The Dissertation Committee for Yen-Ting Chen Certifies that this is

the approved version of the following dissertation:

A Study of Electrical and Material Characteristics of High-k / III-V MOSFETs and SiO_2 RRAMs

Committee:

_____
Jack C. Lee, Supervisor

_____
Ananth Dodabalapur

_____
Dean Neikirk

_____
Emanuel Tutuc

_____
Chih-Kang Ken Shih

A Study of Electrical and Material Characteristics of High-k / III-V MOSFETs and SiO_2 RRAMs

by

Yen-Ting Chen, B.B.A.; B.S.E.;M.S.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2012

Dedicated to

My wife and my family

Acknowledgements

I would like to gratefully and sincerely thank Prof. Jack C. Lee for his guidance, constant support, encouragement, and most importantly, his friendship throughout my doctoral research at the University of Texas at Austin. His mentorship was providing a free and supportive research environment so that I was able to develop my own individuality and merit. I truly appreciate the Ph. D. training under his supervision, which not only enriches my professional knowledge, but also polish me to become an independent researcher.

I would like to acknowledge my committee members, Prof. Ananth Dodabalapur, Prof. Dean Neikirk, Prof. Emanuel Tutuc, and Prof. Chih-Kang Ken Shih for their valuable comments and suggestions to strengthen this work.

I would also like to express my gratitude to our collaborators in SEMATECH. In particular, I would like to thank Dr. Paul Kirsch, Dr. Niti Goel, Dr. Jeff Huang, Dr. Wei-E Wang, and Dr. Wei-Yip Loh for their beneficial discussions and insights that guided me to explore some interesting topics of high-k/III-V gate stacks.

It would not have been possible to write this doctoral thesis without the help and support of our group members. I really appreciate the help and suggestions from our senior group members, Dr. Injo Ok, Dr. Hyoung-Sub Kim, Dr. Feng Zhu, and Dr. Han Zhao. Even though I had a short overlap with Injo, Hyoung-Sub, and Feng, they kindly shared their experience and ideas with me when I just joined this group. My special thanks to Han, who not only trained me on the device fabrication and characterization, but also gave me continuous inspired suggestions and encouragement. I also appreciate

all the helpful support from the current group members, Yanzhen Wang, Fei Xue, Fei Zhou, Yao-Feng Chang, and Pai-Yu Chen. Also, I remember the enjoyable moments in working with all MER students, especially, Chen-Guan, En-Shao, Che-Yun, Ping-Chun, Wei-Cheng, William, Sherry, Jong-Yeon, and Michael for their sincere friendship, genuine caring and earnest advice.

I also appreciate the support from PrivaTran, especially Glenn Mortland, Burt Fowler, and Brad Bringhurst for their valuable discussion and financial support.

I am also grateful to the entire MER staff, especially Jean Toll, Domingo Ferrer, Bill Ostler, Jesse James, Gabriel Glenn, Joyce Kobes, and Gerlinde Sehne for their assistance and support.

Finally, I would like to express my deepest gratitude to my family for their everlasting love. My wife, Nikki Wen, has been standing beside me since the first day we met. She has been my support and inspiration for my life. Our two sons' smiles are my best gifts ever. My parents and my parents-in-law have been our financial and mental support. I also thank Bernie and Sally for leading me to God. Last but not least, I thank God who created me, guided me, blessed me, and has always been with me.

A Study of Electrical and Material Characteristics of High-k / III-V MOSFETs and SiO_2 RRAMs

Yen-Ting Chen, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Jack C. Lee

Aggressive downscaling of complementary metal-oxide-semiconductor (CMOS) transistors has pushed Si-based transistors to their limit. III-V materials have much higher electron mobility compared to Si, which can potentially provide better device performance. Therefore, III-V semiconductor materials have been actively investigated as alternative channel materials, which can extend Moore's law on CMOS scaling beyond the 22 nm node not only by relying on scaling. Meanwhile, conventional silicon dioxide cannot easily meet the requirement for the scaling of the equivalent oxide thickness; as a result, various high dielectric constant (high-k) materials have been incorporated onto the III-V semiconductor substrate. Nevertheless, the key challenges for high-k/III-V MOSFETs still need to be solved in order to implement high performance high-k/III-V MOSFETs. Those challenges are the lack of high quality and thermodynamically stable insulators that passivate the gate dielectric/III-V interface, compatible III-V p-type MOSFETs, and reliability issue of III-V MOSFETs, etc.

The main focus of this dissertation is to develop proper fabrication processes and structures for III-V MOSFETs devices that result in good interface quality and high device performance. Firstly, we studied the effect of interfacial chemistry on ZrO_2/InGaAs gate stack comprehensively, comparing ALD ZrO_2 with H_2O vs. O_3 as the oxidizer. We found that the amount of oxygen is critical to form a good interface. Excessive oxygen concentration, e. g. using O_3 as the ALD precursor, induces III-V native oxides at the interface.

The second part of this dissertation focuses on the III-V MOSFETs with various IPLs. Various IPLs have been demonstrated, for example, a thin PVD Si IPL, and ALD Al_2O_3, $HfAlO_x$, and $ZrAlO_x$. Those IPLs are demonstrated to be effective interfacial dielectric layers to improve device performance, including frequency dispersion, SS, I_{on}, effective channel mobility, and reliability.

The third part of this study highlights a novel CF_4 post-gate plasma treatment on III-V MOSFETs. Fluorine incorporation was demonstrated on various high-k/III-V gate stacks and achieved significant improvements, including $Al_2O_3/In_{0.53}Ga_{0.47}As$, Al_2O_3/InP, $HfO_2/In_{0.53}Ga_{0.47}As$, and HfO_2/InP. Detailed physical analysis, electrical characterization and device performance were carried out. With F incorporation, we have successfully developed excellent interface quality of high-k/III-V MOSFETs. As a result, high-performance III-V MOSFETs have been realized.

Finally, emerging non-volatile memories, RRAMs, have been demonstrated. We addressed its conducting mechanism by conducting various experiments and purposed a model for SiO_x RRAMs: the conducting filament is randomly formed within the SiO_x at

the sidewall edge, depending on pre-existing defects. Moreover, the rupture/recovery could occur anywhere along the conducting filament, depending on a random process that determines the location of the weak spot along the conducting filament. In addition, we improved SiO_2-based RRAM by incorporating a thin silicon layer onto its sidewall. This technique significantly reduced the electroforming voltage and instability of HRS current of SiO_2-based RRAMs. Consequently, a tri-state pulse endurance performance over 10^6 cycles has been demonstrated and the data stored had good read disturb immunity and thermal disturbance.

Table of Contents

List of Tables xii

List of Figures xiii

Chapter 1 Introduction 1

1.2 III-V materials 5

1.3 An emerging non-volatile memories -- RRAM 9

1.4 Outline 11

Chapter 2 Investigation of H_2O vs. O_3 based high-k on $In_{0.53}Ga_{0.47}As$ 16

2.1 III-V Gate Stack Interface Improvement Using H_2O Oxidation of ZrO_2 and TiN Cap Layers 16

2.2 Comparison of H_2O vs. O_3 based high-k on InGaAs 23

2.3 Summary 28

Chapter 3 Interfacial passivation layers on InP and $In_{0.53}Ga_{0.47}As$ MOSFETs 30

3.1 PVD Si IPL on InP MOSFETs with ALD HfO_2 30

3.2 Various IPL on $In_{0.53}Ga_{0.47}As$ MOSFETs with ALD ZrO_2 37

3.3 Summary 44

Chapter 4 Fluorine Incorporation into high-k/III-V MOSFETs for enhanced device characteristics 46

4.1 Motivation 46

4.2 Improved electrical characteristics of TaN / Al_2O_3 / $In_{0.53}Ga_{0.47}As$ MOSFETs by F incorporation 46

4.3 Effects of fluorine incorporation on the electrical properties of atomic-layer-deposited Al_2O_3 gate dielectric on InP substrate 55

4.4 Optimization of Fluorine Plasma Treatment for Interface Improvement on HfO_2/$In_{0.53}Ga_{0.47}As$ MOSFETs 68

4.5 Fluorinated HfO_2 gate dielectric engineering on $In_{0.53}Ga_{0.47}As$ metal-oxide-semiconductor field-effect-transistors 87

4.6 Fluorine incorporation into HfO_2 gate dielectric InP metal-oxide-semiconductor field-effect-transistors 94

4.7 Summary 98

Chapter 5 SiO_x-based resistive switching random access memories 100

5.1 Fundamental understanding of SiO_x-based RRAMs 100

5.2 Working principle of SiO_x-based RRAMs 101

5.3 Size effects and oxide thickness effects on SiO_x-based RRAMs 103

5.4 Conducting filaments formation and switching mechanism of SiO_x-based RRAMs 107
5.5 Reduced electroforming voltage and enhanced programming stability in resistive switching of SiO_2 thin films 121
5.6 Summary 131
Chapter 6 Summary and future work 133
6.1 Summary 133
6.2 Suggestions for future work 135
6.2.1 Surface channel III-V MOSFETs 135
6.2.2 SiO_2-based RRAMs 136
Bibliography 137

List of Tables

Table 1.1 Logic potential solutions from ITRS 2011 3

Table 1.2 Comparison of HP, LOP, LSTP, and III-V/Ge Technologies 4

Table 1.3 Comparison of various III-V compounds and silicon 7

Table 1.4 Comparison of emerging non-volatile memories with flash memories 10

Table 2.1 Comparison of O-Al_2O_3, H-Al_2O_3, Lo-O-ZrO_2, and H-ZrO_2. 29

Table 3.1 Effects of Si IPL on MOSCAPs and MOSFETs characteristics..... 36

Table 4.1 Effects of post-gate CF_4 plasma treatment on $In_{0.53}Ga_{0.47}As$ nMOSFET 54

Table 4.2 Comparison of the electrical and interfacial properties of this work with some recently reported paper. 86

Table 4.3 Comparison of the effects of F incorporation on 4 different gate stacks. 99

Table 5.1 Summary of devices with different shapes, area and perimeter..... 103

List of Figures

Figure 1.1 CMOS scaling trend from 2003 to 2011 and beyond 2

Figure 1.2 Example of non-classical CMOS structures. Combination of III-V and Ge channel on Si substrate. 3

Figure 1.3 (Left) Measured drive current versus drain voltage comparing the 80 nm $In_{0.7}Ga_{0.3}As$ QWFET and 40nm Strained Si MOSFET at matched DIBL, SS and Ioff. $In_{0.7}Ga_{0.3}As$ QWFET exhibits 20-30% drive current gain over Strained Si MOSFET for a wide range of V_{GS}-V_T. (Right) Measured drive current versus operating voltage (V_{cc}) comparing 80 nm $In_{0.7}Ga_{0.3}As$ QWFET and 40nm Strained Si MOSFET for constant Ioff = 100 nA/μm. Despite 2.5X lower capacitance and 60% higher R_S+R_D, $In_{0.7}Ga_{0.3}As$ QWFET exhibits drive current gain over strained Si MOSFET for all V_{cc}, with greater gains at lower V_{cc}..... 7

Figure 1.4 Comparison of InGaAs MOSFET [13] and InP MOSFET [14] with the identical high-k gate dielectric. 8

Figure 2.1 Multi-frequency CV curves of (a) O-ZrO_2, (b) Lo-O-ZrO_2, and (c) H-ZrO_2 (8 nm thick, all with a TiN cap layer). (d): Normalized 1MHz CV comparison of (a), (b), and (c) 19

Figure 2.2 Photoemission data for H-ZrO_2 and Lo-O-ZrO_2 (w/o TiN capping). AsO_x, GaO_x, and In_2O_3 are detected with Lo-O-ZrO_2 whereas no native oxides are detected with H-ZrO_2..... 19

Figure 2.3 Photoemission data for TiN/Lo-O-ZrO_2 and TiN/H-ZrO_2. 20

Figure 2.4 HRTEM of O-ZrO_2/$In_{0.53}Ga_{0.47}As$ (w/o TiN cap). Tracing the crystal lattice, an amorphous layer between crystalline ZrO_2 and $In_{0.53}Ga_{0.47}As$ layer is identified as a III-V native oxide layer, which may consist of AsO_x, GaO_x, and In_2O_3. 21

Figure 2.5 HRTEM image of TiN/O-ZrO_2/InGaAs. No interfacial layer is observed, which confirms that TiN gettering of AsO_x, GaO_x and In_2O_3 bonding is associated with the reduction of the interfacial layer. 22

Figure 2.6 Photoemission data of TiN/O-Al_2O_3 and TiN/H-Al_2O_3..... 22

Figure 2.7 (a) CV Hysteresis of H-Al_2O_3 and O-Al_2O_3. Less hysteresis of the sample with H-Al_2O_3, even under higher electric field, suggests lesser bulk traps in H-Al_2O_3. (b) Similar improvement in H-ZrO_2. 24

Figure 2.8 Charge pumping spectra of decreasing gate pulse frequencies. 24

Figure 2.9 Mean D_{it} of O-Al_2O_3, H-Al_2O_3, Lo-O-ZrO_2, and H-ZrO_2. (a) and (b) : Charge pumping method with varying t_R/t_F time of the gate pulse under 200KHz. H_2O-based oxides has less charge pumping current compared to O_3-based oxides. (c) : mean D_{it} fitting according. (d): Mean D_{it} values of the 4 high-k oxides. 26

Figure 2.10 Improvements in normalized I_d and G_m and SS confirm that H_2O-based precursor chemistry is effective for InGaAs-based devices. (I_{off} is similar)..... 27

Figure 2.11 Normalized I_d improvements are attributed to better high-k/III-V gate stack quality of H_2O-based oxides. 27

Figure 3.1 Hysteresis and frequency dispersion behavior of InP MOSCAPs with different gate dielectrics: (a) 10Å Si IPL /51Å HfO_2 stacked gate dielectric. Inset: its hysteresis characteristics. (b) 70Å HfO_2 without Si IPL. Inset: its hysteresis characteristics. Device area:$1.23x10^{-4}cm^2$. 33

Figure 3.2 Drain currents and extrinsic transconductance vs gate voltage. 34

Figure 3.3 I_d-V_d characteristics as a function of V_g at $V_g - V_t$ = 0, 0.3 V, 0.6 V, 0.9 V, and 1.2 V for In MOSFETs with 70Å HfO_2 (open) and 10Å Si IPL /51Å HfO_2 stacked gate dielectric (solid). 34

Figure 3.4 Gate leakage current density vs gate voltage for InP MOSFETs with 70Å HfO_2 (circle) and 10Å Si IPL/51Å HfO_2 stacked gate dielectric (square). Inset: D_{it} comparison extracted by full conductance method. 35

Figure 3.5 Frequency dispersion of MOSCAPs on InP substrate with different gate dielectrics: (a) single ZrO_2, (b) Al_2O_3/ZrO_2, (c) $HfAlO_x/ZrO_2$, (d) $ZrAlO_x/ZrO_2$. Frequency dispersion is calculated by the capacitance difference between 1MHz and 10 KHz at gate voltage=1.0 V. 39

Figure 3.6 Interface trap density (D_{it}) distribution of: a single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks. The D_{it} distribution is extracted by conductance method. 40

Figure 3.7 Log scale I_d-V_g charactristic at V_d=50mV of InP with a single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$. 41

Figure 3.8 Effective channel mobility versus invertion charge density for InP MOSFETs with different gate dielectrics: single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$. 42

Figure 3.9 I_d versus V_d as a function of V_g at V_g-V_{th} from 0V to 2V with 0.5V step. MOSFETs with different gate stacks (W/L=600μm/5μm): (a) single ZrO_2, (b) Al_2O_3/ZrO_2, (c) $HfAlO_x/ZrO_2$, (d) $ZrAlO_x/ZrO_2$. I_d at V_d = 2.5V,V_g-V_{th} =2V increases about four times with insertion of passivation layers. 43

Figure 3.10 Reliability performance with a constant stress of effect electrical field 10MV/cm on InP MOSFETs with single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ (a)V_{th} shift and (b) SS increase. 44

Figure 4.1 AFM images on (a) Al_2O_3 / $In_{0.53}Ga_{0.47}As$ (b) Al_2O_3 / $In_{0.53}Ga_{0.47}As$ after CF_4 plasma treatment 3min. (c) surface roughness RMS values measured at different steps. 49

Figure 4.2 SIMS depth profile of Al_2O_3 / $In_{0.53}Ga_{0.47}As$ structure for fluorine, oxygen, and aluminum, atoms distribution. Red solid line: w/o F plasma. Red dash line: w/ F plasma. 50

Figure 4.3 Drain currents and extrinsic transconductance versus gate voltage. 51

Figure 4.4 Effective channel mobility vs. inversion charge density for $In_{0.53}Ga_{0.47}As$ MOSFETs w/ F plasma treatment (solid) and w/o F plasma treatment (open). Inset (a): split-CV hysteresis for the samples w/ F plasma treatment. Inset (b): split-CV hysteresis for the samples w/o F plasma treatment. 51

Figure 4.5 I_d-V_d characteristics as a function of various $V_g - V_t$ for $In_{0.53}Ga_{0.47}As$ MOSFETs w/ F plasma treatment (solid) and w/o F plasma treatment (open). 52

Figure 4.6 D_{it} distribution for $In_{0.53}Ga_{0.47}As$ MOSFETs w/ F plasma treatment (solid) and w/o F plasma treatment (open). 53

Figure 4.7 Maximum extrinsic transconductance (G_m) and drain current as a function of CF_4 plasma rf power (W/L= 600 μm/20 μm at V_d = 50 mV). CF_4 plasma treatment time: 3 min. 58

Figure 4.8 Effective channel mobility versus inversion charge density as a function of CF_4 plasma rf power. CF_4 plasma treatment time: 3 min. 59

Figure 4.9 SS and EOT as a function of CF_4 plasma rf power. CF_4 plasma treatment time: 3 min. 59

Figure 4.10 I_d-V_d characteristics as a function of various $V_g - V_{th}$ for InP MOSFETs w/o F plasma treatment (solid) and w/ F plasma treatment 20W 3min (open)..... 60

Figure 4.11 SIMS depth profile of the Al_2O_3 / InP structure for fluorine, oxygen, and aluminum atoms distribution. Red solid line: w/o F plasma. Square: w/ F plasma 20W for 3 min. Triangle: w/ F plasma 20W for 5 min. 61

Figure 4.12 Maximum extrinsic transconductance (G_m) and drain current as a function of CF_4 plasma treatment time (W/L= 600 μm/20 μm at V_d = 50 mV). rf power : 20W. 62

Figure 4.13 Maximum effective channel mobility as a function of CF_4 plasma treatment time. rf power : 20W. 62

Figure 4.14 SS and EOT as a function of CF_4 plasma treatment time. rf power : 20W. 63

Figure 4.15 The drain current and extrinsic transconductance (G_m) versus $V_g - V_{th}$ as a function of CF_4 plasma treatment time. rf power : 20W..... 64

Figure 4.16 Effective channel mobility versus inversion charge density as a function of CF_4 plasma treatment time. rf power : 20W. 64

Figure 4.17 I_d-V_d characteristics as a function of various $V_g - V_{th}$ for InP MOSFETs w/o F plasma treatment, w/ F plasma treatment 20W 3 min and 20W 5 min. 64

Figure 4.18 D_{it} distribution for InP MOSFETs w/o F plasma treatment and w/ F plasma treatment 10-40W. CF_4 plasma treatment time: 3 min. 65

Figure 4.19 D_{it} distribution for InP MOSFETs w/o F plasma treatment and w/ F plasma treatment 1 min, 3 min and 5 min. rf power : 20W. 66

Figure 4.20 V_{FB} versus EOT for the Al_2O_3/InP stack w/o F plasma treatment and w/ F plasma treatment 3 min and 5 min. rf power : 20W. 67

Figure 4.21 The device cross-sectional structure of HfO_2/$In_{0.53}Ga_{0.47}As$ gate stack. Samples were treated in a mixed flow of CF_4 and O_2 gas (ratio ~ 10:1) with varied RF power and treatment time. 72

Figure 4.22 XPS analysis of the Hf 4*f* electronic spectra for the samples with and without F incorporation. 73

Figure 4.23 XPS analysis of the F 1*s* electronic spectra for the control sample and fluorinated sample..... 74

Figure 4.24 SIMS depth profiles of F, S, Hf, O, In, Ga, and As in the fluorinated HfO_2/$In_{0.53}Ga_{0.47}As$ gate stack w/ and w/o PDA. Movement of F atoms is clearly observed. F atoms piled up at the HfO_2/$In_{0.53}Ga_{0.47}As$ interface after PDA. The control sample with PDA only is also presented..... 74

Figure 4.25 EOT versus physical thickness for the samples with and without F incorporation. Inset: the gate leakage current comparison of samples with and without F incorporation..... 75

Figure 4.26 Maximum G_m and I_d as a function of CF_4 plasma RF power (W/L= 600 μm/20 μm at V_d = 50 mV and 0.5V). CF_4 plasma treatment time: 3 min..... 76

Figure 4.27 SS as a function of CF_4/O plasma RF power. CF_4 plasma treatment time: 3 min. Inset: I_d-V_g comparison of the control sample and the sample with CF_4 plasma treatment for 30W/3min. 77

Figure 4.28 Maximum G_m and I_d as a function of CF_4 plasma treatment time (W/L= 600 μm/20 μm at V_d = 50 mV and 0.5 V). RF power: 30W..... 77

Figure 4.29 SS as a function of CF_4 plasma treatment time. RF power : 30W. Inset: I_d-V_g comparison of the control sample and the sample with CF_4 plasma treatment for 30W/5min. 78

Figure 4.30 Effective channel mobility versus inversion charge density as a function of CF_4 plasma treatment time. RF power: 30W..... 78

Figure 4.31 Split-CV curves of the fluorinated sample. Frequency dispersion at V_g = 1V is reduced from 12.7% to 5.7% with F treatment. Inset: Split-CV curves of MOSFETs 79

Figure 4.32 I_d comparison under high field (V_d = 0.5 V) and low field (V_d = 50 mV) for the sample w/ and w/o F incorporation. Significant improvements have been achieved in I_d and SS with F incorporation. Inset: their corresponding G_m curves under high field..... 79

Figure 4.33 I_d-V_d comparison at $V_g = V_{th}$ to V_{th} + 2 V for the control and the fluorinated sample. 80

Figure 4.34 Series resistance extracted from the total resistance versus gate length at V_g - V_{th} = 0.5, 1, and 1.5 V for devices with 600 μm width. 80

Figure 4.35 DC and pulsed I_d-V_g at V_d=50mV w/ and w/o F incorporation. Smaller I_d-V_g hysteresis indicates less charge trapping after CF_4 plasma treatment. Inset: Configuration of pulsed I_d-V_g measurement. 82

Figure 4.36 Multiple frequency charge pumping spectra for the samples w/ and w/o F incorporation. Q_{br} can be scanned by reducing measured frequency to probe deeper into HfO_2 bulk..... 83

Figure 4.37 Charge pumping measurement with rise/fall time dependence. Samples with F incorporation show much smaller I_{cp}. V_{base}: -2.3 V to 1 V in a step of 50 mV and the pulse amplitude is 1 V. 85

Figure 4.38 Q_{cp} (=I_{cp}/f) vs. $\ln[(t_R \times t_F)^{1/2}]$. The mean D_{it} value is extracted by linear fitting according to [86 - 87]. Samples with F incorporation show less D_{it} value of 4.9×10^{11}/eVcm^2. 85

Figure 4.39 Cross section of the HfO_2 / $In_{0.53}Ga_{0.47}As$ MOSFETs for (a) the post-F sample, and (b) the mid-F sample. 89

Figure 4.40 SIMS depth profile of HfO_2 / $In_{0.53}Ga_{0.47}As$ structure. Straight line: the control sample, circle: the post-F sample, and square: the mid-F sample. 91

Figure 4.41 CV curves for the MOSCAPs of TaN / HfO_2 / $In_{0.53}Ga_{0.47}As$ structure. Filled square: the control sample, circle: the post-F sample, triangle: the mid-F sample. 91

Figure 4.42 I_d and G_m characteristics as a function of various $V_g - V_{th}$ for the control sample, the post-F sample, and the mid-F sample. Inset: the corresponding SS. 92

Figure 4.43 Effective channel mobility for the control sample, the post-F sample, and the mid-F sample. 93

Figure 4.44 I_d-V_d at $V_g = V_{th}$ to V_{th} + 2 V for the control sample, the post-F sample, and the mid-F sample. 94

Figure 4.45 Cross-sectional TEM images of TaN / 9.6 nm HfO_2 / InP gate stack (a) without postgate CF_4 plasma treatment and (b) with postgate CF_4 plasma treatment 30W/3min. (c) their corresponding F *1s* XPS spectrums. The scale bars in two TEM images represent 2 nm. 96

Figure 4.46 SS and hysteresis as a function of CF_4 plasma rf powers. CF_4 plasma treatment time: 3 min. Inset (a): hysteresis curves of the control sample, and inset (b): hysteresis curves of the sample with 30W CF_4 plasma treatment. 97

Figure 4.47 The normalized G_m of InP MOSFETs with different rf powers. CF_4 plasma treatment time: 3 min. 98

Figure 5.1 The electroforming process in a SiO_x memristor. Inset: device structure. 102

Figure 5.2 Typical LRS and HRS switching I-V curves for SiO_2-based RRAMs. 102

Figure 5.3 First voltage sweep of different size devices. 103

Figure 5.4 (a) Schematic cross section structure of capacitor type SiO_x RRAMs. (b) Electroforming process of forward and backward sweeps. The numbers aside the I-V curves indicate the corresponding sweep order. 104

Figure 5.5 V_{EF} as a function of SiO_x CETs and device sizes. (Big square: 540×540 μm; medium square: 340 ×340 μm; small square: 140 ×140 μm) 105

Figure 5.6 Hard breakdown of small square devices on 6 nm SiO_x. Inset: illustration of competing effect of hard breakdown and CF formation. 105

Figure 5.7 P_{EF} as a function of SiO_x CETs and device sizes. 106

Figure 5.8 On/off current as a function of SiO_x CETs and device 107

Figure 5.9 Set/reset power as a function of SiO_x CETs and device sizes. 107

Figure 5.10 (a) Device structure of 5 RRAMs in series. The arrows indicate current path from pad A to pad B. (b) Device cross-sectional structure. 108

Figure 5.11 Electroforming sequence of 5 RRAMs in A-B mode. The numbers aside the I-V curves indicate the corresponding sweep order. 109

Figure 5.12 Current status of deices 1 to 5 corresponding to each electroforming sweeps in Fig. 5.11. 109

Figure 5.13 I-V curves of the on/off switching process for the A-B mode. 110

Figure 5.14 Current status of deices 1 to 5 corresponding to each on/off switching in Fig. 5.13. 110

Figure 5.15 Schematic cross-section structure of SiO_x RRAMS. For A-B mode, voltage is applied on device A with device B grounded and voltage of backside (V_c) is monitored simultaneously. For A(B)-C mode, voltage is applied on device A(B) with backside C grounded. 111

Figure 5.16 Electroforming process for A-B mode. The numbers aside the I-V curves indicate the corresponding sweep order. After 5^{th} scan, devices are ready for RRAM operation. The inset shows the corresponding backside voltage (V_c) change during the electroforming process..... 113

Figure 5.17 I-V curves for (a) A-C mode and (b) B-C mode. They are corresponding to each electroforming sweep in Fig. 5.16. 113

Figure 5.18 I-V curves of the LRS/HRS switching process for the A-B mode. (a): It turns the A-B mode to LRS/HRS when sweeping to 5.5 V/10 V. The corresponding A-C and B-C mode are shown in Fig. 5.19 (a) and (b), respectively. (b): It turns the A-B mode to LRS/HRS when sweeping to -5.5 V/-10 V. The corresponding A-C and B-C mode are shown in Fig. 5.19 (c) and (d), respectively. 114

Figure 5.19 I-V curves for A-C mode and B-C mode. (a) and (b): corresponding to A-B mode switching in Fig. 5.18 (a). (c) and (d) : corresponding to A-B mode switching in Fig. 5.18 (b). The insets in (a) to (d) show the measurement setup..... 114

Figure 5.20 Electroforming process for A-B mode. The numbers aside the I-V curves indicate the corresponding sweep order. After 4^{th} scan, devices are ready for RRAMs operation. The inset shows the corresponding backside voltage (V_c) change in the electroforming process 115

Figure 5.21 I-V curves for (a) A-C mode and (b) B-C mode. They are corresponding to each electroforming sweeps in Fig. 5.20. 116

Figure 5.22 I-V curves of the LRS/HRS switching process for the A-B mode. (a): It turns the A-B mode to LRS/HRS when sweeping to 5.5 V/10 V. The corresponding A-C and B-C mode are shown in Fig. 5.23 (a) and (b), respectively. (b): It turns the A-B mode to LRS/HRS when sweeping to -5.5 V/-10 V. The corresponding A-C and B-C mode are shown in Fig. 5.23 (c) and (d), respectively 117

Figure 5.23 I-V curves for A-C and B-C mode. (a) and (b) : corresponding to A-B mode switching in Fig. 5.22 (a). (c) and (d) : corresponding to A-B mode switching in Fig. 5.22 (b). The insets in (a) to (d) show the measurement setup..... 117

Figure 5.24 Illustration of Si random nanocrystal formation. (b) Band diagram of Si nanocrystal/SiO_2/ Si nanocrystal region subjected to 5V bias. (c) Localized switching model for SiO_x RRAMs. 119

Figure 5.25 I-V curves of the SSiRRAM electroforming process for devices with different silicon sputtering times. The control sample I-V curve is plotted for comparison. The inset shows the schematic cross-section of the SSiRRAM..... 124

Figure 5.26 Statistical plot of LRS current at 1V as a function of silicon sputtering time. The inset shows typical LRS and HRS switching I-V curves, and defines parameters I_{Set}, I_{Reset}, V_{Set}, and V_{Reset}..... 126

Figure 5.27 Statistical plot of SSiRRAM reset power. The inset shows the % coefficient of variation of LRS current and LRS average current for SSiRRAMs with different silicon sputtering times. 126

Figure 5.28 I-V curves for 100 switching cycles for SSiRRAM device with 30 sec silicon sputtering time. Multilevel programming is demonstrated by a 10× separation between the three states. The inset shows the resistance plot of the three states (with current read at 1V). 127

Figure 5.29 I-V curves for 100 switching cycles for the control sample. Programmed HRS current spreads-out and becomes mixed with MRS. The inset shows the resistance plot of the three states (with current read at 1V). 127

Figure 5.30 Tri-state pulse switching endurance of 10^6 cycles for SSiRRAM devices. 130

Figure 5.31 The read disturb immunity of LRS, MRS, and HRS by constant voltage stress of 1 V for 1000 seconds. 130

Figure 5.32 Retention properties of LRS, MRS, and HRS under continuous high temperature environment. HRS and MRS became unstable in the 5^{th} hour of baking (200^oC). When cooled to 20^oC, the 3 states recovered to their initial status. 131

Chapter 1 Introduction

1.1 Nowadays background

Driven by the ever increasing need for larger density, higher output performance and lower power consumption, silicon (Si) complementary metal-oxide-semiconductors (CMOS) devices have been aggressively scaled for more than 30 years. For much of that time, semiconductor industry has diligently followed Moore's Law, which states that the number of transistors incorporated in a chip will approximately double every 24 months [1]. In order to continue driving Moore's Law to increase functionality and performance with decreased costs, performance enhancers have been added to continue the density scaling roadmap, e. g. SiGe source/drain and uniaxial strain in the 90 nm and 65 nm nodes [2, 3] and high dielectric constant materials (high-k) along with metal-gate in the 45 nm and 32 nm nodes [4, 5]. Unfortunately, these performance enhancers cannot boost performance beyond 32 nm node. Therefore, a new MOSFETs architecture emerged to meet the requirements for the 22 nm node. In May 2011, Intel Inc. introduced a fundamentally different technology for 22 nm: 3-D Tri-Gate transistors [6]. Compared to conventional planar MOSFETs, Tri-Gate transistor offers improved leakage current, reduced power consumption and switching times, enabling superior efficiency and/or speed.

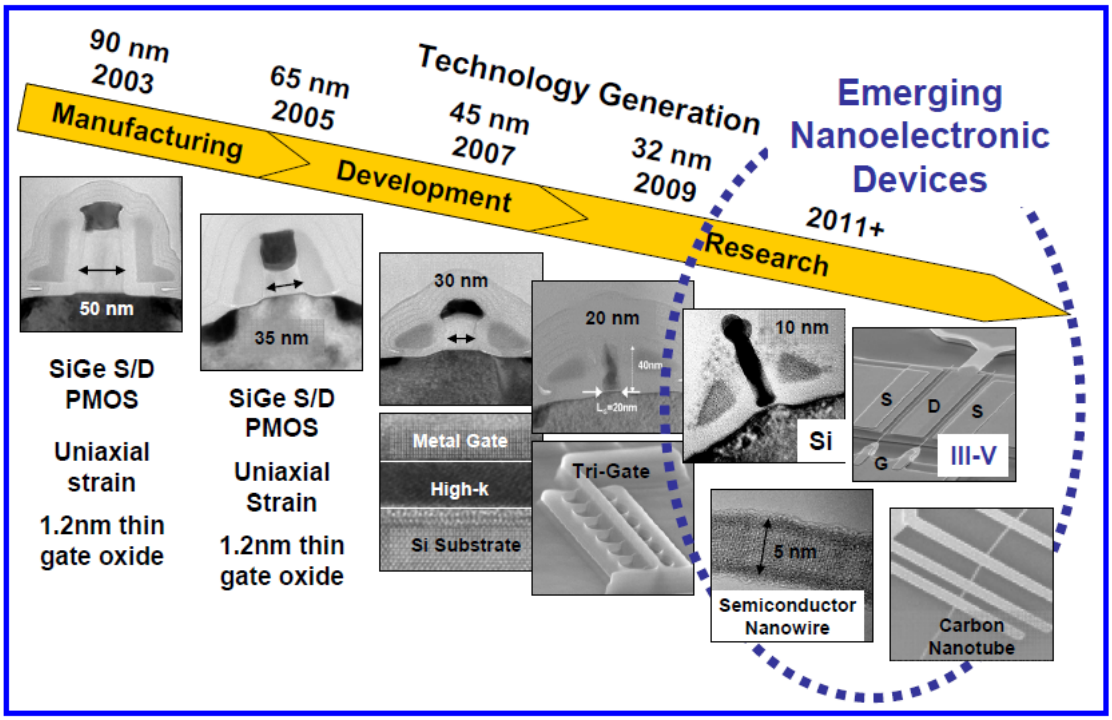


Figure 1.1 CMOS scaling trend from 2003 to 2011 and beyond [7].

Even though Tri-Gate transistor extends silicon lifespan in CMOS industry, the essential property of silicon may hinder further scaling due to its low electron mobility. Recently, high mobility materials such as, III-V and Ge, have drawn great interest for their advantages in high carrier mobility over their Si-based counterparts, as shown in Fig. 1.1. Most III-V materials lack good mobility for p-type carriers. In order to provide a CMOS solution, Ge is projected to be a good choice, even though it adds complexity to the whole process (integration of III-V and Ge on Si substrate). Researchers proposed a combination of a III-V n-channel MOSFET with a p-channel Ge MOSFET to form high-speed and low-power complementary FET pair [8, 9], as shown in Fig. 1.2. It is believed that III-V and Ge materials will be epitaxially grown on Si substrate with a buffer layer (might not be an oxide as shown in Fig. 1.2). The reason behind this is not only to take

advantage of the well-established Si platform, but also to expect that Si components will be included in the same chips.

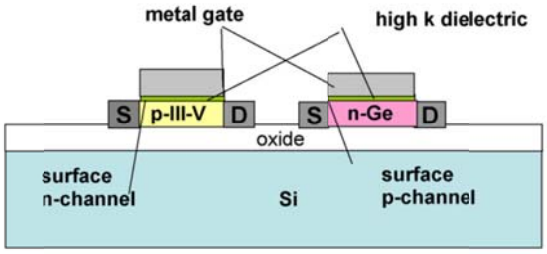


Figure 1.2 Example of non-classical CMOS structures. Combination of III-V and Ge channel on Si substrate.

International Technology Roadmap for Semiconductors (ITRS) has included high-mobility channel III-V/Ge as a technology for both high speed and low dynamic power in their 2011 Edition [10]. ITRS also anticipates that III-V/Ge will be in production in year 2018, as shown in the Table 1.1.

Table 1.1 Logic potential solutions from ITRS 2011 [10]

First Year of IC Production DRAM 1/2 Pitch (nm)	2011 36	2012 31	2013 28	2014 25	2015 22	2016 20	2017 18	2018 16	2019 14	2020 13
Fully depleted SOI MOSFET										
Multi-Gate MOSFETs										
Enhanced transport with alternate channels: III-V or/and Ge										
Enhanced transport with alternate channels: CNT, Nanowire, graphene										
Non-CMOS Logic Devices and Circuits/Architectures										

Research Required
Development Underway
Qualification / Pre-Production
Continuous Improvement

It is well known that V_{dd} is more difficult to scale than other parameters, mainly due to the fundamental limit of the subthreshold slope of ~60 mV/decade. This trend will continue and become more severe when it approaches the regime of 0.6 V. This fact along with the continuing increase of current density causes the dynamic power density (proportional to Vdd^2) to climb with scaling, soon to an unacceptable level. III-V/Ge can provide some relief in this area by allowing more aggressive V_{dd} scaling. Table 1.2 compares high-performance (HP) logic, low operating (dynamic) power (LOP) chips and low standby (static) power (LSTP) chips. Compared to HP, III-V/Ge can provide faster speed and lower dynamic power, which can be adapted into future high performance logic, such as microprocessor unit chips for desktop PCs or servers, etc.

Table 1.2 Comparison of HP, LOP, LSTP, and III-V/Ge Technologies [10]

	HP[a]	**LOP[b]**	**LSTP[c]**	**III-V/Ge**
Speed (I/CV)	1	0.5	0.25	1.5
Dynamic power (CV^2)	1	0.6	1	0.6
Static power (I_{off})	1	5×10^{-2}	1×10^{-4}	1

a: High-performance (HP) logic refers to chips of high complexity, high speed, and relatively high power dissipation, such as microprocessor unit chips for desktop PCs, servers, etc.

b: Low operating (dynamic) power (LOP) chips are typically for relatively high-performance mobile applications, such as laptop computers, where the battery is likely to be of high capacity and the focus is on reduced operating power dissipation.

c: Low standby (static) power (LSTP) chips are typically for lower-performance, lower-cost consumer type applications, such as consumer cellular telephones, with lower battery capacity and an emphasis on the lowest possible static power dissipation, i.e., the lowest possible leakage or off-current.

1.2 III-V materials

The previous section has given a background introduction of current semiconductor industry. III-V/Ge would be a solution for future high speed and low dynamic power applications, which is anticipated to be in production in the year 2018 [10]. This section will further discuss the advantage of III-V materials. First of all, we discuss how III-V MOSFETs can outperform conventional Si MOSFETs in terms of switching speed and dynamic power.

Propagation delay (τ_d) is the time delay for a signal to propagate from one gate to the next in a chain of identical gates. The τ_d of an inverter is the average of the delays of pull-down and pull-up of the next gate and could be written as [11],

$$\tau_d = \frac{1}{2}\text{(pull-down delay + pull-up delay)} \approx \frac{1}{2}\left(\frac{C_{load}V_{cc}}{2I_{onN}} + \frac{C_{load}V_{cc}}{2I_{onP}}\right)$$

$$\text{where pull-down delay} \approx \frac{C_{load}V_{cc}}{2I_{onN}}, \qquad \text{pull-down delay} \approx \frac{C_{load}V_{cc}}{2I_{onP}}$$

For a compatible NFET and PFET, I_{onN} matches I_{onP}. Therefore, the delay is $Q/I = C_{load}V_{cc}/2I_{on}$ and the switching speed is proportional to $I_{on}/C_{load}V_{cc}$. In order to maximize the switching speed, it is clearly important to maximize I_{on}.

Power consumption is an important goal for device design and engineering. In each switching cycle, a charge $C_{load}V_{cc}$ is transferred from the power supply to the load capacitance, C_{load}. The charge taken from the power supply in each second, $kC_{load}V_{cc}f$, is the average current provided by the power supply. Here, f is the clock frequency and k (<1) is an activity ratio that represents the fact that a particular gate in a given circuit is

not switched every clock cycle all the time. Therefore, the dynamic power ($P_{dynamic}$) could be expressed as,

$$P_{dynamic} = V_{cc} \times \text{average current} = k\, C_{load} V_{cc} = k\, C_{load} V_{cc}^{2}$$

The $P_{dynamic}$ dominates the power consumption when the transistor is switched frequently. One way to reduce $P_{dynamic}$ is to lower V_{cc}. It is desirable for a transistor to provide a large on-current at a low V_{cc}, which can reduce circuit switching delay and power consumption simultaneously.

