HIGH DIELECTRIC CONSTANT MATERIALS IN SONOS-TYPE NON-VOLATILE MEMORY STRUCTURES

TAN YAN NY
(B. Eng. and M. Eng., NUS)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE
2007

Acknowledgements

I would like to thank my thesis supervisors, Associate Professor Chim Wai Kin, Associate Professor Cho Byung Jin and Professor Choi Wee Kiong, for their teaching and guidance throughout my candidature.

I would like to thank Mr. Joo Moon Sig, Kim Sun Jung, Mr. Whang Sung Jin, Lau Boon Teck and O Yann Wai Lin for their assistance rendered during device fabrication in Silicon Nano Devices Laboratory. In addition, I would like to thank Walter Lim and Chew Han Guan for their kind assistance while working in Microelectronics Laboratory. In Centre for Integrated Circuits Failure Analysis and Reliability Laboratory, I would like to thank Mrs. Ho Chiew Mooi, Mr. Goh Thiam Peng, Anna Li and Koo Chee Kiong for their assistance in equipment maintenance. My appreciation to fellow postgraduate students, Tsu Hau, Guan Song, Xin Hua, Soon Leng, Soon Huat, Mans Osterberg, Jian Xin, Alfred, Szu Huat, Heng Wah, Shen Chen, Jin Quan and Wen Zhuo for making my stay in NUS an enriching experience.

Last but not least, I would like to thank my family for their constant support and encouragement.

Table of Contents

PAGE

ACKNOWLEDGEMENTS i

TABLE OF CONTENTS ii

SUMMARY v

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF SYMBOLS xvi

CHAPTER 1 **Introduction** **1**

1.1 Background 1

1.2 Motivation for the Project 4

1.3 Research Objectives 5

1.4 Organization of Thesis 6

References 8

CHAPTER 2 **Literature Review** **10**

2.1 History of Nonvolatile Memory Structures 10

2.2 Current and Future Nonvolatile Memories 16

2.3 SONOS Nonvolatile Memory 23

2.3.1 SONOS gate stack scaling 27

2.3.2 Novel SONOS Structures 28

References 34

CHAPTER 3 **Hafnium Oxide as the Charge Storage Layer in SONOS-type Nonvolatile Flash Memory for Minimization of the Over-erase Phenomenon** **40**

3.1 Introduction 40

3.2	Sample Fabrication	41
3.3	Results and Discussion	43
3.4	Summary	51
	References	52
CHAPTER 4	**Hafnium Aluminum Oxide as the Charge Storage Layer in SONOS-type Nonvolatile Memory for High-Speed Operation with Improved Charge Retention and Endurance Performance**	**54**
4.1	Introduction	54
4.2	Sample Fabrication	55
4.3	Results and Discussion	56
4.4	Summary	71
	References	72
CHAPTER 5	**Development of High-κ Blocking Oxide Layer in SONOS-type Nonvolatile Memory**	**76**
5.1	Introduction	76
5.2	Hafnium Aluminum Oxide Blocking Oxide Layer in SONOS-type Nonvolatile Memory for High-Speed Operation	77
	5.2.1 Introduction	77
	5.2.2 Sample Fabrication	78
	5.2.3 Results and Discussion	79
5.3	Evaluation of Lanthanum Aluminum Oxide and Lanthanum Yttrium Aluminum Oxide as the Blocking Oxide Layer in SONOS-type Nonvolatile Memory	93

5.3.1 Introduction 93

5.3.2 Sample Fabrication 94

5.3.3 Results and Discussion 95

(A) Evaluation of $(La_2O_3)_x(Al_2O_3)_{1-x}$ with different composition ratios as blocking oxide 95

(B) Feasibility study of $(LaAl)_xY_{1-x}O_3$ with different composition ratios as blocking oxide for SONOS memory 97

5.4 Summary 102

References 104

CHAPTER 6 **SONOS-type Nonvolatile Memory with Ultra-high-κ Charge Storage Layer and High-κ Tunnel and Blocking Oxide Layers** **108**

6.1 Introduction 108

6.2 Sample Fabrication 111

6.3 Results and Discussion 112

6.4 Summary 121

References 122

CHAPTER 7 **Conclusion** **125**

7.1 Summary of Findings 125

7.2 Recommendations for future work 128

References 130

LIST OF PUBLICATIONS 131

Summary

SONOS (polysilicon-oxide-silicon nitride-oxide-silicon) Flash memory is one of the more attractive candidates to realize FLASH vertical scaling. This work entails finding innovative solutions, using high dielectric constant (high-κ) materials, to overcome the limitations of the conventional floating gate structure as a result of rapidly shrinking device geometries.

The conventional method to increase the programming speed and to lower the operating voltage of SONOS devices is by reducing the tunnel oxide thickness. However, this seriously degrades the charge retention capability of the device. To overcome this limitation, the SOHOS (polysilicon-oxide-high-κ-oxide-silicon) Flash memory has been attempted in this work by replacing the silicon nitride layer with a high dielectric constant material. Basically, due to the higher κ value, the equivalent oxide thickness is reduced for the same physical thickness of the film. Hence, the effect on device performance is expected to be similar to that of scaling the tunnel oxide thickness without the disadvantages that come with smaller physical thicknesses, especially increased tunneling current leakage. SOHOS structure with hafnium oxide (HfO_2) as the charge storage layer demonstrated superior charge storage capability at low voltages, faster programming and less over-erase problem as compared to the conventional SONOS device. However, such a SOHOS device had poorer charge retention capability than SONOS. On the other hand, using aluminum oxide (Al_2O_3) as the charge storage layer resulted in a SOHOS structure with improved charge retention performance, but at the expense of a slower programming speed. By adding a small amount of aluminum to HfO_2 to form hafnium aluminum oxide (HfAlO), the resultant SOHOS structure with HfAlO as a charge storage layer can combine the advantages of both HfO_2 and Al_2O_3, such as fast programming

speed, good charge retention and good program/erase endurance. Hence, the programming speed of the SOHOS device was successfully increased without reducing the tunnel oxide thickness through an appropriate choice of the high-κ charge storage layer.

An alternative method to increase program/erase speed without decreasing the tunnel oxide thickness is by using a high-κ material as the blocking oxide. From electrostatics consideration, the use of a high dielectric constant blocking oxide layer will cause a smaller voltage drop across the blocking oxide and greater voltage drop across the tunnel oxide. This will result in a simultaneous increase of the electric field across the tunnel oxide and reduction of the electric field across the blocking oxide, leading to more efficient program and erase processes. The effect of the κ value and band gap energy of the blocking oxide layer on the program/erase speed and charge retention of SONOS devices was investigated by using $(HfO_2)_x(Al_2O_3)_{1-x}$ with different HfO_2 concentration ratios (x) as the blocking oxide. Other high-κ materials with suitable conduction and valence band offsets were also evaluated.

Finally, the integration of high-κ tunnel and blocking oxides and an ultra-high-κ charge storage layer (TiO_2) was also demonstrated in this project. HfAlO/TiO_2/HfAlO SOHOS capacitors showed much greater flatband voltage shift at lower program/erase voltages compared to the conventional SONOS device after post-deposition and forming gas anneals.

List of Tables

		Pages
Table 2.1	Summary of memory parameters for different types of nonvolatile memories	23
Table 4.1	The split conditions of samples with different HfAlO charge storage layer thicknesses, different tunnel oxide thicknesses and 65 Å blocking oxide. The cell structure is similar to Fig. 4. 1.	67
Table 4.2	Comparison between this work (HfAlO device) and published data. SRO is silicon rich oxide.	70
Table 5.1	Comparison between this work and published data. SRO is silicon rich oxide.	89
Table 5.2	Estimated barrier heights of the TaN/$(LaAl)_xY_{1-x}O_3$ interface and conduction band offsets of $(LaAl)_xY_{1-x}O_3$ with respect to silicon.	101

List of Figures

Pages

Figure 1.1: Two classes of nonvolatile semiconductor memory devices: (a) floating-gate device and (b) charge-trapping device (b) (SONOS device). 2

Figure 2.1: Two classes of nonvolatile semiconductor memory devices: (a) floating-gate device and (b) charge-trapping device (MNOS device). 11

Figure 2.2: First operating floating-gate device: the FAMOS (Floating-gate Avalanche injection MOS) device, introduced by Frohman-Bentchkowsky [3-6]. 12

Figure 2.3: The SAMOS (Stacked gate Avalanche injection MOS) device [7-8]. The device is written like the FAMOS device. Several different erasure mechanisms are possible. 13

Figure 2.4: NOR Flash array equivalent circuit [18]. 17

Figure 2.5: A NOR-structured memory array illustrating the over-erase phenomenon. 18

Figure 2.6: Equivalent circuit of the NAND-structured cell array. 19

Figure 2.7: Basic cross section of a Phase Change Memory [19]. 22

Figure 2.8: Evolution of the SONOS NVSM device [24]. 24

Figure 2.9: Physical operation of a SONOS device [25]. 25

Figure 2.10: Energy band diagrams of the programming mechanisms: (a) Direct tunneling, (b) Modified Fowler-Nordheim tunneling, (c) trap assisted tunneling (d) Fowler-Nordheim tunneling [26]. 26

Figure 3.1: Fabricated SONOS-type memory devices with Si_3N_4 or HfO_2 charge storage layers. 42

Figure 3.2: Flatband voltage shift plotted against the charging (positive) and discharging (negative) gate voltage for SONOS, SOHOS1 and SOHOS2 memory devices. 44

Figure 3.3: (a) Program and (b) erase threshold voltage shift of SOHOS1 (with HfO_2 charge storage layer) and SONOS n-channel MOSFETs for V_g - V_{fb} = +6 V during program and V_g - V_{fb} = -5.3V during erase. 45

Figure 3.4: Program/erase (P/E) cycling data for (a) SONOS and

(b) SOHOS1 (with HfO_2 charge storage layer) n-channel MOSFETs. 46

Figure 3.5: X-ray diffraction spectra of SOHOS1 and SOHOS2 structures. 48

Figure 3.6: Ideal energy band diagrams for (a) SONOS and (b) SOHOS structures. 48

Figure 3.7: Energy band diagram schematic of the SONOS structure with HfO_2 (solid lines) or Si_3N_4 (dashed lines) as the charge storage layer during (a) write (program) and (b) erase operations. 50

Figure 3.8: Charge retention performance of the SOHOS1, SOHOS2 and SONOS devices as characterized by the flatband voltage shift at an applied gate bias (V_g) of 0V after the device has been charged at $V_g = 6V$. 50

Figure 4.1: Fabricated SOHOS (with HfO_2 or HfAlO or Al_2O_3 charge storage layer) and SONOS (Si_3N_4) transistor structures with HfN/TaN gate electrode. 55

Figure 4.2: Flatband voltage shift during charge retention measurements versus time of SONOS-type memory devices with Si_3N_4, Al_2O_3, HfO_2 or HfAlO as the charge storage layer during discharging at a gate bias of -1.45 V below the initial flatband voltage of a charged device. The devices were programmed to an initial V_{fb} shift of 1.1 V before the retention measurements. 56

Figure 4.3: The drain current transients of (a), (b) Al_2O_3 memory devices and (c), (d) HfO_2 memory devices during the application of a read voltage after the application of a program voltage for 20s. The read and program voltages for Al_2O_3 devices were 3.3 V and 9 V, respectively. For HfO_2 devices the read and program voltages were 2.9 V and 7 V respectively. 60

Figure 4.4: Drain current difference during discharging divided by squared temperature (T) versus the inverse of T for HfO_2 and Al_2O_3 memory devices. 61

Figure 4.5: Density of stored charge, extracted from the hysteresis in the C-V curves, and plotted against the gate voltage sweep range for SONOS-type capacitor structures with Si_3N_4, Al_2O_3 or HfAlO as the charge storage layer. 61

Figure 4.6: Flatband voltage shift plotted against the charging/discharging (program/erase) voltage extracted from the hysteresis in the

C-V curves for memory capacitors with Si_3N_4, Al_2O_3 or HfAlO as the charge storage layer. 63

Figure 4.7: (a) Programming (V_g-V_{fb} = 6V) and (b) erasing (V_g-V_{fb} = –6V) characteristics of SONOS and SOHOS transistors with Si_3N_4, HfAlO and Al_2O_3 charge storage layers. 63

Figure 4.8: Program/Erase (P/E) endurance characteristics of SONOS and SOHOS transistors with Si_3N_4 and HfAlO charge storage layers. 64

Figure 4.9: Ideal energy band diagrams of SONOS-type structures (HfN/TaN gate) with (a) Si_3N_4 (conventional SONOS), (b) HfAlO (10% Al_2O_3 concentration) and (c) Al_2O_3 as the charge storage layer. 65

Figure 4.10: Energy band diagram schematic of SONOS-type structures with HfAlO (solid lines) or Si_3N_4 (dashed lines) as the charge storage layer during (a) write (program) and (b) erase operations. 65

Figure 4.11: XRD spectra of (a) HfO_2 and (b) HfAlO. As-deposited HfO_2 was already crystallized while HfAlO remained amorphous up to 800°C. 66

Figure 4.12: (a) Programming characteristics (i.e., threshold voltage shift versus time at a tunnel oxide field of 5 MV/cm) of SOHOS transistors with 40 Å, 75 Å and 125 Å thick HfAlO charge storage layer and 27 Å thick tunnel oxide. (b) Threshold voltage shift of SOHOS transistors after 50 s programming versus thickness of the HfAlO charge storage layer for tunnel oxide fields of 5, 6 and 7 MV/cm during programming. The tunnel oxide is 27 Å thick. The traps are saturated after 50s programming. 68

Figure 4.13: Charge retention characteristics (i.e., threshold voltage versus time) of SOHOS transistors with HfAlO charge storage layer of 40 Å, 75 Å and 125 Å thickness and 27 Å tunnel oxide performed at V_g = 0V with source/drain and substrate grounded. 69

Figure 4.14: (a) Programming (V_g - V_{fb} = 8.5V) and (b) erasing (V_g - V_{fb} = -15V) characteristics of threshold voltage shift versus time of SONOS transistor with 27 Å tunnel SiO_2/75 Å Si_3N_4 charge storage layer and SOHOS transistor with 34 Å tunnel SiO_2/75Å HfAlO charge storage layer. V_{th}(t=0) denoted the V_{th} of uncharged device. 69

Figure 4.15: Graph of V_{th} shift after programming at V_g-V_{fb} = 6V, 1ms against the V_{th} decay rate per decade of retention measurement time. Comparison between this work (HfAlO

device) and published data (refer to Table 4.2). 70

Figure 5.1: (a) Fabricated SONOS Flash transistor structures with HfN/TaN gate electrode. The blocking oxide layer is either SiO_2 or high-κ dielectric. (b) Fabricated SONOS Flash transistor structures with TaN gate electrode. The blocking oxide layer is high-κ dielectric. 79

Figure 5.2: XPS spectra for (a) Al 2*p* core levels, (b) O 1*s* core levels and (c) Hf 4*f* core levels taken from $(HfO_2)_x(Al_2O_3)_{1-x}$ samples (used in the blocking oxide layer), with x values determined to be 0.15 and 0.48. 80

Figure 5.3: Programming transient for (a) $V_g - V_{fb} = 6V$ (b) $V_g - V_{fb} = 7V$ and (c) $V_g - V_{fb} = 9V$ for SONOS devices with SiO_2 (solid symbol) or high-κ (open symbols) blocking oxide layers. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a). 81

Figure 5.4: Threshold voltage shift after programming at $V_g - V_{fb} = 6V$, 7V, 8V and 9V for 100 μs for SONOS devices with HfAlO blocking oxide layer with different HfO_2 mole fraction x. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a). 81

Figure 5.5: Schematic energy band diagrams for SONOS devices with Al_2O_3 [(a) and (c)] and HfO_2 [(b) and (d)] blocking oxide layers in the program mode for low [(a) and (b)] and high [(c) and (d)] gate voltage situations. 83

Figure 5.6: Erasing transient for (a) $V_g - V_{fb} = -6V$ (b) $V_g - V_{fb} = -7V$ and (c) $V_g - V_{fb} = -8V$ for SONOS devices with SiO_2 (solid symbol) or high-κ (open symbols) blocking oxide layers. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a). 84

Figure 5.7: Schematic energy band diagrams for SONOS devices in the erase mode: (a) Comparing SiO_2 (solid lines) and high-κ (e.g., Al_2O_3) (dashed lines) blocking oxide layers, and (b) Comparing Al_2O_3 (solid lines) and HfO_2 (dashed lines) blocking oxide layers. 86

Figure 5.8: Program/Erase (P/E) endurance characteristics of SONOS device with $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ (48% HfO_2) blocking oxide. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a). 86

Figure 5.9: (a) Charge retention characteristics of SONOS devices with SiO_2 (solid symbol) or high-κ (open symbols) blocking oxide layers performed at $V_g = 0V$ with source/drain and substrate grounded. The devices were programmed to an initial V_{th} shift of 1.25V before the retention measurements. (b) The same result as in (a) but with the time scale plotted up to 10^9 seconds. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a). 87

Figure 5.10: Schematic energy band diagrams for SONOS devices with Al_2O_3 (solid lines) and HfO_2 (dashed lines) blocking oxide layer during charge retention measurement. 88

Figure 5.11: Graph of V_{th} shift after programming at V_g-V_{fb} = 6V, 100μs against the V_{th} decay rate per decade of retention measurement time. Comparison between this work and published data (refer to Table 5.1). 89

Figure 5.12: Programming transient for (a) V_g - V_{fb} = 9V (b) V_g - V_{fb} = 11V and (c) V_g - V_{fb} = 13.5V for SONOS devices with HfO_2, $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ or Al_2O_3 blocking oxide layers. The gate stacks of the SONOS devices are 40 Å SiO_2/ 70 Å Si_3N_4/ 120 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (b). 91

Figure 5.13: Erasing transient at V_g - V_{fb} = -12.5V for SONOS devices with HfO_2, $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ or Al_2O_3 blocking oxide layers. The gate stacks of the SONOS devices are 40 Å SiO_2/ 70 Å Si_3N_4/ 120 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (b). 92

Figure 5.14: Charge retention characteristics of SONOS devices with HfO_2, HfAlO or Al_2O_3 blocking oxide layers performed at $V_g = 0V$ and source/drain and substrate grounded. The devices were programmed to an initial V_{th} shift of 2.9V before retention measurements. 92

Figure 5.15: Fabricated SONOS structures with TaN gate electrode. The blocking oxide layer is $(La_2O_3)_x(Al_2O_3)_{1-x}$ with different composition ratios. 94

Figure 5.16: Fabricated $(LaAl)_xY_{1-x}O_3$ capacitor structures with TaN gate electrode. 95

Figure 5.17: High-Frequency Capacitance-Voltage (HFCV) measurements of SONOS capacitors with $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide. The capacitors have dimensions of 200 μm × 200 μm. 97

Figure 5.18: Gate current density versus gate voltage (J_g-V_g) measurements of SONOS capacitors with $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide. The capacitors have dimensions of 200 μm × 200 μm. 97

Figure 5.19: High-frequency capacitance-voltage (HFCV) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions. The capacitors have dimensions of 200 μm × 200 μm. 98

Figure 5.20: Gate-current versus gate voltage (J_g-V_g) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions. The capacitors have dimensions of 200 μm × 200 μm. 99

Figure 5.21: Gate-current density at gate voltage of 3V above the flatband voltage against EOT of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions. The capacitors have dimensions of 200 μm × 200 μm. 99

Figure 5.22: $d(\ln J)/dV$ plotted against V_g for TaN /$(LaAl)_xY_{1-x}O_3$ /n-Si devices. 100

Figure 5.23: High-Frequency Capacitance-Voltage (HFCV) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions after 900°C, 60s, N_2 anneal. The capacitors have dimensions of 200 μm × 200 μm. 102

Figure 5.24: Gate current versus gate voltage (J_g-V_g) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions after 900°C, 60s, N_2 anneal. The capacitors have dimensions of 200 μm × 200 μm. 102

Figure 6.1: TEM micrograph of TiN film on SiO_2 underlayer after 850°C, 10 s anneal in vacuum. EDX analysis revealed formation of $TiSi_2$ after annealing [6]. 110

Figure 6.2: TEM micrograph of 4nm SiO_2/17nm TiO_2 layers after forming gas annealing at 420°C for 30 minutes. 110

Figure 6.3: Fabricated SiO_2/TiO_2 capacitor structures with TaN gate electrode. 112

Figure 6.4: (a) Device structure of fabricated HfAlO/TiO_2/HfAlO capacitor structures with TaN gate electrode (b) Device structure of fabricated HfAlO/AlN/TiO_2/AlN/HfAlO capacitor structures with TaN gate electrode. (c) Ideal energy band diagram of HfAlO/TiO_2/HfAlO capacitor. 112

Figure 6.5: (a), (c) and (e) HFCV and (b), (d) and (f) J_g-V_g graphs of

SiO_2/TiO_2 capacitors; (a) and (b) after forming gas anneal only, (c) and (d) after 700°C, 30 s, O_2 PDA and (e) and (f) after 950°C, 30 s, N_2 anneal. The devices have gate areas of 200 μm × 200 μm. 114

Figure 6.6: (a) and (c) HFCV and (b) and (d) J_g-V_g graphs of HfAlO/TiO_2/HfAlO capacitors; (a) and (b) after 700°C, 30 s, O_2 PDA of the TiO_2 layer and (c) and (d) after 900°C, 30 s, N_2 anneal. The devices have gate areas of 200 μm × 200 μm. 116

Figure 6.7: TEM micrograph of HfAlO/TiO_2/HfAlO capacitors after 900°C N_2 anneal for 30s. 116

Figure 6.8 C-V curves of HfAlO/TiO_2/HfAlO memory capacitors after PDA at 700°C for 30s in O_2 showing counter-clockwise hysteresis for various gate voltage (V_g) sweep ranges as indicated. The capacitance was measured at 100 kHz, with a gate voltage sweep rate of 0.1 V/s. Gate area is 200 μm × 200 μm. 117

Figure 6.9: Flatband voltage shift extracted from the hysteresis C-V curves plotted against the charging (positive) and discharging (negative) gate voltage for 60 Å HfAlO/60 Å TiO_2/120 Å fAlO and 25 Å SiO_2/60 Å Si_3N_4/60 Å SiO_2 memory devices. Gate area is 200 μm × 200 μm. 117

Figure 6.10: Charge retention characteristics of HfAlO/TiO_2/HfAlO memory devices measured with V_g = 0V. The devices were programmed to a V_{fb} shift of 2.7V before retention measurement. 118

Figure 6.11: (a), (c) and (e) are HFCV while (b), (d) and (f) are J_g-V_g graphs of HfAlO/AlN/TiO_2/AlN/HfAlO capacitors; (a) and (b) with only 600°C, 30s, O_2 PDA of the TiO_2 layer, (c) and (d) after 800°C, 30 s, N_2 anneal while (e) and (f) after 900°C, 30 s, N_2 anneal. The devices have gate areas of 200 μm × 200 μm. 119

Figure 6.12: Retention characteristics of HfAlO/AlN/TiO_2/AlN/HfAlO memory devices measured with V_g = 0V. The devices were programmed to a V_{fb} shift of 2.6V before retention measurement. 120

Figure 6.13: Flatband voltage shift extracted from the hysteresis C-V curves plotted against the charging (positive) and discharging (negative) gate voltage for 60 Å HfAlO/60 Å TiO_2/ 120 Å HfAlO and 60 Å HfAlO/20 Å AlN/60 Å TiO_2/20 Å AlN/120 Å HfAlO memory devices. Gate area is 200 μm × 200 μm. 121

Figure 7.1: Conduction band edge diagrams of various tunnel barriers: (a) a typical uniform barrier; (b) idealized crested symmetric barrier; (c) crested, symmetric layered barrier. U is the maximum barrier height, expressed in units of energy. 128

List of Symbols

Å	10^{-10} m
MOS	Metal oxide semiconductor
SONOS	Polysilicon-oxide-silicon nitride-oxide-silicon
EEPROM	Electrically-erasable-programmable-read-only-memory
ONO	Oxide-nitride-oxide
W/L	Transistor gate width to gate length dimensions
C-V	Capacitance-Voltage
V_{FB}	Flatband voltage
P/E	Program/erase
PDA	Post-deposition-anneal
XRD	X-ray diffraction
V_{th}	Threshold voltage
F-N	Fowler-Nordheim
I_D	Transistor drain current
μ	Mobility of charge carrier
C_{ox}	Gate oxide capacitance
V_G	Gate voltage
V_D	Transistor drain voltage
γ	Transistor body effect parameter
ϕ_s	surface potential
Q_G	Charge at the gate
Q_{ot}	Oxide trapped charge
Q_s	Charge in silicon

E_{trap}	Energy level of charge trap
T	Temperature (in Kelvin)
k	Boltzmann constant
J_g	Gate current density

Chapter 1

Introduction

1.1 Background

Since the very first days of the mid-1960s, when the potential of metal-oxide-semiconductor (MOS) technology to realize semiconductor memories with superior density and performance than would ever be achievable with the then commonly used magnetic core memories became known, chip makers have thought of solutions to overcome the main drawback of the MOS memory concept, that is, its intrinsic volatility. The first sound solutions to this problem were the floating gate concept [1] and the metal-nitride-oxide-semiconductor (MNOS) memory device [2] both of which were proposed in 1967. Tremendous progress has been made over the years in realizing the idea of a reliable, high-density reprogrammable read-only-memory (ROM) memory.

New applications and lower memory costs have driven increases in memory chip sales. Flash memory chips permitted cellular phones, audio internet players and digital cameras to be manufactured at a price that is affordable for consumers. The term Flash refers to the fact that the contents of the whole memory array, or of a memory block (sector), is erased in a single step. Low power and high-density dynamic random access memory (DRAM) chips permitted the personal digital assistant to meet low-power battery requirements and to have the capability of performing tasks that were once the domain of desktop personal computers (PCs). Advances in semiconductor lithography will continue to result in increased data storage density and lower costs per unit megabyte of storage. New nonvolatile

memory technologies such as ferroelectric, polymer and magnetoresistive memories will promote new applications for nonvolatile memory and will allow nonvolatile memory to replace volatile memory in PCs, network equipment and cellular phone applications.

The basic operating principle of nonvolatile semiconductor memory devices is the storage of charges in the gate stack structure of a MOS field effect transistor (MOSFET). The charge storage can be realized in two ways, which has led to the subdivision of nonvolatile semiconductor memory devices into two main classes. The first class of devices is based on the storage of charge on a conducting or semiconducting layer that is completely surrounded by a dielectric, usually silicon dioxide (SiO_2), as shown in Fig. 1.1(a). Since this layer acts as a completely electrically isolated gate, this type of device is commonly referred to as a floating-gate device [1]. In the second class of devices, the charge is stored in discrete trapping centers of an appropriate dielectric layer. These devices are, therefore, usually referred to as charge trapping devices. The most successful devices in this category are the MNOS (metal-nitride-oxide-silicon) and SONOS (silicon-oxide-nitride-oxide-silicon) or MONOS (metal-oxide-nitride-oxide-silicon) structures, in which the charge storage layer is a silicon nitride layer on top of a very thin silicon oxide layer. Figure 1.1(b) illustrates the SONOS structure.

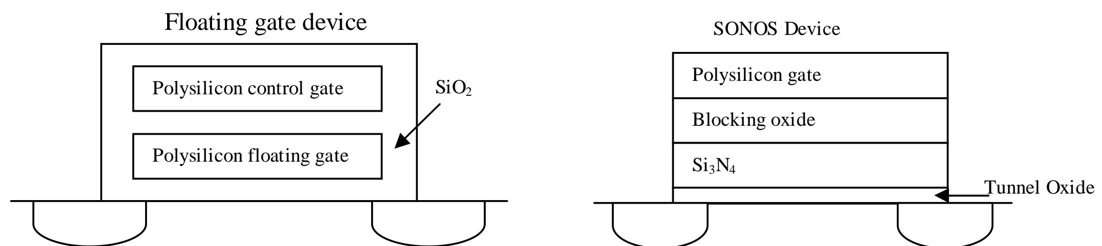


Figure 1.1: Two classes of nonvolatile semiconductor memory devices: (a) floating-gate device and (b) charge-trapping device (SONOS device).

The Semiconductor Industry Association (SIA) International Technology Roadmap for Semiconductors (ITRS) [3] states that the difficult challenge, beyond the year 2005, for nonvolatile semiconductor memories is to achieve reliable, low-power, low-voltage performance.This challenge is formidable since memory program and erase operations are incompatible with aggresively scaled low-voltage devices. The ITRS projection is based on the continued scaling of polysilicon floating-gate nonvolatile semiconductor memory (NVSM) devices, which employ tunnel oxides with thicknesses greater than 7 nm and with concomitant program/erase electric fields in excess of 6 MV/cm [6]. The net result is the need for high-voltage generator charge-pump circuits.

Currently, most Flash electrically erasable and programmable read only memories (EEPROMs) are based on floating-gate devices [4]. However, the floating-gate memory has limitations with respect to scaling the cell size and program/erase voltages [5]. The relatively thick (7-12 nm) tunnel oxide in floating-gate type memories provides good 10-year data retention. However, the high voltage requirement [5] has created a reliability issue, as it has exceeded the voltage limits of the scaled peripheral complementary MOS (CMOS) devices. The concern over the loss of the entire memory charge through a single defect in the tunnel oxide limits vertical scaling of the tunnel oxide thickness [5]. The demand for low power and low voltage electronics has accelerated the pace for NVSM circuit designers to consider SONOS for low voltage, high density EEPROMs. The motivation for the interest in SONOS lies in low programming voltages, endurance to extended write/erase cycling, resistance to radiation and compatibility with high density scaled CMOS technology. As the charges are stored in discrete traps in the insulating charge storage layer for the

SONOS device structure, a single defect in the tunnel oxide will not result in the loss of the entire memory charge.

1.2 Motivation for the Project

Applications for portable data equipment are becoming widespread, and in this field the nonvolatile memory is generating particularly strong interest. Pre-eminent among applications of nonvolatile memory are Flash memory cell structures. The Flash memory is a type of nonvolatile memory based on block erasure of electrically rewriteable EEPROM. Because it has achieved low cost and high integration, this type of memory is being put to a wide range of uses. Currently, most Flash EEPROMs are based on floating-gate devices [4]. However, the floating-gate memory has limitations with respect to scaling the cell size and program/erase voltages [5]. The demand for low power and low voltage electronics has accelerated the pace for NVSM circuit designers to consider SONOS for low voltage, high density EEPROMs. The floating-gate Flash EEPROM is a slow write/erase device because of low tunneling currents in the oxide [6]. Hence, the floating gate NVSM is limited to a rather low number (e.g., 10^5) write/erase cycles due to a low charge-to-breakdown, Q_{BD}, of its relatively thick tunnel oxide. In contrast, an ultra-thin tunnel oxide can conduct a high current for a dramatic increase in the Q_{BD} [6], leading to an improvement in NVSM reliability for scaled SONOS devices. In addition, the better scaling perspective, together with easier integration in a base line CMOS process, makes SONOS an excellent candidate for embedded Flash in the 90 nm technology node and beyond [7]. For example, the embedded SONOS NVSM requires only four additional noncritical masking steps over the base logic process, compared to eleven additional masking steps for the embedded floating-gate NVSM. Hence, SONOS

requires lower production cost. This makes SONOS memory as one of the most attractive candidates to realize Flash vertical scaling.

Increase in programming speed of SONOS devices and lower voltage operation had been accomplished previously by reducing the tunnel oxide thickness [8], [9]. However, this seriously degrades the charge retention capability of the device. To overcome this limitation, the SOHOS (polysilicon-oxide-high-κ-oxide-silicon) Flash memory has been attempted by replacing the silicon nitride layer with a high dielectric constant (high-κ) material. Basically, due to the higher κ value, the equivalent oxide thickness is reduced for the same film physical thickness. Hence, the effect on device performance is expected to be similar to that of tunnel oxide scaling without the disadvantages that come with smaller physical thicknesses.

An alternative method to increase program/erase speed without decreasing the tunnel oxide thickness is by using a high-κ material as the blocking oxide [10-13]. From electrostatics consideration, the use of a high dielectric constant blocking oxide layer will cause a smaller voltage drop across the blocking oxide and greater voltage drop across the tunnel oxide. This will result in a simultaneous increase of the electric field across the tunnel oxide and reduction of the electric field across the blocking oxide leading to more efficient program and erase processes [10-13]. The effect of the κ (dielectric constant) value and band gap energy of the blocking oxide layer on the program/erase speed and charge retention of SONOS devices is also investigated.

1.3 Research Objectives

The objective of this project is to find innovative solutions, using high dielectric constant materials in the SONOS memory structure, to overcome the limitations of conventional floating-gate NVSM as a result of fast shrinking device geometries.

SONOS type memory devices with suitable high-κ charge storage layers to replace Si_3N_4 (SOHOS structure) will be fabricated and characterized. Different types of high-κ materials with different band gaps, valence and conduction band offsets with respect to silicon, κ-value, crystallization temperature and other material properties will be evaluated. By using materials with higher dielectric constant compared to Si_3N_4 will result in lower program/erase voltages due to higher tunnel oxide coupling ratio. In addition, by using materials with suitable band gap and valence and conduction band offsets, with respect to silicon, may reduce hole tunneling and over-erase effects.

In addition, the use of high-κ blocking oxide in the SONOS memory device will be evaluated. The effect of the κ value and band gap energy of the blocking oxide layer on the program/erase speed and charge retention of SONOS devices is investigated by using hafnium aluminium oxide, or $(HfO_2)_x(Al_2O_3)_{1-x}$, with different concentration ratios (x) as the blocking oxide. Other high-κ materials with suitable conduction and valence band offsets will also evaluated.

Finally, the integration of high-κ tunnel and blocking oxides and ultra-high-κ charge storage layer will also be demonstrated in this project.

1.4 Organization of Thesis

Chapter 2 reports the key findings in the literature on SONOS memory devices with an emphasis on the use of high-κ material in the SONOS memory structure.

Chapter 3 investigates the use of a hafnium oxide (HfO_2) high-κ charge storage layer in SONOS memory devices in order to increase the programming speed without reducing the tunnel oxide thickness. By using HfO_2 instead of Si_3N_4 in the SONOS device structures, faster programming speed and over-erase reduction are achieved.

Chapter 4 presents the results on SOHOS devices using hafnium aluminum oxide (HfAlO) as the charge storage layer. The SOHOS structure, with HfO_2 as the charge storage layer, demonstrates faster programming and less over-erase problem as compared to the conventional SONOS device using Si_3N_4 as the charge storage layer. However, such a SOHOS device has poorer charge retention capability than SONOS and also poor program/erase endurance. On the other hand, using aluminum oxide (Al_2O_3) as the charge storage layer results in a SOHOS structure with improved charge retention performance, but at the expense of a slower programming speed. By adding a small amount of aluminum to HfO_2 to form HfAlO, it will be demonstrated that the resultant SOHOS structure with HfAlO as the charge storage layer can combine the advantages of both HfO_2 and Al_2O_3, such as fast programming speed, good charge retention capability and good program/erase endurance.

Chapter 5 investigates the use of a high-κ blocking oxide in SONOS memory devices. The effect of the κ (dielectric constant) value and band gap energy of the blocking oxide layer on the program/erase speed and charge retention of SONOS devices is investigated by using $(HfO_2)_x(Al_2O_3)_{1-x}$ with different HfO_2 concentration ratios (x) as the blocking oxide. Other high-κ materials with suitable conduction and valence band offsets are also evaluated.

Finally, the integration of high-κ tunnel and blocking oxides and an ultra-high-κ titanium dioxide (TiO_2) charge storage layer into a SONOS memory structure is discussed in chapter 6. HfAlO/TiO_2/HfAlO SOHOS capacitors showed much greater flatband voltage shift at lower program/erase voltages compared to the conventional SONOS device after post-deposition and forming gas anneal. Chapter 7 summarizes and concludes the work presented in this thesis.

References

[1] D. Kahng and S. M. Sze, "A floating gate and its application to memory devices", *Bell Syst. Tech. J.*, vol. 46, pp. 1288-1292, 1967.

[2] H. A. R. Wegener, A. J. Lincoln, H. C. Pao, M. R. O'Connell and R. E. Oleksiak, "The variable threshold transistor, a new electrically alterable non-destructive read-only storage device", in *IEEE IEDM Tech. Dig.*, 1967, p. 70.

[3] The International Technology Roadmap for Semiconductors (ITRS) 2005-Emerging Research Devices.

[4] B. Kim, S. B. Yi and K. Y. Seo, "A scaled SONOS single transistor memory cell for a high density NOR structure with common source lines", *Journal of the Korean Physical Society*, vol. 41, pp. 945-948, 2002.

[5] B. Jiankang and M. H. White, "Design considerations in scaled SONOS nonvolatile memory devices", *Solid State Electronics*, vol. 45, pp. 113-120, 2001.

[6] M. H. White, D. A. Adams and B. Jiankang, "On the go with SONOS", *IEEE Circuits and Devices Magazine*, pp. 22-31, 2000.

[7] C. T. Swift, G. L. Chindalore, K. Harber, T. S. Harp, A. Hoefler, C. M. Hong, P. A. Ingersoll, C. B. Li, E. J. Prinz and J. A. Yater, "An embedded 90 nm SONOS nonvolatile memory utilizing hot electron programming and uniform tunnel erase", in *IEEE IEDM Tech. Dig.*, 2002, p. 927.

[8] M. L. French and M. H. White, "Scaling of multidielectric nonvolatile SONOS memory structures", *Solid State Electronics*, vol. 37, pp. 1913-1923, 1994.

[9] M. L. French, C. Y. Chen, H. Sathianathan and M. H. White, "Design and scaling of a SONOS multidielectric device for nonvolatile memory

applications”, *IEEE Trans. Components, Packaging and Manufacturing Technology.*- Part A, vol. 17, pp. 390-397, 1994.

[10] V.A. Gritsenko, “Design of SONOS memory transistor for Terabit scale EEPROM”, in *IEEE Conf. Electron Device and Solid State Circuits,* 2003, p. 345.

[11] C. H. Lee, K. I. Choi, M. K. Cho, Y. H. Song, K. C. Park and K. Kim, “A novel SONOS structure of SiO_2/SiN/Al_2O_3 with TaN metal gate for multi-gigabit flash memories”, in *IEDM Tech. Dig*., 2003, p. 613.

[12] S. Choi, M. Cho, H. Hwang and J. W. Kim, “Improved metal-oxide-nitride-oxide-silicon-type flash device with high-κ dielectrics for blocking layer”, *J. Appl. Phys.*, vol. 94, pp. 5408-5410, 2003.

[13] C. H. Lee, S. H. Hur, Y. C. Shin, J. H. Choi, D. G. Park and K. Kim, “Charge-trapping device structure of SiO_2/SiN/High-κ dielectric Al_2O_3 for high-density flash memory”, *Appl. Phys. Lett*., vol. 86, pp. 152908 (1-3), 2005.

Chapter 2

Literature Review

2.1 History of Nonvolatile Memory Structures

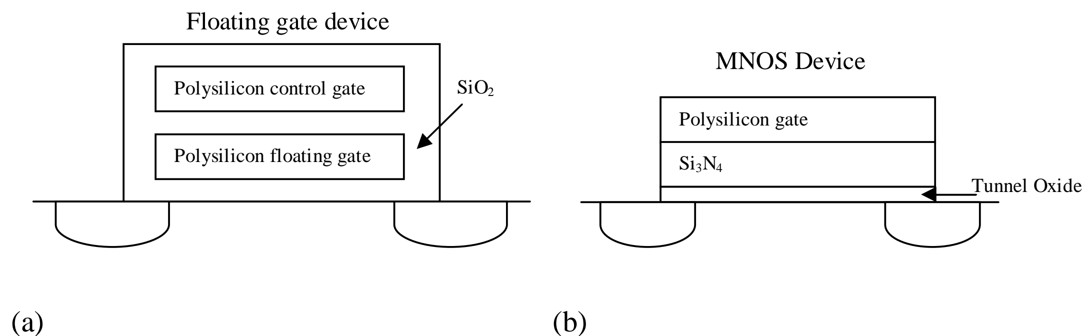
The first nonvolatile metal-oxide-semiconductor (MOS) memory device was introduced in 1967 by D. Kahng and S. M. Sze [1]. Their idea was to use a floating-gate device to store charges. The memory transistor that they proposed started from a basic MOS structure where the gate structure is replaced by a layered structure of a thin oxide, a floating but conducting metal layer, a thick oxide and an external metal gate, as shown in Fig. 2.1. This device is referred to as the MIMIS (metal-insulator-metal-insulator-semiconductor) cell. Electrons were injected into the floating-gate by direct tunneling during programming. To discharge the floating-gate, a negative voltage pulse is applied to the metal gate, removing the electrons by the same direct tunneling mechanism.

The tunnel oxide thickness is limited to less than 5 nm due to the direct tunneling programming mechanism. Hence, any defects in the tunnel oxide will cause all the stored charges in the floating-gate to leak off. Due to technological constraints, the MIMIS cell could not be reliably built at that time. However, the introduction of this device contained several important concepts that have led to the development of both classes of nonvolatile memory devices. The direct tunneling concept has been used in charge trapping devices while the floating-gate concept has led to a whole range of floating-gate memory types.

In order to solve the technological constraint of the MIMIS cell, two approaches are possible: (1) replacing the conducting charge trapping layer with an

insulating one, or (2) increasing the tunnel dielectric thickness and employing other charge injection mechanisms.

The first solution was used in the MNOS (metal-nitride-oxide-semiconductor) cell (Fig. 2.1 (b)), introduced by Wegener *et al.* [2], almost simultaneously with the MIMIS cell. In the MNOS cell, the polysilicon floating-gate is replaced by a nitride layer, which contains numerous electron and hole trapping centers. As the charge storage layer is an insulator, any defects in the tunnel oxide will not cause all the stored charges to leak out.



(a) (b)

Figure 2.1: Two classes of nonvolatile semiconductor memory devices: (a) floating-gate device and (b) charge-trapping device (MNOS device).

The second solution has been used in a wide range of nonvolatile memory devices. The first operating floating-gate device, shown in Fig. 2.2, was introduced in 1971 by Frohman-Bentchkowsky and is known as the Floating-gate Avalanche injection MOS (FAMOS) device [3-6]. In the original p-channel FAMOS cell, the floating-gate is completely surrounded by a thick (~ 100 nm) oxide. Hence the problem of possible shorting paths is reduced. In the FAMOS cell, programming is performed by charge transport to the floating-gate by avalanche injection of electrons from a reverse biased p-n junction. However, no mechanism for electrical erasure exists due to the lack of an external gate. Hence, erasure was done using ultraviolet

(UV) irradiation. The FAMOS device has found wide applications and was the first cell to reach volume manufacturing levels comparable to other semiconductor memory types. FAMOS devices have evolved into a class of memory products called EPROM (electrically-programmable-read-only-memory).

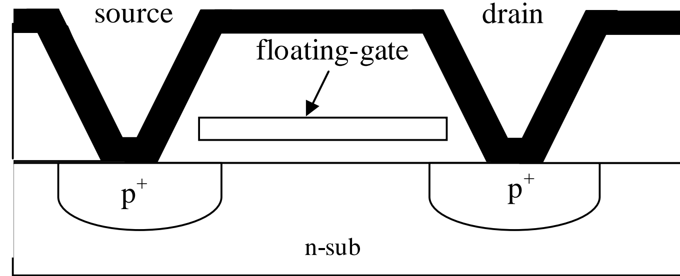


Figure 2.2: First operating floating-gate device: the FAMOS (Floating-gate Avalanche injection MOS) device, introduced by Frohman-Bentchkowsky [3-6].

The drawbacks of the FAMOS device were alleviated in several adapted concepts. In the Stacked gate Avalanche injection MOS (SAMOS) [7-8], as shown in Fig. 2.3, an external control gate is added to improve the writing efficiency by an increased drift velocity of the electrons in the oxide and a field-induced energy barrier lowering at the silicon-silicon dioxide (Si-SiO_2) interface. Electrical erasure also became possible by field emission through the top dielectric due to polyoxide conduction. Consequently, electrically-erasable-programmable-read-only-memory (EEPROM) products became feasible.

These first floating-gate memory devices were all p-channel devices. In n-channel devices, drain avalanche results in hole injection, which is much less efficient due to the higher energy barrier experienced by the holes. Hence, for n-channel devices, several alternative injection mechanisms were proposed and used for

floating-gate applications. These include Fowler-Nordheim (F-N) tunneling through thin oxides (<12 nm) and channel hot-electron (CHE) injection [9].

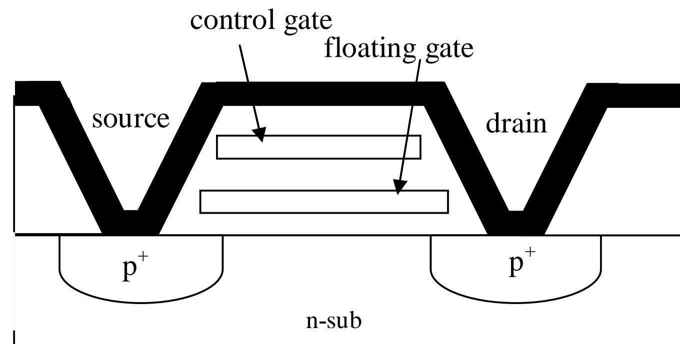


Figure 2.3: The SAMOS (Stacked gate Avalanche injection MOS) device [7-8]. The device is written like the FAMOS device. Several different erasure mechanisms are possible.

F-N tunneling is a field-assisted electron tunneling mechanism. At high electric fields, electrons in the silicon conduction band will see a triangular energy barrier with a width dependent on the applied field. Electrons in the silicon conduction band can tunnel through the triangular energy barrier giving rise to F-N current.

At large drain biases, the minority carriers that flow in the channel of a MOS transistor are heated by the large electric fields seen at the drain side of the channel and their energy distribution is shifted higher. These electrons can collide with the silicon lattice atoms near the drain and generate minority and majority carriers through impact ionization. The majority carriers are normally collected at the substrate contact and form the substrate current. The minority carriers are collected at the drain. Some of these carriers gain enough energy to surmount the Si-SiO_2 energy barrier. If the oxide field favours injection, these carriers are injected over the barrier into the gate insulator and give rise to the so-called hot carrier injection gate current.

The first nonvolatile memory product that can be electrically programmed by the user and erased afterwards is the EPROM device, introduced in the 1970s. Programming can be carried out by channel hot-electron injection while UV light is used to erase the memory. The EPROM cell can consist of a single transistor as it does not need addressing down to the byte level during an erase operation.

Since UV light is used for erasure, a quartz window has to be provided in the EPROM package, which makes this package quite expensive. Reprogramming of the device is also not user friendly. The circuit has to be taken off the circuit board for erasing. The erase operation takes about 20 minutes, and then the whole memory has to be reprogrammed byte by byte. This rather tedious procedure must be performed even if the content of a single byte has to be changed. These drawbacks have been obviated in the EEPROM. As both programming and erasing are controlled by electrical signals, the circuit can be reprogrammed while residing on the circuit board. Each operation, including erasing, can be performed in a byte-addressable way. The EEPROM cell consists of a memory transistor and a select transistor [10], thus leading to the so-called two transistor memory cell. However, the large area requirements and the relatively high operating voltages (15 to 20 V) due to the thick (8 to 10 nm) tunnel oxide limits further scaling down of the EEPROM cell [11]. Charge-trapping, as well as floating-gate devices, are used for EEPROM products.

During the 1980s, a novel nonvolatile memory product was introduced; referred to as the Flash EEPROM [12]. The general idea was to combine the fast programming capability and high density of EPROMs with the electrical erasibility of EEPROMs. The first products were merely the result of adapting EPROMs in such a way that the cell could be erased electrically. Consequently, these devices use channel hot-electron injection for programming and F-N tunneling for erasure. The memory

can be erased electrically but not selectively. The content of the whole memory chip is always cleared in one step. The advantages over the EPROM are the faster (electrical) erasure and the in-circuit reprogrammability, which leads to a cheaper package. Although Flash EEPROM has a higher density compared to traditional EEPROM, many bytes are erased simultaneously instead of a single byte at a time. The Flash memory technology has been a dominant technology for the past two decades.

Other forms of nonvolatile memory technologies that have evolved in the past few decades include the MNOS, SONOS (polysilicon-oxide-nitride-oxide-silicon) and ferroelectric devices. The MNOS devices were invented in 1967 [2] and were the first electrically alterable read only memory (EAROM) devices. The nonvolatile function of these devices is based on the storage of charges in discrete traps in the nitride layer. These charges (electrons or holes) are injected from the channel region into the nitride by quantum mechanical tunneling through an ultra-thin oxide (typically 1.5 to 3nm).

Hole injection from the gate limits the memory window in MNOS devices. The problem becomes more severe for thinner nitride layers. An efficient way to solve this problem is by introducing a top blocking oxide layer in between the silicon nitride and the gate electrode resulting in the SONOS memory structure [13]. The aim of the top oxide is not only to inhibit gate injection, but also to block the charges injected from the silicon substrate at the top oxide-nitride interface. This results in higher trapping efficiency. In this way, the total thickness of the insulator structure can be reduced, and consequently, the programming voltage can be reduced.

Ferroelectric memory devices store information based on polarization state rather than stored charge [14]. Certain crystalline materials show the tendency to polarize spontaneously under the influence of an external field and to remain

polarized after the external field is removed. The polarization can simply be reversed by applying a field of opposite polarity. The ferroelectric material used is a lead-zirconate-titanate compound (Pb[Zr, Ti]O_3, PZT), which is a perovskite-type ceramic. These memories have fast write time (~100 ns) and good endurance [15]. However, the main drawback is the problem of incorporating these materials to mainstream silicon technology [16].

2.2 Current and Future Nonvolatile Memories

The present baseline for non-volatile memory technology is based on both NOR and NAND Flash employing the floating-gate structure [17]. The current Flash technology node, based on the polysilicon half pitch, is at 70 nm for NOR Flash and 64 nm for NAND Flash. According to the International Technology Roadmap for Semiconductors, the difficult challenge for Flash scaling to 32 nm technology and beyond is the non-scalability of the tunnel and interpoly dielectrics [17].

In the coming years, portable systems will demand even more nonvolatile memories, either with high density and very high writing throughput for data storage application or with fast random access for code execution. Although in the past, different types of Flash cells and architectures have been proposed, two of them can be considered as industry standard today. These are the common ground NOR Flash due to its versatility in addressing both the code and data storage segments, and the NAND Flash which is optimized for the data storage market. In code storage, the program or operating system is stored in the Flash memory (usually NOR structure) and is executed by the microprocessor or microcontroller [18]. NOR chips function like a computer's main memory, while NAND works like a hard disk. For example, in

a digital camera, NOR Flash contains the camera's internal software, while NAND Flash is used to store the images.

The "NOR" Flash name is related to the way the cells are arranged in an array, through rows and columns in a NOR-like structure as shown in Fig. 2.4. Flash cells sharing the same gate constitute the so-called word line (WL), while those sharing the same drain electrode (one contact common to two cells) constitute the bit line (BL). In this array organization, the source electrode is common to all of the cells (Fig. 2.4). A NOR Flash memory cell is usually programmed by channel hot electron injection into the floating gate at the drain side and it is erased by means of Fowler-Nordheim electron tunneling through the tunnel oxide from the charge storage layer to the silicon surface.

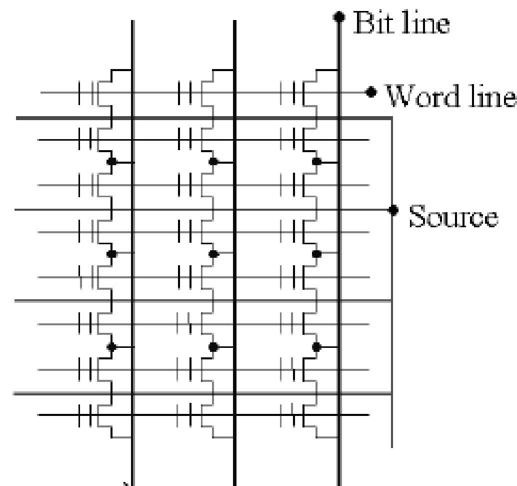


Figure 2.4: NOR Flash array equivalent circuit [18].

In the NOR array, threshold voltage after both program and erase operations are maintained above 0V. The threshold voltage distribution widths are tightly controlled by uniformity in currents and parameters. If one of the memory cells has an erased threshold voltage that is too low, or even negative (over-erase), it will cause

excessive bit line leakage and read failure, as illustrated in Fig. 2.5. During the read operation, positive read voltages are applied to the selected word and bit lines. The unselected bit lines and word lines are floated and grounded, respectively. If the selected memory cell has a high positive threshold voltage (a written cell), current does not flow through the bit line. However, if any of the other memory cells sharing the same bit line has a negative threshold voltage, current will flow through the bit line causing a read failure.

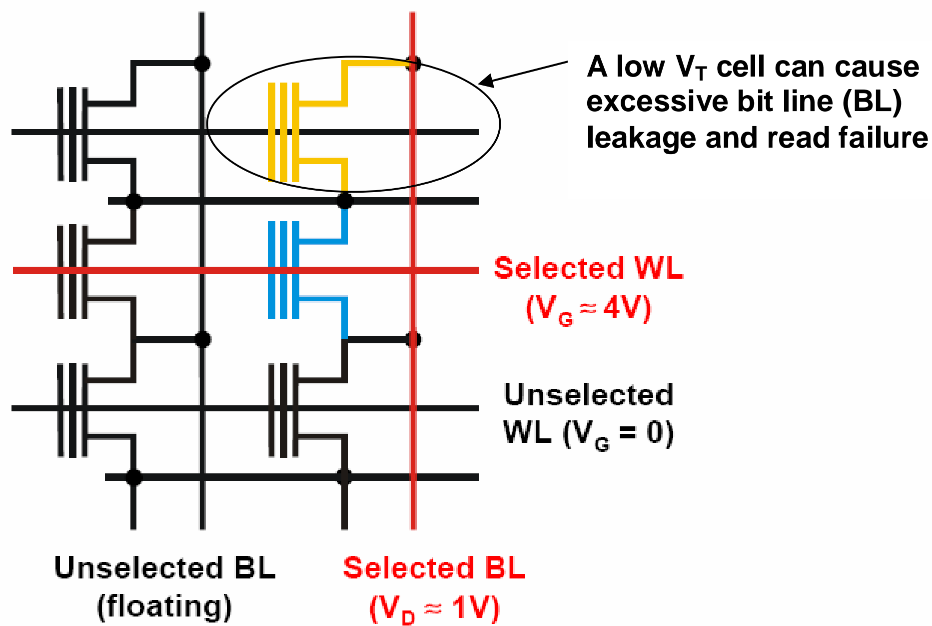


Figure 2.5: A NOR-structured memory array illustrating the over-erase phenomenon.

In the NOR structure, the memory cells are connected to a bit line in a parallel manner. The NAND structure reduces the cell size by connecting the cells in series between a bit line and a source-line, thus reducing the number of bit and source line contact holes [9]. The resulting cell structure occupies 85% of the area of a NOR cell stacked gate array and is easier to scale down. Figure 2.6 shows the equivalent circuit of the NAND-structured cell. As shown in the figure, the NAND-structured cell

arranges eight memory transistors in series, sandwiched between two select gates, select gate 1 (SG1) and select gate 2 (SG2). The first gate (SG1) ensures selectivity, and the second (SG2) prevents the cell current from passing during a programming operation. Program and erase are usually carried out by Fowler-Nordheim tunneling through the tunnel oxide. The reading speed of the NAND structure is slower than that of the NOR-structured array as a number of memory cells are connected in series.

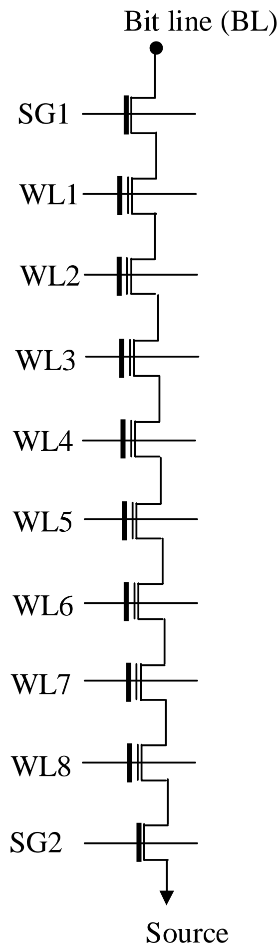


Figure 2.6: Equivalent circuit of the NAND-structured cell array.

In the NAND-structured array, erased cells ("0") have negative threshold voltages while programmed cells ("1") have positive threshold voltages. During the read operation, 0V is applied to the gate of the selected memory cell, while a positive

read voltage is applied to the gate of the other cells. Therefore, all of the other memory transistors serve as transfer gates. As a result, in the case when the selected transistor has a negative threshold voltage (“0”), the memory transistor is in depletion mode and current flows. On the other hand, current does not flow when the selected memory transistor has a high positive threshold voltage (“1”) as it is in the enhancement mode. The state of the cell is detected by a sense amplifier that is connected to the bit line. Due to the different read operation, over-erase is not an issue in the NAND-structured array.

Due to the scaling issues of the baseline nonvolatile memory, future replacements for the floating-gate structure are actively investigated. These include research on new materials and mechanisms in Phase Change Memory, Magnetic Random Access Memory (MRAM), Ferroelectric Random Access Memory (FeRAM) and SONOS memory.

The Phase Change Memory (PCM) is being studied as a candidate for next generation nonvolatile memory technology [19]. PCM consists of a transistor to supply the drive current and a phase change resistor made of a chalcogenide material. The basic phase change material is of the same family of materials used in optical re-writable CD/DVD RW disks (e.g., GeSbTe). In the phase change memory (Fig. 2.7), electric current of different magnitudes are passed from a heater element to the chalcogenide material and local joule heating is used to change the programmable volume around the contact region. Higher current and fast quenching freeze the material to an amorphous state giving high resistance ($> 40 \times$) compared to the lower resistance crystalline state. The time required for switching to an amorphous state is typically less than 10-30 ns. Medium current for longer pulse time is used to re-crystallize the region to a crystalline state, which has low resistance. A much lower

current with essentially no joule heating is used for reading the memory, differentiating between the high (amorphous) and low (crystalline) resistance states [19]. The basic memory cell shows fast programming capability of < 30 ns, good endurance characteristics of up to 10^{12} write/erase cycles and 10 years charge retention. Other advantages include ease of scalability and low fabrication costs. However, a great deal of electrical power is consumed during programming [19, 20]. Hence, one of the main focuses of research into PCM is switching current reduction. The programming current scales with the contact area and improves with lithography scaling. Thus far, reducing the area of the bottom electrode contact has effectively reduced the power consumption, but the required area has always been much smaller than the respective process node [20]. From the viewpoint of rational scaling, the programming power has to be reduced to a level that is compatible with a conventionally sized bottom electrode contact for practical use. The smallest reset current achieved recently is 100 μA, compatible with core MOSFETs used in standard 0.13 μm CMOS technology [20]. The high reset current requirement puts a limit on the minimum width of the transistor used to apply this current, thus resulting in a larger cell size.

Magnetic Random Access Memory (MRAM) devices employ a magnetic tunnel junction (MTJ) as the memory element and a transistor to provide the drive current. An MTJ cell consists of two ferromagnetic materials separated by a thin insulating layer that acts as a tunnel barrier. When the magnetic moment of one layer is switched to align with the other layer (or to oppose the direction of the other layer) the effective resistance to current flow through the MTJ changes. The magnitude of the tunneling current can be read to indicate whether a one or a zero was stored. Advantages of MRAM are fast write and erase speeds (< 50ns), low power

requirements and very high endurance [21]. However, similar to FeRAM and PCM, MRAM also faces significant challenges for integration into mainstream CMOS production.

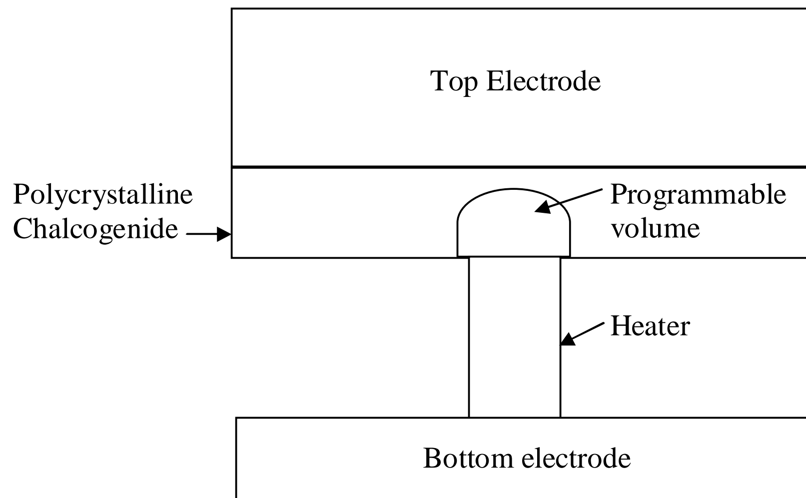


Figure 2.7: Basic cross section of a Phase Change Memory [19].

SONOS memory is considered to be one of the most attractive candidates to replace the conventional floating-gate structure. In SONOS memory, charges are stored within traps of the nitride charge storage layer. As the charges are stored in discrete traps of the insulating charge storage layer, any defect in the tunnel oxide will not cause all the charges to leak out. This is one of the main advantages of SONOS memory devices as compared to the conventional floating-gate structure. The idea behind the distributed charge storage is similar to that of the nanocrystal memory device. However, conventional silicon nanocrystal memory devices have smaller memory window as compared to SONOS due to the relatively small nanocrystal density (10^{10}-10^{11} cm^{-2}) [22]. Control of the nanocrystal size distribution, lateral spacing and shape are additional process challenges.

A summary of the various memory parameters for different types of nonvolatile memories is shown in Table 2.1.

Table 2.1: Summary of memory parameters for different types of nonvolatile memories

	Write /Erase Time	Write/ Erase Voltage	Cell Size (μm^2) (2006)	Endurance	Retention
SONOS	10 μs/ 1 ms	17V/ 17V	0.0098	> 10^5 cycles	> 10 years
NOR Flash	1 μs/ 1 ms	7V/ 7V	0.049	> 10^5 cycles	> 10 years
NAND Flash	200 μs/ 1 ms	17V/ 17V	0.0098	> 10^5 cycles	> 10 years
FeRAM	100 ns/ 100 ns	.8V/ 1.8V	0.34	> 10^{12} cycles	> 10 years
PCRAM	30 ns/ 50 ns	.8V/ 1.8V	0.047	> 10^5 cycles	> 10 years
MRAM	< 50 ns	.8V/ 1.8V	0.19	> 10^{15}cycles	> 10 years

2.3 SONOS Nonvolatile Memory

MNOS memories were invented nearly 30 years ago [2] and were the first electrically-alterable NVSM. Figure 2.8 illustrates schematically the progression of SONOS NVSM devices which has led to the present SONOS device structure. Initial device structures in the early 1970s were p-channel MNOS structures with aluminum gate electrodes and thick (i.e., 45 nm) silicon nitride charge storage layers. Write/erase voltages were typically 25-30V. In the late 1970s and early 1980s, scaling moved to n-channel polysilicon-nitride-oxide-silicon (SNOS) devices with write/erase voltages of 14-18V. The SNOS technology combines the use of a polysilicon gate technology with low pressure chemical vapour deposition (LPCVD) nitride of uniform thickness. The triple dielectric MONOS structure was introduced in 1968 [23]. The blocking oxide minimized charge injection from the gate electrode. In the late 1980s and early 1990s, n- and p-channel SONOS devices with the triple dielectric structure emerged with write/ erase voltages of 5-12 V.

The advantages of the triple dielectric structure are [24]:

(a) lower programming voltages since the blocking action of the top oxide enables easier scaling of the nitride thickness. Previously, nitride scaling is limited by charge leakage to the top electrode;

(b) minimized charge injection to and from the gate electrode;

(c) improved memory retention since the loss of charge to the gate electrode is minimized.

(d) increased endurance to extended erase/write cycling due to reduction in write/erase voltages.

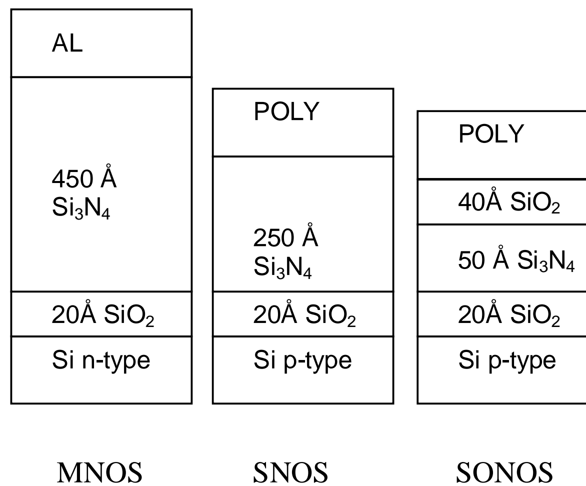


Figure 2.8: Evolution of the SONOS NVSM device [24].

The device operation mechanism is the same for SONOS and MONOS structures. Figure 2.9 shows the write/erase physical operation of a SONOS device [25]. A net positive or negative charge is stored in deep traps within the nitride dielectric depending on whether a negative or positive voltage is applied, respectively, to the gate electrode. In the program (write) operation, electrons quantum-mechanically tunnel from the silicon inversion layer through an energy barrier of 3.1 eV into an ultra-thin oxide which is typically less than 2.5 nm thickness (Fig. 2.9(b)).

The electrons tunnel through the ultra-thin oxide into the silicon nitride film and are stored in deep traps [25]. During the erase operation, trapped electrons in the nitride film are detrapped and these tunnel through the tunnel oxide into the silicon substrate. In addition, holes are injected from the p-substrate into the silicon nitride valence band where they are trapped in a manner similar to electrons (Fig. 2.9(a)). Thus in summary, for SONOS device operation both carrier types are involved in the transport process.

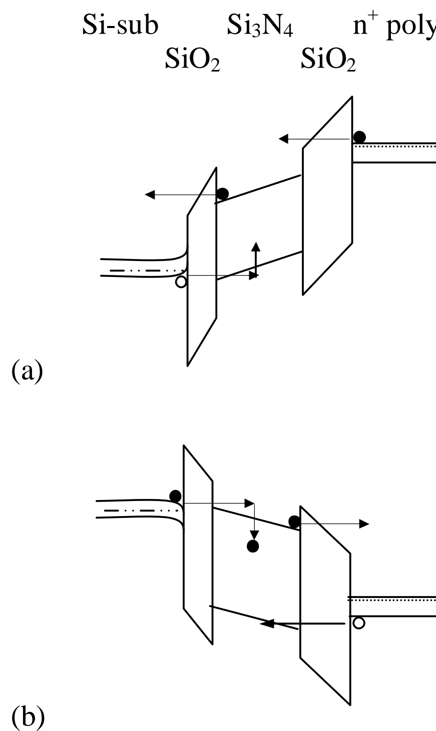


Figure 2.9: Physical operation of a SONOS device [25].

The main programming mechanisms in a SONOS device are direct band-to-band tunneling (DT), modified Fowler-Nordheim tunneling (MFN), trap-assisted tunneling (TAT) and Fowler-Nordheim (FN) tunneling which are illustrated in Figs. 2.10(a), (b), (c) and (d), respectively [26]. In addition, the requirement for the electric

field across the tunnel oxide for each mechanism is also stated in Fig. 2.10. The description of the various parameters in Fig. 2.10 is as follows:

E_{OX} = electric field across the tunnel oxide,

T_{OX} = tunnel oxide thickness,

T_N = Si_3N_4 thickness,

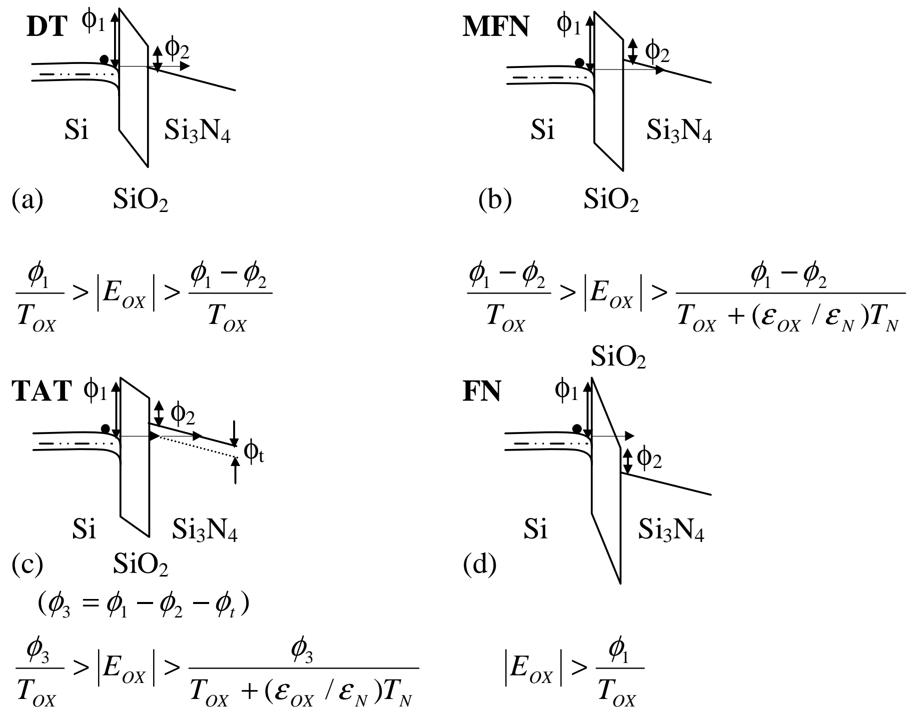
ε_{OX} = Dielectric constant of SiO_2,

ε_N = Dielectric constant of Si_3N_4,

ϕ_1 = potential barrier as a result of the conduction band discontinuity at the Si-SiO_2 interface,

ϕ_2 = potential barrier as a result of the conduction band discontinuity at the Si_3N_4-SiO_2 interface,

ϕ_t = trap energy level



$$\frac{\phi_1}{T_{OX}} > |E_{OX}| > \frac{\phi_1 - \phi_2}{T_{OX}}$$

$$\frac{\phi_1 - \phi_2}{T_{OX}} > |E_{OX}| > \frac{\phi_1 - \phi_2}{T_{OX} + (\varepsilon_{OX} / \varepsilon_N) T_N}$$

$$(\phi_3 = \phi_1 - \phi_2 - \phi_t)$$

$$\frac{\phi_3}{T_{OX}} > |E_{OX}| > \frac{\phi_3}{T_{OX} + (\varepsilon_{OX} / \varepsilon_N) T_N}$$

$$|E_{OX}| > \frac{\phi_1}{T_{OX}}$$

Figure 2.10: Energy band diagrams of the programming mechanisms: (a) Direct tunneling, (b) Modified Fowler-Nordheim tunneling, (c) trap assisted tunneling (d) Fowler-Nordheim tunneling [26].

In general, for high speed SONOS devices with good long term retention characteristics, the initial programming mechanism is DT [26]. As charge storage occurs in the nitride layer, the tunnel oxide electric field relaxes and the tunneling mechanism becomes MFN.

The main research on SONOS memory devices includes investigating ways to improve the device performance by gate stack scaling, optimization of the process and fabrication of novel devices using different structures or materials for the tunnel oxide, charge storage layer, blocking oxide and gate electrode. In the following sections, scaling of the SONOS gate stack would first be investigated. Subsequently, the use of different structures or materials for the tunnel oxide, charge storage layer, blocking oxide and gate electrode would be discussed. Finally, the use of high-κ materials in the SONOS gate stack to extend the scaling limits of the conventional SONOS memory is investigated.

2.3.1 SONOS gate stack scaling

The evolution of high-density EEPROMs continuously imposes a demand on reducing power consumption while improving data retention and endurance. The demand for low power and low voltage electronics has accelerated the pace for NVSM circuit designers to consider SONOS for low voltage, high density EEPROMs.

One method to improve the device performance is by scaling the oxide-nitride-oxide (ONO) stack. However, the scaling process is complex since varying the thickness of each dielectric layer can influence both the programming speed and charge retention. Previous MNOS/SONOS scaling scenarios keep the electric field across the tunnel oxide or nitride layer nearly constant and assuming zero charge in

the nitride layer [27]. To maintain the same write/erase speed when the programming voltage is scaled, the effective thickness must be scaled in accordance with the reduction in programming voltage. Scaling the effective thickness corresponds to decreasing the thickness of either the tunnel oxide, nitride or blocking oxide layers.

Decreasing the tunnel oxide thickness from 1.8nm to 1.1nm, while maintaining the thicknesses of the other layers constant (5 nm Si_3N_4 and 4 nm blocking oxide), effectively increases the program and erase speed [27]. However, the charge retention is severely degraded by the decrease in tunnel oxide thickness. On the other hand, scaling the blocking oxide thickness from 4 nm to 3.3 nm increases the programming speed for small threshold voltage (V_{th}) shift. However, for long programming pulse duration and large V_{th} shift, the V_{th} window is reduced [27]. This is probably due to the reduction in the blocking oxide capability to prevent charge transfer to and from the gate electrode. Decrease of the Si_3N_4 layer thickness from 13 nm to 4.5 nm with a simultaneous increase in the tunnel oxide (1.8 nm to 2 nm) and blocking oxide (3 nm to 5 nm) thicknesses to maintain the same initial electric field resulted in a lower programming voltage [28]. However, a smaller memory window was observed for the thinner nitride layer device. Hence, scaling the tunnel oxide is more effective than scaling the nitride or blocking oxide to improve programming speed.

2.3.2 Novel SONOS Structures

From the previous section, it was shown that scaling the gate stack to improve the programming speed by reducing the tunnel oxide thickness has the inevitable trade-off of charge retention degradation. Other methods to improve the device

performance are by process optimization or the use of different materials for the ONO layer and the gate electrode.

SONOS nonvolatile memory devices annealed in deuterium instead of the conventional forming gas or hydrogen anneal showed improved charge retention and endurance to program/erase cycling [29]. Interface state generation is reduced under program/erase cycling and charge retention is improved in deuterium annealed samples compared to their hydrogen-annealed counterparts. Interface states may provide an additional shift in the device V_{th} and degrade long-term charge retention by increasing the so-called back-tunneling current. During deuterium annealing, the atomic deuterium may diffuse to the Si-SiO_2 interface where they attach to silicon dangling bonds to terminate the electrically active interface traps, similar to the hydrogen case. It was proposed that the Silicon-Deuterium bonds may actually be stronger than the Silicon-Hydrogen (Si-H) bonds, resulting in reduction in interface state density generation.

Minami *et al.* proposed the use of a blocking SiO_2 deposited by chemical vapour deposition (CVD) instead of a thermally grown oxide [30]. The potential barrier of the CVD oxide may therefore be sharper than that of the thermally oxidized top oxide layer that includes oxynitride (SiON). This is due to the abrupt composition change from Si_3N_4 to CVD oxide compared to the gradual composition change from Si_3N_4 to SiON to SiO_2 in the thermally oxidized top oxide case. The sharper potential barrier increases the blocking capability of the top oxide, resulting in better charge retention.

Reisinger *et al.* have demonstrated a novel n-channel SONOS structure with p^+ polysilicon gate instead of the conventional n^+ polysilicon gate [31]. In the erase mode, the p^+ gate prevents the F-N tunneling of electrons from the conduction band of

the gate into the Si_3N_4. By bringing the Fermi level down to the valence band, the effective barrier for electron tunneling is increased by 1 eV. This improves the erase speed. However, the fabrication process is more complicated compared to the conventional n^+ gate due to the need of additional masking steps.

Improvement in data retention of MONOS memory devices could also be achieved by depositing Si_3N_4 with NH_3 and $SiCl_4$ (silicon tetrachloride, STC) instead of with NH_3 and the conventionally used $SiCl_2H_2$ (dichlorosilane, DCS) [32]. This was attributed to the reduction of Si-H bond density in the STC silicon nitride as compared to that of the DCS silicon nitride. The Si-H bond density in the STC Si_3N_4 was less than 1% of the DCS Si_3N_4, as measured by Fourier transform infrared spectroscopy. However, there is also a significant reduction in V_{th} window. This indicates that Si_3N_4 with a lower Si-H bond density has fewer carrier traps. The improvement in charge retention capability was attributed to suppression of the leakage of trapped electrons through shallow traps which are related to Si-H bonds with activation energy of 0.1-0.2 eV.

SONOS memory devices with band-gap-engineered Si_3N_4 charge trapping layer showed good endurance and superior charge retention capability compared to conventional Si_3N_4 devices [33, 34]. Band gap engineering was achieved by varying the silicon/nitrogen ratio from high to low during the deposition process through gas flow rate control of $SiCl_2H_2/NH_3$. Silicon-rich nitride has an abundance of shallow trapping levels attributed to silicon dangling bonds [35] while the standard nitride has deeper trapping levels [33, 34]. During programming, electrons can be easily captured by the shallow traps. Subsequently, the injected electrons are transferred to adjacent deeper levels by lateral hopping. However, the erasing speed of the band-gap-engineered device is slower compared to both conventional and silicon-rich Si_3N_4

devices. This may be due to the difficulty of discharging the electrons from deeper trapping levels near the blocking oxide.

Another method to improve device performance is by using alternative materials such as high-κ dielectrics as part of the gate stack. Basically, due to the higher dielectric constant or κ value, the equivalent oxide thickness is reduced for the same physical thickness of the film. Hence, the effect on device performance is expected to be similar to that of ONO stack scaling without the disadvantages that come with smaller physical thicknesses [36].

In 1978, tantalum oxide (Ta_2O_5) was investigated as a possible replacement for Si_3N_4 in MNOS structures [37]. The threshold voltage of the MTOS (metal-tantalum oxide-silicon dioxide-silicon) memory capacitors can be shifted using lower gate voltages than are needed for a comparable MNOS device. However, the charge retention characteristics of the MTOS devices have not been fully investigated.

High quality Si_3N_4 formed by rapid thermal nitridation was investigated as the tunnel dielectric in SONOS memory [38]. The control devices fabricated have conventional thermal SiO_2 tunnel dielectric. The tunnel nitride and tunnel silicon dioxide thicknesses investigated were 26 Å and 17 Å, respectively. Due to the lower barrier heights for electrons and holes for Si_3N_4, the physically thicker Si_3N_4 tunnel dielectric device has comparable programming speed as the control SiO_2 tunnel dielectric devices. In addition, the Si_3N_4 tunnel dielectric devices show comparable charge retention performance compared to the control devices due to the larger physical thickness of the tunnel Si_3N_4. The main advantage of using high quality Si_3N_4 as the tunnel dielectric is superior endurance characteristics, attributed to less interface trap generation. This may be due to the superior quality of the thermal

nitride and lower electric field in the higher-κ thermal nitride during programming/erasing.

Two-bit cell SONOS type flash memories using high-κ charge trapping layers were investigated by Sugizaki *et al.* [39]. The high-κ layers investigated were Si_3N_4, aluminum oxide (Al_2O_3) and hafnium oxide (HfO_2). Among these films, Al_2O_3 showed superior charge retention characteristics compared to Si_3N_4 devices while HfO_2 devices showed poor retention characteristics. The memory devices fabricated used hot-carrier injection for programming and hot-hole injection for erase as the tunnel oxide used is 70 Å thick.

Program/erase speed can also be improved by using high-κ blocking oxide [40-43]. From electrostatics consideration, the use of a high-κ blocking oxide layer will cause a smaller voltage drop across the blocking oxide and greater voltage drop across the tunnel oxide. This will result in a simultaneous increase of the electric field across the tunnel oxide and reduction of the electric field across the blocking oxide leading to more efficient program and erase processes. Hence, it allows the use of a thicker tunnel oxide layer leading to improved charge retention [41].

Several SONOS structures that use high-κ material for tunnel, charge storage and blocking oxide layers have also been proposed. A novel MONOS-type nonvolatile memory using HfO_2 as tunnel and blocking oxide layers and Ta_2O_5 as the charge trapping layer was proposed by Wang *et al.*[44]. The devices can be programmed as fast as 1 μs and erased in 10 ns at an 8-V gate bias. This may be due to the smaller conduction band offset with respect to the silicon substrate of the tunnel HfO_2 compared to that of conventional tunnel SiO_2. The charge retention obtained is also comparable to that of conventional SONOS devices. This is attributed to the thicker

physical thickness of the tunnel HfO_2 (48 Å) compared to the conventional tunnel SiO_2 (≤ 25 Å).

Wang *et al.* has demonstrated a novel SONOS structure with IrO_2/ HfAlO/ HfSiO/ HfAlO gate stack [45] with good charge retention. HfSiO charge storage layer has lower operation voltage than Si_3N_4 and better retention than HfO_2. The use of high work function IrO_2 as the metal gate leads to lower erasing voltage due to a higher barrier to minimise electron tunneling from the gate electrode.

References

[1] D. Kahng and S. M. Sze, "A floating-gate and its application to memory devices", *Bell Syst. Tech. J.*, vol. 46, pp. 1288-1292, 1967.

[2] H. A. R. Wegener, A. J. Lincoln, H. C. Pao, M. R. O'Connell and R. E. Oleksiak, "The variable threshold transistor, a new electrically alterable non-destructive read-only storage device", in *IEEE IEDM Tech. Dig.*, 1967, p. 70.

[3] D. Frohman-Bentchkowsky, "A fully decoded 2048-bit electrically programmable MOS-ROM", in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers.*, 1971, p. 80.

[4] D. Frohman-Bentchkowsky, "Memory behaviour in a floating-gate avalanche injection MOS (FAMOS) structure", *Appl. Phys. Lett.*, vol. 18, pp. 332-334, 1971.

[5] D. Frohman-Bentchkowsky, "A fully decoded 2048 bit electrically programmable FAMOS read-only memory", *IEEE J. Solid State Circuits*, vol. SC-6, pp. 301- 306, 1971.

[6] D. Frohman-Bentchkowsky, "FAMOS-A new semiconductor charge storage device", *Solid State Electronics*, vol. 17, pp. 517-528, 1974.

[7] H. Iizuka, T. Sato, F. Masuoka, K. Ohuchi, H. Hara, H. Tango, M. Ishikawa, and Y. Takeishi, "Stacked gate avalanche injection type MOS (SAMOS) memory", in *Proc. 4th Conf. Solid State Devices, Tokyo*, 1972; *J. Japan. Soc. Appl. Phys.*, vol. 42, 1973, p. 158.

[8] H. Iizuka, F. Masuoka, T. Sato, and M. Ishikawa, "Electrically alterable avalanche injection type MOS read-only memory with stacked gate structure", *IEEE Trans. Electron Devices*, vol. ED-23, pp. 379-387, 1976.

[9] "Nonvolatile Semiconductor Memory Technology, A Comprehensive Guide to Understanding and Using NVSM Devices", Edited by William D. Brown, Joe E. Brewer, IEEE Press, pp. 9-245, 1998.

[10] G. Yaron, S. J. Prasad, M. S. Ebel and B. M. K. Leong, "A 16K EEPROM employing new array architecture and designed-in reliability features", *IEEE J. Solid State Circuits*, vol. SC-17, pp. 833-840, 1982.

[11] S. K. Lai, V. K. Dham and D. Guterman, "Comparison and trends in today's dominant E^2 technology", in *IEEE IEDM Tech. Dig.*, 1986, p. 580.

[12] F. Masuoka, M. Asano, H. Iwahashi, and T. Komuro, "A new Flash EEPROM cell using triple polysilicon technology", in *IEEE IEDM Tech. Dig.*, 1984, p. 464.

[13] E. Suzuki, H. Hiraishi, K. Ishi and Y. Hayashi, "A low voltage alterable EEPROM with metal-oxide-nitride-oxide-semiconductor (MONOS) structure", *IEEE Trans. Electron Devices*, vol. 30, pp. 122-128, 1983.

[14] J. T. Evans and R. Womack, "An experimental 512-bit nonvolatile memory with ferroelectric storage cell", *IEEE J. Solid State Circuits*, vol. 23, pp. 1171-1175, 1988.

[15] R. Moazzami and C. Hu, "A ferroelectric DRAM cell for high-density NVRAM's", *IEEE Electron Device Letters.*, vol. 11, pp. 454-456, 1990.

[16] "Gate Dielectrics and MOS ULSIs: Principles, Technologies and Applications", Edited by W. Engl, Springer Series in Electronics and Photonics, pp. 315-316, 1998.

[17] International Technology Roadmap for Semiconductors, 2005.

[18] R. Bez, E. Camerlenghi, A. Modelli and A. Visconti, "Introduction to Flash Memory", *Proceedings of the IEEE*, vol. 91, pp. 489-501, 2003.

[19] S. Lai, "Current status of the phase change memory and its' future", in *IEEE IEDM Tech. Dig.*, 2003, p. 255.

[20] N. Matsuzaki, K. Kurotsuchi, Y. Matsui, O. Tonomura, N. Yamamoto, Y. Fujisaki, N. Kitai, R. Takemura, K. Osada, S. Hanzawa, H. Moriya, T. Iwasaki, T. Kawahara, N. Takaura, M. Terao, M. Matsuoka, and M. Moniwa, "Oxygen-doped GeSbTe phase-change memory cells featuring 1.5-V/100-μA standard 0.13-μm CMOS Operations", in *IEEE IEDM Tech. Dig.*, 2005, p. 738.

[21] Stuart S. P. Parkin, "Spintronic materials and devices: Past, present and future!", in *IEEE IEDM Tech. Dig.*, 2004, p. 903.

[22] K-H. Joo, X. Wang, J. H. Han, S-H. Lim, S-J. Baik, Y-W. Cha, J. W. Lee, I-S. Yeo, Y-K. Cha, I. K. Yoo, U-I. Chung, J. T. Moon, and B-I. Ryu, "Novel transition layer engineered Si nanocrystal flash memory with MHSOS structure featuring large V_{th} window and fast P/E speed", in *IEEE IEDM Tech. Dig.*, 2005, p. 166.

[23] B. V. Keshevan and H. C. Lin, "MONOS memory element", in *IEEE IEDM Tech. Dig.*, 1968, p. 140.

[24] M. H. White, L. Yang, A. Purwar and M. L. French, "A low voltage SONOS nonvolatile semiconductor memory technology", *IEEE Trans. on Components, Packaging and Manufacturing Technology*, Part A, vol. 20, pp. 190-195, 1997.

[25] M. H. White, D. A. Adams, B. Jiankang, "On the go with SONOS", *IEEE Circuits and Devices*, pp. 22-31, 2000.

[26] B. Jiankang and M. H. White, "Design considerations in scaled SONOS nonvolatile memory devices", *Solid State Electronics*, vol. 45, pp. 113-120, 2001.

[27] M. L. French, C. Y. Chen, H. Sathianathan and M. H. White, "Design and scaling of a SONOS multidielectric device for nonvolatile memory applications", *IEEE Trans. on Components, Packaging and Manufacturing Technology*, Part A, vol. 17, pp. 390-397, 1994.

[28] B. Jiankang and M. H. White, "Retention reliability enhanced SONOS NVSM with scaled programming voltage", in *IEEE Aerospace Conf. Proc.*, vol. 5, 2002, p. 2383.

[29] B. Jiankang and M. H. White, "Improvement in retention reliability of SONOS nonvolatile memory devices by two-step high temperature Deuterium anneals", in *IEEE Int. Reliability Phys. Symp. Proc.*, 2001, p. 52.

[30] S. Minami and Y. Kamigaki, "A novel MONOS nonvolatile memory device ensuring 10-year data retention after 10^7 erase/ write cycles", *IEEE Trans. Electron Devices*, vol. 40, pp. 2011-2017, 1993.

[31] H. Reisinger, M. Franosch, B. Hasler and T. Bohm, "A novel SONOS structure for nonvolatile memories with improved data retention", in *VLSI Tech. Symp. Proc.*, 1997, p. 113.

[32] K. Nomoto, G. Asayama and T. Kobayashi, "Data retention improvement of metal-oxide-nitride-oxide-semiconductor memories using silicon-tetrachloride-based silicon nitride with ultra low Si-H bond density", *Jap. J. Appl. Phys.*, vol. 43, pp. 2211-2216, 2004.

[33] T. S. Chen, K. H. Wu, H. Chung and C. H. Kao, "Performance improvement of SONOS memory by bandgap engineering of charge-trapping layer", *IEEE Electron Device Letters*, vol 25, pp. 205-207, 2004.

[34] K. H. Wu, H. C. Chien, C. C. Chan, T. S. Chen and C. H. Kao, "SONOS device with tapered bandgap nitride layer", *IEEE Trans. Electron Devices*, vol. 52, pp. 987-992, 2005.

[35] M. Tanaka, S. Saida, Y. Mitani, I. Mizushima and Y. Tsunashima, "Highly reliable MONOS devices with optimized silicon nitride film having deuterium terminated charge traps", in *IEEE IEDM Tech. Dig.*, 2002, p. 237.

[36] M. Specht, H. Reisinger, M. Stadele, F. Hofmann, A. Gschwandtner, E. Landgraf, R. J. Luyken, T. Schulz, J. Hartwich, L. Dreeskornfeld, W. Rosner, J. Kretz and L. Risch, "Retention times of novel charge trapping memories using Al_2O_3 dielectrics", in *33rd Conf. on European Solid-State Device Research*, 2003, p. 155.

[37] R. L. Angle and H. E. Talley, "Electrical and charge storage characteristics of the tantalum oxide-silicon dioxide device", *IEEE Trans. Electron Devices*, vol. 25, pp. 1277-1283, 1978.

[38] M. She, H. Takeuchi and T. J. King, "Silicon-nitride as a tunnel dielectric for improved SONOS-type flash memory", *IEEE Electron Device Letters*, vol. 24, pp. 309-311, 2003.

[39] T. Sugizaki, M. Kobayashi, M. Ishidao, H. Minakata, M. Yamaguchi, Y. Tamura, Y. Sugiyama, T. Nakanishi and H. Tanaka, "Novel multi-bit SONOS type flash memory using high-κ charge trapping layer", in *VLSI Tech. Symp. Proc.*, 2003, p. 27.

[40] V. A. Gritsenko, “Design of SONOS memory transistor for terabit scale EEPROM”, in *IEEE Conf. on Electron Devices and Solid-State Circuits*, 2003, p. 345.

[41] S. Choi, M. Cho, H. Hwang and J. W. Kim, “Improved metal-oxide-nitride-oxide-silicon-type flash device with high-κ dielectrics for blocking layer”, *J. Appl. Phys.*, vol. 94, pp. 5408-5410, 2003.

[42] C. H. Lee, K. I. Choi, M. K. Cho, Y. H. Song, K. C. Park and K. Kim, “A novel SONOS structure of SiO_2/SiN/Al_2O_3 with TaN metal gate for multi-giga bit flash memories”, in *IEDM Tech. Dig.*, 2003, p. 613.

[43] C. H. Lee, S. H. Hur, Y. C. Shin, J. H. Choi, D. G. Park and K. Kim, “Charge-trapping device structure of SiO_2/SiN/High-κ dielectric Al_2O_3 for high-density flash memory”, *Appl. Phys. Lett.*, vol. 86, pp. 152908 (1-3), 2005.

[44] X. Wang, J. Liu, W. Bai and D. L. Kwong, “A novel MONOS-type nonvolatile memory using high-κ dielectrics for improved data retention and programming speed”, *IEEE Trans. Electron Devices*, vol. 51, pp. 597-602, 2004.

[45] Y. Q. Wang, P. K. Singh, W. J. Yoo, Y. C. Yeo, G. Samudra, A. Chin, J. H. Chen, S. J. Wang and D. L. Kwong, “Long retention and low voltage operation using IrO_2/HfAlO/HfSiO/HfAlO gate stack for memory application”, in *IEDM Tech. Dig.*, 2005, p. 162.

Chapter 3

Hafnium Oxide as the Charge Storage Layer in SONOS-type Nonvolatile Flash Memory for Minimization of the Over-erase Phenomenon

3.1 Introduction

The polysilicon-oxide-silicon nitride-oxide-silicon (SONOS) structure has recently drawn attention for application in electrically-erasable-programmable-read-only-memories (EEPROMs) due to superior charge retention performance compared to the conventional polysilicon floating-gate type EEPROMs [1]. Since the SONOS device stores charge in the spatially isolated deep-level traps, a single defect in the tunnel oxide will not cause the discharge of the memory cell [2, 3].

The current demand for low power and low voltage electronics has accelerated the pace for SONOS gate stack (ONO) scaling. The most effective way to improve the programming speed is by reducing the tunnel oxide thickness. However, this will inevitably result in charge retention degradation. An alternative method to improve device performance is by using high-κ material as part of the gate stack. Due to the higher dielectric constant of the film, the equivalent oxide thickness is reduced for the same physical thickness of the film. Hence, the effect on device performance is expected to be similar to that of ONO gate stack scaling without the disadvantages that come with smaller physical thicknesses.

In addition, SONOS devices are susceptible to over-erase, in which the threshold voltage of the erased device becomes more negative than the uncharged device. During the write (program) operation of a SONOS n-channel transistor

device, electrons tunnel through the ultra-thin oxide into the silicon nitride film and are stored in deep-level traps [4]. During the erase operation under negative gate bias, trapped electrons in the nitride film are detrapped and these tunnel through the tunnel oxide into the silicon substrate. In addition, holes are injected from the p-type substrate into the silicon nitride valence band during the erase operation, where they are trapped in deep-level hole traps. Thus, for SONOS device operation, both carrier types are involved in the transport process. This makes threshold voltage control during the erase operation difficult. If the electrical erase continues beyond a specified point, due to excessive hole injection, it will result in more positive charges on the silicon nitride causing an over-erase problem [5]. The threshold voltage of the erased device will be more negative than the uncharged device. The over-erase phenomenon in the SONOS device can short out the column of memory cells which the SONOS device is connected to in an electrically-programmable read-only memory array-like structure, such as in the NOR array [5].

In the subsequent sections, improvement in programming speed and the reduction of the over-erase phenomenon by using hafnium oxide instead of silicon nitride as the charge storage layer will be demonstrated. These are attributed to the differences in band offset and crystallinity of the charge storage layer.

3.2 Sample Fabrication

We have fabricated three different SONOS-type memory capacitors with a triple dielectric stack structure. These three different structures exhibit different severity of the over-erase phenomenon. The structure which uses silicon nitride (Si_3N_4) as the charge storage layer is the conventional SONOS device. The other two structures use a hafnium oxide (HfO_2) film as the charge storage layer and are denoted

as SOHOS1 and SOHOS2 devices. The capacitors were fabricated using 4-8 Ω-cm (100) p-type silicon substrates. In the triple dielectric stack formation, the 25 Å tunnel oxide was grown by rapid thermal oxidation at 1000^0C. After tunnel oxide formation, 60 Å of Si_3N_4 was deposited by low-pressure chemical-vapor-deposition (LPCVD) for the SONOS structure. For SOHOS1 structures, 60 Å of HfO_2 was deposited by metal-organic-chemical-vapor-deposition (MOCVD) at 400 °C under a pressure of 0.4 Torr, followed by post-deposition-annealing (PDA) at 700 °C in a nitrogen ambient. For SOHOS2 structures, the deposition and PDA of 30 Å HfO_2 was carried out twice to obtain a total thickness of 60 Å. As the PDA was carried out twice, the HfO_2 film in SOHOS2 devices is expected to be more crystallized than that in SOHOS1 devices [6]. Finally, 55 Å thick of blocking oxide was deposited as LPCVD TEOS ($Si(OC_2H_5)_4$). All capacitor structures have similar gate areas of 800 × 800 μm^2. The resulting structure is shown in Fig. 3.1.

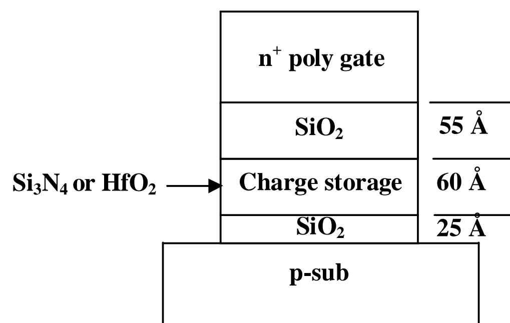


Figure 3.1: Fabricated SONOS-type memory devices with Si_3N_4 or HfO_2 charge storage layers.

In addition, SONOS-type memory n-channel transistors with Si_3N_4 and HfO_2 (SOHOS1) charge storage layers with the same gate stack thickness as the capacitor structures were also fabricated. Transistor source/drain annealing was performed at

950 ^{o}C for 30s. The transistor structures tested have gate width to gate length dimensions of W/L = 100 µm/20 µm.

3.3 Results and Discussion

The flatband voltage shift during charging and discharging of the memory capacitor structures were extracted from high-frequency capacitance-voltage (C-V) measurements, which showed a counter-clockwise hysteresis in the C-V curves of the p-type substrate capacitors. The quasi-neutral C-V curve (i.e, the uncharged condition of the memory device) was obtained by restricting the bias during the forward and reverse C-V sweeps to a small gate voltage range to avoid charging up the capacitors. The flatband voltage shift with respect to the quasi-neutral condition was plotted against the charging/discharging voltage (positive/negative gate voltage) in Fig. 3.2. The SONOS memory capacitor did not show any saturation behavior in flatband voltage shift for both charging (write) and discharging (erase) operations, unlike that of the SOHOS1 and SOHOS2 memory capacitors. From Fig. 3.2, it can be seen that during erase, the flatband voltage of the SONOS device was shifted negatively with respect to the quasi-neutral condition. The negative flatband voltage shift in the SONOS device increased monotonically as the erase (negative gate) voltage increased, as holes were injected into the nitride layer resulting in the over-erase phenomenon [5]. In contrast, the flatband voltage of both SOHOS structures became saturated after the erase operation, and was almost similar to that of the uncharged device. The flatband voltage shift for the SOHOS capacitors saturated at about –1 V for SOHOS1 device and 0 V for SOHOS2 device after the erase operation. Hence, the over-erase phenomenon that was present in the SONOS device was reduced in both SOHOS structures.

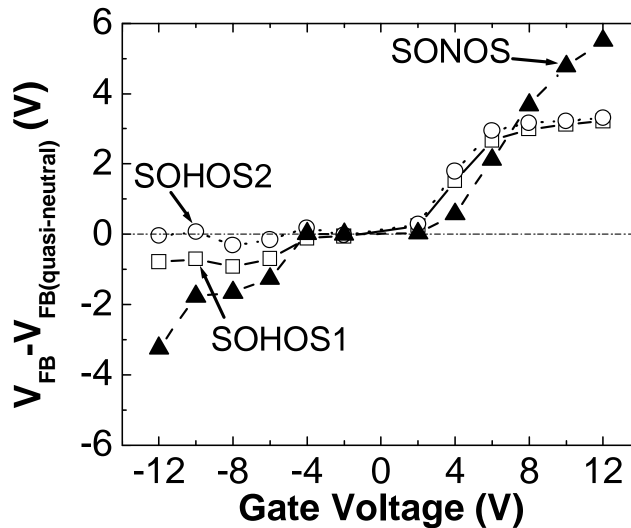


Figure 3.2: Flatband voltage shift plotted against the charging (positive) and discharging (negative) gate voltage for SONOS, SOHOS1 and SOHOS2 memory devices.

The program/erase results of SOHOS1 (with HfO_2 charge storage layer) and SONOS n-channel MOSFETs for $V_g - V_{fb} = +6V$ during program and $V_g - V_{fb} = -5.3V$ during erase are illustrated in Figs. 3.3(a) and (b), respectively. The SOHOS device had a faster programming speed compared to SONOS. It can be seen that the SOHOS device showed little over-erase as the erase threshold voltage saturated at the threshold voltage value of an uncharged device (i.e., $V_{th}(t=0)$). On the other hand, the SONOS device did not show any saturation in its erase threshold voltage.

In addition, the program/erase cycling data, or endurance results, for SONOS and SOHOS1 n-channel MOSFETs with Si_3N_4 and HfO_2 charge storage layers are shown in Figs. 3.4 (a) and (b), respectively. Threshold window closure was observed after 400 program/erase cycles for the SOHOS1 device with HfO_2 charge storage layer. For the SONOS device, no significant threshold window degradation was observed after 10,000 program/erase cycles. Hence, SOHOS1 device with HfO_2

charge storage layer had much poorer endurance characteristics as compared to SONOS. It is imperative to improve the endurance characteristics as nonvolatile memories have to be reprogrammed frequently during normal operation. The endurance characteristics can be improved by using hafnium aluminum oxide as the charge storage layer which will be shown in the subsequent chapter.

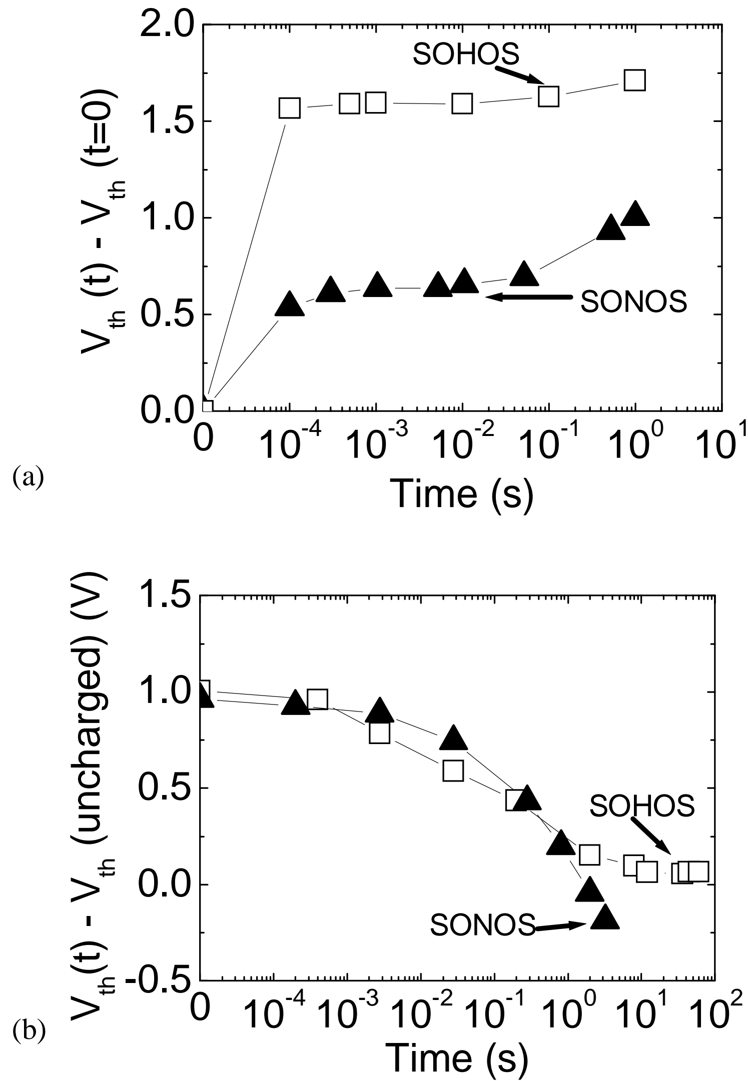


Figure 3.3: (a) Program and (b) erase threshold voltage shift of SOHOS1 (with HfO_2 charge storage layer) and SONOS n-channel MOSFETs for V_g - V_{fb} = +6 V during program and V_g - V_{fb} = -5.3V during erase.

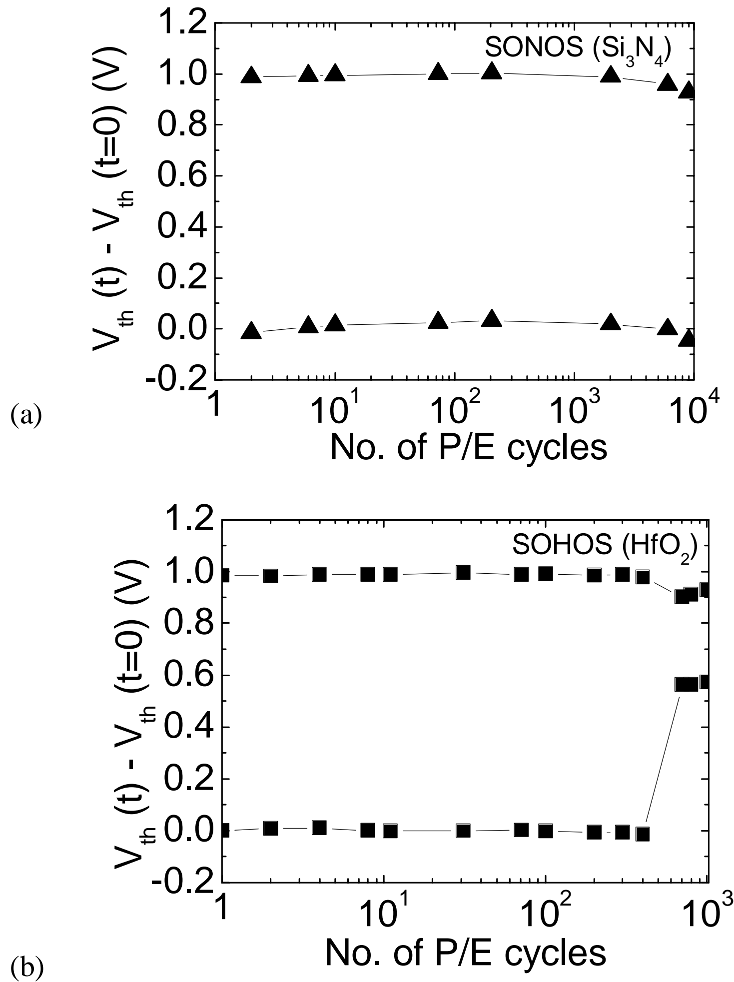


Figure 3.4: Program/erase (P/E) cycling data for (a) SONOS and (b) SOHOS1 (with HfO_2 charge storage layer) n-channel MOSFETs.

The major structural difference between SOHOS (with HfO_2 charge storage layer) and SONOS devices is the difference in band offset and the crystallinity of the charge storage layer. The nitride film was in an amorphous state, while the HfO_2 film was well crystallized by the high temperature PDA process [6-8]. The degree of crystallization of the HfO_2 film in the SOHOS2 device will be higher as compared to that in the SOHOS1 device, as the SOHOS2 device underwent the PDA process twice [6]. X-ray diffraction (XRD) results in Fig. 3.5 show that SOHOS2 was more crystallized as evidenced by the presence of (1,1,1) and (0,0,2) crystalline peaks of

HfO_2 in the XRD spectrum of SOHOS2 but not in that of SOHOS1. For SOHOS devices, the charges may be trapped in electron and hole traps in the HfO_2 layer or by charge confinement in the quantum well, similar to SONOS devices. From the ideal energy band diagrams of SONOS and SOHOS structures shown in Figs. 3.6 (a) and 3.6 (b), respectively, the quantum well formed by the conduction band is deeper for the SOHOS structure as compared to the SONOS structure (1.6 eV compared to 1.05 eV) [9, 10]. Therefore, at the same gate bias where modified Fowler-Nordheim (F-N) tunneling dominates, the electrons must tunnel through a thicker energy barrier in SONOS to the conduction band of the charge storage layer (Si_3N_4) as compared to SOHOS. The conduction band offset of Si_3N_4 with respect to silicon is 2.05 eV, as compared to a 1.5 eV conduction band offset of HfO_2 with respect to silicon. This is illustrated in Fig. 3.7 (a), where the modified F-N tunneling consists of direct tunneling through the thin tunnel oxide layer and F-N tunneling through the charge storage layer. The flatband voltage shift with respect to gate voltage is higher for SOHOS devices in the lower voltage region between 0 V and 7 V, as illustrated in Fig. 3.2. In addition, the programming speed results for the SOHOS (with HfO_2 charge storage layer) and SONOS transistors are shown in Fig. 3.3 (a). It can be seen that the SOHOS device charged up much faster compared to SONOS.

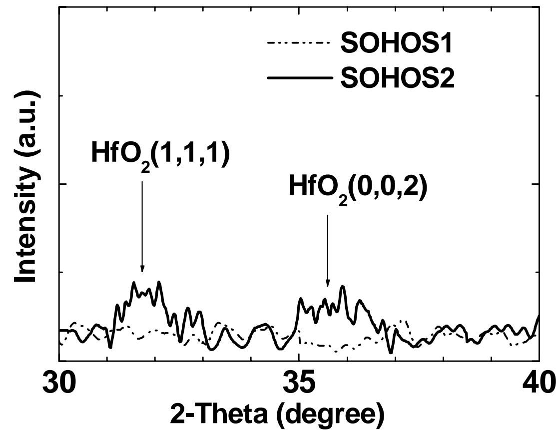


Figure 3.5: X-ray diffraction spectra of SOHOS1 and SOHOS2 structures.

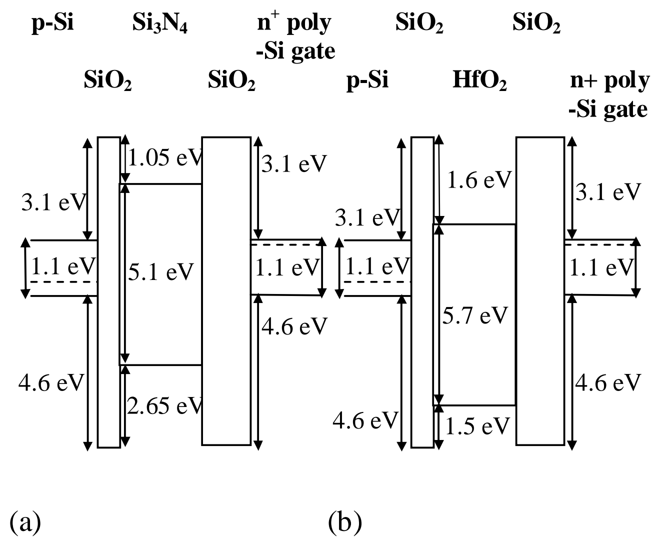


Figure 3.6: Ideal energy band diagrams for (a) SONOS and (b) SOHOS structures.

During the erase operation, the electrons in the HfO_2 quantum well and electron traps tunneled through the tunnel oxide to the p-type silicon (p-Si) substrate. Due to the larger barrier height for holes (4.6 eV) compared to that of electrons (3.1 eV), hole tunneling from the p-Si substrate to the HfO_2 quantum well was minimal. In addition, SOHOS has a shallower quantum well for holes (1.5 eV) as compared to SONOS (2.65 eV). Hence, at the same gate bias where modified F-N tunneling

dominates, the holes would have to tunnel through a much thicker barrier to the HfO_2 valence band, as compared to the Si_3N_4 case. The valence band offset of Si_3N_4 with respect to silicon is 1.95 eV, as compared to a 3.1 eV valence band offset of HfO_2 with respect to silicon. This is illustrated in Fig. 3.7 (b). Since holes were minimally involved in the erase operation, the erase flatband voltage shift of the SOHOS2 devices did not become more negative than that of the uncharged device, i.e. there was no net positive charge on the HfO_2 layer after erase. Hence the over-erase problem was minimized in SOHOS devices if the HfO_2 was well crystallized with minimum amount of charge trapping sites. In this case, the over-erase phenomenon would be dependent on the degree of crystallization of the HfO_2 film. This was verified by comparing between the erase flatband voltage shifts of SOHOS1 and SOHOS2 devices as shown in Fig. 3.2. The SOHOS1 device, which was less crystallized, showed a more negative flatband voltage shift than that of the uncharged device, which means that holes have tunneled through the tunnel oxide into hole traps in HfO_2 during erase.

In addition, it had been shown by Yeo *et al.* that multistep deposition of HfO_2 resulted in lower gate leakage current as compared to a single step deposition for HfO_2 capacitor structures [6]. For multistep deposited films, the grain boundaries and pinholes may be offset from one layer to another to block leakage current paths [6].

However, it was also found that SOHOS has poorer charge retention performance than SONOS as shown in Fig. 3.8. There is a possibility that crystallization of HfO_2 will generate grain boundaries which can act as current leakage paths [8]. Hence there will be an increase in lateral conduction which can result in poor charge retention.

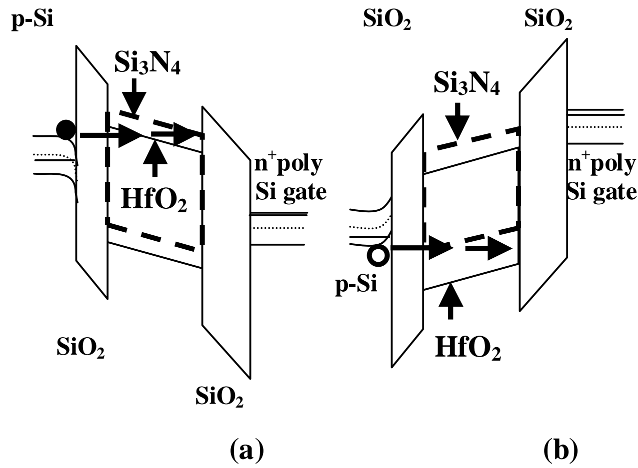


Figure 3.7: Energy band diagram schematic of the SONOS structure with HfO_2 (solid lines) or Si_3N_4 (dashed lines) as the charge storage layer during (a) write (program) and (b) erase operations.

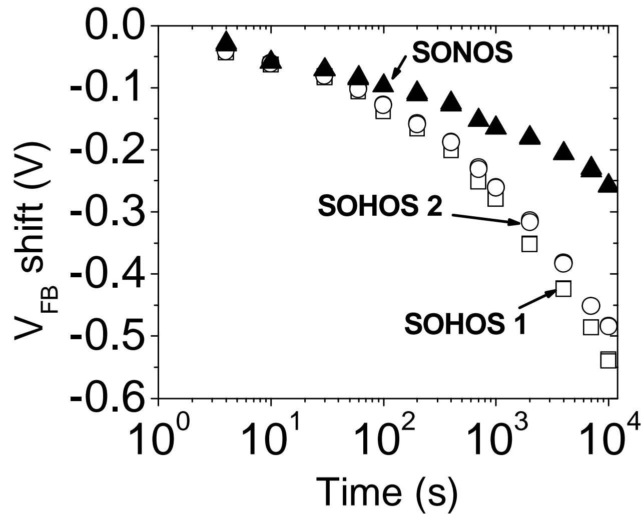


Figure 3.8: Charge retention performance of the SOHOS1, SOHOS2 and SONOS devices as characterized by the flatband voltage shift at an applied gate bias (V_g) of 0V after the device has been charged at V_g = 6V.

3.4 Summary

We have shown that the over-erase phenomenon in SONOS memory structures can be minimized by replacing silicon nitride with hafnium oxide as the charge storage layer. The reduction in over-erase in the SOHOS structure as compared to the conventional SONOS structure was attributed to the difference in band offset and degree of crystallization of the charge storage layer. SOHOS structure with HfO_2 charge storage layer showed faster programming than conventional SONOS. However, it had poorer retention and endurance characteristics. In the next chapter, it will be shown that the addition of aluminum to HfO_2 to form hafnium aluminum oxide will greatly improve the device retention and endurance characteristics.

References

[1] B. Jiankang and M. H. White, "Design considerations in scaled SONOS nonvolatile memory devices", *Solid State Electronics*, vol. 45, pp. 113-120, 2001.

[2] H. Wann and C. Hu, "High endurance ultra-thin tunnel oxide in MONOS device structure for dynamic memory applications", *IEEE Electron Device Lett.*, vol. 16, pp. 491-493, 1995.

[3] Y. Kamagaki, S. I. Minami, T. Hagiwara, K. Furusawa, T. Furuno, K. Uchida, M. Terasawa and K. Yamazaki, "Yield and reliability of MNOS EEPROM products", *IEEE J. of Solid State Circuits*, vol. 24, pp. 1714- 1722, 1989.

[4] M. H. White, D. A. Adams and B. Jiankang, "On the go with SONOS", *IEEE Circuits and Devices*, pp. 22-31, 2000.

[5] "Nonvolatile Semiconductor Memory Technology, A Comprehensive Guide to Understanding and Using NVSM Devices", Edited by William D. Brown and Joe E. Brewer, IEEE Press, pp. 193-309, 1998.

[6] C. C. Yeo, B. J. Cho, M. S. Joo, S. J. Whoang, D. L. Kwong, L. K. Bera, S. Mathew and N. Balasubramaniam, "Improvement of electrical properties of MOCVD HfO_2 by multistep deposition", *Electrochemical and Solid State Lett.*, vol. 6, pp. F42-F44, 2003.

[7] W. J. Zhu, T. Tamagawa, M. Gibson, T. Furukawa and T. P. Ma, "Effect of Al inclusion in HfO_2 on the physical and electrical properties of the dielectrics", *IEEE Electron Device Lett.*, vol. 23, pp. 649-651, 2002.

[8] M. S. