Atomistic simulation study of high- κ oxide defects for understanding gate stack and RRAM reliability

Gu, Chenjie

2015

Gu, C. (2015). Atomistic simulation study of high- κ oxide defects for understanding gate stack and RRAM reliability. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/62224>

<https://doi.org/10.32657/10356/62224>

**ATOMISTIC SIMULATION STUDY OF HIGH-K
OXIDE DEFECTS FOR UNDERSTANDING GATE
STACK AND RRAM RELIABILITY**

GU CHENJIE

School of Electrical & Electronic Engineering

Nanyang Technological University

2015

**ATOMISTIC SIMULATION STUDY OF HIGH-K
OXIDE DEFECTS FOR UNDERSTANDING GATE
STACK AND RRAM RELIABILITY**

GU CHENJIE

School of Electrical & Electronic Engineering

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

2015

ACKNOWLEDGEMENT

My sincere gratitude to my supervisor, *Prof. Ang Diing Shenp* from Nanyang Technological University (NTU), for his kindness that offer me the opportunity to study in NTU. His sound technical guidance, inspiring ideas, constructive discussions and constant encouragement are always the most important driving force throughout the entire period of my Ph.D study. 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. Ang D. S.* 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. Ang D. S.*'s supervision.

I would like to extend my sincere appreciate to Dr. Wang Xinpeng, Dr. Fang Zheng and Wen Baobing in the Institute of Micro-Electronics for their sample preparation, device measurement support and fruitful discussion.

I would like to acknowledge all the lab staffs and technicians in NTU for their technical support on the prober station and measurement equipment.

Many thanks to my group members (*Mr. Hu Youzhou, Mr. Du Guoan, Mr. Teo Zhiqiang, Mr. Ho Junjie, Mr. Gao Yuan, Miss. Boo Ann Ann, Mr. Yew Kuang Sing, Miss. Duan Tianli, Mr. Tham Wai Hoe, Mr. Zhang Haizhong, Mr. Zhou Yu, and Mr. Tung Zhi Yan*) 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, parents in-law and wife (*Miss Chen Jiali*). Their endless love and encouragement 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	X
LIST OF TABLES	XXVI
LIST OF ABBREVIATIONS	XXVII
LIST OF SYMBOLS	XXIX
<i>CHAPTER 1 INTRODUCTION</i>	<i>1</i>
1.1 Background.....	1
1.1.1 Transistor Scaling and Gate Dielectrics Evolution.....	1
1.1.2 Bias Temperature Reliability of the Advanced Gate Stack	3
1.1.3 High- κ Material for Emerging Memory Application	7
1.2 Motivation of the Study.....	8
1.3 Objectives of the Study.....	12
1.4 Organization of the Thesis.....	13
1.5 Original Contributions.....	14
<i>CHAPTER 2 LITERATURE REVIEW</i>	<i>17</i>
2.1 Introduction	17
2.2 An Overview of the Bias Temperature Instability in the MOSFET.....	18
2.3 Reliability Issue of the SiON Gate Stack	19
2.3.1 The Early Understanding of NBTI Degradation.....	19
2.3.2 The Impact of Nitrogen Incorporation on NBTI Degradation.....	21
2.3.2 First-principles Simulation for BTI study in the MOSFET	34
2.3.3 The Oxide Switching Traps Model for the NBTI in the SiON Gate Dielectrics	40
2.4 Reliability Study of the Hafnium Oxide Gate Stack	45

2.4.1	Brief Introduction of the Hafnium Oxide as the Gate Stack	45
2.4.2	Poor Interface and Bulk Quality of the Hafnium Oxide Gate Stack	48
2.4.4	Bias Temperature Instability in the Hafnium Oxide Gate Stack	53
2.4.5	The Impact of La Doping on the Reliability Issue in the HfO ₂ Gated Device	61
2.5	The Implementation of Hafnium Oxide as the RRAM Stack	65
2.5.1	The Limitation of the FLASH Memory	65
2.5.2	An Overview of RRAM	67
2.5.3	Current Understanding of RRAM Switching Mechanism	72
2.5.4	The Effect of Capping Layer on the HfO ₂ Based RRAM	75
2.5.5	The Impact of the Electrode Material on the HfO ₂ Based RRAM Switching	77
2.5.6	First-principles Simulation on the RRAM	80
2.6	Summary	83
CHAPTER 3 FIRST-PRINCIPLES SIMULATION AND METHODOLOGY		84
3.1	Introduction	84
3.2	Density Functional Theory	85
3.3	Band Gap Correction	87
3.4	Defect Formation Energy and Charge Transition Level	90
3.5	Summary	92
CHAPTER 4 ATOMIC VIEW OF RECOVERABLE AND PERMANENT COMPONENT		94
4.1	Impact of Two-fold Coordinated Nitrogen on the Generation of Deep-level Hole Traps under Negative-bias Temperature Stressing	94
4.1.1	Introduction	94
4.1.2	Experimental Details	97
4.1.3	Deep-level Hole Trap Generation	97
4.1.4	Simulation Details	106
4.1.5	Discussion on Simulation Results	108
4.1.6	Conclusion	121
4.2	Structure, Property and Dynamics of K _N Center in SiON Dielectrics	122
4.2.1	Introduction	122

4.2.2	First-principles Modeling and Dynamics of K_N Center in SiON Dielectric	122
4.2.3	Results and Discussion	125
4.2.3	Conclusion	129
CHAPTER 5 PROPERTY AND DYNAMICS OF OXIDE DEFECTS IN HIGH-κ GATE STACK		130
5.1	The Evolution of Switching Oxide Traps in the High- κ Gate Stack under NBTI and PBTI	130
5.1.1	Introduction	130
5.1.2	Experiment and Simulation Details	131
5.1.3	The Evolution of Hole Trapping in the High- κ gate p-MOSFET	133
5.1.4	The Evolution of Electron Trapping in the High- κ gate n-MOSFET	140
5.1.5	Atomic Modeling and Possible Oxide Defect Discussion	147
5.1.6	Conclusion	158
5.2	Impact of Lanthanum on Positive-Bias Temperature Instability-Insight from First-principles Simulation	160
5.2.1	Introduction	160
5.2.2	La Induced Slowing Down of PBTI Recovery	161
5.2.3	First-principles Simulation and Discussion	166
5.2.4	Conclusion	174
CHAPTER 6 NEW UNDERSTANDING OF VACANCY GENERATION MECHANISM FOR IMPROVED RESISTANCE SWITCHING IN THE RRAM DEVICE		175
6.1	Introduction	175
6.2	Experiments Details	176
6.3	Experimental Results and Discussion	177
6.4	First-principles Simulation and Model Discussion	180
6.4.1	Simulation Results Discussion	180
6.4.2	Physical Model Discussion	183
6.5	Conclusion	186
CHAPTER 7 CONCLUSION AND RECOMMENDATIONS		188
7.1	Conclusion	188

7.2	Recommendations	190
	LIST OF PUBLICATIONS	193
	BIBLIOGRAPHY	194

ABSTRACT

Bias temperature instability (BTI) is widely recognized as a critical reliability issue for the state-of-the-art complementary metal-oxide-semiconductor (CMOS) device. Generally, BTI is result from a progressive building up of the interface states and bulk oxide charges under the combining effect of the gate voltage and an elevated temperature. These charged defects will shift the threshold voltage (V_t) of the metal-oxide-semiconductor field-effect transistor (MOSFET) and degrade the inversion charge mobility, reducing the device driving current. Worse case, when the shift of V_t exceeds the design tolerance, functional failure of integrated circuits may result.

Previous studies of the NBTI on the SiON gated device show that the threshold voltage degradation is typically characterized by a much weaker dependence on the stress time and temperature as compared to that of the SiO₂ counterpart. Moreover, in a recent NBTI study, Ang *et al.* reveals the presence of deep-level hole traps (DLHTs) after negative-bias temperature stressing. Their experimental results show that the hole traps have energy states above the Si Fermi potential, which enable them have the ability to retain the positive charges for a very long time after the negative gate biasing. However an unambiguous understanding of the role of nitrogen on the DLHTs generation is still not available. Therefore, in this work, we investigate the impact of nitrogen on the trap level generation via first-principles modeling and simulation. Our results show that the coordination of the neighboring nitrogen atom plays a paramount role on the structural relaxation of V_O following the capture of a hole. In particular, a two-fold coordinated nitrogen atom is shown to induce very significant structural relaxation of V_O . The resultant trapped-hole site has a very deep charge transition level

in the SiO₂ band gap (DLHTs). On the other hand, in the absence of the nitrogen atom or if the neighboring nitrogen atom is three-fold coordinated (*e.g.*, the dangling nitrogen bond is terminated by a hydrogen atom). The resultant trapped-hole site would have a much shallower charge transition level. This set of simulation results show a clear correlation between nitrogen in the gate oxide and the DLHT generation by negative-bias temperature stress. Besides, in Ang's work, a broad energy distribution of stress induced positive trap states across over the band gap were detected, whereas knowledge on how the neighboring nitrogen atoms affect the defect levels in the SiON is quite limited. Hence, in the subsequent modeling and simulation work, we modeled various oxide defects with different nitrogen configurations in the SiON gate stack. Our simulation results from these models show a wide-spread of the defects levels within the SiON gap due to the different dynamics of these defect morphors. Distinctive relaxation characteristics of deep-level hole traps, as compared to those of the interface states, are also observed.

As the scaling of the transistors continuously moves on, the SiO₂ or SiON dielectric becomes too thin to sustain the insulator property and causes high gate leakage current due to the direct tunneling. Therefore, hafnium based high dielectric constant oxides are introduced as a replacement of the conventional SiO₂/SiON dielectrics in sub-65 nm technology nodes. However, the inherently higher bulk defect density as well as the poor interface between the HfO₂ and Si brings more severe reliability challenges. In Gao *et al.* dynamic BTI study on the n-MOSFET (DNBTI) and p-MOSFET (DPBTI), the results reveal a group of defects that can cyclically be charged and discharged under low stress voltage- switching oxide traps (SOTs), but when the stress voltage/temperature is increased to certain value, there is a conversion of the cyclic component to the relatively permanent one. Although preliminary studies

ascribed this phenomenon to the evolution of the oxygen vacancies under stress condition, the stress induced leakage current (SILC) as well as the BTI relaxation behavior can't be fully covered. Hence, in this work, we propose a novel defect morphor-- V_O-O_i , by taking account into the coupling of the oxygen vacancy and the interstitial oxygen in the high- κ material. Through calculating the charge coupling force and evolution barrier of the V_O-O_i defect pair, our simulation results give an unambiguous picture of the dynamics of the recoverable and permanent component in the HfO_2 . The physical procedure of the conversion between these two components is entirely reproduced by this defect model. Besides, the PBTI degradation study on the La doped HfO_2 n-MOSFET show an extremely slow recovery rate compared to those pure HfO_2 or SiON gated device. In view of this reliability issue, the V_O-O_i defect pair model is then used to study the effect of lanthanum (La) doping on the PBTI of the HfO_2 gate stack. Our simulation results show that the presence of the La eases the generation of the V_O-O_i pairs, and the generated La associated V_O-O_i pair typically shows deep electron trap characteristics, which ultimately gives the physical explanation for the slow recovery phenomenon.

Finally, we extend our understanding of the hafnium oxide to the RRAM application. Since the optimization of the RRAM performance and the RRAM switching mechanism are still quite controversial, therefore we perform both electrical measurement and simulation work to build up the understanding on the RRAM working principle. Our results show that the first forming step is the most critical process, which determines the switch capability of the generated conductive filament in the RRAM. Using of Pt as the bottom electrode mediates the conductive filament generation, which significantly improves the resistance switching characteristics of the TiN/ HfO_2 /Pt device.

List of Figures

Figure		Page
1.1	Typical MOSFET structure for (a) p-MOSFET and (b) n-MOSFET, respectively. When voltage applied to the gate the electrode, inversion layer will be formed at the channel surface, connecting the source (S) and drain (D) terminals together.	4
1.2	The effect of the applied voltage on the saturated value of N_{FB} after BT treatment with the metal electrode negatively biased [14].	5
2.1	Si/SiO ₂ interface structural model. The P _b center is labeled in the figure.	20
2.2	(a) The V_{th} shift due to the NBTI for pure and nitrided oxide. A large V_{th} shift was observed for nitrided oxide. (b) The extrapolated device life time for different stress voltage [9].	21
2.3	(a) Initial gate leakage of the device, nitridation doesn't increase the initial leakage. (b) The increment of the gate leakage after the NBTI stress, large increment of the nitrided sample is observed. [9].	22
2.4	NBT-stress-time dependence of the ΔV_{TH} (a) and temperature-dependence of ΔV_{TH} ; (b) for p-MOSFET having SiON/SiO ₂ film. The oxide thickness is 2.3 nm. It is found that NBT degradation of SiON is more remarkable than that of SiO ₂ , and the activation energy of ΔV_{TH} for SiON ($E_a \sim 0.1$ eV) is lower than that for SiO ₂ ($E_a \sim 0.2$ eV). These results are consistent with Kimizula N.'s work [8].	23
2.5	NBT-stress-time dependence of ΔV_{TH} for p-MOSFET having SiON and SiO ₂ film. BT stress was performed at 140 °C and under $E_{stress} = -7.5$ MV/cm [8].	24
2.6	Correlation between interface-state generation (ΔD_{it}) and threshold voltage shift (ΔV_{TH}) in both cases of SiON and SiO ₂ films. Note that ΔD_{it} - ΔV_{TH} correlations of SiON and SiO ₂ films are clearly split. Moreover, it is found that respective ΔD_{it} - ΔV_{TH} correlations are independent of stress condition, temperature and fluorine dosage [8].	25

2.7	Schematic diagram for mechanism of NBTI degradation. It is inferred that two kinds of origins contribute to NBT degradation. One is the hydrogen-related process such as Si-H bond-breaking (a) the other is hydrogen-unrelated structure such as two-fold coordinated N contributes to NBT degradation (b) [8].	26
2.8	C-V shift (ΔV_{FB}) and increase in interface state density (ΔD_{it}) of H ²⁻ and D ²⁻ annealed plasma-nitrided SiON/Si diodes plotted against duration of NBTS [60].	27
2.9	ESR signals (first derivative) of the P _b centers detected for (a) the NBTS-applied plasma-nitrided SiON/Si, (b) thermally depassivated plasma-nitrided SiON/Si, and (c) thermally depassivated SiO ₂ /Si samples. The magnetic field was set parallel to the [100] axis [60].	27
2.10	(a) Optimized SDR of NBT-induced defects in 7.5 nm SiO ₂ and 2.3 nm PNO-based devices as a function of g-value. NBTI generates different defects in SiO ₂ and PNO devices. (b) SDT ($V_F = -0.2$ V) and SDR ($V_F = +0.2$ V) spectra of PNO defect both measurement clearly show the generation of a single dominating signal at $g = 2.00020$. The spectrometer gain is 100 times greater for the $V_F = -0.2$ V spectrum [61].	28
2.11	K_N center as proposed by Campbell, J. P. <i>et al.</i> The number of nitrogen atom attached to the Si depends on the specific defect configuration.	30
2.12	Threshold voltage shift $ \Delta V_{th} $ of device A, stressed at -2.6 V gate bias for 5×10^4 s, as a function of $(\kappa T)^{-1}$ (open circles), showing the presence of two distinct degradation mechanism (I and II). Also shown for comparison is the Arrhenius plot of a 12-nm SiO ₂ gate p-MOSFET (L/W=0.6/0.4 $\mu\text{m}/\mu\text{m}$), stressed at comparable oxide field. Degradation due to mechanism I (activation energy ~ 0.037 eV), is not observable in this case. Eliminating the contribution of mechanism I in Device A yields the Arrhenius plot (filled circles) for mechanism II, whose activation energy (0.24 eV) agrees very well with that of the SiO ₂ gate device. Inset: schematic energy-band diagram of the p-MOSFET under negative gate bias, depicting transfer of electronic charge at deep defect states introduced by nitrogen to holes in the inversion layer [25].	32
2.13	Arrhenius plots for NBT stress-induced threshold voltage shift $ \Delta V_{th} $ of device A and B. The circles denote the respective Arrhenius plots of devices A and B after eliminating the contribution of mechanism I [25].	33
2.14	Cluster model of the Si/SiO ₂ and Si/SiO _x N _y interface structures.	35

	The reaction energies (E_R) for the trapping of the hydrogen ion (H^+) are obtained from first-principle calculations after performing geometry optimization on both reactants and products [59].	
2.15	Relative reaction energies ΔE_R of NBTI reaction at pure and nitride oxide interface with different numbers of N as 1st-nearest neighbor for hydrogen-originated hole trapping nitrogen centers [65].	35
2.16	Relative reaction energies ΔE_R of NBTI reaction at oxide interface with different No. of N as 1 st -nearest neighbor for hydrogen-originated hole trapping nitrogen centers. Namely, (a) vacancy defects and (b) oxygen [65].	36
2.17	Relative reaction energies ΔE_R vs. the No. of N atoms as 1 st -nearest neighbor for three different types of hydrogen trapping centers, namely vacancy defect, oxygen and nitrogen [65].	37
2.18	Atomic structures of the (a) N_i , (b) N_i-O_i , (c) N_i-O_{2i} , and (d) $N_i-O_i-O_r$ defects at the interface [70].	38
2.19	Threshold-voltage shift of the p-MOSFET $ \Delta V_{th} $ versus the stress induced interface-state density ΔN_{it} extracted from CP current (I_{cp}) measurement. I_{cp} was measured using a 1-MHz +1/-1-V gate-voltage pulse, with a rise/fall time of 100 ns. The dashed line denotes $ \Delta V_{th} $ ($=q\Delta N_{it}/C_{ox}$; C_{ox} is the gate oxide capacitance) calculated based on the extracted ΔN_{it} . Static stress was carried out with a gate voltage of -2.6 V, corresponding to an oxide field of ~8.8MV/cm. Bipolar stress was achieved using a 100-kHz 50% duty cycle gate pulse, with a 50-ns rise/fall time. The positive (negative) level was set at +1.5 V (-2.6 V). Stress temperature was 125 °C. The inset compares the ΔN_{it} and $ \Delta V_{th} $ of the static and bipolar stresses at equivalent stress time T_{eq} ($=50\%$ of total clock time in the case of bipolar stress) [25].	42
2.20	Amount of recovery per cycle, δI_{cpr} is observed experimentally (open circle) to remain unchanged after many stress/recovery cycles. The same observation holds over a wide range of V_{gs} . (not shown). On the other hand, the <i>R-D model</i> predicts a steady decrease of δI_{cpr} by ~50% after 30 stress/recovery cycles (filled circle); results obtained via simulation using the same stress conditions and device geometry). In the R-D context, δI_{cpr} decreases because hydrogen is transported further away from the Si/SiO ₂ interface after numerous stress/recovery cycle; (b) δI_{cpr} increases with V_{gs} , indicating that it is stress induced [48].	43
2.21	δI_{cpr} (circle) is observed to be independent of t_s/t_r . This result	44

	doesn't corroborate a transport-limited mechanism of N_{it} generation and recovery as stipulated by the <i>R-D model</i> [47].	
2.22	(a) Dynamic NBTI after subtracting away the hole-trapping component; (b) the evolution of $ \Delta V_{t,r} $ during DNBTI; (c) Simulated evolution of $ \Delta V_t $, using a diffusion framework for molecular hydrogen; (d) comparison of $ \Delta V_{t,r} $ and $ \Delta V_{t,s} $ [47].	44
2.23	Variation of dielectric constant with band gap of high- κ oxides [79].	47
2.24	C-V characteristic of Al_2O_3 , HfO_2 and SiO_2 NGETs. Significant hysteresis is observed for the HfO_2 dielectric layer [81].	48
2.25	$(\alpha h\nu)^{1/2}$ vs $h\nu$ plot of the samples prepared at different temperature [14].	49
2.26	Dependence of the absorption coefficient and annealing time and temperature in oxygen ambient. Absorption coefficient of the extra peak was integrated over photon energy for the quantification, as shown in the inset [14].	49
2.27	Schematic illustration of the back-bonded HfO_2 silicon interface dangling bond defect [83].	50
2.28	(a) Oxygen vacancy; (b) oxygen interstitial defect in the amorphous HfO_2 .	51
2.29	Summary of the defect energy levels of oxygen vacancy and interstitial O under various charge states in the HfO_2 [79].	51
2.30	(a) p-MOSFET I_d - V_g characteristics before and after NBTI. V_t shift and slight G_m degradation are observed. (b) V_t and G_m as functions of time during NBTI for without SN. G_m degradation is less significant than V_t degradation [88].	53
2.31	Dependence of NBTI-induced threshold voltage shift (ΔV_{th}) on stressing conditions for HfO_2 p-FETs with tungsten gates. Symbols indicate measured data and solid lines are fits using model (7). (a) dependence of ΔV_{th} versus stressing tie curves at 125 °C; (b) dependence of ΔV_{th} versus stressing time curves on stress temperature at $V_s=-1.5V$ [52].	53
2.32	Dependence of NBTI induced V_t shift for various pFETs with FUSI NiSi gate: (a) $SiO_2/HfO_2/NiSi$, (b) $SiO_2/NiSi$, and (c)	53

- SiON (~5% N)/poly-Si; symbols are measured and solid line are model fits. (d) The extrapolated life time for various devices under different voltage [52].
- 2.33 PBTI for NiSi gated nFETs with SiO₂, SiO₂/HfSiO & SiO₂/HfO₂ as gate dielectrics; symbols are measurements and solid lines are model fitting. PBTI becomes worse with increasing Hf content in the dielectric layer [52]. 54
- 2.34 Comparison between PBTI (open symbols) and NBTI (filled symol) for SiO₂/HfO₂/NiSi FETs. (a) ΔV_t versus stress time at constant stress field; (b) Predicted ΔV_t at 10 years of stressing as a function stress field; the predicted values are estimated from model fitting. PBTI is worse than NBTI in HfO₂/NiSi FETs [52]. 55
- 2.35 The wave form used for DPBTI. Each stressing and recovery phase was set to 1000 s, and totally 30 cycles are carried. 56
- 2.36 The evolution of R as a function of the number of stress/relax cycles under various oxide stress fields at a given temperature of 100 °C. R is constant at 5.5 MV/cm but shows a progressive decrease at higher oxide stress fields [27]. 56
- 2.37 The increase in ΔV_t^{eos} , with respect to the ΔV_t^{eos} of the first stress/relax cycle (denoted as $\delta\Delta V_t^{\text{eos}}$), are comparable for the three cases but the increase in ΔV_t^{eor} , with respect to the ΔV_t^{eor} of the first stress/relax cycle (denoted as $\delta\Delta V_t^{\text{eor}}$) is much larger for the 7 and 8.5 MV/cm oxide stress field as compared to that of the 5 MV/cm oxide stress field. This indicates that the decrease in R at higher oxide stress fields is due to a part of R being transformed into more permanent ΔV_t [27].. 57
- 2.38 (a) Relaxation traces measured after AC and DC stress with tress time changing from 1 sec to 10000 sec. notice that the AC relaxations always start with lower V_t shift and shallower slope, but they gradually merges with the corresponding DC relaxation at longer relaxation times; (b) Comparison of relaxation traces after DC stress and AV stress with different duty cycles. All the relaxations merge into each other after longer relaxation times [89]. 58
- 2.39 (a) Occupancy level of defects after AC stress with frequency 1Hz, duty factor 99% (ON/OFF ratio = 100) for a stress time $t_{\text{AC}} = 100$ ks; the status is taken after a full charging cycle with a 1 μ s measuring (=recovery) delay. The diagonal line marks $\tau_c = \tau_e$, the shifted diagonal is shifted by a factor 100 corresponding 60

to the ON/OFF ratio. The white cross marks the AC period-time; the green dashed rectangle marks the charging and discharging periods of the AC signal. The dotted green trajectories mark lines of const. charging times τ 1 s and 100 ks. Right hand labels denote the filling state corresponding to the color code. The black horizontal and vertical dash/dotted lines denote the net stress time and the recovery time (= measuring delay), respectively. (b) Recovery curves all taken after a net stress time of 100 s with various duty factors. The arrows mark the position where the AC-curves and the DC-curve (100%) are expected to merge [99].

2.40	Final profile of the La atom after annealing through EELS [94].	62
2.41	Both the La-doped HfO ₂ gate stack and the control HfO ₂ had the same final physical thickness (~2.5 nm); EOTs were also similar (~0.95 nm), as determined from CVC simulation. The CV result shows the leftwards shift (~300 mV) of the flat-voltage [94].	62
2.42	I_d - V_g characteristics and V_{th} uniformity for samples with and without La; (b) effective electron mobility of La doped HfO ₂ MOSFET. The EOT of the samples are 1.5 nm. We can observe an obvious effective mobility improvement with the La doping [92].	63
2.43	Gate-leakage characteristics for gate injection and substrate injection. The La-doped devices show a higher V_{bd} for both accumulation and inversion regimes [94].	63
2.44	Time evolution of ΔV_{th} in PBTI. As an effect of La incorporation, ΔV_{th} were suppressed in the same over-drive stress [93].	64
2.45	ΔV_{th} in PBTI as a function of the extent of conduction electron injection for (a) HfSiON and (b) HfLaSiON. ΔV_{th} depends on stress bias. This indicates that trap sites were increased with stress [93].	64
2.46	Typical floating gate FLASH memory cell. By controlling the number of the trapped electron in the floating gate, the threshold voltage of the cell can be modified.	66
2.47	The experimental i - v curves for the Pt-TiO _{2-x} -Pt device under voltage sweep [97].	68

2.48	(a) Typical RRAM cell with MIM structure. Usually use the transition metal oxide as the insulator layer, but the top/bottom electrodes can be various conductive materials. (b) Typical switching curves for bipolar mode (circle) and unipolar mode (triangle).	69
2.49	Classification of the resistive switching effects which are considered for non-volatile memory applications. The switching mechanisms based on thermal, chemical, and electronic/electrostatic effects are concisely summarized [99].	72
2.50	Typical current-voltage characteristic of Ag/Ag-Ge-Se/Pt electrochemical metallization cell using a triangular voltage sweep. The “ON” conductance is limited by a compliance current of 25 μ A. The curve in the compliance current state is shown as a dashed line because the actual voltage does not correspond to the voltage shown on the axis. The insets A to D show the different stages of the switching procedure [99].	73
2.51	Schematic switching process of the valence change RRAM stack. (a) pristine RRAM stack, a few pre-existing defects randomly locate in the HfO_x layer; (b) In the forming step, a large number of oxygen vacancies is generated, creating the conductive filament (c) the reset step that back drive the mobile oxygen ion to the vacancies, and eliminate the conductive filament.	74
2.52	XTEM micrographs of the TiN/AlCu/HfO _x /TiN device. (a) Before PMA. (b) After PMA [34].	75
2.53	Bi-stable switching characteristic of TiN/AlCu/HfO _x /TiN MIMC device with a cell size of $65 \times 65 \mu\text{m}^2$. (a) Typical I–V curve of bipolar resistance switch. The inset shows the multilevel operation with different clamped ISET. (b) Uniformity of operation voltages (20 devices) [34].	76
2.54	(a) Comparison of the bipolar resistance switching characteristics between the TiN/TiO _x /HfO _x /TiN devices in 1R and 1T1R configuration. (b) Tight distribution of R_{low} and R_{high} in TiN/TiO _x /HfO _x /TiN device with 1T-1R configuration [34].	77
2.55	(a) Schematic device structure of the HfO _x based RRAM devices. The HfO _x resistive switching layer and the TaTiN top electrode (TE) are deposited by reactive PVD, respectively. (b) Process flow of the device fabrication [102].	78

2.56	Typically I - V characteristics of forming process measured using current sweep and the first reset process after forming [102].	78
2.57	Both bipolar and unipolar resistive switching behaviors were measured in the HfO_x based RRAM. (a) Measured I - V curves for bipolar switching. (b) Measured I - V curves for unipolar switching [103].	79
2.58	Typical unipolar reset and set switching traces obtained for a cell diameter of 100 μm and for both the 5 nm (cross and open symbols) and 15 nm (closed symbols) HfO_2 thicknesses; for the 5 nm thickness, trace 1 (diamonds) programs the cell into the LRS state, then trace 2 (open Squares) is performed from 0 to 1.2 V, then trace 3 (triangles) is performed from 0 to 1.4 V, and finally trace 4 (crosses) is performed from 0 to 1.8 V. Typical consecutive bipolar switching loops between the LRS state and an I_R [107].	80
2.59	Process of V_O diffusion in $\text{HfO}_{1.97}$ during a time-accelerated molecular dynamic at 1573 K. the cyan and red atoms corresponded to Hf and oxygen, respectively. The grey densities correspond to the localization of electronic densities bound to the oxygen vacancies [104].	82
3.1	Calculated versus measured single-particle band gaps for 15 different materials. PBE: open disks, PBE0: filled disks [117].	89
3.2	Formation energy of interstitial hydrogen as a function of Fermi level in to different semiconductors, illustrating the qualitatively different position of the $\epsilon(+/-)$ level. In GaN (a), $\epsilon(+/-)$ lies within the band gap, whereas in ZnO (b) $\epsilon(+/-)$ lies above [118].	92
4.1	(a) Evolution of threshold voltage shift ΔV_t during negative-bias temperature stressing and relaxation. Gate voltages applied in the respective phases are as indicated. During relaxation, the gate voltage was first switched to 0 V and maintained at that level for 4×10^3 s. It was then switched to +1.5 V for 1×10^3 s (but PIV measurement was carried out with a 0-V offset) after which it was returned to 0 V for another 4×10^3 s. (b) ΔV_t evolution during the first 0 V and the subsequent +1.5 V relaxation phases plotted on a linear-log scale. The slopes of the regression line are as shown.	98
4.2	Schematic energy band diagrams of the MOS structure under negative gate biasing. (a) Hole capture by a near-interface	99

defect state where the trapping probability is greater than the de-trapping (dashed arrow) probability. (b) In an ultra-thin gate oxide, hole capture by a defect state located further from the Si/SiO₂ would be less likely because of a correspondingly large probability of de-trapping to the gate, *i.e.*, the defect is neutral and does not contribute to NBTI. In both (a) and (b), structural relaxation of the defect is ignored. (c) Significant lattice relaxation (via interaction with phonon) upon hole capture results in an “upward” shift of the positively charged defect state (*i.e.*, negative U behavior) and the creation of a deep-level hole trap.

4.3	(a) Comparison of $ \Delta V_{t,r} $, the decrease in the magnitude of ΔV_t per relaxation cycle as a function of the number of stress/relax cycles between experiment and simulation based on the <i>R-D model</i> . For the ease of comparison with the simulation data, $ \Delta V_{t,r} $ is normalized by the value for the first relaxation cycle. (b) Comparison of $ \Delta V_{t,r} $ and $ \Delta V_{t,s} $, the increase in the magnitude of ΔV_t per stress cycle. $ \Delta V_{t,EoR} $ refers to the magnitude of ΔV_t at the end of each relaxation cycle [48].	100
4.4	Dependence of threshold voltage shift (extracted after 1×10^4 s stress) contributed by shallow and deep hole traps on the nitrogen concentration at the polysilicon/SiO ₂ interface in atomic percent.	104
4.5	Evolution of threshold voltage shift $ \Delta V_t $ arising from shallow hole traps (SHTs) and deep-level hole traps (DLHTs) for DUTs having varying nitrogen concentration at the polysilicon/SiO ₂ interface.	105
4.6	Supercell comprising 8 α -cristobalite unit cells used in our simulation study. The bigger blue spheres denote Si atoms while the red spheres denote oxygen atoms.	106
4.7	(i) Oxide bonding network in the vicinity of the neutral oxygen vacancy defect, depicting the distance between Si(1) and Si(2). (ii) Electron density analysis showing the distribution of electronic charge (yellow) in the same vicinity of the oxide network. The dotted circle highlights the bonding between Si(1) and Si(2).	109
4.8	(i) Oxide bonding network in the vicinity of the oxygen vacancy defect following the capture of a hole. (ii) Electron density analysis showing the distribution of electronic charge (yellow) in the same vicinity of the oxide network. The dotted circle highlights the electronic charge density between Si(1) and Si(2). The dotted circle is not shown in part (c)(ii) as there is no electronic charge between Si(1) and Si(2) in this case.	109

4.9	Electron localization function analysis in the vicinity of the oxygen vacancy defect for the case of the twofold coordinated nitrogen atom, before (a) and after (b) the capture of a hole. The yellow regions denote iso-surfaces having a value of 0.85 (a value of 1 means the electronic charge is completely localized). An enlargement of the iso-surface between N and Si(1) following the capture of a hole signifies increased electronic charge localization between the two atoms. There is also a change in the structure of the Si(1)O ₂ N cluster from tetrahedral to planar after the hole capture.	111
4.10	Formation energy versus Fermi level ($E_F = 0$ corresponds to the oxide valence band edge) characteristic for a positively charged oxygen vacancy defect. Intersection with the $E^f = 0$ line gives the charge transition level (CTL), which denotes the electron chemical potential energy for transition from neutral to a positively charged state. For $E_F < \text{CTL}$, the positively charged state is stable.	113
4.11	Schematic diagram illustrating the underestimation of the oxide band gap using the semilocal density functional scheme (left), which can be corrected by the hybrid functional scheme (right). After correction, the charge transition levels (CTLs) are more accurately located with respect to the oxide and Si band edges. The oxygen vacancy defect with a twofold neighboring N gives rise to the deepest CTL (<i>i.e.</i> , furthest above the oxide valence band edge).	114
4.12	Evolution of threshold voltage shift $ \Delta V_t $ during negative-bias temperature stressing and relaxation. Unlike in Fig. 4.1 where the +1.5 V gate voltage was interrupted and set to 0 V during PIV measurement, the gate voltage in this case was maintained at +1.5 V throughout the relaxation (<i>i.e.</i> , PIV measurements were made at a dc offset of +1.5 V). A step-like increase in $ \Delta V_t $ is evident when the gate voltage was subsequently switched to 0 V (big upward arrow). (b) A magnitude view of the $ \Delta V_t $ increase at the +1.5 V to 0 V transition. Such a step-like change is not observed in a pristine device.	118
4.13	Pure SiO ₂ ; (b) Si-O-N with one N attached to the Si(1); (c) Si-O-2N(i) with two N atoms attached to the same Si(1); (d) Si-O-2N with two N atoms attached to Si(1) and Si(2) separately (ii); (e) Si-O-3N(i) with three N atoms attached to the same Silicon; (f) Si-O-3N(ii) with two N attached to Si (1) and the other to Si(2).	124
4.14	sMSD results for (a) Pure SiO ₂ ; (b) Si-O-N; (c) Si-O-2N(a); (d) Si-O-2N(b); (e) Si-O-3N(a); (f) Si-O-3N(b).	125

4.15	V_O structure relaxation cases (a) significant structure change after the hole trapping; (b) mild structure change after the hole trapping.	126
4.16	Extracted CTLs for (a) Pure SiO_2 ; (b) Si-O-N ; (c) Si-O-2N(i) ; (d) Si-O-2N(ii) ; (e) Si-O-3N(i) ; (f) Si-O-3N(ii) .	127
5.1	Evolution of the threshold voltage of an n-MOSFET under certain dynamic PBTI cycles. R denotes the recoverable component of a specific cycle and is defined as the difference of the threshold voltage shift at the end of the stress phase and at the end of the following relaxation phase. P denotes the remnant ΔV_t and is a measure of the part of the SOTs which did not discharge at the end of each relaxation phase. Similar definitions apply for the case of DNBTI.	134
5.2	Evolution of NBTI recovery per relaxation interval of a (a) HfO_2 gate; (b) HfSiON gate p-MOSFET subjected to stress/relaxation cycling at different oxide stress fields. The oxide stress field is calculated based on the equivalent oxide thickness of the gate stack. The exact oxide stress field in the 4-nm thick HfO_2 is smaller than that in the 2-nm thick HfSiON .	135
5.3	Evolution of NBTI recovery (denoted as R) per relaxation interval of a (a) HfO_2 gate; (c) HfSiON gate p-MOSFET subjected to stress/relaxation cycling at different temperatures. (b) and (d) show the evolution of the corresponding R, normalized with respect to that for the first relaxation interval.	136
5.4	Increase of threshold voltage shift, $ \Delta V_t $ measured at the end of stress (eos) and end of relaxation interval of each dynamic NBTI cycle, with reference to the respective $ \Delta V_t ^{\text{eos}}$ and $ \Delta V_t ^{\text{eor}}$ values of the first cycle. The (a) HfO_2 and (b) HfSiON gate p-MOSFETs are the same as those in Fig. 5.3.	137
5.5	Gate current I_g versus gate voltage V_g curves for the same (a) HfO_2 and (b) HfSiON gate p-MOSFET in Fig. 5.2, before and after dynamic NBTI stressing which yields a (i) constant; (ii) decreasing recovery.	137
5.6	Gate current I_g versus gate voltage V_g curves of a HfO_2/TiN p-MOSFET showing an increased current only when there is a decrease of $ \Delta V_t $ recovery during dynamic negative bias temperature stressing. I_g was measured after the 30th relaxation interval (Fig. 5.1). The oxide stress fields are as indicated: (a) 5.5 MV/cm; (b) 7 MV/cm; (c) 8.5 MV/cm.	138
5.7	Threshold voltage shift, ΔV_t of the HfO_2/TiN gate n-MOSFET subjected to positive bias temperature stress (at +1.8 V, 100 °C), followed by relaxation at 0 V and then -1 V gate bias. The	141

- second abrupt decrease of ΔV_t may be attributed to the emission of electrons trapped at energetically deeper oxide traps (Fig. 5.8).
- 5.8 Schematic energy band diagram of the HfO₂/TiN gate stack under (a) 0 V and (b) -1 V gate bias. Under 0 V gate bias, an electron trapped at an energetically deep oxide trap could not be readily emitted due to the energy barrier ϕ between the trap level and the Si conduction band edge. A negative gate bias ‘raises’ the trap level (thus reducing the barrier ϕ), allowing the trapped electron to be now easily emitted to the Si conduction band. 141
- 5.9 Recovery of threshold voltage shift, ΔV_t per relaxation interval as a function of the number of positive bias temperature stress and relaxation cycles applied on the HfO₂/TiN gate n-MOSFET. The oxide stress fields are as indicated. The decrease of ΔV_t recovery in the case of 7 and 8.5 MV/cm stress is clearly suppressed under a negative gate relaxation voltage. 142
- 5.10 Change of ΔV_t^{eos} and ΔV_t^{eor} (with respect to corresponding values of the first stress/relax cycle) during PBTI stress/relax cycling under different oxide stress fields. A much faster increase of ΔV_t^{eor} as compared to ΔV_t^{eos} is evident at the 7 MV/cm stress. The HfO₂/TiN n-MOSFETs are the same as those in Fig. 5.9. 143
- 5.11 PBTI stressed induced shift of shallow to deep electron trapping ‘deactivates’ the oxide traps involved, resulting in reduced cyclical ΔV_t fluctuation in subsequent stress/relax cycling. The application of an opposite (negative) gate relaxation polarity could ‘reactivate’ most of the traps. The sequence of tests applied on a HfO₂/TiN gate n-MOSFET is as follows: Phase I: The n-MOSFET was subjected to PBTI stress (@ 5.5 MV/cm) and relaxation (@ 0 V) for 5 cycles and the ΔV_t recovery for each cycle noted; Phase II: The device was then subjected to PBTI stress (@ 7 MV/cm) and relaxation (@ 0 V) for 30 cycles (evolution of ΔV_t recovery in this case is similar to that shown in Fig. 5.9(b)); Phase III: A repeat of phase I gave a smaller ΔV_t recovery (downward arrow); Phase IV: The device was relaxed under -1 V gate bias; Phase V: A repeat of phase I saw a restoration of ΔV_t recovery (upward arrow). 144
- 5.12 The $|\Delta V_t|$ recovery per cycle for dynamic NBTI stress (oxide stress fields are as indicated) is clearly increased under a positive gate relaxation voltage (+1 V), and the result may be ascribed to the discharge of deep-level hole traps. But unlike PBTI (Fig. 5.8), the positive gate relaxation voltage is found to 145

- have no apparent impact on the decrease of $|\Delta V_t|$ recovery per cycle observed at comparable oxide stress field (7, 8.5 MV/cm). The decrease of $|\Delta V_t|$ recovery is nearly the same as that obtained under a 0 V gate relaxation voltage, implying that the transformed hole traps are more permanent.
- 5.13 The impact of PBTI stressing on SILC or bulk oxide trap generation in the HfO_2/TiN gate n-MOSFET. The gate current measurement was performed at the 30th relaxation interval (Fig. 12). Unlike the case of NBTI stress (Fig. 6), only very marginal SILC generation is observed at the highest oxide stress field of 8.5 MV/cm, indicating the absence of any significant generation of bulk oxide traps over the range of oxide stress field studied. 146
- 5.14 (a) Ball and stick diagram of a portion of the defect-free amorphous HfO_2 supercell used in our simulation study. The supercell consists of 32 Hf atoms and 64 O atoms and was generated by the melt-and-quench scheme; (b) An O atom in-between the two Hf atoms as shown was removed to simulate the oxygen vacancy defect (arrow). 148
- 5.15 Charge transition level or CTL for the (a) neutral-to-positive ($0/+$) state transition; (b) neutral-to-negative ($0/-$) state transition of all 64 cases of oxygen vacancy defects in the amorphous HfO_2 supercell used in our simulation study. The upper and lower dashed lines denote the Si conduction and valence band edge, respectively. The corresponding band offsets are also shown. 150
- 5.16 (a) Ball and stick diagram illustrating an oxygen vacancy cum oxygen interstitial ($V_{\text{O}}\text{-}O_i$) defect in the amorphous HfO_2 used in our simulation study. The O atom in-between two Hf atoms is manually removed and placed randomly in an interstitial position (O_i) in a nearby region, simulating the creation of the $V_{\text{O}}\text{-}O_i$ defect under electrical stress. (b) Schematic energy band diagram of the HfO_2/Si stack showing the charge transition levels for ten arbitrarily chosen O_i positions. Closely spaced CTLs result in a thick line. The CTL for a V_{O} defect is also shown for comparison. 153
- 5.17 (a) Ball and stick diagram of a $V_{\text{O}}\text{-}O_i$ defect yielding a shallow electron trap (CTL = 3.78 eV) in the HfO_2 band gap. The yellow shades denote positive partial electron density obtained after introducing an additional electronic charge into the supercell. The additional electron can be seen to be trapped at the V_{O} site. (b) A different $V_{\text{O}}\text{-}O_i$ defect which yields a deep electron trap (CTL = 1.58 eV). Part of the electron trapping can be seen to occur at the interstitial oxygen O_i . 155