G. Dewey et al., [12] have reported that the $In_{0.7}Ga_{0.3}As$ quantum well field effect transistor (QWFET) exhibits 20-30% higher drive current compared to strained Si MOSFET for a wide range of V_{GS}-V_T, as shown in Fig. 1.3, which depicts measured drive current versus V_{cc} comparing $In_{0.7}Ga_{0.3}As$ QWFET and strained Si MOSFET for constant I_{off} = 100 nA/μm. $In_{0.7}Ga_{0.3}As$ QWFET exhibits higher drive current compared to strained Si MOSFET for all V_{cc}, with greater gains at lower V_{cc}. The meaning behind this is that $In_{0.7}Ga_{0.3}As$ QWFET can provide the same drive current at lower V_{cc} while maintains the same I_{off}. Therefore, with lower V_{cc}, $In_{0.7}Ga_{0.3}As$ QWFET increases switching speed and reduces dynamic power.

III-V materials are compound materials and have high variety, especially for ternary compound semiconductor materials with different element composition. Table 1.3 lists several common III-V compounds, comparing with silicon.

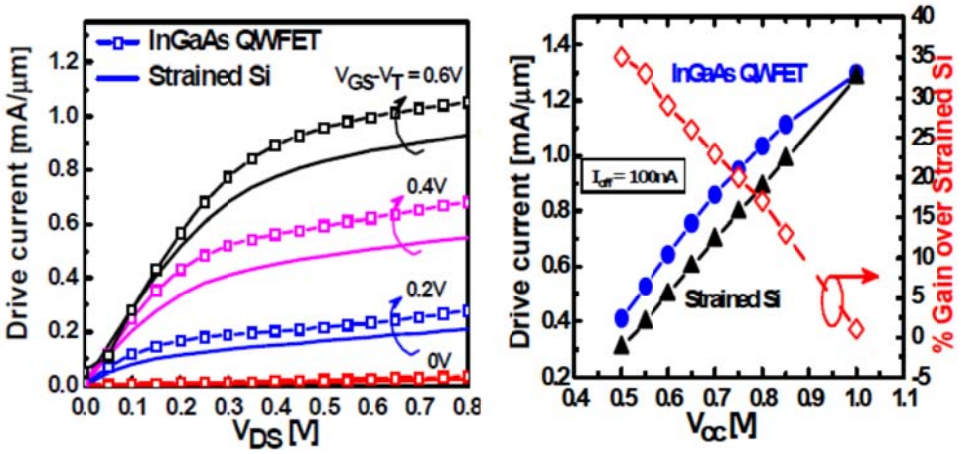


Figure 1.3 (Left) Measured drive current versus drain voltage comparing the 80 nm $In_{0.7}Ga_{0.3}As$ QWFET and 40nm Strained Si MOSFET at matched DIBL, SS and Ioff. $In_{0.7}Ga_{0.3}As$ QWFET exhibits 20-30% drive current gain over Strained Si MOSFET for a wide range of V_{GS}-V_T. (Right) Measured drive current versus operating voltage (V_{cc}) comparing 80 nm $In_{0.7}Ga_{0.3}As$ QWFET and 40nm Strained Si MOSFET for constant Ioff = 100 nA/μm. Despite 2.5X lower capacitance and 60% higher R_S+R_D, $In_{0.7}Ga_{0.3}As$ QWFET exhibits drive current gain over strained Si MOSFET for all V_{cc}, with greater gains at lower V_{cc}.

Table 1.3 Comparison of various III-V compounds and silicon

	Si	GaSb	GaAs	InP	$In_{0.53}GaAs$	$In_{0.7}GaAs$	InAs
Lattice Constant (A)	5.431	6.096	5.653	5.869	5.869	5.937	6.058
Electron Effective Mass (m^*/m_0)	0.19	0.041	0.067	0.077	0.041	0.034	0.023
Electron Affinity (eV)	4.05	4.06	4.07	4.38	4.5	4.65	4.9
Band gap (eV)	1.12	0.726	1.42	1.35	0.74	0.58	0.35
n_i at R. T (cm^{-3})	$1x10^{10}$	$1.5x10^{12}$	$2.1x10^{6}$	$1.3x10^{7}$	$6.3x10^{11}$	$1.1x10^{13}$	$1x10^{15}$
N_C (cm^{-3})	$3.2x10^{19}$	$2.1x10^{17}$	$4.7x10^{17}$	$5.7x10^{17}$	$2.1x10^{17}$	$1.5x10^{17}$	$8.7x10^{16}$
N_V (cm^{-3})	$1.8x10^{19}$	$1.8x10^{19}$	$9x10^{18}$	$1.1x10^{19}$	$7.7x10^{18}$	$7.5x10^{18}$	$6.6x10^{18}$
Electron mobility (cm^2/Vs)	1500	5500	8500	4600	12000	20000	33000
Hole mobility (cm^2/Vs)	450	850	400	150	300	400	460

III-V materials have much higher electron mobility compared to silicon, which can potentially provide higher drive current, better switching speed, and lower dynamic

power consumption. Among these III-V compounds, GaSb aims for p-III-V FET, which is still an open issue to have a compatible pFET for III-V; GaAs has difficulties in fabrication, which requires much higher thermal budget to fully activate source/drain implantation dopants; Fermi level pins in the conduction band of InAs, resulting in high off current in InAs FET. Only InP and In_xGaAs are promising for future application.

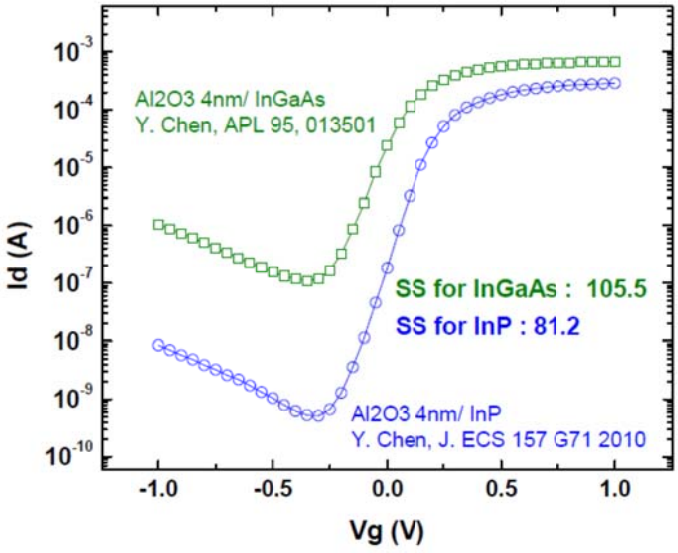


Figure 1.4 Comparison of InGaAs MOSFET [13] and InP MOSFET [14] with the identical high-k gate dielectric.

Figure 1.4 compares InGaAs MOSFET [13] and InP MOSFET [14] with the identical high-k gate dielectric. It shows that InP MOSFET has lower off current, which could be due to reduced junction leakage by wider band gap of InP (1.35 eV). As a result, the subthreshold slope (SS) is steeper in InP MOSFET compared to InGaAs MOSFET. Whereas InGaAs MOSFET conducts higher drive current because electron has higher mobility in InGaAs (12000 cm^2/Vs) compared to InP (4600 cm^2/Vs). Therefore, InP can aim for high switching application and InGaAs targets for high driving capability

application. In this work, we investigated InP and InGaAs and implemented high performance InP and InGaAs MOSFETs.

1.3 An emerging non-volatile memories -- RRAM

Flash memories have dominated non-volatile memory in stand-alone and embedded chips with great commercial success for more than 20 years. Aggressive scaling for larger density and higher output performance has push flash memories dimension down to 22 nm. However, flash memories are approaching the scaling limit, owing to its relatively long programming time (> 1 μs) and limited cycle endurance and high programming voltages (> 10 V) [15, 16]. Therefore, dense, fast, and low-power non-volatile memory will become highly desirable and emerging new non-volatile memory technologies are investigated to achieve high scalability with excellent performance beyond the 22 nm node.

Since the ultimate scaling limitation for charge storage devices (e.g. flash memories) is too few electrons, devices that provide memory states without electric charges are promising to scale further. Several non-charge-based non-volatile memories have been extensively investigated, such as Magnetic RAM (MRAM) [17], Ferroelectrics RAM (FeRAM), [18] Phase-Change RAM (PCRAM) [19] and random access memories (RRAMs) [20].

Due to a short history of RRAMs (the first memristor reported in 2008 [21]), the understanding and application of RRAMs are still under investigation. However, RRAMs

have shown promising properties, such as simple selector circuit (1D1R) with unipolar RRAMs, small cell area, fast programming time and read time, and good endurance performance. Table 1.4 compares these non-charge-based non-volatile memories with flash memories [22-24].

Table 1.4 Comparison of emerging non-volatile memories with flash memories [22-24]

Memory	Flash		FeRAM	MRAM	PCRAM	RRAM
Attributes	NOR	NAND				
Knowledge level	Mature		Prototypical products			Early state
Cell elements	1T		1T1C	1T1R	1T1R	1T1R/ 1D1R
Cell area (F^2)	10	5	22	20	4.8	4[23]
Write/Erase time	1us/10ms	1ms/100us	10ns/10ns	20ns/20ns	50ns/120ns	0.3ns[24]
Write voltage (V)	12	15	0.9 – 3.3	1.5	3	<3
Endurance	$< 10^6$		10^{12}	$>10^{14}$	10^{12}	10^{10}[24]
Scalability	Major technological barriers		Poor	Poor	Moderate	Excellent

To realize high performance RRAMs, development efforts have investigated many possible oxide-based resistive switching materials, such as TiO_2, NiO, MnO_2, TaO_x, HfO_2 and SiO_2 [25-34]. Among these oxides, SiO_2-based RRAMs [31-34] have drawn increasing attention due to their *unique* unipolar behavior: SiO_2-based RRAMs have set voltage [switching from the high resistance state (HRS) to low resistance state (LRS)] smaller than reset voltage (switching from LRS to HRS) and thus self-compliant properties. Therefore, SiO_2-based RRAMs may provide a better alternative than conventional unipolar RRAMs, which require current compliance to protect the device from complete breakdown when switching to the LRS. Moreover, SiO_2-based RRAMs have shown promising performance [31], including over 10^4 switching cycles, high switching speed (50 ns for set process and 80 ns for reset process) and robust nonvolatile

properties with extrapolated lifetime beyond 10 years. In addition, the selector circuit for unipolar SiO_2-based RRAMs could simplify the circuit design to one diode and one RRAM, potentially leading to better operational control of cross-bar array structures [30]. Furthermore, SiO_2 is a conventional material with high availability and ease of integration with current and future technology platforms. Therefore, we choose SiO_2-based RRAMs as our research target.

1.4 Outline

The motivation of this work contains two parts. The first part is to explore possible techniques to improve high-k/III-V gate stack quality and, in turn, to realize high performance III-V MOSFETs with high-k gate dielectrics. To achieve this goal, high-k/III-V interface chemistry were analyzed, comparing H_2O vs. O_3 as the oxidizer. Excessive oxidation was reduced by using H_2O as the oxidizer, rather than using O_3. Therefore, high-k/III-V interface quality was better with H_2O as the oxidizer. Insertion of interfacial passivation layers (IPLs) between high-k and the III-V substrate was investigated to further improve the interface quality. Various IPLs were studied, such as, a thin PVD Si IPL, a thin ALD Al_2O_3 or $HfAlO_x$, etc. Beyond IPLs, we incorporated fluorine into high-k/III-V gate stack to further improve III-V MOSFETs performance. The advantage of fluorine incorporation is the exclusion from IPLs, and it can provide even better high-k/III-V gate stack quality, which is very promising to implement high performance III-V MOSFETs with good scalability.

The second part of this work is to explore a possible candidate for future non-volatile memories application. Flash memories have dominated non-volatile memories for more than 20 years. Essential characters of flash memories, including big cell area, long programming time, and relative poor endurance, hinder its further scaling. A new non-volatile memories technique is highly desirable to further improve the performance and lower the cost. RRAMs, with small cell area, fast programming time, and good endurance, are promising for the next generation non-volatile memories. In this part, we studied the conducting mechanism of SiO_2-based RRAMs. The conducting path randomly formed in the SiO_2 and the switching happened in a weak spot along this conducting path. SiO_2-based RRAMs performance was further improved by sputtering a thin silicon layer onto the RRAMs' sidewall. Conducting current variation was reduced and multilevel application was realized with endurance up to 10^6 times.

Chapter 2 comprehensively discusses my studies of the effect of interfacial chemistry on ZrO_2/InGaAs gate stack, comparing ALD ZrO_2 with H_2O vs. O_3 as the oxidizer. We found that the amount of oxygen is critical to form a good interface. Excessive oxygen concentration, e. g. using O_3 as the ALD precursor, induces III-V native oxides at the interface. In the following experiments, we clear identify the correlation between the interface chemistry (elemental As, AsO_x, GaO_x, and In_2O_3) and electrical parameters, such as hysteresis, interface trap density (D_{it}), SS, transconductance (G_m), and drive current (I_{on}). By varying ALD oxidizer (O_3, H_2O) and clean up (TMA) we form, remove and prevent III-V oxides respectively. Highly interface sensitive physical analysis (synchrotron radiation photoemission at Brookhaven National Laboratory, and HRTEM) unveil the interface conditions. Moreover, applying Ti-based

materials as a capping layer between the high-k and gate metal is useful. An understanding of these interfaces helps us to control the chemistry at the high-k/III-V interface. A Ti-capping layer is able to getter excessive oxygen at the high-k/III-V interface to mitigate native oxide formation, resulting in a better interface quality and MOSFETs performance.

Chapter 3 is dedicated to the III-V MOSFETs with various IPLs. First, we investigated the effect of a thin Si IPL on the scalability of ALD HfO_2 on InP MOSCAPs, based on hysteresis and frequency dispersion analysis. 10Å Si IPL /51Å HfO_2 stacked gate dielectric was found to provide equivalent oxide thickness (EOT) of ~18Å while maintaining good interface with InP substrate at the same time. MOSFETs with 10Å Si IPL /51Å HfO_2 stacked gate dielectric show more than three times higher G_m and improved SS than the ones with a single 70Å HfO_2 gate dielectric. Next, we have investigated and compared device performance of InP MOSFETs with a single ZrO_2 gate dielectric and with stacked gate dielectrics using various interfacial dielectric layers between ZrO_2 and InP substrate. Al_2O_3, $HfAlO_x$, and $ZrAlO_x$ are demonstrated to be effective interfacial dielectric layers to improve device performance, including frequency dispersion, SS, I_{on}, effective channel mobility, and reliability.

Chapter 4 discusses my research of a novel CF_4 post-gate plasma treatment on III-V MOSFETs, which improves both high-k bulk and interface quality. Fluorine incorporation was demonstrated on various high-k/III-V gate stacks and achieved significant improvements, including $Al_2O_3/In_{0.53}Ga_{0.47}As$, Al_2O_3/InP, $HfO_2/In_{0.53}Ga_{0.47}As$, and HfO_2/InP. Detailed physical analysis, electrical characterization and device performance were carried out. With F incorporation, we have successfully developed

excellent interface quality of high-k (Al_2O_3 or HfO_2) directly on III-V substrate ($In_{0.53}Ga_{0.47}As$, or InP) without using interface passivation layer. For $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack, fluorinated samples exhibit low D_{it} of 4.9×10^{11} $cm^{-2}eV^{-1}$, which is the lowest value over prior reported $HfO_2/In_{0.53}Ga_{0.47}As$ gate stacks. Furthermore, a simple technique by a "two-step" high-k deposition with incorporating F between two high-k depositions was proposed to manipulate the F profile in the HfO_2 gate dielectric. As a result, we can further increase F concentration in the high-k/III-V gate stack and further improve gate stack quality.

Chapter 5 focuses on my studies of emerging non-volatile memories, RRAMs. In order to understand the conducting mechanism, various experiments were conducted. The device area effects and SiO_x thickness effects were also demonstrated. We found that there might be only one CF responsible for LRS/HRS switching. Moreover, devices with larger perimeter can be formed more easily compared to ones with small perimeter due to the higher possibility of containing weak spot at the edge. Testing structures of 5-RRAMs-chain and two adjacent devices in series help to understand the conducting and switching mechanism. We purposed a model for SiO_x RRAMs: the conducting filament is randomly formed within the SiO_x at the sidewall edge, depending on pre-existing defects. The conducting filament normally grows from both electrodes and it could be formed quickly in one electrode, which could be due to more pre-existing defects near that electrode. Moreover, the rupture/recovery could occur anywhere along the conducting filament, depending on a random process that determines the location of the weak spot along the conducting filament. In addition, we have investigated the effects of incorporating a thin silicon layer into a SiO_2-based RRAM. This technique significantly

reduced the electroforming voltage and instability of HRS current of SiO_2-based RRAMs. Consequently, a middle-state can be programmed between LRS and HRS, and remained distinguishable with tri-state pulse endurance performance over 10^6 cycles. The data stored had good read disturb immunity and thermal disturbance.

Chapter 6 summarizes my Ph.D. work. In addition, the future work is also discussed.

Chapter 2 Investigation of H_2O vs. O_3 based high-k on $In_{0.53}Ga_{0.47}As$

2.1 III-V Gate Stack Interface Improvement Using H_2O Oxidation of ZrO_2 and TiN Cap Layers

This chapter comprehensively discusses the characteristics of $ZrO_2/In_{0.53}Ga_{0.47}As$ gate stack. In physical analysis, we have investigated the interfacial chemistry reaction between atomic-layer deposition (ALD) precursors and the substrate ($In_{0.53}Ga_{0.47}As$). We found that the amount of oxygen is a critical factor to form a good interface. Excessive oxygen concentration, e. g. using O_3 as the ALD precursor, induces III-V native oxides at the interface. Even though adding a TiN cap layer on top of ZrO_2 can relieve native oxides formation through oxygen gettering effect, the residue of native oxides degrades MOSFETs performance significantly. In the following experiments, we clear identify the correlation between the interface chemistry (elemental As, AsO_x, GaO_x, and In_2O_3) and electrical parameters, such as hysteresis, D_{it}, SS, G_m, and I_{on}. By varying ALD oxidizer (O_3, H_2O) and clean up (TMA) we form, remove and prevent III-V oxides, respectively. Highly interface sensitive physical analysis (synchrotron radiation photoemission at Brookhaven National Laboratory, and HRTEM) unveil the interface conditions.

High mobility III-V semiconductors along with high-k dielectrics are projected to be key ingredients in future complementary metal oxide semiconductor technology [35]. Such devices aimed towards the sub-22 nm node will require good control of short channel effects, which in turn imposes stringent requirements for lower capacitive equivalent thickness (CET), reduced D_{it}, charge trapping, and a conformal high-k deposition process. At the same time, ALD is a desirable technique due to its precise and

uniform control of the dielectric layer to a fraction of a monolayer and ease of manufacturability. To lower the CET of the device without degrading device performance, any sub-oxidation of the III-V surface needs to be minimized. Theoretically, interfacial O maybe an issue; i.e., the III-V Fermi level can be pinned by <1% of a monolayer of chemisorbed O [36]. It is, therefore, critical to understand and favorably control reactions at the high-k/III-V interface.

An earlier report on Al_2O_3 has demonstrated that using O_3 as the oxidizer initiates excess interfacial oxidation of the $In_{0.53}Ga_{0.47}As$ surface compared to using H_2O [37]. However, the report did not include correlation of electrical data with chemical reactions. Moreover, the utilization of H_2O needs to be explored with higher-k-value materials because Al_2O_3 might impede CET scaling. Titanium (Ti)-based materials have been widely used on resistive random access memory technology due to their ability to getter oxygen [38 - 40]. Applying Ti-based materials as a capping layer between the high-k and gate metal is useful. An understanding of these interfaces helps us to control the chemistry at the high-k/III-V interface. In this chapter, we not only demonstrate the effect of interface chemistry by comparing ALD oxidizers (H_2O vs. O_3) for $ZrO_2/In_{0.53}Ga_{0.47}As$ film systems, but also correlated the device electrical characteristics and chemical reactions at the high-k/$In_{0.53}Ga_{0.47}As$ interface. The effect of adding a TiN capping layer has also been studied.

N-type (Si-doped, $9\times10^{17}/cm^3$) InP wafers were used as the starting substrates. N-type (Si-doped, $5\times10^{17}/cm^3$) $In_{0.53}Ga_{0.47}As$, 100 nm, was grown by molecular beam epitaxy as a buffer layer, followed by a 150 nm n-type $In_{0.53}Ga_{0.47}As$ layer (Si-doped, $1\times10^{17}/cm^3$). A thin InP was then grown as a protection layer. For device fabrication, InP

was removed by an HCl-based solution, followed by an appropriate surface clean. An 8 nm thick ZrO_2 layer was deposited by ALD with various oxidizers (O_3, low dose O_3 and H_2O), followed by an *in situ* ALD 2 nm TiN layer. Some device representative thinner ZrO_2 blanket film samples (both with and without the TiN cap) were processed for X-ray photoemission spectroscopy (XPS) measurements at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory (BNL). Subsequently, 200 nm TaN was sputtered and patterned for the gate electrode. The backside contact was deposited by e-beam evaporation, followed by annealing at 400°C for 30 s in N_2 ambient.

Figure 2.1 shows the multi-frequency capacitance-voltage (CV) curves for 8 nm thick ZrO_2 capacitors when using O_3 (O-ZrO_2), low dose O_3 (Lo-O-ZrO_2), and H_2O (H-ZrO_2) (all processed with a thin TiN cap layer). The CV curves for O-ZrO_2 are severely stretched out, indicative of a high D_{it}. The use of O_3 as the oxidizer for $ZrO_2/In_{0.53}Ga_{0.47}As$ capacitors produces deleterious interface effects as it oxidizes excessively. This finding is in good agreement with the work of Brennan *et. al*. [37]. Reducing the O_3 dose mitigates excessive interface oxidation which improves CV stretch-out as illustrated in Fig 2.1(b) and it can be further improved by using H_2O as the oxidizer (Fig. 2.1(c)). Frequency dispersions in accumulation for O-ZrO_2 and Lo-O-ZrO_2 are 27% and 50%, respectively. The frequency dispersion is further reduced to 13% in H-ZrO_2, indicative of fewer border traps [41].

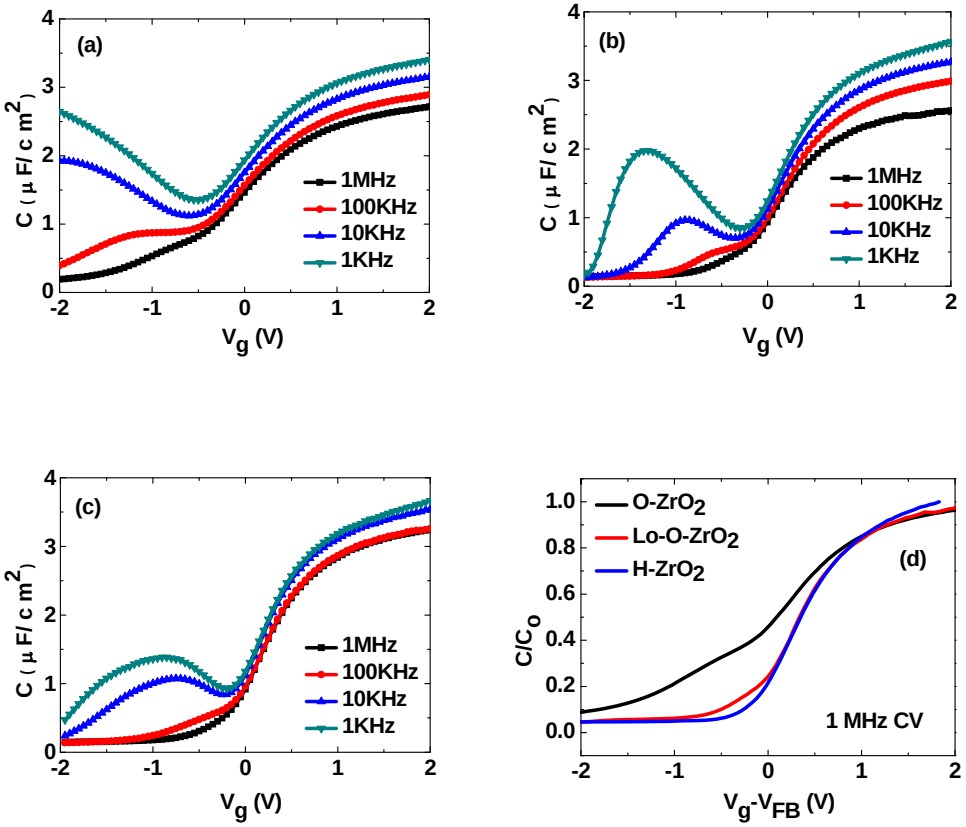


Figure 2.1 Multi-frequency CV curves of (a) O-ZrO_2, (b) Lo-O-ZrO_2, and (c) H-ZrO_2 (8 nm thick, all with a TiN cap layer). (d): Normalized 1MHz CV comparison of (a), (b), and (c)

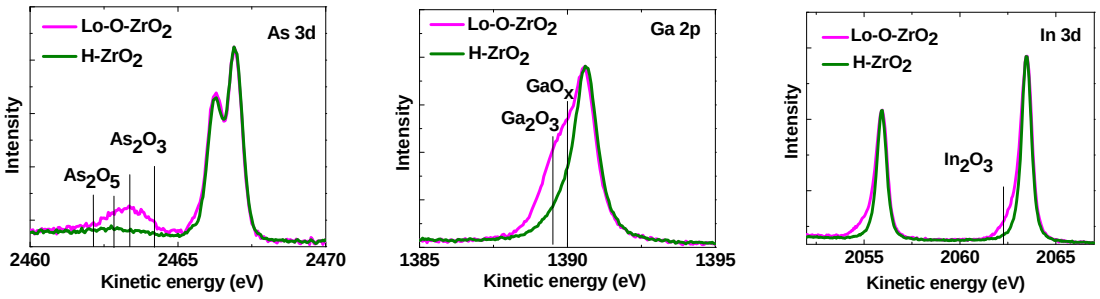


Figure 2.2 Photoemission data for H-ZrO_2 and Lo-O-ZrO_2 (w/o TiN capping). AsO_x, GaO_x, and In_2O_3 are detected with Lo-O-ZrO_2 whereas no native oxides are detected with H-ZrO_2.

Photoemission data for As *3d*, Ga *2p*, and In *3d* for Lo-O-ZrO_2 and H-ZrO_2 both with and without a TiN capping layer are shown in Fig. 2.2. Even with a lower O_3 dose (Lo-O-ZrO_2), III-V native oxides of As_2O_5, As_2O_3, Ga_2O_3, GaO_x, and In_2O_3 were clearly observed. However, these oxides were significantly reduced by a TiN capping layer, as shown in Fig. 2.3. TiN can getter oxygen from the interface. When compared with capacitor data (Fig. 2.1), this change at the interface reduces D_{it} in H_2O-based high-k. Given that TiN/Lo-O-ZrO_2 and TiN/H-ZrO_2 had negligible AsO_x and In_2O_3 to start with (Fig. 2.3) yet we observed changes in D_{it} in their capacitor samples (CV stretch-out, Fig. 2.1(d)), we can safely assume that the D_{it} in these two cases was primarily affected by GaO_x-based interfacial oxides and possibly by the presence of As-As bonds, which may be formed by As-O decomposition (As $3d_{3/2}$ increase). Note that AsO_x and In_2O_3 are easily removed by a metal precursor and low ozone dielectric chemistries, but eliminating GaO_x is more difficult. Excess GaO_x, however, can be reduced by using a H_2O-based precursor.

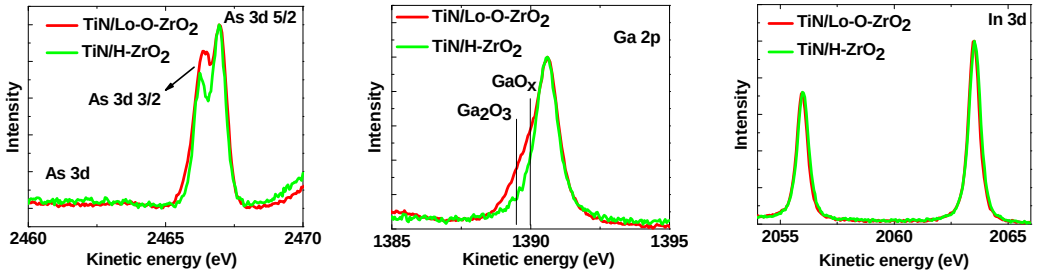


Figure 2.3 Photoemission data for TiN/Lo-O-ZrO_2 and TiN/H-ZrO_2.

The ability of TiN to getter oxygen is further confirmed by high resolution transmission electron microscopy (HRTEM). Figure 2.4 shows the HRTEM image of an O-ZrO_2/$In_{0.53}Ga_{0.47}As$ gate stack (without a TiN capping layer). A thin 0.8~1 nm amorphous interlayer is observed, which is possibly the product of excess oxidation when using O_3 as the oxidizer. The interlayer is believed to be composed of AsO_x, GaO_x, and In_2O_3 (Fig. 2.2). On the other hand, in the TiN/O-ZrO_2/$In_{0.53}Ga_{0.47}As$ gate stack, as shown in Fig. 2.5, the crystalline structure of ZrO_2 can be traced to the $In_{0.53}Ga_{0.47}As$ substrate, indicating that no obvious interlayer is formed. This confirms that the interlayer is associated with AsO_x, GaO_x, and In_2O_3 bonds (Figs. 2.2 and 2.3) and TiN is effective in gettering these detrimental bonds. Ga_2O_3 and GaO_x residues at the interface (some of them might up diffuse into the dielectric) are still detected by photoemission scans (Fig. 2.3), but the amount is insufficient to form an interlayer.

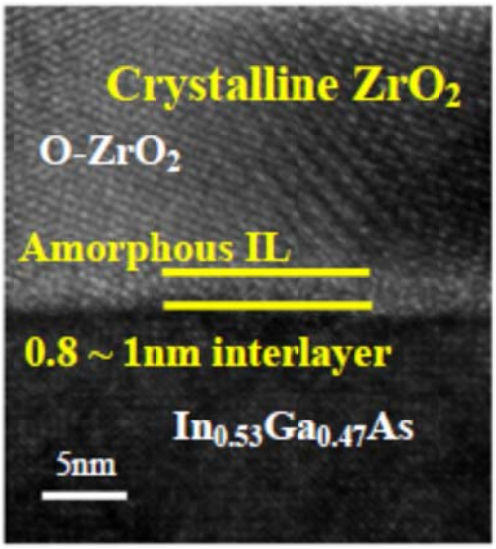


Figure 2.4 HRTEM of O-ZrO_2/$In_{0.53}Ga_{0.47}As$ (w/o TiN cap). Tracing the crystal lattice, an amorphous layer between crystalline ZrO_2 and $In_{0.53}Ga_{0.47}As$ layer is identified as a III-V native oxide layer, which may consist of AsO_x, GaO_x, and In_2O_3.

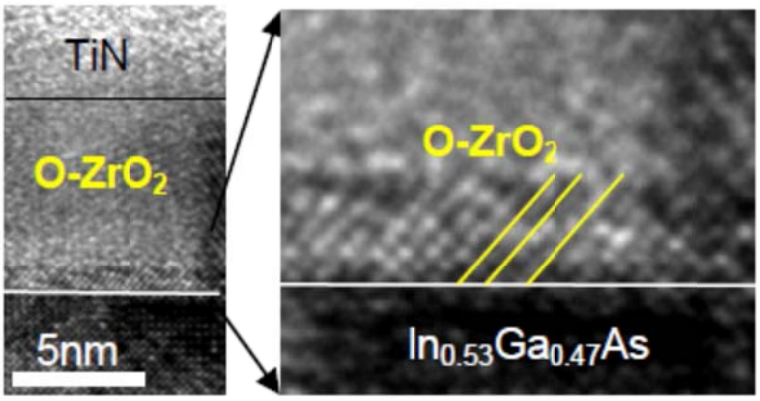


Figure 2.5 HRTEM image of TiN/O-ZrO_2/InGaAs. No interfacial layer is observed, which confirms that TiN gettering of AsO_x, GaO_x and In_2O_3 bonding is associated with the reduction of the interfacial layer.

At this point, we observed that TiN helps to reduce interface native oxides. In addition, it is important to demonstrate TiN gettering effect on other high-k material, i.e. H_2O based Al_2O_3 (H-Al_2O_3) vs. O_3 based Al_2O_3 (O-Al_2O_3). In TiN/O-Al_2O_3, TiN also helps to reduce AsO_x at the interface leaving As-As bonds compared to the sample without TiN. As 3d $^3/_2$ increase in O-Al_2O_3 is possibly due to As-O decomposition thus forming As-As bonds. In comparison with TiN/H-Al_2O_3, TiN/O-Al_2O_3 has higher interfacial oxide contribution (As-As, AsO_x, GaO_x, and In_2O_3).

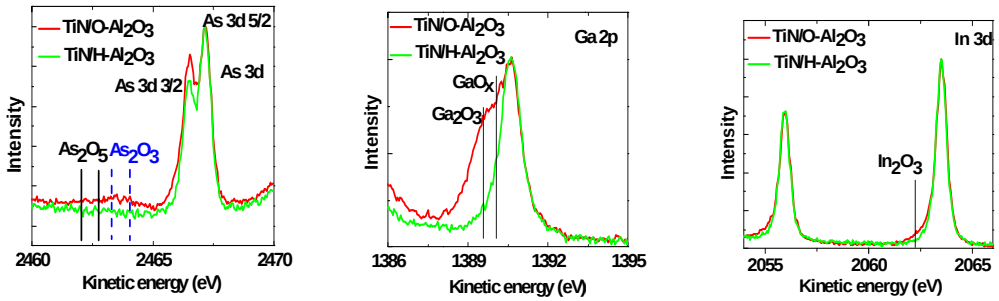


Figure 2.6 Photoemission data of TiN/O-Al_2O_3 and TiN/H-Al_2O_3.

The effects on the interface chemistry of varying ALD oxidizer (H_2O vs. O_3) and adding a TiN capping layer demonstrated that using O_3 as the oxidizer results in too much oxidation and forms an interlayer. This interlayer, clearly observed by HRTEM (Fig. 2.4), is composed of As_2O_5, As_2O_3, Ga_2O_3, GaO_x, and In_2O_3. The poor interface quality of O-ZrO_2 results in CV stretch-out. Adding a TiN capping layer reduces AsO_x, GaO_x, and In_2O_3 and ultimately decomposes interlayer due to gettering by TiN. Excess Ga_2O_3 and GaO_x can be further reduced by using a H_2O-based precursor. Our results suggest that a H_2O-based ALD along with a TiN cap layer improves the III-V interface quality.

2.2 Comparison of H_2O vs. O_3 based high-k on InGaAs

The previous section has showed the physical analysis of H_2O vs. O_3 based high-k on $In_{0.53}Ga_{0.47}As$ and the TiN oxygen gettering effect. Reduced native oxide formation can be obtained by adding a TiN cap layer and using H_2O as the oxidizer. This section will focus on comparing electrical performance of H_2O vs. O_3 based high-k (Al_2O_3 and ZrO_2) on $In_{0.53}Ga_{0.47}As$ (all with TiN) by detailed characterization.

Smaller C-V hysteresis, as shown in Fig. 2.7, suggests lesser bulk traps in the H_2O-based high-k. Evaluation of border traps (Q_{br}) of samples with H_2O vs. O_3 based high-k was carried out by charge pumping measurement with low gate pulse frequencies. Q_{br} locate inside the high-k (near interface) with long time constants as they interact with the conduction band electrons through tunneling, therefore, Q_{br} are unable to follow high frequency signal (1MHz). With low gate pulse frequencies, the conduction band electrons are allowed to tunnel deeper into high-k bulk and then, trapped by Q_{br}. This contributes to

the increase of traps density from 1MHz to 20KHz [41]. Figure 2.8 illustrates the reduced Q_{br} by using H_2O as the oxidizer. Q_{br} for O-Al_2O_3, Lo-O-ZrO_2, H-ZrO_2 and H-Al_2O_3 were estimated to 3.3×10^{20}, 3.6×10^{19}, 2.7×10^{19}, and 1.1×10^{19} cm^{-3}, respectively, indicating better high-k bulk quality with H_2O based high-k.

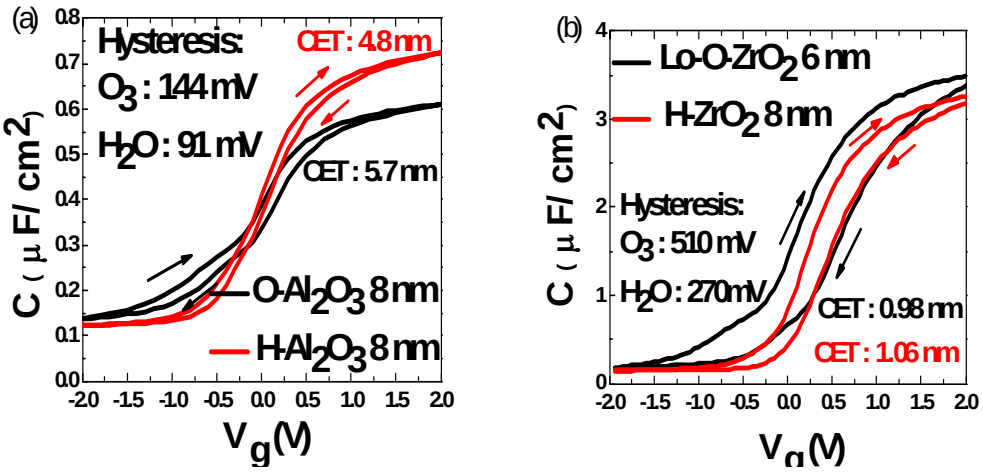


Figure 2.7 (a) CV Hysteresis of H-Al_2O_3 and O-Al_2O_3. Less hysteresis of the sample with H-Al_2O_3, even under higher electric field, suggests lesser bulk traps in H-Al_2O_3. (b) Similar improvement in H-ZrO_2.

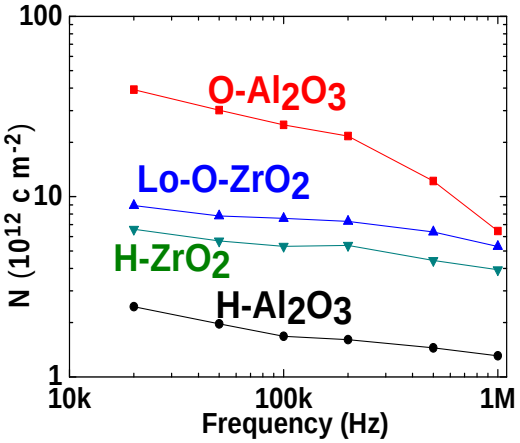


Figure 2.8 Charge pumping spectra of decreasing gate pulse frequencies.

To accurately characterize the interface quality of MOSFETs with H_2O vs. O_3 based high-k, charge pumping measurement were employed (at room temperature). Source and drain (S/D) were grounded while sweeping the base level (V_{base}: -1.5 V to 1 V in a step of 50 mV) of gate pulse (with a constant-amplitude, 1 V) at different frequencies. The region of the bandgap probed was from electron emission energy level to hole emission energy level, which was around the midgap. The charge pumping current (I_{cp}) is proportional to D_{it}. Figure 2.9 (a) and (b) exhibits lower I_{cp} for samples with H_2O based high-k, indicative of reduced D_{it}. The mean D_{it} values were extracted by changing rise time (t_R) and fall time (t_F), according to the following equation [42 - 43],

$$\frac{I_{cp}}{f} = 2qD_{it}AkT\left\{\ln\sqrt{t_R t_F} + \ln\left(\frac{\left|V_{fb}-V_t\right|}{\left|\Delta V_g\right|}V_{th}n_i\sqrt{\sigma_n\sigma_p}\right)\right\}$$

The mean D_{it} values for O-Al_2O_3, Lo-O-ZrO_2, H-ZrO_2 and H-Al_2O_3 were estimated to 4.71×10^{13}, 1.15×10^{13}, 8.14×10^{12}, and 4.29×10^{12} $eV^{-1}cm^{-2}$, respectively, as shown in Fig. 2.9 (d). O-Al_2O_3 has a greater D_{it} than Lo-O-ZrO_2, which is consistent with the photoemission data in Figs. 2.3 and 2.6 showing that O-Al_2O_3 has stronger AsO_x, GaO_x and In_2O_3 signals. Both H-Al_2O_3 and H-ZrO_2 show reduced D_{it} compared to alternative samples with dielectrics deposited using O_3.

