ SiO_x in the respective samples. It is observed that the UVO-only sample exhibits a higher intensity ratio relative to the as-dep sample. Since the Hf $4f$ spectra of the UVO-only and the as-dep

samples are identical (i.e. no obvious Hf silicate formation), an increase in the intensity ratio for the UVO-only sample must be necessarily contributed by a higher HfO_2 content rather than a reduced Hf silicate content. A further increase in the ratio is observed when the post-UVO RTA step was performed, as shown by the UVO-RTA sample. This indicates more Hf-O bonds were formed via an incorporation of oxygen atoms (thus removal of oxygen vacancies) in HfO_2 film by the UVO-RTA method. The RTA-UVO sample shows a comparable intensity ratio with UVO-only sample due to a higher Hf silicate/ SiO_x content. In Fig. 5.16(b), two broad peaks are observed in the XPS core-level spectra of Si 2p. The lower BE peak centered at ~ 99 eV is attributed to the photoelectrons from the Si-Si bonds in substrate whereas the higher BE peak centered at ~ 103 eV is attributed to the photoelectrons from SiO_2 and Hf silicate. Except for the RTA (1000 °C) sample, most of the samples exhibit nearly identical trend, implying a stable oxidation state of Si. The Hf-O-Si peak centered at ~ 102.5 eV for the RTA (1000 °C) is attributed to the presence of a thicker Hf silicate layer.

The inference is also borne out in the electron energy-loss spectroscopy (EELS) study. Fig. 5.17(a) shows the background corrected O K edge spectra of the HfO_2 layer for the various gate samples, which exhibit the characteristic double peaks at the edge onset (one centered at ~ 533.5 eV and the other centered at ~ 537.8 eV), consistent with that reported in the literature [171]. The ratio of the two peaks with comparable intensity heights is nearly identical for most of the samples, except for the UVO-RTA sample, which shows ~ 40 % taller peak (centered at 537.8 eV) relative to that at the shoulder (centered at 533.5 eV). From the O K edge spectra of SiO_2 , only a single peak centered at ~ 537.8 eV is observed and this peak is attributed to the multiple scattering of the ejected O 1s electron from a given oxygen atom to its six second-

Fig. 5.15: The XPS core-level spectra of (a) Hf 4f and (b) O 1s obtained from the HfO₂/IL gate stacks annealed at various conditions. All the spectra were referenced to C 1s peak binding energy (BE) at 285 eV.

Fig. 5.16: (a) $I_{\text{Hf-O}} / I_{\text{(Hf-O-Si, Si-O)}}$ intensity ratios of the respective gate stacks obtained from the O 1s core level spectra in Fig. 5.15(b). The intensity of the peaks centered at BE ~530 eV and that at the shoulder BE ~531-532 eV correspond to the relative contents of HfO₂ and Hf silicate/SiO_x respectively. (b) The XPS core-level spectra of Si 2p obtained from the HfO₂/IL gate stacks annealed at various conditions.

nearest neighboring oxygen atoms (first O shell) [11]. Therefore, the 537.8 eV peak observed in the O K edge spectra of the HfO₂ film may be ascribed to the similar mechanism. The peak intensity reflects the amount of local oxygen atoms (i.e. the intensity will drop when there is a deficiency of oxygen). Relative to the as-dep sample, the ~537.8 eV peak intensity is significantly enhanced by the UVO-RTA anneal, followed by the UVO-only anneal which exhibits comparable intensity with that shown by the RTA-UVO anneal. This confirms that the UVO-RTA method exhibits the most efficiency in incorporating oxygen atoms into the HfO₂ film as compared with the rest. On the other hand, an almost similar trend is also observed for the O K edge spectra of the underlying SiO₂ IL (Fig. 5.17(b)). The peak centered at 537.8 eV can be seen growing evidently for the two-step annealed samples (both UVO-RTA and RTA-UVO samples show comparable maximum oxygen intensity counts) followed by the UVO-only sample. More interestingly, the intensity of the peak centered at ~560 eV is enhanced by the two-step anneal. The 560 eV peak was also observed in the SiO₂ EELS results reported in literature and was attributed to the stretched O-O bonds [172]. This implies that, apart from the high- κ , the quality of the underlying SiO₂ IL was also improved by the UVO anneal, thereby contributing to a larger reliability improvement margin as compared with the similar gate stacks without intentional IL. In addition, a small peak (centered at ~533 eV) with a very low intensity can be seen at the shoulder of the much larger peak (centered at ~537.8 eV) as oppose to the SiO₂ EELS result (which shows no peak at 533.5 eV). This is ascribed to the possible interference signal from the HfO₂ film in the IL EELS results.

Fig. 5.17: Background corrected O K edge EELS spectra of (a) 2.5 nm HfO_2 film, and (b) 1 nm SiO_2 IL for the various gate stacks. The peak intensity corresponds to the amount of local oxygen atoms (i.e. a lower peak intensity indicates a higher oxygen deficiency). Double peaks characteristic can be seen in the O K edge spectra of HfO_2 , consistent with the literature [171]. A growing 537.8 eV peak intensity is observed in both HfO_2 film and SiO_2 IL for the samples annealed with UVO.

5.3.5 Reliability of ultrathin sub-1nm EOT HKMG gate stack

Gate oxide with less than 1 nm EOT is required to meet the sub-45 nm technology node [173]. This can be achieved by limiting the IL re-growth via high- κ last integration scheme. However, without IL (which has a higher BD electric field as compared with the high- κ layer), the reliability margin of the entire gate stack is substantially reduced since IL is the layer that controls the BD. Optimization of the high- κ film quality, therefore, is crucial for sub-1nm EOT device. 7.5 Å EOT high- κ stacks with gate leakage density J_g on the order of 1 A/cm² have been reported [167, 174-175], however, the J_g is much higher than what optimized gate-first processes could provide [176-178]. Some studies were made under the condition where a high quality thermal oxide was already formed prior to the high- κ deposition, thus limiting the EOT to above 1 nm [147, 167]. Studies on realizing a low leakage and reliable HKMG stack with nearly zero IL thickness for high- κ last integration are limited [179]. In this section, we demonstrate an excellent performance and reliability of 7.5 Å EOT TiN/HfO₂ gate stack, achieved via multistep high- κ deposition cum two-step annealing (comprising a RTUVO anneal and a subsequent RTA at 400 °C - 500 °C) method.

A 150 mm prime grade p-Si (100) wafer with substrate resistivity < 10 ohms/cm was cleaned with standard RCA. The wafer was then dipped with diluted hydrofluoric (HF) acid to remove the native oxide and to produce H-terminated Si surface prior to HfO₂ film growth under exactly the same condition (i.e. multistep deposition cum two-step annealing) as described in Section 5.3.1. The HF-last cleaning has helped limit the IL growth to ~3 Å at the ALD temperature, as revealed by the high-resolution transmission electron micrographs in Fig. 5.18. The almost identical IL thicknesses in the as-dep, UVO-only, RTA-UVO, UVO-RTA and UVO-RTA500°C samples implies

Fig. 5.18: High resolution cross-sectional transmission electron micrographs of the (a) as-dep; (b) UVO-only; (c) UVO-RTA; (d) UVO-RTA500°C; (e) UVO-RTA600°C; (f) RTA-UVO gate stacks. All the samples show almost identical HfO_2 (~ 2.5 nm) and IL (~ 0.3 nm) thicknesses, except for the UVO-RTA600°C sample, which exhibits a thicker IL (~ 0.5 nm).

that no IL re-growth occurred during multistep annealing (even at RTA temperature as high as 500 °C), as corroborated by the nearly identical EOTs (~ 7.5 Å) extracted from the corresponding high-frequency (1 MHz) capacitance-voltage (C - V) curves of the respective samples (Fig. 5.19) [180]. This observation is in general agreement with [167], which showed that a 400-450 °C anneal in O_2 would not result in additional IL growth.

Each of the C - V curves in Fig. 5.19 is an average of more than ten devices measurements. The gate stack work function is estimated to be 4.7 eV and the average κ -value of the gate stack is ~ 15 (based on the extracted EOT of 7.5 Å). The accumulation capacitance is lesser (~ 2.6 $\mu\text{F}/\text{cm}^2$ rather than ~ 4.6 $\mu\text{F}/\text{cm}^2$ @ -1.1 V) due to the non-negligible Si capacitance of the lowly doped ($\sim 6 \times 10^{15} \text{ cm}^{-3}$) p-substrate. Although the physical thickness and the EOT of the as-dep sample are identical to that of the other samples, a high gate leakage current is observed for the as-dep sample as shown by the capacitance “roll-off” in the accumulation region [167, 175] and a relatively high gate current density J_g of 0.2 A/cm^2 at $V_{\text{FB}} - 1$ V; V_{FB} is flat

Fig. 5.19: High-frequency (1 MHz) gate capacitance versus gate voltage (C_g - V_g) curves of the various TiN/HfO₂ gate stacks. Each C_g - V_g curve is an average data of more than ten devices. The gate stack work function is estimated to be 4.7 eV, typical of the TiN/ HfO₂ stack. The average κ -value of the gate stack is 15, based on the extracted EOT of 7.5 Å (which corresponds to 4.6 $\mu\text{F}/\text{cm}^2$ oxide capacitance). The lesser accumulation capacitance ($\sim 2.6 \mu\text{F}/\text{cm}^2$ @ -1.1 V) is because of the non-negligible Si capacitance of the lowly doped ($\sim 6 \times 10^{15} \text{ cm}^{-3}$) p-substrate (Debye length $\sim 500 \text{ Å}$).

band voltage (see Fig. 5.20(a)). Since the J_g measured from the as-dep sample is comparable to those obtained under similar thermal budget [167, 174-175] and no difference in the ALD rate was observed, as compared with the case where deposition was made on a pre-grown thermal SiO₂ (which serves as a nucleation layer prior to the high- κ deposition) as shown in Section 5.3.1 (Fig. 5.10), we reckon that the high gate leakage current is not due to the effect of poor nucleation on the HF-passivated Si surface, but is mainly a result of a higher density of as-grown oxide defects. In the RTA-only (i.e. only RTA 420 °C was carried out in each deposition step) sample, the capacitance “roll-off” in the accumulation region is eliminated and the J_g is reduced to 60 mA/cm² at $V_{\text{FB}} - 1 \text{ V}$ (denoted by filled circle). These changes may be ascribed to the purging of precursor-related impurities and film densification by the RTA

performed after each high- κ deposition step [78], leading to improved interface and “bulk” properties.

A marked reduction of J_g is achieved in the UVO-RTA sample. The $J_g \sim 8 \text{ mA/cm}^2$ measured at $V_{FB} - 1 \text{ V}$ for the UVO-RTA sample (~ 25 times lower than that of the as-dep sample) is lower than the 0.1 A/cm^2 level generally attained under optimized gate-first integration [176-178] and is comparable to that of a high- κ last gate stack with a 3-Å nitrided thermal oxide as the IL [179]. This result shows that the two-step annealing (UVO-RTA) method is much more effective (than RTA-only or UVO-only) in enhancing the interface and “bulk” properties of a high- κ film, consistent with that obtained in Fig.5.11(b). It is interesting to note that for the case when the order of the annealing steps was reversed (i.e. RTA-UVO sample) and for the case when only UVO anneal was performed (i.e. UVO-only sample), a lesser J_g reduction is achieved relative to the UVO-RTA sample. This result, again indicates that the RTA step, when performed after the UVO annealing in each high- κ deposition step, plays an important role in the overall enhancement achieved. Just like O_2 annealing [167], UVO annealing eliminates oxygen vacancy (V_o 's) defects in the high- κ oxide [74]. It is believed that the post-UVO RTA step not only drives off precursor-related impurities (thereby resulting in high- κ film densification) [78], it also provides the thermal activation required for V_o “repair”. We expect this latter role to remain relevant in the case of a non-organic Hf precursor, although the exact improvement remains to be ascertained. Considering that no optimization was carried out in our case (e.g. RTA and UVO anneal were performed *ex situ*, no further treatment of the HF-last Si surface was made prior to ALD, a relatively low ozone concentration was used throughout anneal process, UVO anneal time and RTA temperature were arbitrarily fixed, etc.), there is potential for further J_g reduction using the UVO-RTA method. For instance, by

increasing the post-UVO RTA temperature to 500 °C, J_g was further reduced to 3.5 mA/cm² without any EOT penalty (Fig. 5.20(a)). However, a further increase to 600 °C results in an additional IL growth (from 0.3 nm to 0.5 nm), as evidenced in the TEM images in Fig. 5.18.

Fig. 5.20: (a) Average gate current density versus gate voltage (J_g - V_g) characteristics of the various TiN/HfO₂ gate stacks measured under room temperature (300 K). More than ten devices were measured for each split. A significant reduction in J_g is observed for the UVO-RTA sample (~ 8 mA/cm² @ $V_{FB} - 1$ V), further reduction is achieved by increasing the RTA temperature to 500 °C as shown by the UVO-RT500 °C sample (~ 3.5 mA/cm² @ $V_{FB} - 1$ V). (b) Gate current density J_g versus EOT at 300 K in comparison with the poly-Si/SiO₂ benchmark and the J_g reported in literature [174, 181].