5.18	(a) Ball and stick diagram for the same V_O-O_i defect in Fig. 5.1.6.4(a) functioning as a deep hole trap (CTL = 4.26 eV) in the HfO_2 band gap. The yellow shades denote negative partial electron density obtained after removing an electronic charge from the supercell. The hole is trapped at the V_O site. (b) Ball and stick diagram for the V_O-O_i defect in Fig. 5.17(b) functioning as a shallow hole trap (CTL = 2.56 eV).	155
5.19	Another example of a V_O-O_i defect functioning as a shallow hole trap (CTL = 1.85 eV). The yellow shades denote negative partial electron density obtained after removing an electronic charge from the supercell. Unlike the example shown in Fig. 5.18(b), hole trapping in this case occurs at the O_i site.	156
5.20	(a) (i) A deep oxide trap due to the V_O-O_i defect is not an active trap-assisted-tunneling (TAT) center under positive gate bias, since an electron, once trapped, could be easily be emitted. (ii) The same trap could, however, function as an active TAT center, facilitating the tunneling of holes from the Si valence band towards the metal gate under a negative gate bias. (b) A shallower V_O-O_i oxide trap could function as an active TAT center under both gate polarities, similar to the V_O defect shown in (a).	156
5.21	(a) Linear drain current degradation ΔI_d (in percent), measured at two different delays, of a La-doped n-MOSFET and undoped p-MOSFET (on the same wafer) as a function of stress time. (b) Typical logarithmic time dependence of ΔI_d recovery (expressed as a percentage of the total I_d degradation at the end of stress) of a variety of n- and p-MOSFETs with different gate stacks, subjected to PBTI and NBTI stress, respectively. The number in the bracket refers to the dielectric thickness [29].	162
5.22	(a) Schematic energy band diagram for the TiN/ HfO_2 /Si structure showing (i) a detrapping barrier ϕ of a deep electron trap under $V_g = 0$ and (ii) reduction of the barrier by a negative gate bias, which facilitates emission of the trapped electron. (b) Schematic depiction of a deep hole trap, where emission of the trapped hole (via the capture of an electron from the Si substrate) is hindered by the barrier ϕ . Under a positive gate bias, the hole trap is lowered below the Si conduction band and is neutralized by an injecting electron.	164
5.23	(a) Evolution of threshold voltage shift ΔV_t during a cycle of PBTI stress and relaxation of an undoped HfO_2 /TiN gate n-MOSFET. ΔV_t is expressed as a percentage of the total ΔV_t at the end of the stress phase. (b) Recovery of linear drain current degradation ΔI_d , expressed as a percentage of the total degradation at the end of stressing, for the La-doped n-	165

	MOSFET and undoped p-MOSFET under different gate relaxation voltages.	
5.24	Recovery of oxide defects at and near the SiO_x/Si interface (as probed by the charge pumping method at 1 MHz) of the La-doped n-MOSFET and undoped p-MOSFET following PBTI and NBTI stressing, respectively. ΔI_{cp} is expressed as a percentage of the total ΔI_{cp} at the end of stress [29].	165
5.25	(a) Ball and stick diagram of part of the defect-free HfO_2 supercell used in our simulation study. (b) Two Hf atoms are replaced by La atoms to simulate La doping of the HfO_2 . (c) The O atom in between the two Hf atoms was manually removed to create an oxygen vacancy (V_{O}) defect. (d) The V_{O} defect is formed in between two La atoms in the La doped HfO_2 .	167
5.26	Schematic energy band diagram illustrating the charge transition level (akin to the trap level) of the oxygen vacancy (V_{O}) and vacancy-interstitial ($V_{\text{O}}\text{-}O_{\text{i}}$) defect in the (a) undoped HfO_2 and (b) La-doped HfO_2 .	169
5.27	(a) Ball and stick representation of the amorphous HfO_2 supercell with a $V_{\text{O}}\text{-}O_{\text{i}}$ defect, created by removing one of the O atoms in between two Hf atoms (see labels) and placing it randomly in an interstitial position (O_{i}) close to the V_{O} defect. (b) A similar representation of the $V_{\text{O}}\text{-}O_{\text{i}}$ defect in the La-doped HfO_2 supercell. In this case, the O_{i} was originally located in between the two La atoms.	170
6.1	Schematic device layer structures for different samples. (a) TiN/10-nm HfO_2/TiN ; (b) TiN/5-nm HfO_2/TiN ; (c) TiN/5-nm $\text{HfO}_x(\text{O-deficient})/\text{TiN}$; (d) TiN/5-nm HfO_2/Pt ; (e)) TiN/10-nm HfO_2/Pt ; (f) Ti-rich TiN/10-nm HfO_2/Pt .	176
6.2	(a) TiN/10-nm HfO_2/TiN - (i) A marginal decrease in the current during reset and the resultant memory window is small. (ii) A large difference between the current in the high resistance state (HRS) and the pre-forming state. (b) TiN/5-nm HfO_2/TiN (nearly no switching); (c) TiN/5-nm HfO_x/TiN with a memory window around 10.	178
6.3	(a) With one of the TiN electrodes replaced by Pt and the Pt made the cathode during forming/set, more significant current switching is obtained and the current after reset is returned to the pre-forming level reset (b) Negative forming/set only shows marginal memory window; (c) Behavior similar to the TiN/5-nm HfO_2/Pt device is observed for the TiN/10-nm HfO_2/Pt device; (d). Negative forming/set shows negligible memory window.	179

6.4	Ball-and-stick representation of (a) the amorphous HfO_2 supercell generated using the melt-and-quench scheme. (b) The incorporation of Pt is simulated by replacing an Hf atom with a Pt atom.	180
6.5	(i)(a),(ii)(a) The generation of vacancy (V_O). (i)(b),(ii)(b) Defect site after full structural relaxation under negatively charged condition. (i)(c),(ii)(c) Defect site after full structural relaxation under neutral condition.	181
6.6	Evolution of system energy during V_O generation under negatively charged condition. (b) Pt induced meta-stable state.	182
6.7	Ball-and-stick illustration of the impact of Pt incorporation into the HfO_2 on the surrounding Hf-O bond length. Compared to the case when Pt is absent (a), the Hf-O bond length in the vicinity of the Pt atom is increased (b)	183
6.8	Schematic illustration of (a) vacancy V_O generation during forming/set; (b) recombination between V_O and nearby interstitial oxygen O_i during reset in the TiN/5-nm HfO_2 /Pt device.	184
6.9	(a) In a thicker HfO_2 , more V_O 's are needed to form a percolation path. (b) Although the proximity of O_i 's to the V_O sites in the Pt influence zone could lead to a near-complete removal of the V_O 's, this could not happen as easily in zone X, due to the greater physical separation between the V_O 's and O 's.	185
6.10	(a) Current-voltage curves of the Ti-rich TiN/ HfO_2 /Pt device in the order of forming, reset and set. The current compliance was set to 1×10^{-5} A and 1×10^{-3} A for forming and set, respectively. (b) The significant current switching between forming/set and reset can be consistently observed under repeated cycling.	186

LIST OF TABLES

2.1	Measured magnetic-resonance parameters of important silicon dangling-bond defects. The K center and K_N center parameters are nearly the same [61].	29
2.2	Single-particle defect levels (E_{KS}), defect transition levels ($E-/0$) between neutral and negatively charged states, and Si-O distance for Ni-related defects. E_c indicates the conduction band minimum of Si [70].	39
4.1	A comparison of Si-N bond length before and after the capture of a hole by the oxygen vacancy defect.	111

LIST OF ABBREVIATIONS

BD	Breakdown
BTS	Bias temperature stress
BTI	Bias Temperature Instability
BE	Bottom electrode
C-V	Capacitor-voltage
CMOS	Complementary metal-oxide semiconductor
CTL	Charge transition level
CVS	constant voltage stress
DNBTI	Dynamic Negative Bias Temperature Instability
DPBTI	Dynamic Positive Bias Temperature Instability
DLHT	Deep level hole trap
DFT	Density functional theory
E_a	Reaction energy
E_{stress}	Stress electric field
ESP	Electron spin paramagnetic
ESR	Electron spin resonance
EDMR	Electrically detected magnetic resonance
EELS	Electron energy loss spectroscopy
EOT	Equivalent Oxide thickness
EoR	End of recovery
EoS	End of stress
HKMG	High- κ metal gate
IL	Interfacial thickness
MOSFET	Metal-oxide-semiconductor field-effect transistor
MIM Cap	Metal-insulator-metal capacitor

NVM	Non-volatile memory
PVD	Physical vapor deposition
PMA	Post metal annealing
SiON	Silicon oxynitride
SDR	Spin dependent recombination
SdT	Spin dependent tunneling
SILC	Stress-induced leakage current
TAT	Trap-assisted tunneling
TEM	Transmission electron microscopy
TMO	Transition metal oxide
TE	Top electrode
V_{FB}	Flat band voltage
VM	Volatile memory

List of Symbols

κ or ϵ	dielectric constant
μ	micro
k	Boltzmann's constant
q	electronic charge
β	Weibull slope
$\text{\AA}$	Angstrom
φ	wave function
V_O	Oxygen vacancy
O_i	Interstitial oxygen
V_O-O_i	Oxygen vacancy-interstitial pair
H^+	Positively charged hydrogen ion
N_{it}	Number of interface states
D_{it}	Interface state density
E_R	Reaction energy
C_{ox}	Oxide thickness
E_f	Formation energy
E_F	Fermi energy
V_{th}/V_{TH}	Threshold voltage
E_{xc}	Exchange-correlation energy
ρ	Electron density

V_{set}	Set voltage
V_{reset}	Reset voltage
V_{gs}	Gate stress voltage
I_{cpr}	Charge pumping recovery current
t_{r}	Recovery time
t_{s}	Stress time

Chapter 1 INTRODUCTION

1.1 Background

1.1.1 Transistor Scaling and Gate Dielectrics Evolution

It has been more than 50 years since Kahng and Attala adopted the Si as the substrate and the SiO₂ as the gate dielectric to fabricate the metal-oxide-semiconductor field effect transistor (MOSFET) [1, 2]. With optimized fabrication process, a high quality interface between the Si-SiO₂ can be formed with relatively low interface state density ($\sim 10^{10}$ cm⁻²/eV) [3], which guarantees the field effect. In addition, due to the covalent property of the SiO₂, the fabricated bulk oxide layer also shows fabulous thermal and chemical stability. Excellent electrical properties, *e.g.*, low electron/hole bulk trapping centers and large breakdown (BD) electric field (>10 MV/cm) can be achieved [4, 5]. Therefore, the Si/SiO₂ MOSFET has grown to be the most fundamental and important electronic device in the micro-electronics field.

In the past decades, academia and industry have spent enormous resources in pursuing superior device performance, and the device size scaling is proved to be the most effective method in achieving density, speed and power improvement. Moreover, due to the rapid progress of the fabrication technologies, *e.g.*, lithographic technique, ion implantation and atomic layer deposition, *etc.*, the device density on the chip has followed an exponential increment, that is the number of transistors on

the integrated circuits doubles every 18-24 months with the shrinking of the transistor's critical feature size, which is known as Moore's Law [6]. However, in order to follow the shrinking scheme, the device feature size in particular the gate length of the MOSFET has been scaled from micron to nanometer range (14 nm from Intel). Moreover, according to logic CMOS road map published by the ITRS [7], the scaling will be continued towards 10 and 7 nm nodes in the next five years, whereas in order to meet this roadmap, the thickness of the gate dielectrics also has to be scaled accordingly to maintain the gate capacitance. For example, in the case of SiO₂ gate dielectric layer, when the channel length reaches sub-micron regime, the thickness of the SiO₂ has to be reduce to nanometer region. However, at this stage, the thin SiO₂ is no longer able to function as an insulator due to the unacceptably high direct tunneling current [8, 9]. Therefore, an alternative solution has to be found in order to keep the gate capacitance unchanged but reduce the direct tunneling current. In view of this dilemma, researchers trace back to the physical definition of the capacitance in *Eq. (1.1)*:

$$C = \frac{\kappa \epsilon_0 A}{d} \text{ --- Eq. (1.1)}$$

where d is the dielectric thickness, κ is the material dielectric constant, ϵ_0 is dielectric constant for the vacuum, and A is the capacitor area. It can be seen that the capacitance is determined by three key parameters, namely the dielectric thickness, the dielectric constant and the capacitor area. Meanwhile, it is also known that the direct tunneling rate strongly depends on the thickness of the insulator. Thus, in order to reduce the direct tunneling current while maintaining the capacitance, a thicker layer with higher dielectric constant material can be used to replace the pure SiO₂. Guided by this idea, among various dielectric materials, nitrided SiO₂ (SiON), which evolves from the SiO₂,

is first proven to be a suitable high dielectric constant material to replace SiO_2 [8, 9]. Experimental results show that the implementation of the SiON not only successfully in reducing the gate tunneling leakage, but also in suppressing the boron penetration from the poly-silicon gate [8, 9]. However, with the aggressive scaling scheme continuously moves on, the gate length of the MOSFET device comes into the sub-nanometer range. At this stage, the scaling of the thickness of the SiON reaches a critical point as it becomes less capable in functioning as a good gate insulator in suppressing the direct tunneling leakage [10, 11]. Hence, it is necessary to seek another alternative material with higher permittivity to replace the SiON. Considering all the essential requirements that a novel high permittivity gate stack material should have, hafnium (Hf) based oxide is found to exhibit superior performance compared to the other candidates, *e.g.*, the interface state ($D_{it} \sim 10^{11} - 10^{12}$ states/cm²), oxide trap density ($\sim 10^{11}$ cm⁻²), $I_{ds}-V_g$ characteristics and EOT achievement (< 1 nm) [10, 12-14]. Therefore it has been commonly accepted as the third generation of gate dielectric for the MOSFET.

1.1.2 Bias Temperature Reliability of the Advanced Gate Stack

The MOSFET relies on the gate capacitor to control the conductivity of the channel, as shown in the Fig. 1.1. For a p-MOSFET under working condition, usually a negative bias is applied to the gate electrode. When the applied bias voltage is higher than the MOSFET's threshold voltage ($|V_{bias}| > |V_{th}|$), inversion holes will accumulate near the channel region, thereby connecting the source and drain terminals together, and the device is "ON". However, experimental results show that under this situation, the inversion holes in the channel will penetrate into the Si/SiO₂ interface with the help of the electric field and high temperature. As a result, the hole trapped at the

Fig. 1.1: Typical MOSFET structure for (a) p-MOSFET and (b) n-MOSFET, respectively. When voltage applied to the gate electrode, inversion layer will be formed at the channel surface, connecting the source (S) and drain (D) terminals together.

interface or in the bulk region will consequently cause the undesired threshold voltage shift or even the circuit failure [15]. Since the above reliability issue is facilitated by the negative gate bias and the high temperature, therefore it is named as Negative Bias Temperature Instability (NBTI) on the p-MOSFET. Instead, for the n-MOSFET, inversion electrons may penetrate into the gate stack and cause the threshold voltage shift. Since the n-MOSFET is usually positively biased under the working condition, therefore the reliability issue on the n-MOSFET is named as Positive Bias Temperature Instability (PBTI).

The threshold voltage instability issue on SiO₂ MOSFET was first observed by Miura and Matukura in 1966 [15]. In the experiment, a 300 nm thick SiO₂ grown on a n-type silicon substrate MOS structure was used. Their observation shows a fast increase of the accumulated electron density (N_{FB}) at the Si/SiO₂ interface under the high temperature and gate bias stressing conditions, as shown in the Fig. 1.2. However, at that point of time, the major suspicion was the migration of the mobile ions under the stress field. An electrochemical generation of mobile ions was considered to

happen when the sample was stressed under the high temperature and stress field. But the BTI phenomenon didn't attract many researchers' interests at that early stage. It -

Fig. 1.2: The effect of the applied voltage on the saturated value of NFB after BT treatment with the metal electrode negatively biased [1].

is because the MOSFET had a buried channel that separates the inversion carrier away from the interface and thereby mitigating the carrier injection. Moreover, the threshold voltage was relatively high ($>5\text{V}$), therefore small variation of threshold voltage resulted from the BTI effect don't significantly affect the MOSFET's operation.

Whereas with the rapid development of the process technology, in the past decades, the aggressive device scaling has pushed the MOSFET feature size into the submicron region, which results in the change of the device structure from a buried- to surface-channel, as well as a thinner gate oxide with higher gate oxide electric field. These substantial changes cause the device more vulnerable to the BTI. In addition, the SiON is used to replace the SiO₂ in order to suppress the gate leakage and boron penetration. Many experimental results have shown that the presence of nitrogen in the gate oxide has substantially worsens the BTI performance [8, 9, 16-22]. Kimizuka *et*

al. are those pioneers that conducting the NBTI study on the SiON device [9, 23]. Their results first demonstrate the significant degradation of the threshold voltage in the SiON gated device as compared to that of the SiO₂ device under the same NBTI measurement. Besides, other researchers were also performed electrical characterization to study the NBTI on p-MOSFETs by employing the ultra-thin oxynitride (SiON) gate oxide [8, 16, 22, 24, 25]. The results show that the degradation of the SiON gated p-MOSFET is typically characterized by a much weaker dependence on stress time and temperature as compared to that of the SiO₂ counterpart. However the understanding on the effect of the nitrogen is not clear although the critical impact of nitrogen on the NBTI is realized, hence intensive study of BTI degradation mechanisms in the MOSFET revives. Particularly the NBTI on the SiON p-MOSFETs has received a considerable amount of attention from both academia and industry, but few attentions have been paid to the PBTI in the SiO₂/SiON MOSFET, because the PBTI is relatively negligible in the SiO₂/SiON device if compared to NBTI [26]. Furthermore, as the scaling continues, the gate stack evolves to the hafnium oxide. Due to the poor Si/HfO₂ interface quality as well as the relatively high intrinsic defect density in the HfO₂ bulk, plethora of observations show much more severe degradation of both NBTI and PBTI in the p-MOSFET and n-MOSFET, respectively [26-28]. Other than that, on the La doped HfO₂ n-MOSFET, the PBTI study reveals an unexpected slow recovery rate compared to the nominal PBTI recovery curve in the HfO₂ n-MOSFET [29], which is unable to be explained by the current PBTI model.

In a summary, the BTI remains a highly controversial reliability issue and more efforts are needed to shed light on the fundamental understanding of the BTI degradation mechanism.

1.1.3 High- κ Material for Emerging Memory Application

It is well known that hafnium (Hf) based oxide will replace the traditional SiO_2/SiON to suppress the gate leakage and save the device scaling, whereas similar dilemmas are also happening to the FLASH memory device. Previously, the FLASH memory has pronounced advantages in terms of the device integration density, power consumption, and Si-based fabrication process compatibility over the other kind of memory device [30, 31], therefore it has been widely studied during the past decades and attributed to the rapid progress of the microelectronic fabrication techniques, industry has pushed its feature size towards to sub-20 nm. However, at this stage, great challenges are encountered on both process integration (*e.g.*, lithography) and device reliability (*e.g.*, dielectric breakdown, short channel effect, cross-talking, direct tunneling and *etc.*) [30, 32-34], therefore it is necessary to develop a novel emerging memory to evade these issues.

In view of the above dilemma, a promising new memory should have the characteristics of scalability beyond 20 nm, compatible with the current CMOS process stream, low power consumption, acceptable retention time (>10 years), fast working speed and good reliability/endurance. Although various novel/improved NVM prototypes have been proposed or studied, for example, ferroelectric random-access memory (FeRAM), phase-change random-access memory (PCRAM), magnetoresistive random-access memory (MRAM) and HfO_2 based resistive random-access memory (RRAM), however, among all these devices, only the RRAM is able to meet the aforementioned criteria, that is because PCRAM suffers a high power consumption and cross-talking [35], whereas MRAM/FeRAM is not compatible with the standard

CMOS fabrication process [36-38]. Therefore, HfO₂ based RRAM is recognized as the most promising candidate to replace the FLASH memory.

1.2 Motivation of the Study

In the previous part, we have given a brief introduction of the BTI issue on the SiON and the hafnium metal gate MOSFET. Although in the past decades, plenty of observations and explanations were presented by the researchers, the BTI still remains a relatively obscure yet important reliability issue, especially when the device size reaches nanometer regime and the threshold voltage tolerance becomes more and more critical. Therefore more efforts are needed to shed light on the interpretation of the experimental results as well as the fundamental mechanism understanding.

Since in the MOSFET fabrication process, the hydrogen is inevitably used to repair the Si/SiO₂ interface. Therefore, in the early days, some of the scholars deduced that the H atom/ion will move out of the interface region under the BTI stressing, which subsequently generate the interface states and cause the threshold voltage shift [39], and based on this thinking, they depicted a physical degradation picture as follows:

The Si-H bond is broken by the injected hole, and then the H⁺ is driven away by the electric field, leaving the Si dangling bond near the interface as the interface states.

Since the above degradation process includes the reaction and diffusion steps, hence a model named *Reaction-Diffusion model (R-D model)* is proposed, and the degradation curve is deduced to follow a power-law time exponent tendency [39-43].

Later on as the gate dielectric is replaced by the SiON, much more severe threshold voltage degradation under the BTI measurement is observed. Through various experimental results, it is shown that the NBTI degradation magnitude in the

SiON device is around one orders higher than that of the Si/SiO₂ under the same stress condition. Moreover the degradation of the SiON gated p-MOSFET typically shows a much weaker dependence on the stress time and the temperature as compared to that of the SiO₂ counterpart [8, 9, 23]. As the early studies attribute the degradation of the threshold voltage to the dissociation of H near the Si/SiO₂ interface in the SiO₂ gated MOSFET, hence some of the researchers extended the *R-D model* to the Si/SiON interface, and claim that the presence of the nitrogen element weaken the interface structure and promote the interface states generation, causing an even worse degradation [40, 41, 43-45].

The above hydrogen related explanation of the NBTI in the SiO₂ and SiON device has dominated the NBTI mechanism study for a long time, until recently certain new evidences were revealed [19, 20, 22, 46, 47]. Through the improved measurement methodology, the stress induced hole trapping in the dielectric layer was proved to be another important component that causes the threshold voltage degradation [22]. In Ang *et al.* work, they present a phenomenological relationship between the energy distribution of the stress-induced hole traps and the power-law time exponent of the NBTI [25]. Their experimental results indicate that increased generation of deep-level hole traps (DLHTs) *i.e.*, trap-energy levels are near and/or above the Si conduction-band edge, results in a small exponent (< 0.2), and annealing the DLHTs will bring the exponent to ~ 0.3 . Measurement on the control sample of n-MOSFET (in which the effect of DLHTs is suppressed) shows an exponent of ~ 0.4 - 0.5 for sole interface-state generation. This implies that the relatively small exponent (~ 0.3) of the p-MOSFET is due to remnant DLHTs which charge-up positively again when the device subjected to negative gate biasing during measurement. This set of new evidences calls for the re-consideration of the notion that as-measured exponents of

~ 0.14 - 0.17 are experimental proof of H_2 -diffusion-driven interface-state generation. Moreover, dynamic NBTI (DNBTI) results from the SiON device indicate that the threshold voltage shift ($|\Delta V_t|$) comprises of a *Recoverable* and a relatively *Permanent* component [20, 46, 48, 49]. When dynamic NBTI (~ 30 Stress/Relax cycles) is carried out, the recoverable component shows a “non-Reaction and Diffusion” characteristic, which is found to remain constant during the cycles [48]. This observation implies that the amount of $|\Delta V_t|$ increase during stress is almost entirely recovered in the subsequent relaxation cycle, which conflict with the *R-D model*. Therefore, in view of the above two inferences as well as the results from the other publications [21, 50, 51], the NBTI degradation mechanism should not be limited to the hydrogen related interface state generation, whereas the hole trapping in the dielectric is also an important component. However, the unambiguous process of the hole trapping and de-trapping, as well as the origin of cyclic behavior is still under debate, thus a comprehensive understanding of the experimental results, as well as possible physical models are urgently desired for the BTI study.

In addition, when the hafnium oxides are implemented as the novel gate dielectric, the BTI faces new challenges. Generally, apart from an inherently high defect density in the bulk, the interface quality between the Si substrate and the HfO_2 layer is worse than that of the Si/SiO₂ or Si/SiON. Therefore, studies of the NBTI in the HfO_2 gated device show incredible high degradation. Other than that, measurement of the PBTI in the SiO₂ and SiON device shows almost negligible degradation if compared to that of the NBTI [26, 52]. However in the Hf oxide gated stacks, much more severe PBTI is reported which is the new concern of the n-MOSFET reliability [26, 53]. For this reason, it is essential to conduct more experiments to investigate both the NBTI and PBTI mechanisms in the HfO_2 gated MOSFET.

Besides, it is widely believe that the HfO_2 based RRAM possesses the potential to be the next generation NVM or universal memories. Preliminary results show excellent operating characteristics: 1) fast writing/erasing (W/E) speed (~ 10 ns); 2) a simple device architecture (1 Transistor + MIM Capacitor) can be integrated with the logic circuit; 3) low operating voltage (1-2 V), which leads to a low power consumption; 4) CMOS process compatibility cuts the fabrication cost. However, the implementation of RRAM is still facing various challenges. *e.g.*, 1) the switching mechanism in the HfO_2 has not been fully understood yet, 2) the discrepancies of the cell switching parameters variability from device-to-device, die-to-die and wafer-to-wafer are still quite unacceptable, 3) the device reliability, especially the device retention has not been fully studied [54, 55]. Therefore it will be helpful to extend the understanding of the HfO_2 in the gate stack to RRAM.

Finally, the development of the computing technology enables the investigation of the electronic structure (principally the ground state) of many-body systems, particularly atoms, molecules, and the condensed materials through atomic modeling and theoretical calculation. Since the BTI is caused by the material properties change under the bias condition, therefore the simulation work based on the density functional theory is an alternative way to understand the BTI degradation mechanism in the SiON and HfO_2 . In our simulation work, the VASP software package adopting the density functional theory is used [56]. It allows the researchers to build the defect models, and solve the Schrodinger equation within the scope of density functional theory which enables us to study the material properties change under the bias stressing. Hence the application of the atomic modeling and theoretical calculation in the interpretation of the defect characteristics is an advantage.

1.3 Objectives of the Study

The aim of this work is to study the electrical characteristics of the local defects in the SiON and hafnium oxides stacks through electrical measurement and atomic modeling, thereafter the understanding on the hafnium oxide defect properties is extended to the RRAM application. The switching mechanisms of the RRAM device with various electrodes are studied. Detailed objectives are listed as below:

- To understand the impact of the nitrogen related defect on the generation of deep-level hole traps under negative-bias temperature stressing. Experimental results and specified atomic defect model will be discussed.
- To obtain an understanding on the structure, property and dynamic of K_N center in SiON dielectric. Various defect forms with different nitrogen configurations in the SiON will be demonstrated and their structure dynamics as well as impacts on the hole trapping will be discussed.
- To study the evolution of switching oxide traps in the high- κ gate stack under NBTI and PBTI experiment. The DNBTI measurement will be implemented and the unique degradation behavior of NBTI and PBTI under different bias voltage will be studied through atomic modeling.
- To unveil the impact of the lanthanum doping on the device PBTI characteristics. The slow recovery of PBTI resulted from La doping will be examined through the first-principle simulation.
- To uncover the improved RRAM switching characteristics with different electrodes. Experimental characterization and first-principles simulation are performed and a novel model will be proposed to support the improved RRAM switching characteristic.

In all the experiment, SiON/Hafnium oxide gated device are used for NBTI and PBTI measurement. For the RRAM study, Metal-Insulator-Metal (MIM) stack is prepared, in which hafnium oxide is used as the insulator layer, and different types of metal materials are deposited as the electrodes. Our first-principles simulation work is performed by using VASP.

1.4 Organization of the Thesis

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

- *Chapter One* highlights the background, motivation, objectives and original contributions of the research work.
- *Chapter Two* presents the literature review of the SiON and hafnium oxide gate stack reliability issues in terms of NBTI and PBTI. In addition, the current experimental observations and the proposed switching mechanism of the hafnium oxide based RRAM device are summarized.
- *Chapter Three* presents the density functional theory, band gap correction scheme, and the charge transition level extraction method.
- *Chapter Four* discusses the atomic property of recoverable and permanent component in the BTI study. In the first part, the two-fold coordinated nitrogen on the generation of the deep-level hole traps under NBTI is elaborated. In the second part, a detailed discussion on the structure, property and dynamic of the K_N center in SiON dielectrics is included.
- *Chapter Five* presents a synthesized interpretation of the BTI results through the property and relaxation of the oxygen defects in the hafnium oxide gated device. The first part discusses the dynamics of the oxygen vacancies and the vacancy-interstitial defect pairs in the hafnium oxide under both NBTI and PBTI measurements. The second part

presents an insight of the impact of La doping on the PBTI through first-principles simulation.

- *Chapter Six* presents a study on the resistive switching memory of the hafnium oxide stack. The switching mechanism and the optimization of the RRAM stack are proposed.
- *Chapter Seven* gives the conclusion of this thesis and the recommendations for the future research works

1.5 Original Contributions

The major contributions of this research lie in the characteristics study of the SiON/high- κ stacks in MOSFET and RRAM device through electrical measurements and first- principles simulation. The atomic modeling and simulation combined with the electrical measurements give unambiguous physical pictures of the material property changes under the bias condition. The original contributions are listed as follows:-

- We show via first-principles simulation based on the density functional theory, the paramount role of the neighboring nitrogen atom on the structural relaxation of V_O following the capture of a hole in the SiON dielectric. Particularly, a two-fold coordinated nitrogen atom is shown to induce very significant structural relaxation of V_O . The resultant trapped-hole site has a very deep charge transition level in the SiO₂ band gap. On the other hand, in the absence of the nitrogen atom or if the neighboring nitrogen atom is three-fold coordinated (*e.g.*, the dangling nitrogen bond is terminated by a hydrogen atom), the resultant trapped-hole site would have a much shallower charge transition energy level.

- We show a comprehensive study of the properties of nitrogen involving oxygen vacancy defects by utilizing the first-principles simulation. The dynamic and relaxation characteristics of different nitrogen associated defects are studied (K_N centers), our results illustrate that nitrogen related defects can have various behaviors and consequently generate a wide range of defect energy levels throughout the band gap, which perfectly reproduce the experimental observations.
- We show that the transient charge trapping in the high- κ oxide defects can be transformed into more permanent trapped charge when the applied electrical cum thermal stress exceeds a certain threshold. In addition, it is found that when the transformation of hole trapping occurs under negative-bias temperature stress, a correlated increase of the gate current is always observed, which points to the transformation process being the origin for bulk oxide trap generation. However, when the transformation of electron trapping occurs under positive-bias temperature stress, an increase of the gate current is not always observed. From *ab-initio* simulation, we show that an intrinsic oxide defect - the oxygen vacancy-interstitial (V_O-O_i) - could consistently explain the above experimental observations, *e.g.*, generation of deep-level trapped charges as well as transformation of recoverable (or shallow) to permanent (or deep) charge trapping.
- We provide new insights of the impact of La doping on reliability issues in the HfO_2 gate stacks. Through atomic modeling, the electronic properties of the La associated oxygen vacancy (V_O) defect and the related vacancy-interstitial (V_O-O_i) paired defect are discussed. Our modeling work enables us to explain the PBTI improvement, the TDDB degradation, and the slow recovery of PBTI in the La-doped $HfSiO_x$ gate n-MOSFET.

- Our results elaborate the distinctive switching characteristics of the RRAM stacks with different electrodes. We show the critical role of the forming process in creating the first conductive filament in the fresh device. The Pt mediated oxygen vacancy formation during the forming step is modeled, which explains the excellent switching characteristics of the RRAM stacks when Pt is used as cathode.

Our research works have been published in established international conferences and peer reviewed journals, such as the *ECS Transactions (ECST)*, *Solid States Devices and Materials symposium (SSDM)*, *Microelectronic Reliability*, and *IEEE Electronic Device Meeting (IEDM)*.

Chapter 2 LITERATURE REVIEW

2.1 Introduction

This chapter is divided into four main sections. The first three sections will give a comprehensive review of the bias temperature reliability issues arising from the implementation of the SiON and high- κ film as gate oxide. In section 2.2, we give a general description of the bias temperature instability issue in both SiON and high- κ gated MOSFET. In section 2.3, a detailed review of the reliability issue in the SiON gate stack is elaborated. Literatures show that the SiON gate stacks suffer a much more severe NBTI degradation than that of the traditional SiO₂ device and the degradation typically shows a much weaker dependence on stress time and temperature as compared to that of the SiO₂ counterpart. However the understanding of the nitrogen associated degradation mechanism remains controversial. Therefore, in this section, we will present a systematic review of the published experimental results as well as the mechanism proposed. In the subsequent section 2.4, we will review the new reliability challenges in the hafnium oxide based gate stack. Unlike the conventional gate stack, the hafnium oxide shows worse interface quality compared to the Si/SiO₂ or SiO₂/SiON ones. In addition, the bulk hafnium oxide inherently contains more intrinsic defects that are able to capture the carriers. Early experimental results have already illustrated that hafnium oxide gated devices underwent both

significant NBTI and PBTI degradation for p-MOSFET and n-MOSFET, respectively, however the mechanism explanation of the NBTI and PBTI are quite controversial. For the PBTI degradation, even no clear model is available. Therefore, in this section, the electrical performances as well as the simulation modeling are summarized. Finally, in the Section 2.5, the extended application of the hafnium oxide gate stack in the RRAM will be addressed. In the first place, the limitation of the current FLASH memory and the necessity of developing a novel NVM will be presented. After that detailed discussion on the impact of the electrodes and the current understanding of the RRAM switching mechanism are included.

2.2 An Overview of the Bias Temperature Instability in the MOSFET

It is no doubt that bias temperature instability (BTI) is one of the most critical topics in the field of MOSFET reliability study. Revival of the concern of NBTI issue is largely due to Kimizuka *et al.* constructive work on SiON device [9, 23], in which they illustrate the predominance of NBTI over the channel hot-electron injection for sub-100 nm CMOS technologies. After that, various experimental results discussing the nitrogen related NBTI are shown in the literature, most of them ascribe the degradation to the generation of positive charged interface states, and the presence of the nitrogen promotes the generation rate [8, 9, 16, 40, 41]. However, in the recent publications, researchers propose some new insights on the NBTI [19, 20, 22]. In Ang *et al.* work, they present a phenomenological relationship between the energy distribution of stress-induced hole traps and the power-law time exponent of NBTI, in which the experimental results are shown to contradict with the traditional inference [25]. Moreover, results from dynamic NBTI study on the SiON device illustrate that a

constant part of the threshold voltage degradation is fully recoverable even with the cycles going up, which is also discrepant with the classical NBTI mechanism understanding [48, 57, 58]. Therefore, it is necessary to re-consider the nature of the defects and the trapping/de-trapping procedures. Furthermore, for the SiO₂ or SiON gated device, PBTI is rather insignificant compared to the NBTI, therefore fewer study has been carried out pertaining to the PBTI. But with the implementation of the high- κ material, both NBTI and PBTI on the hafnium oxides based device show significant degradation [26, 27]. And the high intrinsic defect densities as well as the poor interface quality are considered to be the main culprits. However the understanding of the exact degradation mechanism is still quite limited. Therefore, in the next several sections, we will present a detailed literature review pertaining to the BTI measurement as well as the mechanism discussion, to help us grip a picture of the current research status.

2.3 Reliability Issue of the SiON Gate Stack

2.3.1 The Early Understanding of NBTI Degradation

Before the implementation of the ultra-fast measurements technique, it has been widely discussed that the interface states generation at Si/SiO₂ or Si/SiON interface during NBTI stress was one of the dominant physical process that contributes to the threshold voltage degradation. This inference comes from the recognition of defective Si/SiO₂ or Si/SiON interface structure. Since from the material study, crystal Si possess a tetrahedral structure, in which the central Si atom bonded to the nearby four Si atoms. However, during the oxidation, due to the lattice mismatch between the Si and SiO₂, the bonding configuration at the surface will be changed, as shown in the

Fig 2.1, where most of the Si atoms are bonded to oxygen atoms while some of the remaining Si bonds are left dangling ((100) orientation). As a result, those unpassivated bonds will form the interface defect states, which are able to capture/release the carriers under the stressing condition, causing the threshold voltage shift [41, 44]. Therefore, those interface traps are apparently undesirable, and in order to eliminate them, a common way called hydrogen forming gas annealing during the device fabrication will be introduced to passivate the dangling bonds. In view of the presence of the H atoms near the interface region, some pioneers in BTI study proposed that under the NBTI stress, the Si-H bonds would be broken by the injected-

Figure 2.1: Si/SiO₂ interface structural model. The P_b center is labeled in the figure.

holes, and the newly released hydrogen species from the broken Si-H covalent bonds will diffuse away from the interface, leaving behind a positively charged interface state (Si⁺) which is responsible for the threshold voltage shift [43]. On the other hand, when the gate stress is removed during the recovery process, the dangling Si- bonds are considered to be re-passivated by the available hydrogen species which is near/at the interface instantaneously. While those remaining hydrogen which are far away from the interface will slowly diffuse back and recombine with the broken Si-H bonds at a slower rate. Since the above process involves the reaction, diffusion and re-forming of

the Si-H bond, therefore it is termed as the *Reaction- Diffusion model (R-D model)* [42, 44].

2.3.2 The Impact of Nitrogen Incorporation on NBTI Degradation

As the scaling scheme continues, the CMOS technology reaches the sub-100 nm nodes. Boron penetration and unacceptable high gate leakage current due to the direct tunneling turn out to be the critical concerns for the ultra-thin SiO₂ gated device reliability as well as the low power application. Awareness of the above mentioned dilemmas, nitrogen (N) was introduced into the gate oxide (*nitridation*) to increase the dielectric constant as well as suppress the boron penetration. This was proved to be an effective way to eliminate the aforementioned critical issues [8, 16, 23]. However, in view of the complex valence property of nitrogen element and increased lattice mismatch between the gate dielectric and the silicon substrate, the SiON dielectric layer contains more intrinsic defects (interface states and bulk traps) if compared to the

Fig. 2.2: (a) The V_{th} shift due to the NBTI for pure and nitrided oxide. A large V_{th} shift was observed for nitrided oxide. (b) The extrapolated device life time for different stress voltages [9].

traditional SiO₂ oxide, as a result, it is more sensitive to the bias temperature stress [9, 23]. There are many papers that talking about the NBTI degradation after nitridation. Kimizula *et al.* are those who first study the effect of nitrogen incorporation on NBTI degradation [9, 23]. By applying the conventional slow measurement, they addressed that nitridation degraded the V_{th} unambiguously compared to that of the pure SiO₂, as illustrated in the Fig. 2.2 (a). With the presence of nitrogen in the gate dielectrics, the magnitude of the V_{th} degradation of the low nitrogen concentration device is almost five times larger than that of the pure SiO₂ device. As a result, nitridation degrades the lifetime of the device by up to two decades, as shown in the stress dependent measurement in Fig. 2.2(b). Other than that, the temperature dependent study of the V_{th} shift on large scale device unveils much lower activation energy for the nitrated oxide as compared to that of the pure oxide, which could be another reason contributing to much more severe V_{th} shift even the SiON and SiO₂ devices are biased under the same stress condition. Moreover, in the leakage current study, experiment result shows that-

Fig. 2.3: (a) Initial gate leakage of the device, nitridation doesn't increase the initial leakage. (b) The increment of the gate leakage after the NBTI stress, large increment of the nitrated sample is observed [9].

after the NBTI stress, the leakage current through the nitride oxide around V_{FB} is unusually high, which implies that much more interface states are generated in the oxide during the stress, as shown in Fig. 2.3.

In view of the above experiment evidences, Kimizula proposes a possible degradation mechanism that hydrogen ion is released from hydrogen terminated silicon bond due to BTI stress and then they move away from the interface region, leaving the interface states. For the dissociated H^+ , they will be captured on either nitrogen or oxygen atom. However, experiment results indicate that the H^+ is more likely to be trapped on nitrogen due to the lower activation energy [59], and ultimately cause more severe degradation in the SiON.

At the same time, Mitanni *et al.* conducted the nitrogen related BTI degradation mechanism study through the Fluorine (F) doping experiment in the SiON

Fig. 2.4: NBT-stress-time dependence of the ΔV_{TH} (a) and temperature-dependence of ΔV_{TH} ; (b) for p-MOSFET having SiON/SiO₂ film. The oxide thickness is 2.3 nm. it is found that NBT degradation of SiON is more remarkable than that of SiO₂, and the activation energy of ΔV_{TH} for SiON ($E_a \sim 0.1$ eV) is lower than that for SiO₂ ($E_a \sim 0.2$ eV). these results are consistent with Kimizula N.'s work [8].

device [8, 16]. For the first step, they measured the degradation curves of SiO_2 and SiON under different temperatures. The activation energies were extracted, as shown in the Fig. 2.4. Similar results, as revealed in Kimizula's work, are obtained that nitrogen degrades the NBTI performance and the degradation activation energy of the SiON is much lower than that in the SiO_2 (see fig. 2.4). Secondly, the F is introduced in to the SiO_2 and SiON , respectively, and the impacts of F doping effects are studied. Their results in Fig. 2.5 show that the incorporation of F in the SiO_2 can improve the NBTI performance and the improvement is attributed to the replacement of the H element by the F element, which harden the Si/ SiO_2 interface. However, the incorporation of F element shows no obvious improvement in the SiON device and the degradation curves of the threshold voltage remains unchanged, as shown in the Fig. 2.5. This set of results imply that the defects response for the NBTI degradation in-

Fig. 2.5: NBT-stress-time dependence of ΔV_{TH} for p-MOSFET having SiON and SiO_2 film. BT stress was performed at 140 °C and under $E_{\text{stress}} = -7.5 \text{ MV/cm}$ [8].

SiON are hydrogen independent, and the doping of the F element in the SiON won't change the defects property that accounts for enhanced NBTI. Thereafter, in order to

understand more on the defects property and the degradation mechanism, the correlation between the threshold voltage shift and the interface state generation is investigated. The results reveal that both SiO_2 and SiON threshold voltage degradation have a tight relation to the interface state generation, as shown in Fig. 2.6. However, the $\Delta V_{th} - \Delta D_{it}$ correlation is independent of the F incorporation and temperature for both SiO_2 and SiON , which implies the defects in the SiON cause severe NBTI performance locate in the bulk. Finally, on the basis of the F incorporation experiment as well as the relation between the $\Delta V_{TH} - \Delta D_{it}$, the authors proposes the N related defect generation process, as shown in Fig. 2.7. For the case in SiO_2 , as indicated in-

Fig. 2.6: Correlation between interface-state generation (ΔD_{it}) and threshold voltage shift (ΔV_{TH}) in both cases of SiON and SiO_2 films. Note that $\Delta D_{it} - \Delta V_{TH}$ correlations of SiON and SiO_2 films are clearly split. Moreover, it is found that respective $\Delta D_{it} - \Delta V_{TH}$ correlation are independent of stress condition, temperature and fluorine dosage [8].