It should be noted that TiN/Lo-O-ZrO_2 and TiN/H-ZrO_2 both had negligible AsO_x and In_2O_3 to start with, Fig. 2.3, and yet we observed changes in D_{it} in their charge pumping measurements (Fig. 2.9), it can safely be assumed that the D_{it} in these two cases were primarily affected by GaO_x based interfacial oxides and possibly by the presence of As-As bonds. Note that AsO_x and In_2O_3 are easily removed by metal precursor and low

ozone dielectric chemistries but GaO_x prevention is more difficult. Excess GaOx can be further reduced by using H_2O based precursor.

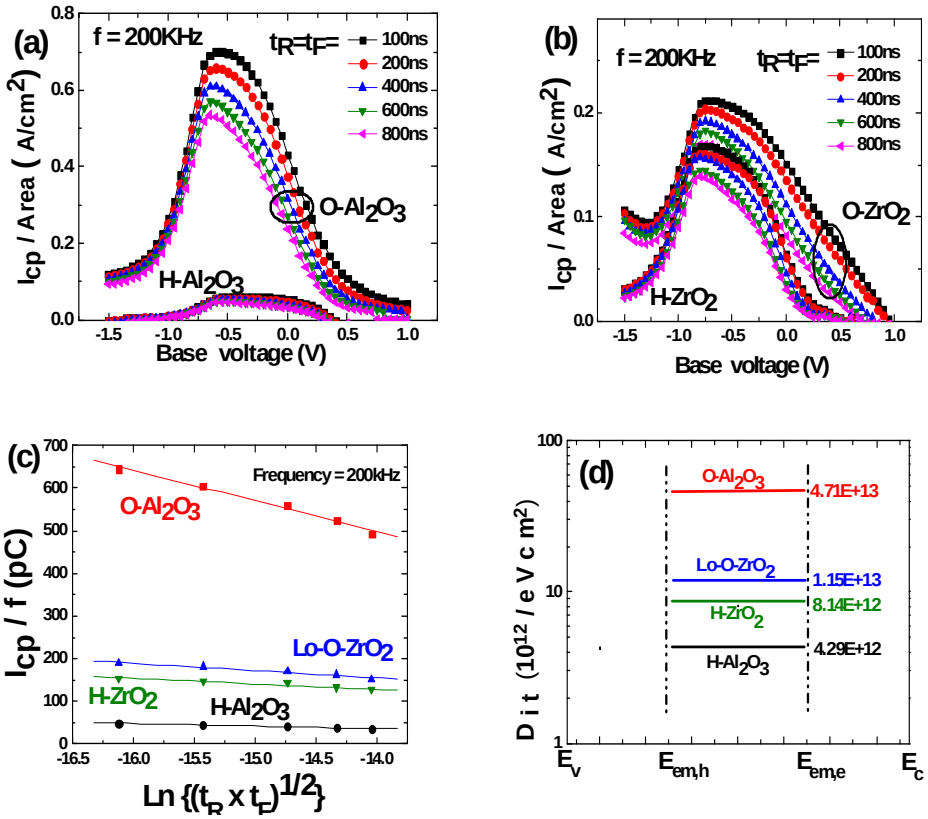


Figure 2.9 Mean D_{it} of O-Al_2O_3, H-Al_2O_3, Lo-O-ZrO_2, and H-ZrO_2. (a) and (b) : Charge pumping method with varying t_R/t_F time of the gate pulse under 200KHz. H_2O-based oxides has less charge pumping current compared to O_3-based oxides. (c) : mean D_{it} fitting according. (d): Mean D_{it} values of the 4 high-k oxides.

With better high-k bulk quality and interface quality of H_2O based high-k, improved electrical performance from MOSFETs with H_2O based high-k can be expected. Figure 2.10 shows the normalized I_d vs. V_g-V_{th} for O-Al_2O_3, H-Al_2O_3, Lo-O-ZrO_2, and H-ZrO_2. SS, indicating of interface quality, has been reduced. 32% higher I_d and 25% higher G_m were achieved with H_2O-based high-k. 17% and 14% higher drive currents

were also obtained with H-ZrO_2 and H-Al_2O_3, respectively, as shown in Fig. 2.11. These improvements are attributed to better gate stack quality.

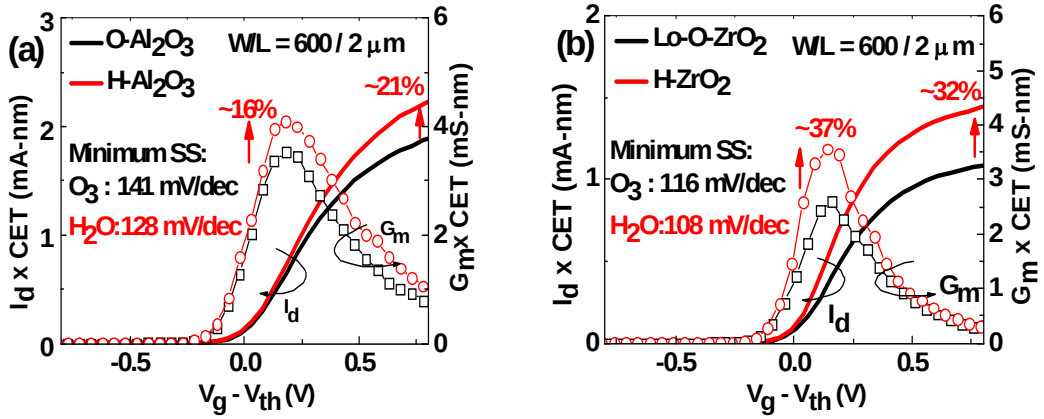


Figure 2.10 Improvements in normalized I_d and G_m and SS confirm that H_2O-based precursor chemistry is effective for InGaAs-based devices. (I_{off} is similar).

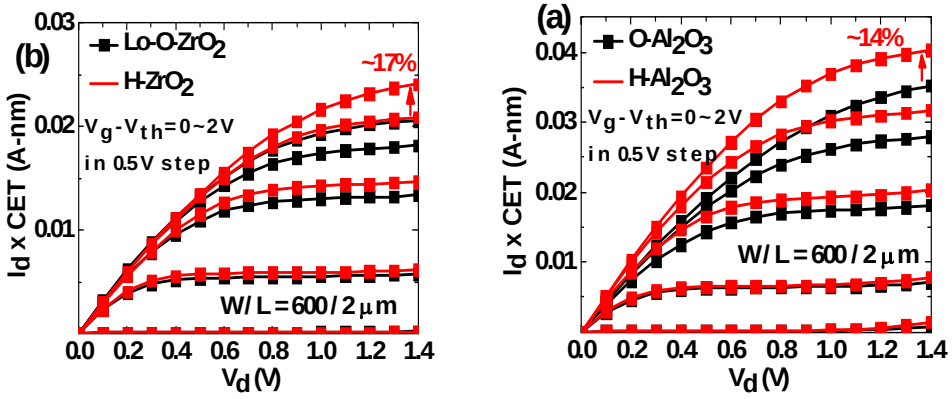


Figure 2.11 Normalized I_d improvements are attributed to better high-k/III-V gate stack quality of H_2O-based oxides.

2.3 Summary

The effects on the interface chemistry of varying ALD oxidizer (H_2O vs. O_3) and adding a TiN capping layer have been demonstrated. Using O_3 as the oxidizer results in too much oxidation and forms an interlayer. This interlayer, clearly observed by HRTEM, is composed of As_2O_5, As_2O_3, Ga_2O_3, GaO_x, and In_2O_3. The poor interface quality of O-ZrO_2 results in CV stretch-out. Adding a TiN capping layer reduces AsO_x, GaO_x, and In_2O_3 and ultimately decomposes this interlayer due to gettering by TiN. Excess Ga_2O_3 and GaO_x can be further reduced by using a H_2O-based precursor. We clearly linked these native oxides to D_{it} formation by careful characterization. By using H_2O as the oxidizer, native oxide formation minimized and resulted in better interface quality compared to oxides formed by O_3 as the oxidizer. Improved electrical performance was achieved with H_2O-based high-k, including reduced SS, higher drive current and larger G_m. Our results suggest that a H_2O-based ALD along with a TiN cap layer improves the III-V interface quality. Table 2.1 provides a detailed comparison of H_2O-based high-k and O_3-based high-k.

Table 2.1 Comparison of O-Al_2O_3, H-Al_2O_3, Lo-O-ZrO_2, and H-ZrO_2.

	O-Al_2O_3	Lo-O-ZrO_2	H-ZrO_2	H-Al_2O_3
Interface chemistry	As_2O_3 As_2O_5 GaO_x Ga_2O_3 In_2O_3 As-As	GaO_x Ga_2O_3 As-As	Less As-As	Less As-As
K-value	7	31-33	33-35	7.5
CV Hys. (mV)	144	510	270	91
Q_{br} (cm^{-3})	3.3E20	3.6E19	2.7E19	1.1E19
Mean D_{it} ($cm^{-2}eV^{-1}$)	4.71E13	1.15E13	8.14E12	4.29E12
G_m×CET (mS-nm)	3.53	2.59	3.56	4.1
I_d×CET (mA-nm)	1.9	1.09	1.44	2.3

Chapter 3 Interfacial passivation layers on InP and $In_{0.53}Ga_{0.47}As$ MOSFETs

3.1 PVD Si IPL on InP MOSFETs with ALD HfO_2

III-V based transistors have been attractive due to their high electron mobility and breakdown field compared to their Si counterparts. A tremendous effort has been done to implement III-V based transistors with high-k gate dielectrics, including GaAs MOSFETs with ALD HfO_2 dielectrics [44], or ALD Al_2O_3 dielectrics [45], or with e-beam evaporated Ga_2O_3 (Gd_2O_3) dielectrics [46]. InP MOSFETs have been implemented with ALD Al_2O_3 dielectrics [47], or ALD HfO_2 dielectrics [48] and InGaAs MOSFETs have also been presented with ALD Al_2O_3 or HfO_2 dielectrics [48 - 49]. However, the main challenge for III-V based MOSFETs is the absence of high-quality, thermodynamically stable dielectrics that can effectively passivate the interface states and prevent Fermi level pinning at III-V gate dielectric interface [50]. Therefore, several interface passivation techniques have been suggested to improve MOSFETs and metal-oxide-semiconductor capacitors (MOSCAPs) performance by employing Si IPL [48, 51], Ge IPL [52], or AlN surface passivation [53]. Recently, extensive studies have been conducted on InP-based MOSFETs [47, 48, 54 - 56] with ALD Al_2O_3 as gate dielectric, InP MOSFETs have showed the capability of high drive current density [55] and much smaller off current density due to its larger bandgap (1.34 eV) compared to InGaAs (0.74 eV for $In_{0.53}Ga_{0.47}As$). Also, InP MOSFETs with $HfAlO_x/HfO_2$ stacked gate dielectric showed impressive SS [e.g., 83 mV /dec [54]] compared to larger SS in InGaAs MOSFETs [e.g., 330 mV /dec [57]] and GaAs MOSFETs [e.g., 162 mV /dec [58]]. These

characteristics make InP a promising alternative material which should be studied further for future low-power logic applications.

Among the developed high-k dielectrics, ALD Al_2O_3 appears to be a promising candidate for III-V based transistors due to its good compatibility with III-V substrate, achieving an unpinned Fermi level at the oxide/ III-V interface [49, 59]. However, its lower k value (8-10) always hinders MOSFETs from further scaling down. ALD HfO_2 takes advantage with k value more than 20 and is more attractive in the application of MOSFETs even though ALD HfO_2 usually has higher interface state density than ALD Al_2O_3 with III-V materials [60 - 62]. In previous work, it has been demonstrated effective passivation of InP using an amorphous Si IPL layer [48]. However, very little work has been performed on gate oxide scaling down below EOT of 20 Å on MOSCAPs or MOSFETs on InP substrate. This section will discuss the effect of inserting a thin layer of Si IPL between ALD HfO_2 and InP substrate. The interface state density was reduced and the well-behaved MOSFETs with EOT of 18Å was realized.

10Å Si IPL /51Å HfO_2 stacked gate dielectric was found to provide EOT of ~18Å while maintaining good interface with InP substrate at the same time. In order to fairly demonstrate and compare the effect of a thin Si IPL, 70Å HfO_2 with EOT of ~17.4Å was fabricated directly on InP substrate without Si IPL. We also fabricated n-channel InP MOSFETs using gate last process, and the performance of transistors with 10Å Si IPL /51Å HfO_2 stacked gate dielectric and 70Å HfO_2 gate dielectric was compared. MOSFETs with 10Å Si IPL /51Å HfO_2 stacked gate dielectric show more than three times higher G_m and improved SS than the ones with a single 70Å HfO_2 gate dielectric. Maximum G_m of 2.3 mS/mm, drive current of 25.9 mA/mm (at V_d=2V, V_g-V_t=1.2V for 5

μm gate length and 600 μm gate width) and SS of 102.8mV/dec were achieved by MOSFETs with 10Å Si IPL /51Å HfO_2 stacked gate dielectrics.

MOSCAPs were fabricated on n-type InP (100) wafer doped with sulfur ($2\times10^{17}/cm^3$). The surface native oxides were removed with the 1% dilute HF solution, followed by 20% $(NH_4)_2S_x$ dip [63], resulting in a clean S-passivated InP surface. After surface treatment, ~10 Å Si IPL was deposited by dc magnetron sputtering in Ar ambient at 400°C, followed by *ex situ* ALD 51Å HfO_2 deposition using tetrakis (dimethylamino) hafnium (TDMAH) [$Hf(NMe_2)_4$] and H_2O as the precursors. For comparison, MOSCAPs having 70Å HfO_2 as a gate dielectric without Si IPL were fabricated. These two samples have similar EOT of ~18Å and ~17.4Å, respectively. After gate dielectric deposition, the post-deposition annealing (PDA) was performed at 500°C for 1min in N_2 ambient, then physical vapor deposited (PVD) TaN was used for gate electrode. After gate patterning, AuGe/Ni/Au alloy was deposited using E-beam evaporation for the backside contact. The samples were then annealed at 420°C for 20sec in N_2. N-channel MOSFETs were fabricated on semi-insulating (SI) InP (100) substrate with a ring-type structure [51] by gate-last process. The surface treatment was performed on SI-InP as same as MOSCAPs, then 100Å Al_2O_3 (dummy capping layer) was deposited by ALD at 200°C. After 35keV, $5\times10^{14}/cm^2$ Si ion implantation at the source and drain region, S/D activation was performed at 750°C for 15s. The Al_2O_3 dummy capping layer was then removed by BOE. After the same surface treatment on these InP samples, 10Å Si IPL /51Å HfO_2 stacked gate dielectric and single 70Å HfO_2 gate dielectric was deposited on separate samples. TaN gate electrode and AuGe/Ni/Au source/drain ohmic contact were deposited by PVD and E-beam evaporation, respectively. C-V curves and MOSFETs output characteristics

were measured by HP 4194 LCR meter and HP 4156 semiconductor parameter analyzer respectively. EOT values were extracted from measured C-V curves using NCSU CVC program [64] with consideration of the quantum mechanism.

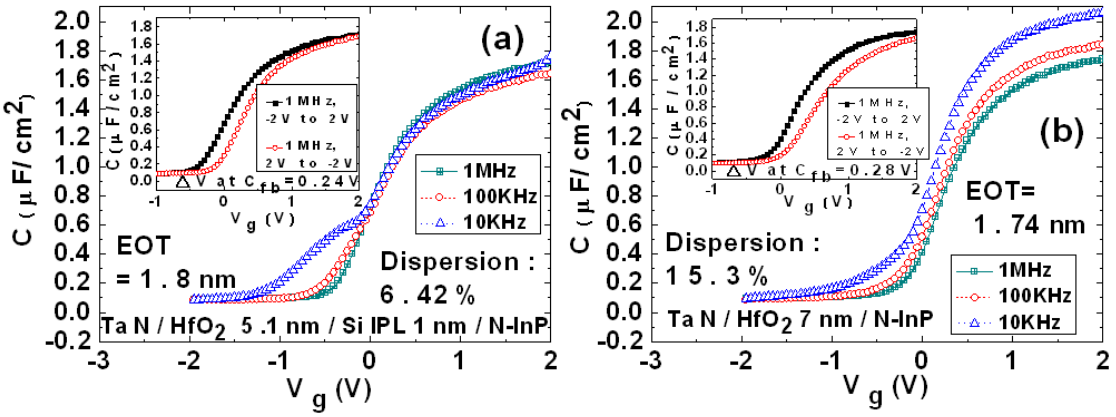


Figure 3.1 Hysteresis and frequency dispersion behavior of InP MOSCAPs with different gate dielectrics: (a) 10Å Si IPL /51Å HfO_2 stacked gate dielectric. Inset: its hysteresis characteristics. (b) 70Å HfO_2 without Si IPL. Inset: its hysteresis characteristics. Device area:$1.23x10^{-4}cm^2$.

Figure 3.1 compares typical C-V curves of InP MOSCAPs with 10Å Si IPL /51Å HfO_2 stacked gate dielectrics and single 70Å HfO_2 dielectric. The sample without Si IPL reveals poor frequency dispersion as large as 15.3% from 1MHz to 10KHz, while the 10Å Si IPL /51Å HfO_2 stacked gate dielectric illustrates a much smaller frequency dispersion of 6.42%, which indicates a much better interface with InP [54]. MOSCAP with Si IPL shows better hysteresis of 0.24V [Fig. 3.1(a) inset], compared to 0.28V for the MOSCAP without Si IPL [Fig. 3.1 (b) inset].

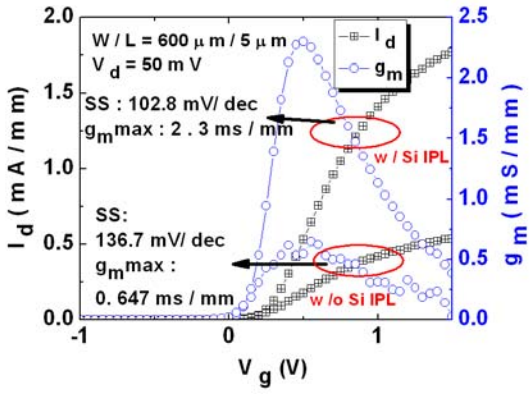


Figure 3.2 Drain currents and extrinsic transconductance vs. gate voltage.

We compare the characteristics of MOSFETs with same gate dielectrics as MOSCAPs in Fig. 3.2 to Fig. 3.4. Figure 3.2 shows the drain current and G_m versus gate voltage at V_d =50mV, where the gate width (W) is 600 μm and the gate length (L) is 5 μm. The maximum G_m of 10Å Si IPL /51Å HfO_2 stacked gate dielectric (2.3 ms/mm) is more than three times higher than the G_m value (0.647 ms/mm) of single 70Å HfO_2 gate dielectric. For 70Å HfO_2 gate dielectric without Si IPL, the SS is 136.7mV/dec, and the SS value reduces to 102.8mV/dec for 10Å Si IPL /51Å HfO_2 stacked gate dielectric.

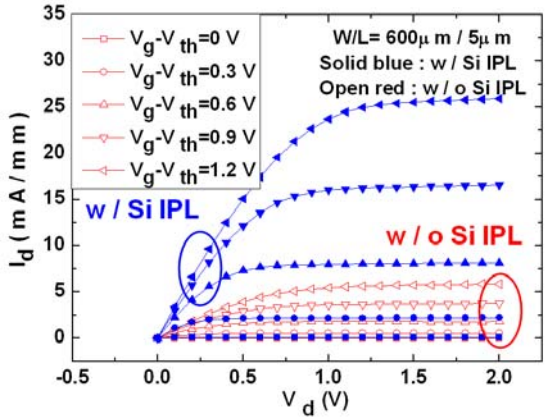


Figure 3.3 I_d-V_d characteristics as a function of V_g at $V_g - V_t$ = 0, 0.3 V, 0.6 V, 0.9 V, and 1.2 V for In MOSFETs with 70Å HfO_2 (open) and 10Å Si IPL /51Å HfO_2 stacked gate dielectric (solid).

I_d-V_d characteristics of MOSFETs with these two gate stacks are shown in Fig. 3.3. MOSFET with 10Å Si IPL /51Å HfO_2 stacked dielectric shows much higher current driver capability than the one with 70Å HfO_2 without Si IPL. The maximum driver current under V_d = 2V and V_g-V_t =1.2V for the gate dielectric with and without Si IPL are 25.9 mA/mm and 5.9 mA/mm, respectively (W/L=600 μm/5 μm). Figure 3.4 illustrates the gate leakage current of MOSFETs with these two gate stacks. The gate leakage current for 70Å HfO_2 gate dielectric is about 3.94×10^{-2} A/cm^2 at V_g=1V while employing Si IPL with similar EOT, gate leakage current reduces over one order to 1.32×10^{-3} A/cm^2. The D_{it} extracted by full conductance method is estimated to be ~ 3-8 × 10^{12} eV^{-1}cm^{-2} for 10Å Si IPL /51Å HfO_2 stacked gate dielectric and ~ 2-9 × 10^{13} eV^{-1}cm^{-2} for single 70Å HfO_2 gate dielectric [Fig. 3.4 inset]. We summarize the comparison of MOSCAPs and MOSFETs characteristics in Table 3.1.

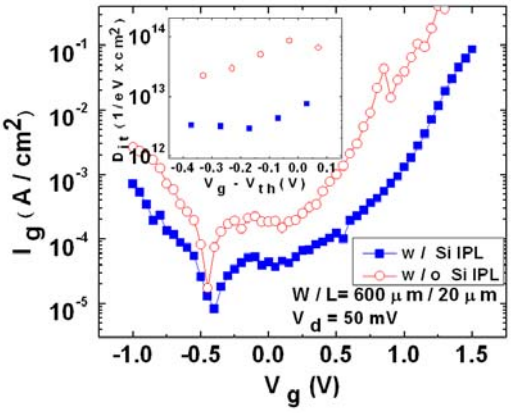


Figure 3.4 Gate leakage current density vs. gate voltage for InP MOSFETs with 70Å HfO_2 (circle) and 10Å Si IPL/51Å HfO_2 stacked gate dielectric (square). Inset: D_{it} comparison extracted by full conductance method.

Table 3.1 Effects of Si IPL on MOSCAPs and MOSFETs characteristics

	MOSCAPs Characteristics		MOSFETs Characteristics				
Gate stacks	Frequency Dispersion (%)	Hysteresis (V)	Id at Vd=2V Vg-Vt=1.2V (W/L=600um/5um) (mA/mm)	Gmmax at Vd = 0.05 V (mS/mm)	S.S. (mV/decade)	Dit ($eV^{-1}cm^{-2}$)	Gate leakage current density (A/cm^2) at Vg = 1V
10Å Si IPL/51Å HfO_2	6.42	0.24	25.9	2.3	102.8	$3\text{-}8\times10^{12}$	1.32×10^{-3}
70Å HfO2	15.3	0.28	5.9	0.647	136.7	$2\text{-}9\times10^{13}$	3.94×10^{-2}

3.2 Various IPL on $In_{0.53}Ga_{0.47}As$ MOSFETs with ALD ZrO_2

The previous section has been demonstrated that a thin Si IPL could improve interface quality. However, during the following process (ALD high k deposition), a part of Si IPL oxidizes and becomes SiO_x, which increases overall EOT and hinders device further scale down. Thus, an IPL with another species that have higher k value is highly desired.

In high k/metal-gate generation, ZrO_2, relieving the issue of thermal instability with poly-Si gate electrode [65], regains focus as a promising candidate to accomplish further scaling down on III-V substrates. Some groups have shown good performance of ZrO_2/InGaAs MOSFET [65 - 66]. However, not much research has been conducted on ZrO_2/InP MOSFET. This section will discuss and compare device performance of InP MOSFETs with a single ZrO_2 gate dielectric and with stacked gate dielectrics using various interfacial dielectric layers between ZrO_2 and InP substrate. Al_2O_3, $HfAlO_x$, and $ZrAlO_x$ are demonstrated to be effective interfacial dielectric layers to improve device performance, including frequency dispersion, SS, current driving capability, effective channel mobility, and reliability.

N-channel MOSFETs were fabricated on semi-insulating InP (100) substrate with a ring-type structure by gate-last process. The surface native oxides were removed with the 1% dilute HF solution, followed by 20% $(NH_4)_2S$ dip, resulting in a clean S-passivated InP surface. A 100Å ALD Al_2O_3 encapsulation layer was then deposited at 200°C. For device fabrication, source and drain regions were selectively implanted with a

Si dose of $2\times10^{14}/cm^2$ at 35 keV. After source and drain activation annealing (performed at 750 °C for 15 s in a N_2 ambient), the encapsulation layer was removed using buffered oxide etch solution. After the same surface treatment, various gate stacks were then deposited, including (a) 48Å ZrO_2, (b) 10Å Al_2O_3/30Å ZrO_2, (c) 10Å $HfAlO_x$/37Å ZrO_2, and (d) 10Å $ZrAlO_x$/42Å ZrO_2. Different thickness of top ZrO_2 layer was deposited in sample (b), (c), and (d) to make overall EOT around 1.2 nm. PDA was then performed for all the samples at 500°C for 60 s in a N_2 ambient. Subsequently, a 200 nm TaN gate electrode was sputtered and followed by gate patterning. The source and drain ohmic contacts were made by an E-beam evaporation of AuGe/Ni/Au and a liftoff process, followed by a 400°C annealing for 30 s in a N_2 ambient. MOSCAPs with the same gate stacks were also fabricated for CV analysis.

Figure 3.5 illustrates the CV characteristics of MOSCAPs on sample (a) to (d). All the samples have similar capacitance value in accumulation region. After consideration of quantum mechanism effect by using the NCSU CVC program [64], a similar EOT of 1.2 nm has been obtained so that fairly comparison is easily conducted. A single ZrO_2 dielectric on InP wafers shows high frequency dispersion of 36.5% (Fig. 3.5(a)) from the capacitance of 1MHz to 10KHz. While the frequency dispersion reduces substantially with insertion of interfacial dielectric layers between ZrO_2 and InP substrate (Fig. 3.5(b), (c), and (d)), indicating that a better interfacial quality has been obtained with thin Al_2O_3, $HfAlO_x$, and $ZrAlO_x$ layers (will discuss later). Note that in Fig. 3.5 (b), a CV "shoot-up" (at frequency of 10KHz) under high positive gate voltage is probably due to large gate leakage current (data not shown). Because Al_2O_3 is a relative low k material compared to $HfAlO_x$, and $ZrAlO_x$, in order to scale down EOT to 1.2 nm, only

30Å ZrO_2 is allowed to deposit on top of Al_2O_3 so that the overall thickness of sample (b) is too thin, causing large gate leakage current. Although the electrical performance of sample (b) seems reasonable at this thickness (MOSFETs performance will discuss later), further scale down beyond 1.2 nm might malfunction due to large gate leakage current.

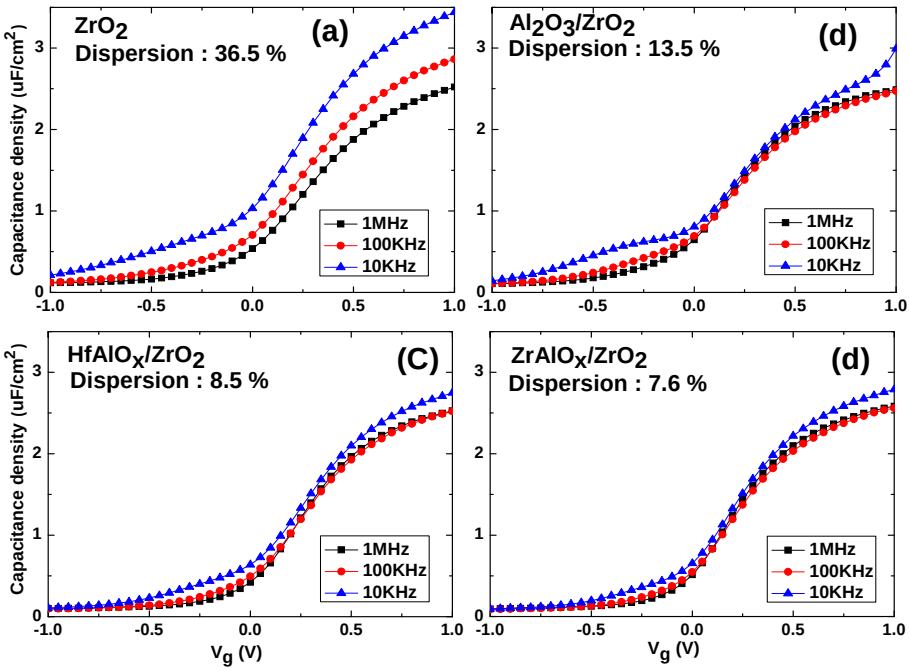


Figure 3.5 Frequency dispersion of MOSCAPs on InP substrate with different gate dielectrics: (a) single ZrO_2, (b) Al_2O_3/ZrO_2, (c) $HfAlO_x/ZrO_2$, (d) $ZrAlO_x/ZrO_2$. Frequency dispersion is calculated by the capacitance difference between 1MHz and 10 KHz at gate voltage=1.0 V.

We investigate the interface quality of these four samples by conductance method. As shown in Fig. 3.6, a single ZrO_2 has high D_{it}, suggesting a poor interface quality when ZrO_2 directly contact with InP substrate. Those interface traps can respond with low frequency CV (<1MHz) and contribute to total capacitance in the form of interface trap capacitors. This results in higher total capacitance measured by CV meter, which could

be observed when MOSCAPs are biased in accumulation region. This is consistent with the high CV frequency dispersion in Fig. 3.5 (a). Insertion of interface dielectric layers (Al_2O_3, $HfAlO_x$, and $ZrAlO_x$) prevents the direct contact of ZrO_2 and InP substrate. Meanwhile, three interface dielectric layers are able to provide good contact quality with InP substrate. In Fig. 3.6, D_{it} reduces more than one order of magnitude with interfacial dielectric layers. Accordingly, MOSCAPs with interfacial dielectric layers have less CV frequency dispersion (Fig. 3.5(b)-3.5(d)). Note that not much difference in D_{it} and frequency dispersion is shown between these three interface dielectric layers.

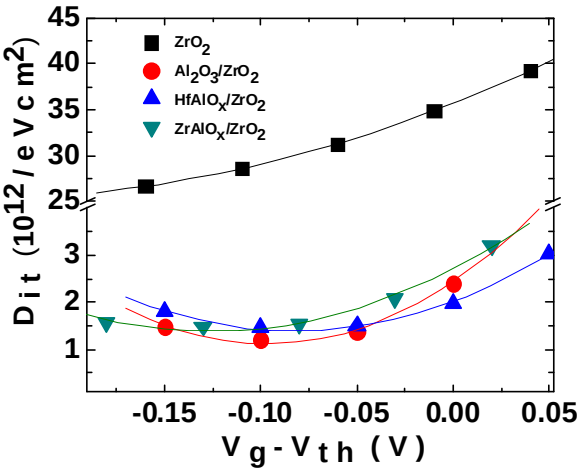


Figure 3.6 Interface trap density (D_{it}) distribution of: a single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks. The D_{it} distribution is extracted by conductance method.

Figure 3.7 illustrates log-scaled I_d-V_g characteristic of InP with four gate stacks. For a single ZrO_2, InP MOSFETs have much higher SS of 112mV/dec, while MOSFETs with Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$, SS improves to 80, 82, and 83, respectively. Better SS performance with interfacial dielectric layers directly suggests a

better interface quality between bilayer gate stacks and InP substrate, which is consistent with low D_{it} in Fig. 3.6.

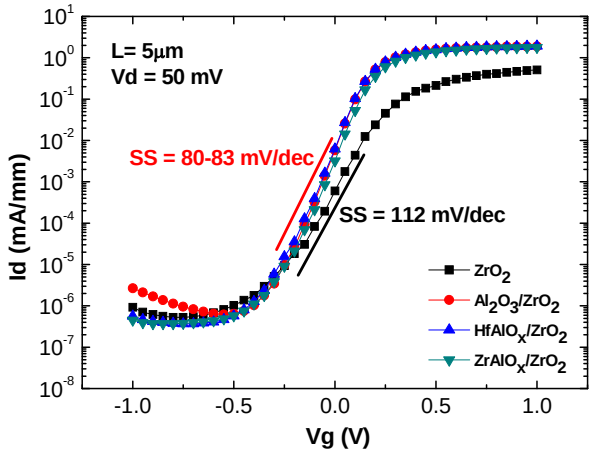


Figure 3.7 Log scale I_d-V_g characteristic at V_d=50mV of InP with a single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$.

Figure 3.8 shows the effective channel mobility of MOSFETs with different gate dielectric stacks. The channel mobility is increased from 126 cm^2/Vs in a single ZrO_2 gate dielectric to 533 cm^2/Vs in the $ZrAlO_x/ZrO_2$ gate stack. Low channel mobility is probably due to high interface trap density at the ZrO_2/InP interface. In N-MOSFETs, the interface traps in the upper band gap will have more significant effect on turn-on performance (Fermi level moves toward conduction band when substrate is being inverted). The interface traps near the conduction band side is believed to be acceptor-like traps, which are negatively charged when they are occupied by electrons. Those negatively charged traps act as scattering centers to the underneath inverted channel. While electrons are transporting in the channel, they would be scattered by those

negatively charged traps resulting in less electron mobility. On the other hand, MOSFETs with interfacial dielectric layers have fewer scattering centers so that electron mobility increases substantially. In the case of Al_2O_3 and $HfAlO_x$ passivation layer, channel mobility achieves 619 cm^2/Vs and 634 cm^2/Vs, respectively.

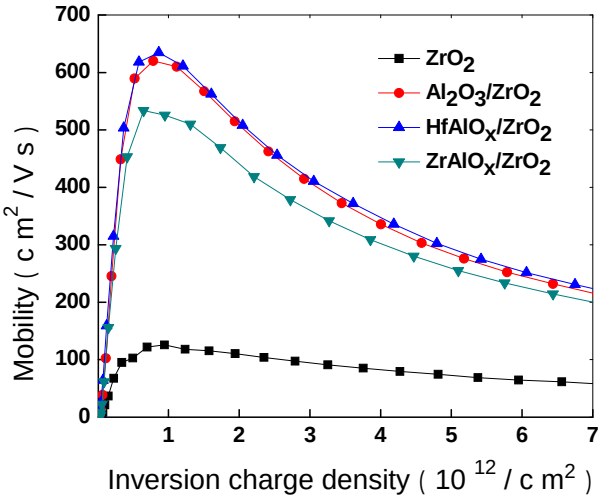


Figure 3.8 Effective channel mobility versus inversion charge density for InP MOSFETs with different gate dielectrics: single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$.

The I_d-V_d curves (gate length 5μm) of these four gate stacks are shown in Fig. 3.9. Low drive current in the MOSFETs with a single ZrO_2 is because of poor electron mobility mentioned above. With a better interface quality, MOSFETs with insertion of interfacial dielectric layers have higher electron mobility so that drive current increases (about four times). The maximum saturation drain currents of InP MOSFETs with Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks are similar (48mA/mm, 50mA/mm and 48mA/mm, respectively). Compared to a single ZrO_2, Al_2O_3/ZrO_2,

$HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks improve electrical performance substantially, suggesting that a thin Al-based passivation layer is required in ZrO_2/InP MOSFETs.

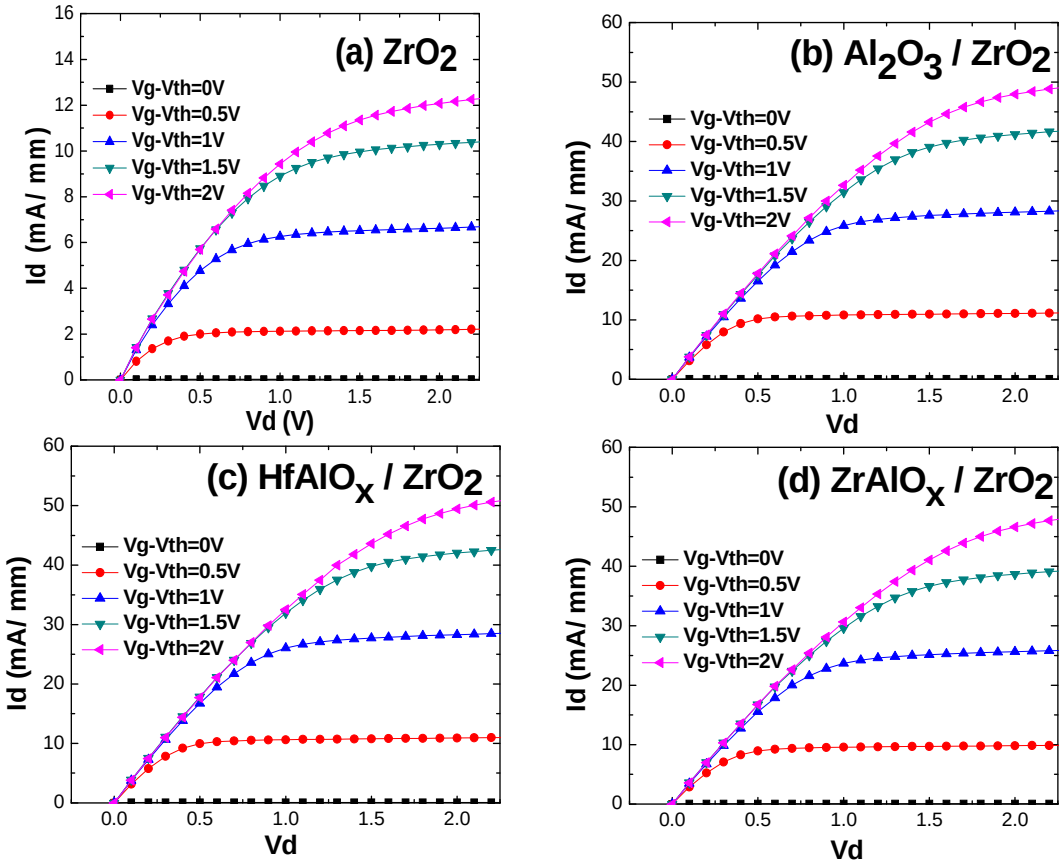


Figure 3.9 I_d versus V_d as a function of V_g at V_g-V_{th} from 0V to 2V with 0.5V step. MOSFETs with different gate stacks (W/L=600μm/5μm): (a) single ZrO_2, (b) Al_2O_3/ZrO_2, (c) $HfAlO_x/ZrO_2$, (d) $ZrAlO_x/ZrO_2$. I_d at V_d = 2.5V,V_g-V_{th} =2V increases about four times with insertion of passivation layers.

To examine the positive bias instability characteristics, we applied a constant dc electrical stress on the MOSFETs. Figure 3.10 (a) shows the SS increase after electrical field stress for these four samples. Larger SS increase for a single ZrO_2 indicates more interface degradation during the stress compared to other three gate stacks. A better interface quality for the bilayer gate dielectrics is in good agreement with Fig. 3.6 and Fig.

3.7. Figure 3.10 (b) shows the threshold voltage (V_{th}) shift under electrical stress field of (V_g - V_{th0})/EOT = 10 MV/cm lasting for 500s. V_{th0} is the threshold voltage of the fresh devices. A large V_{th} shift (200 mV) is obtained for the MOSFETs with single ZrO_2, while V_{th} shift reduces to 164 mV, 124 mV, and 107 mV for the MOSFETs with Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks, respectively. It is believed to be due to more traps generated in a single ZrO_2 during the electrical stress. While MOSFETs with Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks are more robust and less traps are generated.

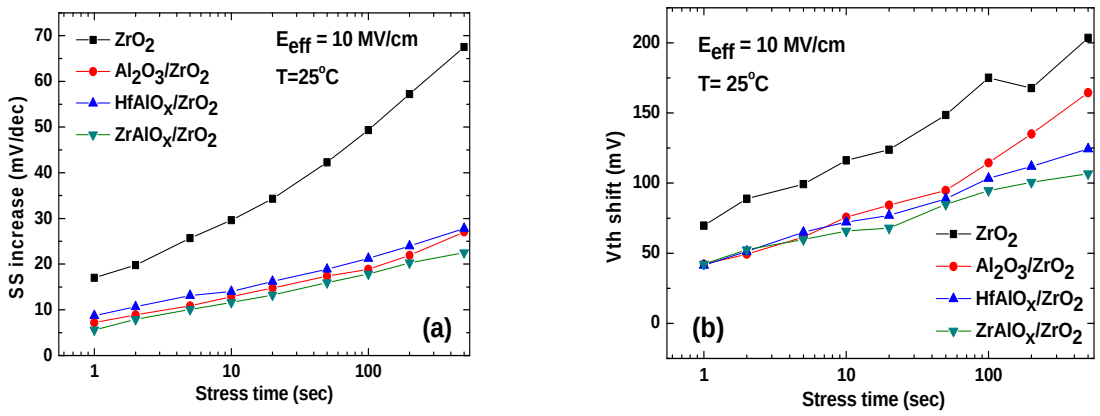


Figure 3.10 Reliability performance with a constant stress of effect electrical field 10MV/cm on InP MOSFETs with single ZrO_2, Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ (a)V_{th} shift and (b) SS increase.

3.3 Summary

MOSCAPs and MOSFETs have been fabricated on InP using ALD HfO_2 as gate dielectric with and without Si IPL. MOSCAPs with 10Å Si IPL /51Å HfO_2 stacked gate

dielectric exhibit EOT of 18Å, and they also show much better interface with InP substrate than MOSCAPs with a single 70Å HfO_2 gate dielectric, demonstrated by 8.9% less frequency dispersion. The characteristics of the transistors are also compared, more than three times higher G_m, four times higher drive current density, and 34.7mV/dec smaller SS are obtained by MOSFETs with 10Å Si IPL /51Å HfO_2 gate stack than MOSFETs with single 70Å HfO_2 gate dielectric. MOSCAPs and MOSFETs with Si IPL have been scaled down to EOT ~ 18Å.

Moreover, InP MOSFETs using various gate dielectrics including a single ZrO_2, stacked Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ were fabricated and their performances were compared. Using Al_2O_3/ZrO_2, $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ as the interfacial dielectric layers achieves smaller SS, larger drive current, and channel mobility compared to a single ZrO_2 layer. Moreover, reliability performance has also been improved. These improvements are attributed to a better interface quality brought by the interfacial dielectric layers. Because $HfAlO_x$ and $ZrAlO_x$/ have higher k value compared to Al_2O_3, InP MOSFETs with $HfAlO_x/ZrO_2$, and $ZrAlO_x/ZrO_2$ gate stacks are more reliable for further scaling down.

Chapter 4 Fluorine Incorporation into high-k/III-V MOSFETs for enhanced device characteristics

4.1 Motivation

Chapter 2 discussed the possibility of high-k/III-V gate stack without IPLs. We have studied the impact of interface chemistry on the interface quality of high-k/III-V MOSFETs. Our results showed that high-k grown by H_2O outperformed that grown by O_3, which pointed out that, without interface protection, using H_2O as the oxidizer is more reliable to implement high performance high-k/III-V MOSFETs. Chapter 3 have presented improved interface quality by applying a thin IPL and significant improvements have been achieved. However, these IPLs usually have lower k values, which hinder further EOT scaling. For future III-V MOSFETs application, a high quality high-k gate stack without IPL is urgently needed.

This chapter will demonstrate an improved performance by a post-treatment of CF_4 plasma. Without IPLs, a further improved interface quality has been achieved by fluorine (F) incorporation. The application of F incorporation leads to better interface quality without any interfacial passivation layers. In this Chapter, the improvements of MOSFETs on $Al_2O_3/In_{0.53}Ga_{0.47}As$, Al_2O_3/InP, $HfO_2/In_{0.53}Ga_{0.47}As$, and HfO_2/InP gate stacks by F incorporation will be presented.

4.2 Improved electrical characteristics of TaN / Al_2O_3 / $In_{0.53}Ga_{0.47}As$ MOSFETs by F incorporation

This section will discuss a post-gate CF_4 plasma treatment on $In_{0.53}Ga_{0.47}As$ channel MOSFETs. F has been incorporated into the ALD Al_2O_3 gate dielectric by post-gate CF_4 plasma treatment. A smaller SS and reduced D_{it} has been achieved with F passivation, suggesting a better interface quality. With CF_4 plasma treatment, drive

current, G_m and effective channel mobility has been shown to increase by 13.9%, 12.5%, and 29.6%, in comparison to the control devices respectively.

Fabrication of MOSFETs on III-V-based compounds has been attractive for their advantages over their Si-based counterparts in high electron mobility. The key challenge in implementing III-V materials is the development of a compatible gate dielectric stack [67]. Various high-k materials have been studied, including in-situ MBE deposition of Ga_2O_3-Gd_2O_3 mixture or Gd_2O_3 on GaAs and InGaAs substrates [68 - 70], silicon or germanium as interfacial passivation layer between GaAs, InP, or InGaAs substrates and dielectrics [51, 72 - 73], and ALD of Al_2O_3, HfO_2, and ZrO_3 directly on III-V substrates [60, 74 - 75].

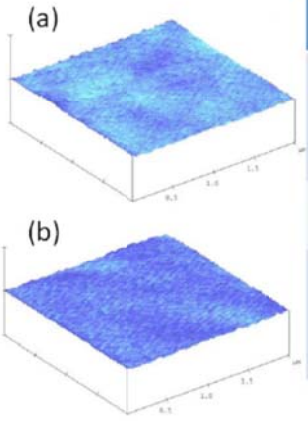
Recently, F incorporation into the high-k gate dielectrics has been shown to improve device performance and reliability on Si substrate [75 - 77] and Ge substrate [78]. K. Seo et al. incorporated F into high-k gate dielectric by F_2 annealing in the presence of UV radiation [76], passivating high-k bulk and high-k /Si interface defects. Y. Mitani et al. introduced F into gate dielectrics by F implantation and followed by annealing [77]. Moreover, F can be incorporated by CF_4 plasma treatment before or after gate dielectric deposition [75, 71]. It has been demonstrated that F tends to segregate to the HfO_2/SiO_2 and the SiO_2/Si interfaces passivating oxygen vacancy and interface traps by forming stronger Hf-F and Si-F bonds [75]. The similar mechanism has been reported on Ge substrate that Ge-F and Hf-F bonds formation at high-k / Ge interface and in HfO_2 gate dielectric, resulting in reduction in interface and bulk traps, respectively [78]. In this section, we explore the possibility of its inclusion as the gate dielectric in InGaAs-based devices. We demonstrate the impact of post-gate CF_4 plasma treatment on $In_{0.53}Ga_{0.47}As$

nMOSFET. With CF_4 plasma treatment, improved device characteristics have been obtained, including the peak G_m of 2.8 mS/mm (L=20 μm, V_d=50 mV), drive current of 32.5 mA/mm (L=20μm, Vd=50mV), and the peak effective channel mobility of 1372 cm^2/V-S.

$In_{0.53}Ga_{0.47}As$ MOSFETs were fabricated on 300nm p-type thick $In_{0.53}Ga_{0.47}As$ (Be-doped, $5\times10^{16}/cm^3$) epitaxially grown on p-InP substrate with ring-type structure [48]. The native oxides on InGaAs were removed with 1% diluted HF solution, followed by 20% $(NH_4)_2S$ dip, resulting in a clean S-passivated surface. Then, 100Å ALD Al_2O_3 dummy capping layer was deposited. After 35 keV $2\times10^{14}/cm^2$ Si ion implantation, the source and drain activation annealing was performed at 700 °C for 10s. Then, the dummy capping layer was removed using buffered oxide etch. After the same surface preparation (HF and $(NH_4)_2S$), 4nm Al_2O_3 was deposited by ALD. Some samples were treated by CF_4 plasma (rf power of 20W) with pressure of 100 mTorr and flow rate of 50 SCCM for 3 minutes. To avoid possible carbon contamination, O_2 with a flow rate of 5 SCCM was also introduced into the plasma. Post-deposition annealing was then performed for both CF_4-plasma treated samples and control samples (without CF_4-plasma treatment) at 500°C in a N_2 ambient for 90s. After that, a 200 nm TaN gate electrode was sputtered and followed by gate patterning. AuGe/Ni/Au by E-beam evaporation was deposited for source and drain ohmic contact and Cr/Au for backside contact.

AFM images of Al_2O_3 before and after CF_4 plasma treatment are shown in Fig. 4.1(a) and Fig. 4.1(b), respectively. The RMS values at different steps are summarized in Fig. 4.1(c). The surface roughness (RMS) of Al_2O_3 after CF_4 plasma treatment was 0.085

nm, compared to that of the control sample (RMS ~ 0.097 nm), this suggests that no damage was caused by the plasma treatment.



Samples	RMS (nm)
Bare $In_{0.53}Ga_{0.47}As$ after HF (1%) clean	0.199
(a) ALD Al_2O_3 60 cycles	0.097
(b) ALD Al_2O_3 60 cycles + F plasma 3min	0.085

(C)

Figure 4.1 AFM images on (a) Al_2O_3 / $In_{0.53}Ga_{0.47}As$ (b) Al_2O_3 / $In_{0.53}Ga_{0.47}As$ after CF_4 plasma treatment 3min. (c) surface roughness RMS values measured at different steps.

In order to explore the F distribution in the Al_2O_3 / $In_{0.53}Ga_{0.47}As$ structure, the secondary ion mass spectrometry (SIMS) technique was adopted (Fig. 4.2). In the control samples, the F atoms observed in the SIMS profile are possibly the residue due to the dilute HF dip. With CF_4 plasma treatment, an obvious increase of F concentration is observed, indicating that F atoms have been incorporated into Al_2O_3 / $In_{0.53}Ga_{0.47}As$ stack. All other elements in the Al_2O_3 / $In_{0.53}Ga_{0.47}As$ stack remain the same with and without CF_4 plasma treatment.

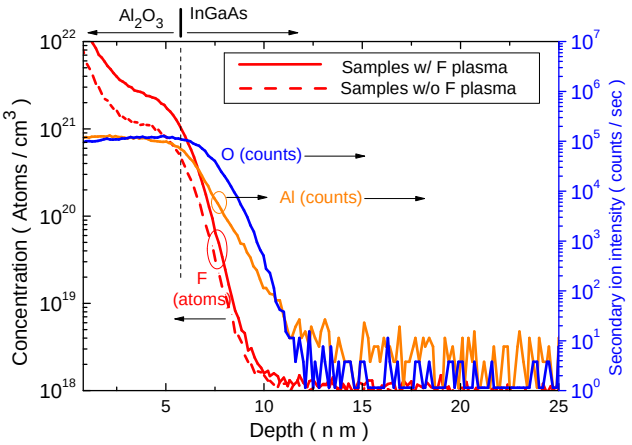


Figure 4.2 SIMS depth profile of Al_2O_3 / $In_{0.53}Ga_{0.47}As$ structure for fluorine, oxygen, and aluminum, atoms distribution. Red solid line: w/o F plasma. Red dash line: w/ F plasma.

Figure 4.3 compares the drain current and G_m versus gate voltage at V_d =50mV, where the gate width (W) is 600 µm and the gate length (L) is 20 µm. The drain current of the control sample was 1.14 mA/mm and after CF_4 plasma treatment, the drain current has been increased by 13.9% to the value of 1.3 mA/mm. A similar trend has been observed in the maximum G_m. With CF_4 plasma treatment, the maximum G_m improved from 2.5 mS/mm to 2.8 mS/mm, increasing by 12.5%. For samples without CF_4 plasma treatment, the SS was 105.5 mV/dec, and the SS value reduced to 97.9mV/dec for the CF_4 treated samples. This indicates that fluorinated Al_2O_3 device has a better interface characterization.

We have also extracted the effective channel mobility of $In_{0.53}Ga_{0.47}As$ MOSFETs with and without CF_4 plasma treatment using split-CV method and plotted it in Fig. 4.4. The peak effective channel mobility for the CF_4 treated samples (1372 cm^2/Vs) is 29.6% higher than that of the control samples (1059 cm^2/Vs). This is believed to be due to the

improved interface quality by post-gate CF_4 plasma treatment. Note that the mobility value of the control samples is consistent with the reported values [80]. Fig. 4.4 inset (a) and (b) show split-CV hysteresis for MOSFETs with and without CF_4 plasma treatment. The hysteresis slightly improves after CF_4 plasma treatment.

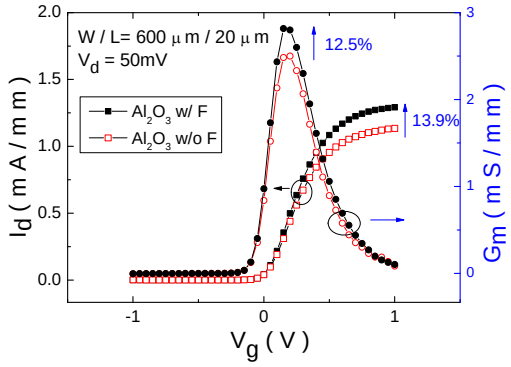


Figure 4.3 Drain currents and extrinsic transconductance versus gate voltage.

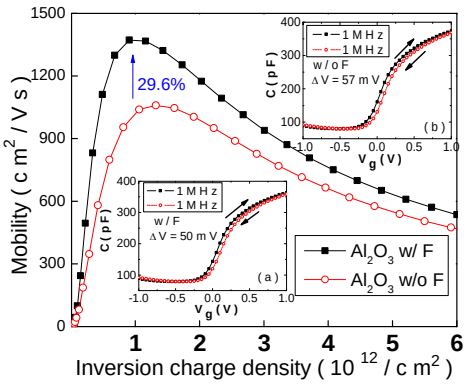


Figure 4.4 Effective channel mobility vs. inversion charge density for $In_{0.53}Ga_{0.47}As$ MOSFETs w/ F plasma treatment (solid) and w/o F plasma treatment (open). Inset (a): split-CV hysteresis for the samples w/ F plasma treatment. Inset (b): split-CV hysteresis for the samples w/o F plasma treatment.

I_d-V_d characteristics of MOSFETs with and without CF_4 plasma treatment are shown in Fig. 4.5. The maximum drive current under V_d = 2.5 V and V_g-V_t =2.5 V for the gate dielectric with and without CF_4 plasma treatment were 32.5 mA/mm and 25.7 mA/mm, respectively (W/L=600 μm/20 μm). The 26.4 % increase in drive current is attributed to the improved effective channel mobility. The D_{it} distribution extracted by full conductance method is shown in Fig. 4.6. It is clear that D_{it} for samples with CF_4 plasma treatment is smaller. The smaller D_{it} after CF_4 plasma treatment is consistent with the smaller SS mentioned earlier.

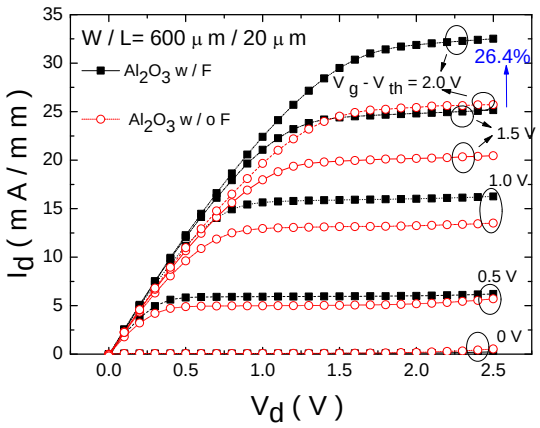


Figure 4.5 I_d-V_d characteristics as a function of various V_g – V_t for $In_{0.53}Ga_{0.47}As$ MOSFETs w/ F plasma treatment (solid) and w/o F plasma treatment (open).

We summarize the comparison of the MOSFETs characteristics with and without CF_4 plasma treatment in Table 4.1.

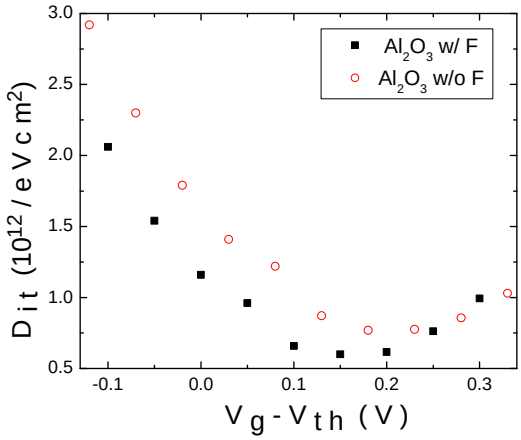


Figure 4.6 D_{it} distribution for $In_{0.53}Ga_{0.47}As$ MOSFETs w/ F plasma treatment (solid) and w/o F plasma treatment (open).

Table 4.1 Effects of post-gate CF_4 plasma treatment on $In_{0.53}Ga_{0.47}As$ nMOSFET

	MOSFETs Characteristics							
Gate stacks 4 nm Al_2O_3 / $In_{0.53}Ga_{0.47}As$	EOT (nm)	V_{th} (V)	Hysteresis in split-CV (mV)	I_d at V_d=2.5V V_g-V_t=2.5V (W/L=600μm/20μm) (mA/mm)	G_m max at V_d = 0.05 V (mS/mm)	Mobility (cm^2/Vs)	SS (mV/decade)	D_{it} min ($eV^{-1}cm^{-2}$)
w/o CF_4 plasma treatment	2.1	0.02	57	25.7	2.5	1059	105.5	7.7×10^{11}
w/ CF_4 plasma treatment	2.2	0.02	50	32.5	2.8	1372	97.9	6×10^{11}

In summary, we have investigated the effects of post-gate CF_4 plasma treatment on Al_2O_3 / $In_{0.53}Ga_{0.47}As$ nMOSFET. The incorporation of F into Al_2O_3 / $In_{0.53}Ga_{0.47}As$ stack is confirmed with SIMS analysis. With F atoms passivation, a better interface quality has been obtained which improves drive currents, transconductance, and mobility by 13.9%, 12.5%, and 29.6%, respectively.

4.3 Effects of fluorine incorporation on the electrical properties of atomic-layer-deposited Al_2O_3 gate dielectric on InP substrate

The previous section has presented the improvements of Al_2O_3/$In_{0.53}Ga_{0.47}As$ MOSFETs. It would be interesting to explore the possibility of its inclusion on another high-k/III-V gate stack. This section will present the impact of F incorporation into Al_2O_3 /InP gate stack. Wider bandgap makes InP immune from the issues related with band-to-band-tunneling, impact ionization, and low On/Off ratio. These characteristics make InP a promising alternative material for continuing the complementary metal-oxide-semiconductor roadmap beyond the 22 nm node. In addition, we will systematically engineer the CF_4 plasma condition. The optimum condition of CF_4 plasma condition has been found, which improves MOSFETs performances best. A possible mechanism will also be discussed.

An experimental investigation on the impact of post-gate CF_4 plasma treatment on InP metal-oxide-semiconductor field-effect transistors is presented. F has been incorporated into the atomic-layer-deposited Al_2O_3 gate dielectric by post-gate CF_4 plasma treatment. The effects of rf power and plasma treatment time were systematically studied. An appropriate amount of F incorporated in Al_2O_3 gate dielectric brings great improvement in electrical characterization. The treatment of rf power of 20W for 5 min is

found to be the optimum condition for the Al_2O_3/InP stack. It is believed to be due to the great reduction of the oxide fixed charge in Al_2O_3 bulk.

N-channel MOSFETs were fabricated on semi-insulating InP (100) substrate with a ring-type structure. The native oxides on InP were removed with 1% diluted HF solution, followed by 20% $(NH_4)_2S$ dip, resulting in a clean S-passivated surface. Then, 10 nm ALD Al_2O_3 dummy capping layer was deposited. After 35 keV $2\times10^{14}/cm^2$ Si ion implantation, the source and drain activation annealing was performed at 750 °C for 15 s. Then, the dummy capping layer was removed using buffered oxide etch. After the same surface preparation (HF and $(NH_4)_2S$), 4 nm Al_2O_3 was deposited by ALD. Some samples were treated by CF_4 plasma, and control samples without CF_4 plasma treatment were also fabricated as references. The flow rate of CF_4 plasma was 50 SCCM and, to avoid possible carbon contamination, O_2 with a flow rate of 5 SCCM was also introduced into the plasma with the pressure of 100 mTorr. PDA was then performed for both CF_4-plasma treated samples and the control samples at 500°C in a N_2 ambient for 60 s. Subsequently, a 200 nm TaN gate electrode was sputtered and followed by gate patterning. The source and drain ohmic contacts were made by an E-beam evaporation of AuGe/Ni/Au and a liftoff process, followed by a 400°C annealing in a N_2 ambient for 30s.

One objective of our research is to investigate the condition of CF_4 plasma post-gate treatment. Two crucial factors, rf power of CF_4 plasma and the plasma treatment time, affect the electrical characteristics significantly. Insufficient plasma treatment might not improve the gate dielectric quality and excessive plasma treatment possibly causes plasma damage and corrodes the improvement. To study the effects of rf power on Al_2O_3 gate dielectrics, some samples were treated in CF_4 plasma for 3 min at different rf power

in the range from 10 to 40W. It was found that the samples which were treated by the power of 20W improved most in terms of mobility performance. With the fixed rf power of 20W, we continued to study the effect of CF_4 plasma treatment time on Al_2O_3 gate dielectrics. Some samples were treated in CF_4 plasma at different treatment times ranging from 1 to 9 min with a fixed rf power of 20W. It was found that with 5 min plasma treatment great improvement was accomplished.

Effect of CF_4 plasma rf power

The drain current and extrinsic transconductance (G_m) as a function of different rf power (CF_4 plasma treatment time: 3 min) are shown in Fig. 4.7, where the gate width (W) is 600 μm and the gate length (L) is 20 μm at V_d = 50 mV. The drain current and maximum G_m of the control sample are 0.49 mA/mm and 0.97 mS/mm, respectively. The drain current and maximum G_m, with the rf power of 20W, reach 0.64 mA/mm and 1.24 mS/mm, respectively, whereas the drain current and maximum G_m roll back with power larger than 20W. Effective channel mobility of InP MOSFETs with different rf powers is plotted in Fig. 4.8. The peak effective channel mobility for the rf power of 20W (777 cm^2/Vs) is 27.4% higher than that of the control samples (610 cm^2/Vs). The improvement in the drain current, maximum G_m, and effective channel mobility is believed to be due to the passivation of oxygen vacancy and fixed charge by an appropriate number of F atoms after introducing CF_4 plasma post-gate treatment [75, 81]. The mechanism will be discussed later in this article. Note that the mobility value of the control samples is consistent with the reported values [80].

SS and equivalent oxide thickness (EOT) with different rf power are shown in Fig. 4.9. MOSFETs have similar EOT (~ 2 nm) with different rf power. SS for the control sample, samples with 10W treatment, and samples with 20W treatment are 81.2, 82.3, 82.4 mV/dec, respectively. SS remains similar for rf power ≦ 20W, suggesting CF_4 plasma treatment might not significantly improve the interface quality. For the rf power of 40W, SS increases to 85.1 mV/dec, indicating that CF_4 plasma treatment even degraded the interface quality. As a result, electrical characteristics are worse compared to the control samples in terms of drain current, maximum G_m, and effective channel mobility. This is possibly due to the damage by excessive plasma while the rf power was 40W.

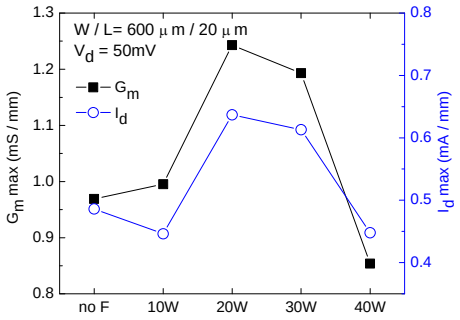


Figure 4.7 Maximum extrinsic transconductance (G_m) and drain current as a function of CF_4 plasma rf power (W/L= 600 μm/20 μm at V_d = 50 mV). CF_4 plasma treatment time: 3 min.

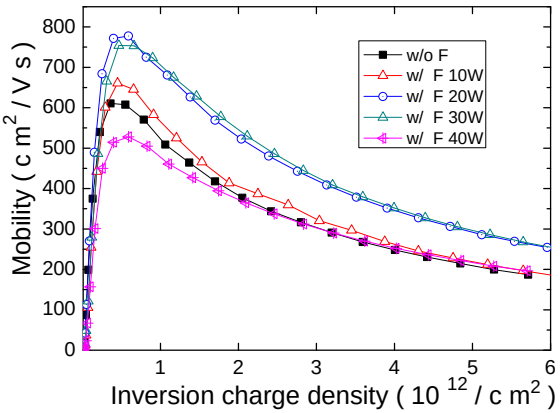


Figure 4.8 Effective channel mobility versus inversion charge density as a function of CF_4 plasma rf power. CF_4 plasma treatment time: 3 min.

With the rf power of 20W, CF_4 plasma treatment improved electrical characteristics most in terms of effective channel mobility while the interface quality remained similar. I_d-V_d characteristics of MOSFETs with and without 20W/3 min CF_4 plasma treatment are shown in Fig. 4.10. The maximum drain current under V_d = 2.5 V and V_g-V_t = 2 V for the gate dielectric with and without CF_4 plasma treatment are 16.7 mA/mm and 12.6 mA/mm, respectively (W/L= 600 μm/20 μm). The 32 % increase in drain current is attributed to the improved effective channel mobility.

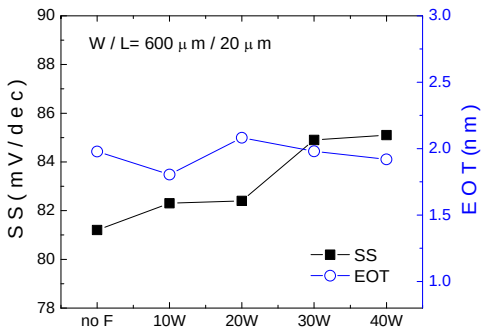


Figure 4.9 SS and EOT as a function of CF_4 plasma rf power. CF_4 plasma treatment time: 3 min.

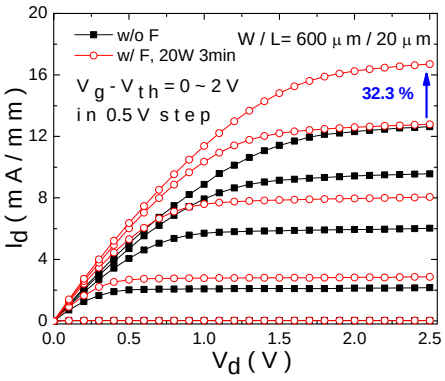


Figure 4.10 I_d-V_d characteristics as a function of various $V_g - V_{th}$ for InP MOSFETs w/o F plasma treatment (solid) and w/ F plasma treatment 20W 3min (open).

Effect of CF_4 plasma treatment time

In order to explore the F distribution in the Al_2O_3 / InP structure with different treatment time (rf power: 20W), the secondary ion mass spectrometry (SIMS) technique was adopted (Fig. 4.11). In the control samples, the F atoms observed in the SIMS profile were possibly the residue of dilute HF dip or the contamination during PDA process. With the increase of CF_4 plasma treatment time, an obvious increase of F concentration was obtained, indicating that F atoms had been incorporated into the Al_2O_3 / InP stack. All other elements in the Al_2O_3 / InP stack remained the same for the samples with and without CF_4 plasma treatment.

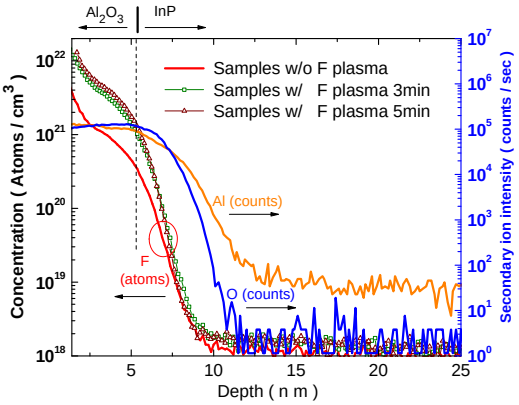


Figure 4.11 SIMS depth profile of the Al_2O_3 / InP structure for fluorine, oxygen, and aluminum atoms distribution. Red solid line: w/o F plasma. Square: w/ F plasma 20W for 3 min. Triangle: w/ F plasma 20W for 5 min.

Figure 4.12 compares the drain current and maximum G_m as a function of plasma treatment times ranging from 1 to 9 min with a fixed rf power of 20W (W/L= 600 μm/20 μm at V_d = 50 mV). The peak values of drain current and maximum G_m are at 5 min plasma treatment, indicating that 5 min is optimal for the Al_2O_3 / InP stack. A similar trend has been obtained in effective channel mobility (Fig. 4.13). With CF_4 plasma treatment 20W for 5 min, mobility has been improved from 610 cm^2/Vs to 912 cm^2/Vs, an increase of 49.5%. SS and EOT as a function of plasma treatment time are shown in Fig. 4.14. MOSFETs have similar EOT (~ 2 nm) with different plasma treatment times. With treatment time less than 7 min, SS values are similar. However, SS value increases after 7 min and 9 min plasma treatment, suggesting that interface quality might be damaged if excessive plasma treatment time is employed.

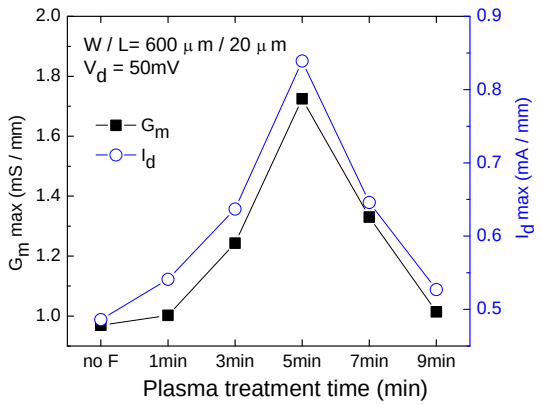


Figure 4.12 Maximum extrinsic transconductance (G_m) and drain current as a function of CF_4 plasma treatment time (W/L= 600 μm/20 μm at V_d = 50 mV). rf power : 20W.

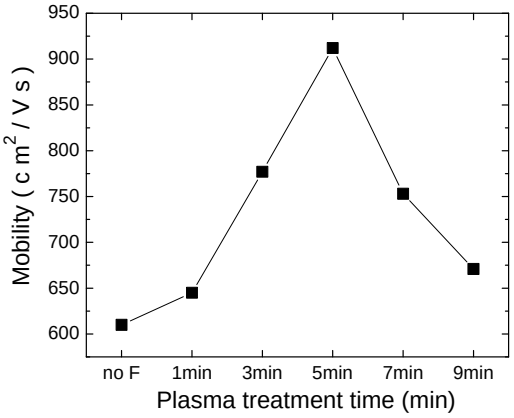


Figure 4.13 Maximum effective channel mobility as a function of CF_4 plasma treatment time. rf power : 20W.

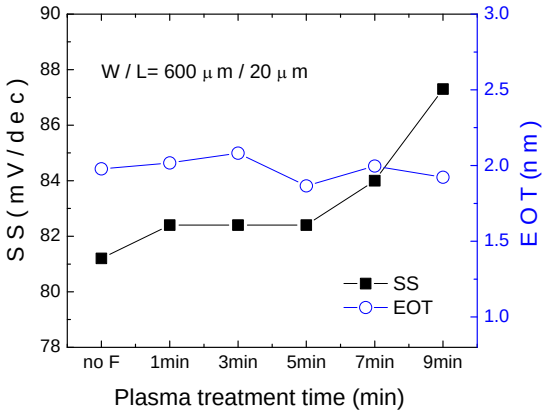


Figure 4.14 SS and EOT as a function of CF_4 plasma treatment time. rf power : 20W.

We compare the drain current and G_m for the sample without and with plasma treatment 1, 3, and 5 min in Fig. 4.15. Both the drain current and G_m increased with the increase of the plasma treatment time. The G_m values of MOSFETs without and with plasma treatment 1, 3, and 5 min are 0.97, 1, 1.24, and 1.73 mS/mm, respectively (W/L = 600 μm/20 μm at V_d = 50 mV). Fig. 4.16 shows the effective channel mobility of InP MOSFETs without and with plasma treatment 1, 3, and 5 min. The peak mobility of MOSFETs without and with plasma treatment 1, 3, and 5 min are 610, 645, 777, and 912 cm^2/Vs, respectively. The increase in mobility is still significant even under high electric field. I_d-V_d characteristics of MOSFETs without and with CF_4 plasma treatment 3 and 5 min are shown in Fig. 4.17. The maximum drain current under V_d = 2.5 V and V_g-V_t =2 V for the MOSFETs without and with CF_4 plasma treatment 3 and 5 min are 12.6 mA/mm 16.7 mA/mm, and 26.2 mA/mm, respectively (W/L = 600 μm/20 μm).

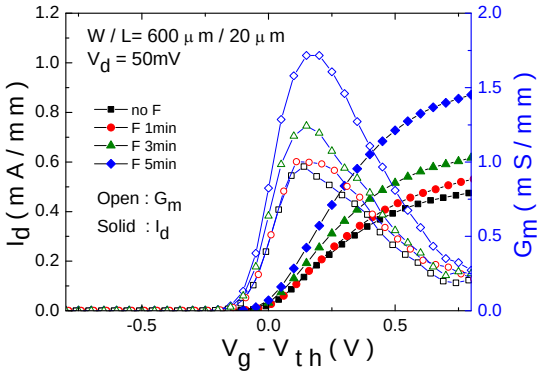


Figure 4.15 The drain current and extrinsic transconductance (G_m) versus $V_g - V_{th}$ as a function of CF_4 plasma treatment time. rf power : 20W.

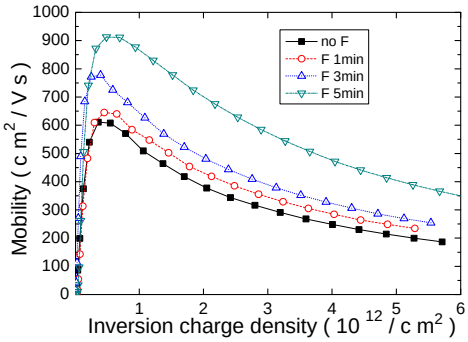


Figure 4.16 Effective channel mobility versus inversion charge density as a function of CF_4 plasma treatment time. rf power : 20W.

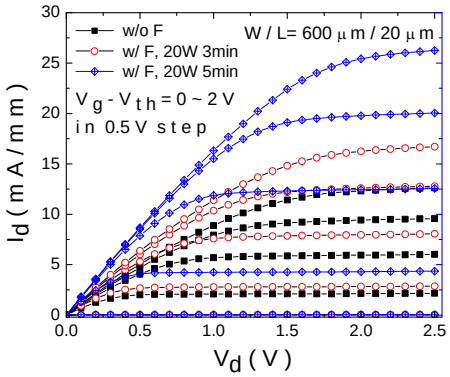


Figure 4.17 I_d-V_d characteristics as a function of various $V_g - V_{th}$ for InP MOSFETs w/o F plasma treatment, w/ F plasma treatment 20W 3 min and 20W 5 min.

Interface quality and oxide fixed charge

We have shown great improvement in electrical characteristics brought by CF_4 plasma treatment. The D_{it} distribution has been extracted using the full conductance method and plotted in Fig. 4.18 and Fig. 4.19. With the rf power of 10W and 20W (CF_4 plasma treatment time: 3 min, Fig. 4.18), D_{it} values remain a similar level with control sample. While the power increases to 30W and 40W, D_{it} values increase slightly. Moreover, in terms of treatment time (rf power: 20W, Fig. 4.19), with plasma treatment D_{it} values remain similar and are independent of treatment time. From D_{it} extraction, it seems that CF_4 plasma treatment might not improve the interface quality on InP substrate whereas previous reports show that CF_4 plasma treatment has improved the interface quality on Si [75 - 77], Ge [78], and InGaAs [81] substrate.

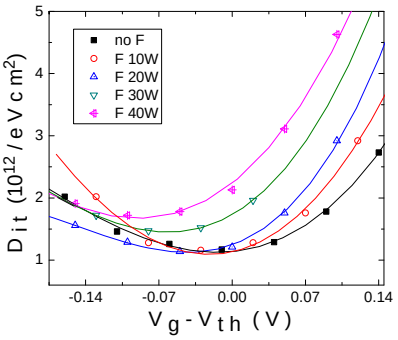


Figure 4.18 D_{it} distribution for InP MOSFETs w/o F plasma treatment and w/ F plasma treatment 10-40W. CF_4 plasma treatment time: 3 min.

We address a possible picture of the interface between Al_2O_3 /InP substrate after CF_4 plasma treatment. It is expected that F will be an attractive alternative passivant for InP because it has much higher binding energy with In (5.25 eV) and P (4.56 eV) compared to In-H bond (2.52 eV) and P-H bond (3.08 eV) [82]. However, after CF_4

plasma treatment 20W for 5 min, SS remains similar (81.2 and 82.4 mV/dec for the control sample and the 20W/5 min treatment, respectively), suggesting that passivation of interface traps might be not significant. Due to the exceptional SS and D_{it} value in the control sample, we believe that Al_2O_3 (without F) has good interface quality with InP substrate and there might be less In- and P- dangling bonds that need to be passivated. As a result, although F still passivates some dangling bonds, the improvement is not significant.

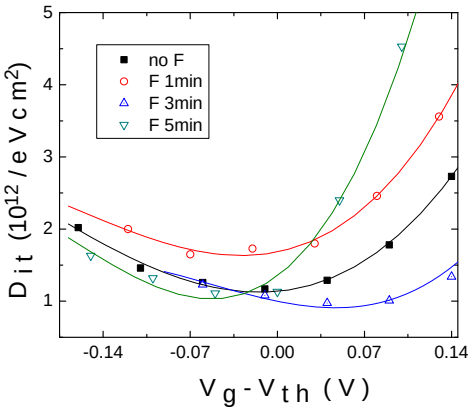


Figure 4.19 D_{it} distribution for InP MOSFETs w/o F plasma treatment and w/ F plasma treatment 1 min, 3 min and 5 min. rf power : 20W.

In order to extract fixed charge in Al_2O_3 bulk, metal-oxide-semiconductor capacitors with varying thickness of Al_2O_3 layer were fabricated. The magnitude of the fixed charge located near the Al_2O_3 /InP interface is related to V_{FB} as follows:

$$V_{FB} = \phi_{MS} - \frac{Q_f}{\kappa \varepsilon_0} t_{ox}$$

where Q_f is the oxide fixed charge in Al_2O_3 bulk, ϕ_{MS} is the difference in work function between the gate metal and the substrate, and t_{ox} is the thickness of the Al_2O_3

gate dielectric. The slope of the V_{FB} versus EOT plot, shown in Fig. 4.20, gives the magnitude and sign of the fixed charge. The extracted fixed charge from the slope is Q_f= + 2.42 × 10^{12} /cm^2 for the sample without CF_4 plasma treatment whereas Q_f= + 1.22 × 10^{12} /cm^2 was obtained for the sample with 3 min CF_4 plasma treatment. With 5 min CF_4 plasma treatment, fixed charge was further reduced to the value of Q_f= 6.33 × 10^{11} /cm^2. The great reduction of oxide fixed charge was obtained with 5 min CF_4 plasma treatment, which probably explains the improvement in electrical characteristics.

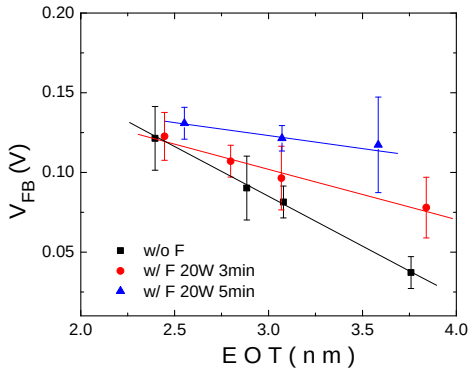


Figure 4.20 V_{FB} versus EOT for the Al_2O_3/InP stack w/o F plasma treatment and w/ F plasma treatment 3 min and 5 min. rf power : 20W.

In conclusion, we have systematically investigated the effects of post-gate CF_4 plasma treatment on Al_2O_3/InP nMOSFET by varying the rf power of CF_4 plasma and the plasma treatment time. The incorporation of F into the Al_2O_3/InP stack has been confirmed with SIMS analysis. By varying the rf power with a fixed 3 min, MOSFETs with 20W plasma treatment improved most significantly. By varying plasma treatment time with a fixed 20W power, 5 min plasma treatment was found to be optimal for the Al_2O_3/InP stack. In our discussion above, significant improvement was not accomplished

with insufficient plasma power or shorter treatment time, whereas plasma damage might result with excessive plasma power or over-treatment time. By applying the appropriate condition, with 20W CF_4 plasma treatment 5 min, MOSFETs had great improvement in drain current, G_m, and effective channel mobility. The improvement is believed to be due to the reduction of oxide fixed charge in high . bulk. Further work, especially a detailed investigation of the physical and electrical behavior of F atoms on InP substrate, is necessary in order to fully understand the effect of F atoms on the Al_2O_3/InP interface.

4.4 Optimization of Fluorine Plasma Treatment for Interface Improvement on $HfO_2/In_{0.53}Ga_{0.47}As$ MOSFETs

The previous section has demonstrated that F incorporation into Al_2O_3/ $In_{0.53}Ga_{0.47}As$ and Al_2O_3/InP gate stacks significantly improves MOSFETs performance. One drawback of Al_2O_3 is its relative lower k value, which cannot meet the requirement of aggressive scale down. Thus, HfO_2, with higher k value, is more desired for future III-V MOSFETs and it is worth to demonstrate fluorine incorporation on HfO_2/III-V gate stacks.

This section will present that notable improvements in the $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack could be achieved by a post-HfO_2 fluorine plasma treatment, including excellent interface quality of low equivalent oxide thickness HfO_2 (1.4 nm) directly on $In_{0.53}Ga_{0.47}As$ without using interface passivation layer, ~5x reduction in interface trap density from 2.8×10^{12} to 4.9×10^{11} $cm^{-2}eV^{-1}$, ~10x reduced border traps from 1.6×10^{19} to 1.6×10^{18} cm^{-3}, and ~ 40% less charge trapping centers. As a result, improved electrical performances have been obtained. Frequency dispersion in capacitance-voltage

characteristics has been reduced. SS has been improved from 127 to 109 mV/dec. Effective channel mobility has been enhanced from 826 to 1067 cm^2/Vs. An improved drive current of 123 mA/mm at V_d = 2.5 V and V_g-V_{th} =2 V (5 μm channel length) has also been presented in this section.

High-mobility III-V semiconductors along with high-k gate dielectrics are projected to be key ingredients in future complementary metal-oxide-semiconductor technology. Among these, $In_{0.53}Ga_{0.47}As$ has been intensively studied for their advantages in high electron mobility over their Si-based counterparts. $In_{0.53}Ga_{0.47}As$ MOSFETs have been demonstrated to provide large drive current density [83 - 84] with acceptable SS [85]. Even so, the devices aimed towards the sub 22 nm nodes will require good control of short channel effects which in turn imposes stringent requirements of reduced EOT, reduced interface trap densities and charge trapping in the high-k bulk.

Significant progress of interface improvement was achieved by proper surface pre-treatment and interface passivation layer [86 - 88], which has been presented in Chapter 2 and Chapter 3. Those approaches involve processes performed prior to high-k deposition. Interface states formed during the high-k deposition need to be suppressed by a post-treatment. Hydrogen (H) passivation by forming gas annealing (FGA) is an effective post-treatment for Si-based MOSFETs devices. Several groups have demonstrated FGA on $In_{0.53}Ga_{0.47}As$ and shown that annealing in hydrogen can compensate most of the positive bulk fixed charge, the negative interfacial fixed charge, and interface states [87]. However, one concern of FGA is the weak H-terminated bonds. The trap creation phenomena in the oxide are related to H release by hot electrons from defect sites near gate electrode/oxide interface [88]. This atomic, ionic, or molecular H

then can move toward oxide/substrate interface with some entering the substrate. The mobile species can then de-passivate states on either side of this interface which act as generation-recombination centers [89]. In addition, during electrical stress/operation, weak H bonds are liberated (meanwhile, dangling bonds are re-activated), resulting in oxide degradation and negative bias temperature instability [90 - 91]. Thus, a post-treatment with another species which can provide stronger binding energy is highly desired.

F incorporation has been demonstrated on Si, Ge, and $In_{0.53}Ga_{0.47}As$ substrates that F can passivate high-k bulk traps and interface defects at high-k/substrate (Si, Ge and $In_{0.53}Ga_{0.47}As$) interface [75 - 76], [81], [92 - 93]. Although the insertion of a thin interface passivation layer can improve interface quality [84], those layers usually have relative lower k value [94], which may hinder EOT scaling and may not meet the requirement for the sub 22 nm nodes.

We have demonstrated better electrical performance with CF_4 plasma treatment on $Al_2O_3/In_{0.53}Ga_{0.47}As$ gate stack in the previous sections [81]. However, the use of HfO_2 would provide thinner EOT than Al_2O_3 because of its higher k value. In this paper, we demonstrated the impact of post-HfO_2 CF_4 plasma treatment on $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack through detailed physical analysis, electrical characterization and device performance discussion. With F incorporation, we have successfully developed excellent interface quality of HfO_2 directly on $In_{0.53}Ga_{0.47}As$ without using interface passivation layer. Fluorinated samples exhibit low D_{it} of 4.9×10^{11} $cm^{-2}eV^{-1}$, which is the lowest value over prior reported $HfO_2/In_{0.53}Ga_{0.47}As$ gate stacks. HfO_2 bulk quality is also improved.

Consequently, significant improvements in effective channel mobility (μ_{eff}), drive current (I_d), and SS have been achieved.

Figure 4.21 shows the cross-sectional schematic of the device structure and the illustration of F incorporation. P-type (Zn-doped, $3\times10^{18}/cm^3$) InP wafers were used as the starting substrates. P-type (Be-doped, $5\times10^{16}/cm^3$) $In_{0.52}Al_{0.48}As$ of 100 nm thick was grown by molecular beam epitaxy as a buffer layer, followed by a 300 nm p-type $In_{0.53}Ga_{0.47}As$ layer (Be-doped, $5\times10^{16}/cm^3$). The native oxides were removed with 1% diluted HF solution, followed by 20% $(NH_4)_2S$ sulfur passivation pre-treatment. A 10-nm-thick atomic-layer-deposited (ALD) Al_2O_3 was deposited at a substrate temperature of 200°C as an encapsulation layer. For device fabrication, S/D regions were selectively implanted with a Si dose of $2\times10^{14}/cm^2$ at 35 keV. The S/D activation annealing was performed in nitrogen ambience at 700°C/10s. The encapsulation layer was then removed using buffered oxide etch solution. A 5 nm-thick ALD HfO_2 film was deposited after the same surface preparation (HF and $(NH_4)_2S$). Some samples were treated *ex situ* with CF_4 plasma under low radio-frequency power of 30W for 5 min. A mixed flow of CF_4 and O_2 gas (ratio ~ 10:1) was introduced into the chamber with pressure of 100 mTorr. Control samples without CF_4 plasma treatment were also fabricated as references. PDA was then performed for all the samples at 500°C for 60 s in a nitrogen ambience. Subsequently, a 200 nm TaN gate electrode was sputtered and followed by gate patterning. AuGe/Ni/Au alloy was deposited by E-beam evaporation and a liftoff process for the S/D Ohmic contacts; backside contact was made by E-beam evaporation of Cr/Au, followed by annealing at 400°C for 30 s in a nitrogen ambience.

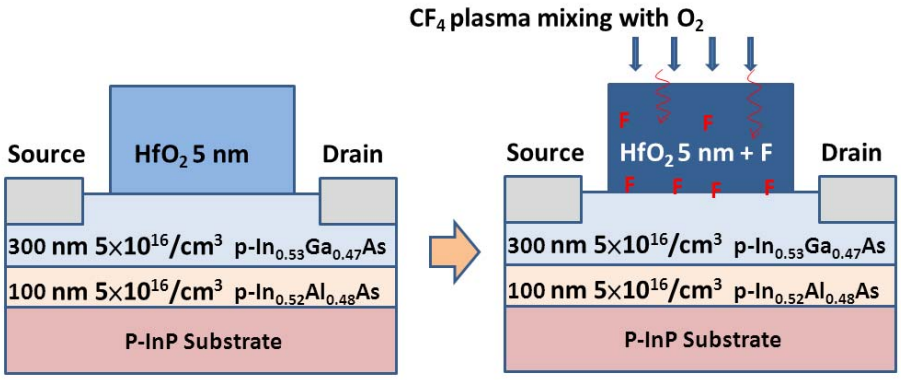


Figure 4.21 The device cross-sectional structure of $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack. Samples were treated in a mixed flow of CF_4 and O_2 gas (ratio ~ 10:1) with varied RF power and treatment time.

Figures 4.22 and 4.23 show the X-ray photoelectron spectroscopy (XPS) spectra of F *1s* and Hf 4*f* for the $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack with and without CF_4 plasma treatment. All the detected binding energy was calibrated by the C 1*s* signal at 284.5eV. The inset of Fig. 2 shows the F 1*s* spectra of sample right after CF_4 plasma treatment (before PDA). The peak signal located at ~ 685eV corresponds to the F bonds in the bulk HfO_2 [76], indicating that F was incorporated into the HfO_2 after CF_4 plasma treatment. After PDA, some amount of F out diffused during the annealing process, resulting in the drop of F 1*s* signal. The F concentration remaining in the HfO_2 is estimated to be 2.7 at. %, which is lower than that used in the low-k technology [95]. Compared to the control sample, the fluorinated sample has an increased binding energy by 0.2eV and 0.3eV for the Hf $4f_{7/2}$ and Hf $4f_{5/2}$ signal, respectively. This suggests that parts of the oxygen vacancies were terminated by the incorporated F atoms to form stronger Hf-F bonds with higher binding energy [96]. To confirm the F distribution/movement during CF_4 plasma treatment, SIMS technique was performed on $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack, as shown in Fig. 4.24. Right after CF_4 plasma treatment (before PDA), F atoms mainly stayed around

1.7 nm below the HfO_2 surface. As being deeper into the HfO_2 bulk, F concentration dropped to the same level as that in the control sample (near the $HfO_2/In_{0.53}Ga_{0.47}As$ interface). However, after PDA process, F atoms were driven into the deeper HfO_2 bulk and piled up at the $HfO_2/In_{0.53}Ga_{0.47}As$ interface. Due to Ga or As dangling bonds and Ga-Ga or As-As like-atom bonds created by oxidation [97], defective bonds generated at the interface. F atoms tended to passivate these defective bonds at the interface, resulting in high F concentration at the $HfO_2/In_{0.53}Ga_{0.47}As$ interface. All other elements in the $HfO_2/In_{0.53}Ga_{0.47}As$ stack remained the same for the samples with and without F incorporation.

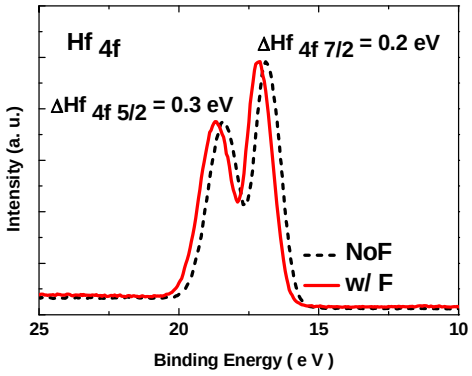


Figure 4.22 XPS analysis of the Hf 4*f* electronic spectra for the samples with and without F incorporation.

One concern of doping F into gate oxide is that dielectric constant of the gate oxide would decrease with heavily F incorporation. From XPS, the F concentration in our fluorinated HfO_2 is estimated to be 2.7 at. %, which is lower than that used in the low-k technology [95]. Therefore, the dielectric constant of HfO_2 remains similar (~17) after F incorporation, as shown in Fig. 4.25. The inset of Fig. 4.25 compares the gate leakage

current of samples with and without F incorporation. The gate leakage current is slightly reduced with F incorporation.

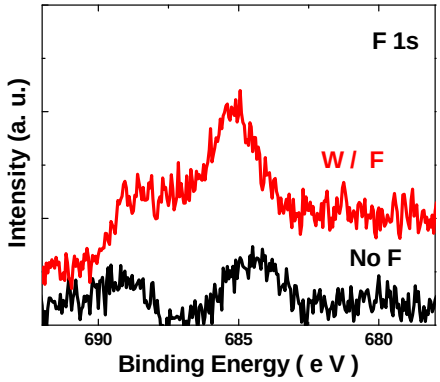


Figure 4.23 XPS analysis of the F 1*s* electronic spectra for the control sample and fluorinated sample.

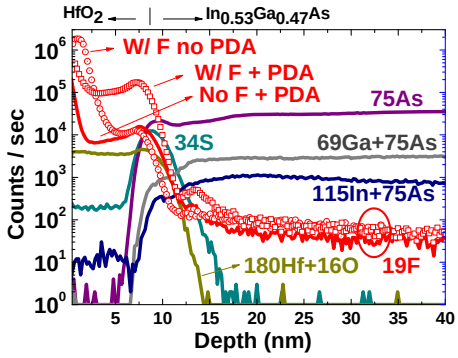


Figure 4.24 SIMS depth profiles of F, S, Hf, O, In, Ga, and As in the fluorinated $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack w/ and w/o PDA. Movement of F atoms is clearly observed. F atoms piled up at the $HfO_2/In_{0.53}Ga_{0.47}As$ interface after PDA. The control sample with PDA only is also presented.

One objective of this section is to investigate the optimum condition of CF_4 plasma. The RF power and the plasma treatment time are two critical factors that affect the electrical characteristics significantly. Insufficient plasma treatment might not improve the gate dielectric quality, whereas excessive plasma treatment possibly causes

plasma damage and corrodes the improvement. To study the effects of RF power on HfO_2 gate dielectrics, some samples were treated in CF_4 plasma for 3 min at different RF power in the range from 20 to 40W. It was found that the samples treated by the power of 30W improved most in terms of G_m, I_d and SS. With the fixed RF power of 30W, we continued to study the effect of CF_4 plasma treatment time. Some samples were treated in CF_4 plasma at different treatment times ranging from 1 to 7 min with a fixed RF power of 30W. It was found that 5 min plasma treatment further improved SS.

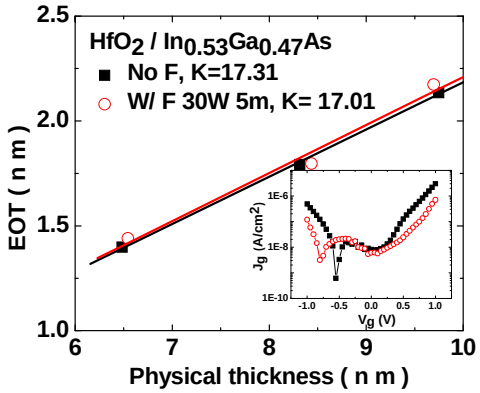


Figure 4.25 EOT versus physical thickness for the samples with and without F incorporation. Inset: the gate leakage current comparison of samples with and without F incorporation.

The maximum G_m and I_d as a function of different RF power (CF_4 plasma treatment time: 3 min) are shown in Fig. 4.26 (W/L = 600/20 μm, at V_d = 50 mV and 0.5 V). The maximum G_m and I_d of the control sample are 3.1 mS/mm and 1.3 mA/mm (V_d = 50 mV), and 20.3 mS/mm and 12 mA/mm (V_d = 0.5 V) respectively. With F plasma treatment of 30W, the maximum G_m and I_d reach 3.7 mS/mm and 1.7 mA/mm (V_d = 50 mV), and 26 mS/mm and 16.2 mA/mm (V_d = 0.5 V), respectively. However, the maximum G_m and I_d roll back with power larger than 30W indicative of possible plasma

damage. SS data with different RF power are shown in Fig. 4.27. $In_{0.53}Ga_{0.47}As$ MOSFETs have similar EOT (~ 1.4 nm) with different RF power treatment. With F incorporation, SS has been improved from 127 to 118.1 mV/dec (as shown in the inset of Fig. 4.27), which suggests that the interface quality has been improved. For the RF power of 40W, SS increases to 127.5 mV/dec, indicating that excessive CF_4 plasma treatment degrades the interface quality.

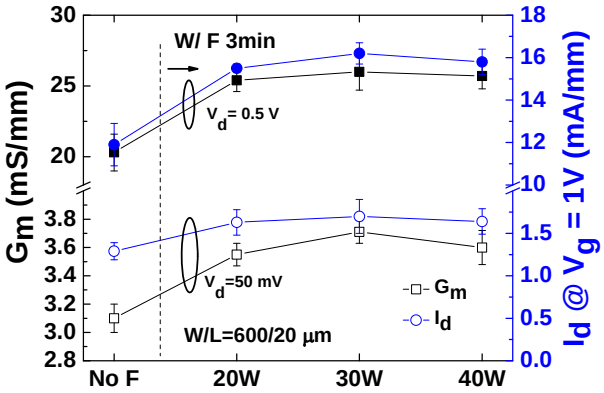


Figure 4.26 Maximum G_m and I_d as a function of CF_4 plasma RF power (W/L= 600 µm/20 µm at V_d = 50 mV and 0.5V). CF_4 plasma treatment time: 3 min.

Figure 4.28 compares the maximum G_m and I_d as a function of plasma treatment times ranging from 1 to 7 min with a fixed RF power of 30W (W/L= 600 µm/ 20 µm at V_d = 50 mV and 0.5 V). 3 min plasma treatment reaches the peak values of the maximum G_m and I_d, whereas 5 min plasma treatment achieves the lowest SS value (Fig. 4.29). The inset of Fig. 4.28 compares the I_d–V_g curves (in log-linear scale) of the control sample and the sample with F treatment 30W/5min. A steeper SS slope is clearly observed.

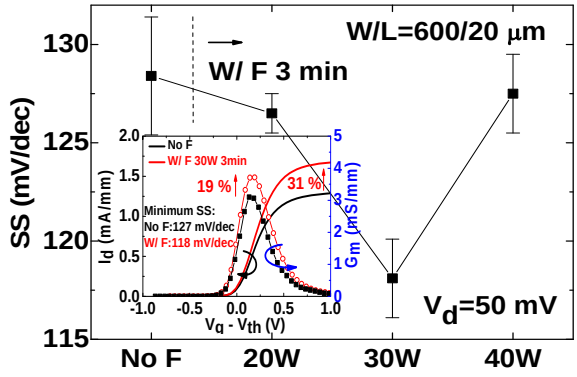


Figure 4.27 SS as a function of CF_4/O plasma RF power. CF_4 plasma treatment time: 3 min. Inset: I_d-V_g comparison of the control sample and the sample with CF_4 plasma treatment for 30W/3min.

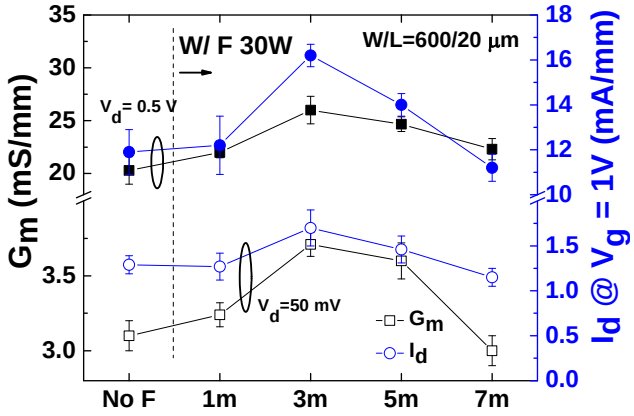


Figure 4.28 Maximum G_m and I_d as a function of CF_4 plasma treatment time (W/L= 600 μm/20 μm at V_d = 50 mV and 0.5 V). RF power: 30W.

Effective channel mobility (μ_{eff}) of $In_{0.53}Ga_{0.47}As$ MOSFETs with different plasma treatment times are plotted in Fig. 4.30. The peak μ_{eff} of 30W/3min reaches 1144 cm^2/Vs, which is 38% improvement compared to the control samples (826 cm^2/Vs). The improvements in the I_d, G_m, SS and μ_{eff} are believed to be due to the improved interface quality by an appropriate amount of CF_4 plasma post-HfO_2 treatment.

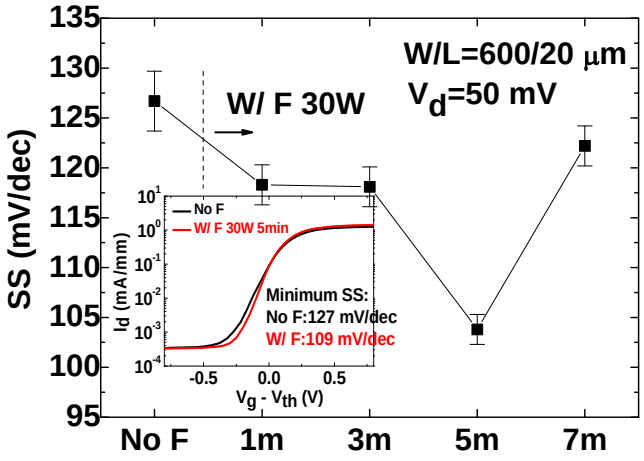


Figure 4.29 SS as a function of CF_4 plasma treatment time. RF power : 30W. Inset: I_d-V_g comparison of the control sample and the sample with CF_4 plasma treatment for 30W/5min.

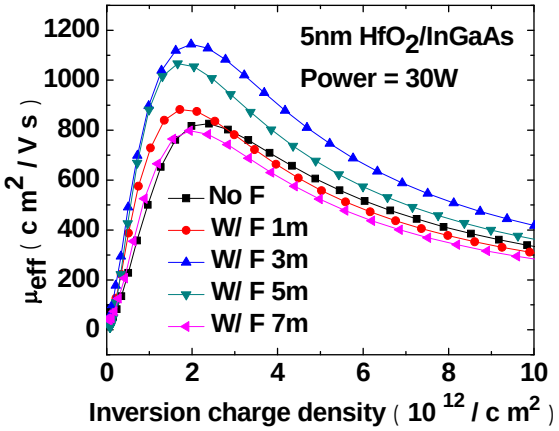


Figure 4.30 Effective channel mobility versus inversion charge density as a function of CF_4 plasma treatment time. RF power: 30W.

We notice that the plasma damage occurred if excessive plasma (either plasma wattage or treatment time) was applied. The plasma damage could come from disordering, surface roughening, and fluorine contamination [98]. The disordering layer contains

dangling bonds and broken bonds, which would scatter the electrons underneath (in the channel) and lower the electron mobility. Severe disordering would consider surface roughness, which results in more dangling bonds and broken bonds. If the F concentration is too high (> 5 at.%, [95]), the dielectric constant of HfO_2 decreases, resulting in lowering drive current.

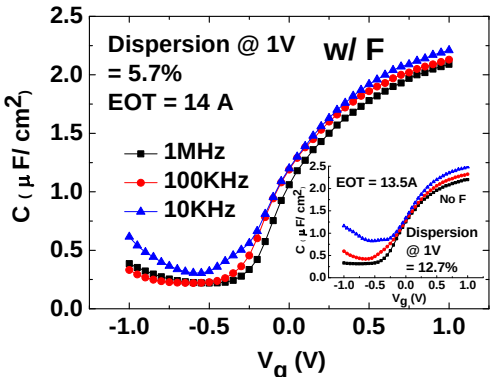


Figure 4.31 Split-CV curves of the fluorinated sample. Frequency dispersion at V_g = 1V is reduced from 12.7% to 5.7% with F treatment. Inset: Split-CV curves of MOSFETs

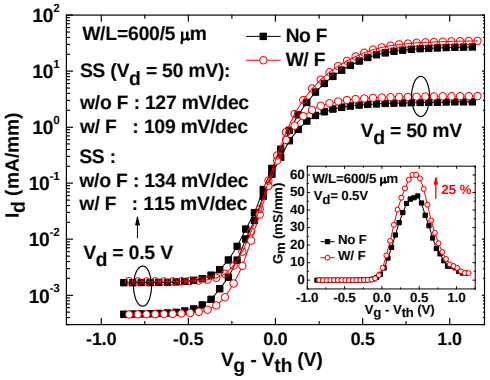


Figure 4.32 I_d comparison under high field (V_d = 0.5 V) and low field (V_d = 50 mV) for the sample w/ and w/o F incorporation. Significant improvements have been achieved in I_d and SS with F incorporation. Inset: their corresponding G_m curves under high field.

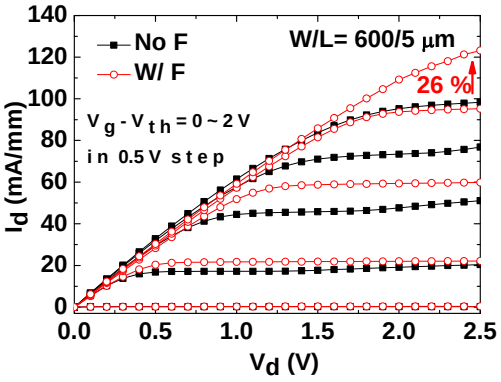


Figure 4.33 I_d-V_d comparison at $V_g = V_{th}$ to V_{th} + 2 V for the control and the fluorinated sample.

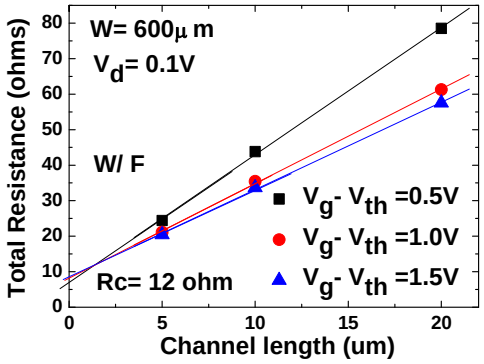


Figure 4.34 Series resistance extracted from the total resistance versus gate length at V_g - V_{th} = 0.5, 1, and 1.5 V for devices with 600 μm width.

Their corresponding multi-frequency split C-V curves are shown in Fig. 4.31. Small amount of F incorporated did not change the dielectric constant of HfO_2 (~17.3, data not shown) and EOT remained similar (1.35 nm for the control sample and 1.4 nm for the fluorinated sample). It has been known that Q_{br} are responsible for large accumulation CV dispersion in multi-frequency CV measurements [99]. When the MOSFET channel is inverted, electrons accumulated in the conduction band can tunnel to the border traps in the HfO_2. After F treatment, frequency dispersion at V_g = 1 V is

reduced from 12.7% to 5.7% indicative of reduced Q_{br}, which is consistent with data shown in Fig. 4.29. The improvement of CV dispersion at V_g = -0.5 V is believed to be due to the reduction of D_{it} [100]. The interface traps can degrade μ_{eff} due to trapping inversion carriers and coulomb scattering, especially at low inversion charge densities/electrical field [41, 101]. F incorporation effectively reduces interface traps, resulting in less trapping inversion carriers and coulomb scattering, and thereby improves μ_{eff}. Since the effective channel mobility of the control sample is in good agreement with our previous report [84], we can ascribe the improvement only to the F incorporation.

Figure 4.32 compares I_d at low field and high field for the fluorinated and non-fluorinated samples (W/L = 600 μm/ 5 μm). With F plasma treatment, I_d and G_m (at V_d = 0.5 V, V_g-V_{th}=1 V) has been improved from 27 to 34.7 mA/mm and from 48 to 60 mS/mm (Inset of Fig. 4.32), respectively. These are attributed to the improved effective channel mobility. Note that the SS was reduced after F plasma treatment (from 127 to 109 mV/dec), indicating improved interface quality. This is consistent with the data shown in the section *B*. I_d-V_d characteristics are shown in Fig. 4.33. The maximum I_d at V_d = 2.5 V and V_g-V_{th} =2 V for the control sample and fluorinated sample are 98 and 123 mA/mm, respectively (W/L = 600 μm/ 5 μm). The S/D series resistance of fluorinated samples was estimated to be about 12 Ω for the 600 μm gate width (i.e., ~ 7200 Ω · μm, and ~ 6600 Ω · μm for the control sample, data not shown), as shown in Fig. 4.34. These values are much higher than those of state-of-the-art devices with optimized S/D activation condition [102 - 103]. A higher drive current can be obtained through careful S/D engineering to minimized S/D contact resistance.

To evaluate the improvement on the HfO_2 bulk quality, pulsed I_d-V_g measurement (setup shown in the inset of Fig. 4.35, W/L = 600 μm/5 μm) was conducted [104]. Note that the rise/fall time of the gate pulse (100 μs) was possibly the fastest pulse time on InGaAs that produced clean data [104]. Instead of DC gate bias, by which almost all the defects follow, by applying pulsed gate bias, only defects with response time applicable for the pulse time used would follow gate bias and degrade pulsed I_d-V_g current. Higher pulsed I_d-V_g current in fluorinated samples suggests that those defects have been reduced with F incorporation. During the high voltage within the pulse (transistor turned on), electrons were trapping in the HfO_2, causing a right shift during backward I_d-V_g pulse scan. The pulsed I_d-V_g hysteresis for the control and fluorinated samples are 0.228 and 0.138 V, respectively. ~ 40% reduction of pulsed I_d-V_g hysteresis indicates reduced charge trapping in the fluorinated HfO_2, as shown in Fig. 4.35. The data suggests that F incorporation reduces charge-trapping centers (possibly fixed charge) in the HfO_2, providing better HfO_2 bulk quality.

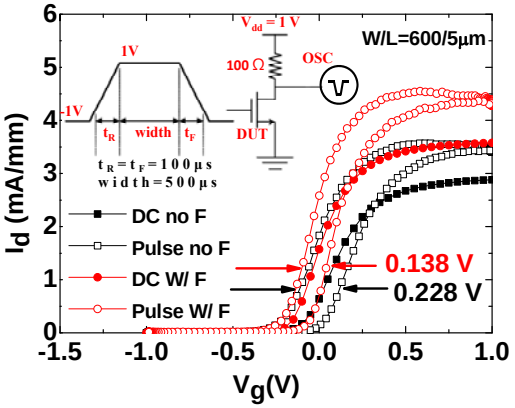


Figure 4.35 DC and pulsed I_d-V_g at V_d=50mV w/ and w/o F incorporation. Smaller I_d-V_g hysteresis indicates less charge trapping after CF_4 plasma treatment. Inset: Configuration of pulsed I_d-V_g measurement.

Evaluation of border traps (Q_{br}) of samples with and without F incorporation was carried out by charge pumping measurement with low gate pulse frequencies. Q_{br} locate inside the high-k (near interface) with long time constants as they interact with the conduction band electrons through tunneling, therefore, Q_{br} are unable to follow high frequency signal (1MHz). With low gate pulse frequencies, the conduction band electrons are allowed to tunnel deeper into HfO_2 bulk and then, trapped by Q_{br}. This contributes to the traps density increase from 1MHz to 10KHz [101]. Figure 4.36 illustrates the reduced Q_{br} by using F (~10x reduction from 1.6×10^{19} to 1.6×10^{18} cm^{-3}), indicating better HfO_2 bulk quality with F incorporation.

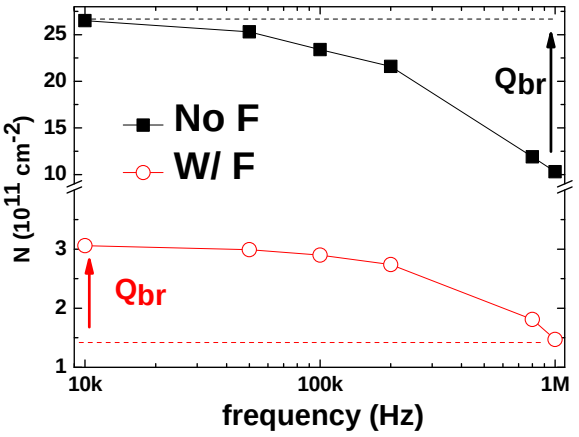


Figure 4.36 Multiple frequency charge pumping spectra for the samples w/ and w/o F incorporation. Q_{br} can be scanned by reducing measured frequency to probe deeper into HfO_2 bulk

To accurately characterize the interface quality of fluorinated and non-fluorinated MOSFETs, charge pumping measurement were employed (at room temperature). S/D were grounded while sweeping the base level (V_{base}: -2.3 V to 1 V in a step of 50 mV) of gate pulse (with a constant-amplitude, 1 V) at different frequencies. The region of the bandgap probed was from electron emission energy level to hole emission energy level,

which was around the midgap. The charge pumping current (I_{cp}) is proportional to D_{it}. Figure 4.37 exhibits lower I_{cp} for the fluorinated sample indicative of reduced D_{it}. The mean D_{it} values were extracted by changing rise time (t_R) and fall time (t_F), according to the following equation [105 - 106],

$$\frac{I_{cp}}{f} = 2qD_{it}AkT\left\{\ln\sqrt{t_R t_F} + \ln\left(\frac{\left|V_{fb} - V_t\right|}{\left|\Delta V_g\right|}V_{th}n_i\sqrt{\sigma_n\sigma_p}\right)\right\}$$

where q is the electronic charge, A is the transistor gate area (2.08×10^{-4} cm^2 in our devices), k is the Boltzmann constant, V_{fb} is the flat band voltage, V_t is the threshold voltage, ΔV_g is the gate pulse amplitude, V_{th} is the thermal velocity of the carriers, n_i is the surface concentration of minority carriers, and σ_n and σ_p are the capture cross sections of electrons and holes, respectively. The mean D_{it} values were extracted from the slope of I_{cp}/f versus $\ln[(t_R\times t_F)^{1/2}]$, as shown in Fig. 4.38. It has been found that the mean D_{it} value was reduced ~5× from 2.8×10^{12} to 4.9×10^{11} $cm^{-2}eV^{-1}$ after F plasma treatment. F atoms possibly passivate dangling bonds and oxygen vacancies in the $HfO_2/In_{0.53}Ga_{0.47}As$ interface and thereby reduce D_{it} value. Table 4.2 compares D_{it} values of high-k/$In_{0.53}Ga_{0.47}As$ gate stacks in recent publications. The D_{it} value reported in this paper has the lowest value in prior reported $HfO_2/In_{0.53}Ga_{0.47}As$ gate stacks. In addition, compared to previous work on insertion of interface passivation layers [85], F incorporation can achieve ~3.6x lower D_{it} value, suggesting that F incorporation is more effective to improve interface quality. Moreover, it has the advantage of EOT scaling.

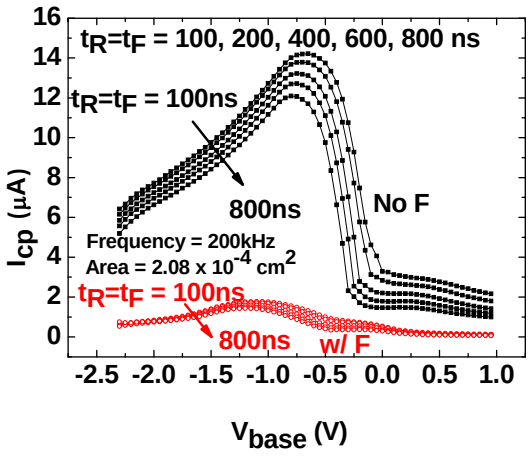


Figure 4.37 Charge pumping measurement with rise/fall time dependence. Samples with F incorporation show much smaller I_{cp}. V_{base}: -2.3 V to 1 V in a step of 50 mV and the pulse amplitude is 1 V.

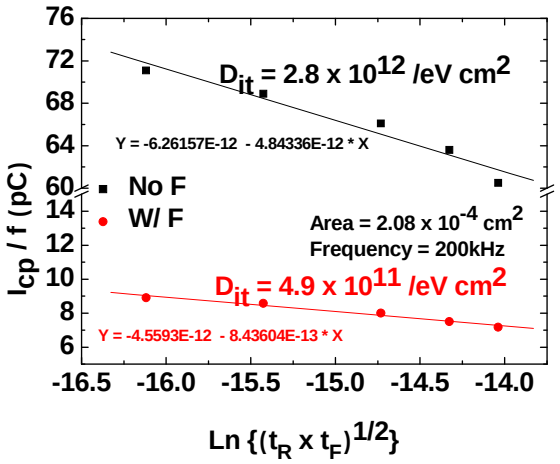


Figure 4.38 Q_{cp} (=I_{cp}/f) vs. ln[(t_R×t_F)$^{1/2}$]. The mean D_{it} value is extracted by linear fitting according to [86 - 87]. Samples with F incorporation show less D_{it} value of 4.9×10^{11}/eVcm^2.

Table 4.2 Comparison of the electrical and interfacial properties of this work with some recently reported paper.

High k	Passivation method	Channel material	L_G	EOT or thickness of high-k	D_{it} ($cm^{-2}eV^{-1}$)	SS (mV/dec)	Ref.
Al_2O_3*	-	$In_{0.53}Ga_{0.47}As$	-	4.2nm	5×10^{11}	-	107
Al_2O_3	-	$In_{0.53}Ga_{0.47}As$	-	10nm of Al_2O_3	2.5×10^{11}	-	42
Al_2O_3	-	$In_{0.53}Ga_{0.47}As$	1.5 μm	8 nm of Al_2O_3	1×10^{12}	>200	108
Al_2O_3	-	$In_{0.53}Ga_{0.47}As$	0.5 μm	30nm of Al_2O_3	1.4×10^{12}	240	60
HfO_2*	-	$In_{0.53}Ga_{0.47}As$	-	2.1nm	1×10^{12}	-	107
HfO_2*	-	$In_{0.53}Ga_{0.47}As$	-	7.8nm of HfO_2	2×10^{12}	-	109
HfO_2*	Al-doped	$In_{0.53}Ga_{0.47}As$	-	8-9nm of HfO_2	6×10^{12}	-	110
HfO_2	PH_3	$In_{0.53}Ga_{0.47}As$	4 μm	1.7nm	8.6×10^{11}	103	102
HfAlO	SiH_4+NH_3	$In_{0.53}Ga_{0.47}As$	2-10 μm	3.8nm	6.5×10^{11}	155-210	86
ZrO_2	$LaAlO_3$	$In_{0.53}Ga_{0.47}As$	5 μm	1.63nm	7.5×10^{11}	116	101
HfO_2	CF_4/O_2 post treatment	$In_{0.53}Ga_{0.47}As$	5-20 μm	1.4nm	4.9×10^{11}	109	This work
HfO_2	Control			1.35nm	2.8×10^{12}	127	This work

* Capacitor Structure.

The effects of post-HfO_2 F treatment on $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack have been extensively studied. XPS and SIMS analysis verified the presence of stronger Hf-F bonds

and higher F concentration at the $HfO_2/In_{0.53}Ga_{0.47}As$ interface. The gate stack quality has been notably improved by F incorporation. The mean D_{it} value has been reduced ~5x from 2.8×10^{12} to 4.9×10^{11} $cm^{-2}eV^{-1}$. Q_{br} have shown ~10x reduction from 1.6×10^{19} to 1.6×10^{18} cm^{-3}, resulting in reduced frequency dispersion for the fluorinated samples. ~40% less I_d-V_g hysteresis in pulsed I_d-V_g measurement indicates less charge trapping centers in the fluorinated HfO_2. Therefore, enhanced electrical performance have been presented: reduced SS from 127 to 109 mV/dec, enhanced μ_{eff} from 826 to 1067 cm^2/Vs, and improved I_d from 98 to 123 mA/mm at V_d = 2.5 V and V_g-V_{th} =2 V (5 μm channel length). These results suggest that the post-HfO_2 F treatment could be a key technique to implement high performance III-V MOSFETs for the sub 22 nm nodes.

4.5 Fluorinated HfO_2 gate dielectric engineering on $In_{0.53}Ga_{0.47}As$ metal-oxide-semiconductor field-effect-transistors

We have understood that F atoms can passivate defects in high-k bulk and at high-k/III-V interface, therefore, improve high-k/III-V MOSFETs performance. Even so, we are not satisfied with our current improvements. The following question would be: If more F atoms are incorporated into high-k/III-V gate stack, will the MOSFETs performance improve even better? How to incorporate more F atoms into high-k/III-V gate stack?

This section will discuss a simple technique by a “two-step” high-k deposition and incorporating F between two high-k depositions. As a result, we can increase F concentration in the high-k/III-V gate stack and further improve gate stack quality. Our experiment design is the following: We introduce F into gate stacks in two ways: CF_4 plasma treatment has been performed after 8 nm HfO_2 deposition or after partial

deposition of HfO_2 4 nm and followed by 4 nm HfO_2 deposition to make the overall thickness 8 nm. The latter one incorporates more F into the HfO_2 and shows great improvements. Fluorine is believed to reduce fixed charge by forming Hf-F bonds in the HfO_2 bulk and also passivate the interface traps in the $HfO_2/In_{0.53}Ga_{0.47}As$ interface.

To engineer the F profile in the HfO_2 gate dielectric, CF_4 plasma treatment was performed after full 8 nm HfO_2 deposition (denoted as post-F) or after partial deposition of HfO_2 4 nm and followed by 4 nm HfO_2 deposition to make the overall thickness 8 nm (denoted as mid-F). In the post-F samples, the maximum drain current density of 128 mA/mm and G_m of 10.6 mS/mm for a 5-μm gate-length MOSFET are achieved, which has a significant improvement over the samples without F treatment. Inspiringly, the maximum drain current density and G_m in the mid-F samples are even higher than those in the post-F samples, which are believed to be due to the F profile in the mid-F samples.

$In_{0.53}Ga_{0.47}As$ MOSFETs were fabricated with ring-type structure on a 300 nm p-doped $In_{0.53}Ga_{0.47}As$ layer (Be-doped, $5\times10^{16}/cm^3$) epitaxially grown on a 2 in. InP P-substrate. The native oxides on InGaAs were removed with 1% diluted HF solution, followed by 20% $(NH_4)_2S$ pretreatment. A 10-nm-thick ALD Al_2O_3 was deposited at a substrate temperature of 200 °C as an encapsulation layer. For device fabrication, source and drain regions were selectively implanted with a Si dose of $2\times10^{14}/cm^2$ at 35 keV. After source and drain activation annealing performed at 700 °C for 10 s in a nitrogen ambient, the encapsulation layer was removed using buffered oxide etch solution. An 8 nm ALD HfO_2 film was deposited after the same surface preparation (HF and $(NH_4)_2S$). Some samples (post-F) were treated by postgate CF_4 plasma at radio-frequency (rf)

power of 30W with pressure of 100 mTorr and flow rate of 50 SCCM for 3 minutes. To avoid possible carbon contamination, O_2 with a flow rate of 5 SCCM was also introduced into the plasma. To manipulate the F profile in the HfO_2 gate dielectric, some samples (mid-F) were treated by CF_4 plasma treatment after partial deposition of HfO_2 4 nm and followed by 4 nm HfO_2 deposition to make the overall thickness of 8 nm. Control samples without any CF_4 plasma treatment were also fabricated as references. PDA was then performed for all the samples at 500°C for 60 s in a nitrogen ambient. Subsequently, a 200 nm TaN gate electrode was sputtered and followed by gate patterning. The source and drain Ohmic contacts were made by an E-beam evaporation of AuGe/Ni/Au and a liftoff process; backside contact was made by an E-beam evaporation of Cr/Au, followed by a 400°C annealing for 30 s in a nitrogen ambient. The schematic cross section of the device structure of post-F and mid-F is shown in Fig. 4.39.

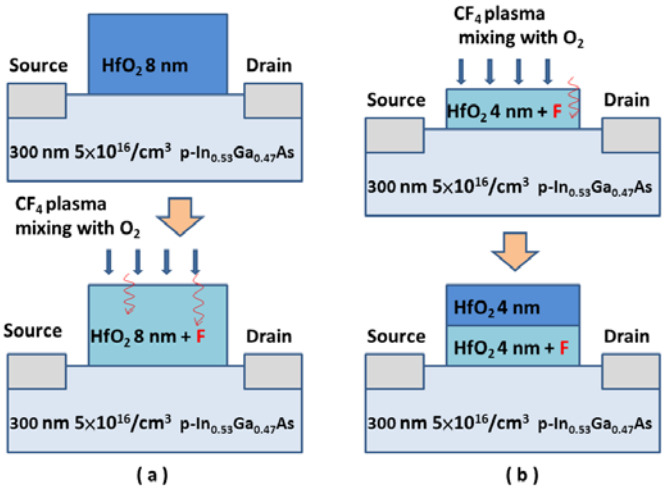


Figure 4.39 Cross section of the HfO_2 / $In_{0.53}Ga_{0.47}As$ MOSFETs for (a) the post-F sample, and (b) the mid-F sample.

Figure 4.40 shows the F concentration depth profile, which was examined by SIMS, in the control sample, the post-F sample, and the mid-F sample. For both post-F

and mid-F samples, it is apparent that F atoms have accumulated mainly at the interface between the HfO_2 and the InGaAs substrate. Remarkably, F concentration in the mid-F sample is higher than that in the post-F sample, indicating that the former one is more efficient to incorporate F and passivate the fixed charge, which has been known that the fixed charge is located near the interface. We speculate that F atoms diffuse out of the HfO_2 during PDA process and the upper half of the HfO_2 in the mid-F sample serves as a blocking layer to prevent F atoms from diffusing, whereas in the post-F sample, part of F atoms in the HfO_2 diffuse out. CV curves of metal-oxide-semiconductor capacitors for the control sample, post-F, and mid-F are shown in Fig. 4.41. A shift toward right is obtained in CV curves for the post-F and the mid-F samples compared to the control sample, indicating a smaller amount of positive fixed charge in the post-F and the mid-F samples. The observed V_{FB} shift (ΔV_{FB}) can be converted to fixed charge density reduction (ΔQ_F) using the relation, $\Delta Q_F = -C_{OX} \Delta V_{FB} /q$. We observe a ΔQ_F of 5×10^{11} cm^{-2} in the post-F sample and a significant reduction in the mid-F (1.3×10^{12} cm^{-2}), suggesting a better oxide quality in the post-F sample and the mid-F sample due to the passivation of fixed charge by forming Hf-F bond in the HfO_2 bulk.

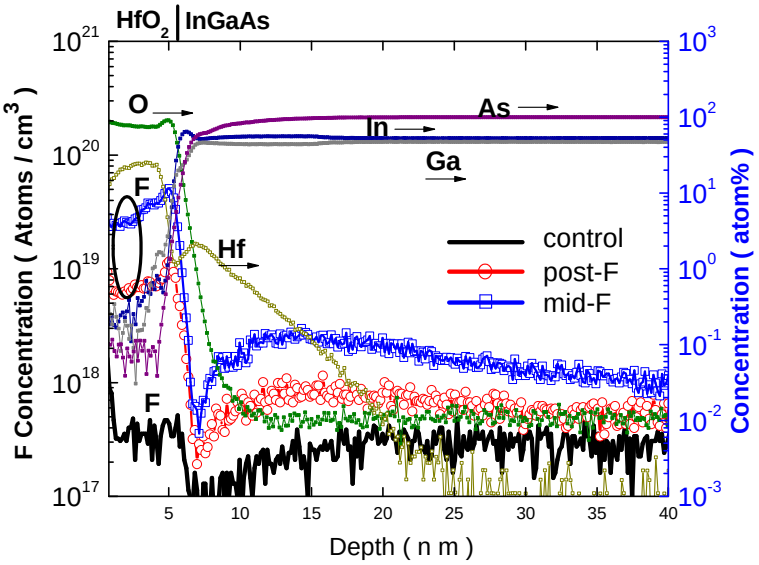


Figure 4.40 SIMS depth profile of HfO_2 / $In_{0.53}Ga_{0.47}As$ structure. Straight line: the control sample, circle: the post-F sample, and square: the mid-F sample.

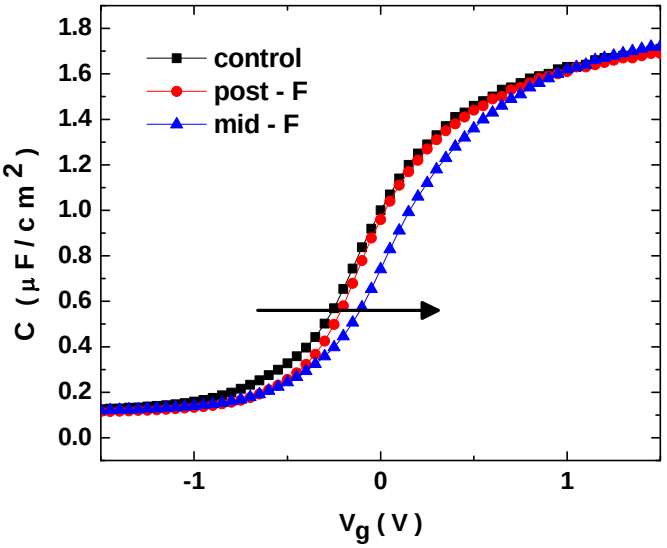


Figure 4.41 CV curves for the MOSCAPs of TaN / HfO_2 / $In_{0.53}Ga_{0.47}As$ structure. Filled square: the control sample, circle: the post-F sample, triangle: the mid-F sample.

Figure 4.42 compares the drain current and G_m versus gate voltage at V_d =50mV, where the W is 600 μm and the L is 5 μm. The maximum G_m of the control sample was

8.2 mS/mm and after CF_4 plasma treatment (post-F), the maximum G_m has been increased by 29.3% to the value of 10.6 mS/mm. Moreover, the maximum G_m achieves 11.9 mS/mm, which is an increase of 45.1%, in the mid-F samples. A similar trend has been observed in the drain current. The inset of Fig. 4.42 is the SS. The SS values for the control sample, post-F, and mid-F sample are 128, 118.9, and 112.3 mV/dec, respectively. A better interface quality for the CF_4 treated samples (post-F and mid-F) is most likely due to the passivation of the interface trap.

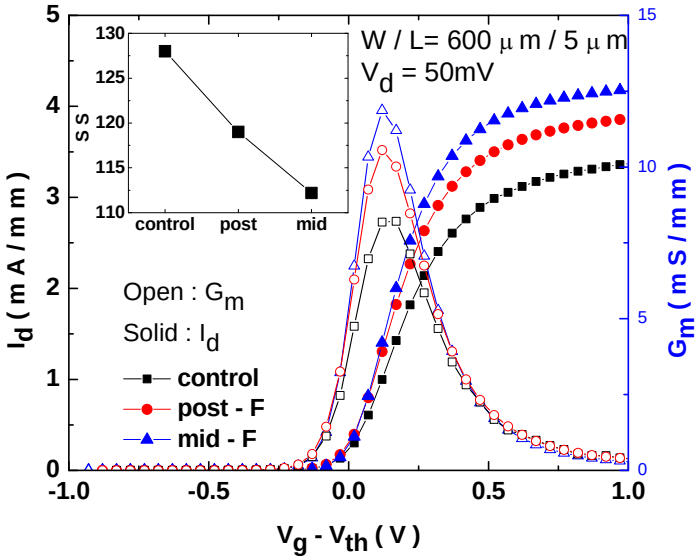


Figure 4.42 I_d and G_m characteristics as a function of various $V_g - V_{th}$ for the control sample, the post-F sample, and the mid-F sample. Inset: the corresponding SS.

The effective channel mobility has been extracted using the split-CV method and plotted in Fig. 4.43. The peak effective channel mobility for the post-F samples (1055 cm^2/Vs) is 27.4% higher than that of the control samples (828 cm^2/Vs). This is believed to be due to the improved interface quality and less fixed charge in the HfO_2/InGaAs gate stack. With even higher F concentration in the HfO_2, passivating more fixed charge and

interface traps, the peak effective channel mobility for the mid-F samples achieves 1305 cm^2/Vs, an increase of 57.6% compared to the control sample. Since the effective channel mobility of the control sample is in good agreement with our previous report [84], we can ascribe the improvement only to the F incorporation step. I_d-V_d characteristics are shown in Fig. 4.44. The maximum drive current under V_d = 2.5 V and V_g-V_{th} =2 V for the control sample, post-F, and mid-F sample are 99, 128, and 145 mA/mm, respectively (W/L = 600 μm/5 μm). The improvement is 29.3% and 46.5% for the post-F and the mid-F samples, respectively, which is attributed to the improved effective channel mobility.

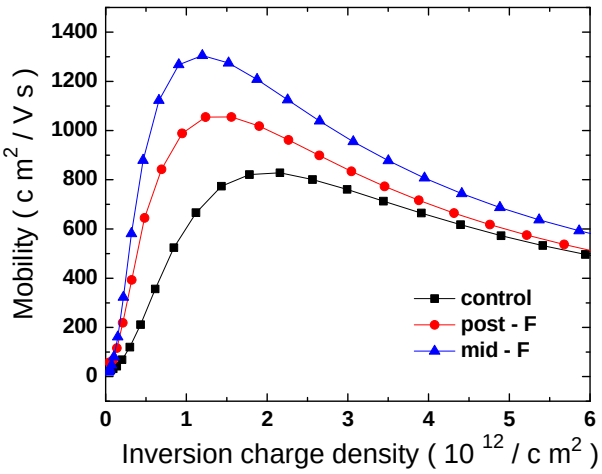


Figure 4.43 Effective channel mobility for the control sample, the post-F sample, and the mid-F sample.

In summary, fluorinated HfO_2 gate dielectric engineering on $In_{0.53}Ga_{0.47}As$ MOSFET has been presented. It has been demonstrated that F incorporation reduces fixed charge and also passivates interface traps, which improves the effective channel mobility of the post-F samples to 1055 cm^2/Vs. By manipulating the F concentration in the HfO_2, we have obtained higher F concentration in the mid-F sample, which has more

F incorporated in the HfO_2 bulk and the HfO_2/ InGaAs interface boosting the effective channel mobility to 1350 cm^2/Vs.

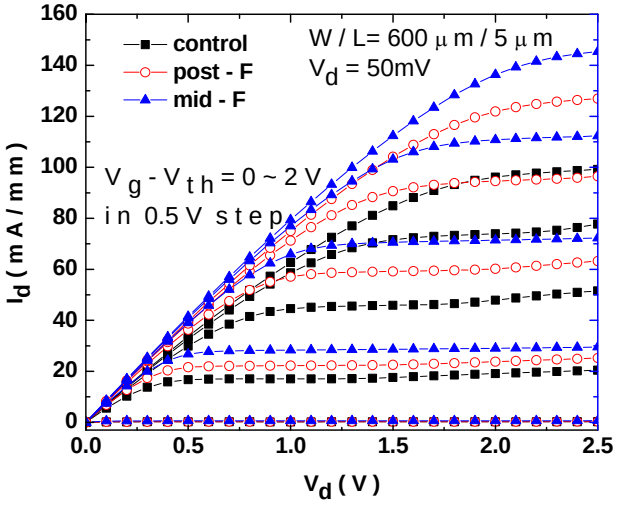


Figure 4.44 I_d-V_d at $V_g = V_{th}$ to V_{th} + 2 V for the control sample, the post-F sample, and the mid-F sample.

4.6 Fluorine incorporation into HfO_2 gate dielectric InP metal-oxide-semiconductor field-effect-transistors

We have demonstrated fluorine incorporation in $Al_2O_3/In_{0.53}Ga_{0.47}As$, Al_2O_3/InP, and $HfO_2/In_{0.53}Ga_{0.47}As$ gate stacks. Significant improvements have been achieved. Finally, we will show improvements of fluorine incorporation in HfO_2/InP gate stack in the following.

This section presents the effects of F incorporation on electrical characteristics of HfO_2/InP gate stack. F had been introduced into HfO_2 gate dielectric by postgate CF_4 plasma treatment, which was confirmed by X-ray photoelectron spectroscopy analysis. Compared to the control sample, fluorinated samples had great improvements in subthreshold swing, hysteresis, and the normalized extrinsic transconductance. These

improvements can be attributed to the reduction of fixed charge in the HfO_2 bulk and less interface trap density at the HfO_2/InP interface.

We fabricated InP MOSFETs on a semi-insulating InP (100) substrate. The native oxides were removed with 1% diluted HF solution, followed by 20% $(NH_4)_2S$ pretreatment. A 10-nm-thick ALD Al_2O_3 was deposited at a substrate temperature of 200° C as an encapsulation layer. For device fabrication, source and drain regions were selectively implanted with a Si dose of $2\times10^{14}/cm^2$ at 35 keV. The source and drain activation annealing was performed in a nitrogen ambient at 750 °C/15 s. The encapsulation layer was removed using buffered oxide etch solution. ALD HfO_2 film was deposited on both substrates after the same surface preparation (HF and $(NH_4)_2S$). Some samples were transferred *ex situ* to a RIE chamber (RIE 80 m., Plasma Technology) and treated by postgate CF_4 plasma with low rf power ranging from 20W to 40W. A mixed flow of CF_4 and O_2 gas (ratio ~ 10:1) was introduced into the chamber with pressure of 100 mTorr. Control samples without CF_4 plasma treatment were also fabricated as references. PDA was then performed for all the samples at 500°C for 60 s in a nitrogen ambient. Subsequently, a 200 nm TaN gate electrode was sputtered and followed by gate patterning. AuGe/Ni/Au alloy was deposited by E-beam evaporation and a liftoff process for the source and drain Ohmic contact, followed by annealing at 400°C for 30 s in a nitrogen ambient. MOSCAPs with the same gate stack were also fabricated for capacitance-voltage analysis.

The cross-sectional transmission electron microscopy (TEM) images of HfO_2/InP gate stack with and without CF_4 treatment are shown in Fig. 4.45 (a) and (b). Both TEM

images show crystalline HfO_2, which is inevitable after the PDA process. Similar thickness of HfO_2 (9.6 nm for the control sample and 9.5nm for the sample with CF_4 treatment) suggests no significant plasma etching during low rf power CF_4 treatment. Fig. 4.45 (c) shows the X-ray photoelectron spectroscopy spectra of F *1s* for the HfO_2/InP gate stack with and without CF_4 treatment. The peak located at ~ 685 eV corresponds to the F bonds in the bulk HfO_2 [76], indicating that F is incorporated into HfO_2 after CF_4 plasma treatment.

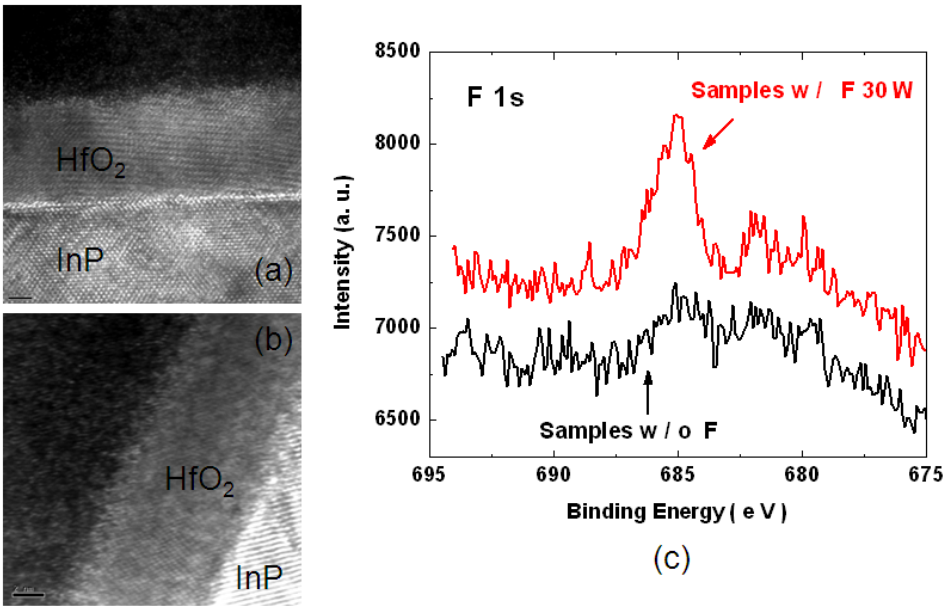


Figure 4.45 Cross-sectional TEM images of TaN / 9.6 nm HfO_2 / InP gate stack (a) without postgate CF_4 plasma treatment and (b) with postgate CF_4 plasma treatment 30W/3min. (c) their corresponding F *1s* XPS spectrums. The scale bars in two TEM images represent 2 nm.

Figure 4.46 compares the SS and hysteresis for the HfO_2/InP gate stack with different CF_4 rf powers for 3 min. SS for the control sample, samples with 20W treatment, and samples with 30W treatment are 120, 106.6, 97.1 mV/dec, respectively. An improved SS suggests less interface trap density, which is believed to be due to the passivation of

interface traps by forming strong fluorine bonds. We reduced hysteresis in the MOSCAPs, which shows a similar trend that hysteresis improves with increasing rf power (rf power ≦ 30W). Hysteresis is caused by the trapping and de-trapping process between the traps in the HfO_2 bulk and the substrate [111]. It reduces with CF_4 plasma treatment, suggesting that the fixed charge in the HfO_2 bulk has been reduced by forming Hf-F bonds. For the samples with rf power of 40W, SS and hysteresis increase compared to the samples with 30W treatment, indicating a possible plasma damage caused by excessive plasma. The inset (a) and (b) of Fig. 4.46 are the hysteresis curves for the control sample and the samples with 30W treatment, respectively.

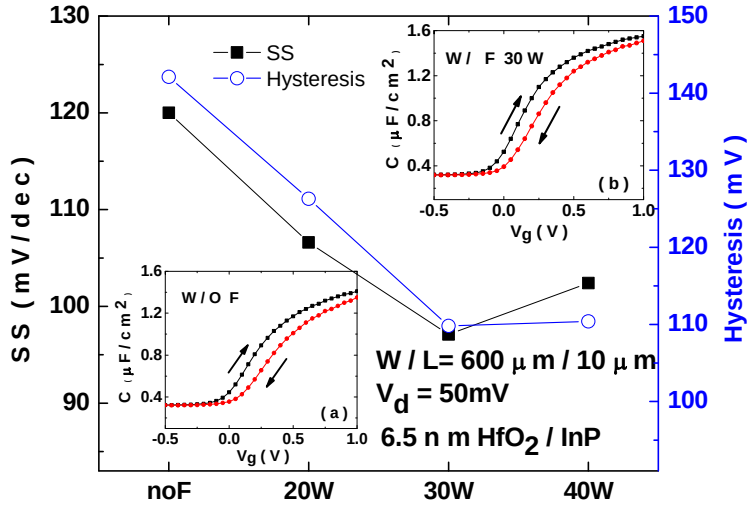


Figure 4.46 SS and hysteresis as a function of CF_4 plasma rf powers. CF_4 plasma treatment time: 3 min. Inset (a): hysteresis curves of the control sample, and inset (b): hysteresis curves of the sample with 30W CF_4 plasma treatment.

Figure 4.47 shows the normalized G_m as a function of different rf powers ranging from 20 to 40W for the HfO_2/InP gate stack, where the W is 600 μm and the L is 10 μm at V_d = 50 mV. The normalized G_m enhances with the increase of rf power, which is

believed to be due to reduced fixed charge in HfO_2 bulk and less interface trap density. It has been found that with rf power of 30W (3 min), HfO_2/InP gate stack has the best performance in SS, hysteresis and G_m.

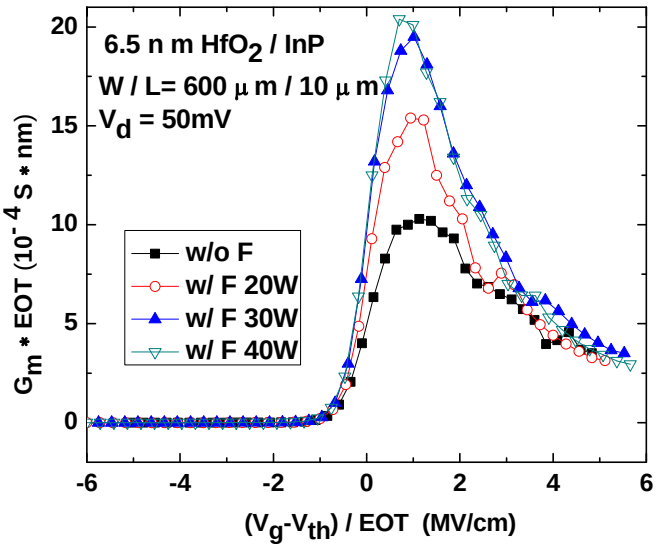


Figure 4.47 The normalized G_m of InP MOSFETs with different rf powers. CF_4 plasma treatment time: 3 min.

In conclusion, the effects of F incorporation on electrical characteristics of HfO_2/InP and gate stack have been investigated. By optimizing the rf power of CF_4 plasma (30W for 3min), great improvements in electrical characteristics have been achieved in SS, hysteresis, the normalized G_m and the normalized drain current. The improvements can be explained by an appropriate amount of F incorporated into the HfO_2 bulk and at the HfO_2/InP interface, passivating fixed charge and interface traps.

4.7 Summary

We have demonstrated improved performances by a post-treatment of CF_4 plasma on various high-k/III-V gate stacks, including $Al_2O_3/In_{0.53}Ga_{0.47}As$, Al_2O_3/InP,

$HfO_2/In_{0.53}Ga_{0.47}As$, and HfO_2/InP gate stacks. We have summarized the effects of F incorporation on these 4 gate stacks in the Table 4.3.

Table 4.3 Comparison of the effects of F incorporation on 4 different gate stacks.

Gate stacks	**Al_2O_3/InP**	**Al_2O_3/InGaAs**	**HfO_2/InGaAs**	**HfO_2/InGaAs**
Optimum F condition	**20W 5min:** Best output characteristics	**20W 3min**	**30W 3min:** Best output characteristics	**30W 3min:** Best output characteristics **30W 5min:** Best interface Quality
Best mobility (cm^2/Vs)	610 → 913	1060 → 1372	NA	821 → 1304
F distribution	Higher F in high k than at the interface	Higher F in high k than at the interface	NA	Pile up at interface
Improve interface	No	Slightly	Yes	Yes
Improve oxide quality	Yes	Yes	Yes	Yes
Plasma damage by excessive plasma	Yes	Yes	Yes	Yes

Chapter 5 SiO_x-based resistive switching random access memories

5.1 Fundamental understanding of SiO_x-based RRAMs

Resistive switching random access memories (RRAMs) have attracted substantial attention for the next generation of nonvolatile memory due to their fast switching speed, high on/off ratio, large integration density, and long retention time [25]. Generally, RRAMs have two types of switching behavior, "unipolar" and "bipolar", which is determined by the gate bias polarity needed when switching devices. Bipolar resistive switching materials, including TiO_2, NiO, MnO_2, TaO_x, HfO_2 [25 - 30], have shown promising resistive switching results. However, bipolar RRAMs could complicate the selector circuit, where one transistor and one RRAM structure is required, while the selector circuit for unipolar RRAMs could potentially simplify the circuit design to one diode and one RRAM and offer better operational control of cross-bar array structures [30]. Therefore, unipolar switching RRAMs are more promising for high-density nonvolatile memory applications. Conventional unipolar RRAMs have set voltage [switching from the high resistance state (HRS) to low resistance state (LRS)] larger than reset voltage (switching from LRS to HRS). Inevitably, conventional unipolar RRAMs require current compliance to protect the device from complete breakdown when switching the device to LRS, which in turn increases circuit area, complexity and cost. Recently, RRAMs from silicon oxide (SiO_x) [31 - 34] have drawn increasing attention. These devices exhibit a *unique* unipolar behavior: SiO_x-based RRAMs have set voltage smaller than reset voltage and thus self-compliant properties. Moreover, SiO_x-based RRAMs have shown promising performance [31], including over 10^4 switching cycles

with high switching speed (50 ns for set process and 80 ns for reset process) and robust nonvolatile properties with extrapolated lifetime beyond 10 years. Therefore, unipolar SiO_x-based RRAMs, which could potentially operate without current compliance, are superior for future applications and scalability. Besides, SiO_x also has high availability and ease of integration with current and future technology platforms. So far, only a few reports have been published on demonstrating SiO_x RRAMs.

5.2 Working principle of SiO_x-based RRAMs

SiO_x-based RRAMs are simply capacitor structures with sidewall exposed. Before operating as resistive switching devices, SiO_x-based RRAMs need to be electroformed by several high voltage sweeps. Figure 5.1 exhibits the typical current-voltage (I-V) curves of the electroforming process of SiO_x-based RRAMs. In the first high voltage sweep, a sudden current increase up to ~ 10^{-7} to 10^{-6} A, was induced at around 20V, accompanied by current fluctuations. We define this voltage as the electroforming voltage (V_{EF}). In the subsequent voltage sweeps (curves 2-4), the current fluctuations gradually increased with current jumps at smaller voltage values. When the current increased sharply at 3 to 4 V, followed by a current plateau, the electroforming process was finished.

Then, this device can operate as a resistive switching device. Typical current-voltage (I-V) curves for LRS/HRS switching are illustrated in the Fig. 5.2. The device is switched to the LRS by applying a linear voltage sweep to ~4V, where the set window is around 3~5 V. With the voltage stop in the set window, the device becomes LRS. On the

other hand, the device is switched to the HRS by applying a sweep to > ~6V, where the reset window is around 5~10 V. With the voltage stop in the reset window, the device becomes HRS. Interestingly, the low-voltage (~1V) conductance in the HRS is a strong function of the previous applied voltage, with conductance decreasing for higher applied voltage. Therefore, SiO_x-based RRAMs has high potential for multilevel storage application (will discuss later).

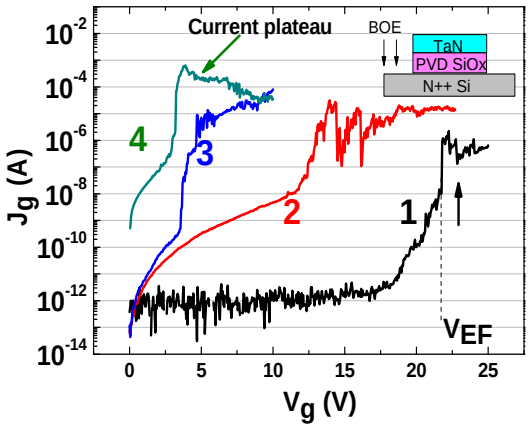


Figure 5.1 The electroforming process in a SiO_x memristor. Inset: device structure.

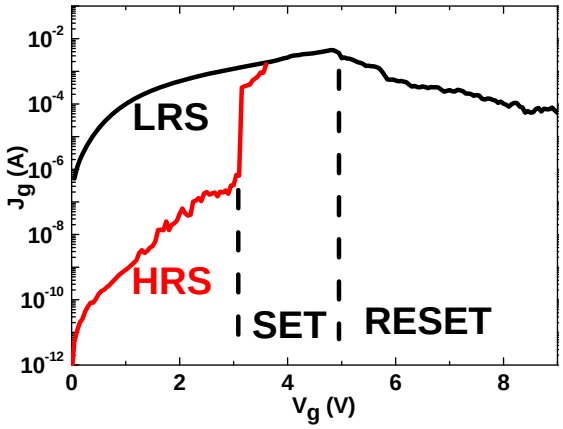


Figure 5.2 Typical LRS and HRS switching I-V curves for SiO_2-based RRAMs.

5.3 Size effects and oxide thickness effects on SiO_x-based RRAMs

In order to understand the conducting filament (CF) properties, we have fabricated SiO_x-based RRAMs with different pattern. Figure 5.3 shows the first voltage sweep (electroforming) of different size devices. Devices with large perimeter can reduce V_{EF}. Comb-shaped devices have small V_{EF} ~ 5V. This suggests that these devices possibly contain more electrically weak spots in the edge, which need less energy (voltage) to form a silicon filament. Once the devices are electroformed, the set and reset voltages are independent of device shape, area and perimeter, as summarized in Table 5.1.

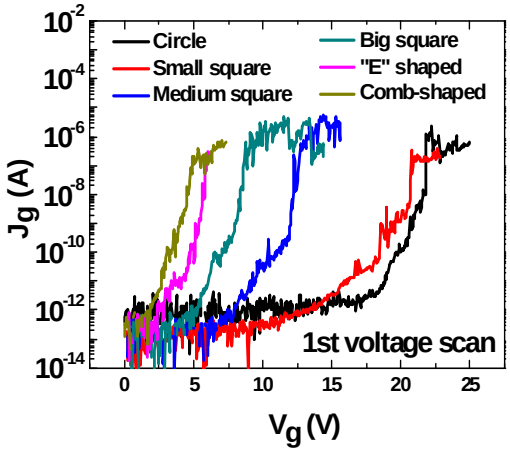


Figure 5.3 First voltage sweep of different size devices.

Table 5.1 Summary of devices with different shapes, area and perimeter

	Area (cm^2)	Perimeter (um)	P/A ratio	$V_{EF,MAX}$ (V)	V_{SET}/V_{RESET} (V)
Circle	1.54×10^{-4}	439.6	35	21.7	3.3 / 5
Small square	3.61×10^{-4}	760	47.5	21	3.1 /4.2
Medium square	1.94×10^{-3}	1760	110	12	3.2 / 5
Big square	2.9×10^{-3}	2160	134	8.5	2.7 / 4
“E” shaped (⋿)	2.24×10^{-3}	4261	53	5.7	2.8 / 4.7
Comb shaped (⩍)	4.43×10^{-3}	8220	54	5	3.1 / 4.5

We also examined the CF behavior in different CETs (CETs: 6, 15, 32, and 53 nm, extracted by capacitance-voltage measurements, data not shown) SiO_x RRAMs in a conventional capacitor structure, as shown in Fig. 5.4 (a). Instead of several voltage sweeps (Fig. 5.1), we applied a forward, followed by a backward voltage sweep, as shown in Fig. 5.4(b). Once the forward voltage sweep reaches a certain value (V_{EF}), which induces a current vibration at the 10^{-7} to 10^{-6}A level, the followed backward voltage sweep would conduct high current and turn the device into on state, i.e., the CF were formed successfully. The next sweep would start from on state (curve 2 in Fig. 5.4(b)). Then the device can operate with LRS/HRS switching as a RRAM. The mechanism behind this could be due to the competition of the field and the thermal effect. [5] Here, we also defined electroforming current (I_{EF}) and the Set Current (I_{Set}) /Reset Current (I_{Reset}) and V_{Set}/V_{Reset} are the values of the current and voltage detected at the beginning of the resistance switching from a HRS/LRS state to a LRS/HRS state, respectively, as shown in Fig. 5.4(b).

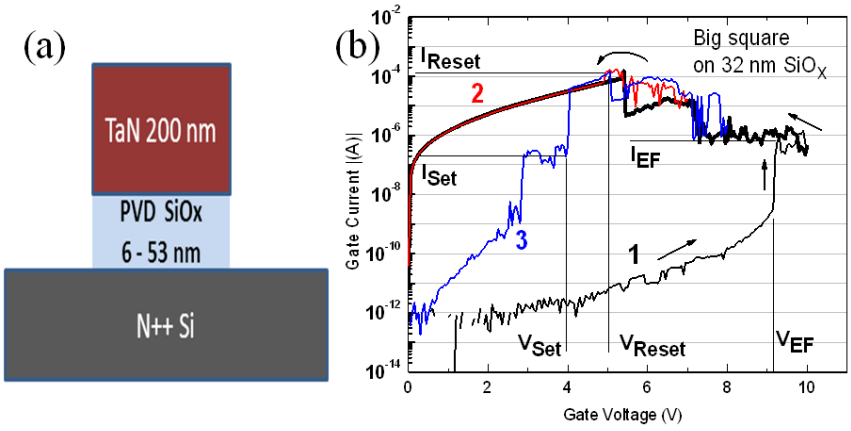


Figure 5.4 (a) Schematic cross section structure of capacitor type SiO_x RRAMs. (b) Electroforming process of forward and backward sweeps. The numbers aside the I-V curves indicate the corresponding sweep order.

Figure 5.5 shows the V_{EF} change as a function of SiO_x CETs and device sizes (540 µm×540 µm, big square; 340 µm×340 µm, medium square; 140 µm×140 µm, small square). Reduced V_{EF} for the devices with large area (perimeter) suggests that large perimeter devices possibly contain more weak spots at the edge, which need less energy (voltage) to form a CF. This trend is consistent with our results mentioned in the earlier this section. V_{EF} also decreases with thinner CET, which could be due to enhanced effective electric field in the SiO_x with thinner CET. Small square devices of 6 nm-SiO_x RRAMs could not be electroformed due to hard breakdown, as shown in Fig. 5.6. The device hard breakdown voltages were from 6 V to 8.6 V. No current vibration at 10^{-7} to 10^{-6}A level was observed for all the devices under test. We extrapolated the V_{EF} required for the small square devices of 6 nm-SiO_x RRAMs to be 10V (Fig. 5.5), which is larger than the hard breakdown voltage. This suggests a competition between CF formation and oxide hard breakdown during the electroforming process (the inset of Fig. 5.6).

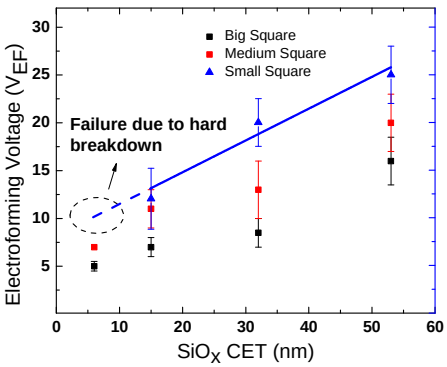


Figure 5.5

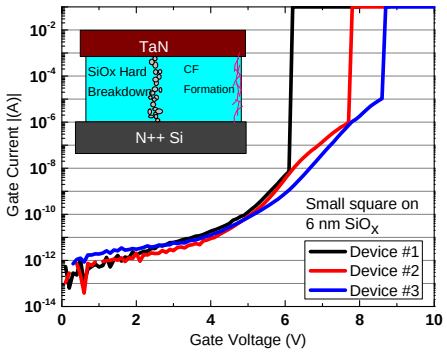


Figure 5.6

Figure 5.5 V_{EF} as a function of SiO_x CETs and device sizes. (Big square: 540×540 µm; medium square: 340 ×340 µm; small square: 140 ×140 µm).

Figure 5.6 Hard breakdown of small square devices on 6 nm SiO_x. Inset: illustration of competing effect of hard breakdown and CF formation.

In the electroforming process, electroforming power (P_{EF}, product of V_{EF} and I_{EF}) is a key factor, which indicates the energy (power) needed to initiate the CF formation. Figure 5.7 shows the P_{EF} change with the change of SiO_x CETs and device sizes. Thick CET RRAMs need higher P_{EF} than thin CET RRAMs to maintain sufficient power density (in a vertical direction) to initiate CF formation. Big size RRAMs consume more power than small size RRAM, which could be due to larger leakage in big size devices. Figure 5.8 and Fig. 5.9 show the LRS/HRS current and set/reset power as a function of SiO_x CETs and device size, respectively. On current and set/reset power were independent with CETs and device sizes. This implies that there might be only a single CF responsible for the LRS/HRS switching. Once the single CF is formed, the rupture/recovery is localized to the weak spot of the CF. Therefore, the power needed to break/restore the CF is independent of SiO_x CETs and device size.

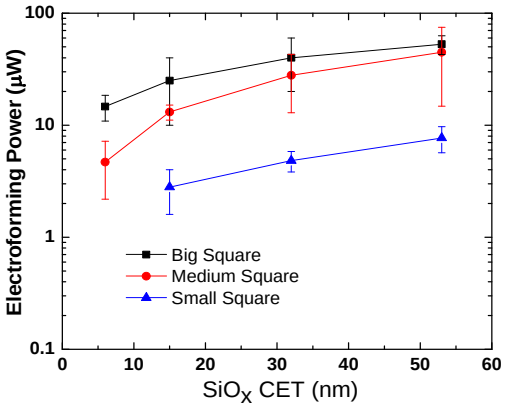


Figure 5.7 P_{EF} as a function of SiO_x CETs and device sizes.

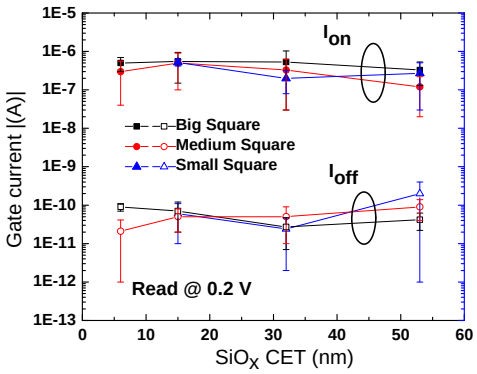


Figure 5.8

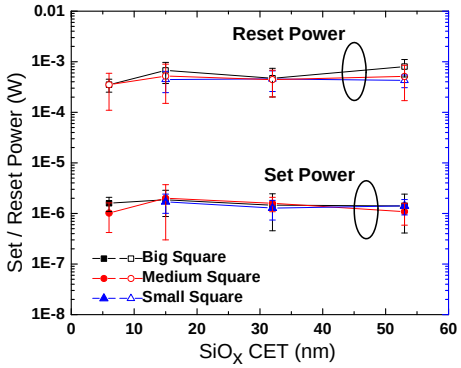


Figure 5.9

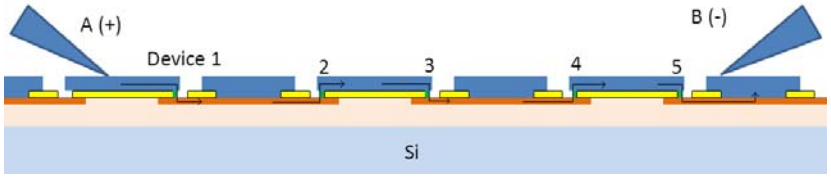
Figure 5.8 On/off current as a function of SiO_x CETs and device

Figure 5.9 Set/reset power as a function of SiO_x CETs and device sizes.

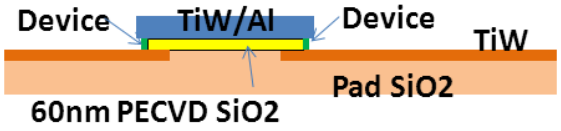
Based on our measurements, we found that there might be only one CF responsible for LRS/HRS switching. Moreover, devices with larger perimeter can be formed easier compared to ones with small perimeter. This is because that a large perimeter device possibly contains more weak spots at the edge, which need less energy (voltage) to form a CF. We will discuss more detail about the CF rupture/recovery behavior in the next section. A possible model will also be purposed.

5.4 Conducting filaments formation and switching mechanism of SiO_x-based RRAMs

This section will discuss a model of CF formation and switching for SiO_x-based RRAMs. We examine the SiO_x-based RRAMs switching behavior in two designed testing structures.



(a)



(b)

Figure 5.10 (a) Device structure of 5 RRAMs in series. The arrows indicate current path from pad A to pad B. (b) Device cross-sectional structure.

Figures 5.10 (a) and (b) show the cross-sectional structure and equivalent circuit of 5 RRAMs in series, so-called 5-RRAMs-chain, where pad A was positively biased with pad B grounded (denoted as A-B mode). Arrows in Fig. 5.10 (a) indicate the current path in A-B mode. Figure 5.11 shows the electroforming process in A-B mode (to form devices 1 to 5 simultaneously). In the first high voltage sweep, a sudden current increase (up to ~ 10^{-7} to 10^{-6} A, indicating soft breakdown) was induced at a V_{EF} accompanied by current fluctuations. In the subsequent voltage sweeps (curves 2-4), the currents and fluctuations gradually increased with current jumps at smaller voltage value. When the current jumped at 3 to 4V, followed by a current plateau, the electroforming process was finished. After each electroforming sweep, the current states of devices 1 to 5 were monitored individually by low voltage scan (0 to +1V, as shown in Fig. 5.12). After the first and second sweep, no clear trend of current states change was observed, suggesting that the electroforming process is possibly a random process, depending on the location of weak spots. After the third and fourth sweeps, the devices near two electrodes (devices

1, 2, 4, and 5) electroformed (reached high current) earlier than the device in the center (device 3), suggesting that the CF possibly grew from two electrodes toward the center. We further investigated the switching behavior of these 5 RRAMs. After on/off switching in A-B mode 20 times to make the CF more stable, we switched in A-B mode to on state and off state and monitored the current status change of devices 1 to 5, as shown in Fig. 5.13. It shows that the switching happens in only one device (device 5, Fig. 5.14), while the others remained highly conductive, suggesting that the CF rupture does not occur over the whole CF, but only in a localized region instead.

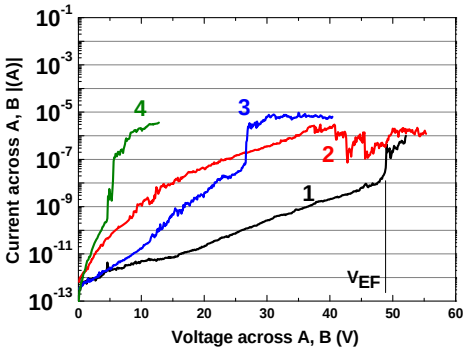


Figure 5.11

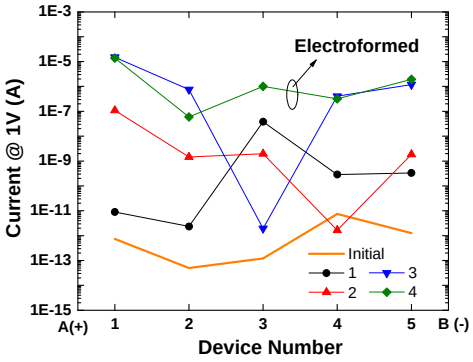


Figure 5.12

Figure 5.11 Electroforming sequence of 5 RRAMs in A-B mode. The numbers aside the I-V curves indicate the corresponding sweep order.

Figure 5.12 Current status of deices 1 to 5 corresponding to each electroforming sweeps in Fig. 5.11.

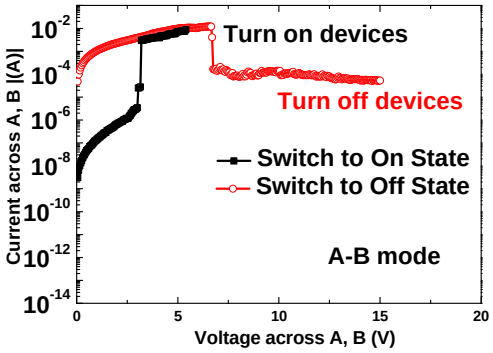


Figure 5.13

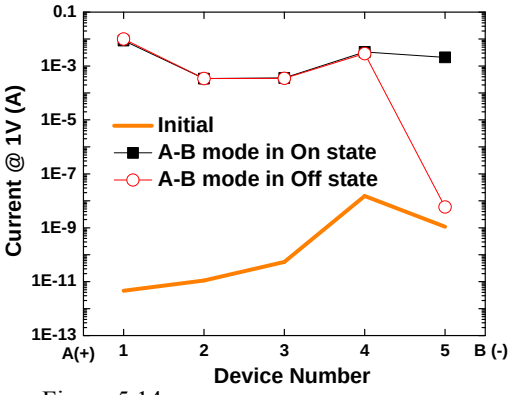


Figure 5.14

Figure 5.13 I-V curves of the on/off switching process for the A-B mode.

Figure 5.14 Current status of deices 1 to 5 corresponding to each on/off switching in Fig. 5.13.

Even though in this case the rupture/recovery happened near the cathode end, this is not always true. We will present the other testing structure in the following, which shows that rupture/recovery is a random process and would occur either near cathode end or anode end.

In the second testing structure, adjacent devices were selected to examine the electroforming and switching characteristics of the serially connected structure. SiO_x RRAMs were fabricated on heavily doped n-type (n^{++}) (100) Si (1-7×10^{19} cm^{-3}) with resistivity of 0.001-0.005 ohm-cm. The surface native oxides were removed by 1% dilute HF solution. Then, 30 nm-thick SiO_x was deposited by magnetron sputtering of silicon target in Ar and O_2 ambient at 200°C, followed by PDA at 500°C for 5min in O_2 ambient. Then, TaN was sputtered and patterned for gate electrode. CF_4-based chemistry was used

to etch TaN; and CHF_3 mixed with O_2 was used to etch SiO_x. The device cross-sectional view is shown in Fig. 5.15. Electrical characteristics were measured by an Agilent B1500A semiconductor device analyzer in vacuum ambient ($<1\times10^{-3}$ mbar).

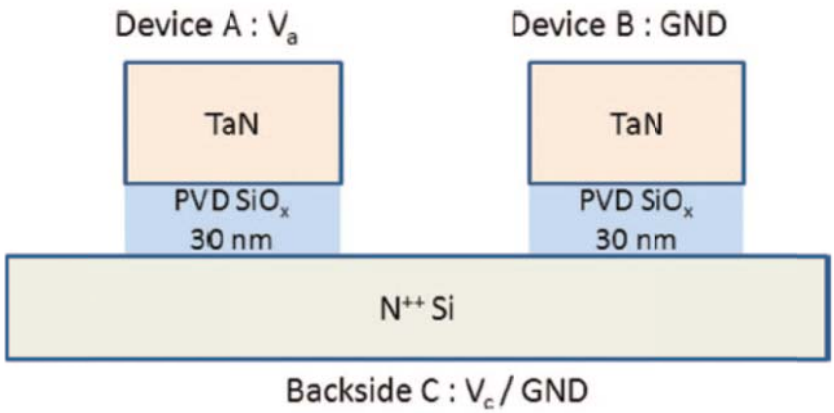


Figure 5.15 Schematic cross-section structure of SiO_x RRAMS. For A-B mode, voltage is applied on device A with device B grounded and voltage of backside (V_c) is monitored simultaneously. For A(B)-C mode, voltage is applied on device A(B) with backside C grounded.

To investigate the mechanism of CF growth during the electroforming process, the second testing structure was configured as follows. Adjacent devices A and B were selected to examine the electroforming and switching characteristics of the serially connected structure (TaN/SiO_x/n^{++}Si/SiO_x/TaN) where device A was positively biased with device B grounded and the backside voltage (V_c) monitored (denoted as, A-B mode) [26]. During electroforming sweeps, electrons were injected from device B and transported to device A through the SiO_x beneath device B, n^{++} Si substrate, and the SiO_x beneath device A. V_c was monitored simultaneously to understand the voltage change across both devices A and B, where V_c could be considered the effective applied voltage on device B and the difference between V_a and V_c could be regarded as the effective applied voltage on device A. After each electroforming sweep, the conductive state of

device A and device B were detected individually by positively biasing device A/device B with the backside grounded (denoted as, A-C mode and B-C mode, respectively).

Figure 5.16 exhibits the current-voltage (I-V) curves of the electroforming process for A-B mode and the inset of Fig. 5.16 shows the corresponding V_c change during each electroforming sweep. In the first high voltage sweep, a sudden current increase up to ~ 10^{-7} to 10^{-6} A, was induced at around 20 V, accompanied by current fluctuations. In the subsequent voltage sweeps (curves 2-5), the current fluctuations gradually increased with current jumps at smaller voltage values. When the current increased sharply at 3 to 4 V, followed by a current plateau, the electroforming process was finished. Note that the V_c to V_a ratio was around 0.45 with V_a < 16V, it increased to 0.56 at V_a = 19.8 V, and dropped abruptly to 0.06 at V_a = 20.4 V. This implies that, in the beginning, the applied voltage on the top electrode of device A was divided roughly equally across devices A and B, and then, just before V_c decreased sharply, more than half of the applied voltage was dropped across device B. The subsequent V_c drop to near 0V suggests that the CF in device B was formed first.

After each electroforming sweep, the conductivities of devices A and B were monitored individually in A-C and B-C mode, as shown in Fig. 5.17. The pristine states of devices A and B were in the HRS. After the first electroforming sweep in A-B mode, device A was still in the HRS while device B moved to the LRS. This is consistent with our observation in the inset of Fig. 5.16 that the CF in device B was formed right after the first electroforming sweep. In the following electroforming sweeps, device B was still in the LRS and device A was moving toward the LRS.

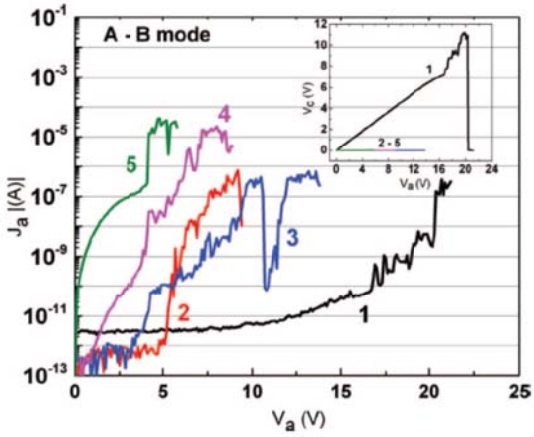


Figure 5.16 Electroforming process for A-B mode. The numbers aside the I-V curves indicate the corresponding sweep order. After 5th scan, devices are ready for RRAM operation. The inset shows the corresponding backside voltage (V_c) change during the electroforming process

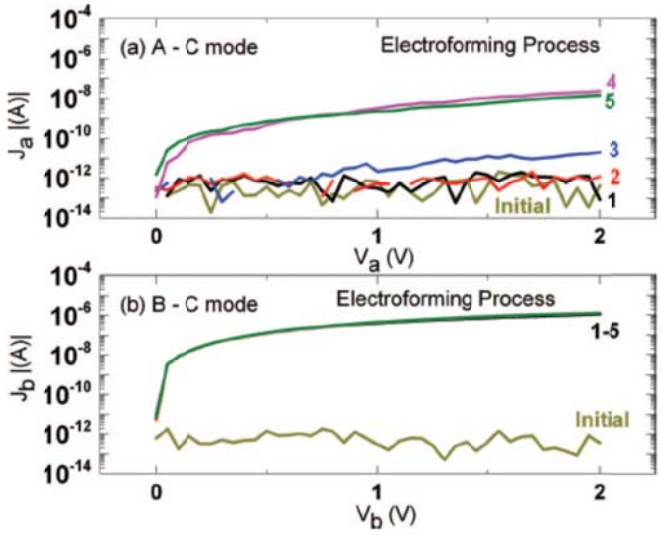


Figure 5.17 I-V curves for (a) A-C mode and (b) B-C mode. They are corresponding to each electroforming sweep in Fig. 5.16.

As shown in Fig. 5.18, we further switched A-B mode with positive or negative bias to LRS/HRS to observe the conductivity changes in A-C and B-C modes (Figs. 5.19 (a) – (d)). We found that rupture/recovery was independent of the A-B mode bias polarity

and it occurred in only A-C mode, whereas B-C mode remained unaffected. Note that during the A-B mode switching, V_c stayed close to 0 V, as shown in the insets of Figs. 5.18 (a) and (b), suggesting that the CF was intact in B-C mode.

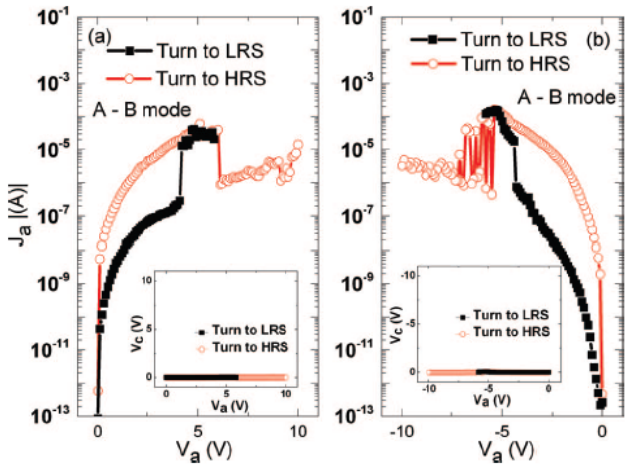


Figure 5.18 I-V curves of the LRS/HRS switching process for the A-B mode. (a): It turns the A-B mode to LRS/HRS when sweeping to 5.5 V/10 V. The corresponding A-C and B-C mode are shown in Fig. 5.19 (a) and (b), respectively. (b): It turns the A-B mode to LRS/HRS when sweeping to -5.5 V/-10 V. The corresponding A-C and B-C mode are shown in Fig. 5.19 (c) and (d), respectively.

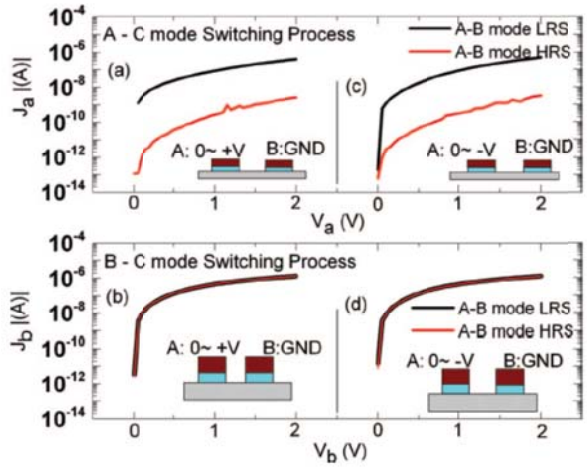


Figure 5.19 I-V curves for A-C mode and B-C mode. (a) and (b): corresponding to A-B mode switching in Fig. 5.18 (a). (c) and (d) : corresponding to A-B mode switching in Fig. 5.18 (b). The insets in (a) to (d) show the measurement setup.

We point out here that another CF growth and switching behavior has also been observed. Figure 5.20 shows the electroforming process of another device A and B set (fresh devices). Likewise, in the first high voltage sweep, there was a sudden current increase up to ~ 10^{-7} to 10^{-6} A at around 12 V, accompanied by current fluctuations. In contrast with the inset of Fig. 5.16, V_c did not drop but remained at high voltage (0.51~0.65×V_a), as shown in the inset of Fig. 5.20. Note that the initial soft breakdown fields in Figs. 5.16 and 5.20 are smaller than 10 MV/cm, which is a typical value for thermally grown SiO_2 [113]. In order to lower the electroforming voltage, sputtered SiO_x was employed due to its high as-deposited defect density, leading to an electrically weak material which essentially accelerates the electroforming process. In addition, etching the SiO_2 material forms an exposed sidewall that is expected to promote soft breakdown at low voltage [31 - 34]. With the subsequent voltage sweeps (curves 2-4), A-B mode could still be electroformed while V_c remained high.

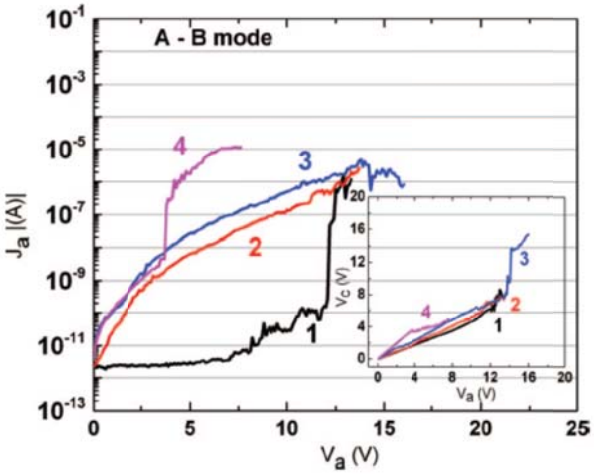


Figure 5.20 Electroforming process for A-B mode. The numbers aside the I-V curves indicate the corresponding sweep order. After 4th scan, devices are ready for RRAMs operation. The inset shows the corresponding backside voltage (V_c) change in the electroforming process

Similarly, after each electroforming sweep, the conductivities of devices A and B were monitored individually in A-C and B-C mode, as shown in Fig. 5.21. The initial states of devices A and B were in the HRS. After the first electroforming sweep in A-B mode, both the currents in A-C and B-C mode increased. In the following electroforming sweeps, the currents in A-C and B-C mode gradually increased toward the LRS. Note that in the third electroforming sweep in A-B mode, the V_c to V_a ratio jumped up to 0.97 at V_a = 14.2 V (the inset of Fig. 5.20), suggesting that the CF in device A was formed, therefore, V_c was pulled up close to V_a. At this point, the A-C mode became LRS (Fig. 5.21 (a)).

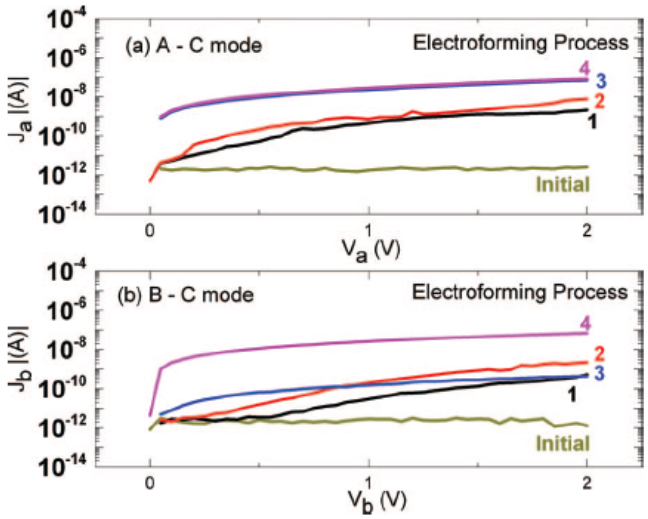


Figure 5.21 I-V curves for (a) A-C mode and (b) B-C mode. They are corresponding to each electroforming sweeps in Fig. 5.20.

We further switched A-B mode with positive or negative bias to LRS/HRS, as shown in Fig. 5.22. Their corresponding current changes in A-C and B-C mode were shown in Figs. 5.23 (a) – (d). We found that the rupture/recovery mechanism was independent of the A-B mode bias polarity, and occurred in only B-C mode whereas A-C mode remained unaffected. Note that during the A-B mode switching, V_c stayed high (>

0.5 V_a, the insets of Figs. 5.22 (a) and (b)), suggesting that the CF was undamaged in A-C mode.

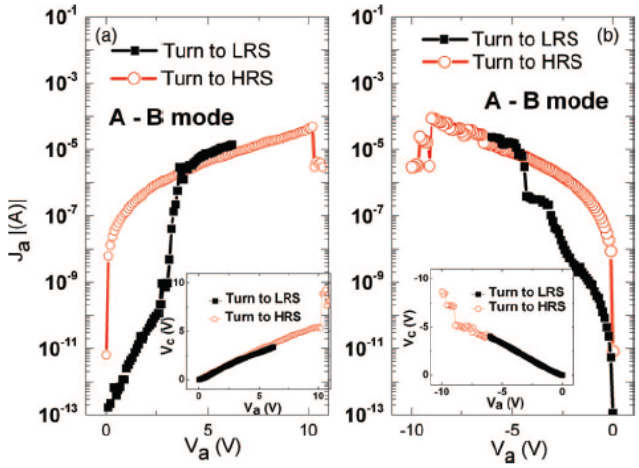


Figure 5.22 I-V curves of the LRS/HRS switching process for the A-B mode. (a): It turns the A-B mode to LRS/HRS when sweeping to 5.5 V/10 V. The corresponding A-C and B-C mode are shown in Fig. 5.23 (a) and (b), respectively. (b): It turns the A-B mode to LRS/HRS when sweeping to -5.5 V/-10 V. The corresponding A-C and B-C mode are shown in Fig. 5.23 (c) and (d), respectively

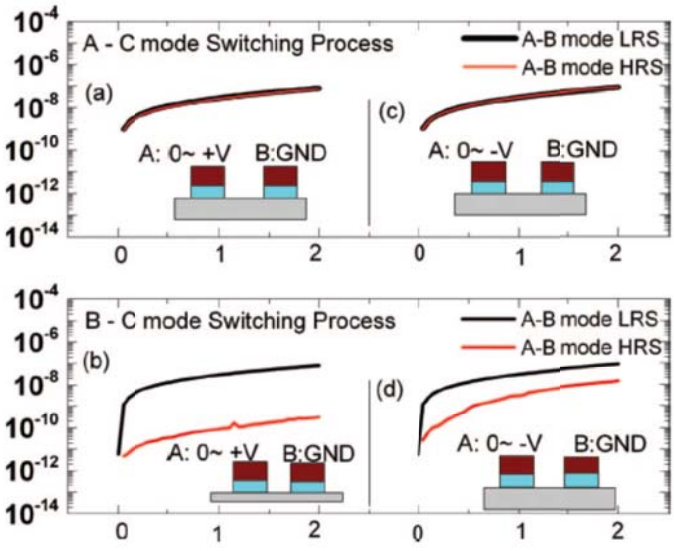


Figure 5.23 I-V curves for A-C and B-C mode. (a) and (b) : corresponding to A-B mode switching in Fig. 5.22 (a). (c) and (d) : corresponding to A-B mode switching in Fig. 5.22 (b). The insets in (a) to (d) show the measurement setup.

We propose a possible mechanism for the random electroforming process in SiO_x RRAMs: Hole or electron injection and charge trapping by oxygen vacancy defects have been linked to bias temperature instability in MOSFETs [114], and models involving hydrogen release and capture by oxygen vacancy defects to form electrically-active hydrogen bridge defects have been correlated with soft breakdown and time-dependent dielectric breakdown in thin oxides [115]. It is reasonable to expect that similar defects may be present in the SiO_x RRAM device, where current-induced localized heating and injected charge carriers help to break either Si-O bonds or Si-H bonds. Then, the Si atoms can cluster to form Si nanocrystals, providing a CF between electrodes. The Si nanocrystals are presumed to be randomly distributed, with spacing determined by the pre-existing defect distribution within the SiO_2, as shown in Fig. 5.24 (a). The band diagram of a localized Si nanocrystral/ SiO_2 /Si nanocrystal is shown in Fig. 5.24 (b). Injected holes or electrons could react with a variety of oxygen vacancy defect types to generate metastable charge states leading to structural and electrical changes in the oxide. For example, when charged to a positive, metastable state, the hydrogen bridge can convert to a pair of isolated defects: an inactive Si-H; and a positively-charged, 3-fold coordinated oxygen atom that can diffuse or drift away [116]. This is one possible pathway that could initiate the oxygen reduction reaction ($SiO_2 \rightarrow SiO_x$) and lead to nanocrystal Si growth during the electroforming process.

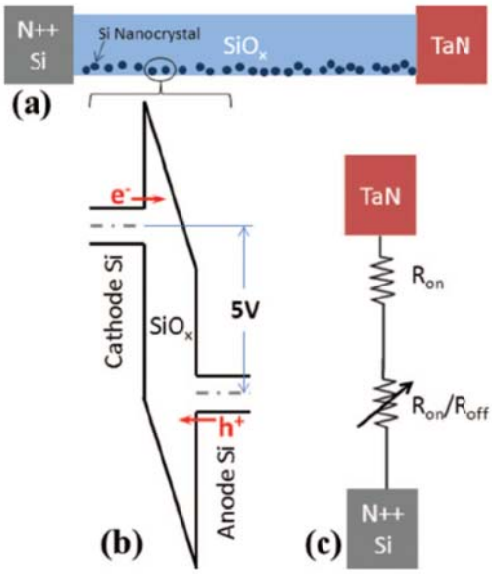


Figure 5.24 Illustration of Si random nanocrystal formation. (b) Band diagram of Si nanocrystal/SiO_2/ Si nanocrystal region subjected to 5V bias. (c) Localized switching model for SiO_x RRAMs.

Based on our observations, in the first electroforming high voltage sweep, the applied voltage equally stresses the SiO_x film below both electrodes, suggesting that the CF grows from both electrodes (Fig. 5.21 (a) and (b)) and the CF could be formed quickly in one electrode (Fig. 5.17 (a) and (b)), which could be due to more pre-existing defects near that electrode. In terms of switching behavior, the rupture/recovery could occur anywhere along the CF, depending on a random process that determines the location of the weak spot. We noticed that the magnitudes of the currents in A-C and B-C mode were slightly different, and the switching always happened in the device with smaller current. The smaller current suggests a narrower CF with higher resistance in that area, which is more favorable to rupture/recovery as compared to a thick CF in other regions. Once the weak spot is established in the CF, the rupture/recovery is localized in that weak spot (other regions of CF are still intact) and applied voltage polarity does not change the location of the weak spot. Moreover, monitoring V_c simultaneously measures

how the CFs change in both devices. As a result, it is concluded that there is only one localized spot responsible for resistance switching, where the resistance can be switched to on state resistance (R_{on}) or off state resistance (R_{off}) by the applied voltage. Other parts of the CF are not affected by the applied voltage and their resistance remains R_{on}, leading to the localized switching model that has been developed for SiO_x RRAMs based on these results, as shown in Fig. 5.24 (c).

These results indicate that SiO_x RRAM forming and switching may arise from different mechanisms than those present in other resistance switching materials, such as TiO_2 or NiO. A similar experiment with n-type TiO_2 showed that rupture and recovery of the CF occurs mostly near the anode side [113, 117], while in p-type NiO, the switching happens near the cathode side [114, 117]. A previous report also showed improved reset switching time by stacking TiO_2 and NiO films [118]. Interestingly, SiO_x RRAMs are different from those materials and have no preference of switching location (which can be either near the anode or cathode side). This is thought to be due to a different CF formation mechanism. In TiO_2 and NiO films, CF formation relies on migration of oxygen vacancies, oxygen ions or metal interstitials to form a conical-shaped CF, where the CF at one side has a smaller diameter than that at the other side [117]. This results in a preference of rupture location in TiO_2 and NiO RRAMs. However, if the CF in SiO_x RRAMs is formed by silicon nanocrystals [119], it might not be conical-shaped. The weak spot could be randomly formed and it might occur in any location along the CF. In addition, because the SiO_x RRAM architecture introduces a sidewall exposed to ambient air, the potential effects of water, and water decomposition products, within the SiO_x layer near the sidewall must be taken into account. Since water decomposition in

amorphous SiO_2 and at oxygen vacancies lead to formation of Si-OH and Si-H groups [119], ionic species such as OH^-, H^+ and H_3O^+ could lead to different CF formation and switching mechanisms in SiO_x RRAM as compared to other materials.

5.5 Reduced electroforming voltage and enhanced programming stability in resistive switching of SiO_2 thin films

This section will discuss the effects of incorporating a thin silicon layer into a SiO_2-based resistive switching random access memory. The thin silicon layer was deposited onto the sidewall region of the device by physical vapor deposition. It is found that this thin silicon layer effectively reduces the electroforming voltage and stabilizes device current in both low- and high-resistance programmed states. It is concluded that the improved performance is due to formation of a more robust, more uniform conducting filament. As a result of this advantage, stable tri-state programming can be achieved in the SiO_2-based resistive memory device.

Recently, RRAMs have shown capability of multilevel storage in many oxides [120 - 121]. By setting a current compliance during the set process, or by controlling the amplitude of the voltage pulse during the reset process, multilevel LRS and HRS have been achieved [32, 113, 122]. J. Yao *et al*. have demonstrated multilevel switching in SiO_2-based RRAMs by changing the magnitude of the reset voltage sweep [32, 122]. However, SiO_2-based RRAMs show large variation in HRS current (under the same voltage sweep conditions) [122], which can lead to instability in the programmed HRS current and bit error rates when operated as a multilevel memory.

This section will present a fabrication method that results in lower V_{EF} and improved HRS current stability of SiO_2-based RRAMs. By sputtering a thin silicon layer on the exposed sidewall of SiO_2-based RRAMs (SSiRRAMs), a discontinuous layer comprised of silicon nanocrystals is introduced between the top and bottom electrodes. Given that the electroforming process produces silicon nanocrystals embedded into the amorphous SiO_2 layer, which potentially support formation of CFs [119], depositing a thin, discontinuous silicon layer onto the SiO_2 sidewall provides an external silicon source that may accelerate CF formation. Compared to the random formation of CFs in a high-voltage electroformation process [123], the addition of a sputtered-silicon thin layer on the sidewall may promote formation of more robust or more uniform CFs, which acts to effectively reduce HRS current instability. With the improved HRS current stability in our SSiRRAMs, tri-state programming has been realized.

SiO_2 RRAMs were fabricated on a heavily doped n-type (n^{++}) (100) Si ($1\text{-}7\times10^{19}$ cm^{-3}) 4-inch wafer with resistivity of 0.001-0.005 ohm-cm. The surface native oxide was removed by 1% dilute HF solution. A 30 nm-thick thermal SiO_2 layer was then grown in a high temperature furnace at 950°C. Then, TaN was sputtered and patterned for gate electrode. The SiO_2 was etched in 6:1 buffered oxide etch to expose a sidewall between the top TaN electrode and bottom n+ Si electrode. The wafer was cleaved into several pieces and transferred to a high vacuum physical vapor deposition chamber (< 10^{-7} Torr). A thin silicon layer was deposited by dc magnetron sputtering in Ar ambient at room temperature with various sputter times ranging from 5s, 10s, 30s, and 1 m. The silicon deposition rate was estimated to be 0.25 Å/sec. Control samples without any sputtered silicon layer were fabricated as reference devices. The device cross-sectional view is

shown in the inset of Fig. 5.25. Despite the undercut in the SiO_2 layer, silicon is still expected to be deposited onto the sidewall due to the large angle of incidence during the sputtering process, although the thickness of the silicon layer on the sidewall will be substantially less that the thickness deposited onto planar regions. Electrical characteristics were measured by an Agilent B1500A semiconductor device analyzer with the devices in a vacuum ($<1\times10^{-3}$ mbar) probe chamber. The thin layer of silicon on top of the TaN gate electrode was easily penetrated by the probe tip in order to make good electrical contact, and contact to the bottom n+ substrate electrode was achieved through the wafer chuck.

Figure 5.25 compares the electroforming process (the first voltage sweep) of several SSiRRAM devices and the control sample. For the control sample, in the first high voltage sweep, a sudden current increase up to ~ 10^{-7} to 10^{-6} A was induced at around 30V, the subsequent backward voltage sweep induces high current and switches the device into a LRS and the CF was successfully formed. The voltage where the large current jump occurs is defined as V_{EF} (30 V for the control sample). The SSiRRAMs with 5s sputter time showed reduced V_{EF} of 23 V, and, with longer sputter times, V_{EF} was further reduced. Even though it was difficult to define a precise V_{EF} for devices with longer sputtering times, a very clear trend is observed where longer sputtering times make it easier to electroform the device. As a result of the electroforming process, silicon nanocrystals have been observed to form within the SiO_2 layer, and to generally align along the direction of current flow, thus creating a conducting path between the two electrodes [31]. By sputtering a thin silicon layer on the sidewall, which provides an external silicon source, Fig. 5.25 demonstrates that the electroforming process was

significantly accelerated. Once electroformed, the device is switched to the LRS by applying a linear voltage sweep to ~4V. The device is switched to the HRS by applying a sweep to a voltage > ~6V. The low-voltage (~1V) conductance in the HRS is a strong function of the previous applied voltage, with conductance decreasing for higher applied voltage. For the devices used in this study, applying 6V resulted in a medium resistance state (MRS), and applying 9V resulted in a HRS with lower conductance (at 1V bias) than the MRS.

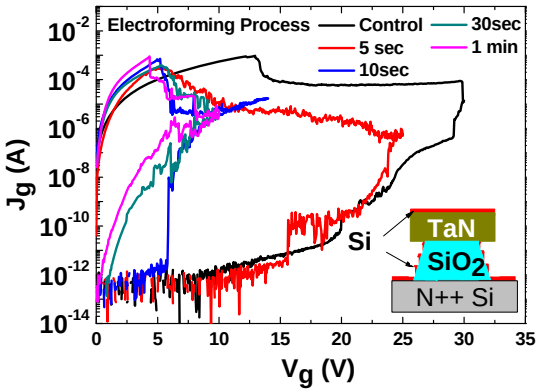


Figure 5.25 I-V curves of the SSiRRAM electroforming process for devices with different silicon sputtering times. The control sample I-V curve is plotted for comparison. The inset shows the schematic cross-section of the SSiRRAM.

Possible reasons for why the external silicon layer may result in a more robust CF include the formation of an increased oxygen vacancy concentration within the SiO_2 layer very near the sidewall surface. Oxygen vacancies tend to cluster in higher concentration near SiO_2/silicon interfaces [116]. As a result of the discontinuous silicon layer deposited onto the sidewall, we expect vacancies to cluster near each nanocrystal on the sidewall, thereby producing a modulated vacancy concentration along the sidewall that can potentially support a higher percolation current. Compared to control devices without the

sputtered silicon layer, electroformation can then proceed with lower applied voltage and a more robust CF can form with a larger cross-section and better uniformity across the sidewall. This scenario is consistent with a defect-driven switching mechanism [33].

We examined the set and reset process characteristics for SSiRRAMs, such as LRS current, HRS current, set power, and reset power, as defined in the inset of Fig. 5.26. Device current was read at 1 V. The Set Current (I_{Set}) /Reset Current (I_{Reset}) and Set Voltage (V_{Set})/ Reset Voltage (V_{Reset}) are the values of the current/voltage detected at the beginning of the resistance switching from a HRS/LRS state to a LRS/HRS state, respectively. These statistical results were collected after switching each sample 30 times between LRS and HRS. The set power ($I_{Set}V_{Set}$) and reset power ($I_{Reset}V_{Reset}$) describe the power needed to initiate recovery and rupture of the CF, respectively. Figure 5.26 shows that LRS current increased with increasing silicon sputtering time. This could be due to formation of more robust CF with increasing sputtering time, which is then able to conduct higher current. The variance in LRS current, shown in the inset of Fig. 5.27, was reduced in samples with longer sputtering time, suggesting a more uniform CF. Figure 5.27 shows the reset power as a function of silicon sputtering time. Due to a more robust CF in the samples with longer sputtering time, higher power was required to rupture the CF. The inset of Fig. 5.27 shows that LRS average current increased with increasing sputtering time, while the coefficient of variation followed the opposite trend (coefficient of variation was defined as the standard deviation divided by the average). HRS current and the set power are similar for each sample (data not shown).

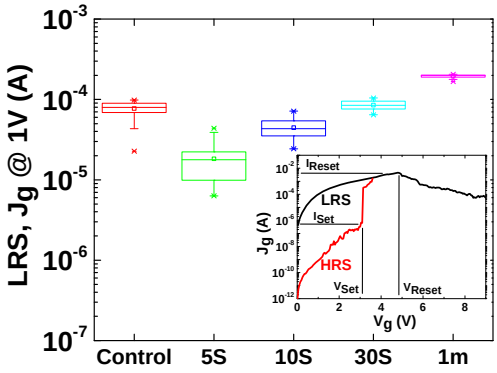


Figure 5.26 Statistical plot of LRS current at 1V as a function of silicon sputtering time. The inset shows typical LRS and HRS switching I-V curves, and defines parameters I_{Set}, I_{Reset}, V_{Set}, and V_{Reset}.

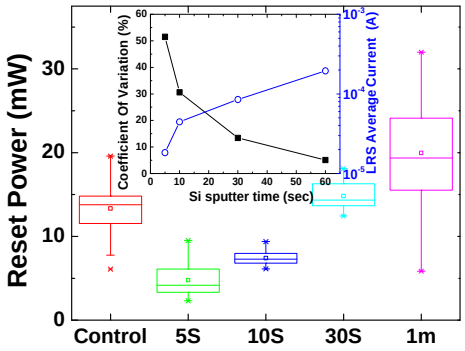


Figure 5.27 Statistical plot of SSiRRAM reset power. The inset shows the % coefficient of variation of LRS current and LRS average current for SSiRRAMs with different silicon sputtering times.

By changing the span of the reset voltage sweep, tri-state switching was realized in the SSiRRAM, as shown in Fig. 5.28. To program the device to LRS, MRS and HRS, the voltage was swept to 4.5 V, 6 V and 9 V, respectively. The devices were cycled 100 times between the 3 states. In devices with longer Si sputtering times, HRS current variability was substantially reduced, presumably due to a more robust, more uniform CF

as compared to the control sample. Therefore, it was possible to program a MRS in between LRS and HRS with a 10× separation between each state [124].

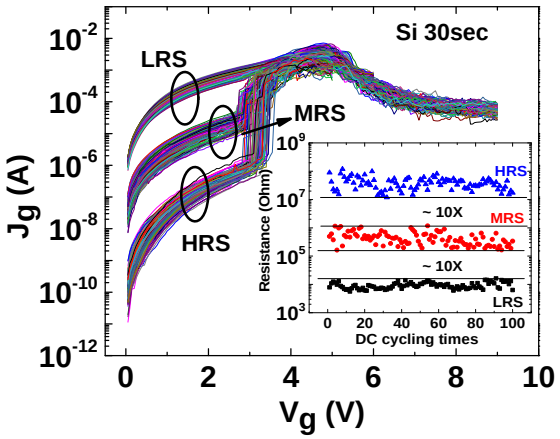


Figure 5.28 I-V curves for 100 switching cycles for SSiRRAM device with 30 sec silicon sputtering time. Multilevel programming is demonstrated by a 10× separation between the three states. The inset shows the resistance plot of the three states (with current read at 1V).

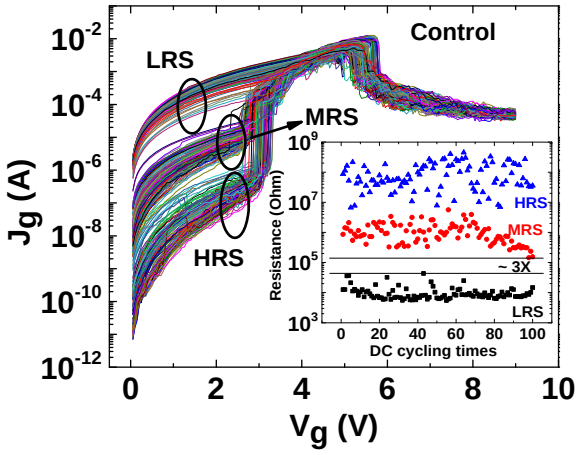


Figure 5.29 I-V curves for 100 switching cycles for the control sample. Programmed HRS current spreads-out and becomes mixed with MRS. The inset shows the resistance plot of the three states (with current read at 1V).

As a reference, the control sample was also examined for tri-state programming, as shown in Fig. 5.29. The large variance in programmed current, which has been observed in previous reports [122], results in the inability to distinguish the MRS from the HRS, and there is only ~ 3× separation between LRS and MRS.

One SiRRAM sample was dipped in a Tetramethylammonium Hydroxide (TMAH) solution to remove the sidewall silicon layer. The SiRRAM still operated normally without any additional electroforming process (data not shown). This suggests that the CF was not on the sidewall silicon layer but within the SiO_2 bulk of SiRRAM. Moreover, a vacuum environment and an etched sidewall are required to form a CF in SiO_2-based RRAMs [122]. We speculate that the CF is in the SiO_2 bulk but very near the sidewall surface. Possible mechanisms for more robust CF in SiSRAM include the formation of an increased oxygen vacancy concentration within the SiO_2 layer very near the sidewall surface. Oxygen vacancies tend to cluster in higher concentration near SiO_2/silicon interfaces [116]. As a result of the discontinuous silicon layer deposited onto the sidewall, we expect vacancies to cluster near the silicon layer on the sidewall, thereby producing a modulated vacancy concentration along the sidewall that can potentially support a high percolation current. Note that, compared to the control sample, electroformation in SiRRAM can then proceed with lower applied voltage and a more robust CF can form with a larger cross-section and better uniformity along the sidewall.

Defect clustering near the sputtered silicon layer potentially form defect-rich clusters in the near-surface regions along the SiO_2 sidewall. In the LRS, high-conductance defects may form a continuous filament, whereas, in the HRS, defects in a certain region of the CF may be converted to a low-conductance state, forming a

conductance “gap” that reduces current flow to ~1nA. In the MRS, fewer defects may be converted to the low-conductance state, possibly leading to a narrower gap region containing a mix of low- and high-conductance defects that is only capable of sustaining a medium-level current ~1μA. Electron traps such as the dimer oxygen vacancy (E'_{δ} center) and the oxygen vacancy/hydrogen molecule complex, both having a shallow thermodynamic energy level near the a-SiO_2 valence band-edge [116], are examples of defects considered to have low conductance, whereas the gamma oxygen vacancy (E'_{γ} center) and the oxygen vacancy/hydrogen atom complex (hydrogen bridge) both have a thermodynamic energy level near the middle of the a-SiO_2 energy band-gap so that they can easily exchange charge with device electrodes [116] to support high conductance. As a result, defects along the CF can potentially be converted between low-conductance defects (E'_{δ} center or oxygen vacancy/hydrogen molecule complex) and high-conductance defects (E'_{γ} center or hydrogen bridge) to accomplish reversible switching.

With stable HRS current characteristics, tri-state pulsed endurance testing demonstrated SSiRRAM performance up to 10^6 cycles for the 3 programmed states (Fig. 5.30). Pulse voltages of 6 V, 9 V, and 12V (with pulse width of 500 ns) were used to switch the device into LRS, MRS, and HRS, respectively. Figure 5.31 shows the read disturb immunity of programmed data during constant voltage stress at 1 V, where no degradation is observed in the three SSiRRAM resistance states during the stress time of 10^3 seconds.

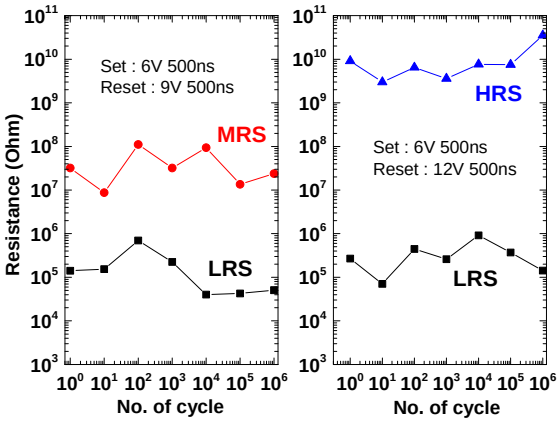


Figure 5.30 Tri-state pulse switching endurance of 10^6 cycles for SSiRRAM devices.

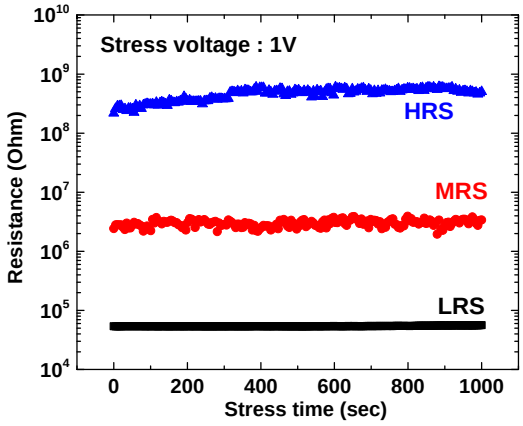


Figure 5.31 The read disturb immunity of LRS, MRS, and HRS by constant voltage stress of 1 V for 1000 seconds.

Figure 5.32 demonstrates the thermal stability of SSiRRAM devices. Devices were electroformed and set to LRS, MRS and HRS in vacuum. Then the devices were baked on a hot chuck (in air) for continuous high temperature data retention testing. The first 2 hours of 100°C baking and the next 2 hours of 150°C baking did not affect the 3 states. During the fifth hour of baking at 200°C, the HRS and MRS became somewhat unstable and a decrease in resistance was measured. However, after cooling the devices

to room temperature, both HRS and MRS recovered to their initial status and the devices could operate normally in vacuum (data not shown). This indicates that even though a vacuum environment is required to electroform and program SSiRRAM devices, the stored data can survive exposure to high temperature and remains nonvolatile in air ambient.

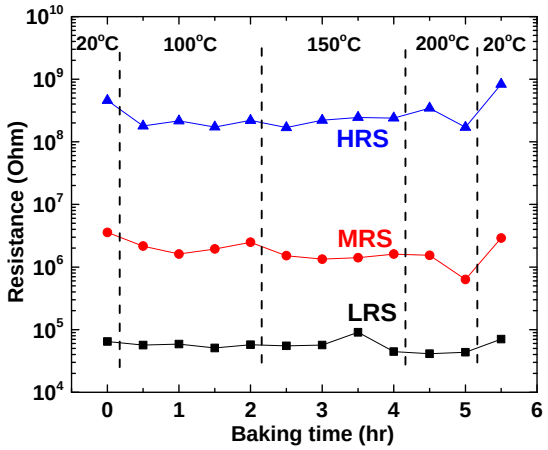


Figure 5.32 Retention properties of LRS, MRS, and HRS under continuous high temperature environment. HRS and MRS became unstable in the 5th hour of baking (200°C). When cooled to 20°C, the 3 states recovered to their initial status.

5.6 Summary

In conclusion, CF growth and resistance switching of SiO_x RRAMs have been studied. It was found that CFs grew from both electrodes and sometimes the CF can be formed quickly near a single electrode, potentially due to a higher concentration or weaker pre-existing defects. The switching is induced by rupture and recovery in a randomly formed weak spot along the CF where the CF is possibly narrower so that less current is conducted as compared to other regions of the CF. Therefore, it is concluded

that the weak spot is favorable to rupture/recovery and that reversible switching in SiO_x RRAM is a localized phenomenon. In contrast to other materials, the weak spot formation in SiO_x is a random process, and can occur in any location along the CF.

Moreover, the effects of incorporating a thin silicon layer onto the sidewall of SiO_2-based RRAMs were presented. The V_{EF} was significantly reduced and instability of HRS current was substantially improved in the SSiRRAM devices. With longer sputtering time, a more robust and more uniform CF was potentially formed, resulting in higher LRS average current with small variance. Meanwhile, the instability of LRS and HRS current was also significantly improved in the SSiRRAM devices. The addition of a sputtered-silicon thin layer on the sidewall may promote formation of more robust or more uniform CFs, which acts to effectively reduce HRS current instability. Consequently, a MRS (medium resistive state) can be programmed between LRS and HRS, and remained distinguishable with tri-state pulse endurance performance over 10^6 cycles. Stored SSiRRAM tri-state data show read disturb immunity at constant 1V stress for 1000 seconds, and the programmed states can be sustained during 150^{o}C thermal disturbance. These results suggest that SSiRRAMs are promising for future multilevel memory applications.

Chapter 6 Summary and future work

6.1 Summary

With Si CMOS roadmap approaching its physical limit, tremendous research has been conducted to identify promising technologies to extend Moore's Law. This Ph. D. dissertation investigated the device characteristics of high-k/III-V MOSFETs with various novel techniques including, adapting H_2O as the oxidizer in ALD deposition, inserting a thin passivation layer, and performing post-gate fluorine plasma treatment. These techniques target our ultimate goal, i.e., to realized high performance III-V MOSFETs, which could be a solution for future high speed and low power application.

First, the proper fabrication process for high-k/InGaAs MOSFETs has been identified by comparing physical analysis and device performance from H_2O-based vs. O_3-baesd high-k. We found that excessive oxygen concentration (using O_3 as the ALD precursor) induces III-V native oxides at the interface, leading to high interface trap density. Moreover, we identify the correlation between the interface chemistry (elemental As, AsO_x, GaO_x, and In_2O_3) and electrical parameters, such as hysteresis, D_{it}, SS, G_m, and I_{on}. We found that the interface chemistry degrades electrical performance and H_2O, as the oxidizer, is favorable for InGaAs MOSFETs. In addition, it has been found that Ti-based materials as a capping layer on top of the high-k layer helps to mitigate oxidation at the interface, resulting in a better interface quality and MOSFETs performance.

Secondly, III-V MOSFETs with various IPLs have been investigated to improve interface quality and therefore, increase electron mobility and drive current. Si IPL was

found to be an effective IPL to improve interface quality of InP MOSFETs. On the other hand, Al_2O_3, $HfAlO_x$, and $ZrAlO_x$ are demonstrated to be effective interfacial dielectric layers to improve device performance, including frequency dispersion, SS, current driving capability, effective channel mobility, and reliability.

Thirdly, a novel CF_4 post-gate plasma treatment was employed on III-V MOSFETs to improve both high-k bulk and interface quality. Fluorine incorporation was demonstrated on various high-k/III-V gate stacks, including $Al_2O_3/In_{0.53}Ga_{0.47}As$, Al_2O_3/InP, $HfO_2/In_{0.53}Ga_{0.47}As$, and HfO_2/InP. Significant improvements on interface quality and high-k bulk quality have been achieved. With F incorporation, we have successfully developed excellent interface quality of high-k (Al_2O_3 or HfO_2) directly on III-V substrate ($In_{0.53}Ga_{0.47}As$, or InP) without using interface passivation layer. For $HfO_2/In_{0.53}Ga_{0.47}As$ gate stack, fluorinated samples exhibit low D_{it} of 4.9×10^{11} $cm^{-2}eV^{-1}$, which is the lowest value over prior reported $HfO_2/In_{0.53}Ga_{0.47}As$ gate stacks. Furthermore, a "two-step" high-k deposition with incorporating F between two high-k depositions was able to further increase F concentration in the high-k/III-V gate stack. Therefore, further improvements on gate stack quality have been obtained.

Finally, the conducting mechanism of SiO_x RRAMs has been investigated. From device area effects and SiO_x thickness effects, we found that there might be only one CF responsible for on/off switching. Moreover, devices with larger perimeter can be formed easier compared to ones with small perimeter due to the higher possibility of containing weak spot at the edge. We purposed a model for SiO_x RRAMs: the conducting filament is randomly formed within the SiO_x at the sidewall edge, depending on pre-existing defects. The rupture/recovery could occur anywhere along the conducting filament, depending on

a random process that determines the location of the weak spot along the conducting filament. In addition, we have improved SiO_2-based RRAM performance by incorporating a thin silicon layer onto its sidewall. This technique significantly reduced the electroforming voltage and instability of HRS current of SiO_2-based RRAMs. As a result, tri-state application has been demonstrated and the pulse endurance test can sustain 10^6 cycles. The data stored had good immunity of read and thermal disturbance.

6.2 Suggestions for future work

6.2.1 Surface channel III-V MOSFETs

Even though we have demonstrated excellent interface quality of HfO_2/InGaAs MOSFETs by fluorine incorporation, we would never be satisfied with the progress that we have achieved. Further improvement will be highly required for realizing high performance high-k/III-V MOSFETs. We still need to explore novel high-k materials, which might have good interface quality with III-V. For example, Intel has been using $TaSiO_x$ as high-k gate dielectrics on III-V substrate. However, Intel never discloses the composition and analysis of $TaSiO_x$. It is worth investigating $TaSiO_x$ further to have a better understanding.

Another important issue on high-k/III-V MOSFETs is the reliability. Since high-k/III-V MOSFETs are getting matured and closer to mass production, reliability issue is a major concern. However, reliability has not been addressed completely. Several reliability issues of high-k/III-V MOSFETs, mostly adapted from Si, need to be solved accurately. For example, TDDB or BTI testing on high-k/III-V MOSFETs, V_{th} shift issue, gate oxide integrity, stress induced leakage current, and electrostatic discharge.

6.2.2 SiO_2-based RRAMs

The conducting mechanism of SiO_2-based RRAMs is not fully understood yet. The random formation model we purposed in this dissertation is a macro-description of the filament formation and switching. A detail and micro-oriented model of SiO_2-based RRAMs is still under debate. More electrical characterization and physical analysis are required to unveil the conducting filament property. For example, conductive-AFM might be able to observe the conducting filament status while it is formed and switching.

In terms of SiO_2-based RRAMs performance, the switching speed needs to be improved to be compatible with HfO_2-based RRAMs. The switching speed used in our study was 500 ns, which is far away from the state-of-the art of 0.3 ns in HfO_2-based RRAMs. We need to engineer the SiO_2 quality, composition, or adding another element into the SiO_2, to stabilize the conducting filament and further speed up the switching. Deuterium annealing helps to reduce electroforming voltage, and might be able to increase the switching speed.

Finally, another issue of SiO_2-based RRAMs is their relative high V_{set} and V_{reset} compared to HfO_2-based RRAMs. High V_{set} and V_{reset} mean that it takes more power to switch LRS/HRS, which is detrimental for SiO_2-based RRAMs to be practically used in production. So far, V_{set} is always around 3~4 V and V_{reset} is always around 5~7 V no matter the thickness of SiO_2, the device perimeter, composition of SiO_2, … etc. This phenomenon seems highly related to the conducting mechanism of the SiO_2-based RRAMs. We need to understand the conducting mechanism first, then, we could find a solution to reduce V_{set} and V_{reset}.

Bibliography

[1] http://www.intel.com/about/companyinfo/museum/exhibits/moore.htm

[2] S. Thompson, et al., "A 90 nm logic technology featuring 50 nm strained silicon channel transistors, 7 layers of Cu interconnects, low-k ILD, and 1 um2 SRAM cell", Tech. Dig. –Int. IEEE Electron Devices Meet. 20042 61

[3] P. Bai, et al., "A 65nm logic technology featuring 35nm gate lengths, enhanced channel strain, 8 silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% b-free packaging", Tech. Dig. –Int. IEEE Electron Devices Meet. 2004 657

[4] K. Mistry, et al., "A 45nm logic technology with high-k + metal gate transistors, strained silicon 9 Cu interconnect layers, 193nm dry patterning, and 100% b-free packaging", Tech. Dig. –Int. IEEE Electron Devices Meet. 2007 247

[5] S. Natarajan, et al., "A 32 nm logic technology featuring 2nd-generation high-k + metal gate transistors, enhanced channel strain and 0.171um2 SRAM cell size in a 291 MB array", Tech. Dig. –Int. IEEE Electron Devices Meet. 2008 941

[6] http://newsroom.intel.com/community/intel_newsroom/blog/2011/05/04/intel-reinvents-transistors-using-new-3-d-structure/

[7] R. Chau, presentation in the 14th biennial Conference on Insulating Films on Semiconductors, INFOS 2005

[8] S. Takagi, M. Sugiyama, T. Yasuda, and M. Takenaka, "Ge/III-V channel engineering for future CMOS", ECS Trans. 19 9, 2009

[9] H. Hasegawa, M. Akazawa, A. Domanowska, and B. Adamowicz, "Surface passivation of III-V semiconductors for future CMOS devices-past research, present status and key issues for future", Applied Surface Science 256, pp. 5698, 2010

[10] http://www.itrs.net/Links/2011ITRS/2011Chapters/2011PIDS.pdf

[11] K. Chen, C. Hu, el al., "Optimizing sub-quarter micron CMOS circuit speed considering interconnect loading effects", *IEEE Trans. Electron Devices*, vol. 44, no. 9, pp. 1556, 1997

[12] G. Dewey, R. Kotlyar, R. Pillarisetty, M. Radosavljevic, T. Rakshit, H. Then, and R. Chau, "Logic performance evaluation and transport physics of schottky-gate II-V compound semiconductor quantum well field effect transistors for power supply voltages (Vcc) ranging from 0.5V to 1.0V", Tech. Dig. –Int. IEEE Electron Devices Meet. 2009 487

[13] Y. -T. Chen, H. Zhao, J. H. Yum, Y. Wang, F. Xue, F. Zhou, and J. C. Lee, Appl. Phys. Lett. 95, 013501 (2009)

[14] Y.-T. Chen, H. Zhao, J. H. Yum, Y. Wang, F. Xue, F. Zhou, and J. C. Lee, J. Electrochem. Soc. 157 G71 (2010)

[15] iRAP report "ADVANCED SOLID-STATE MEMORY SYSTEMS AND PRODUCTS: EMERGING NON-VOLATILE MEMORY TECHNOLOGIES, INDUSTRY TRENDS AND MARKET ANALYSIS", Publish Date: *Apr 2011*, Pages: 146, Report Code: et-114

[16] W. Zhao, L. Torres, Y. Guillemenet, L. V. Cargnini, Y. Lakys, J.-O. Klein, D. Ravelosona, G. Sassatelli, and C. Chappert, "Design of MRAM based logic circuits and its applications", Proceedings of the 21st edition of the great lakes symposium on Great lakes symposium on VLSI, pp. 431, 2011

[17] C. J. Lin, S. H. Kang, Y. J. Wang, K. Lee, X. Zhu, W. C. Chen, X. Li, W. N. Hsu, Y. C. Kao, M. T. Liu, W. C. Chen, Y. Lin, M. Nowak, N. Yu, and L. Tran, "45nm low power CMOS logic compatible embedded STT MRAM utilizing a reverse-connection 1T/1MTJ cell", Tech. Dig. –Int. IEEE Electron Devices Meet. 2009 279

[18] M. Qazi, M. Clinton, S. Bargling, A. P. Chandrakasan, "A low-voltage 1Mb FRAM in 0.13 um CMOS featuring time-to-digital sensing for expanded operation margin", Solid-State Circuits, vol. 47, 1, pp. 141, 2012

[19] A. Gyanathan, and Y.-C. Yeo, "Novel multi-level PCRAM cell with Ta_2O_5 barrier layer in between a graded $Ge_2Sb_2Te_5$ stack", VLSI Technology, Systems and Applications (VLSI-TSA), 2011

[20] Y. H. Tseng, C.-E. Huang, C.-H. Kuo, Y.-D. Chih, and C. J. Lin, "High density and ultra-small cell size of contact ReRAM (CR-RAM) in 90nm CMOS logic technology and circuits", Tech. Dig. –Int. IEEE Electron Devices Meet. 2009 109

[21] D. B. Strukov, et al., *Nature* , vol. 453, pp. 80-83, May 1, 2008

[22] http://www.itrs.net/Links/2010ITRS/Home2010.htm 2010 Future Memory Devices Workshop Summary

[23] C.-H. Wang, Y.-H. Tsai, K.-C. Lin, M.-F. Chang, Y.-C. King, C.-J. Lin, S.-S. Sheu, Y.-S. Chen, H.-Y. Lee, F. T. Chen, and M.-J. Tsai, "Three-dimensional $4F^2$ ReRAM cell with CMOS logic compatible process", Tech. Dig. –Int. IEEE Electron Devices Meet. 2010 664

[24] H. Y. Lee, Y. S. Chen, P. S. Chen, P. Y. Gu, Y. Y. Hsu, S. M. Wang, W. H. Liu, C. H. Tsai, S. S. Sheu, P. C. Chiang, W. P. Lin, C. H. Lin, W. S. Chen, F. T. Chen, C. H. Lien, and M.-J. Tsai, "Evidence and solution of over-reset problem for HfOx based resistive memory with sub-ns switching speed and high endurance", Tech. Dig. –Int. IEEE Electron Devices Meet. 2010 460

[25] H. Y. Lee, P. S. Chen, T. Y. Wu, Y. S. Chen, C. C. Wang, P. J. Tzeng, C. H. Lin, F. Chen, C. H. Lien, and M.-J. Tsai, Tech. Dig. –Int. IEEE Electron Devices Meet. 2008, 297

[26] K. M. Kim, B. J. Choi, S. J. Song, G. H. Kim, and C. S. Hwang, J. Electrochem. Soc. 156 G213 (2009)

[27] M. K. Yang, J-W. Park, T. K. Ko, and J-K. Lee, Appl. Phys. Lett. 95, 042105 (2009)

[28] J. J. Yang, M.-X. Zhang, J. P. Strachan, F. Miao, M. D. Pickett, R. D. Kelly, G. Medeiros-Ribeiro, and R. S. Williams, Appl. Phys. Lett. 97, 232102 (2010)

[29] L. Zhang, R. Huang, M. Zhu, S. Qin, Y. Kuang, D. Gao, C. Shi, and Y. Wang, IEEE Electron Device Lett. 31, 966 (2010)

[30] Y. Y. Chen, G. Pourtois, X. P. Wang, C. Adelmann, L. Goux, B. Govoreanu, L. Pantisano, S. Kubicek, L. Altimime, M. Jurczak, J. A. Kittl, G. Groeseneken, and D. J. Wouters, International Memory Workshop, 2011

[31] J. Yao, Z. Sun, L. Zhong, D. Natelson, and J. M. Tour, Nano Lett. 10, 4105 (2010)

[32] J. Yao, L. Zhong, D. Natelson, and J. M. Tour, Appl. Phys. Lett. 93, 253101 (2008)

[33] Y. Wang, Y.-T. Chen, F. Xue, F. Zhou, Y.-F. Chang, B. Fowler, and J. C. Lee, Appl. Phys. Lett. 100,083502 (2012)

[34] J. Yao, Z. Sun, L. Zhong, D. Natelson, and J. M. Tour, Scientific Reports, 2:242 (2012)

[35] ITRS for Semiconductor, 2010 update. http://www.itrs.net/Links/2010ITRS/Home2010.htm.

[36] J. Robertson, Appl. Phys. Lett. 94, 152104 (2009).

[37] B. Brennan, M. Milojevic, H. C. Kim, P. K. Hurley, J. Kim, G. Hughes, and R. M. Wallace, Electrochem.Solid-State Lett. 12, H205 (2009).

[38] Y. S. Chen, H. Y. Lee, P. S. Chen, P. Y. Gu, C. W. Chen, W. P. Lin, W. H. Liu, Y. Y. Hsu, S. S. Sheu, P. C. Chiang, W. S. Chen, F. T. Chen, C. H. Lien, and M. J. Tsai, Tech. Dig. – Int. Electron Devices Meet. 2009, 1.

[39] C. Choi, and J. C. Lee, J. Appl. Phys. 108, 064107 (2010).

[40] S.-Y. Wang, D.-Y. Lee, T.-Y. Tseng, and C.-Y. Lin, Appl. Phys. Lett. 95, 112904 (2009).

[41] Y. Yuan, L. Wang, B. Yu, B. Shin, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, IEEE Electron Device Lett. 32, 485 (2011).

[42] R. Xie, T. H. Phung, W. He, Z. Sun, M. Yu, Z. Cheng, and C. Zhu, " High mobility high-k/Ge pMOSFETs with 1 nm EOT -New concept on interface engineering and interface characterization," in *IEDM Tech. Dig*., 2008, pp. 393-396

[43] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. Dekeersmaecker, "A reliable approach to charge-pumping measurements in MOS transistors," *IEEE Trans. Electron Devices*, vol. 31, no. 1, pp. 42-53, Jan. 1984

[44] S. J. Koester, E. W. Kiewra, Yanning Sun, D. A. Neumayer, J. A. Ott, D. K. Sadana, D. J. Webb, J. Fompeyrine, J. –P. Locquet, M. Sousa, and R. Germann, IEEE 64th Device Research Conference, 43, (2007)

[45] P. D. Ye, G. D. Wilk, J. Kwo, B. Yang, H.-J. L. Gossmann, M. Frei, S. N. G. Chu, J. P. Mannaerts, M. Sergent, M. Hong, K. K. Ng, and J. Bude, IEEE Electron Device Lett., 24, 209, (2003)

[46] C. P. Chen, T. D. Lin, Y. J. Lee, Y. C. Chang, M. Hong, and J. Kwo, Solid-State Electronics 52, 1615 (2008)

[47] H. Zhao, D. Shahrjerdi, F. Zhu, M. Zhang, H.-S. Kim, I. Ok, J. H. Yum, S. Park, S. K. Banerjee, and J. C. Lee, Appl. Phys. Lett., 92, 233508, (2008)

[48] I. Ok, H. Kim, M. Zhang, F. Zhu, S. I. Park, J. Yum, H. Zhao, D. Garcia, P. Majhi, N. Geol, W. Tsai, C. K. Gaspe, M. B. Santos, and J. C. Lee, Appl. Phys. Lett., 92, 202903, (2008)

[49] Y. Xuan, Y. Q. Wu, H. C. Lin, T. Shen, and P. D. Ye, IEEE Electron Device Lett., 28, 935, (2007)

[50] M. Passlack, J. K. Abrokwah, R. Droopad, Y. Zhiyi, C. Overgaard, S. I. Yi, M. Hale, J. Sexton, and A. C. Kummel, IEEE Electron Device Lett., 23, 508, (2002)

[51] I. OK, H. Kim, M. Zhang, T. Lee, F. Zhu, L. Yu, S. Koveshnikov, W. Tsai, V. Tokranov, M. Yakimov, S. Oktyabrsky, and J. C. Lee, IEEE Electron Devices Meeting, (2006)

[52] H.-S. Kim, I. Ok, F. Zhu, M, Zhang, S. Park, J. Yum, H. Zhao, and J. C. Lee, IEEE 65th Device Research Conference, 99, (2007)

[53] F. Gao, S. J. Lee, R. Li, S. J. Whang, S. Balakumar, D. Z. Chi, C. C. Kean, S. Vicknesh, C. H. Tung, and D. –L. Kwong, IEEE Electron Devices Meeting, (2006)

[54] H. Zhao, D. Shahrjerdi, F. Zhu, H.-S. Kim, I. Ok, M. Zhang, J. H. Yum, S. K. Banerjee, and J. C. Lee, Appl. Phys. Lett., 92, 253506, (2008)

[55] Y. Q. Wu, Y. Xuan, T. Shen, P. D. Ye, Z. Cheng, and A. Lochtefeld, Appl. Phys. Lett., 91, 022108, (2007)

[56] Y. Q. Wu, M. Xu, Y. Xuan, P. D. Ye, J. Li, Z. Cheng, and A. Lochtefeld, University/Government/Industry Micro/Nano Symposium, 49, (2008)

[57] Y. Xuan, Y. Wu, and P. D. Ye, IEEE Electron Device Lett., 29, 294, (2008)

[58] K. Rajagopalan, J. Abrokwah, R. Droopad, M. Passlack, IEEE Electron Device Lett., 27, 959, (2006)

[59] D. Shahrjerdi, M. M. Qye, A. L. Holmes, Jr., and S. K. Banerjee, Appl. Phys. Lett., 89, 043501, (2006)

[60] Y. Xuan, Y. Wu, T. Shen, T. Yang, and P. Ye, IEEE Electron Devices Meeting, 637, (2007)

[61] Y. Wu, Y. Xuan, P. Ye, Z. Cheng, and A. Lochtefeld, IEEE 65th Device Research Conference, 117, (2007)

[62] M. Frank, G. Wilk, D. Starodub, T. Gustafsson, E. Garfunkel, Y. Chabal, J. Grazul, and D. Muller, Appl. Phys. Lett., 86,152904, (2005)

[63] D. Shahrjerdi, E. Tutuc, and S. K. Banerjee, Appl. Phys. Lett., 91, 063501, (2007)

[64] C. A. Richter, A. R. Hefner, and E. M. Vogel, IEEE Electron Device Lett., 22, 35, (2001)

[65] N. Geol, D. Heh, S. Koveshnikov, I. Ok, S. Oktyabrsky, V. Tokranov, R. Kambhampati, M. Yakimov, Y. Sun, P. Pianetta, C. Gaspe, M. Santos, J. Lee, S. Datta, P. Majhi, W. Tsai, Tech. Dig. *Int. Electron Devices Meet.,* 363 (2008)

[66] J. Huang, N. Geol, H. Zhao, C. Y. kang, K. S. Min, G. Bersuker, S. Oktyabrsky, C. K. Gaspe, M. B. Santos, P. Majhi, P.D. Kirsch, H.-H. Tseng, J. C. Lee, and R. Jammy, Tech. Dig. *Int. Electron Devices Meet.,* 335 (2009)

[67] R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz, and M. Radosavljevic, IEEE Trans. Nanotechnol. 4, 153, (2005)

[68] M. Passlack, M. Hong, J. Mannaerts, R. Opila, S. Chu, N. Moriya, and F. Ren, IEEE Trans. Electron Devices 44, 214, (1997)

[69] M. Hong, J. Kwo, A. R. Kortan, J. P. Mannaerts, and A. M. Sergent, Science 283, 1897 (1999)

[70] M. Passlack, P. Zurcher, K. Rajagopalan, R. Droopad, J. Abrokwah, M. Tutt, Y. Park, E. Johnson, O. Hartin, A. Zlotnicka and P. Fejes, IEEE Electron Devices Meeting, 621 (2007)

[71] H. Kim, I. Ok, M. zhang, C. Choi, T. Lee, F. Zhu, G. Thareja, L. Yu, and J. C. Lee, Appl. Phys. Lett., 88, 252906, (2006)

[72] Y.-T. Chen, H. Zhao, J. H. Yum, Y. Wang, and J. C. Lee, Appl. Phys. Lett., 94, 213505, (2009)

[73] Y. Xuan, T. Shen, Y. Wu, M. Xu, and P. Ye, IEEE 66th Device Research Conference, 37, (2008)

[74] S. Koveshnikov,N. Goel, P. Majhi, C. Gaspe, M. Santos, S. Oktyabrsky, V. Tokranov, M. Yakimov, R. Kambhampati, H. Bakhru, F. Zhu, J. Lee, and W. Tsai, IEEE 66th Device Research Conference, 43, (2008)

[75] W.-C.Wu, C-S. Lai, S.-C. Lee, M.-W. Ma, T.-S. Chao, J.-C. wang, C.-W. Hsu, P.-C. Chou, J.-H. Chen, K.-H. Kao, W.-C. Lo, T.-Y. Lu, L.-L. Tay, and N. Rowell, IEEE Electron Devices Meeting, 405, (2008)

[76] K. Seo, R. Sreenivasan, P. C. McIntyre, and K. C. Saraswat, J, IEEE Electron Device Lett., Vol. 27, No. 10, 821, (2006)

[77] Y. Mitani, H. Satake, Y. Nakasaki, and A. Toriumi, IEEE Trans. Electron Devices 50, 2211, (2003)

[78] R. Xie, M. Yu, M. Y. Lai, L. Chan, and C. Zhu, Appl. Phys. Lett., 92, 163505, (2008)

[79] C.S. Lai, W. C. Wu, T. S. Chao, J. H. Chen, J. C. Wang, L. L. Tay and N. Rowell, Appl. Phys. Lett., 89, 072904, (2006)

[80] Y. Xuan, Y. Q. Wu, H. C. Lin, T. Shen, and Peide D. Ye, IEEE Electron Device Lett., Vol. 28, No. 11, 935, (2007)

[81] Y.-T. Chen, H. Zhao, J. H. Yum, Y. Wang, F. Xue, F. Zhou, and J. C. Lee, Appl. Phys. Lett., 95, 013501, (2009)

[82] CRC *Handbook of Physics and Chemistry* (CRC, Boca Raton, 84th, 2003)

[83] Y. Xuan, Y. Wu, and P. Ye, *IEEE Electron Device Lett.*, vol. 29, no. 4, pp. 294-296, Apr. 2008

[84] H. Zhao, Y. Chen, J. H. Yum, Y. Wang and J. C. Lee, in *IEEE 67th Device Research Conference Dig.*, 2009, pp.89-90

[85] D. Shahrjerdi, E. Tutuc, and S. K. Banerjee, *Appl. Phys. Lett.*, vol. 91, no.6, pp. 063501, Aug. 2007

[86] H.-C. Chin, X. Liu, X. Gong, and Y.-C. Yeo, *IEEE Trans. Electron Devices*, vol. 57, no. 5, pp. 973-979, May 2010

[87] H. Shin, J. R. Weber, R. D. Long, P. K. Hurley, C. G. Van de Walle, and P. C. Mclntyre, *Appl. Phys. Lett*., vol. 96, no.15, pp. 152908, Apr. 2010

[88] B. J. DiMaria and J. W. Stasiak, *J. Appl. Phys*., vol. 65, no. 6, pp. 2342-2356, Mar. 1989

[89] D. J. DiMaria, E. Cartier, and D. Arnold, *J. Appl. Phys*., vol. 73, no. 7, pp. 3367-3384, Apr. 1993

[90] E. Efthymiou, S. Bernardini, J. F. Zhang, S. N. Volkos, B. Hamilton, A. R. Peaker, *Thin Solid Films*, vol. 517, no. 1, pp. 207-208, Nov. 2008

[91] M. Houssa, S. De Gendt, J. L. Autran, G. Groeseneken, and M. M. Heyns, *Appl. Phys. Lett*., vol. 85, no.11, pp. 2104, Sep. 2004

[92] R. Xie, W. He, M. Yu, and C. Zhu, Appl. Phys. Lett. 93, 073504 (2008)

[93] H.-C. Chin, X. Gong, L. Wang, and Y.-C. Yeo, *Electrochem.Solid-State. Lett*., vol. 13, no. 12, pp. H440-H442, Sep. 2010

[94] H. Zhao, J. H. Yum, Y.-T. Chen, and J. C. Lee, *J. Vac. Sci. Technol. B*., vol. 27, no. 4, pp. 2024-2027, Jul. 2009

[95] S. Lee, and J-W. Park, *J. Appl. Phys*., vol. 80, no. 9, pp. 5260-5263, Nov. 1996

[96] Y.-Y. Chen and C.-R. Hsieh, *IEEE Electron Device Lett.*, vol. 31, no. 11, pp. 1178-1180, Nov. 2010

[97] J. Robertson, *Microelectronic Engineering* vol. 86, no. 7-9, pp. 1558-1560 Jul.-Sep. 2009

[98] T. Tatsumi, S. Fukuda, S. Kadomura, *Jpn. J. Appl. Phys*. 1993, *32*, 6114–6118

[99] F. Xue, H. Zhao, Y.-T. Chen, Y. Wang, F. Zhou, and J. C. Lee, *J. Vac. Sci. Technol. B*., vol. 29, no. 4, pp. 040601, Jul./Aug. 2011

[100] J. Huang, N. Goel, H. Zhao, C. Y. Kang, K. S. Min, G. Bersuker, S. Oktyabrsky, C. K. Gaspe, M. B. Santos, P. Majhi, P. D. Kirsch, H.-H. Tseng, J. C. Lee, and R. Jammy, in *IEDM Tech. Dig*., 2009, pp. 335-338

[101] D. Kuzum, J.-H. Park, T. Krishnamohan, H-S. Philip Wong, and K. C. Saraswat, *IEEE Trans. Electron Devices*, vol. 58, no. 4, pp. 1015-1022, Apr. 2011

[102] J. Lin, S. Lee, H.-J. Oh, W. Yang, G. Q. Lo, D. L. Kwong, and D. Z. Chi, in *IEDM Tech. Dig*., 2008, pp. 401-404

[103] Y. Q. Wu, M. Xu, R. S. Wang O. Koybasi, and P. D. Ye, in *IEDM Tech. Dig*., 2009, pp. 323-326

[104] H. Zhao, N. Kong, Y. Chen, Y. Wang, F. Xue, F. Zhou, S. K. Banerjee, and J. C. Lee, in *IEEE 68th Device Research Conference Dig.*, 2010, pp.55-56

[105] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. Dekeersmaecker, *IEEE Trans. Electron Devices*, vol. 31, no. 1, pp. 42-53, Jan. 1984

[106] H. Zhao, J. Huang, Y.-T. Chen, J. H. Yum, Y. Wang, F. Xue, F. Zhou, and J. C. Lee, *Appl. Phys. Lett*., vol. 95, no.25, pp. 253501, Dec. 2009

[107] H. C. Chiu, L. T. Tung, Y. H. Chang, Y. J. Lee, C. C. Chang, J. Kwo, and M. Hong, *Appl. Phys. Lett*., vol. 93, no.20, pp. 202903, Nov. 2008

[108] Lin, D.; Brammertz, G.; Sioncke, S.; Fleischmann, C.; Delabie, A.; Martens, K.; Bender, H.;Conard, T.; Tseng, W.H.; Lin, J.C.; *et al*. In *Proceedings of 2009 IEEE International Electron Devices Meeting (IEDM)*, Baltimore, MD, USA, 7–9 December 2009; pp. 1–4.

[109] Y. C. Chang, M. L. Huang, K. Y. Lee, Y. J. Lee, T. D. Lin, M. Hong, J. Kwo, T. S. Lay, C. C. Liao, and K. Y. Cheng, *Appl. Phys. Lett*., vol. 92, no.7, pp. 072901, Feb. 2008

[110] Y. Hwang, V. Chobpattana, J. Y. Zhang, J. M. LeBeau, R. Engel-Herbert, and S. Stemmer, *Appl. Phys. Lett*., vol. 98, no.14, pp. 142901, Apr. 2011.

[111] C. Leroux, J. Mitard, G. Ghibaudo, X. Garros, G. Reimbold, B. Guillaumot, and F. Martin, Tech. Dig. –Int. IEEE Electron Devices Meet. 2004 737

[112] K. M. Kim, B. J. Choi, Y. C. Shin, S. Choi, and C. S. Hwang, Appl. Phys. Lett. 91, 012907 (2007)

[113] Y. Hokari, T. Baba, and N. Kawamura, IEEE Trans. Electron Devices 32, 2485 (1985)

[114] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wanger, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, IEEE Trans. Electron Devices 58, 3652 (2011)

[115] E. Y. Wu, J. Sun, and R.-P. Vollertsen, Dig. –Int. Reliability Physics Symposium. 2009, 708

[116] P. E. Blochl, Physical Review B 62, 6158, (2000)

[117] K. M. Kim, D. S. Jeong, and C. S. Hwang, Nanotechnology, 22, 254002 (2011)

[118] K. M. Kim, S. J. Song, G. H. Kim, J. Y. Seok, M. H. Lee, J. H. Yoon, J. Park, and C. S. Hwang, Adv. Func. Mater. 21 1587 (2011)

[119] T. Bakos, S. N. Rashkeev, and S. T. Pantelides, *Physical Review Letters* **88**-5, p. 055508-1 (2002)

[120] S.-Y. Wang, C.-W. Huang, D.-Y. Lee, T.-Y. Tseng, and T.-C. Chang, J. Appl. Phys. 108, 114110 (2010)

[121] Y.-F. Chang, T.-C. Chang, and C.-Y. Chang, J. Appl. Phys. 110, 053703 (2011).

[122] J. Yao, L. Zhong, D. Natelson, and J. M. Tour, Appl. Phys. A 102, 835 (2011).

[123] Y.-T. Chen, B. Fowler, Y. Wang, F. Xue, F. Zhou, Y.-F. Chang, and J. C. Lee, ECS Journal of Solid State Science and Technology, 1 (3) P1-P4 (2012)

[124] Z. Wei, Y. Kanzawa, K. Arita, Y. Katoh, K. Kawai, S. Muraoka, S. Mitani, S. Fujii, K.Katayama, M. Iijima, T. Mikawa, T. Ninomiya, R. Miyanaga, Y. Kawashima, K. Tsuji, A. Himeno, T. Okada, R. Azuma, K. Shimakawa, H. Sugaya, and T. Takagi, R. Yasuhara, K.Horiba, H. Kumigashira, and M. Oshima, Tech. Dig. –Int. IEEE Electron Devices Meet. 2008, 293