An apparent positive shift in the V_{FB} of the UVO-RTA sample (denoted by solid line), relative to the rest of the annealed samples, can also be observed in Fig. 5.19, corresponding to the marked J_g reduction shown in Fig. 5.20(a). Studies have shown that the work function of the TiN/ HfO₂ gate stack can be increased via an O₂ enriching anneal [182], thus inducing a positive V_{FB} shift such as that shown. However, considering that the same UVO anneal conditions were applied, work

function increase is not likely the main reason for the additional positive V_{FB} shift of the UVO-RTA sample, as compared to the UVO-only and RTA-UVO samples. It is thus believed that the removal of V_o 's, which are known to be a major source of positive trapped charge in the HKMG stack, is the main reason, consistent with the results reported in [167]. Based on the extracted EOT and V_{FB} shift of +65 mV, the effective positive trapped density in this sample is $\sim 2 \times 10^{12} \text{ cm}^{-2}$ lower than the rest of the annealed samples.

The T_{BD} data of the gate stacks subjected to constant gate voltage (-2.2 V) stressing at 125°C (Fig. 5.21a)) show a trend consistent to that achieved in Section 5.3.3. The T_{BD} (@ 63rd percentile) of the UVO-RTA sample is ~ 2.5 , 6 and 9 times longer than that of the RTA-UVO, RTA-only and as-dep samples respectively. From the T_{BD} data in Fig. 5.12(a), however, the UVO-RTA sample exhibits a larger improvement margin, i.e. ~ 12 times longer T_{BD} than that of the as-dep sample. The larger improvement margin is attributed to the improved IL quality as evidenced by the SiO_2 IL EELS result (Fig. 5.17(b)). This shows that the two-step anneal method not only “repair” the V_o related defects in the high- κ , but also those present in the underlying IL. It is observed that the Weibull slopes of the respective distributions are almost similar ~ 1.5 , indicating that a similar mechanism governs the BD. It is believed that more new traps are required to be generated to form a percolation path in the UVO-RTA sample, hence resulting in a longer T_{BD} as compared with the other samples. Positive trapped charge generation in the various gate stacks under constant gate voltage stressing was also examined by monitoring the V_{FB} shift as a function of stress time (Fig. 5.21b)). Relative to the as-dep sample, the RTA-only sample shows the least reduction, followed by the UVO, RTA-UVO and UVO-RTA samples, consistent with that shown in Fig. 5.12(b). The longest T_{BD} and the smallest V_{FB} shift

of the UVO-RTA sample, therefore, confirm that this annealing sequence is most effective in reducing the defect precursors in the HKMG stack.

Fig. 5.21 (a) Time-to-breakdown T_{BD} distributions of the various TiN/HfO₂ gate stacks subjected to constant gate voltage (-2.2 V) stressing at 125 °C. The MOS capacitor area used in the TDDDB test is $20 \times 20 \mu\text{m}^2$. The Weibull slopes β of the respective distributions are ~ 1.5 . The UVO-RTA sample exhibits the longest T_{BD} with ~ 2.5 , 6 , and 9 times longer than that of the RTA-UVO, RTA-only, and as-dep samples respectively. (b) Flat band voltage shift ΔV_{FB} as a function of stress time at 125 °C. The applied electric field across the gate stack is estimated to be ~ 10 MV/cm. The slowest positive trap generation is observed for the UVO-RTA sample.

5.4 Summary

The TiN/HfZrO/SiO_x gate stack, formed via multistep deposition cum room-temperature UVO anneal, was examined by using STM and pulse C - V measurement and the results were compared to that of the as-dep and RTA (600 °C in N₂ for 30 s) samples. Evidence shows that a large part of the improvement seen in the multistep deposited cum UVO annealed sample, relative to the RTA sample, is a direct consequence of the suppressed crystallization of the HfZrO and hence a reduction in the density of the GB-related defects. Compared to the as-dep sample, the observed

improvement due to UVO annealing is marginal although further improvement, ascribed to the elimination of oxygen vacancy defects, may be achieved by either increasing the UVO anneal time after each deposition step or increasing the number of deposition steps. We propose a new method, via multistep high- κ deposition cum two-step annealing (comprising a RTUVO anneal and a subsequent RTA at 420°C) approach, for enhancing the performance and reliability of 7.5 Å EOT HfO₂/TiN gate stack in a high- κ last process. The UVO-RTA sample consistently exhibits the lowest gate current density J_g , positive oxide trapped charge generation and the longest T_{BD} as compared to the as-dep, UVO-only, RTA-only and RTA-UVO samples. The J_g obtained is among the lowest reported to date for high- κ last processes and it could be further reduced ~50 % by increasing the post-UVO RTA temperature to 500 °C without EOT penalty. The improvement is attributed to a more efficient incorporation of oxygen atoms (thereby eliminating the oxygen vacancies) into the HfO₂ film by the UVO-RTA method, as evidenced in the chemical structure analysis by EELS and XPS. The findings provide a promising means for realizing low-leakage and reliable sub-1 nm EOT HfO₂/TiN stacks for high- κ last integration.

Chapter 6 BIMODAL WEIBULL DISTRIBUTION OF HIGH-K/METAL GATE STACK – INSIGHTS BY STM

6.1 Introduction

Time-dependent dielectric breakdown (TDDB) is a crucial reliability test for assessing the lifetime of a gate dielectric film operating at nominal conditions. A high stress gate voltage (typically 0.5 V lower than the BD voltage, V_{BD}) is applied to a device at high temperature (typically 125 °C) in order to accelerate the trap generation rate in an oxide. BD occurs whenever an oxide generates sufficient traps to form a conducting path connecting the metal gate and the Si substrate, allowing huge electron/hole carriers flow through the oxide. BD is an irreversible process resulting from the thermal damage and microstructural defects originating from the high localized leakage current density and temperature. The time-to-breakdown, T_{BD} , data of a set of samples can be plotted as a failure distribution, with the distribution slope reflects the critical number of traps needed to be generated in order to construct a conducting path as predicated in the percolation BD model [183]. Compared with SiO₂/SiON, the TDDB results of high- κ devices yield interesting results: (i) a smaller Weibull slope, and (ii) a non Weibull distribution, which could not be explained by

using the existing percolation BD model. It was reported that two different Weibull slopes (instead of a single Weibull slope) present in an area-scaled TDDB distribution obtained from the high- κ gate stack BDs. Therefore, nano-characterization is important to resolve the microscopic changes occur in high- κ BD. In this chapter, we perform localized constant voltage stressing (CVS), via STM probe, to acquire the T_{BD} data at high- κ grain and GB respectively. Subsequently, we examine the correlation of the gate area size and the bimodal Weibull characteristic.

6.2 Experiments

A (2 nm) HfO_2 /(1 nm) SiO_2 bi-layer gate stack was formed on p-Si (100). The HfO_2 film and SiO_2 IL were grown via ALD and rapid thermal oxidation respectively. The SiO_2 IL was intentionally formed to ensure good interfacial properties with the Si substrate. The gate stack was subjected to PDA in N_2 at a temperature less than 700 °C for 30 s, followed by sputter-deposition of TiN film (corresponds to gate electrode) and ~1000 °C spike anneal in N_2 to mimic the source/drain dopant activation anneal. The TiN was then removed by wet etching. STM study was carried out in UHV ($\sim 10^{-10}$ Torr) under controlled temperature, this will eliminate the possible electro-deposition on the HfO_2 surface and minimize the thermal drift at the probe tip. The V_s was applied to the Si substrate and the STM probe was grounded. CCI and CITS were performed to acquire the topography and tunneling spectra of the high- κ stack respectively. CCI was acquired at a feedback setting of $V_s = 3$ V, $I_t = 25$ pA. To acquire the T_{BD} data, the probe was positioned over a leakage site and the feedback circuit was momentarily switched off to hold the probe at a fixed position for localized CVS.

6.3 Area-scaling Weibull distribution of high- κ /metal gate stack

Some research groups [27-28, 47] have reported the observation of two Weibull slopes in a single area-scaled TDDDB failure distribution obtained from the high- κ device BDs. The upper percentile regime of the distribution exhibits a shallow Weibull slope $\beta \sim 0.6$ and increases to ~ 1.5 towards the lower percentile regime. The shallow Weibull slope ($\beta \sim 0.6$) is typically associated with the small area device BDs whereas the steep Weibull slope ($\beta \sim 1.5$) is mainly related with the large area device BDs. Attempts to clarify the origin of this phenomenon have been made and various hypotheses have been proposed. For instance, Nigam *et al* [47] applied kinetic Monte Carlo to simulate a dual layer percolation model (to mimic a bi-layer high- κ /IL gate stack) to investigate the effect of the defect generation rate in high- κ and IL on the TDDDB failure distribution. The author ascribed the bimodal Weibull characteristic to a faster (by 3000 times) defect generation rate in the high- κ layer compared to that in the SiO_x IL. Ribes *et al* [27] argued that a high- κ BD would occur preferentially above the IL-initiated BD path and therefore its location may not correspond to the high- κ weakest point. Consequently, the failure distribution does not follow Poisson area scaling. It was reported in [46, 48] that a bimodal Weibull characteristic was observed in the failure distribution obtained from the polycrystalline HfO_2 device BDs but not from the amorphous HfSiON device BDs, suggesting that the high- κ phase change is involved in the non-Weibull characteristic. It is interesting to note that only a single Weibull slope is observed when the T_{BD} data collected from the relatively large area (i.e. $100 \mu\text{m}^2$ to $100000 \mu\text{m}^2$ rather than $0.006 \mu\text{m}^2$ to $0.6 \mu\text{m}^2$) devices were used for the area-scaled TDDDB plot (as shown in Fig. 2.7 (b)) [42]. The area-dependent

Weibull slope seems to point to the role of high- κ defects but the details are unclear since the data acquired from the standard characterization method is typically an average over an area.

In order to clarify the origin of the bimodal Weibull distribution, the microscopic details of high- κ BD need to be resolved. By using STM, we present a new phenomenological explanation for the bimodal Weibull distribution of the high- κ gate stack. Localized constant voltage stress is applied, via STM probe, at GB and grain separately in order to acquire the T_{BD} statistics at the respective locations. The relative influence of the GBs on the BD distribution of the high- κ gate stack in devices with different areas is discussed.

6.4 TDDB characteristic of grain and grain boundary

The topographical image of the HfO_2/IL gate stack, acquired via CCI mode, is shown in Fig. 6.1(a). The V_s applied to the Si substrate was 3 V and the I_t was 25 pA during the image acquisition. Under this bias condition, the I_t is primarily due to electrons tunneling from the STM probe. Therefore, the I_t is sensitive to the variations of the probe/high- κ tunneling barrier, which is associated with the high- κ structure [99, 127]. A change in the I_t when the STM probe was scanned across the high- κ surface was sensed by a feedback circuit which then instantaneously adjusted the vacuum gap between the probe and the high- κ surface to keep the I_t constant at the preset value. The fluctuations in the probe/high- κ gap spacing across the scanned area therefore give rise to the contrast in the STM topographical image. The bright shades generally correspond to the micro-protrusions whereas the dark shades correspond to the valleys on the high- κ surface [99, 104, 127]. The tunneling spectra of the HfO_2/IL gate stack

corresponding to the topographical image in Fig. 6.1(a) were extracted via CITS, which was carried out simultaneously with the CCI mode [99, 104, 127]. CITS essentially involves the measurement of the local I_t - V_s curve at all the positions on the scanned area with the feedback circuit turned off momentarily to hold the STM probe at a fixed distance from the high- κ surface. The tunneling spectra (at a given V_s) at all the locations on the HfO_2 surface can be mapped to form a current image as shown in Fig. 6.1(b). The contrast in the current image denotes the variations of the I_t across the high- κ gate stack. The bright shades correspond to the higher leakage sites in relative to the other locations in the scanned area. A change in the I_t caused by the presence of structural defects (i.e. GBs) can be distinguished via reference to the current image.

An “inverted contrast” of the CCI and the CITS images clearly shows that the (bright regions) granular features (diameter ~ 20 nm) in the topographical image (Fig. 6.1(a)) arise due to the high- κ physical protrusions instead of the presence of electronic traps, implying that the HfO_2 layer is polycrystalline. This is supported by

Fig. 6.1: (a) $50 \times 50 \text{ nm}^2$ constant-current STM image of the $\text{HfO}_2/\text{SiO}_x$ gate stack at a set point of $V_s = 3 \text{ V}$ and $I_t = 25 \text{ pA}$. Under positive V_s , the polycrystalline HfO_2 layer is imaged, as verified by the granular features (bright regions of diameter ~ 20 nm) in the image. (b) Corresponding current image (under substrate injection bias at $V_s = -3 \text{ V}$) showing (bright shades) “leaky” GBs encompassing the less conductive grains.

the current map in Fig. 6.1(b), which shows that the (brighter circumference) conductive GBs encompassing the (dark shades) less conductive granular features (i.e. grains) [127]. Study shows that the oxygen vacancies preferentially segregate near a GB and the concentration can be as high as $2.5 \times 10^{14} \text{ cm}^{-2}$. Therefore, oxygen vacancy is believed to be the main source of electronic traps responsible for an increased I_t at GBs [138]. Due to a higher local I_t , one may expect a faster wear-out process at GBs (i.e. they are the potential BD locations under gate stressing).

To examine the BD strength of GBs in the $\text{HfO}_2/\text{SiO}_x$ gate stack, we applied CVS at the GBs via the STM probe. The closed-loop feedback circuit was momentarily disabled so that the probe was kept at a constant height from the high- κ surface irrespective of an increase in the I_t resulting from the local oxide degradation under stress. The V_s was increased and kept constant at 4 V. This simulates a CVS typically applied on the MOS structure, except for a very small “gate area” of several tens of $\text{\AA}^2$ in this case. The voltage across the gate stack was estimated to be 2 V, based on a gate stack capacitance of $1.9 \text{ }\mu\text{F}/\text{cm}^2$ and a vacuum gap of 5 $\text{\AA}$. The evolution of the I_t until BD was monitored by an oscilloscope (attached to the current output) throughout the stress. For comparison, a similar CVS was also applied on the grains to assess the degradation rate at both locations. Fig. 6.2(a) shows the evolution of the I_t until BD for both GB and grain. In this experiment, the maximum current measurable by the STM probe is 10 nA due to the current compliance of IVP-300 pre-amplifier. As expected, the T_{BD} for the local stressing applied at the GB is considerably shorter than that for the grain. This clearly shows that the pre-existing defects at the high- κ GBs have an important impact on the corresponding TDDB distribution.

The T_{BD} statistics of grain and GB BDs are spatially resolved and plotted as individual Weibull distributions (Fig. 6.2(b)). The T_{BD} at the GB is about two orders of

magnitude shorter than that at grain for all the positions examined. As discussed in Chapter 4, the IL beneath a GB is highly susceptible to degradation under stress. This result further confirms that the final BD of a high- κ gate stack is determined by the “weakest link” assisted by GBs [64]. Since a critical number of defect density has to be reached in order to induce an oxide BD (based on the percolation BD model), the higher pre-existing defects density at a GB means that a generation of fewer traps is needed to form the BD percolation path. This means a shorter stress time is needed to reach the critical defect density. Moreover, the “voltage loading” at the IL beneath a low resistive GB is expected to be higher [127]. These effects thus result in a much shorter T_{BD} at GBs.

Fig. 6.2: (a) Evolution of the tunneling current I_t under localized constant voltage stressing at 4 V, applied via the STM probe, on GB and grain respectively (Fig. 6.1). As expected, the gate stack BD occurs much earlier for the stressing applied at the GB than that applied at the grain. (b) Weibull distributions of the T_{BD} for the stressing applied at the GBs ($\beta = 1.45$) and grains ($\beta = 1.36$). T_{BD} is defined as the time when the I_t increases abruptly to the compliance setting (10 nA).

A further analysis on the respective failure distributions of GB and grain yields an interesting result. It is observed that both T_{BD} distributions exhibit almost the same Weibull slope in spite of an initial difference in the defect density. On one hand, this result is expected, since the same number of traps is needed to form a conducting path to induce a BD, regardless of GB or grain for a given physical gate stack thickness. On the other hand, there is no obvious indication that in the case of the GB, the underlying ultrathin IL is “controlling” the gate stack BD [47, 64]. Kinetic Monte Carlo (kMC) simulations based on a dual layer percolation model [47] has shown that in the presence of a high defect density in the high- κ layer, the Weibull slope for T_{BD} would be reduced since it is the IL which determines the overall gate stack BD. This prediction is, however, not borne out by the result in Fig. 6.2(b). A possible reason is that the quality of the IL is already compromised by the GB above, i.e. both layers degrade at comparable rates. Simulation [64-65, 113] has shown that the density of oxygen vacancies in the underlying IL tends to be higher due to the presence of oxygen deficient GBs. For a given large area device, compared with the Weibull distribution of the amorphous gate stack, the polycrystalline gate stack is expected to exhibit (i) similar Weibull slope; since the same number of traps is needed to induce BD regardless of the dielectric film state, and (ii) shorter T_{BD} ; since GB BD dominates the failure distribution. However, the area-scaled Weibull distribution of the polycrystalline gate stack exhibits two Weibull slopes, rather than one Weibull slope shown by the amorphous gate stack, in a single TDDB failure distribution due to the presence of grain and GB BD which governs the failure distribution of small and large area device respectively.

6.5 Bimodal Weibull distribution

The failure distributions at the respective locations of GB and grain are spatially resolved using nanoscale CVS via the STM probe. It is clearly shown that the T_{BD} at GB is indeed shorter than that at grain, implying that the impact of the GB on the failure distribution is non-negligible. The two different slopes observed in the bimodal Weibull distribution are strongly associated with the gate area size. However, it is unclear the origin of this phenomenon since the standard characterization metrology is unable to resolve the inhomogeneity issue in the defects density of the high- κ gate stack. In this section, we investigate the relationship between the GB-related defects density and the gate area size and its correlation to the bimodal Weibull characteristic.

6.5.1 Correlation of the GB-related defect density and gate area size

To study the relative influence of grain and GB BD on the T_{BD} of the high- κ gate stack as a function of area scaling, we divide a $200 \times 200 \text{ nm}^2$ STM current map into smaller regions (Fig. 6.3). The number of regions having a certain percentage of the area occupied by grain is depicted in Fig. 6.4. An analysis on the $200 \times 200 \text{ nm}^2$ region reveals that about 75 % of the area is occupied by grain. However, as we divide this area into smaller regions, it can be observed that some of the regions now have $> 75 \%$ of the area occupied by grain, even though the area is still bigger than the grain size. For instance, three out of four $100 \times 100 \text{ nm}^2$ regions examined show at least 75 % and above of the area are occupied by grain, thirteen out of sixteen $50 \times 50 \text{ nm}^2$ regions examined show at least 75 % and above of the area are occupied by grain. The number of regions having almost the entire area occupied by grain increases substantially as the area of the region approaches the grain size. This means that grain BD would

become increasingly important as the gate area is scaled down. However, it should be noted that there is still a non-negligible number of regions with $> 50\%$ of the area occupied by GBs, particularly when the size of a region is scaled down to $20 \times 20 \text{ nm}^2$ and below. Therefore, the T_{BD} statistics of the small area devices would be determined by the BDs at both grain and GBs.

Fig. 6.3: $200 \times 200 \text{ nm}^2$ STM current map (at $V_s = -3 \text{ V}$) of the $\text{HfO}_2/\text{SiO}_x$ gate stack acquired via constant imaging tunneling spectroscopy (CITS). The image is divided into smaller regions to analyze the percentage of the area occupied by (bright shades in the current map) GBs as a function of gate area size.

Fig. 6.4: Number of “devices” (obtained by dividing the current map in Fig. 6.3 into smaller regions) having a certain percentage of the gate area occupied by grain. As the gate area size approaches $20 \times 20 \text{ nm}^2$, the grain BD becomes increasingly important as the number of regions having almost the entire area occupied by grain increases substantially. At the same time, it is observed that there are a small number of regions with $> 50 \%$ of the area occupied by GBs, implying the co-existence of the GB and grain BD influence on the T_{BD} statistics of the small area devices.

6.5.2 The role of GB in bimodal Weibull distribution

To illustrate the impact of the mixture of grain and GB BD modes on the T_{BD} distribution of the small area devices, we consider an example based on Fig. 6.3 and Fig. 6.4, where 75 % of the gate stack area is occupied by grain. One would expect the fraction of regions having the entire area occupied by grain to reach 75 % as the $200 \times 200 \text{ nm}^2$ current map in Fig. 6.3 is further sub-divided into the smaller regions. Based on the Weibull distribution for grain and GB BDs in Fig. 6.2(b), a combined T_{BD} distribution for 75 % grain and 25 % GB BDs can be obtained as shown in Fig. 6.5(a).

Fig. 6.5: (a) Example of a T_{BD} distribution due to a mixture of grain and GB BDs in the small area devices. The distribution is derived from the Weibull distribution for grain and GB BDs in Fig. 6.2(b), by taking the percentage of the grain and GB BDs to be 75 % and 25 %, respectively (cf. Fig. 6.4). (b) Combination of the two T_{BD} distributions in Fig. 6.2(b) yields a bimodal Weibull distribution, with the steep (shallow) Weibull slope at the lower (upper) percentile predominated by the shorter (longer) T_{BD} at the GBs (grains).

Fig. 6.6: Weibull distributions of the T_{BD} for the enhanced field (at $V_s = 4.5$ V) accelerated stressing applied at the GBs (filled triangle) and grains (open triangle). T_{BD} is defined as the time when the I_t increases abruptly to the compliance setting (10 nA).

The Weibull slope of this combined T_{BD} distribution is evidently reduced (< 1) as compared to that of the Weibull distribution for the grain BD (~ 1.36) in Fig. 6.2(b). In addition, the GB BDs introduce a “tail” having a higher slope, similar to that seen in the experimental T_{BD} distribution for the small area devices [28, 47]. The result implies that it may not be possible to observe the intrinsic grain BD distribution even at device areas comparable to the average grain size, since the early failures due to GB BD in a fraction of these devices would decrease the slope of the overall distribution. It should be mentioned that the relative influence of grain and GB BDs on the combined T_{BD} distribution of the small area devices may vary with the gate stress voltages because of the potentially different voltage accelerations of grain and GB BDs. This prediction is borne out in the T_{BD} distributions (Fig. 6.6) of the grain and GB BDs obtained under relatively higher voltage (at 4.5 V) stressing. It is observed that the Weibull slopes shown by the grain BDs obtained under 4 V and 4.5 V stressing are almost similar (~ 1.36 and ~ 1.40 respectively). In contrast, a reduced (relative to that stressed at 4 V) β is observed for the GB stressed at 4.5 V. More importantly, the T_{BD} distribution of the GB (stressed at 4.5 V) shifts evidently to the smaller T_{BD} whereas the grain (stressed at 4.5 V) shows a less significant shift compared to their respective counterparts stressed at 4 V. Therefore, a further reduction in the Weibull slope of the combined T_{BD} distribution is estimated.

Based on the above results, we provide a phenomenological explanation for the bimodal Weibull distribution that resulting from an area-scaling [27-28, 47]. A grain diameter of ~ 20 nm (Fig. 6.1) implies that many GBs would exist in the high- κ dielectric of the large area devices ($\sim 0.6 \mu\text{m}^2$ [27]). The fact that BD occurs preferentially at GBs (as evidenced in Fig. 6.2) means that this would be the dominant BD mode in the large area devices, i.e. it is not likely to observe grain BD since the

T_{BD} at grain is about two orders of magnitude longer than that at GB. This explains why the large area devices account for the lower percentile regime of the area-scaled T_{BD} distribution [27-28, 47]. The Weibull slope of 1.45 of the T_{BD} distribution for GB BD (Fig. 6.2(b)) is in general agreement with those observed on the large area devices (~ 1 -2) having the gate stacks of comparable thicknesses [27-28, 40, 42, 47]. In the small area devices, however, the ratio of the area occupied by grains to the area occupied by GBs is increased (Fig. 6.4), i.e. the relative importance of the GB-related BD is expected to be reduced. The T_{BD} statistics in this case would be determined by the BDs at both GBs and grains. The former is likely to increase the fraction of the early failures, causing the Weibull slope of the combined distribution to decrease with respect to that due purely to BD at grains. The T_{BD} distribution in Fig. 6.5(b), obtained by combining the individual distributions of GBs and grains in Fig. 6.2(b), lends a strong support for the inference. A change of the Weibull slope from 1.43 to 0.61, corresponding to a transition from a regime dominated by the GB-related BDs to one determined by both GB- and grain-related BDs, is clearly observed.

6.6 Summary

Local electrical stress, applied via the STM probe, on GBs and grains of a high- κ gate stack, allows the individual T_{BD} distributions to be spatially resolved. The T_{BD} at GB is about two orders of magnitude shorter than that at grain, implying that the “weakest link” that determines the final gate stack BD is correlated to the high- κ GBs. Both T_{BD} distributions exhibit almost similar Weibull slopes in spite of an initial difference in defect density. Based on the results, an explanation for the bimodal Weibull distribution resulting from an area-scaling is proposed. It is shown that (i) GB-related BD determines the steep Weibull slope at the lower percentile of the T_{BD}

distribution (corresponds to the T_{BD} distribution of the large area devices); (ii) the combined effect of grain- and GB-related BDs determines the shallow Weibull slope at the upper percentile of the T_{BD} distribution (corresponds to the T_{BD} distribution of the small area devices in which the relative importance of the GB-related BDs is reduced compared to the large area devices).

Chapter 7 CONCLUSION & RECOMMENDATION

7.1 Conclusion

A systematic reliability study, comprising microscopic (via STM) and device level electrical characterizations as well as physical analysis (via TEM, EELS and XPS), of ultrathin (< 5 nm) high- κ dielectric provides useful insights into the performance and reliability of advanced high- κ MOSFETs. The key findings achieved in this research are summarized as follows:-

- STM nanoscale characterization of the $\text{HfO}_2/\text{SiO}_x$ gate stack reveals that the degradation and BD of ultrathin SiO_x IL depends on the quality of the thicker high- κ film. Observation shows that the presence of pre-existing or stress-induced leakage sites in the high- κ film results in an evident degradation of the underlying IL (as indicated by a significant increase in the local tunneling spectra under substrate injection bias).
- Most of the high- κ leakage sites preferentially concentrate near GBs in a polycrystalline high- κ film. During stress, GBs present low resistive paths propagating through the high- κ layer, thereby causing a large voltage drop across the underlying IL. Since the IL region adjacent to GB is highly susceptible to degradation, this location is a preferential BD spot of the entire gate stack.

- The T_{BD} statistics at the respective locations of GBs and grains were acquired from localized constant gate voltage stressing via STM probe. The TDDB distributions of GB and grain show almost similar Weibull slopes in spite of an initial difference in the defects density, suggesting that the critical number of traps needed to induce BD is identical, regardless at GBs or grains for a given physical gate stack thickness.
- The multistep deposition-cum-annealing (D&A) method could significantly improve the thermal stability of a high- κ film. A detailed STM study shows that a HfSiO_x gate stack, formed via multistep D&A method, retains a nanocrystalline structure (rather than transforms into polycrystalline phase) even after subjected to high temperature (1000 °C) anneal. The gate stack exhibits significant improvements in the electrical uniformity and BD strength, which are attributed to the reduced probability (due to smaller grains) of a single GB propagating through the entire high- κ layer, thus suppressing the formation of high- κ conductive paths. The STM results are in good agreement with the device measurements.
- The STM and pulse $C-V$ measurements show that a reduced J_g in the multistep-deposited and UVO-annealed HfZrO gate stack, relative to the RTA(600 °C) sample, is primarily a direct consequence of the reduced GB-related defects density by virtue of low-thermal-budget. Compared with the as-dep sample, however, an increasing number of deposition-cum-UVO annealing step can further improve the performance and reliability of a gate stack due to a greater V_O density reduction in a thinner high- κ film deposited in each step.
- A practical two-step annealing (comprising a RTUVO anneal and a subsequent RTA at 420 °C in N_2) method for enhancing the performance and reliability of

a gate-last 7.5 Å EOT HfO₂/TiN gate stack is demonstrated. The two-step annealed gate stack consistently shows a reduced J_g , positive oxide trapped charge generation and a longer T_{BD} as compared with the as-dep and the single-step annealed gate stacks. The subsequent RTA step is believed not only drives off carbon impurities, but also provides thermal activation for an efficient high- κ structural “repair” by the absorbed O₃. An increase in the RTA temperature to 500 °C could further reduce the J_g without EOT penalty.

- Bimodal Weibull distribution, resulting from the area-scaled TDDB distributions of the high- κ device BD, is investigated at nanoscale level. Localized TDDB tests at GBs and grains reveal that: (i) the steep Weibull slope at the lower percentile of the T_{BD} distribution (corresponds to the T_{BD} distribution of the large-area devices) is primarily governed by the GB-related BD, and (ii) the shallow Weibull slope at the upper percentile of the T_{BD} distribution (corresponds to the T_{BD} distribution of the small-area devices) is a result of the mixture mode of the grain- and GB-related BDs.

7.2 Recommendations

Unlike the SiO₂ gate dielectric, which has been extensively studied for the past 4 decades, the application of high- κ metal oxide as a transistor’s gate dielectric is deemed new and therefore the fundamental understanding of the high- κ reliability physics is still lacking. This work could be further extended to enhance the understanding of the high- κ reliability physics. Possible future studies are listed as follows:-

- First-principle simulations reported that V_o 's tend to segregate along a GB in a polycrystalline high- κ due to a lower diffusion barrier at the low resistive path [64]. Our characterization work, via STM, provides direct physical evidence that the tunneling leakage current preferentially concentrates at GB. It would be interesting to subject a polycrystalline high- κ film to room-temperature UVO annealing to "repair" the V_o 's. The changes in the electrical characteristics (e.g. I - V , TDDB, etc) and chemical structures of the UVO-annealed GB can be investigated via electrical measurement, STM, EELS and XPS respectively.
- The TDDB failure distributions, acquired individually at GB and grain via localized constant gate voltage stressing, exhibit almost similar Weibull slopes in spite of an initial different defect density. Although this observation can be explained by using percolation BD model, the inference can be further consolidated by performing nanoscale TDDB test over GB and grain of the high- κ film of different thicknesses.
- The multistep deposition cum two-step annealing method shows a promising means in improving the performance and reliability (i.e. reduced J_g and enhanced BD strength) of high- κ last gate stack. Investigating the engineered high- κ stack from NBTI/PBTI perspective offers useful insights into the interface/bulk trap generation and recovery characteristic. In addition, the correlation of the TDDB and NBTI/PBTI data can be examined since both types of degradation mechanisms are induced by oxide trap generation.
- For the gate stack formed via multistep deposition cum two-step annealing, the lowest J_g (without EOT penalty) is achieved when the RTA is increased to 500 °C. Further increase in the temperature to 600 °C, however, results in an undesirable IL re-growth. This method can be further optimized by fine-tuning the preceding

UVO annealing recipe. For instance, the UVO annealing duration can be prolonged and the ozone concentration can be adjusted in order to improve the oxygen incorporation efficiency.

- Studies [138] have proposed the association of resistive switching (RS) characteristic to the presence of high- κ GBs. However, the role of GBs in the RS mechanism is unclear since the standard characterization tools could not resolve the microscopic changes in high- κ film. The correlation of GBs and the RS mechanism can be investigated at nanoscale level by direct probing at GB via STM.

LIST OF PUBLICATIONS

JOURNAL PAPERS

- [1] **K. S. Yew**, D. S. Ang, L. J. Tang, K. Cui, G. Bersuker, and P. S. Lysaght, “Scanning tunneling microscopy study of the multi-step deposited and annealed HfSiO_x gate dielectric”, Journal of The Electrochemical Society, vol. 158, no. 10, pp. H1021-H1026, Aug. 2011.
- [2] **K. S. Yew**, and D. S. Ang, “Study of the multi-step deposited and UV-Ozone annealed HfZrO gate stack by scanning tunneling microscopy and pulse *C-V* measurement”, IEEE Transactions on Electron Devices, vol. 59, no. 8, pp. 2268-2272, Aug. 2012.
- [3] **K. S. Yew**, D. S. Ang, and G. Bersuker, “Bimodal weibull distribution of metal/high- κ gate stack TDDb – insights by scanning tunneling microscopy”, IEEE Electron Device Letter, vol. 32, no. 2, pp. 146 -148, Feb. 2012.
- [4] **K. S. Yew**, D. S. Ang, and L. J. Tang, “A new method for enhancing high- κ /metal-gate stack performance and reliability for high- κ last integration”, IEEE Electron Device Letter, vol. 34, no. 2, pp. 295-297, Feb 2013.
- [5] **K. S. Yew**, D. S. Ang, C. H. Tang, L. J. Tang, and J. S. Pan, “Study of the electrical and chemical properties of the multi-step deposited and two-step (ultraviolet ozone cum rapid thermal) annealed HfO₂ gate stack”, IEEE Transactions on Electron Devices, *to be submitted*.
- [6] L. Wu, H. Y. Yu, **K. S. Yew**, J. S. Pan, D. S. Ang, W. J. Liu, T. T. Le, T. L. Duan, C. H. Hou, X. F. Yu, D. Y. Lee, K. Y. Hsu, J. Xu, H. J. Tao and M. Cao, “Multideposition multiroom-temperature annealing via ultraviolet-ozone for HfZrO high- κ and integration with TiN metal gate in a gate-last process”, IEEE Transactions on Electron Devices, vol. 58, no. 7, pp. 2177-2181, Jul. 2011.
- [7] Y. H. Tan, **K. S. Yew**, K. H. Lee, Y. J. Chang, K. N. Chen, D. S. Ang, E. Fitzgerald, and C. S. Tan, “Al₂O₃ interface engineering of germanium epitaxial layer grown directly on silicon”, IEEE Transactions on Electron Devices, vol. 60, no. 1, pp. 56-62, Jan 2013.

INTERNATIONAL CONFERENCES

- [1] **K. S. Yew**, Y. C. Ong, D. S. Ang, K. L. Pey, G. Bersuker, P. S. Lysaght, and D. Heh, “Nannoscle characterization of $\text{HfO}_2/\text{SiO}_x$ gate stack degradation by scanning tunneling microscopy”, in Proc. 2009 International Conference on Solid State Devices and Materials, pp. 314-315. (Japan).
- [2] **K. S. Yew**, D. S. Ang, K. L. Pey, G. Bersuker, P. S. Lysaght, and D. Heh, “Enhanced electrical uniformity and breakdown of multi-step deposited and annealed HfSiO – insight by scanning tunneling microscopy”, in Proc. 2010 International Conference on Solid State Devices and Materials, pp. 673-674. (Japan).
- [3] **K. S. Yew**, C. H. Tang, and D. S. Ang, “Multi-step deposition and low-temperature two-step (ultraviolet ozone cum rapid thermal) annealing as a promising means for gate-last high- κ /metal gate application”, in Proc. 2012 International Conference on Solid State Devices and Materials. pp. 32-33. (Japan).
- [4] **K. S. Yew**, C. H. Tang, and D. S. Ang, “Chemical analysis of multi-step deposited and two-step (ultraviolet ozone cum rapid thermal) annealed sub-1 nm EOT HfO_2/TiN gate stack for high- κ last integration”, in Proc. 2013 International Conference on Solid State Devices and Materials, *accepted*. (Japan).
- [5] L. Wu, **K. S. Yew**, D. S. Ang, W. J. Liu, T. T. Le, T. L. Duan, C. H. Hou, X. F. Yu, D. Y. Lee, K. Y. Hsu, J. Xu, H. J. Tao, M. Cao, and H. Y. Yu, “Novel multi deposition multi room-temperature annealing technique via ultraviolet-ozone to improve high- κ /metal (HfZrO/TiN) gate stack integrity for a gate-last process”, in IEDM Tech. Dig., 2010, pp. 273 -276
- [6] G. Bersuker, D. Heh, C. D. Young, I. Morassi, A. Padovani, L. Larcher, **K. S. Yew**, Y. C. Ong, D. S. Ang, K. L. Pey, and W. Taylor, “Mechanism of high- κ dielectric-induced breakdown of interfacial SiO_2 layer”, in Proc. 2010 International Reliability Physics Symposium, pp. 373-378. (USA).

BIBLIOGRAPHIES

- [1]. G. E. Moore, "Cramming more components onto integrated circuits," in *Proceedings of the IEEE*, vol. **86**, no. 1, pp. 82-85, 1998
- [2]. J. Robertson, "High dielectric constant gate oxides for metal oxide Si transistors," *Reports on Progress in Physics*, vol. **69**, no. 2, pp. 327-396, 2006
- [3]. Wikipedia. http://en.wikipedia.org/wiki/Transistor_count last modified on 7 March 2013.
- [4]. A. B. Joshi, D. L. Kwong, and S. Lee, "Improvement in performance and degradation characteristics of MOSFETs with thin gate oxides grown at high temperature," *IEEE Electron Device Letters*, vol. **12**, no. 1, pp. 28-30, 1991
- [5]. M. Depas, B. Vermeire, P. W. Mertens, R. L. Vanmeirhaeghe, and M. M. Heyns, "Determination of tunneling parameters in ultra-thin oxide layer Poly-Si/SiO₂/Si structures, " *Solid-State Electronics*, vol. **38**, no. 8, pp. 1465-1471, 1995
- [6]. S. H. Lo, D. A. Buchanan, Y. Taur, and W. Wang, "Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFET's," *IEEE Electron Device Letters*, vol. **18**, no. 5, pp. 209-211, 1997
- [7]. R. M. Wallace, and G. D. Wilk, "High- κ dielectric materials for microelectronics," *Critical Reviews in Solid State and Materials Sciences*, vol. **28**, no. 4, pp. 231-285, 2003
- [8]. J. H. Stathis, "Reliability limits for the gate insulator in CMOS technology," *IBM Journal of Research and Development*, vol. **46**, no. 2-3, pp. 265-286, 2002
- [9]. D. Adler, "Quantum-mechanics for applied physics and engineering," *Physics Today*, vol. **36**, no. 11, pp. 76, 1983
- [10]. G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High- κ gate dielectrics: Current status and materials properties considerations," *Journal of Applied Physics*, vol. **89**, no. 10, pp. 5243-5275, 2001
- [11]. D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, "The electronic structure at the atomic scale of ultrathin gate oxides," *Nature*, vol. **399**, no. 6738, pp. 758-761, 1999
- [12]. D. J. Dimaria, and E. Cartier, "Mechanism for stress-induced leakage currents in thin silicon dioxide films," *Journal of Applied Physics*, vol. **78**, no. 6, pp. 3883-3894, 1995

-
- [13]. D. J. Dimaria, E. Cartier, and D. Arnold, "Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon," *Journal of Applied Physics*, vol. **73**, no. 7, pp. 3367-3384, 1993
- [14]. E. Harari, "Dielectric-breakdown in electrically stressed thin-films of thermal SiO₂," *Journal of Applied Physics*, vol. **49**, no. 4, pp. 2478-2489, 1978
- [15]. M. Houssa. "High- κ gate dielectrics," London: Institute of Physics, 2003
- [16]. H. R. Huff, and D. C. Gilmer. "High dielectric constant materials," Springer, 2005
- [17]. E. P. Gusev, D. A. Buchanan, E. Cartier, A. Kumar, D. DiMaria, S. Guha, A. Callegari, S. Zafar, P. C. Jamison, D. A. Neumayer, M. Copel, M. A. Gribelyuk, H. Okorn-Schmidt, C. D'Emic, P. Kozlowski, K. Chan, N. Bojarczuk, L. A. Ragnarsson, P. Ronsheim, K. Rim, R. J. Fleming, A. Mocuta, and A. Ajmera, "Ultrathin high- κ gate stacks for advanced CMOS devices," in *2001 IEEE International Electron Devices Meeting (IEDM 2001)*, pp. 20.1.1-20.1.4
- [18]. K. Shubhakar, K. L. Pey, S. S. Kushvaha, S. J. O'Shea, M. Bosman, M. Kouda, K. Kakushima, and H. Iwai, "Localized degradation and breakdown study of cerium-oxide high- κ ; gate dielectric material using scanning tunneling microscopy," in *Proceedings of the 17th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA 2010)*, pp. 1-5
- [19]. H. Iwai, S. Ohmi, S. Akama, C. Ohshima, A. Kikuchi, I. Kashiwagi, J. Taguchi, H. Yamamoto, J. Tonotani, Y. Kim, I. Ueda, A. Kuriyama, Y. Yoshihara, and I. Electronic Devices Society Of, "Advanced gate dielectric materials for sub-100 nm CMOS," in *2002 IEEE International Electron Devices Meeting (IEDM 2002)*, pp. 625-628
- [20]. E. P. Gusev, E. Cartier, D. A. Buchanan, M. Gribelyuk, M. Copel, H. Okorn-Schmidt, and C. D'Emic, "Ultrathin high- κ metal oxides on silicon: processing, characterization and integration issues," *Microelectronic Engineering*, vol. **59**, no. 1-4, pp. 341-349, 2001
- [21]. G. Vellianitis, J. Petry, J. C. Hooker, A. Delabie, and S. De Gendt, "Reliability degradation of HfSiO gate dielectric layers: influence of nitridation," *Microelectronic Engineering*, vol. **84**, no. 9-10, pp. 1972-1975, 2007
- [22]. W. J. Zhu, T. P. Ma, S. Zafar, and T. Tamagawa, "Charge trapping in ultrathin hafnium oxide," *IEEE Electron Device Letters*, vol. **23**, no. 10, pp. 597-599, 2002
- [23]. H. Kim, P. C. McIntyre, and K. C. Saraswat, "Effects of crystallization on the electrical properties of ultrathin HfO₂ dielectrics grown by atomic layer deposition," *Applied Physics Letters*, vol. **82**, no. 1, pp. 106-108, 2003
- [24]. C. M. McGilvery, S. McFadzean, M. MacKenzie, F. T. Docherty, A. J. Craven, D. W. McComb, and S. De Gendt, "Nucleation, crystallisation and phase

- segregation in HfO_2 and HfSiO_x ," in *Microscopy of Semiconducting Materials 2007*, Springer-Verlag Berlin: Berlin. pp. 325-328
- [25]. C. J. Forst, C. R. Ashman, K. Schwarz, and P. E. Blochl, "The interface between silicon and a high- κ oxide," *Nature*, vol. **427**, no. 6969, pp. 53-56, 2004
- [26]. M. Houssa, M. Naili, C. Zhao, H. Bender, M. M. Heyns, and A. Stesmans, "Effect of O_2 post-deposition anneals on the properties of ultra-thin $\text{SiO}_x/\text{ZrO}_2$ gate dielectric stacks," *Semiconductor Science and Technology*, vol. **16**, no. 1, pp. 31-38, 2001
- [27]. G. Ribes, P. Mora, F. Monsieur, M. Rafik, F. Guarin, G. Yang, D. Roy, W. L. Chang, and J. Stathis, "High- κ gate stack breakdown statistics modeled by correlated interfacial layer and high- κ breakdown path," in *2010 IEEE International Reliability Physics Symposium (IRPS 2010)*, pp. 364-368
- [28]. A. Kerber, E. Cartier, B. P. Linder, S. A. Krishnan, and T. Nigam, "TDDDB failure distribution of metal gate/high- κ CMOS devices on SOI substrates," in *2009 IEEE International Reliability Physics Symposium (IRPS 2009)*, pp. 505-509
- [29]. D. H. Triyoso, M. Ramon, R. I. Hegde, D. Roan, R. Garcia, J. Baker, X. D. Wang, P. Fejes, B. E. White, and P. J. Tobin, "Physical and electrical characteristics of HfO_2 gate dielectrics deposited by ALD and MOCVD," *Journal of the Electrochemical Society*, vol. **152**, no. 3, pp. G203-G209, 2005
- [30]. "Savannah 100 & 200 Atomic Layer Deposition System User Manual," *Cambridge NanoTech Inc.*, 2007
- [31]. G. Bersuker, B. H. Lee, and H. R. Huff, "Novel dielectric materials for future transistor generations," *Frontiers in Electronics*, vol. **41**, pp. 221-239, 2006
- [32]. D. Triyoso, R. Liu, D. Roan, M. Ramon, N. V. Edwards, R. Gregory, D. Werho, J. Kulik, G. Tam, E. Irwin, X. D. Wang, L. B. La, C. Hobbs, R. Garcia, J. Baker, B. E. White, and P. Tobin, "Impact of deposition and annealing temperature on material and electrical characteristics of ALD HfO_2 ," *Journal of the Electrochemical Society*, vol. **151**, no. 10, pp. F220-F227, 2004
- [33]. S. Ferrari, and G. Scarel, "Oxygen diffusion in atomic layer deposited ZrO_2 and HfO_2 thin films on $\text{Si}(100)$," *Journal of Applied Physics*, vol. **96**, no. 1, pp. 144-149, 2004
- [34]. N. Miyata, M. Ichikawa, T. Nabatame, T. Horikawa, and A. Toriumi, "Thermal stability of a thin HfO_2 /ultrathin SiO_2 /Si structure: Interfacial Si oxidation and silicidation," *Japanese Journal of Applied Physics Part 2-Letters*, vol. **42**, no. 2B, pp. L138-L140, 2003
- [35]. T. L. Duan, H. Y. Yu, L. Wu, Z. R. Wang, Y. L. Foo, and J. S. Pan, "Investigation of HfO_2 high- κ dielectrics electronic structure on SiO_2/Si

- substrate by x-ray photoelectron spectroscopy," *Applied Physics Letters*, vol. **99**, no. 1, pp. 2011
- [36]. L. A. Ragnarsson, L. Pantisano, V. Kaushik, S. I. Saito, Y. Shimamoto, S. De Gendt, and M. Heyns, "The impact of sub monolayers of HfO₂ on the device performance of high- κ based transistors," in *2003 IEEE International Electron Devices Meeting (IEDM 2003)*, no. pp. 87-90, 2003
 - [37]. M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, "Effective electron mobility in Si inversion layers in metal-oxide-semiconductor systems with a high- κ insulator: The role of remote phonon scattering," *Journal of Applied Physics*, vol. **90**, no. 9, pp. 4587-4608, 2001
 - [38]. S. Saito, D. Hisamoto, S. Kimura, and M. Hiratani, "Unified mobility model for high- κ ; gate stacks [MISFETs]," in *2003 IEEE International Electron Devices Meeting (IEDM 2003)*, pp. 33.3.1-33.3.4
 - [39]. R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, "A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides," in *1995 IEEE International Electron Devices Meeting (IEDM 1995)*, pp. 863-866
 - [40]. Y. H. Kim, K. Onishi, C. S. Kang, H. J. Cho, R. Choi, S. Krishnan, M. S. Akbar, and J. C. Lee, "Thickness dependence of Weibull slopes of HfO₂ gate dielectrics," *IEEE Electron Device Letters*, vol. **24**, no. 1, pp. 40-42, 2003
 - [41]. M. Houssa, L. Pantisano, L. A. Ragnarsson, R. Degraeve, T. Schram, G. Pourtois, S. De Gendt, G. Groeseneken, and M. M. Heyns, "Electrical properties of high- κ gate dielectrics: Challenges, current issues, and possible solutions," *Materials Science & Engineering R-Reports*, vol. **51**, no. 4-6, pp. 37-85, 2006
 - [42]. T. Kauerauf, R. Degraeve, E. Cartier, C. Soens, and G. Groeseneken, "Low Weibull slope of breakdown distributions in high- κ layers," *IEEE Electron Device Letters*, vol. **23**, no. 4, pp. 215-217, 2002
 - [43]. R. Degraeve, A. Kerber, P. Roussel, E. Cartier, T. Kauerauf, L. Pantisano, and G. Groeseneken, "Effect of bulk trap density on HfO₂ reliability and yield," in *2003 IEEE International Electron Devices Meeting (IEDM 2003)*, pp. 935-938
 - [44]. G. Bersuker, D. Heh, C. Young, H. Park, P. Khanal, L. Larcher, A. Padovani, P. Lenahan, J. Ryan, B. H. Lee, H. Tseng, and R. Jarnmy, "Breakdown in the metal/high- κ gate stack: Identifying the "weak link" in the multilayer dielectric," in *2008 IEEE International Electron Devices Meeting (IEDM 2008)*, pp. 791-794
 - [45]. R. Degraeve, T. Kauerauf, A. Kerber, E. Cartier, B. Govoreanu, P. Roussel, L. Pantisano, P. Blomme, B. Kaczer, and G. Groeseneken, "Stress polarity dependence of degradation and breakdown of SiO₂/high- κ stacks," in *41st Annual Proceedings: International Reliability Physics Symposium (IRPS 2003)*, pp. 23-28

-
- [46]. K. L. Pey, N. Raghavan, X. Li, W. H. Liu, K. Shubhakar, X. Wu, and M. Bosman, "New insight into the TDDDB and post breakdown reliability of novel high- κ gate dielectric stacks," in *2010 International Reliability Physics Symposium (IRPS 2010)*, pp. 354-363
- [47]. T. Nigam, A. Kerber, and P. Peumans, "Accurate model for time-dependent dielectric breakdown of high- κ metal gate stacks," in *2009 IEEE International Reliability Physics Symposium (IRPS 2009)*, pp. 523-530
- [48]. N. Raghavan, K. L. Pey, K. Shubhakar, and M. Bosman, "Modified percolation model for polycrystalline high- κ gate stack with grain boundary defects," *IEEE Electron Device Letters*, vol. **32**, no. 1, pp. 78-80, 2011
- [49]. F. Crupi, R. Degraeve, A. Kerber, D. H. Kwak, and G. Groeseneken, "Correlation between stress-induced leakage current (SILC) and the HfO₂ bulk trap density in a SiO₂/HfO₂ stack," in *2004 IEEE International Reliability Physics Symposium (IRPS 2004)*, pp. 181-187
- [50]. H. S. P. Wong, "Beyond the conventional transistor," *IBM Journal of Research and Development*, vol. **46**, no. 2-3, pp. 133-168, 2002
- [51]. G. Bersuker, N. Chowdhury, C. Young, D. Heh, D. Misra, R. Choi,, "Progressive breakdown characteristics of high- κ /metal gate stacks," in *2007 IEEE International Reliability Physics Symposium (IRPS 2007)*, pp. 49-54
- [52]. J. F. Conley, Y. Ono, R. Solanki, G. Stecker, and W. Zhuang, "Electrical properties of HfO₂ deposited via atomic layer deposition using Hf(NO₃)₄ and H₂O," *Applied Physics Letters*, vol. **82**, no. 20, pp. 3508-3510, 2003
- [53]. Y. C. Yeo, T. J. King, and C. M. Hu, "Direct tunneling leakage current and scalability of alternative gate dielectrics," *Applied Physics Letters*, vol. **81**, no. 11, pp. 2091-2093, 2002
- [54]. T. Z. Ma, S. A. Campbell, R. Smith, N. Hoilien, B. Y. He, W. L. Gladfelter, C. Hobbs, D. Buchanan, C. Taylor, M. Gribelyuk, M. Tiner, M. Coppel, and J. J. Lee, "Group IVB metal oxides high permittivity gate insulators deposited from anhydrous metal nitrates," *IEEE Transactions on Electron Devices*, vol. **48**, no. 10, pp. 2348-2356, 2001
- [55]. B. Brar, G. D. Wilk, and A. C. Seabaugh, "Direct extraction of the electron tunneling effective mass in ultrathin SiO₂," *Applied Physics Letters*, vol. **69**, no. 18, pp. 2728-2730, 1996
- [56]. J. Robertson, O. Sharia, and A. A. Demkov, "Fermi level pinning by defects in HfO₂-metal gate stacks," *Applied Physics Letters*, vol. **91**, no. 13, pp. 2007
- [57]. K. Kita, and A. Toriumi, "Intrinsic origin of electric dipoles formed at high- κ /SiO₂ interface," in *2008 IEEE International Electron Devices Meeting (IEDM 2008)*, pp. 1-4
- [58]. A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H. E. Maes, U. Schwalke, and Ieee, "Characterization of the

- V_T instability in $\text{SiO}_2/\text{HfO}_2$ gate dielectrics," in *41st Annual Proceedings: International Reliability Physics Symposium (IRPS 2003)*, pp. 41-45
- [59]. Z. Q. Teo, D. S. Ang, and C. M. Ng, "Separation of hole trapping and interface-state generation by ultrafast measurement on dynamic negative-bias temperature instability," *IEEE Electron Device Letters*, vol. **31**, no. 7, pp. 656-658, 2010
- [60]. Y. Gao, A. A. Boo, Z. Q. Teo, and D. S. Ang, "On the evolution of the recoverable component of the SiON, HfSiON and HfO₂ P-MOSFETs under dynamic NBTI," in *2011 IEEE International Reliability Physics Symposium (IRPS 2011)*, pp. 935
- [61]. G. Ribes, M. Muller, S. Bruyere, D. Roy, M. Denais, V. Huard, T. Skotnicki, and G. Ghibaudo, "Characterization of V_t instability in hafnium based dielectrics by pulse gate voltage techniques," in *Proceedings of the 34th European Solid-State Device Research Conference (ESSDERC 2004)*, pp. 89-92
- [62]. X. Li, C. H. Tung, and K. L. Pey, "The radial distribution of defects in a percolation path," *Applied Physics Letters*, vol. **93**, no. 26, 2008
- [63]. R. Ranjan, K. L. Pey, C. H. Tung, L. J. Tang, G. Groeseneken, L. K. Bera, and S. De Gendt, "A comprehensive model for breakdown mechanism in HfO₂ high- κ gate stacks," in *2004 IEEE International Electron Devices Meeting (IEDM 2004)*, pp. 725-728
- [64]. K. McKenna, and A. Shluger, "The interaction of oxygen vacancies with grain boundaries in monoclinic HfO₂," *Applied Physics Letters*, vol. **95**, no. 22, 2009
- [65]. G. Bersuker, D. Heh, C. D. Young, L. Morassi, A. Padovani, L. Larcher, K. S. Yew, Y. C. Ong, D. S. Ang, K. L. Pey, and W. Taylor, "Mechanism of high- κ dielectric-induced breakdown of the interfacial SiO₂ layer," in *2010 IEEE International Reliability Physics Symposium (IRPS 2010)*, pp. 373-378
- [66]. G. Bersuker, C. S. Park, J. Barnett, P. S. Lysaght, P. D. Kirsch, C. D. Young, R. Choi, B. H. Lee, B. Foran, K. van Benthem, S. J. Pennycook, P. M. Lenahan, and J. T. Ryan, "The effect of interfacial layer properties on the performance of Hf-based gate stack devices," *Journal of Applied Physics*, vol. **100**, no. 9, pp. 2006
- [67]. B. J. O'Sullivan, P. K. Hurley, C. Leveugle, and J. H. Das, "Si(100)-SiO₂ interface properties following rapid thermal processing," *Journal of Applied Physics*, vol. **89**, no. 7, pp. 3811-3820, 2001
- [68]. Y. Lee, S. Kim, J. Koo, I. Kim, J. Choi, H. Jeon, and Y. Won, "Effect of nitrogen incorporation in HfO₂ films deposited by plasma-enhanced atomic layer deposition," *Journal of the Electrochemical Society*, vol. **153**, no. 4, pp. G353-G357, 2006

- [69]. J. Choi, S. Kim, J. Kim, H. Kang, H. Jeon, and C. Bae, "Effects of N₂ remote plasma nitridation on the structural and electrical characteristics of the HfO₂ gate dielectrics grown using remote plasma atomic layer deposition methods," *Journal of Vacuum Science & Technology A*, vol. **24**, no. 4, pp. 900-907, 2006
- [70]. B. H. Lee, L. G. Kang, R. Nieh, W. J. Qi, and J. C. Lee, "Thermal stability and electrical characteristics of ultrathin hafnium oxide gate dielectric reoxidized with rapid thermal annealing," *Applied Physics Letters*, vol. **76**, no. 14, pp. 1926-1928, 2000
- [71]. K. Kukli, T. Pilvi, M. Ritala, T. Sajavaara, J. Lu, and M. Leskela, "Atomic layer deposition of hafnium dioxide thin films from hafnium tetrakis(dimethylamide) and water," *Thin Solid Films*, vol. **491**, no. 1-2, pp. 328-338, 2005
- [72]. R. K. Nahar, V. Singh, and A. Sharma, "Study of electrical and microstructure properties of high dielectric hafnium oxide thin film for MOS devices," *Journal of Materials Science-Materials in Electronics*, vol. **18**, no. 6, pp. 615-619, 2007
- [73]. W. M. Tang, M. G. Helander, M. T. Greiner, G. Dong, W. T. Ng, and Z. H. Lu, "Enhanced performance for OTFT on glass with HfO₂ as gate dielectric by UV-ozone treatment," in *2009 IEEE International Conference of Electron Devices and Solid-State Circuits (EDSSC 2009)*, pp. 513-516
- [74]. Y. Zhao, K. Kita, K. Kyuno, and A. Toriumi, "Suppression of leakage current and moisture absorption of La₂O₃ films with ultraviolet ozone post treatment," *Japanese Journal of Applied Physics Part 1-Regular Papers Brief Communications & Review Papers*, vol. **46**, no. 7A, pp. 4189-4192, 2007
- [75]. P. C. Yang, T. C. Chang, S. C. Chen, H. H. Su, J. Lu, H. C. Huang, D. S. Gan, and N. J. Ho, "Improvement on low-temperature deposited HfO₂ film and interfacial layer by high-pressure oxygen treatment," *Solid-State Electronics*, vol. **62**, no. 1, pp. 128-131, 2011
- [76]. S. Ramanathan, G. D. Wilk, D. A. Muller, C. M. Park, and P. C. McIntyre, "Growth and characterization of ultrathin zirconia dielectrics grown by ultraviolet ozone oxidation," *Applied Physics Letters*, vol. **79**, no. 16, pp. 2621-2623, 2001
- [77]. W. J. Song, S. K. So, D. Y. Wang, Y. Qiu, and L. L. Cao, "Angle dependent X-ray photoemission study on UV-ozone treatments of indium tin oxide," *Applied Surface Science*, vol. **177**, no. 3, pp. 158-164, 2001
- [78]. T. Nabatame, K. Iwamoto, H. Ota, K. Tominaga, H. Hisamatsu, T. Yasuda, K. Yamamoto, W. Mizubayashi, Y. Morita, N. Yasuda, M. Ohno, T. Horikawa, and A. Toriumi, "Design and proof of high quality HfAlO_x film formation for MOSCAPs and nMOSFETs through Layer-by-Layer Deposition and Annealing process," in *2003 Symposium on VLSI Technology*, 2003, pp. 25-26

-
- [79]. J. F. Conley, D. J. Tweet, Y. Ono, and G. Stecker, "Interval annealing during alternating pulse deposition," in *Integration of Advanced Micro-and Nanoelectronic Devices-Critical Issues and Solutions*, 2004, pp. 5-10
- [80]. J. F. Conley, Y. Ono, and D. J. Tweet, "Densification and improved electrical properties of pulse-deposited films via in situ modulated temperature annealing," *Applied Physics Letters*, vol. **84**, no. 11, pp. 1913-1915, 2004
- [81]. E. Miranda, J. Molina, Y. Kim, and H. Iwai, "Tunneling in sub-5nm La_2O_3 films deposited by E-beam evaporation," *Journal of Non-Crystalline Solids*, vol. **352**, no. 1, pp. 92-97, 2006
- [82]. M. Lanza, M. Porti, M. Nafria, G. Benstetter, W. Frammelsberger, H. Ranzinger, E. Lodermeier, and G. Jaschke, "Influence of the manufacturing process on the electrical properties of thin ($< 4\text{nm}$) hafnium based high- κ stacks observed with CAFM," *Microelectronics Reliability*, vol. **47**, no. 9-11, pp. 1424-1428, 2007
- [83]. S. J. Oshea, R. M. Atta, M. P. Murrell, and M. E. Welland, "Conducting atomic-force microscopy study of silicon dioxide breakdown," *Journal of Vacuum Science & Technology B*, vol. **13**, no. 5, pp. 1945-1952, 1995
- [84]. M. P. Murrell, M. E. Welland, S. J. Oshea, T. M. H. Wong, J. R. Barnes, A. W. McKinnon, M. Heyns, and S. Verhaverbeke, "Spatially resolved electrical measurements of SiO_2 gate oxides using atomic force microscopy " *Applied Physics Letters*, vol. **62**, no. 7, pp. 786-788, 1993
- [85]. H. Watanabe, K. Fujita, and M. Ichikawa, "Observation and creation of current leakage sites in ultrathin silicon dioxide films using scanning tunneling microscopy," *Applied Physics Letters*, vol. **72**, no. 16, pp. 1987-1989, 1998
- [86]. K. Fujita, H. Watanabe, and M. Ichikawa, "Scanning tunneling microscopy study on the surface and interface of $\text{Si}(111)/\text{SiO}_2$ structures," *Journal of Applied Physics*, vol. **83**, no. 7, pp. 3638-3642, 1998
- [87]. N. Miyata, and M. Ichikawa, "Electrical characterization of atomic-scale defects in an ultrathin Si oxynitride layer," *Japanese Journal of Applied Physics Part 2-Letters*, vol. **40**, no. 12A, pp. L1271-L1273, 2001
- [88]. L. Zhang, Y. Mitani, and H. Satake, "Visualization of progressive breakdown evolution in gate dielectric by conductive atomic force microscopy," *IEEE Transactions on Device and Materials Reliability*, vol. **6**, no. 2, pp. 277-282, 2006
- [89]. M. Lanza, L. Aguilera, M. Porti, M. Nafria, and X. Aymerich, "Improving the electrical performance of a CAFM for gate oxide reliability measurements," in *2009 Spanish Conference on Electron Devices (CDE 2009)*, pp. 234-237
- [90]. T. Yasue, Y. Yoshida, H. Koyama, T. Kato, and T. Nishioka, "Dielectric breakdown of silicon oxide studied by scanning probe microscopy," *Journal of Vacuum Science & Technology B*, vol. **15**, no. 6, pp. 1884-1888, 1997

-
- [91]. P. Fiorenza, W. Polspoel, and W. Vandervorst, "Conductive atomic force microscopy studies of thin SiO₂ layer degradation," *Applied Physics Letters*, vol. **88**, no. 22, 2006
- [92]. M. Porti, L. Aguilera, X. Blasco, M. Nafria, and X. Aymerich, "Reliability of SiO₂ and high- κ gate insulators: A nanoscale study with conductive atomic force microscopy," *Microelectronic Engineering*, vol. **84**, no. 3, pp. 501-505, 2007
- [93]. E. S. Daniel, J. T. Jones, O. J. Marsh, and T. C. McGill, "Macroscopic and microscopic studies of electrical properties of very thin silicon dioxide subject to electrical stress," *Journal of Vacuum Science & Technology B*, vol. **15**, no. 4, pp. 1089-1096, 1997
- [94]. K. E. Johnson, and T. Engel, "Direct measurement of reaction-kinetics for the decomposition of ultrathin oxide on Si(001) using scanning tunneling microscopy," *Physical Review Letters*, vol. **69**, no. 2, pp. 339-342, 1992
- [95]. Y. Wei, R. M. Wallace, and A. C. Seabaugh, "Void formation on ultrathin thermal silicon oxide films on the Si(100) surface," *Applied Physics Letters*, vol. **69**, no. 9, pp. 1270-1272, 1996
- [96]. V. Iglesias, M. Porti, M. Nafria, X. Aymerich, P. Dudek, T. Schroeder, and G. Bersuker, "Correlation between the nanoscale electrical and morphological properties of crystallized hafnium oxide-based metal oxide semiconductor structures," *Applied Physics Letters*, vol. **97**, no. 26, 2010
- [97]. Y. C. Ong, D. S. Ang, S. J. O'Shea, K. L. Pey, S. J. Wang, C. H. Tung, and X. Li, "Scanning tunneling microscopy study of nitrogen incorporated HfO₂," *Journal of Applied Physics*, vol. **104**, no. 6, 2008
- [98]. K. Kyuno, K. Kita, and A. Toriumi, "Evolution of leakage paths in HfO₂/SiO₂ stacked gate dielectrics: a stable direct observation by ultrahigh vacuum conducting atomic force microscopy," *Applied Physics Letters*, vol. **86**, no. 6, 2005
- [99]. Y. C. Ong, D. S. Ang, K. L. Pey, S. J. O'Shea, K. E. J. Goh, C. Troadec, C. H. Tung, T. Kawanago, K. Kakushima, and H. Iwai, "Bilayer gate dielectric study by scanning tunneling microscopy," *Applied Physics Letters*, vol. **91**, no. 10, 2007
- [100]. K. Meinel, A. Eichler, K. M. Schindler, and H. Neddermeyer, "STM, LEED, and DFT characterization of epitaxial ZrO₂ films on Pt(111)," *Surface Science*, vol. **562**, no. 1-3, pp. 204-218, 2004
- [101]. K. Shubhakar, K. L. Pey, S. S. Kushvaha, M. Bosman, S. J. O'Shea, N. Raghavan, M. Kouda, K. Kakushima, Z. R. Wang, H. Y. Yu, and H. Iwai "Nanoscale electrical and physical study of polycrystalline high- κ dielectrics and proposed reliability enhancement techniques," in *2011 IEEE International Reliability Physics Symposium (IRPS 2010)*, pp.