Fig. 2.7 (a), the Si-H breakage dominates the degradation process. However, in the SiON , a new defect form, namely Si-N-Si dimmer, as shown in Fig. 2.7 (b) may exist. Under the negative bias condition, the accumulated hole will immerse the interface region. At this point, Si-N-Si will capture one hole and subsequently cause the

Fig. 2.7: Schematic diagram for mechanism of NBTI degradation. It is inferred that two kinds of origins contribute to NBT degradation. One is the hydrogen-related process such as Si-H bond-breaking (a) the other is hydrogen-unrelated structure such as two-fold coordinated N contributes to NBT degradation (b) [8].

threshold voltage shift. This model proposed by Mitanni *et al.* is absolutely hydrogen independent, and thus the incorporation of the F won't change the NBTI performance of SiON.

Besides various electrical measurement performed on device to study the BTI degradation characteristics, in Fujieda *et al.* work, physical characterization using electron spin resonance (ERS) measurement was carried out to explore the NBTI process [60]. Since the interfacial dangling bonds (P_{b1} and P_{b0} center) are the direct index of the density of interface traps, therefore the change of P_{b1} and P_{b0} center signals are monitored during NBTI test by the ERS.

To conduct the characterization, two p^+ poly-Si/SiON/Si (100) kinds of diodes with H^{2-} or D^{2-} annealing were prepared, respectively. The devices were then subjected to the NBTIS at 150 °C under -2V bias first, and after that the samples were cooled down to room temperature to perform the C-V measurement. As shown in the Fig. 2.8,

Fig. 2.8: C-V shift (ΔV_{FB}) and increase in interface state density (ΔD_{it}) of H^{2-} and D^{2-} annealed plasma-nitrided SiON/Si diodes plotted against duration of NBTS [60].

Fig. 2.9: ESR signals (first derivative) of the P_b centers detected for (a) the NBTS-applied plasma-nitrided SiON/Si, (b) thermally depassivated plasma-nitrideed SiON/Si, and (c) thermally depassivated SiO₂/Si samples. The magnetic field was set parallel to the [100] axis [60].

ΔV_{FB} and ΔD_{it} for the D^{2-} annealed devices are 0.7-0.8 times lower than that of H^{2-} annealed devices. Based on this observation, the authors claim that the presence of D near the interface strongly affects the dissociation process. In other words, D^{2-}

annealing retards the generation of interface states. Thereafter, the ESP measurement is conducted on the prepared samples. From the ESR intensity spectrum, two distinctive P_b center signals, namely P_{b0} and P_{b1} , which are determined by the orientation of the dangling bond, are detected as indicated in the Fig. 2.9. By comparing the P_b density with the ΔD_{it} , the author found that they are quite similar, which implies that the dominated defects generated by the NBTS are P_{b1} centers. Moreover, their detailed analysis of the ESR spectrum shows that in the plasma nitride diode, the P_{b1}/P_{b0} density ratio (1.72) is much larger than that in the pure SiO_2 case (0.56). The effective-

Fig. 2.10: (a) optimized SDR of NBT-induced defects in 7.5 nm SiO_2 and 2.3 nm PNO-based devices as a function of g-value. NBTI generates different defects in SiO_2 and PNO devices. (b) SDT ($V_F = -0.2 \text{ V}$) and SDR ($V_F = +0.2 \text{ V}$) spectra of PNO defect both measurement clearly show the generation of a single dominating signal at $g=2.00020$. The spectrometer gain is 100 times greater for the $V_F = -0.2 \text{ V}$ spectrum [61].

g-value of the P_{b1} is 2.0028 in the plasma nitride SiON whereas in the pure SiO_2 , it is 2.0065, but for the g-value of P_{b0} center in the plasma nitride SiON is 2.0064, which is similar to the g-value of the P_{b0} center in the pure SiO_2 . The predominance of the P_{b1} center as well as the change of the g-value in the SiON is also found in the NO-annealed oxide, which is supposed to introduce the N element near the Si/SiON interface region. Finally, in view of the above evidences, the authors conclude that the

N in the plasma nitride SiON changes the interface atomic configuration and thereafter degrades the NBTI performance.

Similar experimental results are also shown in Yonamoto *et al.* work, in which they conclude that NBTI is due to an increase of the P_{b1} -center rather than the P_{b0} -center [17]. However, in Campbell *et al.* work, new morphor of defect, named K_N center, is found to be an important component that causes the threshold voltage degradation [61, 62]. As shown in the ultra-sensitivity SDR spectrum in the Fig 2.10(a). After NBTIS, on the SiO_2 device, two distinguished SDR signals corresponding to the P_{b0} and P_{b1} are observed, which are consistent with the other literatures [17]. However, for the SiON device, a much broader SDR signal at $g = 2.0020 \pm 0.0003$ is detected, which strongly indicates the existence of new type defect. Then, in the study of the impact of source/drain to substrate forward bias on the EDMR sensitivity to the SiON defect, the authors find similar SDR and SDT spectrum, as shown in Fig. 2.10(b). Both the SDT and SDR show the dominating signal at $g=2.0020$. Thereafter in order to obtain an

TABLE 2.1: measured magnetic-resonance parameters of important silicon dangling-bond defects. The K center and K_N center parameters are nearly the same [61].

	Centerline Width (G)	g-value*	Hyperfine Splitting (G)
P_{b0} [36]	$\cong 3$	2.0059	105
E' [57]	$\cong 2$	2.0005	424
K center [56]	$\cong 14$	2.0028	358
NBTI Defect (K_N center)	$\cong 14$	2.0020	349

*For the E' , K , and K_N centers this is simply the zero-crossing g-value; but for P_{b0} , it represents the g-value at a specific magnetic field orientation. In these measurements $H \parallel \langle 100 \rangle$ surface normal.

Fig. 2.11: K_N center as proposed by Campbell, J. P. *et al.* The number of nitrogen atom attached to the Si depends on the specific defect configuration.

overall picture of the defect types present in the dielectric, they compare their experimental results with some important silicon dangling-bond defects, as summarized in the Table 2.1. We can see that the g -value of the P_b and E' center are located at 2.0059 and 2.0005, respectively, whereas the g -value of the K center, which represents the silicon dangling-bond defects in the silicon nitride, is located at 2.0058. Interestingly, a similar g -value for the NBTI defect is detected at 2.0020. Besides, the centerline width (G) and hyperfine splitting (G) of K center and NBTI defect are also quite similar. Based on the above evidences, the authors conclude that the SiON defect is a silicon dangling bond with back-bonded to the nitrogen atoms, namely K_N center, as shown in Fig. 2.11, and the number of the back-bonded nitrogen atoms depends on the specific local atomic structure.

Furthermore, in view of the non-hydrogen related degradation evidence, intensive studies are performed by various researchers to investigate the physical degradation mechanism [8, 22, 61, 63]. In Ang *et al.* work, they show evidence that two distinctive degradation mechanism coexist in the SiON device under NBTI measurement [25]. In the experiment, poly-silicon gate p-MOSFET with dry oxinitride

SiON are used, and SiO₂ as well as SiN gate dielectrics device are prepared as the control samples. For the first step of the characterization, temperature dependent measurements are performed, and the Arrhenius plot of the V_t degradation is shown in the Fig. 2.12. As expected, the V_t shift of the SiON is larger than that of the pure SiO₂ device even under the same measurement condition. However, the extracted activation energy based on the measurement results under different temperatures, show some discrepancy (see Fig. 2.12). For the SiON device, the activation energy at low temperature regime is ~ 0.037 eV (E_{a1}), but it goes up to ~ 0.14 eV (E_{a2}) when the temperature increases. For the case of SiO₂ device, the activation energy remains at 0.27 eV within the entire temperature range. The above observation indicates that the degradation in the SiON device is temperature dependent, and two distinctive mechanisms with different temperature coefficients co-existing in the SiON device, which contribute to the threshold voltage degradation coincidentally. Other than that, another analysis by extrapolating the Arrhenius plot of the E_{a1} to the high temperature regime, and subtracting it from the measured curve at high temperature regime (dashed line) are performed. This mathematical operation helps them obtain an E_a that has the activation energy equal to 0.24 eV, which is within the conventional SiO₂ range. At this stage, they infer that besides the conventional degradation mechanism (noted as mechanism 2) as shown in the SiO₂, a new mechanism should exist in the SiON, which will be noted as mechanism 1 here. With reference to Fig. 2.12, it can be observed that mechanism 1 contributes to the degradation within 75 °C -125 °C, and mechanism 2 only shows up above 125 °C. Although the above investigation strongly implies the existence of nitrogen related degradation--mechanism 1, however, it can't roles out the impact of tunnel current which is temperature independent because during the measurement, the SiON device shows higher tunneling leakage. Regarding

of this issue, the authors perform comparison study of the SiON and SiN device. SiN devices possess $\sim 10\times$ lower direct tunneling leakage than the SiON devices are-

Fig. 2.12: Threshold voltage shift $|\Delta V_{th}|$ of device A, stressed at -2.6 V gate bias for 5×10^4 s, as a function of $(\kappa T)^{-1}$ (open circles), showing the presence of two distinct degradation mechanism (I and II). Also shown for comparison is the Arrhenius plot of a 12-nm SiO₂ gate p-MOSFET (L/W=0.6/0.4 $\mu\text{m}/\mu\text{m}$), stressed at comparable oxide field. Degradation due to mechanism I (activation energy $\sim 0.037 \text{ eV}$), is not observable in this case. Eliminating the contribution of mechanism I in Device A yields the Arrhenius plot (filled circles) for mechanism II, whose activation energy (0.24 eV) agrees very well with that of the SiO₂ gate device. Inset: schematic energy-band diagram of the p-MOSFET under negative gate bias, depicting transfer of electronic charge at deep defect states introduced by nitrogen to holes in the inversion layer [25].

and then similar measurements are performed as what they do for the SiO₂ and SiON devices in the first part, after that the activation energies for SiON/SiN are extracted respectively. The obtained results are shown

Fig. 2.13: Arrhenius plots for NBT stress-induced threshold voltage shift $|\Delta V_{th}|$ of device A and B. The circles denote the respective Arrhenius plots of devices A and B after eliminating the contribution of mechanism I [25].

in Fig. 2.13. Surprisingly, the SiN devices suffer much higher degradation, especially at low temperature range than that of the SiON device. The extracted E_a of the SiN device at low temperature is ~ 0.019 eV, even lower than that of SiON. This implies that the direct tunneling current should not be the culprit that causes the higher degradation in the SiON, since in the SiN device which has lower tunneling current, whereas higher degradation is observed. In other words, the nitrogen related degradation mechanism is not due to direct tunneling current. Moreover, if conducting the same operation that extrapolating and subtracting the component of mechanism 1, the obtain curve which is contributed by mechanism 2 is parallel to the one obtained in the SiON case, and the E_a is almost identical. This confirms that the nitrogen mechanisms in both SiON and SiN are similar although the gate material and fabricate process are different. And the much more significant degradation observed in the SiN device can be subscribed to the higher concentration of nitrogen associated defects that related to the mechanism 1.

To summarize, the NBTI characteristics and mechanism in the SiO₂/SiON are reviewed. In the early days, the hydrogen related degradation model dominated the NBTI field, however, with more and more new non-hydrogen related evidences revealed through physical characterization as well as novel electrical measurement, intensive debate on the NBTI mechanism revives.

2.3.2 First-principles Simulation for BTI study in the MOSFET

As addressed in the previous section, early days' studies on the NBTI degradation mechanism attribute the dissociation of the hydrogen species on the N/Si atom as one of the dominate reason that causes the severe degradation in the nitrated oxide. Based on this preliminary understanding, some of the researchers performed first-principles simulation to explore this electrochemical process. Tan *et al.* worked on the modeling and simulation of the hydrogen related defects in the SiON material, and gave a detailed description of the degradation process [59, 64]. On the basis of the *R-D model*, they first investigated the Si/SiO₂ and Si/SiO_xN_y interface structures, as shown in the Fig. 2.14. The mobile hydrogen is considered to be trapped on the oxygen/nitrogen atom, respectively, and the reaction energies for the hole capturing are calculated. As we can see that, for both SiO₂ and SiON, the reaction energies are negative, indicating that the H⁺ trapping process is exothermic ($E_R < 0$) and the electrochemical reaction is energy favorable. However, for the case in the pure SiO₂, $E_R = -8.51$ eV, but when the trapping happens on the nitrogen, the reaction energy decreases to -9.54 eV, which implies that the H⁺ trapping near the Si/SiON interface is more ready to occur, and consequently leads to the worse NBTI performance. In addition, in Yang *et al.* work, they extended their models to different nitrogen concentration with three different hydrogen trapping positions, namely-

oxygen vacancy, oxygen and nitrogen [65]. For the case that H^+ trapping on the nitrogen, four kinds of SiON models with different numbers of nearby nitrogen atoms are studied, as shown in the Fig. 2.15. Initially the H attaches to Si atom, and forms the Si-H bond. Under the negative bias condition, injected hole from the inversion layer causes the breaking of the Si-H bond, which is recognized as the reaction step. And then the H^+ moves towards the nitrogen site, forming the N- H^+ cluster. The activation energies of the electrochemical process are calculated for each model and labeled in-

(a) NBTI at defective interface (vacancy defect as reaction site)		$\Delta E_{R,NBTI}$
		-0.02 eV
		-0.40 eV
		-0.67 eV
		-1.00 eV
(b) NBTI at defect-free interface (O as reaction site)		
		0.0 eV
		-0.27 eV
		-0.45 eV
		-0.61 eV

Fig. 2.16: Relative reaction energies ΔE_R of NBTI reaction at oxide interface with different No. of N as 1st-nearest neighbor for hydrogen-originated hole trapping nitrogen centers. Namely, (a) vacancy defects and (b) oxygen [65].

the figure. As we can see that, with the increase of the nitrogen concentration, the reaction energy decreases. Based on the above simulation result, the authors then provide a possible explanation that due to the lone pair on the nitrogen, it can act as an effective donor, and hence the defect cluster associated with nitrogen is more effective to act as the hole capturing center. Besides, similar tendencies are found when the H^+ trapping happens in vacancy site or oxygen, as summarized in the Fig. 2.16, and the same explanations are applied to these models. Furthermore, the authors also compare the reaction energies difference when the H^+ is trapped on the different sites (oxygen

vacancy, oxygen and nitrogen) under the same nitrogen concentration condition, as shown in the Fig. 2.17. Results indicate that the lowest reaction energy is obtained when the H^+ is trapped on the nitrogen, which implies that nitrogen is a more effective hydrogen-trapping center as compared to the oxygen vacancy as well as the oxygen, and this is attributed to lone pair on nitrogen that can act as a more effective donor to attract the H^+ .

Fig. 2.17: Relative reaction energies ΔE_R vs. the No. of N atoms as 1st-nearest neighbor for three different types of hydrogen trapping centers, namely vacancy defect, oxygen and nitrogen [65].

However, either in Tan *et al.* or Yang *et al.* work, they haven't considered the reaction barrier for the Si-H dissociation near the Si-SiO₂ interface, various experiment and simulation work show that the Si-H bond is quite strong (>2 eV) [66, 67]. Under the NBTI stress field as well as the hole inversion condition, the Si-H is unlikely to be broken, therefore their studies are not complete. In view of this, Tsetseris and Pantelides *et al.* proposes an alternative hydrogen source- from dopant sites in the depletion region of the Si substrate, they also identified the dissociation of

Si-H bonds by protons as a reaction that can account for the observed generation of interface traps [68, 69].

Other than the simulation work that focuses on the *R-D model*, Lee *et al.* conducted the first-principles simulation in the SiON gate stack with the nitrogen interstitial (N_i) related defects [70]. Several defect models are generated depending on the local atomic configuration of the N_i as shown in the Fig.2.18. In the Fig. 2.18(a), the interstitial nitrogen (N_i) locates near the interface region, and connects to three Si atoms. A full structure relation is conducted to obtain the stable atomic configuration. As we can see, the stable defect structure includes a trivalent N and Si (Si_I) atom, whereas Si_{II} and Si_{III} bond with other three Si atoms besides the N atom. Since the-

Fig. 2.18: Atomic structures of the (a) N_i , (b) N_i-O_i , (c) N_i-O_{2i} , and (d) $N_i-O_i-O_r$ defects at the interface [70].

generated Si_I is under coordinate, therefore it is considered as a chemically active center, however the valance shell of the nitrogen is full, and thus it won't be able to capture the holes. Besides, partially oxidized is energetically favorable, therefore in

Fig. 2.18(b)(c)(d), one or two oxygen atoms are introduced near the N_i to form the Si-O-Si structure. After the model preparation, the defect levels are calculated and listed

TABLE 2.2: Single-particle defect levels (E_{KS}), defect transition levels ($E_{-/0}$) between neutral and negatively charged states, and Si-O distance for Ni-related defects. E_c indicates the conduction band minimum of Si [70].

Defect	E_{KS} (eV)	$E_{-/0}$ (eV)	Distance (Å)	
			Si_I-O_I	Si_I-O_{II}
N_i	$E_c-0.4$	$E_c-0.31$		
N_i-O_i	$E_c-0.29$	$E_c-0.20$	3.36	
$N_i-O_i-O_r$	E_c	E_c	1.85	3.33
N_i-O_{2i}	$E_c-0.19$	$E_c-0.12$	3.32	3.47

in the TABLE 2.2. As we can see, all the N_i related defects generated defect levels are located in the up part of the Si band gap. For the cases of $N_i-O_i-O_r$ and N_i-O_{2i} , due to enhancement of the interaction between the N and nearby O, the defect level are pushed towards the Si conduction band. Moreover, the analysis of the atomic structure implies that the shift of the defect level doesn't come from the direct interaction between the N_i and O, but from the interaction with three-fold Si atom. The above simulation work provides an alternatively view of the device degradation that independent of the hydrogen.

To summarize, early first-principle calculations has modeled various possible reaction and diffusion equation. The simulation results show that the reaction energy decreases with the increase of the number of the nitrogen atom in the model. However, in view of the relative strong Si-H bond, researcher proposed a new hydrogen source, which located in the substrate. Meanwhile, from the study of the different trapping points, namely oxygen vacancy, oxygen and nitrogen, it shows that nitrogen is a more

effective hole capturing center. However, recent experimental evidences show the non-hydrogen characteristic of the NBTI. Hence the simulation work on the N_i reveals another possible mechanism that causes the significant degradation.

2.3.3 The Oxide Switching Traps Model for the NBTI in the SiON Gate Dielectrics

In the previous parts, we reviewed the NBTI characteristics in the SiON devices, and different degradation models were presented. In view of the V_{th} degradation, commonly it is classified in to two parts [71-76]:

Part I: the fast transient part that results from the rapid charge and discharge of the hole trapping and de-trapping;

Part II: the slow transient stage that is ascribed to the dissociation/repassivation of the interface states by the hydrogen species.

However, more and more evidences in the experiment show that the slow recovery part doesn't fully origin from the dissociation/passivation of the interface states, the bulk defects also contribute to part of it by building up the deep level traps [74, 75]. In Aichinger *et al.* work, an ultra-fast temperature change technique is used to study the NBTI recovery mechanism. After stressing, the dedicated devices are recovered under different conditions, *e.g.*, temperature cycling or bias polarity change. The set of results obtained show that recovery of the degradation is both bias and temperature dependent, but don't match elastic tunneling or interface states passivation mechanism. And the unique recovery behavior under temperature or bias change leads the author concludes the inelastic phonon assisted tunneling [77]. Besides, as shown in Ang *et al.* work, they find that the measured ΔV_t through static I_V is unbelievably

larger than the value calculated from the ΔN_{it} , which implies that ΔN_{it} alone is not the solely component that accounts for the threshold voltage degradation, although the accuracy of the charge pumping measured N_{it} is questionable [24]. Since in the ref. [78], it addresses that a rapid repassivation of ΔN_{it} may happen once the stress is removed, and the recovery can be accelerated by a reverse polarity bias, therefore, in Ang's study, a bipolar measurement scheme is implemented to investigate the evolution of ΔN_{it} . To be clear, if the rapid repassivation of the ΔN_{it} is existing during the measurement and ΔN_{it} is the only component that responses for the V_{th} degradation, the experiment results of the bipolar cases will show that both the ΔN_{it} and ΔV_{th} of the bipolar stress shall be lower than that of the static stress cases. However, the actual measurement results don't show like the above inference, as we can see in the Fig. 2.19 inset. For the static case, the V_{th} is larger than that of the bipolar measurement case, whereas an almost similar ΔN_{it} is obtained for both. Moreover, from the Fig. 2.19, we can find that the ΔV_{th} versus ΔN_{it} of the bipolar one is closer to the calculated one. This observation strongly implies that the significant reduction of the ΔV_{th} in the bipolar experiment is due to the de-trapping of the deep level hole traps which are unable to be detected by the charge pumping measurement, and hence give a clear indication that the slow transient V_{th} degradation comprises of two components, ΔN_{it} , as well as the deeply trapped holes. Besides, in one of the other measurements carried by the same group, the deep level hole trapping phenomenon is consistently observed [25]. Through the DC current-voltage measurement, they compare the stress induced evolution of positive trap (ΔN_{pt}) and interface state (ΔN_{it}), the results indicate that ΔN_{pt} is much larger than that the ΔN_{ita} (~5 times), which can't be explained by the usual thinking that hole trapping levels are located at the upper portion of the Si band gap

and they are uniformly distributed. The only possible inference is that NBTI induced a broad distribution of the defect levels, and most of them are near the Si valance band-

Fig. 2.19: Threshold-voltage shift of the p-MOSFET $|\Delta V_t|$ versus the stress induced interface-state density ΔN_{it} extracted from CP current (I_{cp}) measurement. I_{cp} was measured using a 1-MHz +1/-1-V gate-voltage pulse, with a rise/fall time of 100 ns. The dashed line denotes $|\Delta V_t|$ ($=q\Delta N_{it}/C_{ox}$; C_{ox} is the gate oxide capacitance) calculated based on the extracted ΔN_{it} . Static stress was carried out with a gate voltage of -2.6 V, corresponding to an oxide field of ~ 8.8 MV/cm. Bipolar stress was achieved using a 100-kHz 50% duty cycle gate pulse, with a 50-ns rise/fall time. The positive (negative) level was set at +1.5 V (-2.6 V). Stress temperature was 125 °C. The inset compares the ΔN_{it} and $|\Delta V_t|$ of the static and bipolar stresses at equivalent stress time T_{eq} ($=50\%$ of total clock time in the case of bipolar stress) [25].

(relatively deep) which are not be able to be detected by the charge pumping measurement. Moreover, in the recent publication, Teo *et al.* subject the device to dynamic NBTI study [48]. His impressive results elaborate a non-Reaction-Diffusion picture. Since for the *R-D model*, it is built on the hypothesis that:

- (a) Fast generation (recovery) ascribe to spontaneous depassivation (repassivation) of

Si-H (Si-) bonds by hole (hydrogen) at the interface;

- (b) Slow generation (recovery) rate-limited by hydrogen diffusion away from (towards) the interface.

Therefore, in the charge pumping experiment, the amount of the ΔI_{cp} recovered after each cycle would decrease due to the hydrogen diffusing away from the interface region. In addition, the diffusing length is supposed to be time dependent, and with the

Fig. 2.20: Amount of recovery per cycle, δI_{cpr} is observed experimentally (open circle) to remain unchanged after many stress/recovery cycles. The same observation holds over a wide range of V_{gs} . (not shown). On the other hand, the *R-D model* predicts a steady decrease of δI_{cpr} by $\sim 50\%$ after 30 stress/recovery cycles (filled circle); results obtained via simulation using the same stress conditions and device geometry). In the R-D context, δI_{cpr} decreases because hydrogen is transported further away from the Si/SiO₂ interface after numerous stress/recovery cycle. (b) δI_{cpr} increases with V_{gs} , indicating that it is stress induced [48].

change of t_s/t_r , the ΔI_{cp} should be changed. However, in Teo's experiments, both the abovementioned phenomenon don't show, whereas constant I_{cp} is observed during the

cyclic measurement, as shown in the Fig. 2.20. Moreover, different t_r/t_s combinations are implemented in the Fig. 2.21, results illustrate that there is no ΔI_{cp} changed even-

Fig. 2.21: δI_{cpr} (circle) is observed to be independent of t_s/t_r . This result does not corroborate a transport-limited mechanism of N_{it} generation and recovery as stipulated by the *R-D model* [48].

Fig. 2.22: (a) Dynamic NBTI after subtracting away the hole-trapping component; (b) the evolution of $|\Delta V_{t,r}|$ during DNBTI; (c) simulated evolution of $|\Delta V_t|$, using a diffusion framework for molecular hydrogen; (d) comparison of $|\Delta V_{t,r}|$ and $|\Delta V_{t,s}|$ [48].

after 30 S/R cycles. These set of experiment results strongly indicates that the transport-limited mechanism of N_{it} generation and recovery which stipulated by the *R-D model* is not true. In the subsequent experiment, Teo studies the amount of recovery of the threshold voltage shift ΔV_{th} under the dynamic NBTI experiment. As illustrated in Fig. 2.22, $\Delta V_{t,s}$ is defined as the increase of $|\Delta V_t|$ in the prior stress period, $\Delta V_{t,r}$ is the recoverable component of the total ΔV_{th} , and $\Delta V_{t,Eor}$ represents the relatively permanent component of the total ΔV_{th} degradation (the residual after 1000 s recover at 0 V), DNBTI results within 30 cycles show that there is no change of the recoverable component - $\Delta V_{t,r}(x)$ (see fig. 2.22(b)), in other words, the $\Delta V_{t,r}(x)$ shows a cyclic behavior in the DNBTI measurement, and the above experiment results further expose the limitation of the *R-D model*. Worth to mention, similar work are also performed by Kaczer, and Grass *et al.* [20], in which the ubiquity of the threshold voltage relaxation is demonstrated through the measure-stress-measure method, and the property of recoverable component under various condition are discussed that contribute to the physical modeling.

2.4 Reliability Study of the Hafnium Oxide Gate Stack

2.4.1 Brief Introduction of the Hafnium Oxide as the Gate Stack

The implement of the SiO_2 as the gate dielectric layer has been a great success since the advent of the MOSFET era. To be the dielectric layer of MOSFET, Si/SiO_2 can form a nearly defect free interface under proper fabrication process, which guarantees the performance of the MOSFET. However in the past decades, the CMOS fabrication technology has followed the path of device scaling to achieve density, speed and power improvements, and the gate dielectrics thickness has been scaled

accordingly to obtain sufficient gate capacitance. But pure SiO₂ is a relatively low dielectric constant material, in order to obtain enough gate capacitance to preserve the drain current, the thickness of the SiO₂ now almost reaches their physical limitation (~1.3 nm), beyond which would lead to an unacceptably high tunneling current. Therefore, it is necessary to find a way to reduce the tunneling current yet keep the scaling scheme continuing. Considering the physical definition of the capacitor and the equivalent oxide thickness (EOT) as shown in the Eq. (2.1) (2.2),

$$C = \frac{\kappa \cdot S}{d} \text{ --- Eq.(2.1)}$$

$$EOT = d_{H,\kappa} \cdot \kappa_{SiO_2} / \kappa_{H,\kappa} \text{ --- Eq.(2.2)}$$

Where d is the thickness of the dielectric material, S is the area of the capacitor, and κ is the dielectric permittivity of dielectric material, respectively. The capacitance is determined by these three key parameters. Moreover, as we also know, the direct tunneling current is strongly thickness dependent, therefore to obtain the same capacitance whereas can reduce the high tunneling current, one effective way is to use a thicker but higher dielectric constant material to replace the SiO₂/SiON, by which the capacitance can be preserved meanwhile the tunneling current can be suppressed. However in order to find a suitable high- κ material to replace the SiO₂/SiON as the novel gate dielectrics, certain requirements have to be satisfied [79, 80].

- Thermodynamically stable when the high- κ material deposited on the silicon substrate;
- CMOS process compatible during the fabrication;
- Large enough band offset(>1eV for both conduction and valence band) with the Si to prevent the electron/hole tunneling;

- The high- κ material is able to form a high quality interface structure with the Si substrate.

In view of the above criterions, the selections of the materials are limited to HfO_2 , ZrO_2 , Al_2O_3 , Y_2O_3 and La_2O_3 . But for the ZrO_2 , the thermodynamic data is not accurate enough that hampers the process integration during the fabrication. Moreover ZrO_2 is slightly reactive with the silicon, which is not desired. Whereas for the Y_2O_3 and Al_2O_3 , as we can see in the Fig. 2.23, their dielectric constants are around 13 and 10, respectively. Both of them are less than 20, which is not high enough to support more than one generations of the gate dielectric in the future road map. Therefore, the promising candidates now are restricted to HfO_2 and La_2O_3 . For the gate dielectric application, usually the amorphous form of material is preferred, unfortunately HfO_2 and La_2O_3 possess relatively low crystallization temperature, which pushes the thermal budget of the fabrication process to a much more restrict region. Meanwhile, current microelectronics fabrication process tends to be able to deposit high quality of HfO_2 -

Fig. 2.23: Variation of dielectric constant with band gap of high- κ oxides [79].

gate dielectric whereas for La_2O_3 may has higher intrinsic defect density, hence most of the high- κ gate stack studies focuses on HfO_2 , and various optimized MOSFETs by implementing the HfO_2 as gate dielectrics layer with superior EOT as well as low leakage are demonstrated.

2.4.2 Poor Interface and Bulk Quality of the Hafnium Oxide Gate Stack

The success of SiO_2 as the gate dielectric relies on its low interface state while grown on the Si substrate ($D_{it} \sim 10^{10}$ states/ cm^2) [3], however it will be a big challenge

Fig. 2.24: C - V characteristic of Al_2O_3 , HfO_2 and SiO_2 NGETs. Significant hysteresis is observed for the HfO_2 dielectric layer [81].

to achieve the comparable high quality interface for the Si/ HfO_2 structure, since there is no common element between Si and HfO_2 . During the formation of the transition layer, silicon atom has to bond with either oxygen atom or hafnium atom, which will induce significant lattice -mismatch. Early days of studies show that the interface states of the Si/ HfO_2 structure can reach $\sim 10^{11}$ - 10^{12} states/ cm^2 , which will cause severe near-interface charge trapping, as shown in the Fig. 2.24 [81]. Besides, Takeuchi *et al.*

carried the physical characterization of the HfO_2 on Si by using the spectroscopic ellipsometry [14]. In the experiment, strong absorption peaks within the range of 4.25 eV to 5.75 eV is detected, as shown in the Fig. 2.25. The detected three unusual peaks are considered to be promoted by the transition of the electron from the valance band-

Fig. 2.25: $(\alpha hv n)^{1/2}$ vs $h\nu$ plot of the samples prepared at different temperature [14].

Fig. 2.26: Dependence of the absorption coefficient and annealing time and temperature in oxygen ambient. Absorption coefficient of the extra peak was integrated over photon energy for the quantification, as shown in the inset [14].

to the defect levels, and the discrepancy of the defect energy levels are ascribed to the process variation during the deposition. Therefore, to further clarify the nature of the

defect, the authors prepare sets of samples with different annealing temperatures and time. By subjecting those samples to spectroscopic ellipsometry measurement, they find that the higher oxidation temperature and longer oxidation time tends to weaken defect-related absorption peaks, as illustrated in the Fig. 2.26. Since generally higher temperature and longer time annealing will make the oxidation more completely, thus the observed absorption peaks can be attributed to the high density of oxygen vacancy defects in the HfO_2 . In addition, Kang *et al.* [82] propose that the defect near the Si/HfO_2 interface is P_b -like silicon dangling bond, however they emphasize that Si/HfO_2 P_b -like dangling bond and that of the Si/SiO_2 P_b dangling bond is not identical due to the fact that g tensor value is largely different. Whereas at the same time, Lenahan *et al.* carried ESR/EDMR measurement on the HfO_2 gate dielectric [83]. Through the experiment results, they conclude that unique peak in the ESR plot is most likely due to the back-bonding silicon dangling bond as shown in Fig. 2.27-

Fig. 2.27: Schematic illustration of the back-bonded HfO_2 silicon interface dangling bond defect [83].

2.4.3 First-principles Simulation of the Defects in the Hafnium Oxide Material

Other than the experimental and physical evidences of the poor quality, theoretical physicists perform first-principles simulation to explore the material

properties of HfO_2 in an atomic view. Robertson and Xiong *et al.* are the first who conducted the first-principles modeling by taking into account both interstitial oxygen and oxygen vacancy defects, as shown in Fig. 2.28 [79, 84]. Defect energy levels are calculated (shown in Fig. 2.29) by using a screened exchange scheme which accurately reproduces the experiment band gap. It is shown that interstitial oxygen atom can be both positively and negatively charged, and gives the levels above the oxide valence band. However for oxygen vacancy, the energy levels lie above the Si-

Fig. 2.28: (a) Oxygen vacancy; (b) oxygen interstitial defect in the amorphous HfO_2 .

Fig. 2.29: Summary of the defect energy levels of oxygen vacancy and interstitial O under various charge states in the HfO_2 [79].

conduction edge, particularly when charged. These sets of results are quite consistent with experiment result [81, 85, 86], and compared to the case in the SiO_2 , the oxygen vacancy in the HfO_2 will tend to be positively charged, since it can transfer electrons into the Si.

At the same time, Foster *et al.* also carried the density-functional theory calculation to study the structure and electronic properties of oxygen vacancy and interstitial defects in monoclinic hafnium oxide [87]. They studied the stable configuration of neutral oxygen vacancies, interstitial oxygen atoms and molecular oxygen in the bulk HfO_2 . Both singly and doubly positively/negatively charged states were investigated. Their results show that atomic oxygen incorporation is energetically favored over molecular incorporation, and the charged defect species are more stable than the neutral species when electrons are available from the hafnium conduction band. The calculated defect levels with respect to the bottom of the silicon conduction band indicate that interstitial oxygen atom, molecules and positively charged oxygen vacancies can trap electrons injected from silicon substrate.

In summary, defects in Hf oxide material retain more complicated morphors, and their behaviors under positive/negative varies strongly depend on local atomic configurations of the particular defects. Among all of the defect types, oxygen vacancy is considered as the dominant one in the HfO_2 , together with the interface state (Si dangling bond to Hf) and interstitial oxygen, degrade the dielectrics property aggressively under the bias condition.

2.4.4 Bias Temperature Instability in the Hafnium Oxide Gate Stack

The low quality of the interface structure, as well as the inherent high defect density in the bulk make the bias temperature instability in the hafnium oxide gated MOSFET as the most critical reliability issue.

In Onishi *et al.* work, they first report the drain current degradation and the $G_{m,max}$ shift in the sub-15-Å EOT p-MOSFET device with HfO₂ gate stack, as shown in Fig. 2.30 [88]. The extracted threshold voltage indicates that after 1000 sec bias -

Fig. 2.30: (a) p-MOSFET I_d - V_g characteristics before and after NBTI. V_t shift and slight G_m degradation are observed. (b) V_t and G_m as functions of time during NBTI for without SN. G_m degradation is less significant than V_t degradation [88].

Fig. 2.31: Dependence of NBTI-induced threshold voltage shift on stressing conditions for HfO₂ p-FETs with tungsten gates. Symbols indicate measured data and solid lines are fits using model (7). (a) dependence of ΔV_{th} versus stressing time curves at $125^\circ C$; (b) dependence of ΔV_{th} versus stressing time curves on stress temperature at $V_s = -1.5$ V [52].

stress at 125 °C, significant increase of the threshold voltage is observed. Meanwhile in Zafar *et al.* work, they prepare p-MOSFET with HfO₂ gate dielectric stacks for the NBTI study [52]. By applying the stressing and measuring scheme, the NBTI under various temperature as well as bias voltages are extracted and shown in the Fig. 2.31. We can see that the absolute value of the threshold voltage continuously increases with the stress time going up, and the degradation of the threshold voltage shows significant stress time and temperature dependence. Under higher temperature or higher stress -

Fig. 2.32: Dependence of NBTI induced V_t shift for various pFETs with FUSI NiSi gate: (a) SiO₂/HfO₂/NiSi, (b) SiO₂/NiSi, and (c) SiON(~5% N)/poly-Si; symbols are measured and solid line are model fits. (d) The extrapolated life time for various devices under different voltage [52].

Fig. 2.33: PBTI for NiSi gated nFETs with SiO₂, SiO₂/HfSiO & SiO₂/HfO₂ as gate dielectrics; symbols are measurements and solid lines from model fitting. PBTI becomes worse with increasing Hf content in the dielectric layer [52].

voltage, the threshold voltage shows much more severe degradation, which is quite comparable to the NBTI characteristic in the Si/SiO₂ device. Similar experimental results are obtained in a more detailed comparable study of the NBTI in Si/HfO₂, Si/SiO₂ and Si/SiON devices, as shown in the Fig. 2.32. Moreover, a comprehensive study of the PBTI in the high- κ gate device is performed by the same scholar. SiO₂, HfO₂, and HfSiO n-MOSFETs are used as the experiment samples. The measurement is performed under various stress voltages, and last for several decades. The threshold voltage shifts are extracted accordingly, as shown in Fig. 2.33. Unlike the situation in the SiO₂ and SiON, the PBTI degradation is almost negligible, whereas on the Hf oxide device, PBTI measurement show severe V_{th} degradation, and the degradation increase with the increase of Hf content, in other words, the V_{th} shift in the pure HfO₂ is higher than that in the HfSiO case (shown in the Fig. 2.33). Meanwhile, compared to the NBTI degradation, ten years life prediction shows that the PBTI becomes the main challenge for device reliability, as illustrated in the Fig. 2.34.

Fig. 2.34: Comparison between PBTI (open symbols) and NBTI (filled symbol) for SiO₂/HfO₂/NiSi FETs. (a) ΔV_t versus stress time at constant stress field; (b) Predicted ΔV_t at 10 years of stressing as a function stress field; the predicted values are estimated from model fitting. PBTI is worse than NBTI in HfO₂/NiSi FETs [52].

In the recent BTI study, Gao *et al.* conducted the PBTI studies of the high- κ device through the dynamic BTI (DNBTI) measurement, the V_t shift of the recoverable as well as the permanent component under repeated stressing and relaxation are examined [27]. In the experiment, n-MOSFET with 3 nm ALD HfO₂ and 1nm SiO_x device was used. To conduct the DNBTI, fresh devices were subjected to the stressing and recovery cycles up to 30 cycles, and each stressing and recovery phase was set to 1000 s, the detailed wave form is shown in the Fig. 2.35. During the experiment,

Fig. 2.35: the wave form used for DPBTI. Each stressing and recovery phase was set to 1000s, and totally 30 cycles are carried.

Fig. 2.36: the evolution of R as a function of the number of stress/relax cycles under various oxide stress fields at a given temperature of 100 °C. R is constant at 5.5 MV/cm but shows a progressive decrease at higher oxide stress fields [27].

different stress fields were implemented, and the evolution of the recoverable and the permanent component were continuously examined. As shown in the Fig. 2.36, the evolutions of the Recoverable component (R component) under different stress conditions show quite different behaviours. Under relatively low stress field, the R component keeps constant within 30 cycles of stress and recovery, however when the stress field increase to 7 MV/cm, the recoverable component slowly decrease with the cycles goes, and moreover, when the e-field increase to 8.5 MV/cm, a more severe decrease of the recoverable component is observed. Therefore, in order to expose the origin of the reduction of the recoverable component, the evolution of the total degradation right after stress phase (ΔV_t^{eos}) and the residual degradation after 1000 s of recovery (ΔV_t^{eor}) are monitored. As illustrated in the Fig. 2.37, when the stress is low (5.5 MV/cm), the increment of the V_t^{eos} is comparable to the V_t^{eor} . However, when the

Fig. 2.37: The increase in ΔV_t^{eos} , with respect to the ΔV_t^{eos} of the first stress/relax cycle (denoted as $\delta \Delta V_t^{\text{eos}}$), are comparable for the three cases but the increase in ΔV_t^{eor} , with respect to the ΔV_t^{eor} of the first stress/relax cycle (denoted as $\delta \Delta V_t^{\text{eor}}$) is much larger for the 7 and 8.5 MV/cm oxide stress field as compared to that of the 5 MV/cm oxide stress field. This indicates that the decrease in R at higher oxide stress fields is due to a part of R being transformed into more permanent ΔV_t [27].

e- field increase to 7 MV/cm or 8.5 MV/cm, a much faster increment of the ΔV_t^{cor} shown compared to the ΔV_t^{cos} , this result indicates the decrease in recoverable component under high e-field is ascribed to the conversion of the recoverable component to the relatively permanent component. Moreover, a temperature accelerated study performed by Pobegen *et al.*, shows the similar results. They demonstrate the impact of the temperature-acceleration on both degradation and recovery, making it possible to perform long-term NBTI test in a fraction of the time usually required.

Furthermore, other researchers studied the PBTI relaxation under AC and DC stress on the high- κ metal gate p-MOSFET [89]. For the AC stress, $T_s = N \times C \times T$, where N is the total number of the stress cycles, C is the duty cycle and T is the period.

Fig. 2.38: (a) Relaxation traces measured after AC and DC stress with stress time changing from 1 sec to 10000 sec. notice that the AC relaxations always start with lower V_t shift and shallower slope, but they gradually merges with the corresponding DC relaxation at longer relaxation times; (b) Comparison of relaxation traces after DC stress and AC stress with different duty cycles. All the relaxations merge into each other after longer relaxation times [89].

Whereas for the DC measurement, a fresh device is stressed for T_s . In both cases, the drain current I_d is monitored at sense bias set approximately equal to the initial V_t . At

the end of recovery, the V_t shift will be extracted from the I_d degradation. As illustrated in the Fig. 2.38(a), the relaxation curve for both DC and AC stress under different stress time are plotted. We can see that for the first ΔV_t point, under AC stress condition, it is always smaller than that under the DC stress condition. However, after a long relaxation phase, the degradation curve for AC and DC gradually merges with each other. Similar results are shown in the Fig. 2.38(b) that under different duty cycles, the degradation of the AC stress is smaller than the DC one although with the increase of the duty cycle. However, at the end of stress, they merge with each other. This set of experiment shows that the trapping and de-trapping process in the HfO_2 are strongly correlated, and the electron trap depth is only tied to the stress time. Whereas, in the above discussion, some intuitive questions about the AC BTI recovery curve and the DC BTI recovery curve are not addressed:

- 1) AC recovery rate is less than DC recovery rate;
- 2) AC and DC recovery traces with the same net stress time as the DC “merge After a certain recovery time;
- 3) At the merging point, the initially small AC recovery rate increase gain.

Therefore, in Reisinger *et al.* research work, they perform a more comprehensive and analytic AC BTI study, and present a model for AC NBTI which based on the capture and emission of charges in and out of oxide border traps [77, 90]. As charging and discharging of the defects in the oxide are considered as the main reason that contribute to the BTI, and moreover the defects are characterized by their –field- and temperature dependent- time constants for charge capturing and emission τ_c and τ_e . Thus, the authors create a defect filling map to illustrate the defect states under AC -

Fig. 2.39: (a) Occupancy level of defects after AC stress with frequency 1Hz, duty factor 99% (ON/OFF ratio = 100) for a stress time $t_{AC} = 100$ ks; the status is taken after a full charging cycle with a $1 \mu s$ measuring (=recovery) delay. The diagonal line marks $\tau_c = \tau_e$, the shifted diagonal is shifted by a factor 100 corresponding to the ON/OFF ratio. The white cross marks the AC period-time; the green dashed rectangle marks the charging and discharging periods of the AC signal. The dotted green trajectories mark lines of const. charging times $\tau = 1$ s and 100 ks. Right hand labels denote the filling state corresponding to the color code. The black horizontal and vertical dash/dotted lines denote the net stress time and the recovery time (= measuring delay), respectively. (b) Recovery curves all taken after a net stress time of 100 s with various duty factors. The arrows mark the position where the AC-curves and the DC-curve (100%) are expected to merge [99].

stress, to help understand the defect state under different AC stress condition, as shown in Fig. 2.39[a]. Based on this map, we can see that the filling status of the respective defect can be interfered with the stressing/recovery time elapsing. Moreover, for the BTI recovery behavior of both AC and DC stress condition as discussed in the ref. [89], we can induce that there is no short term recovery after AC-stress because this short term recovery has already been anticipated during the stress phase prior to

the real recovery phase. Hence, in Fig. 2.39[b], we can expected that after long enough recovery period, when the recovery is not sensitive to the beginning recovery time difference, the AC BTI degradation curve tends to merge with the DC ones.

2.4.5 The Impact of La Doping on the Reliability Issue in the HfO₂ Gated Device

Hafnium oxide material has been intensively studied as the gate dielectrics beyond 45-nm node, various experimental results indicate both the chemical and the physical characteristics of the hafnium oxide material meet the principle requirements as the next generation gate stack very well, however the successful implementation of the hafnium oxide based gate dielectric still faces some challenges:

- 1) Dopant penetration from the poly-silicon gate;
- 2) Fermi level pinning ;
- 3) Poor reliability.

Although metal gate such as TiN has been tried to be the gate electrode, the controllability of the threshold voltage as well as the pursuing of the low threshold voltage are still the deep concerns for the hafnium oxide gated MOSFET.

Early research work done by Kita *et al.* shows that an intrinsic dipole layer at HfO₂ and SiO₂ interface are formed due to the oxygen density difference, and consequently mitigating the threshold voltage tuning [91]. Upon realizing this physical phenomenon, researcher has tried many ways to eliminate the impact of this intrinsic dipole, and a simple rare earth doping in the dielectrics layer is proved to be an effective way to compensate it [92, 93]. Among various rare earth elements, lanthanum (La) has illustrated superior performance [92, 94], *e.g.*, good EOT scaling, carrier mobility and BTI reliability, hence it has been widely used to tune the threshold voltage. To introduce the La into the hafnium oxide layer, usually a thin La₂O₃ would

be deposited on the top of the dielectrics layer, then subject the wafer to annealing process. Physical characterization results has shown that the annealing process help the La diffuse in to $\text{SiO}_2/\text{HfO}_2$ interface region, as illustrated in the Fig. 2.40. After the S/D activation annealing, EELS results show that the La deposited on the high- κ layer piles up at the interface between the SiO_x and high- κ layer. Electrical measurements-

Fig. 2.40: Final profile of the La atom after a high-temperature annealing. Results showed that the La deposited on the high- κ layer piled up at the interface region [94].

Fig. 2.41: Both the La-doped HfO_2 gate stack and the control HfO_2 had the same final physical thickness ($\sim 2.5\text{nm}$); EOTs were also similar ($\sim 0.95\text{ nm}$), as determined from CVC simulation. The CV result shows the leftwards shift of the flat-voltage [94].

are carried on the La doped MOSFET, the results show that with the presence of La near the interface region, the flat voltage is shift leftwards $\sim 300\text{mV}$, which implies the presence of reverse dipole after the doping of La near the interface region (Fig. 2.41). Meanwhile, due to the left shift of the flat voltage, the drain current is significantly-

Fig. 2.42: I_d - V_g and V_{th} uniformity for samples with and without La; (b) effective electron mobility of La doped HfO_2 MOSFET. The EOT of the samples are 1.5 nm. We can observe an obvious effective mobility improvement with the La doping [92, 94].

Fig. 2.43: Gate-leakage characteristics for gate injection and substrate injection. The La-doped devices show a higher V_{bd} for both accumulation and inversion regimes [94].

increased if comparing to the case without La doped, and the improvement in threshold voltage uniformity and effective electron mobility are also observed, as

shown in the Fig. 2.42 [92, 93]. Moreover measurement of the gate leakage on the La doped n-MOSFET show significant reduction of the leakage when bias negatively, as illustrated in the Fig. 2.43, which further confirming the case without La doped, and the improvement in threshold voltage uniformity and effective electron mobility are also observed, as shown in the Fig. 2.42 [92, 93]. Moreover measurement of the gate leakage on the La doped n-MOSFET show significant reduction of the leakage when-

Fig. 2.44: Time evolution of ΔV_{th} in PBTI. As an effect of La incorporation, ΔV_{th} were suppressed in the same over-drive stress [93].

Fig. 2.45: ΔV_{th} in PBTI as a function of the extent of conduction electron injection for (a) HfSiON and (b) HfLaSiON. ΔV_{th} depends on stress bias. This indicates that trap sites were increased with stress [93].

bias negatively, as illustrated in the Fig. 2.43, which further confirming the presence of the dipole induced by the La that raises the electron injection barrier. Although the merits of the La doping are satisfactory, however the reliability study of the La doped n-MOSFET show some unique behaviors. Sato et al. first study the PBTI on La doped n-MOSFET, their early results show that the PBTI was largely suppressed at relatively low stress voltage due to the La doping, as illustrated in the Fig. 2.44 [93, 95]. However, in the subsequent voltage dependent measurement, the experimental results in the Fig. 2.45, show that degradation of the threshold are strongly stress voltage dependent. At relative lower bias voltage (1.6 V), the PBTI of the La doped n-MOSFET is much lower than the device without La doping, however when the bias goes up to 2 V, the degradation rate of the V_{th} with La is much faster than the one without La doping, and the authors ascribe this PBTI behavior to the increase of the defect site under high bias. Besides, simulation work on the La doping is done by Umezawa *et al.* as well [96]. Their computational results indicate that the formation of the oxygen vacancy of La doped HfO_2 is effectively suppressed due to the increase of the formation energy, and accordingly is used to explain the improvement of the PBTI performance of the La doped device at low electric field.

2.5 The Implementation of Hafnium Oxide as the RRAM Stack

2.5.1 The Limitation of the FLASH Memory

Depending on whether the memory cell needs the power supply to retain the data, there are two kinds of memories, namely volatile memory (VM) and non-volatile

memory (NVM). For the VM, once the power supply is switched off, the data stored in the memory will be lost, whereas for the NVM, even when the power supply is -

Fig. 2.46: Typical floating gate FLASH memory cell. By controlling the number of the trapped electron in the floating gate, the threshold voltage of the cell will be modified.

switched off, the data stored in the memory will be retained for future reading. Currently, the most widely used NVM is flash memory with the floating gate structure, a typical flash memory is shown in the Fig. 2.46. By changing the polarity of the control gate bias, the electrons are trapped or de-trapped in the floating gate, and consequently change the threshold voltage of the device to denoting the memory state of “1” or “0”. In the past decades, the industries have invested enormous resources to pursue the memory density improvement, and device scaling is considered to be the most effective way, which has been continuously applied since the invention of flash memory. However, the aggressive scaling brings out numbers of challenges [7, 30, 31]:

- (1) Scaling limitation: currently the minimum feature size has reached 19nm, to fabricate such a small device, the lithography will be one of the biggest issues;
- (2) Fewer electron storing: this will cause severe issue in terms of the data retention;
- (3) High capacitive coupling of the control gate with floating gate and coupling of adjacent cells: the coupling of the parasitic capacitor will hinder the further

increase in the density;

- (4) Breakdown between neighboring word lines: since the write/erase of the data in the flash device need relatively high voltage, the breakdown between the word lines will cause serious device reliability issues;

In view of the above issues, researchers still have no ideal solutions to solve any of them. Besides, due to multiple gate oxide-floating gates, the fabrication process for flash memory are quite different from the one for logic devices, therefore integrating the flash memory with the logic circuit as the embedded memory is still a big challenge in the state-of-the-art fabrication technology.

2.5.2 An Overview of RRAM

Revival of RRAM study

Resistor, capacitor and inductor are three fundamental passive elements for the electrical circuits, however in 1971, Leon *et al.* proposed the possible existence of the fourth fundamental element namely memristor through theoretical deduction, but the concept of memristor hasn't been presented in a physical model. Until in the year of 2007, Stanley *et al.* demonstrated that the Pt/TiO₂/Pt MIM structure can have memristor characteristic though electronic and ionic coupling under external bias [97], a typical I - V curve are shown in the Fig. 2.47. By applying different polarities of bias on the top Pt electrode while the bottom electrode is always grounded, the TiO₂ conductivity can be changed between two resistive states, namely High Resistive State (HRS) and Low Resistive State (LRS), and act as a memristor. Although similar switching phenomenon is observed by Hickmott in the Al/Al₂O₃/Al structure (MIM) in the 1962 [98], and he also proposed the possible application of this kind of stack as the memory, however until the demonstration of the memorisitor in ref. [97], the dopant or

Fig. 2.47: the experimental i - v curves for the Pt-TiO_{2-x}-Pt device under voltage sweep [97].

impurity motion dominated resistive change in the device resistance revived as the promising candidates for novel NVM.

RRAM structure and performance check lists

The extremely simple MIM structure, as well as the potential for high speed, low power, and high integration density application, makes the RRAM the most promising candidate for the next generation novel NVM or even a universal memory. A typical MIM structure of RRAM is shown in Fig. 2.48(a). The top and bottom electrodes usually are metals. However they are not necessary to be the same material, strongly depending on the design. For the insulator-switching layer, various materials have been tried, *e.g.*, perovskite, transition metal oxides and even SiO₂. Experimental results show that with proper electrodes, all of them are able to show the hysteretic i - v curve as illustrated in the Fig. 2.48(b), either in a bipolar or unipolar mode. Usually the RRAM has two distinctive states, namely HRS -

Fig. 2.48 (a) Typical RRAM cell with MIM structure. Usually use the transition metal oxide as the insulator layer, but the top/bottom electrodes can be various conductive materials. (b) Typical switching curves for bipolar mode (circle) and unipolar mode (triangle).

which represents the “OFF” and LRS which represents the “ON”. The resistive change between the HRS and LRS is called the resistive switching (RS) phenomenon. As shown in the Fig. 2.48(b), the resistive change from LRS to HRS is called “RESET”. On the contrary, the switching from HRS to LRS is called “SET”. Only when high enough voltage or current is applied to the electrodes will change the resistance states, and the voltage or current values that leading to the change of the resistive are named SET Voltage/Current (V_{set}/I_{set}) or Reset Voltage (V_{reset}/I_{reset}), respectively. However, when the applied voltage/current is much lower than the V_{set}/I_{set} or V_{reset}/I_{reset} , the resistance of memory cell can retain its original state, and represents the logic value “0” or “1”.

To set benchmarks for the performance of the RRAM, several important parameters have to be considered, including operating voltage, resistance ratio, endurance, retention time, device yield:

(1) Operating voltage/current:

Both research and industries are pursuing low operating voltage/current for the low power memory application. Since if the operating voltage is high, power consumption will increase and reliability will be a challenge, moreover with lower voltage/current, the integration density can be increased without heat dissipation problem. Therefore, comparing to the FLASH memory, the operating voltages for RRAM devices is requested to be only a few volts.

(2) ON/OFF ratio:

The ON/OFF Ratio is a substantial parameter that demonstrates the difference between memory states. Usually for RRAM, it is determined by the resistance ratio of HRS to LRS at 200 mV. ON/OFF ratio directly influences the accuracy of data reading. In terms of periphery circuit design, a minimum ON/OFF ratio of 10 is required [38].

(3) Programing speed:

Programing speed or in other words, read/write speed is defined as the minimum time for writing or reading a RRAM cell. For the classical FLASH memory, the programing speed usually in the range of micro-second, however, in the latest publication, Lee *et al.* demonstrated the fastest switching RRAM cell on the TiN/TiO_x/HfO_x/TiN stack [34]. By applying the pulse measurement, 5 ns of the operating speed was obtained, whereas for the typical application of the RRAM, the write time <100 ns is expected.

(4) Endurance:

Endurance depicts the maximum number of RS cycles that an RRAM cell can reach. Since the RS is based on the switching layer characteristic change under the e-field, but the switching layer will wear out ultimately that no more RS is available, thus good endurance is a significant requirement for the commercial application of RS memory devices. Usually FLASH memory has demonstrated a maximum number of write cycles between 10^3 and 10^7 cycles, based on the device design. Therefore, RRAM

should possess an endurance of greater than 10^7 cycles, or even better to replace the DRAM in the future.

(5) Retention time:

Retention time describes the length of time a RRAM cell can keep its programmed state. Since most of the memory working environment will be around 85°C , thus the retention test should be done at similar temperature, and to be quality nonvolatile available memories, it is expected to retain the value for at least 10 years.

(6) Device uniformity:

Device uniformity indicates the similarity of the device parameters for the integrated memory cells. Due to the process fluctuation, the forming voltage, V_{set} , V_{rest} and I_{max} may have a statistical variation, however to help the design of the periphery control circuit, it is necessary to restrict the above parameters within certain range, especially when the integration density goes up, the control of these parameters become more challenge.

(7) Device yield:

Device yield talks about the percentage of the devices those are within the design specs on the fabricated chip. Since the RRAM switching is based on the change of the characteristic of the insulator layer, the fluctuation of the process may lead to failure of the device. Therefore, certain methods have been developed to improve the yield. By doping a third element or suitable electrode material in the insulator layer are proved to be efficient.

To summarize, RRAM processes extremely simple structure, however in order to meet the NVM application requirements, it should be able to satisfy the above criteria.

2.5.3 Current Understanding of RRAM Switching Mechanism

In the material study, a large number of materials are known to have the conductivity switching property under the e-field, and various physical models are proposed to explain the switching mechanism as shown in the Fig. 2.49 [99]. However, for the RRAM application, early days studies indicate that electronic and ionic coupling under external voltage contributes to the switching, an unambiguous picture -

Fig. 2.49: Classification of the resistive switching effects which are considered for non-volatile memory applications. The switching mechanisms based on thermal, chemical, and electronic/electrostatic effects are concisely summarized [99].

of the fundamental principles of the switching mechanism is still not available. Currently, based on the electrode and insulator layer materials used in the RRAM cell, the switching mechanisms can be categorized into several important types.

(1) Electrochemical Metallization Mechanism:

In the Electrochemical metallization system (ECM), usually an electrochemically reactive metal, such as Cu, Ni, or Ag, is used as the anode and an inert metal (Pt/W) is

taken as the cathode. During the measurement, an electrochemical metal filament formation and dissolution happens under the inference of the external e-field in the dielectric layer, and consequently cause the resistive switching. Thus in some literatures, it is also named as Solid-electrolyte cell (SE) or Conductive Bridging cell (CB). Electrochemical metallization is related to cation-migration induced redox reaction. A detailed switching example is shown in the Fig. 2.50. As we can see, the RRAM is fabricated in to a Ag/Ge_{0.3}Se_{0.7}/Pt structure, during the measurement, the -

Fig. 2.50: Typical current-voltage characteristic of Ag/Ag-Ge-Se/Pt electrochemical metallization cell using a triangular voltage sweep. The “ON” conductance is limited by a compliance current of 25 μ A. The curve in the compliance current state is shown as a dashed line because the actual voltage does not correspond to the voltage shown on the axis. The insets A to D show the different stages of the switching procedure [99].

bias is always applied to the top electrode-Ag whereas the bottom electrode - Pt is grounded. For the first step, the electroforming process, a positive voltage is applied to the Ag electrode, since Ag is an electrochemically reactive element, an anodic dissolution of the Ag will happen and generate the positive charged Ag⁺ cation. These Ag⁺ then will be driven to the cathode side by the e-field, and accumulated towards the

anodes, as indicated in the Fig. 2.50(a). At the end of this process, the metal filament will ultimately contact the cathode and anode together, and turns “ON” the cell. Since the metal filament is relatively stable, so to reset the cell, an adequate voltage of opposite polarity is needed to dissolve the metal atoms from the edge of metal filament and eventually annihilates the filament (Fig. 2.50(c)), which finally reset the cell to “OFF” state.

(2) Valence Change System:

Valence Change Mechanism (VCM) refers to the transportation of the anions in the dielectric layer that causes the change of the cation valence, and subsequently led to considerable resistive change. The main difference between the VCM and ECM is that

Fig. 2.51: Schematic switching process of the valence change RRAM stack. (a) pristine RRAM stack, a few pre-existing defects randomly locate in the HfO_x layer; (b) In the forming step, a large number of oxygen vacancies is generated, creating the conductive filament (c) the reset step that back drive the mobile oxygen ion to the vacancies, and eliminate the conductive filament.

for the VCM, the electrodes won't inject metal cations in to the insulator layer that form the filament, whereas it is the intrinsic anions movement. Usually the valence

change phenomenon happens in the transition metal oxide (TMO), because in the TMO, the oxygen related defect-oxygen vacancy/oxygen interstitial are much more mobile than the transition metal cations. An enhancement or depletion of the defect under the inference of the e-field will affect the valence state of the transition metal oxides and cause the resistive switching. A typical schematic switching process is depicted in the Fig. 2.51. As we can see that in the pristine device, a one-time forming step is needed to generate the filament, as illustrated in the Fig. 2.51(b), when a positive bias is applied to the top electrode, the oxygen atoms in the oxides are supposed to be pulled out the lattice and driven to the anode, whereas the remaining accumulated in the dielectric layer, and ultimately form a metallic conductive filament. To reset the device, a reverse polarity of the bias is needed to drive the oxygen specials back to the lattice position, and oxide the metal filament again to break the conductive filament, (Fig. 2.51(c)).

2.5.4 The Effect of Capping Layer on the HfO_2 Based RRAM

Although the RRAM switching characteristics is successfully demoed, however for the massive integration, the forming voltage as well as the reset/set

Fig. 2.52: XTEM micrographs of the TiN/AlCu/HfO_x/TiN device. (a) Before PMA. (b) After PMA [34].

Fig. 2.53: Bi-stable switching characteristic of TiN/AlCu/HfO_x/TiN MIMC device with a cell size of 65×65 μm². (a) Typical I–V curve of bipolar resistance switch. The inset shows the multilevel operation with different clamped I_{SET}. (b) Uniformity of operation voltages (20 devices) [34].

voltage of the device still suffers the poor uniformity. To improve the performance, Lee *et al.* prepared the TiN/AlCu/HfO_x/TiN RRAM stack, in which AlCu was purposely deposited as the capping layer, a XTEM figure is shown in the Fig. 2.52 [34]. Result shows that after the PMA, the thickness of the AlCu layer increases from 6 nm to 8 nm, which is ascribed to the absorption of the oxygen from the HfO₂ layer and forming the sub-oxide, causing the volume expansion. Meanwhile, due to the loss of the oxygen, the HfO₂ becomes oxygen deficient, and makes the forming step more controllable. A typical switching curve with the cumulative probability is shown in the Fig. 2.53 [34]. Stable bipolar switching under different current compliances is observed, and the set/reset voltage shows good uniformity. Besides, similar improvement in their TiN/Ti/HfO_x/TiN RRAM is observed, in which a thin Ti layer is inserted as the capping layer (see fig. 2.54). Other than the Ti and AlCu insertion are used as the capping layer, Yu *et al.* and Govoreanu *et al.* demonstrated the improved uniformity by using the thin Al/Hf layer as the capping layer [100, 101]. Compared to the case without capping layer (TiN/HfO_x/TiN), their results show significant

improvements on the uniformity, such as the forming voltage, SET/RESET voltage. They ascribed the improved performance to the diffusion of Al into the HfO₂ that reducing the oxygen vacancy formation energy, and subsequently stabilizing the conduction filament. Besides, similar effect is observed by doping the Al, Ge or Gd into HfO₂ layer. A similar explanation that the presence of dopants in the HfO₂ eases the formation of the oxygen vacancy, and meanwhile reduces the oxygen ion migration barrier is addressed.

Fig. 2.54: (a) Comparison of the bipolar resistance switching characteristics between the TiN/TiO_x/HfO_x/TiN devices in 1R and 1T1R configuration. (b) Tight distribution of R_{low} and R_{high} in TiN/TiO_x/HfO_x/TiN device with 1T-1R configuration [34].

2.5.5 The Impact of the Electrode Material on the HfO₂ Based RRAM Switching

Various experimental results indicate that electrodes strongly affect the switching characteristic of the RRAM device, in this sub-chapter, we will review the impact of electrodes on RRAM performance. Chen *et al.* first demonstrated the co-existing of the unipolar and bipolar resistive switching in the TaTiN/HfO_x/Pt RRAM stack [102]. For the stack fabrication, 100 nm of Pt was deposited on the Ti/SiO₂/Si substrates as the bottom electrode, and then 10 nm HfO₂ was deposited by PVD in the O₂ ambient,

RTN at 450 °C in the N₂ atmosphere for 30 s was added after the deposition, and after that a 50 nm TaTiN top electrode was deposited by PVD, finally the memory cells were patterned. A schematic structure is shown in the Fig. 2.55. To activate the -

Fig. 2.55: (a) schematic device structure of the HfO_x based RRAM devices. The HfO_x resistive switching layer and the TaTiN top electrode (TE) are deposited by reactive PVD, respectively. (b) Process flow of the device fabrication [102].

Fig. 2.56: Typically I - V characteristics of forming process measured using current sweep and the first reset process after forming [102].

switching, a forming step is needed, as shown in the Fig. 2.56. First, a positive voltage sweep is applied to the TE whereas the BE is grounded. When the voltage reached 3.5 V, the fresh device suffers a soft breakdown, and consequently switches the device to

LRS. A subsequent negative voltage sweep can reset the resistance to the HRS, as indicated in the Fig. 2.57. Stable SET/RESET between the LRS and HRS can be observed by changing the voltage polarities (positive-SET/negative-RESET). However, if fixing the BE (Pt) to be grounded, and applying the negative voltage at TE, the HRS device can be set to the LRS again, as shown in the Fig. 2.57. After the set, a subsequent negative voltage sweep can bring the device back to the HRS state. The above mentioned switching mode is called unipolar, which is found coexisted with bipolar switching mode. Similar behaviors are observed in Goux *et al.* work [103]. They fabricated TiN/HfO₂/Pt kind of RRAM stacks, the devices were measured under DC sweep. As shown in the Fig. 2.58, when a positive voltage is applied to the Pt -

Fig. 2.57: Both bipolar and unipolar resistive switching behaviors were measured in the HfO_x based RRAM. (a) Measured I-V curves for bipolar switching. (b) Measured I-V curves for unipolar switching [102].

electrode, the devices possess unipolar switching. However, if the Pt is grounded, and a positive voltage is applied to the TiN electrode, the device would be set to LRS. Under this condition, only a negative voltage applied to the TiN electrode can reset the device to the HRS, for which is considered as bipolar switching. Besides, others

Fig. 2.58: Typical unipolar reset and set switching traces obtained for a cell diameter of $100 \mu\text{m}$ and for both the 5 nm (cross and open symbols) and 15 nm (closed symbols) HfO_2 thicknesses; for the 5 nm thickness, trace 1 (diamonds) programs the cell into the LRS state, then trace 2 (open Squares) is performed from 0 to 1.2 V , then trace 3 (triangles) is performed from 0 to 1.4 V , and finally trace 4 (crosses) is performed from 0 to 1.8 V . Typical consecutive bipolar switching loops between the LRS state and an I_R [103].

researchers also fabricated the $\text{M}/\text{HfO}_x/\text{TiN}$ stack, in which M stands for Pt , Cu , W and Au , respectively [104, 105]. Results indicate that after the forming step, the device can have both bipolar and unipolar switching modes.

2.5.6 First-principles Simulation on the RRAM

HfO_2 based resistor random access memory (RRAM) takes advantage of oxygen-related defects in its principle of operation. However a clear atomic picture of the resistive switching phenomenon is not available yet. Clima *et al.* conducted an *ab-initio* molecular dynamics study, in which the oxygen diffusivity in amorphous hafnia

with various oxygen concentrations are discussed [106]. For his work, a crystalline stoichiometric HfO_2 supercell contains 64 Oxygen atoms and 32 Hf atoms were implemented to generate the amorphous model by molecular dynamic simulation. The amorphization process was achieved by the melting-quenching scheme. The obtained atomic configuration of amorphous structured met the experimental measurement results very well. Since for a well-designed valence-change bipolar RRAM stack, it possesses two distinctive resistance levels, named HRS and LRS. The understanding of the switching mechanism between the HRS to LRS depends on the generation/recovery of oxygen vacancy defects. When sufficient amount of oxygen vacancies are generated and aligned to be a conductive filament, the RRAM switches to the LRS. On the other hand, when the oxygen vacancies are annihilated by the mobile oxygen ions, the conductive filament is annihilated, and the device switches to the HRS. Because the switching of the RRAM strongly relies on the oxygen movement in the HfO_2 , therefore, the oxygen ion diffuse barrier is calculated. One example with the oxygen vacancy diffusion in the $\text{HfO}_{1.97}$ is shown in the Fig. 2.59. As we can see, with the time going, the oxygen vacancies diffuse towards the stable state through the minimum energy path (MEP) by the thermal induced driven force. The movements of the vacancies dramatically affect the local electron density, and subsequently change the conductivity of the oxides. Other than that, Wang *et al.* studied the role of Ge doping in improving the RRAM uniformity through the first-principles calculation [107]. Since the physical characterization (EELS, XPS and EDX) results indicating the oxygen deficiency in the Ge doped oxides, thus the oxygen-vacancy formation energy with the presence of Ge was studied. Monoclinic HfO_2 supercell with 96 atoms is selected, and the doping of Ge is achieved by replaced first-principles calculation [107]. Since the physical characterization (EELS, XPS and EDX)

results indicating the oxygen deficiency in the Ge doped oxides, thus the oxygen-vacancy formation energy with the presence of Ge was studied. Monoclinic HfO_2 supercell with 96 atoms is selected, and the doping of Ge is achieved by replaced-

Fig. 2.59: Process of V_O diffusion in $\text{HfO}_{1.97}$ during a time-accelerated molecular dynamic at 1573 K. the cyan and red atoms corresponded to Hf and oxygen, respectively. The grey densities correspond to the localization of electronic densities bound to the oxygen vacancies [107].

one Hf atom by Ge atom, and one oxygen atom near the Ge was removed to introduce the oxygen vacancy. Through the structure relaxation and system energy calculation, the formation energy calculated based on Eq. (2.3)

$$\Delta H_f^{(O,0)} = E(O,0) - E(0) + (\Delta\mu_O + \mu_O^{solid}) \quad \text{--- Eq.(2.3)}$$

are 5.78 eV for the pure HfO_2 and 3.6 eV for the Ge doped case, respectively. Based on the above simulation results, the author infer that the reduction of the formation energy that results in more vacancies generated at the vicinity of the dopant, and consequently reduced the randomness of the process of the conduction filament formation and annihilation. Similar result that in the Al doped case the formation energy of the oxygen vacancy near the dopant will reduced is obtained by Gao *et al.* [108].

2.6 Summary

Bias temperature instability (BTI) is a critical reliability issue results from the defective Si/SiO₂, Si/SiON or Si/HfO₂ interface structure, as well as the bulk defects. Although in the early days, the BTI phenomenon on the large scale, buried channel SiO₂ device didn't get much attention due to insignificant influence on the device reliability, however the significant BTI degradation observed in the ultra-thin SiON gate stack results in intensive study on this reliability issue. In the previous several sub-chapters, we have summarized various experimental results on the SiON gated device, which show that the presence of nitrogen in the gate oxide severely worsened the NBTI performance. Meanwhile we have discussed a large number of physical models based on researchers' experiment or simulation work, to explore the degradation procedure. However, BTI still remains as a highly controversial issue, which needs to put more efforts to understand the exact degradation. Moreover, we extend our literature review to the BTI study on the hafnium oxide gated device. First we review the experimental evidences that address the low quality of the Si/HfO₂ interface as well as the inherently high defect density in the bulk HfO₂. Thereafter, we discuss the PBTI and NBTI impact on the device reliability. Although plethora of observations, explanations, as well as possible physical models are available in literature for the hafnium oxide gated device, however more creative studies are still need in order to reach a unified conclusion. Finally, we discuss the application of hafnium oxide in the RRAM. The effect of the ion doping and the roles of the electrodes are elaborated.

Chapter 3 FIRST-PRINCIPLES SIMULATION AND METHODOLOGY

3.1 Introduction

In the previous chapter, we reviewed the reliability issues on the SiON and hafnium oxide gated device in terms of BTI, and meanwhile we also discussed the application of the hafnium oxide as the switching layer in the RRAM. Currently, most of the BTI reliability issues are ascribed to defect generation/evolution in the material during the stressing. And for the RRAM application, the switching mechanisms are also attributed to the defect generation/recombination. Therefore, study the defects properties in an atomic level will be an advantage to understand the exact mechanism.

This chapter is divided into three main sections. Section 3.1 briefly presents the density functional theory (DFT). Since the current DFT is built on two Hohenberg–Kohn theorems [91, 92], so in the first part, we will review these two fundamental theorems. However, the original DFT suffers the band gap underestimation due to the introduction of an exchange-correlation term, therefore, in the second part-section 3.2, we will present a method based on the hybrid function scheme to perform the band gap correction. In section 3.3, we will review the definition of defect formation energy as well as the charge transition level (CTL).

3.2 Density Functional Theory

The very fundamental Schrödinger equation allows the physicist to describe the many-body system (atoms, molecules, and condensed materials) in a theoretical way (Eq. (3.1)),

$$\hat{H}\psi_i(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N, \vec{R}_1, \vec{R}_2, \dots, \vec{R}_M) = E_i\psi_i(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N, \vec{R}_1, \vec{R}_2, \dots, \vec{R}_M) \quad \text{--- Eq.(3.1)}$$

Where $\hat{H}$ is the Hamilton operator of the many-body system that consists of M nuclei and N electrons in the absence of magnetic or electric fields.

In order to obtain the solution of the many-body system, it is essential to decouple the Hamilton operator into the electron-electron Coulomb interactions and the complex structures of the nuclei that emerge from the combined effects of all the interactions. Therefore the $\hat{H}$ can be re-written as

$$\hat{H} = \hat{H}_e + \hat{H}_N + \hat{H}_{e-N} \quad \text{--- Eq.(3.2)}$$

Where

$$\hat{H}_e(r) = \hat{T}_e(r) + \hat{V}_e(r) = -\sum_i \frac{\hbar^2}{2m} \nabla_{r_i}^2 + \frac{1}{2} \sum_{i,i'} \frac{e^2}{|r_i - r_{i'}|} \quad \text{--- Eq.(3.3)}$$

The first term is the kinetic energy of the electrons, and second one is the Coulombic interaction between all the electrons except itself.

$$\hat{H}_N(R) = \hat{T}_N(R) + \hat{V}_N(R) = -\sum_j \frac{\hbar^2}{2M_j} \nabla_{R_j}^2 + \frac{1}{2} \sum_{j,j'} V_N(R_j - R_{j'}) \quad \text{--- Eq.(3.4)}$$

The first term is the kinetic energy of the nuclei, and second one is the Coulomb interaction between all the nucleus except itself. And

$$\hat{H}_{e-N}(r, R) = -\sum_{i,j} \hat{V}_{e-N}(r_i - R_j) \quad \text{--- Eq.(3.5)}$$

describes the electron and nuclei interaction.

However the method to solve the above equations is the most challenge task. The critical issue is the development of methods to treat electronic correlations with sufficient accuracy that one can predict the diverse array of phenomena exhibited by matter. Although many expressions have been developed [109, 110], the density functional theory (DFT) is considered to be the most successful one to give an almost perfect description of the many body system.

Generally, DFT is a quantum mechanical modeling method used in physics and chemistry to investigate the electronic structure (principally the ground state) of many-body systems, in particular atoms, molecules, and the condensed phases, and it is first put on a firm theoretical footing by the two Hohenberg–Kohn theorems (H–K) [111, 112].

(1) For any system of interacting particles in an external potential V_{extr} , the potential V_{extr} is determined uniquely, except for a constant, by the ground state particle density $\rho_0(\mathbf{r})$.

➤ Corollary1: since the Hamiltonian is thus fully determined, except for a constant shift of the energy, it follows that the many-body wave functions for states (ground or excited) are determined. Therefore all properties of the system are completely determined given only the ground state density $\rho_0(\mathbf{r})$.

(2) A universal functional for the energy $E[\rho]$ in terms of the density $\rho(\mathbf{r})$ can be defined, valid for any external potential V_{extr} . For any particular V_{extr} , the exact ground state energy of the system is the global minimum value of this functional, and the density $\rho(\mathbf{r})$ that minimizes the functional is the exact ground state density $\rho_0(\mathbf{r})$.

- The functional $E(\rho)$ alone is sufficient to determine the exact ground state energy and density. In general, excited states of the electrons must be determined by other means.

Within the framework of above theorems, the intractable many-body problem of interacting electrons in a static external potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential. The effective potential includes the external potential and the effects of the Coulomb interactions between the electrons, *e.g.*, the exchange and correlation interactions. And the complete many-body Schrodinger equation can be re-write in to the following three equations:

$$\{-\nabla^2 + \hat{V}_{KS}[\rho(r)]\}\psi_i(r) = E_i\psi_i(r) \dots Eq.(3.6)$$

$$\hat{V}_{KS}[\rho(r)] \equiv v(r) + \hat{V}_{Coul}[\rho(r)] + \hat{V}_{xc}[\rho(r)] = v(r) + \int dr' \frac{\rho(r')}{|r-r'|} + \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \dots Eq.(3.7)$$

$$\rho(r) = \sum_{i=1}^N |\psi_i(r)|^2 \dots Eq.(3.8)$$

where, E_{xc} the exchange and correction energy, which can be obtained through the local density approximation (LDA) or generalized gradient approximation method (GGA).

3.3 Band Gap Correction

Defects in the condensed matter strongly affect the characteristic of the material. For example, phosphorous doping in the Si will significantly changes the conductivity. Thus a theoretical understanding of the role of the defects is quite desirable, and luckily, in the past years, we have obtained remarkable knowledge about the properties of various defects, *e.g.*, self-doping in the GaN, and oxygen

vacancy in the SiO₂/HfO₂ [113]. However, the increasing complexity of the material composition and the complicated interface structure between different materials make the defects properties harder to be detected through simple physical characterization tools, since usually these are deep level defects and possess ionization energies much higher than the thermal energies. First-principles calculation, based on the density functional theory, is considered to be an essential simulation method to study these defects, to identify the defect levels in the band gap. But due to the discontinuity of the functional derivative of the exchange and correlation energy, the band gap of an insulator calculated from the density functional theory by using the local density approximation (LDA) or generalized gradient approximation (GGA) suffers the band gap error-smaller than the theoretical value, and thereafter affects the position of the defect level in the band gap, therefore it is necessary to add in some physical correction to obtain the correct band gap, and a hybrid function scheme originally introduced by Axel Becke in 1993 is found to be effective and acceptable for the simulation results meet the experimental data [114-116].

The hybrid exchange-correction functionals are a class of approximations to the exchange-correction energy functional in the density functional theory that incorporate a portion of exact exchange from Hartree-Fock theory with the implicit density functionals. Usually it is constructed as a linear combination of the Hartree-Fock exact exchange function (E_x) and any number of exchange and correction explicit density function. Here in our calculation, we consider the class of hybrid functionals by replacing a fraction- α of the Perdew-Burke-Ernzerhof approximation (PBE) exchange with exact exchange, as shown below:

$$E_x^{hybrid} = \alpha E_x^{exact} + (1 - \alpha) E_x^{PBE} \quad (3.9)$$

Usually, the correlation potential is taken unmodified from the generalized gradient approximation (GGA), whereas the correction coefficient α is assigned to 0.25 and mix with the PBE to get the hybrid potential, which noted as PBE0, whereas the semilocal one is denoted as PBE. The value of α has been rationalized in the case of molecular systems and is considered to be a good comprise for many systems, however there is no firm theoretical justification for this choice and the optimal value is admittedly system or even property dependent.

To verify the correction scheme, a set of the calculated band gap is listed in Fig. 3.1 [117], as we can see that improvement is essentially evident for materials such as Ge or InN which have a vanishing or even negative band gap in semilocal approximation. However, it is also evident that the improvement of the PBE0 over the PBE is not systematic. In the PBE0, band gaps are overestimated for low band-gap materials and underestimated for large band-gap materials. It is also clear why the -

Fig. 3.1: Calculated versus measured single-particle bandgaps for 15 different materials. PBE: open disks, PBE0: filled disks [117].

fraction $\alpha=0.25$ is just a good compromise rather than a universal parameter. Anyway, in our calculation of HfO_2 and SiO_2 , by using the $\alpha=0.25$, we can obtain the correct band gap.

3.4 Defect Formation Energy and Charge Transition Level

It is well known that the defects play a very important role in modifying the material properties. However the understanding of static and dynamic properties of defects is still quite limited. Fortunately theoretical calculations enable the researchers to study various defects through atomic modeling. Within all the defect properties obtained through the theoretical calculations, the formation energy is one of the most useful parameters that determined the generation/stability of the defect, therefore we will talk about it in this section.

Generally, in the condense material, the types of defect can be classified as line defects, planar defects, bulk defects and point defect. The formation of these defects strongly depends on the crystal structure, charge status as well as the electron chemical potential. Therefore, the theoretically interpretation of the formation energy of the defect or impurity X in the charge state q is written as:

$$E^f[X^q] = E_{\text{tot}}[X^q] - E_{\text{tot}}[X, \text{bulk}] - \sum_i n_i \mu_i + q[E_F + E_v + \Delta V] - \dots - Eq. (3.10)$$

Where $E_{\text{tot}}[X]$ is the total energy derived from a supercell calculation with one impurity X or defect in the cell, and $E_{\text{tot}}[\text{bulk}]$ is the total energy for the equivalent supercell without any defects. n_i indicate the number of atoms of type i (host atoms or impurity atoms) that have been added to ($n_i > 0$) or removed from ($n_i < 0$) the supercell when the defect or impurity is created, and the μ_i are the corresponding chemical potentials of these species. Usually, we consider these chemical potentials as the

exchange energy with the reservoirs. E_F is the Fermi level, referenced to the valance band maximum in the bulk. Due to the choice of this reference, we have to explicitly put in the energy of the bulk valence-band maximum- E_v , in the expression for formation energies of charged states. Also, we have to add a correction term ΔV , to align the reference potential in our defect supercell with that in the bulk.

Here we show an example to calculate the formation energy of the Mg substitution doping in the GaN bulk [113]:

$$E^f[Mg_{Ga}^0] = E_{tot}[Mg_{Ga}^0] - E_{tot}[GaN, bulk] - \mu_{Mg} + \mu_{Ga} \quad \text{--- Eq.(3.11)}$$

As indicated in the above equation, first we have to calculate the total system energy of the model with the Mg doping and the total energy of the perfect bulk GaN, meanwhile we have to get the chemical potential of the Mg and Ga, respectively, then apply to the Eq. (3.11), we can get the formation energy of the Mg replacement doping in the GaN.

Moreover, as indicated in the Eq. (3.11), the defect structures in the material are always easy to be affected by the charge states. Once the defect structure captures the charges, the atomic configuration will change subsequently in order to compensate the energy increase by capturing the charge, which we call structure relaxation. Therefore it is necessary to investigate the impact of the charges on the defect structure, and the charge transition level (CTL) is the physical parameter identifies the status of the defect under different charge states. It defined as the value of the Fermi level for two different charge states (*e.g.*, neutral and positive in our case) have the same formation energy. As shown below:

$$E_F = \frac{1}{q} \{E_{tot}[X, bulk] - E_{tot}[X^q] + \sum_i n_i \mu_i\} - E_v - \Delta V \quad \text{--- Eq.(3.12)}$$

Here, in Fig. 3.2, an example shows the formation energy and CTL of the hydrogen defects in the GaN is illustrated [118]. As can be seen that, the formation energy is -

Fig. 3.2: Formation energy of interstitial hydrogen as a function of Fermi level in to different semiconductors, illustrating the qualitatively different position of the $\epsilon(+/-)$ level. In GaN (a), $\epsilon(+/-)$ lies within the band gap, whereas in ZnO; (b), $\epsilon(+/-)$ lies above [118].

strongly dependent on the Fermi Level, *e.g.*, for the neutral and negatively charged state, when the E_F is lower than 1.2 eV, H^0 is more stable than H^- , however, when the E_F is raised up to 1.2 eV, the H^- become more stable, hence with the structure adjustment, the defect will hold the electron, and be negatively charged.

3.5 Summary

To summarize, in this chapter, first we discussed the theoretical background of the density functional theory. Although it is easy to have the Schrödinger equation for the many-body system, however solving the Schrodinger equation of the many-body system is the most challenge task, and DFT provides a relative accurate way to study

the material electronic structure through mathematical calculation. Then we discussed the band gap under estimation due to the discontinuity of the functional derivative of the exchange and correlation energy in the DFT, and a band gap correction scheme named hybrid functionals is introduced. Finally, we presented the physical definition of the formation energy as well as the charge transition level in the defect study, which are two useful parameters to identify the defect dynamics under the neutral or charged state.

Chapter 4 ATOMIC VIEW OF RECOVERABLE AND PERMANENT COMPONENT

4.1 Impact of Two-fold Coordinated Nitrogen on the Generation of Deep-level Hole Traps under Negative-bias Temperature Stressing

4.1.1 Introduction

Negative-bias temperature instability is widely recognized as a critical reliability issue of the state-of-the-art CMOS technology. It is due to a progressive generation of interface states and near-interface positive oxide trapped charges under the combined effect of a negative gate voltage and an elevated temperature. The positively charged defects shift the threshold voltage (V_t) of the p-MOSFET negatively and degrade the inversion charge mobility, thereby reducing its drive current. When the shift exceeds a certain design tolerance, functional failure of the integrated circuits may result. NBTI was observed as early as the beginning of the MOS technology in the mid 1960's [119] but it received very little attention at that point of time because of the buried-channel structure of the p-MOSFET. In the buried-channel structure, the holes (which play an important role in the generation of the oxide/interface defects),

are separated from the silicon/silicon dioxide (Si/SiO_2) interface. This has helped mitigate the instability problem to a large extent in older CMOS technologies.

The aggressive scaling of the CMOS technology into the deep sub-micrometer regime at the dawn of this century changed the situation drastically. A switch from a buried- to a surface-channel structure had to be made to keep short-channel effects in check. This move places the holes exactly at the interface, exacerbating the NBTI issue. The problem is further compounded by the increase in oxide field as a result of the reduction in the gate oxide thickness. Another main reason for NBTI becoming so important in recent years lies in the presence of nitrogen in the gate oxide [8, 120]. Nitrogen has been incorporated into the ultra-thin gate oxide because it helps block boron from the heavily doped p-type polysilicon gate (required for a surface-channel structure), thus improving the V_t uniformity of as-processed devices. Nitrogen also increases the dielectric constant of the gate oxide, which in turn allows a thicker SiO_2 to be used for suppressing gate leakage without compromising the effective capacitance. But the presence of nitrogen in the gate oxide has substantially worsened the NBTI problem [8].

In the past decade, there were many electrical characterization studies on the NBTI of p-MOSFETs employing the ultra-thin oxynitride (SiON) gate oxide. The degradation of the p-MOSFET is typically characterized by a much weaker dependence on stress time and temperature as compared to that of the SiO_2 counterpart [8, 121, 122]. This behavior is ascribed to the increased trapping of holes within the ultra-thin gate oxide by defect precursors associated with nitrogen. It is also observed that more trapped holes could be sensed electrically at a given delay after the termination of stress, implying the existence of a larger density of hole traps having deeper energy levels. Ang and other researchers, for the first, proposed the presence of

deep-level hole traps (DLHTs) after negative-bias temperature stressing [22, 24, 25, 57, 74, 123-126]. These hole traps have energy states above the Si Fermi potential, which result in them having the ability to retain their positive trapped charges for a very long time under negative gate biasing, even in a direct-tunneling gate oxide. However, the nature of the hole traps are seldom explored. A recent study on electron paramagnetic resonance has shown that hole trapping sites in the SiON have a crossover g tensor that is different from that of the E' centers ($O_3\equiv Si\bullet + Si\equiv O_3$) in the SiO_2 [61]. The difference is ascribed to the neighboring O atoms being replaced by nitrogen atoms. A defect configuration of the form $N_3\equiv Si\bullet + Si\equiv O_3$ (referred to as a K center) is shown to have a crossover g tensor similar to that observed experimentally. But knowledge on how neighboring nitrogen atoms impact the properties of the oxygen vacancy defect (V_O) remains very limited. In particular, the mechanism behind the generation of DLHTs is not known at this stage.

Therefore, in this chapter, we show via first-principle simulation based on the density functional theory that the coordination of the neighboring nitrogen atom plays a paramount role on the structural relaxation of V_O following the capture of a hole. In particular, a two-fold coordinated nitrogen atom is shown to induce very significant structural relaxation of V_O . The resultant trapped-hole site has a very deep charge transition level in the SiO_2 band gap. On the other hand, in the absence of the nitrogen atom or if the neighboring nitrogen atom is threefold coordinated (*e.g.*, the dangling nitrogen bond is terminated by a hydrogen atom), the resultant trapped-hole site would have a much shallower charge transition energy level. We first present experimental results showing a clear correlation between nitrogen concentration in the gate oxide and the density of DLHT generation by negative-bias temperature stress. This is then

followed by a detailed discussion of the simulation results. Based on the new insights gained, a unified explanation for hole-trap generation and relaxation is given.

4.1.2 Experimental Details

The test devices (DUTs) were p^+ polysilicon gate p-MOSFETs. The SiON gate oxide was formed via decoupled plasma nitridation of an *in-situ* steam generated base oxide. The concentration of nitrogen at the polysilicon/SiO₂ interface varies from ~6 to 13 atomic percent. The equivalent oxide thickness is approximately 1.5 nm, extracted from capacitance-voltage measurement. Negative-bias temperature stressing was carried out at a temperature of 100 °C and at an oxide field of about 10 MV/cm. Device degradation was monitored in terms of threshold voltage shift (ΔV_t), obtained from linear drain current versus gate voltage (I_d - V_g) curves using the constant subthreshold drain current method. An ultra-fast switching method, which is essentially a modified form of the pulsed current-voltage (PIV) method, was employed for I_d - V_g measurement [127]. An advantage of this method is that the very narrow gate pulses (100 ns) enhance probing of shallow hole traps (SHTs) with short de-trapping time constants.

4.1.3 Deep-level Hole Trap Generation

In this section, we present experimental results showing a correlation between DLHT generation and the concentration of nitrogen in the gate oxide. We first outline the experimental procedure for revealing and determining the amount of DLHTs generated by negative-bias temperature stressing before discussing the impact of nitrogen on the generation of DLHTs.

Fig. 4.1(a) depicts the evolution of ΔV_t during the stressing and relaxation phase of a dynamic NBTI cycle. As expected, ΔV_t increases negatively during stressing, indicating the progressive build-up of positive oxide and interface trapped charge. Trapping of holes by oxide/interface defects is believed to be one of the major mechanisms. By virtue of their proximity to the Si/SiO₂ interface, one would expect hole trapping at near-interface defect states to occur spontaneously the moment the gate stress voltage is applied (Fig. 4.2(a)). Hole trapping at defect states further away-

Fig. 4.1: (a) evolution of threshold voltage shift ΔV_t during negative-bias temperature stressing and relaxation. Gate voltages applied in the respective phases are as indicated. During relaxation, the gate voltage was first switched to 0 V and maintained at that level for 4×10^3 s. It was then switched to +1.5 V for 1×10^3 s (but PIV measurement was carried out with a 0-V offset) after which it was returned to 0 V for another 4×10^3 s. (b) ΔV_t evolution during the first 0 V and the subsequent +1.5 V relaxation phases plotted on a linear-log scale. The slopes of the regression line are as shown.

from the Si/SiO₂ interface would take longer time, but in the case of the direct tunneling gate oxide, these defect states are not likely to be positively charged because of the correspondingly large probability of hole de-trapping to the gate (Fig. 4.2(b)).

Based on this viewpoint, the net hole trapping is expected to reach a saturation level very quickly (< 1 s) after the application of the gate stress voltage, with the saturation level limited by the available density of defect states. Some authors attributed the subsequent progressive shift of ΔV_t over prolonged stress period to the generation of interface states, under the framework of the well-known *Reaction-Diffusion model (R-D model)* [42]. We will return to this point at a later stage citing recent experimental evidence which does not support the that the long term evolution of NBTI is driven by hydrogen transport.

Fig. 4.2: Schematic energy band diagrams of the MOS structure under negative gate biasing. (a) Hole capture by a near-interface defect state where the trapping probability is greater than the de-trapping (dashed arrow) probability. (b) In an ultra-thin gate oxide, hole capture by a defect state located further from the Si/SiO₂ would be less likely because of a correspondingly large probability of de-trapping to the gate, *i.e.*, the defect is neutral and does not contribute to NBTI. In both (a) and (b), structural relaxation of the defect is ignored. (c) Significant lattice relaxation (via interaction with phonon) upon hole capture results in an “upward” shift of the positively charged defect state (*i.e.*, negative U behavior) and the creation of a deep-level hole trap.

Just as one expects spontaneous hole trapping to take place at near-interface defect states when the negative gate stress voltage is applied, one also expects spontaneous de-trapping of holes from near-interface defect states when the gate stress voltage is removed, triggering a decrease in the magnitude of ΔV_t . This is evident from the abrupt decrease in $|\Delta V_t|$ at the onset of the relaxation cycle. The decrease is also observed to approach quasi-saturation fairly quickly; after the initial abrupt decrease, the subsequent decrease of $|\Delta V_t|$, over an extended recovery period of 4×10^3 s, is less than 10% of its value at the end of the stress cycle. This quasi-saturation behavior tends to signify the near completion of the hole de-trapping process.

At this juncture, it is essential to point out that some authors ascribed the recovery of ΔV_t to the repassivation of dangling Si bonds by back-diffusing hydrogen, under the framework of the *R-D model* [39, 41, 42]. This school of thought presumes-

Fig. 4.3: (a) comparison of $|\Delta V_{t,r}|$, the decrease in the magnitude of ΔV_t per relaxation cycle as a function of the number of stress/relax cycles between experiment and simulation based on the *R-D model*. For the ease of comparison with the simulation data, $|\Delta V_{t,r}|$ is normalized by the value for the first relaxation cycle. (b) Comparison of $|\Delta V_{t,r}|$ and $|\Delta V_{t,s}|$, the increase in the magnitude of ΔV_t per stress cycle. $|\Delta V_{t,EoR}|$ refers to the magnitude of ΔV_t at the end of each relaxation cycle [48].

that hole trapping saturates very quickly and the subsequent evolution of $|\Delta V_t|$ results from the generation of Si dangling bonds via the dissociation of Si-H bonds, with the rate of generation limited by the transport of hydrogen away from the Si/SiO₂ interface. However, recent experiments on repetitive stressing and relaxation did not support this proposition [18, 20, 46, 48, 63]. In one of our experiment, we found that the amount of ΔV_t recovered per relaxation cycle remain the same regardless of the number of stress and relaxation cycles [48], reproduced here as -Fig. 4.3(a)). This characteristic does not conform to the hydrogen transport model, which stipulates that the amount of interface state repassivation per cycle must decrease with increase in the number of stress and relaxation cycles. This is because repeatedly cycling the DUT between stress and relaxation would result in more hydrogen being transported further away from the Si/SiO₂ interface, thus gradually limiting the amount of interface states which could be repassivated per relaxation cycle. Detailed investigation over many stress and relaxation cycles has instead shown a cyclic behavior of ΔV_t , the increase in the magnitude of ΔV_t in a stress cycle would become reproduced here as Fig. 4.3(b). Our recent results strongly indicate that the ΔV_t fluctuations one typically observed during dynamic NBTI were the consequence of an ensemble of switching hole-trap precursors that were being repeatedly charged (during stressing) and discharged (during relaxation) [128]. Oxygen vacancy defects are a major source of switching oxide traps in the SiO₂ [69, 128] These defects are typically found within the intrinsic sub-stoichiometric oxide region near the Si/SiO₂ interface almost equal to the decrease in the ensuing relaxation cycle reproduced here as Fig. 4.3(b)).

It should be pointed out that the above first-order treatment on hole trapping/detrapping ignores structural relaxation that occurs upon hole capture. This can lead to an underestimation of the extent of hole trapping in an ultra-thin gate oxide

[50]. As depicted in Fig. 4.2(c), if there is significant lattice relaxation at the defect site (via interaction with phonons) following the capture of a hole, it could induce an upward shift of the defect level (, negative U behavior). This upward shift renders a significant increase of the hole emission time. If the defect level remains above the Si Fermi potential after the removal of the negative gate stress voltage, hole emission could not readily take place via the capture of electrons from the Si substrate, yielding a DLHT. This is unlike shallow hole traps which are spontaneously discharged when their energy levels are lowered below the Si Fermi potential upon the termination of the stress.

The existence of these DLHTs can be shown by switching the gate voltage to a moderately positive value during the course of the first 0 V relaxation phase, after quasi-saturation is reached. The positive gate voltage lowers DLHTs below the conduction band edge of Si, promoting the emission of trapped holes via the capture of electrons from the n-Si well/substrate. If the hole de-trapping process has indeed completed (as suggested by the quasi-saturation level), one would not expect the recovery of ΔV_t to be affected by a positive gate voltage. However, this inference is not supported by Fig. 4.1(a), which clearly shows a non-negligible, abrupt decrease in $|\Delta V_t|$ immediately following a switch of the gate voltage from 0 to +1.5 V. This implies that a substantial fraction of the subsequent evolution of $|\Delta V_t|$ following the initial transient-like increase during the stressing phase comprises further trapping of holes but at much deeper energy levels.

In Fig. 4.1(b), we replotted the ΔV_t recovery data of the first 0 V and the subsequent +1.5 V relaxation phases against the relaxation time (defined with respect to the instant when the gate voltage was changed) on a linear-log plot. A logarithmic time dependence spanning many orders of time magnitude is evident, particularly for

the 0 V case, as has also been observed in other studies. A caution about the tendency for one to conclude that recovery has reached saturation based on the “plateau” observed in Fig. 4.1(a). As shown in Fig. 4.1(b), the recovery has in fact not saturated although its rate is decreasing continuously with time. Similar logarithmic time dependence has been observed in the past for the thick gate oxide and has been attributed to the spatial distribution of oxide traps [129]. In our case, the gate oxide thickness is less than 2 nm and the electron tunneling time constant is on the order of microseconds or less. Clearly, the spatial distribution of oxide traps could not explain the wide spread of recovery time constants. In order to account for the observation, it is necessary for one to consider a wide spread in the relaxation energy of the hole traps, whereby relatively deep hole traps are responsible for long recovery time constants.

The decrease in $|\Delta V_t|$ following a switch of the gate voltage to a positive value is permanent, , a subsequent switch of the gate voltage back to 0 V does not return $|\Delta V_t|$ to the level prior to the gate polarity change. A similar statement can be made in regards to the rapid decrease in $|\Delta V_t|$ at the onset of the 0 V relaxation phase, , the level of ΔV_t shift observed in the stressing phase is not seen again unless the DUT is re-stressed at the same negative gate stress voltage for the same period (cf. Fig. 4.3). The decrease in $|\Delta V_t|$ during the transition from 0 to +1.5 V gate recovery voltage is also found to occur only once at the first instant when the gate voltage is made positive; a subsequent switch to a more positive gate voltage yields no further decrease in $|\Delta V_t|$ (not shown). This implies that a significant part of the hole de-trapping process is irreversible, and could possibly involve the reconstruction of broken bonds, which totally eliminate the trapped-hole site. This inference is supported by our simulation work (see later discussion). The remnant ΔV_t is rather permanent, as is evident from the substantially reduced recovery rate under +1.5 V (Fig. 4.1(b)). This permanent ΔV_t

may be attributed to trapped-hole sites which have undergone so substantial structural relaxation that subsequent bond reformation requires extremely long time.

For the purpose of subsequent discussion, we define the difference between $|\Delta V_t|$ at the end of the first 0 V relaxation period and that after the switch to a positive gate voltage was made as the component of $|\Delta V_t|$ due to DLHT ($|\Delta V_t|_{\text{DLHT}}$). The difference between $|\Delta V_t|$ at the end of the stress cycle and that at the end of the first 0 V relaxation period is defined as $|\Delta V_t|$ arising from relatively shallow hole traps ($|\Delta V_t|_{\text{SHT}}$). Since the recovery of ΔV_t in the first 0 V relaxation phase reaches quasi-saturation fairly quickly, varying its period would not significantly affect the values of $|\Delta V_t|_{\text{DLHT}}$ and $|\Delta V_t|_{\text{SHT}}$. Our classification in fact underestimates $|\Delta V_t|_{\text{DLHT}}$ since the permanent fraction is excluded.

Fig. 4.4 shows the dependence of $|\Delta V_t|_{\text{SHT}}$ and $|\Delta V_t|_{\text{DLHT}}$ on the nitrogen concentration in the gate oxide (which corresponds to that at the polysilicon/SiO₂ -

Fig. 4.4: Dependence of threshold voltage shift (extracted after 1×10^4 s stress) contributed by shallow and deep hole traps on the nitrogen concentration at the polysilicon/SiO₂ interface in atomic percent.

interface). As expected, hole trapping increases with nitrogen concentration. More importantly, there is also a distinct increase in the amount of DLHTs, indicating that a higher nitrogen concentration tends to promote greater DLHT generation for a given stress period. The time dependence of $|\Delta V_t|_{\text{SHT}}$ and $|\Delta V_t|_{\text{DLHT}}$ are examined in Fig. 4.5. The former exhibits very weak dependence on the stress time while the stress time dependence of the latter is clearly much stronger. This observation supports the notion that generation of the relatively permanent DLHTs is generally more sensitive to time since it involves more significant structural relaxation. Nevertheless, an increase in the nitrogen concentration tends to reduce the time dependence, implying that the generation of DLHTs is accelerated by more nitrogen. As will be shown later via simulation, a nitrogen atom in a twofold coordinated form facilitates the structural relaxation of an adjacent oxygen vacancy defect.

Fig. 4.5: Evolution of threshold voltage shift $|\Delta V_t|$ arising from shallow hole traps (SHTs) and deep-level hole traps (DLHTs) for DUTs having varying nitrogen concentration at the polysilicon/SiO₂ interface.

4.1.4 Simulation Details

We have chosen the cristobalite SiO_2 for our simulation study, in view that its Kohn-Sham density of states is highly similar to that of the amorphous SiO_2 [130, 131]. A supercell consisting of eight α -cristobalite unit cells was employed to provide sufficient isolation of defects. Fig. 4.6 depicts the supercell used in our study. The dimensions of the supercell are $9.96 \text{ \AA} \times 9.96 \text{ \AA} \times 13.90 \text{ \AA}$. The total number of atoms in the supercell is 96 and the density is 2.32 g/cm^3 . An oxygen vacancy defect was manually created by removing an oxygen atom to induce the formation of the $\text{O}_3\equiv\text{Si}-\text{Si}\equiv\text{O}_3$ dimmer configuration. In the case of the nitride SiO_2 , a neighboring oxygen atom of the Si-Si dimmer was replaced with a nitrogen atom. Since the valency of nitrogen is one less than that of oxygen, a direct replacement would result in the nitrogen atom having an unpaired spin, a dangling bond. Leaving the dangling bond -

Fig. 4.6: Supercell comprising 8 α -cristobalite unit cells used in our simulation study.

The bigger blue spheres denote Si atoms while the red spheres denote oxygen atoms.

unpassivated forms our twofold coordinated nitrogen defect model. Both experimental and simulation data supporting the formation of this two-fold coordinated nitrogen defect in the nitride SiO_2 have been reported [130]. Passivating the dangling bond with a hydrogen atom leads to the threefold coordinated nitrogen defect. As will be shown later, the coordination of the nitrogen atom has a significant impact on the structural relaxation and resultant stability of V_O following the capture of a hole.

The calculation was carried out using the PWSCF code of the QUANTUM ESPRESSO package [132]. Based on the density-functional theory, the electronic structure in the calculation was described by generalized gradient approximation developed by Perdew, Burke, and Ernzerhof (PBE) [133]. The eigenfunctions were expanded on a plane wave basis sets, and ultra-soft pseudo-potentials were applied to each kind of atoms. For all cases, full relaxation of the system was carried out. We tested and found that a cut-off energy of 50 Ry for the wave-function expansion, and a value of 150 Ry for the charge density yielded well converged results. The Γ point was used during the whole relaxation procedure. For the charged system, a uniform background charge with an opposite sign was adopted to maintain neutrality of the whole system. This avoided divergence during the calculation of the total energy.

A well-known issue with the semilocal density-functional approach is the underestimation of the band-gap. This problem prevents an accurate determination of the position of the defect level with respect to the Si band-gap. Fortunately, a recent study has shown that a hybrid functional approach could yield an accurate SiO_2 band-gap [117]. Both approaches, however, give similar result for the charge transition level (< 0.1 eV variation). This finding allows us to make use of the semilocal approach to first determine the charge transition level and then position it accurately with respect

to the Si band edges by making corrections to the Si/SiO₂ band-edge offsets based on the results of the hybrid functional approach. Fig. 4.11 illustrates the details.

4.1.5 Discussion on Simulation Results

Structural Relaxation

As expected, the removal of an oxygen atom to form a neutral oxygen vacancy results in the two affected Si atoms moving closer to one another to form a Si-Si bond as this lowers the overall system energy. The spacing between the two Si atoms (labeled as Si(1) and Si(2)) is similar in all three cases (, pure SiO₂, SiON with a threefold coordinated nitrogen atom next to Si(1), and SiON with a twofold coordinated nitrogen atom next to Si(1)), as shown in Fig. 4.7. From electron density analysis, we confirm that there is a bond formed between the two Si atoms. Moreover, the charge density distribution (dashed circle) between the two Si atoms is similar for all three cases.

However, the neutral oxygen vacancy responds very differently upon the capture of a hole, depending on the coordination of the neighboring nitrogen atom. In the case where the nitrogen atom is absent (, pure SiO₂), the distance between the two Si atoms is increased by ~ 0.13 Å (Fig. 4.8(a)(i)). The charge density between Si(1) and Si(2) has thinned down (Fig. 4.8(a)(ii)) compared to the neutral case (cf. Fig. 4.7(a)(ii)), implying that the electronic charge is lost (or a hole is trapped) at this location. Nevertheless, the remaining unpaired spin is still almost equally shared between the two atoms. This is consistent with the marginal structural relaxation (Fig. 4.8(a)(i)), implying that the original bond between the two Si atoms could readily reform upon the recapture of an electron. In the case of the SiON with a threefold-

Fig. 4.7: (i) oxide bonding network in the vicinity of the neutral oxygen vacancy defect, depicting the distance between Si(1) and Si(2). (ii) Electron density analysis showing the distribution of electronic charge (yellow) in the same vicinity of the oxide network. The dotted circle highlights the bonding between Si(1) and Si(2).

Fig. 4.8: (i) oxide bonding network in the vicinity of the oxygen vacancy defect following the capture of a hole. (ii) Electron density analysis showing the distribution of electronic charge (yellow) in the same vicinity of the oxide network. The dotted circle highlights the electronic charge density between Si(1) and Si(2). The dotted circle is not shown in part (c)(ii) as there is no electronic charge between Si(1) and Si(2) in this case.

coordinated neighboring nitrogen atom, the separation between Si(1) and Si(2) increases to ~ 0.23 Å. Due to the increased spacing, the distribution of the remaining charge between the two atoms is different compared to the pure SiO₂ case. Instead of a continuous “dumbbell-like” shape in the latter (Fig. 4.8(a)(ii)), distinct lobes are observed on each of the two Si atoms (Fig. 4.8(b)(ii)). The similar size of the lobes implies that the remaining unpaired spin is still equally shared between the two Si atoms. But the distinct lobes on each of the Si atoms suggest that the unpaired spin is “hopping” back and forth the two Si atoms continuously. An interesting point is that the N-H bond remains pretty much intact following the hole capture, indicating that the hole is not trapped at this location but is preferentially located in-between the two Si atoms.

The situation for the case of the SiON with a twofold coordinated neighboring nitrogen atom is drastically different compared to the rest. A very significant increase in the spacing between Si(1) and Si(2), by ~ 2.45 Å, is evident (Fig. 4.8(c)(i)). The original tetrahedral structure of Si(1)O₂N has changed into an almost planar form (Fig. 4.8). Compared to the threefold coordinated nitrogen case, the bond length between the Si and N atoms has decreased from the original value of 1.64 Å to 1.58 Å after the hole capture. The former agrees well with that reported for a single Si-N bond while the latter is closer to that of a double Si=N bond (Table III). We reckon that the significant structural relaxation is made possible by a lowering of the total system energy via the interaction between the Si(1) and nitrogen dangling bonds. This is consistent with electron localization function analysis, revealing a shell-like iso-surface (dotted circle, Fig. 4.9(b)) between the Si(1) and N atoms after the hole capture. This implies greater localization of electronic charge between the two atoms. The almost planar structure of Si(1)O₂N suggests that the interaction is facilitated by sp^2

re-hybridization of the now threefold coordinated Si(1) atom. The significant structural relaxation has an important implication in that it makes the resultant trapped-hole site more permanent. The corresponding charge transition level for the defect structure would also be raised as elaborated in the subsequent section.

Fig. 4.9: Electron localization function analysis in the vicinity of the oxygen vacancy defect for the case of the twofold coordinated nitrogen atom, before (a) and after (b) the capture of a hole. The yellow regions denote iso-surfaces having a value of 0.85 (a value of 1 means the electronic charge is completely localized). An enlargement of the iso-surface between N and Si(1) following the capture of a hole signifies increased electronic charge localization between the two atoms. There is also a change in the structure of the Si(1)O₂N cluster from tetrahedral to planar after the hole capture.

TABLE 4.1. A comparison of Si-N bond length before and after the capture of a hole by the oxygen vacancy defect.

Bond Type (in bold)	Vacancy Charge State		Remarks
	0	+1	
O–Si–Si	1.64 Å	1.64 Å	Si–N (1.65 Å) Si=N (1.55 Å) (Ref. (30))
N–Si–Si (N is twofold)	1.64 Å	1.58 Å	
N–Si–Si (N is threefold)	1.73 Å	1.71 Å	

Charge Transition Level

To further probe the defect generation process, we examine the charge transition level (CTL) that corresponds to a change in the charge state of V_O from neutral to positive. The CTL is defined as the value of the electron chemical potential for which two different charge states (neutral and positive in our case) have the same formation energy, E_F . Using the neutral state as the reference [113, 130], the E^f for V_O^+ may be expressed as

$$E^f(+q) = E_{\text{tot}}(+q) - E_{\text{tot}}(0) + q[E_F + E_V + \Delta V] + E_{\text{corr}} - - - Eq. (4.1)$$

where $E_{\text{tot}}(+q)$ is total system energy after the addition of a positive charge $+q$ (q is the electronic charge (1.6×10^{-19} C)), $E_{\text{tot}}(0)$ is total system energy for the neutral state, E_F is Fermi level measured with respect to the valence band maximum of a defect free oxide, ΔV is a correction term describing the potential shift which occurs when the oxide is changed from a defect-free state to one having a defect structure, and E_{corr} is a correction term which accounts for the spurious electrostatic interactions introduced by the finite boundary condition.

A plot of E^f versus E_F is depicted in Fig. 4.10. The intersection of the E^f versus E_F characteristic with the $E^f = 0$ line gives the CTL that corresponds to the transition of V_O from a neutral to a positively charged state. For $E_F < \text{CTL}$, formation of the positively charged state is favorable, V_O^+ is stable since E^f is negative. On the other hand, V_O^+ is unstable (and would have a strong tendency to revert to its neutral state) when $E_F > \text{CTL}$, since E^f is positive in this case. For both the pure SiO_2 and SiON with a threefold coordinated neighboring nitrogen atom, the CTL for V_O^+ are comparable (~ 2.5 -3 eV). A significant increase of the CTL to ~ 5.32 eV is, however, apparent in the case of the SiON with a twofold coordinated neighboring nitrogen atom. This substantial increase in the CTL translates to a wider range of E_F for which the V_O^+

defect state is stable. We reckon that this significant rise in the CTL is made possible by the significant structural relaxation which helps achieve an even lower total system energy after the hole capture. The positions of the various CTLs in the SiO_2 band-gap are shown in Fig. 4.11.

A. Unified Model for Hole-Trap Generation and Relaxation

The simulation results clearly show that the extent of structural relaxation of V_O , as well as the influence exerted by a neighboring nitrogen atom, play a significant role on determining the CTL of the resultant hole trap. Based on this insight, we attempt to provide a unified explanation for some of the seemingly different positive oxide charges reported in the literature. We focus on the recent work of Zhang *et al.* [47], which provides a comprehensive summary on the different electrical characteristics of positive charges in the SiO_2 . Based on the different electrical -

Fig. 4.10: Formation energy versus Fermi level ($E_F = 0$ corresponds to the oxide valence band edge) characteristic for a positively charged oxygen vacancy defect. Intersection with the $E^f = 0$ line gives the charge transition level (CTL), which denotes the electron chemical potential energy for transition from neutral to a positively charged state. For $E_F < \text{CTL}$, the positively charged state is stable.

Fig. 4.11: Schematic diagram illustrating the underestimation of the oxide band gap using the semilocal density functional scheme (left), which can be corrected by the hybrid functional scheme (right). After correction, the charge transition levels (CTLs) are more accurately located with respect to the oxide and Si band edges. The oxygen vacancy defect with a twofold neighboring N gives rise to the deepest CTL (*i.e.*, furthest above the oxide valence band edge).

responses observed, the authors proposed three types of positive oxide charge: 1) as-grown hole traps; 2) cyclic positive charge (CPC); 3) anti-neutralization positive charge (ANPC). In our discussion, we assume V_O as the origin of the positive oxide charge although a similar explanation may also be applied on other types of positive oxide charge. The main essence is that the apparent differences in the electrical behavior may actually be different manifestations of the same defect (V_O) which have undergone varying degrees of structural relaxation after hole capture, depending on the surrounding atomic configurations [47]. This in turn results in a considerable spread in the hole-trap energy distribution.

Through changing the gate voltage, one modulates the Fermi level E_F at the Si/SiO₂ interface, thereby changing the occupancy of near-interface defects. For V_O 's having initially shallow CTLs (, close to the SiO₂ valence band edge), they would only

charge up positively provided the gate bias is negative enough to “lift them up” to the level of the E_F . Once they captured the hole, structure relaxation will happen to compensate the energy increase, which will subsequently determines the CTL of this V_O^+ defect. For those shallow defects, once the negative gate bias is removed, the CTLs would fall below E_F , and the corresponding V_O^+ s would revert to the neutral state (via the capture of an electronic charge). But the overall speed at which this could happen (recovery speed) depends on the distance of a given V_O^+ from the Si/SiO₂ interface, which determines how soon the trapped-hole site captures an electronic charge. In a direct tunneling gate oxide (< 2 nm in thickness), the capture of electronic charge happens almost immediately, since the tunneling time constant is on the order of microseconds or less. Therefore, V_O^+ s with shallow CTLs function as “fast” hole traps which spontaneously discharge when the negative gate bias is removed and one would need a fast electrical measurement method in order to detect them. In a thicker gate oxide, however, the recovery speed involving V_O^+ s situated further away from the interface would be much slower although near-interface V_O^+ s would also recover almost spontaneously. This is because a multi-step of electron injection is needed to neutralize the hole. Or else a relatively high oxide field (*e.g.*, Fowler-Nordheim tunneling) to enhance the electron injection can be applied to promote the former procedure. Overall, by virtue of their shallow CTLs, these V_O 's are not likely to affect device operation under nominal voltage condition (since the voltage would not be large enough to charge them up positively) but are triggered under more severe gate stress voltage or oxide field conditions. But once they are annihilated (via bond reformation), they would no longer be observed under nominal measurement condition since the voltage involved would not be large enough to lift them up again to the E_F . This group of V_O 's fits very well the description of the as-grown hole traps given in ref.

[47]. On the other hand, some of the initially shallow defects may convert to deep level hole traps due to structure change, and their behavior in the subsequent recovery phase will be discussed in the next part.

As we can expect, for V_O 's with initially deeper CTLs, they would charge up positively even under a moderate negative gate voltage. However, a transformation to a stable V_O^+ structure still needs time, since interaction with phonons is needed to create the structural relaxation that stabilizes the V_O^+ state. However, when this happens, part of the V_O^+ will remain a very stable state owing to the relatively high CTL. Removing the negative bias may not return these defects to the neutral state if the CTLs remain above E_F , they behave as DLHTs. Thus, even in a direct tunneling gate oxide, the recovery of these V_O^+ 's could take substantial time, explaining the logarithmic dependence spanning many orders of time under a 0 V gate voltage (Fig. 4.2(b)). In a thick gate oxide, the resultant time dependence may consist of another component arising from the discharging of shallow hole traps located at a distance from the interface. The application of a positive gate voltage lowers some of the CTLs below E_F , thereby neutralizing the positively charged defects, promoting the annihilation of a fraction of the DLHTs (see the abrupt decrease in $|\Delta V_t|$ upon the application of +1.5 V; Fig. 4.2(a)). At this point, it is essential to highlight that some DLHTs possess a very stable deep hole trap structure, therefore their charge states may be “pinned” and remain permanently positive even under a positive gate bias. In Fig. 4.2(a), it can be observed that switching the gate voltage from +1.5 V back to 0 V does not restore $|\Delta V_t|$ to the level prior to the application of the +1.5 V gate voltage. This implies that a fraction of the DLHTs is indeed annihilated by the positive gate voltage. It should be mentioned that although a substantial part of $|\Delta V_t|$ is eliminated almost immediately (implying that most of the bond reformation processes occur rather

quickly), a subsequent gradual decrease in $|\Delta V_t|$ spanning many orders of time remains apparent (Fig. 4.2(b)). For DLHTs which are neutralized, whether reformation of the Si-Si bond would occur clearly depends on the extent of structural relaxation. The rate of Si-Si bond reformation obviously decreases with increase in the structural relaxation. The remaining part of V_O 's could explain the ANPC of ref. [47]. It should be noted that ANPC may also be formed via interaction between trapped holes and H^+ , and this mechanism may co-exist under more severe stress condition (*e.g.*, high charge-injection fluence) which triggers the substantial release of hydrogenous species. Besides the deep level hole trap, we also have some shallow components evolving from the initially deeper CTLs after structure relaxation, it will be recovered once the bias is removed, acting as the aforementioned shallow defect.

For V_O 's which have undergone significant structural relaxation, they may be momentarily charge-compensated under a positive gate voltage and would revert to the more stable V_O^+ state upon removal of the gate bias. At this point, it is necessary for us to clarify the reason for the negligible change in $|\Delta V_t|$ when the gate voltage is switched from +1.5 V to 0 V (Fig. 4.2(a)). This is because even though the device was subjected to a +1.5-V gate voltage during relaxation, the gate bias was interrupted when measurement was made (*i.e.*, PIV measurement was made with a 0-V dc offset [127]). To check this point, we repeated the experiment whereby PIV measurements during relaxation were made by maintaining the gate bias at +1.5 V. The results are depicted in Fig. 4.12. In this case, an increase in $|\Delta V_t|$ is evident following the switch of the gate bias back to 0 V. It is also essential to clarify that the resultant increase of $|\Delta V_t|$ after the gate was switched back to 0 V depends on the relaxation time of the trapped electron in relation to the measurement delay. As measurement of the p-MOSFET I_d - V_g curve requires the gate to be pulsed to the negative polarity, there

would be an inevitable loss of trapped electrons depending on the measurement delay. A reduced delay, crucial for the direct tunneling gate oxide, would help suppress the loss of the trapped electrons. Based on this argument, the quasi-saturation level of $|\Delta V_t|$ at +1.5 V observed experimentally is expected to be always higher than the ideal case for which no trapped electrons is lost during measurement.

Fig. 4.12: Evolution of threshold voltage shift $|\Delta V_t|$ during negative-bias temperature stressing and relaxation. Unlike in Fig. 4.1 where the +1.5 V gate voltage was interrupted and set to 0 V during PIV measurement, the gate voltage in this case was maintained at +1.5 V throughout the relaxation (*i.e.*, PIV measurements were made at a dc offset of +1.5 V). A step-like increase in $|\Delta V_t|$ is evident when the gate voltage was subsequently switched to 0 V (big upward arrow). (b) A magnitude view of the $|\Delta V_t|$ increase at the +1.5 V to 0 V transition. Such a step-like change is not observed in a pristine device.

Now, we turn our attention to the CPC mentioned in ref. [47]. Prior to the alternating positive and negative gate biasing that gave rise to the cyclic behavior, the device was subjected to Fowler-Nordheim injection (FNI) to neutralize the trapped holes. During FNI under a relatively large oxide field, electrons were captured by almost all the hole traps (including those that are relatively deep or ANPC). V_O 's which have undergone minimal structural relaxation would have been annihilated.

Thus, it is reasonable for one to assume that only V_O 's which had undergone substantial structural relaxation, *i.e.*, permanent and relatively deep-level hole traps remained. At the first negative gate bias, electrons trapped at these hole traps would be de-trapped and this might take some time especially if they were situated at some distance from the interface. However, once these electrons were de-trapped, some of the relatively deep hole traps might not be re-neutralized again during the following positive gate biasing phase. This was because the moderate positive gate bias only allowed hole traps within a limited energy range to be re-occupied by electrons. This gave rise to an asymmetric first cycle. The subsequent switching of the gate alternately between the same negative and positive polarities resulted in the repetitive de-trapping and trapping, respectively, of electrons at these relatively permanent hole traps, resulting in a cyclic behavior. Since the gate oxide used was thick (5.5-7 nm), the spatial distribution of the hole traps from the interface also influenced the logarithmic time dependence of electron trapping/detrapping observed, apart from the energy distribution of the hole traps [47].

The cyclic behavior implied that it was the same group of V_O 's that consistently responded to the given alternating negative/positive electrical stimuli. A cyclic behavior of the $|\Delta V_t|$ fluctuations under dynamic NBTI was also observed in our recent studies [49]. The important point here is that the total $|\Delta V_t|$ fluctuation included the fast recovery component that occurs during stress-to-relaxation transition. Moreover, such a cyclic behavior is observed regardless whether the gate voltage is at 0 V or a positive voltage during relaxation. In the former, only relatively shallow hole traps are annihilated whereas both shallow and deep hole traps are removed in the latter (Fig. 4.1). These observations suggest that both the shallow and deep hole traps, corresponding to as-grown hole traps and ANPC in the context of ref. [47], have

similar origins. This inference agrees with the result of ref. [134], in which the interaction between trapped holes and hydrogen was studied. The increase in ANPC after exposure to hydrogen corresponded to the decrease in the amount of as-grown hole traps.

Our simulation clearly shows that V_O 's with a neighboring twofold coordinated nitrogen atom is one type of defect precursors to avoid since they exhibit very deep CTLs. Therefore, these V_O 's would charge up positively even under moderate negative gate voltage and the resultant trapped-hole site would exhibit more significant structural relaxation in a given time compared to the case when the neighboring nitrogen atom is threefold coordinated or is absent. As a consequence, the trapped-hole sites would be permanent, severely impacting the NBTI reliability. Passivating the nitrogen dangling bond with hydrogen (resulting in a three-fold coordinated nitrogen) is shown to significantly lower the CTL, and it appears to be a viable means for suppressing DLHT generation. Zhao and Zhang showed that hydrogen could also result in relatively deep hole traps. It should, however, be emphasized that in ref. [134], the increase in the ANPC only occurred when a device with a substantial amount of trapped holes was exposed to hydrogen. Hydrogen was found to have no observable effect on a pristine device or in a stressed device where the trapped holes were first neutralized by electron injection. This observation suggests that the cracking of hydrogen by trapped hole sites, which leads to the formation of H^+ , is responsible for the increased ANPC. *Ab-initio* simulation has also shown that the positive trap states related to H^+ are relatively deep [135]. With these considerations, hydrogen annealing of as-processed devices, which are expected to contain a low density trapped holes, may remain a viable means for suppressing DLHT generation related to nitrogen.

4.1.6 Conclusion

The issue of DLHT generation under negative-bias temperature stressing is examined in detail both via experiment and simulation. A correlation between DLHT generation and the nitrogen concentration in the gate oxide is demonstrated experimentally. Via first-principles simulation, we show that an oxygen vacancy defect with a neighboring twofold coordinated nitrogen atom is a likely candidate for the DLHT. This defect is found to exhibit very significant structural relaxation upon the capture of a hole, and therefore the resultant positively charged state is relatively stable (corresponding to a deep CTL). On the other hand, the properties of an oxygen vacancy defect with a neighboring threefold coordinated nitrogen atom are shown to be similar to that for the pure SiO_2 (*i.e.*, much shallower CTLs ($\sim 2.5 - 3$ eV) and reduced structural relaxation). These defects are mostly responsible for the fast transients typically observed in dynamic NBTI and they are normally triggered under more negative gate biasing and may not play a significant role under nominal gate voltage condition. Based on the new insights, a unified explanation for hole-trap generation and relaxation is proposed and is shown to be able to account for the different electrical responses of positive oxide charges reported in recent literature.

4.2 Structure, Property and Dynamics of K_N Center in SiON Dielectrics

4.2.1 Introduction

Defects with different atomic configuration in SiON show distinguished behaviors during NBT stress. Early measurements reveal a broad energy distribution of stress induced positive trap states across over the band gap, meanwhile a unique relaxation characteristic of deep-level hole traps, as compared to those of the interface states, is observed [25, 57, 123, 136], however the exact physical mechanism is still unavailable. In this chapter, based on molecular modeling and calculation, we will give a comprehensive picture of the properties of nitrogen related defects by utilizing the first-principles simulation. Results illustrate that nitrogen related defects can have various behaviors and consequently generate a wide range of defect energy levels throughout the band gap.

4.2.2 First-principles Modeling and Dynamics of K_N Center in SiON Dielectric

Negative bias temperature instability (NBTI) of the ultrathin oxynitride gate p-MOSFET is widely recognized as a critical reliability issue for the current CMOS integrated circuits. Previously, an objective and systematic analysis of DC current-voltage data reveals a broad energy distribution of stress induced positive trap states, and meanwhile a distinct relaxation characteristic of deep-level hole trap (DLHT) states, as compared to those of the interface states, is observed [25]. Both aforementioned evidences imply the fundamental difference in the behaviors of the defects in SiON as compared with those in the pure SiO₂. To understand the physical morphor of these trap centers, Campbell *et al.* have studied on the Si oxynitride through SDT and SDR. Their results show that new defects, instead of P_b center in

pure SiO₂, with completely different physical characters, are generated and dominate in Si oxynitrides [61]. Considering the atomic structure of these defects, analysis and comparison on the g value and hyperfine G spectrum of Si₃N₄ and SiON help conclude that the defects in SiON are recognized as the Si atom in the oxygen vacancy that back bond to N atoms instead of O. These kinds of new defect structures in SiON are named as K_N center. However, up to now, knowledge on how this nitrogen related defect gives new observations in the device characterization remains unclear.

Guided by the above mentioned issues, in this chapter, we will focus on the physical mechanism of the NBTI degradation. Atomic level models are proposed and studied by utilizing the first-principles calculations. A comprehensive conclusion to illustrate the role of nitrogen in the device degradation will be given at the end of the discussion.

An amorphous SiO₂ super-cell which contains 64 O atoms and 32 Si atoms is adopted as the skeleton to generate the SiON. Since there are 64 non-identical oxygen sites, therefore we manually remove one oxygen atom away to create an oxygen vacancy (V_O) defect case, and totally 64 possible oxygen vacancy cases are available. Moreover, to nitride the SiO₂, one, two or three nitrogen atoms (nitrogen atomic concentration: 1 at. %, 2 at. % and 3.1 at. %) will replace the oxygen that attached to the Si involving the oxygen vacancy, respectively, and five categories of SiON are generated, as shown in Fig. 4.13. After that, all categories are subjected to VASP to have a full structure relaxation under neutral (0) and positive condition (+1) until reach the minimum system energy [137]. Finally the atom position and system energy are extracted. To have a quantitative description of the structure relaxation, the Mean Square Displacement (MSD) of each atom is extracted according to Eq. (4.2)

$$MSD = \left\langle \left| r_i(+1) - r_i(0) - [r_{cm}(+1) - r_{cm}(0)] \right|^2 \right\rangle \dots \text{Eq.4.2}$$

Where $r_i(0)/r_i(+1)$ is the position of atom i at neutral/positively charged condition. $r_{cm}(0)/r_{cm}(+1)$ represents the position of the center of mass under neutral/positively charged situation.

Fig. 4.13: Pure SiO₂; (b) Si-O-N with one N attached to the Si(1); (c) Si-O-2N(i) with two N atoms attached to the same Si(1); (d) Si-O-2N with two N atoms attached to Si(1) and Si(2) separately (ii);(e) Si-O-3N(i) with three N atoms attached to the same Silicon;

4.2.3 Results and Discussion

We take the sum of each atom's Mean Square Displacement (sMSD) in the respective defect case to investigate the structure relaxation. As illustrated in Fig. 4.14, we can see that sMSD are significantly enhanced in the nitride samples comparing with those in pure SiO_2 , which strongly implies that SiON dielectric is more sensitive-

Fig.4.14: sMSD results for (a) Pure SiO_2 ; (b) Si-O-N ; (c) Si-O-2N(a) ; (d) Si-O-2N(b) ; (e) Si-O-3N(a) ; (f) Si-O-3N(b)

to positive bias. Atom in the material tend to overcome the structure barrier and tune their positions to compensate the energy increase due to the hole injection. As a result, the dielectric property will be vulnerably shifted during device operation. Looking into

the structure relaxation within the same categories (*e.g.*, fig. 4.14(b)), there is remarkable fluctuation of sMSD between each cases. Since our models are generated based on the amorphous structure, thus the discrepancies actually are expected and ascribed to the difference in the local atomic structures. For certain defects cases with surprisingly large sMSD (see fig. 4.15(a)), the capturing of the hole has induced intensive electron-phonon and phonon-lattice interaction, causing significant structure change. Structure analysis indicates that due to the lone electron interaction between Si(1) and N, the distance between Si (1) and Si (2) increases from 1.62 Å to 4.59 Å, and eventually brings the system to the minimum energy level. In addition, we find-

Fig. 4.15: V_O structure relaxation cases (a) significant structure change after the hole trapping; (b) mild structure change after the hole trapping.

that after the structure relaxation, the initial neutral vacancy structure is completely destroyed and extra simulation work also shows that they never can recover back to

the initial structure any more even be neutralized again. Meanwhile, our amorphous models also show cases that possesses quite mild sMSD under positive condition in Fig. 4.15(b). In these cases, the phonon and lattice interaction is almost shielded, the distance between Si (1) and Si (2) only increase a bit, from 2.44 Å to 2.52 Å. Due to the insignificant structure in these cases, once the structures gets the electron back, they are supposed to recover back to initial structure, which is validated by the simulation.

Besides, the Charge Transition Levels (CTLs), extracted from the total system energy, are reviewed, as shown in the Fig. 4.16, the band gap in the plots are corrected

Fig.4.16: Extracted CTLs for (a) Pure SiO₂; (b) Si-O-N; (c) Si-O-2N(i); (d) Si-O-2N(ii); (e) Si-O-3N(i); (f) Si-O-3N(ii)

by the hybrid scheme. We can see the results in the figures that comparing to the CTLs map of SiO_2 , the defect levels shown in the SiON are much more scattering than those in the pure SiO_2 . The wide range of the CTLs values manifests the property of nitrogen related defects can be quite diversity, covering from shallow to extremely deep traps.

At this point, it is worth to view the structure relaxation and CTLs results together, to obtain an overall understanding of the defect behaviors in SiON under positive condition. As mentioned before, due to the different local atomic configuration and nitrogen concentration, the sMSD and the corresponding CTL varies a lot between each case. Those have permanent structure change (large sMSD) with deep CTLs are recognized as unrecoverable deep level hole traps. First because of the deep level of CTLs impede defect to capture the electron, neutralizing the positive charge. Second, the large sMSD prevents the structure recovers back to the previous state, even it is neutralized or negatively charged, which means the defect structure will be permanent, and can act as the hole trapping center again once the bias is applied. On the other hand, for the part with great structure change but with shallow CTLs, although the defect structures are permanent, however they are responsible for the fast recoverable component, since the trapping and de-trapping rate can be very fast due to the shallow CTLs. Whereas worth to mention, we also observe some cases with slight structure change and shallow CTLs upon the capturing of hole, this group of oxygen vacancy cases will also superpose their role on the fast trap centers, more detailed work is need to understand the difference between these two shallow trap centers. Finally, the rest set of cases are those with mild structure relaxation but with deep CTLs. In these cases, the oxygen vacancy tends to recover once get the electron back, are considered to be recoverable deep level hole traps. However, due to the

relative high level of the CTLs, the chance of capturing the electron is very small. Thus under normal recovery condition (*e.g.*, recover @ 0V), they will superpose their trapping effects on the unrecoverable deep level hole traps, shown as quasi-permanent degradation. But, if applying a negatively recover voltage to the gate, these defects will have higher chances to capture the electrons, and recover back to the defect-free states.

In all, based on the above discussion, we can conclude that the incorporation of N apparently expands the defect levels to a wider range, which perfectly reproduce the experiment results. Meanwhile, the increase in structure relaxation raises both shallow and deep level traps simultaneously, and no other than the high concentration of deep level traps that generate the remarkable spectrum in the SDR/SDT.

4.2.3 Conclusion

In summary, we have inspected a large number of defect cases with different local atomic structures. Detailed results manifest that nitride SiO₂ are inherently more sensitive to hole injection and tend to degrade more easily with increase in both the shallow and deep defect level.

Chapter 5 Property and Dynamics of Oxide Defects in High- κ Gate Stack

5.1 The Evolution of Switching Oxide Traps in the High- κ Gate Stack under NBTI and PBTI

5.1.1 Introduction

Bias-temperature instability (BTI) has been a very hot topic since the beginning of the last decade due to its serious impact on the long-term device reliability as the devices are scaled down [8, 9, 24-26, 123]. For the conventional silicon dioxide (SiO_2) or silicon oxynitride (SiON) gate dielectric, nearly all studies were focused on the negative BTI (NBTI) since the impact of the positive BTI (PBTI) was found to be negligible [26]. With the scaling continues, the SiO_2 or SiON dielectric has become too thin and caused excessive gate leakage current as a result of the direct tunneling [138]. Therefore, high- κ /metal gate (HKMG) stacks are introduced as a replacement for the conventional SiO_2 / SiON dielectric in sub-65 nm technology nodes [139]. However, high- κ devices suffers much more severe reliability issues, due to the far more defective interface between the high- κ dielectric and Si as compared to that between the SiO_2 and Si, as well as the inherently higher density of as-grown oxide defects [14, 81, 82, 88]. Preliminary measurement results show that the oxide

defects in high- κ would act as effective trapping centers for both electrons and holes, leading to threshold voltage shift and instability [52, 88]. Thus, unlike the SiO₂/SiON gate dielectric, both the PBTI and NBTI become the reliability concerns in the HKMG stack [27, 140, 141]. Moreover, in the DPBTI/DNBTI study of the high- κ /metal gate transistor, the experimental results exhibit that the oxide defects process a cyclical charging/discharging behavior under moderate stress conditions, and they are rendered more permanently charged when the stress level (either voltage or temperature) is increased, which is similar as we observed in the SiON [49, 142, 143], except that the reduction of the recoverable component in the high- κ material is more sensitive to the electric field/temperature. The above observations imply that the physical structure of oxide defects may be modified when the applied stress exceeds a certain threshold, thus making them more permanent. Therefore, in this chapter, we refer to oxide defects that give rise to threshold voltage fluctuations under BTI testing as switching oxide traps (SOTs) to reflect their transition between the charged and discharged states, and the properties of these SOTs that account for the transformation will be discussed in detail [18, 128, 144].

5.1.2 Experiment and Simulation Details

The experimental results presented in this chapter were obtained on the following test devices: 1) p-MOSFETs with HfSiON/TiN and HfO₂/TiN gate stacks. The respective thickness of the HfSiON and HfO₂ are 2 and 4 nm; the former was deposited using metal-organic chemical vapor deposition and the latter by atomic layer deposition (ALD). Both devices have a 1 nm chemical oxide, prepared under the same condition before high- κ deposition. 2) p- and n-MOSFETs with HfO₂/TiN gate stack. The HfO₂ is 3-nm thick, achieved via ALD. The interfacial layer is a 0.9-nm SiO_x

prepared via rapid thermal oxidation. The EOT of all HKMG stacks is ~ 1.4 nm. Dynamic BTI test comprising alternate stress and relaxation intervals were applied on the test devices. During the stress phase, the applied gate stress voltage was calculated based on the desired average oxide stress field across the gate stack, defined as $(V_g - V_{t0})/\text{EOT}$ (V_{t0} is the threshold voltage of the pre-stress device). During the relaxation phase, the gate voltage was set to either 0 V or at a lower value but an opposite polarity to that of the stress voltage (1 V for NBTI and -1 V for PBTI). For large area test devices, degradation was monitored in terms of threshold voltage shift (ΔV_t), extracted from linear transfer curves measured using the ultra-fast switching method [127]. The very short delay of 60 ns for each drain current measurement enables more accurate probing of the fast BTI recovery than the conventional dc sweep method. The low duty cycle (0.1%) of the gate measurement pulse facilitates measurement of the entire transfer curve without any cumulative recovery effect [145]. However, a current measurement resolution of ~ 5 μA limits its use on small area devices. Degradation of small area test devices was monitored in terms of linear drain current degradation (ΔI_d) measured (at V_{dd}) by an Agilent B1500A parameter analyzer at a delay of ~ 0.75 μs .

For the simulation, all calculations were performed by using VASP which employs the plane wave pseudo-potential methods within the density functional theory [137]. A hybrid density functional scheme was applied to obtain the correct HfO_2 band gap [117]. The ultra-soft pseudo-potentials were used to represent the interactions between the ion core and the valence electrons [146]. The exchange correlation functional was treated within the GGA of Perdew, Burke and Ernzerhof [133]. In all the calculations, the cut-off energy and κ -point were tested to make sure all the calculated results were converged. For the structure optimization, the conjugate

gradient method was used and the ion positions were optimized until the residual force was less than 0.01 eV/Å.

The supercell used in our simulation study was amorphous HfO₂, consisting of 32 Hf atoms and 64 O atoms, with a density of 9.68 g/cm³. The amorphous structure was generated from a cubic HfO₂ structure, via molecular dynamics, following the melt-and- quench scheme [147]. The formation energy E^f of a defect in the HfO₂ was calculated from the total energy E of the defective supercell according to the following formula:

$$E^f(\alpha, q) = E(\alpha, q) - (E_0^0 + n_{\text{Hf}} \mu_{\text{Hf}} + n_{\text{O}} \mu_{\text{O}}) + q(\varepsilon_{\text{F}} + E_{\text{VBM}}) \quad \text{--- Eq. 5.1}$$

where E_0^0 is system energy of the defect-free supercell and E_{VBM} is the valence band maximum of HfO₂. For a defect α in charge state q , E^f is a function of the Fermi level ε_{F} and the respective chemical potential of Hf, and O denoted by μ_{Hf} and μ_{O} respectively. The terms n_{Hf} and n_{O} represent the corresponding number of Hf and O atom(s) added/removed from the supercell to form the defect. The charge transition level (CTL) for neutral-to-negative (0/−) and neutral-to-positive (0/+) state transition of the defect were also calculated and are given by the ε_{F} which corresponds to $E^f(\alpha, q) = 0$. The 0/− and 0/+ CTL, measured with respect to E_{VBM} , are akin to trap levels of the defect in the HfO₂ band-gap [113].

5.1.3 The Evolution of Hole Trapping in the High- κ gate p-MOSFET

Fig. 5.1 shows the zoom-in view of certain NBTI cycles and the definition of the recoverable (R) component for each DNBTI cycle. R is defined as the difference between ΔV_{t}

at the end of each stress phase and at the end of the following relaxation phase. Since NBTI recovery is due to hole detrapping, the R of each relaxation phase is a measure

of the discharge of SOTs that were charged up (filled with holes) in preceding stress phase. Component P denotes the part of ΔV_t remaining and is thus a measure of SOTs which did not successfully discharge at the end of each relaxation phase. Similar definitions are applied for the case of DPBTI.

Referring to our experimental results in the Fig. 5.2, first a salient difference should be noted. For a given oxide stress field or temperature, the extent of hole-trap transformation clearly varies with the type of gate dielectric. We can see that: (i) for the HfSiON device, at a relatively low oxide stress field (-7 MV/cm or -8.5 MV/cm),

Figure 5.1: Evolution of the threshold voltage of an p-MOSFET under certain dynamic NBTI cycles. R denotes the recoverable component of a specific cycle and is defined as the difference of the threshold voltage shift at the end of the stress phase and at the end of the following relaxation phase. P denotes the remnant ΔV_t and is a measure of the part of the SOTs which did not discharge at the end of each relaxation phase. Similar definitions apply for the case of DPBTI.

R is constant, independent of the number of DNBTI cycles, however, for the case of HfO_2 , significant reduction of the R happens at -7 MV/cm (14.9 mV) or -8.5 MV/cm (18.5 mV), respectively; (ii) when the oxide stress field is increased to -10 MV/cm and

above, slight reduction of R happens for the HfSiON device with the number of DBTI cycles (8 mV), whereas, much greater R decreases is found for the HfO_2 gated device-

Fig. 5.2: Evolution of NBTI recovery per relaxation interval of a (a) HfO_2 gate; (b) HfSiON gate p-MOSFET subjected to stress/relaxation cycling at different oxide stress fields. The oxide stress field is calculated based on the equivalent oxide thickness of the gate stack. The exact oxide stress field in the 4-nm thick HfO_2 is smaller than that in the 2-nm thick HfSiON .

(20.8 mV). Other than that, these two devices are subjected to different temperatures measurement for the R component study. As we can find in the Fig. 5.3, HfSiON is more immune to the high temperature, the R keeps constant at 100 °C, and the reduction of R is observed at 150 °C. However, for the HfO_2 gated device, the R component is constant at 30 °C, however, it starts to decrease at 100 °C, which is lower than the situation in the HfSiON gated device. These evidences mean that a fraction of the trapped holes is rendered more permanent as the DNBTI experiment progresses, and Fig. 5.4 supports the inference. For the HfO_2 gated device, plotted in the Fig. 5.4(a), we can see that $|\Delta V_t|^{\text{cor}}$ increases at a faster rate, relative to $|\Delta V_t|^{\text{eos}}$. On the other hand, $|\Delta V_t|^{\text{eos}}$ and $|\Delta V_t|^{\text{cor}}$ both increase at the same rate when R is constant. shows that the decrease is a consequence of $|\Delta V_t|^{\text{cor}}$ increasing at a faster rate than the $|\Delta V_t|^{\text{eos}}$.

Fig. 5.3: Evolution of NBTI recovery (denoted as R) per relaxation interval of a (a) HfO_2 gate; (c) HfSiON gate p-MOSFET subjected to stress/relaxation cycling at different temperatures. (b) and (d) show the evolution of the corresponding R , normalized with respect to that for the first relaxation interval.

Similar situation is shown in the HfSiON , the only difference between the two devices is that a higher temperature is needed for the HfSiON device to trigger the decrease of R . In addition, the SILC measurement results under different DNBTI conditions before stress and after 30 cycles DNBTI cycles are shown in the Fig. 5.5. In the figure, it shows that for both cases, no apparent SILC can be observed when R is constant, in spite of the non-negligible increase in the relatively permanent V_t shift or $|\Delta V_t|^{\text{cor}}$. The absence of SILC means no bulk trap generation despite the substantial cumulative stress time (3×10^4 s). NBTI is generally explained in terms of the generation of interface traps and the trapping of holes at oxide defects [20]. Given that the density of as-grown defects is unchanged when R is constant and there is no bulk trap generation

Fig. 5.4: Increase of threshold voltage shift, $|\Delta V_t|$ measured at the end of stress (eos) and end of relaxation interval of each dynamic NBTI cycle, with reference to the respective $|\Delta V_t|^{eos}$ and $|\Delta V_t|^{eor}$ values of the first cycle. The (a) HfO₂ and (b) HfSiON gate p-MOSFETs are the same as those in Fig. 5.3.

Fig. 5.5: Gate current I_g versus gate voltage V_g curves for the same (a) HfO₂ and (b) HfSiON gate p-MOSFET in Fig. 5.2, before and after dynamic NBTI stressing which yields a (i) constant; (ii) decreasing recovery.

based on the SILC measurement, it also implies that the substantial increase of $|\Delta V_t|^{\text{cor}}$ is related to interface degradation and that interface defect generation is uncorrelated to bulk trap generation. On the other hand, more than an order increase in the gate leakage current is observed when R is decreased. The substantial SILC points to bulk trap generation. In addition, the result implies that bulk trap generation is related to the decrease in R or the transformation of trapped hole sites into a more permanent form. Furthermore, for the HfO_2/TiN p-MOSFET, the devices are also subjected to the positive gate bias relaxation after the DNBTI under different stress fields, which is supposed to inject the electrons into the dielectric layer to neutralize the defects. As illustrated in the Fig. 5.6, significant impact on the SILC in I_g is observed that a -

Fig. 5.6: Gate current I_g versus gate voltage V_g curves of a HfO_2/TiN p-MOSFET showing an increased current only when there is a decrease of $|\Delta V_t|$ recovery during dynamic negative bias temperature stressing. I_g was measured after the 30th relaxation interval (Fig. 5.1). The oxide stress fields are as indicated: (a) 5.5 MV/cm; (b) 7 MV/cm; (c) 8.5 MV/cm.

fraction of the SILC is recovered after the positive bias recovery, which confirms the role of P component in inducing the SILC (see fig. 5.12).

At this stage, it is worth to think that the existence of different ‘threshold’ oxide fields and temperatures beyond which recovery decrease begins to happen for the different gate dielectrics [148]. Since the only difference of the two devices is the gate dielectric material, therefore, we can conclude that that the underlying mechanism is linked to the nature of the gate dielectric. A possible explanation lies in the different ionicity of the gate dielectrics. *Ab initio* simulation has shown that more than 90% of the V_O defects (a major source of hole trapping sites) in the SiO_2 undergo marginal structural relaxation upon the capture of a hole [69, 149, 150]. In fact, the increase in the Si-Si distance is so marginal that the remaining unpaired spin is still equally shared between the two Si atoms [69]. This enables the broken bond to fully reform upon the recapture of an electronic charge (after stress is terminated), thereby annihilating the trapped-hole site. While a greater structural relaxation would render the trapped-hole site more permanent (and thus contributing to a more permanent $|\Delta V_t|$), V_O defects which exhibit such a characteristic constitute a minor fraction of the total V_O defect density in the SiO_2 [149] and therefore their effect may not be significant. Structural relaxation would only become more severe, hence recovery decrease become more apparent, at much higher oxide fields and temperatures which increases the distortion of the oxide network. In the more ionic HfO_2 dielectric, however, the inherently stronger electron-lattice interaction results in more substantial structural relaxation of V_O defects under a given oxide stress field and temperature. This may explain the larger recovery reduction, *i.e.*, greater tendency for permanent trapped-hole formation at a given stress condition and the lower threshold for hole-trap transformation in the HfO_2 as compared to HfSiON and SiON . The faster increase of the permanent $|\Delta V_t|$ as

a consequence of this relative ease of hole-trap transformation may limit the long-term NBTI reliability of HfO_2 gate p-MOSFETs [56, 151].

5.1.4 The Evolution of Electron Trapping in the High- κ gate n-MOSFET

Owing to its impact on n-MOSFET performance, the PBTI effect in the HKMG stack has also attracted considerable attention in recent years (*e.g.*, [19, 28, 53]). But unlike NBTI, interface degradation is a minor issue for PBTI, which arises almost entirely out of electron trapping in high- κ oxide defects [53]. Electron trapping in the HKMG stack under PBTI stress is similar to hole trapping under NBTI stress in many aspects: 1) stress induced ΔV_t rises rapidly in the initial stage, followed by an asymptotic increase, implying the existence of both fast and slow electron trapping [28]; 2) post-stress ΔV_t relaxation spans many orders of time magnitude, indicating the presence of both shallow (fast detrapping) and deep (slow detrapping) electron traps [89]; 3) ΔV_t recovery decreases gradually under stress/relaxation cycling when the applied stress exceeds a certain threshold, pointing to the existence of transient (shallow) to permanent (deep) electron-trap transformation [18, 27].

As in the case of deep hole traps [24, 25], the existence of deep electron traps may be revealed by changing the polarity of V_g midway through post-stress relaxation, as shown in Fig. 5.7. After ΔV_t has reached a ‘quasi-saturation state’ under $V_g = 0$ V, indicating a significant slowing down in the recovery rate, a change in V_g to -1 V is seen to trigger another abrupt ΔV_t reduction. This gate-polarity-dependent ΔV_t recovery implies that a fraction of the trapped electrons are situated at relatively deep states. A long time is needed for an electron at a deep trap to be emitted under $V_g = 0$ V, due to the barrier between the trap energy level and the Si conduction band edge E_C (Fig. 5.1.4.2). However a negative V_g can drastically reduce the emission time, by

Fig. 5.7: Threshold voltage shift, ΔV_t of a HfO_2/TiN gate n-MOSFET subjected to positive bias temperature stress (at +1.8 V, 100 °C), followed by relaxation at 0 V and then –1 V gate bias. The second abrupt decrease of ΔV_t may be attributed to the emission of electrons trapped at energetically deeper oxide traps (Fig. 5.8).

Fig. 5.8: Schematic energy band diagram of the HfO_2/TiN gate stack under (a) post-recovering at 0 V and (b) –1 V gate bias. Under 0 V gate bias, an electron trapped at an energetically deep oxide trap could not be readily emitted due to the energy barrier ϕ between the trap level and the Si conduction band edge. A negative gate bias ‘raises’ the trap level (thus reducing the barrier ϕ), allowing the trapped electron to be now easily emitted to the Si conduction band.

lowering the Fermi level to cut down the electron trapping barrier or even introducing a hole to the defect structure to ease the structure recovery and subsequently lowering the electron trapping barrier. The remnant ΔV_t is ~ 0 , implying that the impact of interface trap generation is far less than electron

trapping in the HfO_2 [53], as also revealed via the post-stress transconductance curve, which is laterally shifted toward more positive V_g with respect to the pre-stress curve but shows no degradation of the peak transconductance (not shown). The lack of degradation of the SiO_x interfacial layer in the HKMG stack under PBTI stressing is in good agreement with the observation of negligible PBTI effect in the SiO_2 gate n-MOSFET.

Evidence for the transformation of shallow to deep electron trapping is given in Fig. 5.9. The ΔV_t recovery per relaxation interval under $V_g = 0$ V can be regarded as a measure of electron emission from a spectrum of relatively shallow traps (cf. Fig. 5.7). When the oxide stress field is 5.5 MV/cm, ΔV_t recovery can be seen to remain approximately constant, independent of the number of stress/relax cycles (Fig. 5.9(a)),

Fig. 5.9: Recovery of threshold voltage shift, ΔV_t per relaxation interval as a function of the number of positive bias temperature stress and relaxation cycles applied on a HfO_2/TiN gate n-MOSFET. The oxide stress fields are as indicated. The decrease of ΔV_t recovery in the case of 7 and 8.5 MV/cm stress is clearly suppressed under a negative gate relaxation voltage.

Fig. 5.10: Change of ΔV_t^{eos} and ΔV_t^{eor} (with respect to corresponding values of the first stress/relax cycle) during PBTI stress/relax cycling under different oxide stress fields. A much faster increase of ΔV_t^{eor} as compared to ΔV_t^{eos} is evident at the 7 MV/cm stress. The HfO_2/TiN n-MOSFETs are the same as those in Fig. 5.9.

implying that it is always a similar group of electron traps that are charged and discharged. However, when a higher oxide stress field (7 MV/cm) is applied, a progressive decrease of ΔV_t recovery is observed (Fig. 5.9(b)). When the oxide field is further increased to 8.5 MV/cm, a greater but also gradual decrease of ΔV_t recovery is obtained (Fig. 5.9(c)). Fig. 5.10 shows that the reduction of ΔV_t recovery is due to ΔV_t^{eor} increasing at a faster rate as compared to ΔV_t^{eos} , indicating that a fraction of the previously recoverable electron trapping is rendered more permanent after each stress interval, similar to the case of NBTI stress (Fig. 5.4). The increased emission time of the electron traps may be ascribed to a shift of the trap states to deeper levels, thus increasing the barrier for electron emission (Fig. 5.8).

However, two important differences between electron and hole trapping evolution should be pointed out. First, the transformation of shallow to deep electron trapping observed under $V_g = 0$ V relaxation can be reversed using an opposite (*i.e.*, negative) gate polarity. When relaxation is performed under $V_g = -1$ V, the ΔV_t

recovery is clearly increased (Fig. 5.9) because electrons trapped at deeper states could also be emitted, in addition to those emitted from shallow trap states under $V_g = 0$ V. More importantly, the gradual decrease in ΔV_t recovery, seen under $V_g = 0$ V relaxation, is no longer observed under $V_g = -1$ V relaxation. A constant ΔV_t recovery is obtained instead. Since a negative V_g could detrap electrons trapped at deep oxide traps (Fig. 5.8), almost the entire spectrum of oxide traps is sensed, regardless if some-

Fig. 5.11: PBTi stressed induced shift of shallow to deep electron trapping ‘deactivates’ the oxide traps involved, resulting in reduced cyclical ΔV_t fluctuation in subsequent stress/relax cycling. The application of an opposite (negative) gate relaxation polarity could ‘reactivate’ most of the traps. The sequence of tests applied on a HfO_2/TiN gate n-MOSFET is as follows: Phase I: The n-MOSFET was subjected to PBTi stress (@ 5.5 MV/cm) and relaxation (@ 0 V) for 5 cycles and the ΔV_t recovery for each cycle noted; Phase II: The device was then subjected to PBTi stress (@ 7 MV/cm) and relaxation (@ 0 V) for 30 cycles (evolution of ΔV_t recovery in this case is similar to that shown in Fig. 5.9(b)); Phase III: A repeat of phase I gave a smaller ΔV_t recovery (downward arrow); Phase IV: The device was relaxed under -1 V gate bias; Phase V: A repeat of phase I saw a restoration of ΔV_t recovery (upward arrow).

shallow trap states have been transformed into deeper trap states. Further evidence in support of the reversibility of shallow-to-deep electron-trap transformation is given in Fig. 5.11. One would expect that once a shallow trap state has transited to a deeper level, it would be rendered inactive and no longer participate in cyclical electron trapping/detrapping. This is indeed observed in Fig. 5.11. An n-MOSFET was first stressed at 5.5-MV/cm oxide field and relaxed at $V_g = 0$ V repeatedly for five cycles, and the ΔV_t recovery was recorded. The device was then stressed at 7-MV/cm oxide -

Fig. 5.12: The $|\Delta V_t|$ recovery per cycle for dynamic NBTI stress (oxide stress fields are as indicated) is clearly increased under a positive gate relaxation voltage (+1 V), and the result may be ascribed to the discharge of deep-level hole traps. But unlike PBTI (Fig. 5.8), the positive gate relaxation voltage is found to have no apparent impact on the decrease of $|\Delta V_t|$ recovery per cycle observed at comparable oxide stress field (7, 8.5 MV/cm). The decrease of $|\Delta V_t|$ recovery is nearly the same as that obtained under a 0 V gate relaxation voltage, implying that the transformed hole traps are more permanent.

field and relaxed at $V_g = 0$ V for 30 cycles. During this test, the ΔV_t recovery per relaxation interval gradually decreased with the number of stress/relax cycles, similar

to the result shown in Fig. 5.9(b). The test at 5.5-MV/cm oxide field and $V_g = 0$ V relaxation was then repeated. ΔV_t recovery is now smaller (by ~ 15 mV), indicating that some of the shallow oxide traps are no longer active after the stress at 7-MV/cm oxide field. At this stage, the device was subjected to $V_g = -1$ V relaxation for 1×10^4 s. As can be seen from Fig. 5.11, the V_t recovery measured during a subsequent test at 5.5-MV/cm oxide field and $V_g = 0$ V relaxation is restored to almost the level before the stress at 7-MV/cm oxide field was applied. On the other hand, hole-trap transformation under NBTI stress could not be reversed with a positive relaxation V_g . Although $|\Delta V_t|$ recovery is increased under positive V_g relaxation, due to the discharge of deep-level

Fig. 5.13: The impact of PBTI stressing on SILC or bulk oxide trap generation in a HfO_2/TiN gate n-MOSFET. The gate current measurement was performed at the 30th relaxation interval (Fig. 5.12). Unlike the case of NBTI stress (Fig. 5.6), only very marginal SILC generation is observed at the highest oxide stress field of 8.5 MV/cm, indicating the absence of any significant generation of bulk oxide traps over the range of oxide stress field studied.

hole traps [24, 57, 123] no apparent suppression of $|\Delta V_t|$ recovery can be observed (Fig. 5.12). The decrease of $|\Delta V_t|$ recovery is almost the same as in the case of $V_g = 0$ V relaxation. This implies that the transformed hole traps are more permanent in the case of NBTI stress at comparable oxide field.

Second, unlike the case of NBTI whereby hole-trap transformation is directly correlated to SILC generation (Figs. 5.5), electron-trap transformation under PBTI stress is not always accompanied by an increase in SILC. As is evident from Fig. 5.13, no apparent SILC growth can be observed after the stress at 7-MV/cm oxide field. But under this stress condition, a non-negligible decrease in ΔV_t recovery (under $V_g = 0$ V relaxation) by ~ 15 mV is obtained (Fig. 5.9 (b)). It should be noted that SILC generation would eventually also occur under PBTI stressing at increased oxide field or high temperature, as observed in other studies [148, 152].

5.1.5 Atomic Modeling and Possible Oxide Defect Discussion

To shed light on the different impacts of hole-trapping and electron-trapping evolution on bulk oxide trap generation discussed above, possible oxide defect models are explored using *ab-initio* simulation. All calculations were performed by VASP, employing plane wave pseudo-potential methods within the density functional theory [137]. A hybrid density functional scheme was applied to obtain the correct HfO_2 band gap [117]. Ultra-soft pseudo-potentials were used to represent the interactions between the ion core and the valence electrons[146]. The exchange correlation functional was treated within the GGA of Perdew, Burke and Ernzerhof [133]. In all calculations, the cut-off energy and k-point were tested to make sure all the calculated results were converged. As for structure optimization, the conjugate gradient method was used and the ion positions were optimized until the residual force was less than $0.01 \text{ eV/\AA}$. The

supercell used in our simulation study was amorphous HfO_2 , generated through melting and quenching scheme[147], in which consists of 32 Hf atoms and 64 O atoms, with a density of 9.68 g/cm^3 . The amorphous structure was generated from a cubic HfO_2 structure, via molecular dynamics, following the melt-and-quench scheme [153].

A. Oxygen Vacancy (V_O) Defect

Many studies have shown that the oxygen vacancy (V_O) defect is a major source of oxide defects in the HfO_2 (e.g., [154]). Thus, our simulation study will first examine whether the electronic properties of the V_O defect could match, at least broadly, the experimental characteristics of PBTI and NBTI outlined above.

To simulate the formation of the V_O defect, the oxygen atom in-between two Hf atoms is manually removed (Fig. 5.14). The resultant oxygen deficient supercell was then allowed to structurally relax, at 0 K, under the neutral as well as positively and negatively charged states (simulating hole trapping and electron trapping, respectively)

Fig. 5.14: (a) Ball and stick diagram of a portion of the defect-free amorphous HfO_2 supercell used in our simulation study. The supercell consists of 32 Hf atoms and 64 O atoms and was generated by the melt-and-quench scheme; (b) An O atom in-between the two Hf atoms as shown was removed to simulate the oxygen vacancy defect (arrow).

until the respective minimum energy state was achieved. The formation energy E^f of a V_O defect in charge state q was calculated from the system energy E of the defective supercell according to the following:

$$E^f(q) = E(q) - (E_0 + n_O \mu_O) + q(E_F + E_{VBM}) \quad \text{--- Eq.5.2}$$

where E_0 is system energy of the defect-free supercell, E_{VBM} is valence band maximum of HfO_2 , E_F is Fermi level, μ_O is chemical potential of oxygen. The term n_O represents the number of oxygen atom(s) removed from the supercell to form the V_O defect. With the E^f for a V_O defect in the neutral ($q = 0$) and charged state ($q = -1$ or $+1$) determined, the charge transition level (CTL) for neutral-to-negative ($0/-$) and neutral-to-positive ($0/+$) state transitions (corresponding to electron and hole trapping) were then calculated as follows:

$$E^f(q) - E^f(0) = E(q) - E(0) + q(E_F + E_{VBM}) \quad \text{--- Eq.5.3}$$

CTL is defined as the $\mathcal{E}_F$ which corresponds to $E^f(q) - E^f(0) = 0$. The $0/-$ and $0/+$ CTLs, measured with respect to E_{VBM} , are akin to energy levels of the negatively and positively charged V_O defect, respectively, in the HfO_2 band gap [117]. Since the supercell consists of 64 O atoms, a total of 64 different V_O defects in the amorphous supercell were studied ($n_O = 1$ for all cases).

Based on Eq. (5.3), the CTL for the $0/-$ and $0/+$ state transition of each of the 64 V_O defects were calculated and summarized in Fig. 5.15 (CTL = 0 and 5.4 eV correspond to the E_{VBM} and E_{CBM} , respectively of the HfO_2). Also shown in Fig. 5.15 are two dashed lines – the top line denotes the E_{CBM} while the bottom line denotes the E_{VBM} of Si. Band offsets between HfO_2 and Si are as indicated. It can be observed that the CTL for the $0/+$ state transition of all V_O 's are situated way above the E_{VBM} of HfO_2 , within the band gap of

Fig. 5.15: Charge transition level or CTL for the (a) neutral-to-positive (0/+) state transition; (b) neutral-to-negative (0/-) state transition of all 64 cases of oxygen vacancy defects in the amorphous HfO_2 supercell used in our simulation study. The upper and lower dashed lines denote the Si conduction and valence band edge, respectively. The corresponding band offsets are also shown.

Si (the average value is 3.43 eV and the standard deviation is 0.31 eV (Fig. 5.15(a)). Holes trapped at these levels may not be readily emitted under a 0 V gate bias, due to the considerable energy barrier for electron capture, *i.e.*, V_O 's in the HfO_2 could function as deep-level hole traps, locking in the trapped holes. Moreover, the distribution of the trap states in the upper half of the Si band gap makes them ideal TAT centers for electrons tunneling from the metal gate to the Si substrate. Thus, V_O defects in HfO_2 could explain the generation of 1) deep-level trapped holes; 2) conversion of transient hole trapping into more permanent trapped holes (when the energy barrier for structural change is successfully overcome) (Figs. 5.2, 5.3); 3) SILC (Figs. 5.5, 5.6) after NBT stressing.

As for SiO_2 , the average CTL for the 0/+ state transition of V_O defects is ~ 2 eV [155], *i.e.*, the trap states are situated near the E_{VBM} of SiO_2 , making them very shallow hole traps. This is primarily because V_O 's in SiO_2 undergo lesser structural relaxation upon hole capture, as compared to V_O 's in HfO_2 , due to the inherently

stronger lattice-electron coupling in the latter. However, our study has shown that the structural relaxation of a V_O defect in SiO_2 could be significantly increased in the presence of an adjacent nitrogen atom, with the resultant CTL shifted above the Si E_{VBM} [155]. Thus, although V_O 's in SiO_2 are shallow hole traps, their properties are modified by nitrogen in the SiON, changing them into deeper hole traps. This result could explain the increased level of stable hole trapping in NBT stressed SiON gate p-MOSFETs [25, 57, 123].

On the other hand, the CTLs for the $0/-$ state transition (electron trapping) are situated close to the E_{CBM} of the HfO_2 (Fig. 5.15(b)). The average CTL is 4.44 eV and the standard deviation is 0.12 eV. This result shows that V_O 's in the HfO_2 behave as very shallow electron traps. Even if a V_O defect could be transformed into a stable state following electron capture and structural relaxation during PBT stressing, emission of the captured electron would readily occur upon removal of the electrical stress (*i.e.*, electron trapping is always transient). Moreover, the positions of the V_O 's trap levels make them ideal TAT centers under positive gate biasing. However, experimental results clearly show the existence of deep level electron trapping, transformation of shallow to deep electron trapping and absence of SILC generation. Thus, the V_O defect model is unable to explain the PBTI results satisfactorily.

B. Vacancy-Interstitial (V_O - O_i) Defect

In this section, we show evidence (from *ab-initio* simulation) that the oxygen vacancy cum oxygen interstitial, *i.e.*, V_O - O_i paired defect in the HfO_2 could function as both shallow and deep electron and hole traps. This model could account for almost all the experimental features of BTI observed to-date. Electron spin resonance measurement has detected both oxygen vacancies and oxygen interstitials (O_i 's) in the

Hf-based high- κ dielectrics [82, 83]. Since high- κ oxides are generally oxygen deficient, the presence of O_i 's alone is not a likely scenario. However, the generally strong ionic character of high- κ oxides implies that field induced displacement of the relatively light oxygen atoms may happen readily in these materials, especially under bias temperature stressing. The observation of O_i 's from ESR measurement suggests that their presence is likely linked to nearby V_O 's. This is the main motivation for our simulation study on the V_O - O_i defect. One could regard the V_O - O_i defect as an intermediate state prior to the formation of the V_O defect, *i.e.*, the V_O defect would be formed when the O_i has moved away from the V_O site.

To create the V_O - O_i defect, the oxygen atom in-between two hafnium atoms was manually removed and placed randomly in an interstitial position near the vacancy site, as illustrated in Fig. 5.16(a). The defective supercell was then subjected to full structural relaxation under the negatively charged condition. Various O_i positions around a given V_O defect site were examined. For some V_O defect sites, the O_i moves back to its original following the full structural relaxation, with minimal change of the supercell structure. These cases may be considered as corresponding to the fast PBTI recovery, *i.e.*, upon removal of the electrical stress, the trapped electrons are spontaneously emitted and the defects eliminated (reverted to the precursor forms). For a typical case where the O_i remains in an interstitial position following full structural relaxation, the CTL corresponding to the $0/-$ transition is depicted in Fig. 5.16(b) for different O_i positions. For certain positions of O_i , the CTLs are relatively shallow (near the E_{CBM} of Si) while for other positions, the CTLs are clearly much deeper. At a certain O_i position, it can be seen that the CTL is very near the E_{VBM} of HfO_2 , *i.e.*, this defect acts as a very deep electron trap. Thus, it is clear from the simulation results that a V_O - O_i defect can give rise to different trap energy levels in the

Fig. 5.16: (a) Ball and stick diagram illustrating an oxygen vacancy cum oxygen interstitial (V_O - O_i) defect in the amorphous HfO_2 used in our simulation study. The O atom in-between two Hf atoms is manually removed and placed randomly in an interstitial position (O_i) in a nearby region, simulating the creation of the V_O - O_i defect under electrical stress. (b) Schematic energy band diagram of the HfO_2/Si stack showing the charge transition levels for ten arbitrarily chosen O_i positions. Closely spaced CTLs result in a thick line. The CTL for a V_O defect is also shown for comparison.

oxide band gap (*i.e.*, function both as a shallow and deep electron trap), depending on the position of the O_i . Fig. 5.17(a) and (b) show examples of the partial electron distribution function in the vicinity of a negatively charged V_O - O_i defect, for the case of a shallow (CTL = 3.78 eV) and deep (CTL = 1.58 eV) electron trapping center, respectively. The partial electron distribution function is obtained by subtracting the electron distribution function of the defect-free supercell from that for the defective supercell, prior to full structural relaxation. It provides an indication on the location where the additional electron is mostly trapped. For the case of the shallow trap, electron trapping can be seen to be localized at the V_O site whereas it is more delocalized (with electron ‘cloud’ also found on the O_i) for the case of the deep trap.

We reckon that under PBTI stress, the applied oxide stress field, in conjunction with lattice vibrations, would randomly displace the oxygen atoms relative to the hafnium atoms in the ionic HfO_2 dielectric. This would then give rise to a broad distribution of trapping states across the HfO_2 band gap. Besides shallow electron traps responsible for fast PBTI recovery, much deeper traps that ‘locks in’ the trapped electrons for long periods are also formed. Some of the deep trap levels may have resulted from initially shallower trap levels, due to a change in the O_i position. Thus, the V_O - O_i defect model could explain 1) the steady build-up of deep level electron trapping 2) PBTI recovery or trapped-electron emission which lasts many orders of time magnitude in our study (from $\sim 10^{-6}$ s (limited by measurement delay) to $\sim 10^8$ s or longer [89]); 3) the transformation of shallow to deep electron trapping – when barrier for structural change is overcome. It should be noted that deep trap levels are, however, not active TAT centers which explain the absence of SILC generation.

As mentioned earlier, SILC generation is eventually observed for PBTI stressing carried out at higher oxide field or temperature. This may be ascribed to the generation of ‘ V_O -like’ or V_O defects, since a higher oxide field or more intense phonon interaction at a higher temperature could cause the O_i ’s to move farther from the V_O sites. *Ab-initio* simulation has shown that an interstitial oxygen atom could propagate through the HfO_2 network via substitution with lattice oxygen atoms [87]. The activation energy is ~ 0.3 eV, which may be readily fulfilled under more severe stress conditions. V_O defects have shallow electron trap levels and are therefore active TAT centers for the Si conduction band electrons in an n-MOSFET.

Interestingly, the V_O - O_i defect could also function as a hole trap, besides functioning as an electron trap. Fig. 5.18 depicts the partial electron distribution function in the vicinity of reconcile the different observations on SILC generation

Fig. 5.17: (a) Ball and stick diagram of a V_O - O_i defect yielding a shallow electron trap (CTL = 3.78 eV) in the HfO_2 bandgap. The yellow shades denote positive partial electron density obtained after introducing an additional electronic charge into the supercell. The additional electron can be seen to be trapped at the V_O site. (b) A different V_O - O_i defect which yields a deep electron trap (CTL = 1.58 eV). Part of the electron trapping can be seen to occur at the interstitial oxygen O_i .

Fig. 5.18: (a) Ball and stick diagram for the same V_O - O_i defect in Fig. 5.17(a) functioning as a deep hole trap (CTL = 4.26 eV) in the HfO_2 bandgap. The yellow shades denote negative partial electron density obtained after removing an electronic charge from the supercell. The hole is trapped at the V_O site. (b) Ball and stick diagram for the V_O - O_i defect in Fig. 5.17(b) functioning as a shallow hole trap (CTL = 2.56 eV).

the same V_O - O_i defects shown in Fig. 5.17. For the V_O - O_i defect that functions as a deep electron trap (CTL = 1.58 eV) – Fig. 5.17(b), it can also function as a shallow hole trap (CTL = 2.56 eV) – Fig. 5.18(b). The resultant hole is trapped at the V_O site. This result may help -

Fig. 5.19: Another example of a V_O - O_i defect functioning as a shallow hole trap (CTL = 1.85 eV). The yellow shades denote negative partial electron density obtained after removing an electronic charge from the supercell. Unlike the example shown in Fig. 5.18(b), hole trapping in this case occurs at the O_i site.

Fig. 5.20: (a) (i) A deep oxide trap due to the V_O - O_i defect is not an active trap-assisted-tunneling (TAT) center under positive gate bias, since an electron, once trapped, could be easily be emitted. (ii) The same trap could, however, function as an active TAT center, facilitating the tunneling of holes from the Si valence band towards the metal gate under a negative gate bias. (b) A shallower V_O - O_i oxide trap could function as an active TAT center under both gate polarities, similar to the V_O defect shown in (a).

following PBT and NBT stress. While the trap level is not an active TAT center in an n-MOSFET, it can clearly facilitates contributing to an increase of the gate current. Fig. 5.18(b) shows that the V_O - O_i defect which hole tunneling from the Si valence band to the metal in the p-MOSFET (Fig. 5.20), functions as a shallow electron trap in Fig. 5.17(a) (CTL = 3.78 eV) can function as a deep hole (CTL = 4.26 eV), with the hole also trapped at the V_O site. The trap level is an active TAT center for electron tunneling from the metal to the Si conduction band and vice versa. It should be mentioned that hole trapping does not always occur at the V_O site. In another V_O - O_i defect studied, hole trapping is found at the O_i site instead of the V_O site. Being an electronegative atom, it is unexpected that hole trapping should happen preferentially at O_i . More studies are required to understand the reason.

Before summarizing, it should be mentioned that in our simulation studies, we have only considered the change in the system energy of an oxide defect upon charge capture (relative to the neutral state) and calculated the resultant trap energy level based on this change. A deep hole or electron trap could result when the decrease in system energy due to structural relaxation offsets the increase due to charge trapping, causing a shift of the trap energy level (towards the oxide E_{CBM} for the hole trap and towards the oxide E_{VBM} for the electron trap). The defect is assumed to be already present in the oxide. The process of defect creation is not examined. Hence, it is essential to clarify that a deep hole or electron trap does not generally mean the defect concerned is also structurally permanent. To determine whether a defect is structurally permanent, one has to examine the change in system energy between the defect-free (or precursor) state and the defective state (outside the scope of our present study). It may be possible that a deep trap is structurally less stable than a shallower one. The result in Fig. 5.12 supports this thinking; a positive gate relaxation voltage eliminates a

non-negligible part of the deep-level hole trapping but has no apparent impact on the transient to permanent trapped-hole transformation. Another example is found in Figs. 5.9 and 5.11, where reversal of deep electron traps to shallower ones is observed under a moderate negative gate relaxation voltage. In the context of the V_O - O_i defect model, an opposite gate polarity may return some of the O_i 's to their previous positions, resulting in the annihilation of the deep electron traps.

5.1.6 Conclusion

Recent experiment evidences on the evolution of charge trapping in the HfO_2 gate dielectrics, subjected to bias-temperature stressing, are reviewed in this chapter. When a test device is subjected to repeated bias-temperature stress and relaxation under increased oxide stress field or temperature, a gradual decrease of the NBTI or PBTI recovery or recoverable charge trapping can be observed. The decrease is found to be due to the relatively permanent V_t shift component (measured after each relaxation interval) increasing faster than the total V_t shift component (measured at the end of each stress interval). The observation shows that a portion of the recoverable charge trapping is progressively transformed into more permanent trapped charge as the stress/relaxation cycling proceeds. Such transformation is, however, not visible despite after many stress/relaxation cycles under reduced oxide stress field or temperature. The results imply that oxide stress field or temperature exceeding certain thresholds may modify the oxide structure, making it more susceptible to permanent defect generation. Due to the apparent absence of such an effect at reduced stress conditions, caution would be needed in interpreting lifetime extrapolation from accelerated stress conditions.

The relationship between transient to permanent charge trapping and generation

of SILC is also examined. For the case of NBTI, the transformation of transient to permanent hole trapping is always accompanied by a correlated increase in SILC. On the other hand, such correlation is generally absent in the case of PBTI, although SILC eventually also occurs under more severe oxide stress field and temperature. *Ab-initio* simulation shows that while the properties of the V_O defect may satisfactorily account for the NBTI results, the shallow electron trapping level fails to explain the existence of deep electron trapping and the transformation of shallow to deep electron trapping in the case of PBTI. An alternate V_O - O_i paired defect model is shown to be able to reconcile the difference observed between NBTI and PBTI.

5.2 Impact of Lanthanum on Positive-Bias Temperature Instability-Insight from First-principles Simulation

5.2.1 Introduction

It is now recognized that the traditional silicon oxynitride (SiON) gate dielectric has to make way for high-permittivity or high- κ metal oxides in sub-nm CMOS technologies. The main reason is that the aggressive scaling down of the SiON thickness has led to an unacceptably high gate leakage current density. Among the many high- κ oxides studied in the past decade, the hafnium dioxide or HfO_2 has emerged as one of the most suitable candidates and has already been adopted in recent commercial CMOS technologies.

However, there are still several difficult issues surrounding high- κ oxide integration. One of the major challenges is the difficulty in realizing transistors with low threshold voltages or V_t 's. For instance, the V_t of a polysilicon/ HfO_2 gate n-channel transistor was ~ 0.7 V. Fermi level pinning at the polysilicon interface is shown to be one of the main reasons for the high V_t [156]. Although the use of metal gates had helped alleviate the problem, the V_t of HfO_2/TiN gate transistors remained high (~ 0.4 V). Recent studies find that dipoles at the $\text{HfO}_2/\text{SiO}_x$ interface also play a major role in the high V_t obtained and the result is successfully explained by a model proposed by Kita and Toriumi [91]. According to this model, the intrinsically higher oxygen areal density of HfO_2 results in the formation of dipoles (positive end on the HfO_2 side) via the transfer of oxygen ions (O^{2-}) to the underlying SiO_x . As a consequence, the flat band voltage is shifted positively. The model also predicts that some rare-earth oxides such as La_2O_3 , which have lower oxygen areal densities, would form dipoles of opposite orientation at the SiO_x interface. Several experimental studies

have exploited this effect to successfully reduce the V_t of n-channel transistors [157-159].

But the incorporation of lanthanum (La) has raised a concern on gate stack reliability. Up to now, mixed observations have been made. For instance, La doping of the HfO_2 is shown to reduce the positive-bias temperature instability or PBTI effect of the n-MOSFET at a low oxide stress field [95, 160]. However, the PBTI effect is worse under a high oxide stress field. Time-dependent dielectric breakdown or TDDB test reveals a similar observation where it is found that degradation of the La-doped HfO_2 scaled more quickly with oxide stress field as compared to that of the undoped HfO_2 [95]. It is recently found that PBTI recovery of a La-doped HfSiO gate n-MOSFET proceeded at a much slower pace as compared to that of an undoped counterpart, indicating that La doping make the PBTI effect more permanent. To-date, however, it remains unclear why La doping has different impacts in the low and high oxide field regimes and why it retards PBTI recovery. In this chapter, we attempt to provide a plausible explanation via a first-principles simulation study of the effect of La on the electronic properties of the oxygen vacancy (V_O) defect as well as the related vacancy-interstitial ($V_O\text{-}O_i$) paired defect in the HfO_2 .

5.2.2 La Induced Slowing Down of PBTI Recovery

Before describing the simulation method and results, we first present a review of recent experimental results which show that La slows down PBTI recovery. The n-MOSFET studied has a $\text{TiN}/\text{HfSiO}/\text{SiON}$ gate stack (of 1.6 nm equivalent oxide thickness) processed under a gate-first flow. La doping was realized by a thin LaO_x cap layer, atomic-layer deposited and subsequently diffused into the HfSiO via the high temperature source/drain annealing step. To ascertain that the observed slow-down of PBTI recovery is not due to a specific process, comparison is made to a

variety of other devices: 1) undoped p-MOSFETs on the same wafer as the La-doped n-MOSFET; 2) n-MOSFETs with undoped 3-nm HfO₂ and 2-nm HfSiON grown via metal-organic chemical vapor deposition over a 1-nm chemical oxide (the gate electrode was polysilicon and TaN, respectively); 3) n- and p-MOSFETs with undoped HfO₂(3 nm)/SiO_x(0.9 nm) gate stacks; 4) p-MOSFETs with the traditional polysilicon gate and SiON gate dielectric (1.8 and 2.2 nm). The n- and p-MOSFETs were subjected to positive-bias and negative-bias temperature stressing, respectively. The observation that has led us to the finding of La induced slow-down in PBTI recovery is shown in Fig. 5.21(a). Stress induced linear drain current degradation ΔI_d (in percent) of the La-doped n-MOSFET and undoped p-MOSFET (on the same wafer), are compared at different measurement delays. When the delay is short (1 μ s), the ΔI_d of

Fig. 5.21: (a) Linear drain current degradation ΔI_d (in percent), measured at two different delays, of a La-doped n-MOSFET and undoped p-MOSFET (on the same wafer) as a function of stress time. (b) Typical logarithmic time dependence of ΔI_d recovery (expressed as a percentage of the total I_d degradation at the end of stress) of a variety of n- and p-MOSFETs with different gate stacks, subjected to PBTI and NBTI stress, respectively. The number in the bracket refers to the dielectric thickness [29].

the p-MOSFET is greater than that of the n-MOSFET. At a much longer delay (1 s), however, a reversal is observed. Both PBTI and NBTI are known to recover upon the termination of electrical stress and a longer delay means a greater recovery would have occurred before the I_d measurement is made. The reversal observed implies that the degraded La-doped n-MOSFET and undoped p-MOSFET are recovering at very different rates. Fig. 5.21(b) confirms that it is the La-doped n-MOSFET which has an abnormally slow recovery. As shown, the recovery of the undoped p-MOSFET is comparable to a variety of other devices with different gate dielectrics and gate electrodes. On the other hand, the La-doped n-MOSFET exhibits a substantially slower recovery as compared to the rest of the undoped devices [29].

Studies have shown that electron and hole trapping (at oxide defects), respectively, is responsible in part for the PBTI and NBTI effect [29]. Recovery due to electron or hole detrapping typically spans a broad range of timing, from as short as microseconds (limited by the measurement speed of the instrument) to beyond 10^6 s [50, 89]. The classical tunneling-front model, which ascribes the dispersion in the recovery times to the emission of filled traps situated at varying distances from the oxide interface, could not explain such a broad range of recovery time (spanning more than 10 orders of magnitude) measured on very thin gate stacks (~ 3 nm or less in thickness) [23, 129, 161]. The result implies that some of the charges are captured by deep oxide traps with relatively permanent defect structure, whose emission time constants are primarily controlled by the potential barrier that exists between the trap state and the conduction band of the Si substrate (as illustrated schematically in Fig. 5.22) and not by the distance of the trap from the oxide interface. This inference is supported by the dependence of degradation recovery on the gate voltage V_g (Fig. 5.23). As shown in Fig. 5.23(a) for the case of PBTI, the V_t shift recovery under $V_g = 0$

Fig. 5.22: (a) Schematic energy band diagram for the TiN/HfO₂/Si structure showing (i) a detrapping barrier ϕ of a deep electron trap under $V_g = 0$ and (ii) reduction of the barrier by a negative gate bias, which facilitates emission of the trapped electron. (b) Schematic depiction of a deep hole trap, where emission of the trapped hole (via the capture of an electron from the Si substrate) is hindered by the barrier ϕ . Under a positive gate bias, the hole trap is lowered below the Si conduction band and is neutralized by an injecting electron.

However, a switch in V_g to -1 V can be seen to trigger another abrupt V_t shift reduction. This rapid second V_t shift recovery may be explained by a lowering of the electron detrapping barrier, as a result of the “rise” in the trap state brought about by defect structure under the negative V_g (Fig. 5.22(a)(ii)). A similar dependence of NBTI recovery on V_g can be seen in Fig. 5.23(b). On the other hand, recovery of the La-doped n V is seen to hit a “quasi-saturation state” soon after the initial steep reduction, indicating that electron detrapping has slowed down substantially MOSFET exhibits a much weaker dependence on V_g . Only marginal reduction in ΔI_d (relative to the overall degradation) is observed at the various instants of V_g change, as compared to that of the undoped HfO₂. This result suggests that the slow-down in PBTI recovery is due to a higher density of deep-level electron traps in the La-doped gate stack.

Fig. 5.23: (a) Evolution of threshold voltage shift ΔV_t during a cycle of PBTI stress and relaxation of a undoped HfO_2/TiN gate n-MOSFET. (b) Recovery of linear drain current degradation ΔI_d , expressed as a percentage of the total degradation at the end of stressing, for the La-doped n-MOSFET and undoped p-MOSFET under different gate relaxation voltages.

Fig. 5.24: Recovery of oxide defects at and near the SiO_x/Si interface (as probed by the charge pumping method at 1 MHz) of the La-doped n-MOSFET and undoped p-MOSFET following PBTI and NBTI stressing, respectively. ΔI_{cp} is expressed as a percentage of the total ΔI_{cp} at the end of stress [29].

Results from high-frequency (1 MHz) charge pumping current (I_{cp}) measurement (Fig. 5.24), which probes defects close to the oxide/Si interface, imply that the deep-level electron traps responsible for the slow-down in PBTI recovery are mostly contained in the high- κ oxide. As can be seen from the recovery characteristics, the percentage recovery of ΔI_{cp} for the La-doped n-MOSFET is actually larger than that for the undoped p-MOSFET. This result cannot explain the much smaller percentage recovery of ΔI_d for the former. Thus, it is inferred that the observed slow-down of PBTI recovery is mostly a bulk effect, in agreement with previous studies asserting that oxide trapping dominates PBTI [53]. As will be shown below, La's presence lowers the formation energy of the V_O - O_i paired defect, which can function as a deep-level electron trap in the HfO_2 .

5.2.3 First-principles Simulation and Discussion

To provide a theoretical support for the above conclusion drawn from the electrical measurement results, *ab-initio* simulation study on the electronic properties of defects in the La-doped and undoped HfO_2 was carried out. The supercell used is amorphous HfO_2 , comprising 32 Hf atoms and 64 O atoms, with a density of 9.68 g/cm³. The dimensions of the supercell are 10.23 Å × 10.23 Å × 10.23 Å. The amorphous structure was generated from a cubic HfO_2 structure, via molecular dynamics, following the melt-and-quench scheme [147]. The cubic HfO_2 structure was “heated” to a temperature of 2500 K and maintained at that temperature for 3 ps to fully liquefy the system. The temperature was then lowered to 300 K in 1 ps. This was followed by a 3-ps period at 300 K to achieve the thermal equilibrium structure. Finally, structural relaxation of atomic coordinates was carried out at 0 K. Fig. 5.25(a) is a ball and stick diagram of part of the defect-free HfO_2 supercell. La doping was

realized by manually replacing two nearby Hf atoms with La atoms and then allowing the system to relax again to the minimum energy state (Fig. 5.25(b)). All calculations were performed using VASP [137], which employs the plane wave pseudo-potential methods within the density functional theory. A hybrid density functional scheme was applied to obtain the correct HfO_2 band gap [117]. The ultra-soft pseudo-potentials were used to represent the interactions between the ion core and the valence electrons [146]. The exchange correlation functional was treated within the GGA of Perdew Burke and Ernzerhof [133]. In all calculations, the cut-off energy and Γ -point were tested. For structure optimization, the conjugate gradient method was used and the ion positions were optimized until the residual force was less than $0.01 \text{ eV/\AA}$.

Fig. 5.25: (a) Ball and stick diagram of part of the defect-free HfO_2 supercell used in our simulation study. (b) Two Hf atoms are replaced by La atoms to simulate La doping of the HfO_2 . (c) The O atom in between the two Hf atoms was manually removed to create an oxygen vacancy (V_O) defect. (d) The V_O defect is formed in between two La atoms in the La doped HfO_2 .

Oxygen Vacancy Defect

The first part of the simulation study was focused on O, a major source of defects in the high-κ oxide. Charge trapping at O's is generally believed to be responsible for the V_t instability observed experimentally [48]. The impact of La is examined by comparing the formation energy E^f of V_O in the La-doped and undoped HfO_2 . The E^f of a defect may be obtained from the difference in the total energy E of the defective and defect-free supercell according to:

$$E^f(\alpha, q) = E(\alpha, q) - (E_0^0 + n_{\text{Hf}}\mu_{\text{Hf}} + n_{\text{O}}\mu_{\text{O}} + n_{\text{La}}\mu_{\text{La}}) + q(E_F + E_{\text{VBM}}) \quad \text{Eq. 5.4}$$

where E_0^0 is system energy of the defect-free supercell and E_{VBM} is the valence band maximum of HfO_2 . For a defect α in charge state q , E^f is a function of the Fermi level E_F and the respective chemical potential of Hf, O and La, denoted by μ_{Hf} , μ_{O} , and μ_{La} , respectively. The terms n_{Hf} , n_{O} and n_{La} represent the corresponding number of Hf, O and La atom(s) added/removed from the supercell to form the defect. In the undoped HfO_2 , a V_O defect was formed by manually removing the O atom between two Hf atoms (Fig. 5.25(c)). The same method was used to form a V_O defect in the La-doped HfO_2 by removing the O atom between two La atoms (Fig. 5.25(d)). The oxygen-deficient supercell in each case was then relaxed until the minimum energy state was attained. Using this method, the E^f for the V_O defect in the undoped and La-doped HfO_2 are determined as 5.85 eV and -3.78 eV, respectively. Clearly, the presence of La has a significant impact on V_O formation in the HfO_2 . The substantially reduced E^f (negative in fact) implies that La doping increases the likelihood of V_O defect formation in the HfO_2 .

Besides E^f , the charge transition level or CTL for neutral-to-negative state transition of the V_O defect was also calculated to examine its role on electron trapping

under PBTI stress. The CTL is defined as the Fermi energy ε_F that corresponds to $E^f(\alpha, q) = 0$ for both the neutral and negative charged V_O defect. The 0/- CTL measured with respect to E_{VBM} of the HfO_2 is akin to the electron trap level of the V_O defect in the HfO_2 band gap [117]. The 0/- CTL of the V_O defect in the undoped HfO_2 (Fig. 5.25(d)) is determined as 4.75 eV. The results are in agreement with previous studies, indicating that V_O 's in the HfO_2 function give rise to shallow electron traps. It should be mentioned that our calculation did not involve only a single V_O defect but all 64 possible cases in the amorphous supercell (comprising 64 O atoms). No difference (statistically) in the CTL for the La-doped and undoped HfO_2 can be observed; the average CTL is ~ 4.7 eV with a spread of ~ 0.3 eV for both cases. Thus, La doping has no apparent impact on the trap energy level. Schematic energy band diagrams showing the positions of the V_O CTL in the undoped and La-doped HfO_2 are given Fig. 5.26. Formation of the V_O defect in the La-doped HfO_2 was also studied via first-principles simulation by Umezawa *et al.* and Liu and Robertson [96, 162]. However, these

Fig. 5.26: Schematic energy band diagram illustrating the charge transition level (akin to the trap level) of the oxygen vacancy (V_O) and vacancy-interstitial ($V_O - O_i$) defect in the (a) undoped HfO_2 and (b) La-doped HfO_2 .

Fig. 5.27: (a) Ball and stick representation of the amorphous HfO_2 supercell with a V_{O} - O_{i} defect, created by removing one of the O atoms in between two Hf atoms (see labels) and placing it randomly in an interstitial position (O_{i}) close to the V_{O} defect. (b) A similar representation of the V_{O} - O_{i} defect in the La-doped HfO_2 supercell.

studies reported an increase in E^{f} (by ~ 0.7 eV), contrary to the significant reduction obtained in our case. We believe the discrepancy is due to the different ways by which the V_{O} defect is generated. In Umezawa *et al.*, the authors removed altogether two O atoms when creating a V_{O} defect in the La-doped HfO_2 . In replacing two Hf atoms by La atoms to simulate La doping, one O atom was also removed to avoid the formation of a hole at the La substitutional sites, due to the difference in the formal charges of La^{3+} and Hf^{4+} . This was followed by the removal of a second O atom to generate a V_{O} defect in the La-doped HfO_2 . The removal of two O atoms may explain the much higher E^{f} of 6.29 eV (more positive than that of an undoped HfO_2) obtained by the authors [96]. In our case, however, such treatment is deemed not necessary by virtue of the amorphous supercell. The negative E^{f} means that our structure is energetically more favorable as compared to those considered in.

The observation of a negative E^{f} for the V_{O} defect in the La-doped HfO_2 implies that this gate dielectric should have a higher density of V_{O} 's as compared to the undoped HfO_2 . Thus, PBTI of the La-doped HfO_2 n-MOSFET should correspondingly be degraded. This appears to contradict experimental studies which

have shown that PBTI of the La-doped n-MOSFET is improved as compared to the undoped counterpart [95]. The discrepancy might have prompted some authors to address the issue of charge neutrality as elaborated above [96, 162]. Although the resultant higher E^f , which implies that V_O formation is suppressed in the La-doped HfO_2 , could explain the reduced PBTI effect observed experimentally, the simulation result could not explain 1) the increased susceptibility of the La-doped HfO_2 to stress induced degradation under a high oxide field; 2) the retardation of PBTI recovery given that the V_O CTL is as shallow as that of the undoped HfO_2 .

The greater ease of V_O formation in the La-doped HfO_2 , as revealed by our simulation result, may be reconciled with the experimental observation of a lesser PBTI effect at a low oxide stress field by referring to a recent report on reduced gate tunneling current in the La-doped n-MOSFET biased in the positive regime [163, 164]. The lower gate tunneling current is attributed to dipoles formed between La and the underlying SiO_x , which increase the barrier at the $\text{HfO}_2/\text{SiO}_x$ interface for electron injection from the n-channel [163, 164]. It should be mentioned that the barrier in question is relevant only under a low positive gate voltage where electron tunneling occurs through both the SiO_x interfacial layer and the La-doped HfO_2 . At a higher positive gate voltage, electron tunneling occurs through the SiO_x only and the benefit associated with the La-dipole-induced $\text{HfO}_2/\text{SiO}_x$ barrier increase should vanish. While the above may explain the lesser PBTI effect observed under a low oxide stress field as well as the greater PBTI effect at a higher oxide stress field, the shallow trap states due to V_O fail to account for the slow-down of PBTI recovery observed in the La-doped HfO_2 .

Vacancy-Interstitial Defect

In explaining the slow-down of the PBTI recovery depicted in Figs. 5.21 and 5.23, the defect concerned should be a relatively deep electron trap. For a non-zero electron detrapping barrier (Fig. 5.22(a)), the trap state should be situated below the Si conduction band minimum E_{CBM} . The V_{O} defect in the HfO_2 fails to meet this requirement due to its rather shallow trap state. Moreover, as shown in Fig. 5.26, the average CTL of ~ 4.7 eV is similar for both the undoped and La-doped HfO_2 , which could not explain the slow-down in PBTI recovery observed for the La-doped HfO_2 n-MOSFET. In this section, we show that the related vacancy-interstitial ($V_{\text{O}}\text{-}O_{\text{i}}$) defect can give rise to a deep electron trap in the HfO_2 band gap and the ease of $V_{\text{O}}\text{-}O_{\text{i}}$ generation in the La-doped HfO_2 is greatly enhanced over the undoped counterpart.

In our simulation study, the $V_{\text{O}}\text{-}O_{\text{i}}$ defect was manually created by removing the O atom in between two Hf or La atoms and placing it at an interstitial position in the vicinity of the V_{O} defect created. The resultant structure was then allowed to relax fully prior to the computation of E^{f} and CTL. Examples of the $V_{\text{O}}\text{-}O_{\text{i}}$ defect in the undoped and La-doped HfO_2 are shown in Fig. 5.27(a) and (b), respectively. Numerous randomly chosen O_{i} positions for a given V_{O} position were investigated for each case.

Interestingly, the CTL of the $V_{\text{O}}\text{-}O_{\text{i}}$ defect is more scattered, and could range from anywhere below the E_{VBM} of the HfO_2 to near the CTL of the V_{O} defect, depending on the position of the O_{i} atom. Some examples are shown schematically in Fig. 5.26. This wide spread in the trap state makes the $V_{\text{O}}\text{-}O_{\text{i}}$ defect a suitable candidate for a deep electron trap. For the undoped HfO_2 , the E^{f} for the $V_{\text{O}}\text{-}O_{\text{i}}$ defect is determined as 6.7 eV. On the other hand, the E^{f} is ~ -2.8 eV for the La doped HfO_2 . The negative E^{f} indicates that $V_{\text{O}}\text{-}O_{\text{i}}$ defects are also energetically more favorable, *i.e.*, they are more easily formed in the La-doped HfO_2 as compared to the undoped HfO_2 . A higher $V_{\text{O}}\text{-}O_{\text{i}}$

defect density in the former may explain the slow-down in the PBTI recovery observed as electrons captured at the relatively deep trap states are more difficult to be emitted in a given time interval.

Previous simulation studies have shown that an interstitial oxygen atom alone can give rise to a deep electron trap with energy state near the E_{VBM} of the HfO_2 . However, the likelihood of O_i existing on its own is low since nearly all high- κ oxides are deficient in oxygen. In our simulation model, the O_i atom comes from an O atom displaced from its original lattice position (in between two Hf atoms). As high- κ oxides have a strong ionic character, electric-field induced atomic displacement is a likely scenario in these materials, considering the relatively high field in thin gate stacks. The CTL of the resultant V_O - O_i defect could lie anywhere in the HfO_2 band gap depending on the position of the O_i atom. Our simulation study has shown that at certain positions of the O_i atom, the CTL is relatively shallow and electron trapping occurs mainly at the V_O site. However, at other O_i positions, electron trapping occurs mainly at the O_i site and the resultant CTL is deeper, *i.e.*, at these positions, we may regard the O_i atom to have effectively deactivated the shallower V_O defect. This behavior may explain the electrical conductance switching of high- κ oxides revealed in recent experimental studies and details will be discussed elsewhere.

Besides being able to explain the slow-down in PBTI recovery of the La-doped HfO_2 , the V_O - O_i defect model could also qualitatively account for the worse PBTI effect observed under a high oxide stress field. The negative E^f implies that the energy barrier for the generation of the V_O - O_i defect in the La-doped HfO_2 is lower than that of the undoped HfO_2 . Thus, under a given oxide field, generation of the V_O - O_i defect could occur more readily in the La-doped HfO_2 . Barrier reduction induced by the applied oxide field could further accelerate the generation of the V_O - O_i defect in the

La-doped HfO_2 , leading to a much worse PBTI effect as compared to that of the undoped HfO_2 .

5.2.4 Conclusion

First-principles simulation performed on an amorphous HfO_2 structure shows that La doping could significantly reduce the E^f (becoming negative in fact) for the V_O defect, indicating that such defects are more likely to be formed in the La-doped HfO_2 as compared to the undoped counterpart. It is thus proposed that the reduced PBTI effect in the La-doped HfO_2 at a low oxide field is not because of the suppressed formation of V_O , as claimed in previous simulation studies, but may be due to the smaller gate current that results from the La induced tunneling-barrier increase at the $\text{HfO}_2/\text{SiO}_x$ interface. Results also show that La doping has no apparent impact on the 0/- CTL of the V_O defect, which remains as shallow as that in the undoped HfO_2 . Thus, the V_O defect model alone could not explain the slow-down in PBTI recovery observed experimentally. We show, however, that the 0/- CTL of a related defect, namely the $V_O\text{-}O_i$ paired defect, could range anywhere from the E_{VBM} to the E_{CBM} of the HfO_2 , depending on the position of the interstitial O atom. Thus, at certain “strategic” positions of the O_i , this defect could function as a deep electron trap. Our choice of the $V_O\text{-}O_i$ defect is guided by the generally strong ionic character of high- κ oxides, which implies that field induced displacement of the relatively light O atoms is a likely scenario in these materials in view of the relatively high field in thin gate stacks. It is shown that the zero-field E^f of the $V_O\text{-}O_i$ defect is also significantly reduced by La doping. Hence, the greater ease of its formation in the La-doped HfO_2 may explain the slow-down in PBTI recovery observed, as well as the increased susceptibility of the La-doped HfO_2 to stress induced degradation at a high oxide stress field.

Chapter 6 New Understanding of Vacancy Generation Mechanism for Improved Resistance Switching in the RRAM Device

6.1 Introduction

Resistive memory devices based on switching defects in high- κ metal oxides have attracted considerable attention. In particular, devices employing the HfO_2 dielectric are promising candidates for embedded flash/DRAM applications, by virtue of their compatibility with the mainstream CMOS technology. In spite of the intense attention, several fundamental issues remain: 1) for simple $\text{TiN}/\text{HfO}_2/\text{TiN}$ device without capping layer, memory window is highly unacceptable if considering the enormous current difference between the high resistance state (HRS) and the fresh device, which consequently indicates that a large number of the forming-induced defects are not switchable[141, 144]; 2) knowledge on the nature of the switching mechanism and the impact of electrode/capping layer is limited [103, 126, 165]. To achieve outstanding resistance switching behavior, it requires basic understanding of the underlying mechanism, thus in this chapter, we will demonstrate both experimental and simulation evidences to support the importance of conductive filament formation, as well as the role

of Pt-mediated vacancy (V_O) generation mechanism in improving the resistance switching of the TiN/HfO₂/Pt device.

6.2 Experiments Details

To investigate the switching characteristics of the RRAM devices, six samples with different stacks/electrodes were studied: A) TiN/10-nm HfO₂/TiN; B) TiN/5-nm

Fig. 6.1: schematic layer structures for different samples. (a) TiN/10-nm HfO₂/TiN; (b) TiN/5-nm HfO₂/TiN; (c) TiN/5-nm HfO_x(O-deficient)/TiN; (d) TiN/5-nm HfO₂/Pt; (e) TiN/10-nm HfO₂/Pt; (f) Ti-rich TiN/10-nm HfO₂/Pt; (the green dots indicate the pre-existing oxygen vacancies; TE-top electrode/BE-bottom electrode).

HfO₂/TiN; C) TiN/5-nm HfO_x(O-deficient)/TiN (A-C are control devices); D) TiN/5-nm HfO₂/Pt; E) TiN/10 nm HfO₂/Pt; F) Ti-rich TiN/10-nm HfO₂/Pt, schematic layer structures are shown in the Fig. 6.1. The HfO₂ was prepared by physical vapor deposition. During the electrical measurement, the bottom electrode (BE) - Pt or TiN (in the case of the control device) electrode was grounded and the voltage bias was applied to the top electrode (TE). All the measurements were carried by the Keithley 4200 semiconductor characterization system.

6.3 Experimental Results and Discussion

As typically observed, in Fig. 6.2(a)(i), the current of the TiN/10-nm HfO₂/TiN device is only decreased marginally after reset and the resultant memory window is less than 10 (HRS/LRS<10), which is too small for NVM application, however large difference between the HRS current and the pre-forming state is observed, as plotted in Fig. 6.2(a)(ii), which indicates that most of the defects generated during the forming step are permanent and not switchable. In the TiN/5-nm HfO₂/TiN device, even worse situation is observed that almost all the defects generated during the forming are permanent, which results in no switching (see fig. 6.2(b)). Whereas for the TiN/5-nm HfO_x/TiN device, due to the inherently high oxygen vacancy concentration in the fresh device, which renders the device formed at lower voltage, therefore the subsequent switching cycle is only able to reset part of the defects, resulting in a memory window around 10 (HRS/LRS<10). At this stage, it is worth to mention the role of the forming step on generating the switchable defects. Since the switching of the RRAM device strongly depends on the property of the generated defects in the conductive filament, therefore a proper forming sweep will determine the memory window. Other than that, for the TiN/5-nm HfO₂/Pt device, a near-complete reset of the current to its pre-

forming level is obtained in Fig. 6.3(a)(i), giving rise to a much larger memory window (~ 600). And referring to the HRS current, as we can see that, it is almost reset to the pre-forming level, which implies that almost the entire percolation path (PP) is eliminated during reset. Interestingly, the near-complete current-reset is only observed when Pt is the cathode during *forming/set* (Figs. 6.3(a)), in Fig. 6.3(b), the reverse -

Fig. 6.2: (a) TiN/10-nm HfO₂/TiN- (i) a marginal decrease in the current during reset and the resultant memory window is small. (ii) A large difference between the current in the high resistance state (HRS) and the pre-forming state. (b) TiN/5-nm HfO₂/TiN (nearly no switching); (c) TiN/5-nm HfO_x/TiN with a memory window around 10 (HRS/LRS ≈ 10).

forming/set only produces a relatively small memory window. In addition, in Fig. 6.3(a)(ii), when Pt is taken as the anode during forming/set, the current for the first reset sweep is lower than that for the reverse sweep during forming/set. This implies that defects formed during the negative forming/set are not stable and have recovered

before the reset sweep, and the generated switchable defects are stable only when the Pt is negatively biased during forming/set.

Currently a general explanation attributes the current-reset to re-oxidation of the PP[99], by the release of oxygen stored in the anode during forming/set, since the top TiN electrode is the anode during the forming/set of the control and TiN/HfO₂/Pt devices, this explanation obviously could not account for the significantly greater current-reset in the latter. Besides, for TiN/10-nm HfO₂/Pt device, a similarly significant current-reset is obtained, as shown in Fig. 6.3(b)(i). In this case, however,

Fig. 6.3: (a) With one of the TiN electrodes replaced by Pt and the Pt made the cathode during forming/set, more significant current switching is obtained and the current after reset is returned to the pre-forming level reset (b) Negative forming/set only shows marginal memory window; (c) Behavior similar to the TiN/5-nm HfO₂/Pt device is observed for the TiN/10-nm HfO₂/Pt device; (d) Negative forming/set shows negligible memory window.

the current during the reverse set-sweep and forward reset-sweep are nearly identical. As defects located near the emitting electrode have a greater impact on the tunneling current, the former implies a higher defect density near the TiN electrode after reset. Meanwhile, the current after reset could not be reset back to the pre-forming level, which implies that the Pt-cathode-induced switchable defects are situated near the Pt itself. Since in a thicker HfO_2 , more defects are needed to build up the conductive filament, thus conductive defects will accumulate near the region close to the top TiN electrode, however the effective length of the Pt electrode is supposed to be limited, thus defects in the upper region are unlikely to be switchable, and become the permanent conductive path.

6.4 First-principles Simulation and Model Discussion

6.4.1 Simulation Results Discussion

To shed light on the Pt-mediated defect generation mechanism, *ab-initio* simulation was performed by using the VASP [137]. A melting-quenching scheme was performed to generate the amorphous HfO_2 supercell [153], which contains 96 Hf -

Fig. 6.4: Ball-and-stick representation of (a) the amorphous HfO_2 supercell generated using the melt-and-quench scheme. (b) The incorporation of Pt is simulated by replacing a Hf atom with a Pt atom.

atoms and 64 oxygen atoms, respectively (see fig. 6.4(a)). To study the impact of Pt, one Hf atom was replaced by the Pt atom, and then subjected the supercell to the full structure relaxation, as illustrated in the Fig. 6.4(b). V_O formation was simulated by displacing a lattice O atom to an arbitrary -interstitial position as shown in Figs. 6.4(b) respectively and the reaction path was examined via the nudged-elastic band method [166]. Ultra-soft pseudo-potentials were used to represent the interactions between the ion core and the valence electrons [146]. The exchange correlation functional was treated within the GGA of Perdew, Burke and Ernzerhof [133]. In all calculations, the cut-off energy and κ -point were tested to make sure all the calculated results were converged. As for structure relaxation, the conjugate gradient method was used and the ion positions were optimized until the residual force was less than 0.01 eV/Å.

Full structure relaxation were performed for the V_O defect of both HfO_2 and HfO_2 -Pt supercell, our interesting results show that when full structural relaxation of the defective supercell was performed under a negatively charged condition, O_i remains in

Fig. 6.5: (i)(a),(ii)(a) The generation of vacancy (V_O). (i)(b),(ii)(b) Defect site after full structural relaxation under negatively charged condition. (i)(c),(ii)(c) Defect site after full structural relaxation under neutral condition.

its interstitial position (Figs. 6.5(i)(b), (ii)(b)). However, when the structural relaxation was carried out under a neutral condition, as illustrated in Figs. 6.5i(c), ii(c)– the V_O site is occupied by a neighboring O atom whose position is in turn filled by O, which consequently eliminate the V_O . This implies that stable V_O s would usually be first formed near the cathode where the defects are negatively charged most of the time. Moreover, the nudged elastic results for structure barrier calculation illustrate that Pt lowers the energy barrier E_B for V_O formation by ~ 1.3 eV as shown in Fig. 6.6(a). This is further confirmed by the observation that the length of Hf-O bonds in the vicinity of the Pt atom is increased in the Fig. 6.7, which means that the Hf-O bonds are now weaker and may be dissociated more easily than in the case when Pt is absent. Finally, in our simulation results, we found Pt introduces an *intermediate* meta-stable V_O-O_i defect, with the interstitial oxygen O_i situated closer to the vacancy site than in the final defective supercell (Fig. 6.5(b)). The E_B for the meta-stable V_O-O_i formation is much smaller (by ~ 60 %) as compared to that for the final V_O formation.-

Fig. 6.6: Evolution of system energy during V_O generation under negatively charged condition. (b) Pt induced meta-stable state.

Fig. 6.7: Ball-and-stick illustration of the impact of Pt incorporation into the HfO₂ on the surrounding Hf-O bond length. Compared to the case when Pt is absent (a), the Hf-O bond length in the vicinity of the Pt atom is increased (b)

6.4.2 Physical Model Discussion

To have a comprehensive understanding of the RRAM switching mechanism, the oxide defect models are discussed by taking the *ab-initio* simulation results into consideration. During the forming step, V_O s are generated near the cathode/HfO₂ interfacial region while O atoms are displaced into interstitial positions and charged negatively as shown in the Fig. 6.5(a). However, after the oxygen atom leave its lattice position, the movement of O^- through the HfO₂ entails a small energy barrier, only around 0.3 eV, which was elaborated in the ref. [167], therefore any delay in stopping the process would render the generated V_O s more permanent when the O^- ions are transported further away. This is especially crucial in the TiN/HfO₂/TiN device for which the E_B for V_O generation is relatively large (Fig. 6.6(a)) and thus forming could only happen at a higher voltage. Since Pt decreases the E_B for V_O generation, forming could thus happen at a lower voltage (reduced electric field) in the TiN/HfO₂/Pt device. Moreover, Fig. 6.6(b) shows that Pt electrode facilitates the formation of a meta-stable

V_O - O_i defect. Therefore, one expects the V_O - O_i separation in this device to be considerably reduced for a given “stopping” delay. Under a reverse (negative) electric field shown in the Fig. 6.8(b), the nearby O_i s could readily recombine with the V_O s,

Fig. 6.8: Schematic illustration of (a) vacancy V_O generation during forming/set; (b)

thus contributing to the near-complete current-reset (see fig. 6.3(a)). Whereas with a thicker HfO_2 layer (see fig. 6.3(b)(i)), different scenario should be considered. To set the RRAM device to LRS, the conductive filament is supposed to penetrate through the whole dielectric layer, thus defects located in the upper portion of the oxide would need to be generated at a given location before a certain current compliance can be reached. While the generation of V_O s near the Pt electrode could still remain Pt-mediated, however once they are generated, the bulk of the initially higher voltage (needed to overcome the E_B for the Pt-mediated V_O generation in a thicker oxide), would be transferred to the upper part of the oxide. This makes the control of switchable V_O generation in this region even more difficult. Based on the above discussion, the switching characteristics of TiN/10-nm HfO_2 /Pt device are elaborated in the Fig. 6.9. One can see that only those defects generated near the Pt cathode during forming could be “switched”, therefore the reset voltage sweep will cure the

switchable defects whereas leave those permanent defects located in the upper layer there, and ultimately cause the current could not completely return to the pre-forming level. To further confirm our model, we purposely modify the defect density in the top half layer of the 10 nm HfO_2 by using a *Ti-rich* TiN top electrode. EDX analysis confirms a higher Ti content in this region. The incorporation of a higher Ti content into HfO_2 is known to increase the pre-existing V_O density [150, 165], as borne out by the now much higher pre-forming current as illustrated in the Fig. 6.8, similar to that-

Fig. 6.9: (a) in a thicker HfO_2 , more V_O 's are needed to form a percolation path. (b) Although the proximity of O_i 's to the V_O sites in the Pt influence zone could lead to a near-complete removal of the V_O 's, this could not happen as easily in zone X, due to the greater physical separation between the V_O 's and O 's.

Fig. 6.10: (a) Current-voltage curves of the Ti-rich TiN/HfO₂/Pt device in the order of forming, reset and set. The current compliance was set to 1×10^{-5} A and 1×10^{-3} A for forming and set, respectively. (b) The significant current switching between forming/set and reset can be consistently observed under repeated cycling.

of the TiN/5-nmHfO₂/Pt device (see fig. 6.3(a)) The forming voltage is also comparable as PP formation now requires only defects in the lower half layer to be generated. Besides, the current after reset is decreased by a similar amount as those for the TiN/5-nm/Pt and TiN/10-nm/Pt devices. These results support the above model: 1) resistance switching is mainly controlled by V_{OS} generated near the cathode; 2) the Pt-mediated generation mechanism renders these V_O 's more "switchable".

6.5 Conclusion

Experimental study on TiN/5-nm HfO₂/Pt resistive memory devices shows near-complete current-reset that leads to a significant enhancement of the memory window. Through *ab-initio* simulation, the effect of Pt on V_O generation in HfO₂ is revealed for the first time. It is shown that Pt lowers the energy barrier (and hence the forming voltage) for V_O generation and introduces a meta-stable defect comprising V_O and O_i that are more closely spaced. These effects are believed to have led to a more controllable forming process (and hence the V_O 's generated are more recoverable) which in turn contributes to the near-complete current-reset observed. The findings

suggest the importance of forming steps in the conductive filament generation and the role of Pt on achieving improved bipolar operation of the TiN/HfO₂/TiN device.

Chapter 7 CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion

A systematic reliability study, comprising device level electrical characterizations as well as first-principles modeling & simulation of SiON and hafnium oxide dielectrics layers, provides useful insights into the performance and reliability of the advanced MOSFETs & RRAM. The key findings achieved in this research are summarized as follows:-

- Study of the role of the neighboring nitrogen atom of V_O following the capture of a hole in the SiON dielectric, via first-principles simulation based on the density functional theory, show the paramount impact on the of the on the structural relaxation. Particularly, a two-fold coordinated nitrogen atom is shown to induce very significant structural relaxation of V_O . The resultant trapped-hole site has a very deep charge transition level in the SiO_2 band gap. On the other hand, in the absence of the nitrogen atom or if the neighboring nitrogen atom is three-fold coordinated (*e.g.*, the dangling nitrogen bond is terminated by a hydrogen atom), the resultant trapped-hole site would have a much shallower charge transition energy level. .

- Insight through the first-principles simulation, a comprehensive picture of the properties of nitrogen involved oxygen vacancy defects is revealed. The dynamic and relaxation characteristic of different nitrogen associated defects are studied, our results illustrate that nitrogen related defects can have various behaviors and consequently generate a wide range of defect energy levels throughout the band gap, which perfectly reproduces the experimental observations.
- Investigation of the transient charge trapping in the high- κ oxide defects via dynamic BTI measurement, we show that the recoverable component of the V_{th} degradation can be transformed into more permanent trapped charge when the applied electrical cum thermal stress exceeds a certain threshold. In addition, we found that when the transformation of hole trapping occurs under negative-bias temperature stress, a correlated increase of the gate current is always observed, which points to the transformation process being the origin for bulk oxide trap generation. However, when the transformation of electron trapping occurs under positive-bias temperature stress, an increase of the gate current is not always observed. Moreover, from *ab-initio* simulation, we elaborate that an intrinsic oxide defect – the oxygen vacancy-interstitial (V_O-O_i) – could consistently explain the above experimental observations. *E.g.*, generation of deep-level trapped charges as well as transformation of transient (or shallow) to permanent (or deep) charge trapping.
- Study of the impact of La doping on the reliability issues in the HfO_2 gate stacks provides one-more step understanding of the La involved dielectric degradation. Through atomic modeling, the electronic properties of the La associated oxygen vacancy (V_O) defect and the related vacancy-interstitial (V_O-

O_i) paired defect are discussed. Our modeling work enables us to explain the PBTI improvement, the TDDB degradation, and the slowing recovery of PBTI in the La-doped $HfSiO_x$ gated n-MOSFET.

- Inspection on the distinctive switching characteristics of the RRAM stacks with different electrodes. We show the critical role of the forming process in creating the first conductive filament in the fresh device. And in addition, the Pt mediated oxygen vacancy formation during the forming step is modeled, which explains the excellent switching characteristics of the RRAM stacks when Pt is used as cathode.

7.2 Recommendations

Unlike the SiO_2 gate dielectric, which has been extensively studied for the past decades, the application of high- κ oxide as the transistor's gate dielectric/RRAM switching layer is deemed new and the fundamental understanding of the high- κ material is still lacking. Moreover, currently the modeling work is only performed under 0K and 0 V/cm electric-field condition with limited number of atoms. Therefore our study work could be further extended to enhance the understanding of the high- κ material properties as well as the atomic modeling. Possible future studies are listed as follows:-

- Studies on DPBTI stressing with different duty cycles show greater transformation of shallow electron-trap states to deeper levels at a higher frequency [18, 89]. To consistently explain the result, it is necessary for the transformation to occur via an intermediate state. However, a clear physical picture of the evolution of the oxide defects (*e.g.*, V_O-O_i) that response the intermediated states is not available yet. Therefore it is necessary to carry more follow-up researches through both

first-principles simulation and electrical measurement to build up the understanding.

- Single defect study on small dimension device show the progressive evolution of the defect charge emission time when the NBTI stress cycles is repeatedly applied. The results confirm our earlier inference of transient to permanent hole trap transformation, derived from the decreasing recovery per cycle of a large area device subjected to NBTI stress/relaxation cycling. While the nature of such defects remains elusive, the study clearly shows that trap parameters such as the emission time constant are not time-invariant but evolve continuously during stress. The evolution is believed to be linked to phonon-facilitated local defect structural changes which occur upon charging/discharging as well as the electric-field assisted barrier lowering. However, currently atomic simulation doesn't comprise the impact of the temperature and electric-field; therefore it is worth to conduct the modeling and simulation work with the temperature and electric fields included [51, 168].
- Although some studies have suggested that hydrogen may play a role on NBTI its effect appears to be insignificant in the ultra-thin nitride SiO_2 or the high- κ oxide in which the charge trapping effect is shown to generally dominate, however experiment results of the deuterium-annealed p-MOSFETs show the improved NBTI performance compared to the hydrogen-annealed devices, indicating the hydrogen-related defects (located near the interface) are involved in the NBTI degradation. And the exact role of hydrogen on NBTI degradation remains elusive to-date. Therefore it is worth to have a more detailed study through electrical measurement and first-principles modeling [60, 169];

- Although the hafnium oxide based RRAM is successfully demonstrated and oxygen vacancies related conduction filament is commonly accepted for explanation of resistive switching behavior, however the controlling of the filament generation, which will impact device uniformity and low power improvement, is still quite challenging. Moreover, based on the previous results, materials oriented of the selection of the capping layer and the electrode material play important role in determining the device performance. Therefore, a more detailed study of the material selection for switching layer and electrodes as well as optimized forming techniques are recommended [34, 170, 171].

LIST OF PUBLICATIONS

- **C. J. Gu** and D. S. Ang, “Impact of Twofold Coordinated Nitrogen on the Generation of Deep-Level Hole Traps under Negative-Bias Temperature Stressing”, ECS Trans., V. 35, Issue 4, 125-143, 2011.
- **C.J. Gu** and D. S. Ang, “Impact of La Doping on the HfO₂ Gate Stack Reliability via First-Principles Simulation”, ECS Trans., V. 53, Issue 3, 193-204, 2013.
- **C. J. Gu** and D. S. Ang, “A New Atomic Defect Model for Positive-Bias Temperature Instability in the High-k Gate n-MOSFET”, Solid State Devices and Materials, Japan, 2013.
- D. S. Ang, **C. J. Gu**, Z. Y. Tung, A.A. Boo and Y. Gao, “Evolution of oxide charge trapping under bias temperature stressing”, Microelectronics Reliability, in press.
- G. A. Du, D. S. Ang and **C.J. Gu**, “Abnormal Slow Recovery Characteristic of La-Doped HfSiO_x N-MOSFET Bias-Temperature Instability”, IEDM Tech. Digest, P. 4.7.1 - 4.7.4, 2010.
- Y. Gao, D. S. Ang and **C. J. Gu**, “On the Evolution of Switching Oxide Traps in the HfO₂/TiN Gate Stack Subjected to Positive- and Negative-Bias Temperature Stressing”, ECS Trans., V. 53, Issue 3, 205-220, 2013.

BIBLIOGRAPHY

1. Atalla, M.M., Tannenbaum, E. and Scheibner, E.J., *Stabilization of Silicon Surfaces by Thermally Grown Oxides*. Bell System Technical Journal, 1959. **38**(3): p. 749-783.
2. Dawon, K., *Silicon-silicon dioxide field induced surface devices*. Bell System Technical Journal, 1961: p. 583-596.
3. Cheng, K., Lee, J. and Lyding, J.W., *Experimental evidence of Si-H bond energy variation at SiO₂-Si interface*. Appl. Phys. Lett. , 2000. **77**(21): p. 3388.
4. Grasser, T., Reisinger, H., Wagner, P.-J., *et al.*, *The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability*. Proc. Int. Reliab. Phys. Symp., 2010: p. 16-25.
5. Pantelides, S.T., Rashkeev, S.N., Buczko, R., *et al.*, *Reactions of hydrogen with Si-SiO₂ interfaces*. IEEE Trans. Nucl. Sci. **47**(6): p. 2262,2268.
6. Moore, G.E., *Cramming More Component onto Integrated Circuits*. Proceedings of the IEEE, 1965. **86**(1): p. 82-85.
7. Chen, A., Das, S., Zhirnov, V., *et al.*, *ITRS Public Conference-Emerging Research devices*. ITRS Report, 2012.
8. Mitani, Y., Nagamine, M., Satake, H., *et al.*, *NBTI Mechanism in Ultra-thin Gate Dielectric-Nitrogen-originated Mechanism in SiON*. IEDM Tech. Digest, 2002: p. 509-511.
9. Kimizuka, N., Yamaguchi, K., Imai, K., *et al.*, *NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10 gate CMOS generation*. Symp. VLSI Tech. Tech. Dig, 2000: p. 92-93.
10. Gusev, E., Buchanan, D.A., Cartier, E., *et al.*, *Ultrathin high-k gate stacks for advanced CMOS devices*. IEDM Tech. Dig., 2001. **20.1.1-20.1.4**.
11. Bersuker, G., Zeitzoff, P., Brown, G., *et al.*, *Dielectrics for future transistors*. Mater. Today, , 2004. **7**: p. 26.
12. Wilk, G.D., Wallace, R.M. and Anthony, J.M., *High-k gate dielectrics: Current status and materials properties considerations* J. Appl. Phys. , 2001. **89**: p. 5243.
13. Chau, R., Datta, S., Doczy, M., *et al.*, *High-k/metal-gate stack and its MOSFET characteristics* IEEE Electron Dev. Lett., 2004. **25**: p. 408.

14. Takeuchi, H., Ha, D. and King, T.-J., *Observation of bulk HfO₂ defects by spectroscopic ellipsometry*. J. Vac. Sci. Technol. A 2004. **22**(4): p. 1337.
15. Miura, Y. and Matukura, Y., *Investigation of Silicon-Silicon Dioxide Interface Using MOS Structure*. Jpn.J.Appl.Phys., 1966. **5**: p. 180.
16. Mitani, Y., Satake, H. and Toriumi, A., *Influence of nitrogen on negative bias temperature instability in ultrathin SiON*. IEEE Trans. Dev. Mater. Reliability, 2008. **8**(1): p. 6-12.
17. Yonamoto, Y. and Akamatsu, N., *Interface traps responsible for negative bias temperature instability in a nitrided submicron metal-oxide-semiconductor field effect transistor*. Appl. Phys. Lett. , 2011. **98**(10): p. 103513.
18. Grasser, T., Kaczer, B., Goes, W., et al., *The paradigm shift in understanding the bias temperature instability: from reaction–diffusion to switching oxide traps*. IEEE Trans. Electron Devices, 2011. **58**(11): p. 3652-3666.
19. Grasser, T., Gos, W., Sverdlov, V., et al., *The universality of NBTI relaxation and its implications for modeling and characterization*. Proc. Int. Reliab. Phys. Symp., 2007: p. 268-280.
20. Kaczer, B., Grasser, T., Roussel, P.J., et al., *Ubiquitous relaxation in BTI stressing—New evaluation and insights*. Proc. Int. Reliab. Phys. Symp., 2008: p. 20-27.
21. Reisinger, H., Vollertsen, R., Wagner, P.-J., et al., *A study of NBTI and short-term threshold hysteresis of thin nitrided and thick non-nitrided oxides*. IEEE Trans. Dev. Mater. Reliability, 2009. **9**(2): p. 106-114.
22. Grasser, T. and Kaczer, B., *Evidence that two tightly coupled mechanisms are responsible for negative bias temperature instability in oxynitride MOSFETs*. IEEE Trans. Electron Devices, 2009. **56**(5): p. 1056-1062.
23. Kimizuka, N., Yamamoto, T., Mogami, T., et al., *The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling*. Symp. VLSI Tech. Tech. Dig, 1999: p. 73-74.
24. Ang, D.S. and Wang, S., *Recovery of the NBTI-Stressed Ultrathin Gate p-MOSFET: The Role of Deep-Level Hole Traps*. IEEE Electron Dev. Lett., 2006. **27**: p. 914-916.
25. Ang, D.S., Wang, S., Du, G.A., et al., *A Consistent Deep-Level Hole Trapping Model for Negative Bias Temperature Instability*. IEEE Trans. Dev. Mater. Reliability, 2008. **8**(1): p. 22-34.
26. Zafar, S., Kim, Y.H., Narayanan, V., et al., *A Comparative study of NBTI and PBTI in SiO₂ HfO₂ stacks with FUSI*. Symp. VLSI Tech. Tech. Dig, 2006.
27. Gao, Y., Ang, D.S., Young, C.D., et al., *Evolution of electron trapping under positive-bias temperature stressing of the HfO₂TiN gate n-mosfet*. Solid State Devices and Materials, 2012.

28. Bersuker, G., Sim, J.H., Park, C.S., *et al.*, *Mechanism of electron trapping and characteristics of traps in HfO₂ gate stacks*. IEEE Trans. Dev. Mater. Reliability, 2007. v. 7: p. 138-145.
29. Du, G.A., Ang, D.S. and Gu, C.J., *Abnormal slow recovery characteristic of La-doped HfSiO_x n-MOSFET bias-temperature instability*. IEDM Tech. Digest, 2010: p. 4.7.1-4.7.4.
30. Bez, R., Camerlenghi, E., Modelli, A., *et al.*, *Introduction to flash memory*. Proceedings of the IEEE 2003. **91**(4): p. 489-502.
31. Pavan, P., Bez, R., Polivo, P., *et al.*, *Flash memory cells-an overview*. Proceedings of the IEEE 1997. **85**(8): p. 1248-1271.
32. Aritome, S., Shiota, R., Hemink, G., *et al.*, *Reliability issues of flash memory cells*. Proceedings of the IEEE 1993. **81**(5): p. 776-788.
33. Chen, A., Haddad, S. and Wu, Y.-C., *Non-volatile resistive switching for advanced memory applications*. IEDM Tech. Digest, 2005.
34. Lee, Y.L., Pang-Shiu, C., Wu, T.-Y., *et al.*, *HfO_x bipolar resistive memory with robust endurance using AlCu as buffer electrode*. IEEE Electron Dev. Lett., 2009. **30**(7): p. 703-705.
35. Rajendran, B., Breitwisch, M., Cheek, R., *et al.*, *Characterization of poly-Silicon emitter BJTs as access devices for Phase Change Memory*. Symp. on VLSI Technol., 2009: p. 31-32.
36. Bersuker, G., Park, C.S., Barnett, J., *et al.*, *The effect of interfacial layer properties on the performance of Hf-based gate stack devices*. J. Appl. Phys. , 2006. **100**(9): p. 094108.
37. Nadimi, E., Roll, G., Kupke, S., *et al.*, *The degradation process of high k SiO₂/HfO₂ gate stacks: a combined experimental and first principle investigation*. IEEE Trans. Electron Devices, 2014. **61**(5): p. 1278-1283.
38. Govoreanu, B., Kar, G.S., Chen, Y.-Y., *et al.*, *10×10nm² Hf HfO_x crossbar resistive RAM with excellent performance, reliability and low-energy operation*. IEDM Tech. Digest, 2011.
39. Alam, M.A., *A critical examination of the mechanics of dynamic NBTI for p-MOSFETs*. IEDM Tech. Digest, 2003: p. 14.4.1 - 14.4.4
40. Varghese, D., Saha, D., Mahapatra, S., *et al.*, *On the dispersive versus Arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: Measurements, theory, and implications*. IEDM Tech. Digest, 2005: p. 684 - 687.
41. Mahapatra, S., Ahmed, K., Varghese, D., *et al.*, *On the physical mechanism of NBTI in silicon oxynitride p-MOSFETs: Can differences in insulator processing conditions resolve the interface trap generation versus hole trapping controversy?* Proc. Int. Reliab. Phys. Symp., 2007: p. 1-9.

42. Jeppson, K.O. and Svensson, C.M., *Negative Bias Stress of MOS Device at High Electric-fields and Degradation of MNOS Device*. J. Appl. Phys. , 1977. **48**(5): p. 2004-2014.
43. Alam, M.A. and Mahapatra, S., *A comprehensive model for PMOS NBTI degradation: Recent progress*. Microelectronic Reliab., 2005. **45**: p. 71-81.
44. Alam, M.A., *A critical examination of the mechanics of dynamic NBTI for PMOSFETs*. IEDM Tech. Dig., 2003: p. 345-348.
45. Chakravarthi, S., Krishnan, A.T., Reddy, V., *et al.*, *A comprehensive framework for predictive modeling of negative bias temperature instability*. Proc. Int. Reliab. Phys. Symp., 2004: p. 273-282.
46. Teo, Z.Q., Ang, D.S. and See, K.S., *Can the Reaction-Diffusion Model Explain Generation and Recovery of Interface States Contributing to NBTI?* IEDM Tech. Digest, 2009: p. 733-739.
47. Zhang, J.F., Zhao, C.Z., Chen, A.H., *et al.*, *Hole traps in silicon dioxides - Part I: Properties*. IEEE Trans. Electron Devices, 2004. **51**(8): p. 1267-1273.
48. Teo, Z.Q., Ang, D.S. and Ng, C.M., *Non-Hydrogen-Transport Characteristics of Dynamic Negative-Bias Temperature Instability*. IEEE Electron Dev. Lett., 2010. **31**(4): p. 269-271.
49. Teo, Z.Q., Boo, A.A., Ang, D.S., *et al.*, *On the Cyclic Threshold Voltage Shift of Dynamic Negative-Bias Temperature Instability*. Proc. Int. Reliab. Phys. Symp., 2011: p. XT.10.1 - XT.10.5.
50. Grasser, T., Kaczer, B., Reisinger, H., *et al.*, *Recent advances in understanding the bias temperature instability*. Proc. Intl. Reliab. Phys. Symp., 2010: p. XT.8.1-XT.8.7.
51. Reisinger, H., Grasser, T., Gustin, W., *et al.*, *The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress*. Proc. Int. Reliab. Phys. Symp., 2010(7-15).
52. Zafar, S., Kumar, A., Gusev, E., *et al.*, *Threshold Voltage instabilities in High k gate dielectrics stacks*. IEEE Trans. Electron Devices, 2005. **5**(1): p. 45-63.
53. Pae, S., Agostinelli, M., Brazier, M., *et al.*, *BTI reliability of 45 nm high-K + metal-gate process technology*. Proc. Intl. Reliab. Phys. Symp., 2008: p. 352-357.
54. Zhang, L. and Huang, R., *Statistical Analysis of Retention Behavior and Lifetime Prediction of HfOx-based RRAM*. Proc. Int. Reliab. Phys. Symp., 2011.
55. Fang, Z., Yu, H.Y., Liu, W.J., *et al.*, *Temperature Instability of Resistive Switching on HfOx-Based RRAM Device*. IEEE Electron Dev. Lett., 2010. **31**(5).

56. Chen, T.-J. and Kuo, C.-L., *First principles study of the structural, electronic, and dielectric properties of amorphous HfO₂*. J. Appl. Phys. , 2011. **110**(6): p. 064105.
57. Ang, D.S. and Wang, S., *Insight Into the Suppressed Recovery of NBTI-Stressed Ultrathin Oxynitride Gate pMOSFET*. IEEE Electron Dev. Lett., 2006. **27**(9): p. 755-758.
58. Ang, D.S. and Wang, S., *Recovery of the NBTI stressed ultra-thin gate P-MOSFET: the role of deep-level hole traps*. IEEE Electron Dev. Lett., 2006. **27**(9): p. 914-916.
59. Tan, S.S., Chen, T.P., Soon, J.M., *et al.*, *Nitrogen-enhanced negative bias temperature instability: An insight by experiment and first-principle calculations*. Appl. Phys. Lett. , 2003. **82**(12): p. 1881.
60. Fujieda, S., Miura, Y., Saitoh, M., *et al.*, *Interface defects responsible for negative-bias temperature instability in plasma-nitrided SiON/Si(100) systems*. Appl. Phys. Lett. , 2003. **82**(21): p. 3677.
61. Campbell, J.P., Lenahan, P.M., Cochrane, C.J., *et al.*, *Atomic-Scale Defects Involved in the Negative Bias Temperature Instability*. IEEE Trans. Dev. Mater. Reliability, 2007. **7**(4): p. 540-556.
62. Campbell, J.P., Lenahan, P.M., Krishnan, A.T., *et al.*, *Identification of the atomic-scale defects involved in the negative bias temperature instability in plasma-nitrided p-channel metal-oxide-silicon field-effect transistors*. J. Appl. Phys. , 2008. **103**(4): p. 044505.
63. Grasser, T., Kaczer, B., Goes, W., *et al.*, *A two-stage model for negative bias temperature instability*. Proc. Int. Reliab. Phys. Symp., 2009: p. 33-44.
64. Tan, S.S., Chen, T.P., Ang, C.H., *et al.*, *Atomic Modeling of Nitrogen Neighboring Effect on Negative Bias Temperature Instability of pMOSFETs*. IEEE Electron Dev. Lett., 2004.
65. Yang, J.B., Chen, T.P., Tan, S.S., *et al.*, *Modeling and Characterization of Nitrogen-Enhanced Negative-Bias Temperature Instability in p-Channel MOSFETs*. J. Electrochem. Soc., 2007. **154**(12): p. G255.
66. Islam, A.E., Kufluoglu, H., Varghese, D., *et al.*, *Critical analysis of short-term negative bias temperature instability measurements: Explaining the effect of time-zero delay for on-the-fly measurements*. Appl. Phys. Lett. , 2007. **90**(8): p. 083505.
67. McPherson, J.W., Khamankar, R.B. and Shanware, A., *Complementary model for intrinsic time-dependent dielectric breakdown in SiO₂ dielectrics*. J. Appl. Phys. , 2000. **88**(9): p. 5351.
68. Tsetseris, L., Zhou, X.J., Fleetwood, D.M., *et al.*, *Hydrogen-Related Instabilities in MOS Devices Under Bias Temperature Stress*. IEEE Trans. Dev. Mater. Reliability, 2007. **7**(4): p. 502-508.

69. Nicklaw, C.J., Lu, Z.Y., Fleetwood, D.M., *et al.*, *The structure, properties, and dynamics of oxygen vacancies in amorphous SiO₂*. IEEE Trans. Nucl. Sci., 2002. **49**(6): p. 2667-2673.
70. Lee, E.-C. and Chang, K.J., *Atomic model for the electrical deactivation of N in Si oxynitrides*. Physica B: Condensed Matter, 2003. **340-342**: p. 974-977.
71. Schroder, D.K., *Negative bias temperature instability: What do we understand?* Microelectronic Reliab., 2007. **47**(6): p. 841-852.
72. Shen, C., Li, M.-F., Foo, C.E., *et al.*, *Characterization and Physical Origin of Fast V_{th} Transient in NBTI of pMOSFETs with SiON Dielectric*. IEDM Tech. Digest, 2006: p. 1-4.
73. Huard, V., Parthasarathy, C., Rallet, N., *et al.*, *New characterization and modeling approach for NBTI degradation from transistor to product level*. IEDM Tech. Digest, 2007: p. 797-800.
74. Grasser, T., Wagner, P.-J., Hehenberger, P., *et al.*, *A Rigorous Study of Measurement Techniques for Negative Bias Temperature Instability*. IEEE Trans. Dev. Mater. Reliability, 2008. **8**(3): p. 526-535.
75. Grasser, T., Th, A., Pobegen, G., *et al.*, *The 'permanent' component of NBTI: Composition and annealing*. Proc. Int. Reliab. Phys. Symp., 2011: p. 6A. 2.1-6A. 2.9.
76. Huard, V., *Two independent components modeling for Negative Bias Temperature Instability*. Proc. Int. Reliab. Phys. Symp., 2010: p. 33-42.
77. Reisinger, H., Grasser, T., Gustin, W., *et al.*, *The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress*. Proc. Int. Reliab. Phys. Symp., 2010: p. 7-15.
78. Yang, T., Li, M.F., Shen, C., *et al.*, *Fast and slow dynamic NBTI components in p-MOSFET with SiON dielectric and their impact on device life-time and circuit application*. Symp. VLSI Tech. Tech. Dig, 2005: p. 92-93.
79. Robertson, J., *Interfaces and defects of high- oxides on silicon*. Solid-State Electronics, 2005. **49**(3): p. 283-293.
80. Wilk, G.D., Wallace, R.M. and Anthony, J.M., *High- κ gate dielectrics: Current status and materials properties considerations* J. Appl. Phys., 2001. **89**: p. 5243.
81. Ma, T.P., *Electrical characterization of High k gate dielectrics*. IEDM Tech. Digest., 2004.
82. Kang, A.Y., Lenahan, P.M. and Conley, J.F., *Electron spin resonance observation of trapped electron centers in atomic-layer-deposited hafnium oxide on Si*. Appl. Phys. Lett. , 2003. **83**(16): p. 3407.

83. Lenahan, P.M. and Conley, J.F., *Magnetic resonance studies of trapping centers in high- κ dielectric films on silicon*. IEEE Trans. Dev. Mater. Reliability, 2005. **5**(1): p. 90-102.
84. Xiong, K., Robertson, J. and Clark, S.J., *Defect energy states in high-K gate oxides*. physica status solidi (b), 2006. **243**(9): p. 2071-2080.
85. Gusev, E.P., Buchanan, D.A., Cartier, E., *et al.*, *Ultrathin High k gates stacks for advanced CMOS devices*. IEDM Tech. Digest., 2001. **01-451**.
86. Oates, A.S., *Reliability issues for high-k gate dielectrics*. IEDM Tech. Digest., 2003. **03-923**.
87. Foster, A., Lopez Gejo, F., Shluger, A., *et al.*, *Vacancy and interstitial defects in hafnia*. Phys. Rev. B, 2002. **65**(17).
88. Onishi, K., Kang, C.S., Choi, R., *et al.*, *Reliability characteristics, including NBTI, of polysilicon gate HfO/sub 2/ MOSFET's*. IEDM Tech. Digest., 2011. **01-659**.
89. Zhao, K., Stathis, J.H., Kerber, A., *et al.*, *PBTI relaxation dynamics after AC vs. DC stress in high-k metal gate stacks*. Proc. Int. Reliab. Phys. Symp., 2010: p. 50-54.
90. Reisinger, H., Grasser, T., Ermisch, K., *et al.*, *Understanding and modeling AC BTI*. Proc. Int. Reliab. Phys. Symp., 2011: p. 6A.1.1-6A.1.8.
91. Kita, K. and Toriumi, A., *Origin of electric dipoles formed at high-k/SiO₂ interface*. Appl. Phys. Lett. , 2009. **94**(13): p. 132902.
92. Kakushima, K. and Okamoto, K., *Impact of Thin La₂O₃ Insertion for HfO₂ MOSFET*. ECS Trans., 2008. **13**(2): p. 29-37.
93. Sato, M. and Kamiyama, S., *Negatively Charged Deep Level Defects Generated by Yttrium and lanthanum Incorporation into HfO₂ for V_{th} adjustment, and the Impact on TDDB, PBTI and 1/f noise*. IEDM Tech. Digest, 2009. **09**: p. 123.
94. Kang, C.Y., Kirsch, P.D., Lee, B.H., *et al.*, *Reliability of La-Doped Hf-based Dielectrics nMOSFETs*. IEEE Trans. Dev. Mater. Reliability, 2009. **9**(2).
95. Sato, M., Umezawa, N., Mise, N., *et al.*, *Physical model of the PBTI and TDDB of la incorporated HfSiON gate dielectrics with pre-existing and stress-induced defects*. IEDM Tech. Digest, 2008: p. 66-67.
96. Umezawa, N., Shiraishi, K., Sugino, S., *et al.*, *Suppression of oxygen vacancy formation in Hf-based high-k dielectrics by lanthanum incorporation*. Appl. Phys. Lett. , 2007. **91**(13): p. 132904.
97. Strukov, D.B., Snider, G.S., Stewart, D.R., *et al.*, *The missing memristor found*. Nature, 2008. **453**(7191): p. 80-3.

98. Hickmott, T.W., *Low - Frequency Negative Resistance in Thin Anodic Oxide Films*. J. Appl. Phys. , 1962. **33**: p. 2669.
99. Waser, R., Dittmann, R., Staikov, G., *et al.*, *Redox-Based Resistive Switching Memories - Nanoionic Mechanisms, Prospects, and Challenges*. Adv. Mater., 2009. **21**(25-26): p. 2632-2663.
100. Goux, L., Chen, Y.Y., Pantisano, L., *et al.*, *On the Gradual Unipolar and Bipolar Resistive Switching of TiN\HfO₂\Pt Memory Systems*. Electrochem. Solid-State Lett., 2010. **13**(6): p. G54.
101. Yu, S., Gao, B., Dai, H., *et al.*, *Improved Uniformity of Resistive Switching Behaviors in HfO₂ Thin Films with Embedded Al Layers*. Electrochem. Solid-State Lett., 2010. **13**(2): p. H36.
102. Chen, B., Gao, B., Fu, Y.H., *et al.*, *Co-existed unipolar and bipolar resistive switching effect of HfO_x-based RRAM*. 2010.
103. Goux, L., Wang, X.P., Chen, Y.Y., *et al.*, *Roles and Effects of TiN and Pt Electrodes in Resistive-Switching HfO₂ Systems*. Electrochem. Solid-State Lett., 2011. **14**(6): p. H244.
104. Hota, M.K., Mahata, C., Mallik, S., *et al.*, *Impact of Top (Pt, Au, and Al) Electrodes on HfAlO_x-Based MIM Capacitors*. J. Electrochem. Soc., 2011. **158**(1): p. H44.
105. Cabout, T., Buckley, J., Cagli, C., *et al.*, *Role of Ti and Pt electrodes on resistance switching variability of HfO₂-based Resistive Random Access Memory*. Thin Solid Films, 2013. **533**: p. 19-23.
106. Clima, S., Chen, Y.Y., Degraeve, R., *et al.*, *First-principles simulation of oxygen diffusion in HfO_x: Role in the resistive switching mechanism*. Appl. Phys. Lett. , 2012. **100**(13): p. 133102.
107. Wang, Z.R. and Zhu, W.G., *Highly uniform, self-compliance, and forming-free ALD HfO₂-based RRAM with Ge doping*. IEEE Trans. Electron Devices, 2012. **59**(4).
108. Gao, B., Kang, J.F., Chen, Y.S., *et al.*, *Oxide-based RRAM Unified microscopic principle for both unipolar and bipolar switching*. IEDM Tech. Digest, 2011. **11**.
109. Slater, J.C., *A Simplification of the Hartree-Fock Method*. Phys. Rev., 1951. **81**: p. 385.
110. Latter, R., *Atomic Energy Levels for the Thomas-Fermi and Thomas-Fermi-Dirac Potential*. Phys. Rev., 1955. **99**: p. 510.
111. Kohn, W. and Sham, L.J., *Self-Consistent Equations Including Exchange and Correlation Effects*. Phys. Rev. , 1965. **140**: p. A1133.

112. Martin, R.M., *Electronic Structure: Basic Theory and Practical Methods*. 2008: Cambridge University Press.
113. Van de Walle, C.G., *First-principles calculations for defects and impurities: Applications to III-nitrides*. J. Appl. Phys. , 2004. **95**(8): p. 3851.
114. Becke, A.D., *Density - functional thermochemistry. III. The role of exact exchange*. J. Chem. Phys., 1993. **98**: p. 5648.
115. Becke, A.D., *Density - functional thermochemistry. II. The effect of the Perdew - Wang generalized - gradient correlation correction*. J. Chem. Phys., 1992. **97**: p. 9173.
116. Becke, A.D., *Density-functional thermochemistry. I. The effect of the exchange-only gradient correction*. J. Chem. Phys., 1992. **96**: p. 2155.
117. Alkauskas, A., Broqvist, P. and Pasquarello, A., *Defect levels through hybrid density functionals: Insights and applications*. physica status solidi (b), 2011. **248**(4): p. 775-789.
118. Van de Walle, C.G. and Neugebauer, J., *Universal alignment of hydrogen levels in semiconductors, insulators and solutions*. Nature, 2003. **423**: p. 626-628.
119. Deal, B.E., Sklar, M., Grove, A.S., et al., *CHARACTERISTICS OF SURFACE-STATE CHARGE (QSS) OF THERMALLY OXIDIZED SILICON*. J. Electrochem. Soc., 1967. **114**(3): p. 266-&.
120. Kimizuka, N., Yamaguchi, K., Imai, K., et al., *NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10- μ m gate CMOS generation*. Symp. VLSI Tech. Tech. Dig, 2000: p. 92-93.
121. Ang, D.S., Wang, S. and Ling, C.H., *Evidence of two distinct degradation mechanisms from temperature dependence of negative bias stressing of the ultrathin gate p-MOSFET*. IEEE Electron Dev. Lett., 2005. **26**(12): p. 906-908.
122. Ang, D.S. and Wang, S., *On the non-Arrhenius behavior of negative-bias temperature instability*. Appl. Phys. Lett. , 2006. **88**(9).
123. Ang, D.S. and Pey, K.L., *Evidence for two distinct positive trapped charge components in NBTI stressed p-MOSFETs employing ultrathin CVD silicon nitride gate dielectric*. IEEE Electron Dev. Lett., 2004. **25**(9): p. 637-639.
124. Vattikonda, R., Wang, W. and Cao, Y., *Modeling and minimization of PMOS NBTI effect for robust nanometer design*. Proc. of Design Automation Conference, 2006: p. 1047-1052.
125. Denais, M., Huard, V., Parthasarathy, C., et al., *Interface Trap Generation and Hole Trapping Under Interface Trap Generation and Hole Trapping Under*. IEEE Trans. Dev. Mater. Reliability, 2004. **4**(4): p. 715-722.

126. Huard, V., Monsieur, F., Ribes, G., *et al.*, *Evidence for hydrogen-related defects during NBTI stress in p-MOSFETs*. Proc. Int. Reliab. Phys. Symp., 2003: p. 178-182.
127. Du, G.A., Ang, D.S., Teo, Z.Q., *et al.*, *Ultrafast Measurement on NBTI*. IEEE Electron Dev. Lett., 2009. **30**(3): p. 275-277.
128. Lelis, A.J. and Oldham, T.R., *Time dependence of switching oxide traps*. IEEE Trans. Nucl. Sci., 1994. **41**(6): p. 1835-1843.
129. Oldham, T.R., Lelis, A.J. and McLean, F.B., *Spatial Dependence of trapped holes determined from tunneling analysis and measured annealing*. IEEE Trans. Nucl. Sci., 1986. **33**(6): p. 1203-1209.
130. Dahinden, P., Broqvist, P. and Pasquarello, A., *Charge transition levels of nitrogen dangling bonds at Si/SiO₂ interfaces: A first-principles study*. Phys. Rev. B, 2010. **81**(8).
131. Yokozawa, A. and Miyamoto, Y., *First-principles calculations for charged states of hydrogen atoms in SiO₂*. Phys. Rev. B, 1997. **55**(20): p. 13783-13788.
132. Giannozzi, P., Baroni, S., Bonini, N., *et al.*, *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*. J. Phys.: Condens. Matter, 2009. **21**(39).
133. Perdew, J.P., Burke, K. and Ernzerhof, M., *Generalized Gradient Approximation Made Simple*. Phys. Rev. Lett., 1996. **77**(18): p. 3865-3868.
134. Zhao, C.Z. and Zhang, J.F., *Effects of hydrogen on positive charges in gate oxides*. J. Appl. Phys. , 2005. **97**(7).
135. Godet, J., Giustino, F. and Pasquarello, A., *Proton-induced fixed positive charge at the Si(100)-SiO₂ interface*. Phys. Rev. Lett., 2007. **99**(12).
136. Ang, D.S., Lai, S.C.S., Du, G.A., *et al.*, *Effect of Hole-Trap Distribution on the Power-Law Time Exponent of NBTI*. IEEE Electron Dev. Lett., 2009. **30**(7): p. 751-753.
137. Kresse, G. and Hafner, J., *Ab initio molecular dynamics for liquid metals*. Phys. Rev. B, 1993. **47**(558-561).
138. Sasaki, H., Ono, M., Yoshitomi, T., *et al.*, *1.5 nm direct-tunneling gate oxide Si MOSFET's*. IEEE Trans. Electron Devices, 1996. **4**(8): p. 1233-1242.
139. Wallace, R.M. and Wilk, G.D., *High- κ dielectric materials for microelectronics*. Crit. Rev. Solid. State. Mater. Sci., 2003. **28**(4): p. 231.
140. Cho, M., Aoulaiche, M., Degraeve, R., *et al.*, *Positive and negative bias temperature instability on sub-nanometer eot high-K MOSFETs*. Proc. Int. Reliab. Phys. Symp., 2010: p. 1095-1098.

141. Lee, K.T., Kang, W., Chung, E.-A., *et al.*, *Technology scaling on High-K & Metal-Gate FinFET BTI reliability*. Proc. Int. Reliab. Phys. Symp., 2013: p. 2D.1.1-2D.1.4.
142. Boo, A.A. and Ang, D.S., *Evolution of hole trapping in the oxynitride gate P-MOSFET subjected to negative-bias temperature stressing*. IEEE Trans. Electron Devices, 2012. **59**(11): p. 3133-6.
143. Gao, Y., Boo, A.A., Teo, Z.Q., *et al.*, *On the evolution of the recoverable component of the SiON, HfSiON and HfO₂ P-MOSFETs under dynamic NBTI*. Proc. Int. Reliab. Phys. Symp., 2011: p. XT.8.1-XT.8.6.
144. Grasser, T., Rott, K., Reisinger, H., *et al.*, *Advanced characterization of oxide traps: The dynamic time-dependent defect spectroscopy*. Proc. Int. Reliab. Phys. Symp., 2013: p. 2D.2.1-2D.2.7.
145. Hu, Y.Z., Ang, D.S. and Teo, Z.Q., *Threshold voltage and mobility extraction by ultrafast switching measurement on NBTI*. IEEE Trans. Electron Devices, 2010. **57**(8): p. 2027-2031.
146. Vanderbilt, D., *Soft self-consistent pseudopotentials in a generalized eigenvalue formalism*. Phys. Rev. B, 1990. **41**(11): p. 7892-7895.
147. Scopel, W.L., Da Silva, A.J.R. and Fazzio, A., *Amorphous HfO₂ and Hf_{1-x}Si_xO via a melt-and-quench scheme using ab initio molecular dynamics*. Phys. Rev. B, 2008. **77**(17): p. 172101.
148. Cartier, E. and Kerber, A., *Stress-induced leakage current and defect generation in nFETs with HfO₂/TiN gate stacks during positive-bias temperature stress*. Proc. Int. Reliab. Phys. Symp., 2009: p. 486-92.
149. Uchino, T. and Yoko, T., *Density functional theory of structural transformation of oxygen-deficient centers in silica during hole trapping: Structure and formation mechanism of the E' center*. Phys. Rev. B, 2006. **74**(12): p. 125203.
150. Kimmela, A., Peter, S., Shlugerb, A., *et al.*, *Positive and Negative Oxygen Vacancies in Amorphous Silica*. ECS Trans., 2009. **19**(2): p. 3-17.
151. Bersuker, G., Park, C.S., Barnett, J., *et al.*, *The effect of interfacial layer properties on the performance of Hf-based gate stack devices*. J. Appl. Phys. , 2006. **100**: p. 094108.
152. Ioannou, D.P., Mittl, S. and La Rosa, G., *Positive bias temperature instability effects in nMOSFETs with HfO₂/TiN gate stacks*. IEEE Trans. Dev. Mater. Reliability, 2009.
153. Scopel, W.L., A.J.R. da Silva and Fazzio, A., *Amorphous HfO₂ and Hf_{1-x}Si_xO via a melt-and-quench scheme using ab initio molecular dynamics*. Phys. Rev. B, 2008. **v. 77**: p. 172101.

154. Xiong, K., Du, Y., Tse, K., *et al.*, *Defect states in the high-dielectric-constant gate oxide HfSiO₄*. J. Appl. Phys. , 2007. v. **101**: p. 024101.
155. Gu, C.J., Ang, D.S. and Teo, Z.Q., *Impact of twofold coordinated nitrogen on the generation of deep-level hole traps under negative-bias temperature stressing*. ECS Trans., 2011. **34**(4): p. 125-43.
156. Wang, X.P., Shen, C., Li, M.F., *et al.*, *Dual Metal Gates with Band-Edge Work Functions on Novel HfLaO High-K Gate Dielectric*. Symp. VLSI Tech. Tech. Dig, 2006: p. 9-10.
157. Sivasubramani, P., Boscke, T.S. and Huang, J., *Dipole Moment Model Explaining nFET V_i Tuning Utilizing La, Sc, Er, and Sr Doped HfSiON Dielectrics*. Symp. VLSI Tech. Tech. Dig, 2007: p. 68-69.
158. Yamamoto, Y., Kita, K., Kyuno, K., *et al.*, *Study of La-Induced Flat Band Voltage Shift in Metal/HfLaOx/SiO₂/Si Capacitors*. Japanese J. Appl. Phys. , 2007. v. **46**: p. 7251-7255.
159. Kamiyama, S., Miura, T. and Kurosawa, E., *Band Edge Gate First HfSiON/Metal Gate n-MOSFETs using ALD-La₂O₃ Cap LAyers Scalable to EOT=0.68 nm for hp 32 nm BULK Device with High Performance and Reliability*. IEDM Tech. Digest, 2007.
160. Kaczer, B., Veloso, A., Aoulaiche, M., *et al.*, *Significant reduction of Positive Bias Temperature Instability in high-k/metal-gate nFETs by incorporation of rare earth metals*. Microelectronic Engineering, 2009. v. **86**(7-9): p. 1894-1896.
161. T. R. Oldham, A. J. Lelis and McLean, F.B., *Spatial Dependence of Trapped Holes Determined from Tunneling Analysis and Measured Annealing*. IEEE Trans. Nucl. Sci., 1986. v. **NS-33**: p. 1203-1209,.
162. Liu, D. and Robertson, J., *Passivation of oxygen vacancy states and suppression of Fermi pinning in HfO₂ by La addition*. Appl. Phys. Lett. , 2009. v. **94**(4): p. 042904.
163. Kang, C.Y., Young, C.D., Huang, J., *et al.*, *The impact of La-doping on the reliability of low V_{th} high-k/metal gate nMOSFETs under various gate stress conditions*. IEDM Tech. Digest, 2008: p. 1-4.
164. Kang, C.Y., Park, C.S., Heh, D., *et al.*, *Performance and reliability characteristics of the band edge high-k/metal gate nMOSFETs with La-doped Hf-silicate gate dielectrics*. Proc. Intl. Reliab. Phys. Symp., 2008: p. 663-664.
165. Larcher, L., Padovani, A., Pirrotta, O., *et al.*, *Microscopic understanding and modeling of HfO₂ RRAM device physics*. IEDM Tech. Digest, 2012. **12**: p. 475.
166. Mills, G., Jonsson, H. and Schenter, G.K., *Reversible work transition state theory: application to dissociative adsorption of hydrogen*. Surf. Sci., 1995. **324**(2-3): p. 305-337.

- 167. Foster, A.S., Shluger, A.L. and Nieminen, R.M., *Mechanism of Interstitial Oxygen Diffusion in Hafnia*. Phys. Rev. Lett., 2002. **89**(22).
- 168. Shluger, A.L. and McKenna, K.P., *Models of oxygen vacancy defects involved in degradation of gate dielectrics*. Proc. Int. Reliab. Phys. Symp., 2013: p. 5A.1.1-5A.1.9.
- 169. Ribes, G., Mitard, J., Denais, M., *et al.*, *Review on high-k dielectrics reliability issues*. IEEE Trans. Dev. Mater. Reliability, 2005. **5**(1): p. 5-19.
- 170. Lin, K.-L., Hou, T.-H., Shieh, J., *et al.*, *Electrode dependence of filament formation in HfO₂ resistive-switching memory*. Journal of Applied Physics, 2011. **109**(8): p. 084104.
- 171. Bersuker, G., Heh, D., Young, C.D., *et al.*, *Mechanism of high-k dielectric-induced breakdown of the interfacial SiO₂ layer*. Proc. Int. Reliab. Phys. Symp., 2010: p. 373-378.