-
- [102]. K. L. Pey, R. Ranjan, C. H. Tung, L. J. Tang, V. L. Lo, K. S. Lim, T. A. L. Selvarajoo, and D. S. Ang, "Breakdowns in high- κ gate stacks of nano-scale CMOS devices," *Microelectronic Engineering*, vol. **80**, no. pp. 353-361, 2005
- [103]. H. Kim, A. Marshall, P. C. McIntyre, and K. C. Saraswat, "Crystallization kinetics and microstructure-dependent leakage current behavior of ultrathin HfO₂ dielectrics: In situ annealing studies," *Applied Physics Letters*, vol. **84**, no. 12, pp. 2064-2066, 2004
- [104]. Y. C. Ong, D. S. Ang, K. L. Pey, Z. R. Wang, S. J. O'Shea, C. H. Tung, T. Kawanago, K. Kakushima, and H. Iwai, "Electronic trap characterization of the Sc₂O₃/La₂O₃ high- κ gate stack by scanning tunneling microscopy," *Applied Physics Letters*, vol. **92**, no. 2, pp. 2008
- [105]. A. H. Carim, M. M. Dovek, C. F. Quate, R. Sinclair, and C. Vorst, "High-resolution electron-microscopy and scanning tunneling microscopy of native oxides on silicon," *Science*, vol. **237**, no. 4815, pp. 630-633, 1987
- [106]. C. C. Yeo, B. J. Cho, M. S. Joo, S. J. Whoang, D. L. Kwong, L. K. Bera, S. Mathew, and N. Balasubramanian, "Improvement of electrical properties of MOCVD HfO₂ by multistep deposition," *Electrochemical and Solid State Letters*, vol. **6**, no. 11, pp. F42-F44, 2003
- [107]. S. N. Magonov, and M. H. Whangbo. "Surface analysis with STM and AFM," 1996, Germany: VCH
- [108]. E. Meyer, J. H. Hans, and B. Roland "Scanning probe microscopy : the lab on a tip, " 2004, Berlin: Springer
- [109]. J. A. Stroscio, and W. J. Kaiser "Scanning tunneling microscopy," vol. 27. 1993, San Diego: Academic press
- [110]. A. Kaneko, Y. Homma, H. Hibino, and T. Ogino, "Ultrahigh vacuum scanning electron microscope system combined with wide-movable scanning tunneling microscope," *Review of Scientific Instruments*, vol. **76**, no. 8, pp. 083709, 2005
- [111]. S. Kerfriden, A. H. Nahle, S. A. Campbell, F. C. Walsh, and J. R. Smith, "The electrochemical etching of tungsten STM tips," *Electrochimica Acta*, vol. **43**, no. 12-13, pp. 1939-1944, 1998
- [112]. E. Meyer, H. J. Hug, and R. Bennewitz, "Scanning probe microscopy : the lab on a tip," *Springer*, 2004
- [113]. G. Bersuker, D. Heh, C. Young, H. Park, P. Khanal, L. Larcher, A. Padovani, P. Lenahan, J. Ryan, B. H. Lee, H. Tseng, and R. Jarnmy, "Breakdown in the metal/high- κ gate stack: identifying the "weak link" in the multilayer dielectric," in 2008 *IEEE International Electron Devices Meeting (IEDM 2008)*, pp. 791-794
- [114]. X. Blasco, J. Petry, M. Nafria, X. Aymerich, O. Richard, and W. Vandervorst, "CAFM characterization of the dependence of HfAlO_x electrical behavior on

- post-deposition annealing temperature," *Microelectronic Engineering*, vol. **72**, no. 1-4, pp. 191-196, 2004
- [115]. N. A. Chowdhury, G. Bersuker, C. Young, R. Choi, S. Krishnan, and D. Misra, "Breakdown characteristics of nFETs in inversion with metal/ WO_2 gate stacks," *Microelectronic Engineering*, vol. **85**, no. 1, pp. 27-35, 2008
- [116]. M. Sato, C. Tamura, K. Yamabe, K. Shiraishi, S. Miyazaki, K. Yamada, R. Hasunuma, T. Aoyama, Y. Nara, and Y. Ohji, "Cathode electron injection breakdown model and time dependent dielectric breakdown lifetime prediction in high- κ /metal gate stack p-type metal-oxide-silicon field effect transistors," *Japanese Journal of Applied Physics*, vol. **47**, no. 5, pp. 3326-3331, 2008
- [117]. R. Degraeve, T. Kauerauf, M. Cho, M. Zahid, L. A. Ragnarsson, D. P. Brunco, B. Kaczer, P. Roussel, S. De Gendt, and G. Groeseneken, " Degradation and breakdown of 0.9 nm EOT SiO_2 /ALD HfO_2 /metal gate stacks under positive constant voltage stress " in 2005 *IEEE International Electron Devices Meeting (IEDM 2005)*, pp. 419-422.
- [118]. Y. C. Ong, D. S. Ang, S. J. O'Shea, K. L. Pey, S. J. Wang, C. H. Tung, and X. Li, "Scanning tunneling microscopy study of nitrogen incorporated HfO_2 ," *Journal of Applied Physics*, vol. **104**, no. 6, 2008
- [119]. Y. C. Ong, D. S. Ang, K. L. Pey, S. J. O'Shea, K. E. J. Goh, C. Troadec, C. H. Tung, T. Kawanago, K. Kakushima, and H. Iwai, "Bilayer gate dielectric study by scanning tunneling microscopy," *Applied Physics Letters*, vol. **91**, no. 10, 2007
- [120]. D. S. Ang, Y. C. Ong, S. J. O'Shea, K. L. Pey, K. Kakushima, and H. Iwai, "Study of trap generation in the $\text{Sc}_2\text{O}_3/\text{La}_2\text{O}_3/\text{SiO}_x$ gate dielectric stack by scanning tunneling microscopy," *Applied Physics Letters*, vol. **93**, no. 24, 2008
- [121]. D. S. Ang, Y. C. Ong, S. J. O'Shea, K. L. Pey, C. H. Tung, T. Kawanago, K. Kakushima, and H. Iwai, "Polarity dependent breakdown of the high- κ / SiO_x gate stack: A phenomenological explanation by scanning tunneling microscopy," *Applied Physics Letters*, vol. **92**, no. 19, 2008
- [122]. S. Lee, W.G. Kim, S.W. Rhee, and K. Yong, "Resistance switching behaviors of hafnium oxide films grown by MOCVD for nonvolatile memory applications," *Journal of the Electrochemical Society*, vol. 155, pg. H92-H96, 2008
- [123]. A. S. Foster, F. L. Gejo, A. L. Shluger, and R. M. Nieminen, "Vacancy and interstitial defects in hafnia," *Physical Review B*, vol. **65**, no. 17, pp. 2002
- [124]. G. Lucovsky, C. C. Fulton, Y. Zhang, Y. Zou, J. Luning, L. F. Edge, J. L. Whitten, R. J. Nemanich, H. Ade, D. G. Schlom, V. V. Afanase'v, A. Stesmans, S. Zollner, D. Triyoso, and B. R. Rogers, "Conduction band-edge states associated with the removal of d-state degeneracies by the Jahn-Teller effect," *IEEE Transactions on Device and Materials Reliability*, vol. **5**, no. 1, pp. 65-83, 2005

-
- [125]. C. D. Young, D. Heh, S. V. Nadkarni, R. Choi, J. J. Peterson, J. Barnett, B. H. Lee, and G. Bersuker, "Electron trap generation in high- κ gate stacks by constant voltage stress," *IEEE Transactions on Device and Materials Reliability*, vol. **6**, no. 2, pp. 123-131, 2006
- [126]. O. Pirrotta, L. Larcher, M. Lanza, A. Padovani, M. Porti, M. Nafria, and G. Bersuker, "Leakage current through the poly-crystalline HfO_2 : Trap densities at grains and grain boundaries," *Journal of Applied Physics*, vol. 114, 134503, 2013.
- [127]. K. S. Yew, Y. C. Ong, D. S. Ang, K. L. Pey, G. Bersuker, P. S. Lysaght, and D. Heh, "Nanoscale Characterization of $\text{HfO}_2/\text{SiO}_x$ Gate Stack Degradation by Scanning Tunneling Microscopy," in 2009 *International Conference on Solid State Devices and Materials (SSDM 2009)*, Japan, pp. 314
- [128]. C. S. Park, M. M. Hussain, J. Huang, C. Park, K. Tateiwa, C. Young, H. K. Park, M. Cruz, D. Gilmer, K. Rader, J. Price, P. Lysaght, D. Heh, G. Bersuker, P. D. Kirsch, H. H. Tseng, and R. Jammy, "A scalable and highly manufacturable single metal gate/high- κ CMOS integration for sub-32nm technology for LSTP applications," in 2009 *Symposium on VLSI Technology*, Japan, 2009, pp. 208-209.
- [129]. G. Bersuker, D. Heh, J. Huang, C. S. Park, A. Padovani, L. Larcher, P. D. Kirsch, and R. Jammy, "Gate leakage current reduction in two-step processed high- κ dielectrics for low power applications," in 2010 *International Conference on Solid State Devices and Materials (SSDM 2010)*, Japan, pp. 1034
- [130]. W. Polspoel, W. Vandervorst, J. Petry, T. Conard, and A. Benedetti, "Comparison of electric properties of ultra-thin thermal and plasma nitrided silicon oxides with different post-deposition treatments using C-AFM," *Microelectronic Engineering*, vol. **80**, pp. 436-439, 2005
- [131]. J. Petry, W. Vandervorst, and X. Blasco, "Effect of N_2 anneal on thin HfO_2 layers studied by conductive atomic force microscopy," *Microelectronic Engineering*, vol. **72**, no. 1-4, pp. 174-179, 2004
- [132]. A. Navrotsky, "Thermochemical insights into refractory ceramic materials based on oxides with large tetravalent cations," *Journal of Materials Chemistry*, vol. **15**, no. 19, pp. 1883-1890, 2005
- [133]. P. S. Lysaght, J. C. Woicik, M. A. Sahiner, B. H. Lee, and R. Jammy, "Characterizing crystalline polymorph transitions in HfO_2 by extended x-ray absorption fine-structure spectroscopy," *Applied Physics Letters*, vol. **91**, no. 12, pp. 2007
- [134]. G. Pant, A. Gnade, M. J. Kim, R. M. Wallace, B. E. Gnade, M. A. Quevedo-Lopez, and P. D. Kirsch, "Effect of thickness on the crystallization of ultrathin HfSiON gate dielectrics," *Applied Physics Letters*, vol. **88**, no. 3, 2006

- [135]. J. H. Gao, R. G. Thompson, and B. R. Patterson, "Computer simulation of grain growth with second phase particle pinning," *Acta Materialia*, vol. **45**, no. 9, pp. 3653-3658, 1997
- [136]. J. Kang, Y.-H. Kim, J. Bang, and K. J. Chang, "Direct and defect-assisted electron tunneling through ultrathin SiO₂ layers from first principles," *Physical Review B*, vol. **77**, no. 19, pp. 2008
- [137]. E. Nadimi, P. Plaenitz, R. Oettking, M. Schreiber, and C. Radehaus, "Single and multiple oxygen vacancies in ultrathin SiO₂ gate dielectric and their influence on the leakage current: an *ab initio* investigation," *IEEE Electron Device Letters*, vol. **31**, no. 8, pp. 881-883, 2010
- [138]. G. Bersuker, D. C. Gilmer, D. Veksler, J. Yum, H. Park, S. Lian, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, A. Shluger, V. Iglesias, M. Porti, M. Nafria, W. Taylor, P. D. Kirsch, and R. Jammy, "Metal oxide RRAM switching mechanism based on conductive filament microscopic properties," in *2010 IEEE International Electron Devices Meeting (IEDM 2010)*, pp. 456
- [139]. C. Hobbs, L. Fonseca, V. Dhandapani, S. Samavedam, B. Taylor, J. Grant, L. Dip, D. Triyoso, R. Hegde, D. Gilmer, R. Garcia, D. Roan, L. Lovejoy, R. Rai, L. Hebert, H. Tseng, B. White, and P. Tobin, "Fermi level pinning at the poly-Si/metal oxide interface," in *2003 Symposium on VLSI Technology*, 2003, pp. 9-10
- [140]. G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high- κ dielectrics reliability issues," *IEEE Transactions on Device and Materials Reliability*, vol. **5**, no. 1, pp. 5-19, 2005
- [141]. R. Kotlyar, M. D. Giles, P. Matagne, B. Obradovic, L. Shrifan, M. Stettler, and E. Wang, "Inversion mobility and gate leakage in high- κ /metal gate MOSFETs," in *2004 IEEE International Electron Devices Meeting (IEDM 2004)*, pp. 391-394
- [142]. K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C. H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McLntyre, P. Moon, J. Neiryneck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, and K. Zawadzki, "A 45nm logic technology with high- κ +metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging," in *2007 IEEE International Electron Devices Meeting (IEDM 2007)*, pp. 247-250
- [143]. S. Hyun, J. Han, H. Park, H. Na, H. J. Son, H. Y. Lee, H. Hong, H. Lee, J. Song, J. J. Kim, J. Lee, W. C. Jeong, H. J. Cho, K. I. Seo, D. W. Kim, S. P. Sim, S. B. Kang, D. K. Sohn, C. Siyoung, H. Kang, and C. Chilhee,

- "Aggressively scaled high- κ last metal gate stack with low variability for 20nm logic high performance and low power applications," in *2011 Symposium on VLSI Technology (VLSIT)*, 2011, pp. 32-33
- [144]. C. L. Hinkle, R. V. Galatage, R. A. Chapman, E. M. Vogel, H. N. Alshareef, C. Freeman, E. Wimmer, H. Niimi, A. Li-Fatou, J. B. Shaw, and J. J. Chambers, "Dipole controlled metal gate with hybrid low resistivity cladding for gate-last CMOS with low V_t ," in *2010 Symposium on VLSI Technology (VLSIT)*, 2010, pp. 183-184
- [145]. K. S. Yew, D. S. Ang, and G. Bersuker, "Bimodal weibull distribution of metal/high- κ gate stack TDDb - insights by scanning tunneling microscopy," *IEEE Electron Device Letters*, vol. **33**, no. 2, pp. 146-148, 2012
- [146]. K. S. Yew, and D. S. Ang, "Study of the multistep-deposited and UV-ozone-annealed HfZrO gate stack by scanning tunneling microscopy and pulse C-V measurement," *IEEE Transactions on Electron Devices*, vol. **59**, no. 8, pp. 2268-2272, 2012
- [147]. L. Wu, K. S. Yew, D. S. Ang, W. J. Liu, T. T. Le, T. L. Duan, C. H. Hou, X. F. Yu, D. Y. Lee, K. Y. Hsu, J. Xu, H. J. Tao, M. Cao, and H. Y. Yu, "A novel multi deposition multi room-temperature annealing technique via ultraviolet-ozone to improve high- κ /metal (HfZrO/TiN) gate stack integrity for a gate-last process," in *2010 IEEE International Electron Devices Meeting (IEDM 2010)*, pp. 11.6.1-11.6.4
- [148]. B. H. Lee, J. Oh, H. H. Tseng, R. Jammy, and H. Huff, "Gate stack technology for nanoscale devices," *Materials Today*, vol. **9**, no. 6, pp. 32-40, 2006
- [149]. E. Cartier, B. P. Linder, V. Narayanan, and V. K. Paruchuri, "Fundamental understanding and optimization of PBT1 in nFETs with SiO₂/HfO₂ gate stack," in *2006 IEEE International Electron Devices Meeting (IEDM 2006), Vols 1 and 2*, pp. 57-60
- [150]. K. Tse, and J. Robertson, "Defects and their passivation in high- κ gate oxides," *Microelectronic Engineering*, vol. **84**, no. 4, pp. 663-668, 2007
- [151]. T. Naito, C. Tamura, S. Inumiya, R. Hasunuma, and K. Yamabe, "Homogeneity improvements in the dielectric characteristics of HfSiON films by nitridation," *Japanese Journal of Applied Physics Part 1-Regular Papers Brief Communications & Review Papers*, vol. **46**, no. 5B, pp. 3197-3201, 2007
- [152]. S. Guha, and V. Narayanan, "Oxygen vacancies in high dielectric constant oxide-semiconductor films," *Physical Review Letters*, vol. **98**, no. 19, 2007
- [153]. H. Sim, H. Chang, and H. Hwang, "Electrical characteristics of ozone-oxidized HfO₂ gate dielectrics," *Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers*, vol. **42**, no. 4A, pp. 1596-1597, 2003
- [154]. L. Wu, H. Yu, K. S. Yew, J. Pan, W. J. Liu, and T. L. Duan, "Multideposition multiroom-temperature annealing via ultraviolet ozone for HfZrO high- κ and

- integration with a TiN metal gate in a gate-last process," *IEEE Transactions on Electron Devices*, vol. **58**, no. 7, pp. 2177-2181, 2011
- [155]. K. S. Yew, D. S. Ang, L. J. Tang, K. Cui, G. Bersuker, and P. S. Lysaght, "Scanning tunneling microscopy study of the multi-step deposited and annealed HfSiO_x gate dielectric," *Journal of the Electrochemical Society*, vol. **158**, no. 10, pp. H1021-H1026, 2011
- [156]. L. C. Lee, G. Black, R. L. Sharpless, and T. G. Slanger, "O 1s yield from O₃ photo-dissociation at 1700-2400 Å," *Journal of Chemical Physics*, vol. **73**, no. 1, pp. 256-258, 1980
- [157]. K. S. Yew, D. S. Ang, L. J. Tang, K. Cui, G. Bersuker, and P. S. Lysaght, "Scanning tunneling microscopy study of the multi-step deposited and annealed HfSiO_x gate dielectric," *Journal of the Electrochemical Society*, vol. **158**, no. 10, pp. H1021-H1026, 2011
- [158]. J. F. Zhang, S. Taylor, and W. Eccleston, "A comparative-study of the electron trapping and thermal detrapping in SiO₂ prepared by plasma and thermal-oxidation," *Journal of Applied Physics*, vol. **72**, no. 4, pp. 1429-1435, 1992
- [159]. G. Groeseneken, L. Pantisano, L. A. Ragnarsson, R. Degraeve, M. Houssa, T. Kauerauf, P. Roussel, S. De Gendt, and M. Heyns, "Achievements and challenges for the electrical performance of MOSFETs with high-κ gate dielectrics," in *Proceedings of the 11th International Symposium on the Physical and Failure Analysis of Integrated Circuits, IPFA 2004*, pp. 147-155
- [160]. W. D. Zhang, B. Govoreanu, X. F. Zheng, D. R. Aguado, M. Rosmeulen, P. Blomme, J. F. Zhang, and J. Van Houdt, "Two-pulse C-V: A new method for characterizing electron traps in the bulk of SiO₂/high-κ dielectric stacks," *IEEE Electron Device Letters*, vol. **29**, no. 9, pp. 1043-1046, 2008
- [161]. G. Puzzilli, B. Govoreanu, F. Irrera, M. Rosmeulen, and J. Van Houdt, "Characterization of charge trapping in SiO₂/Al₂O₃ dielectric stacks by pulsed C-V technique," *Microelectronics Reliability*, vol. **47**, no. 4-5, pp. 508-512, 2007
- [162]. D. S. Ang, S. Wang, and C. H. Ling, "Evidence of two distinct degradation mechanisms from temperature dependence of negative bias stressing of the ultrathin gate p-MOSFET," *IEEE Electron Device Letters*, vol. **26**, no. 12, pp. 906-908, 2005
- [163]. W. D. Zhang, B. Govoreanu, X. F. Zheng, D. R. Aguado, A. Rosmeulen, P. Blomme, J. F. Zhang, and J. Van Houdt, "Two-pulse C-V: A new method for characterizing electron traps in the bulk of SiO₂/high-κ dielectric stacks," *IEEE Electron Device Letters*, vol. **29**, no. 9, pp. 1043-1046, 2008
- [164]. C. Z. Zhao, M. B. Zahid, J. F. Zhang, G. Groeseneken, R. Degraeve, and S. De Gendt, "Threshold voltage instability of p-channel metal-oxide-semiconductor field effect transistors with hafnium based dielectrics," *Applied Physics Letters*, vol. **90**, no. 14, 2007

-
- [165]. Y. Wang, V. Lee, and K. P. Cheung, "Frequency dependent charge-pumping, how deep does it probe?," in *2006 IEEE International Electron Devices Meeting (IEDM 2006), Vols 1 and 2*, pp. 491-494
- [166]. K. Xiong, J. Robertson, and S. J. Clark, "Defect energy states in high- κ gate oxides," *Physica Status Solidi B-Basic Solid State Physics*, vol. **243**, no. 9, pp. 2071-2080, 2006
- [167]. Y. Yi-Lin, Z. Wenqi, C. Chi-Yun, H. Yi-Ping, C. Pin-Tseng, H. Chia-Wei, C. Li-Kong, L. Chien-Ting, H. Che-Hua, L. Chien-Ming, and Y. Wen-Kuan, "Reliability improvement of 28 nm high- κ /metal gate-last MOSFET using appropriate oxygen annealing," *IEEE Electron Device Letters*, vol. **33**, no. 8, pp. 1183-1185, 2012
- [168]. M. H. Cho, Y. S. Roh, C. N. Whang, K. Jeong, S. W. Nahm, D. H. Ko, J. H. Lee, N. I. Lee, and K. Fujihara, "Thermal stability and structural characteristics of HfO₂ films on Si (100) grown by atomic-layer deposition," *Applied Physics Letters*, vol. **81**, no. 3, pp. 472-474, 2002
- [169]. T. P. Smirnova, L. V. Yakovkina, V. N. Kitchai, V. V. Kalchev, Y. V. Shubin, N. B. Morozova, and K. V. Zherikova, "Chemical vapor deposition and characterization of hafnium oxide films," *Journal of Physics and Chemistry of Solids*, vol. **69**, no. 2-3, pp. 685-687, 2008
- [170]. G. Aygun, and I. Yildiz, "Interfacial and structural properties of sputtered HfO₂ layers," *Journal of Applied Physics*, vol. **106**, no. 1, 2009
- [171]. M. P. Agustin, G. Bersuker, B. Foran, L. A. Boatner, and S. Stemmer, "Scanning transmission electron microscopy investigations of interfacial layers in HfO₂ gate stacks," *Journal of Applied Physics*, vol. **100**, no. 2, 2006
- [172]. Z. W. Yuan, S. Csillag, M. A. Tafreshi, and C. Colliex, "High-spatial-resolution extended energy-loss fine-structure investigations of silicon dioxide compounds," *Ultramicroscopy*, vol. **59**, no. 1-4, pp. 149-157, 1995
- [173]. G. He, Z. Q. Sun, G. Li, and L. D. Zhang, "Review and perspective of Hf-based high- κ gate dielectrics on silicon," *Critical Reviews in Solid State and Materials Sciences*, vol. **37**, no. 3, pp. 131-157, 2012
- [174]. L. A. Ragnarsson, S. Severi, L. Trojman, K. D. Johnson, D. P. Brunco, M. Aoulaiche, M. Houssa, T. Kauerauf, R. Degraeve, A. Delabie, V. S. Kaushik, S. De Gendt, W. Tsai, G. Groeseneken, M. Kristin De, and M. Heyns, "Electrical characteristics of 0.8 nm EOT HfO₂/TaN low thermal-budget n-channel FETs with solid-phase epitaxially regrown junctions," *IEEE Transactions on Electron Devices*, vol. **53**, no. 7, pp. 1657-1668, 2006
- [175]. C. Ying-Tsung, I. F. Ssu, C. Wen-Tai, L. Chien-Ting, T. Shih-Hung, W. Shao-Wei, and C. Shouu-Jinn, "Chemical oxide interfacial layer for the high- κ last/gate-last integration scheme," *IEEE Electron Device Letters*, vol. **33**, no. 7, pp. 946-948, 2012

-
- [176]. K. Choi, H. Jagannathan, C. Choi, L. Edge, T. Ando, M. Frank, P. Jamison, M. Wang, E. Cartier, S. Zafar, J. Bruley, A. Kerber, B. Linder, A. Callegari, Q. Yang, S. Brown, J. Stathis, J. Iacoponi, V. Paruchuri, and V. Narayanan, "Extremely scaled gate-first high- κ /metal gate stack with EOT of 0.55 nm using novel interfacial layer scavenging techniques for 22nm technology node and beyond," in *2009 Symposium on VLSI Technology*, 2009, pp. 138-139
 - [177]. T. Ando, M. M. Frank, K. Choi, C. Choi, J. Bruley, M. Hopstaken, M. Copel, E. Cartier, A. Kerber, A. Callegari, D. Lacey, S. Brown, Q. Yang, and V. Narayanan, "Understanding mobility mechanisms in extremely scaled HfO₂ (EOT 0.42 nm) using remote interfacial layer scavenging technique and V_t-tuning dipoles with gate-first process," in *2009 IEEE International Electron Devices Meeting (IEDM 2009)*, pp. 1-4
 - [178]. L. A. Ragnarsson, T. Chiarella, M. Togo, T. Schram, P. Absil, and T. Hoffmann, "Ultrathin EOT high- κ /metal gate devices for future technologies: Challenges, achievements and perspectives (invited)," *Microelectronic Engineering*, vol. **88**, no. 7, pp. 1317-1322, 2011
 - [179]. D. Min, L. Jinping, D. Guo, S. Krishnan, J. F. Shepard, P. Ronsheim, K. Unoh, S. Siddiqui, R. Krishnan, L. Zhengwen, Z. Kai, J. Sudijono, and M. P. Chudzik, "A novel atomic layer oxidation technique for EOT scaling in gate-last high- κ /metal gate CMOS technology," in *2011 IEEE International Electron Devices Meeting (IEDM 2011)*, pp. 28.5.1-28.5.4
 - [180]. J. R. Hauser, and K. Ahmed, "Characterization of ultra-thin oxides using electrical C-V and I-V measurements," in *Characterization and Metrology for Ulsi Technology*, 1998, pp. 235-239
 - [181]. J. Huang, D. Heh, P. Sivasubramani, P. D. Kirsch, G. Bersuker, D. C. Gilmer, M. A. Quevedo-Lopez, M. M. Hussain, P. Majhi, P. Lysaght, H. Park, N. Goel, C. Young, C. S. Park, C. Park, M. Cruz, V. Diaz, P. Y. Hung, J. Price, H. H. Tseng, and R. Jammy, "Gate first high- κ /metal gate stacks with zero SiO_x interface achieving EOT = 0.59nm for 16nm application," in *2009 Symposium on VLSI Technology*, 2009, pp. 34-35
 - [182]. Z. L. Li, T. Schram, T. Witters, J. Tseng, S. De Gendt, and K. De Meyer, "Oxygen incorporation in TiN for metal gate work function tuning with a replacement gate integration approach," *Microelectronic Engineering*, vol. **87**, no. 9, pp. 1805-1807, 2010
 - [183]. R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, and H. E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," *IEEE Transactions on Electron Devices*, vol. **45**, no. 4, pp. 904-911, 1998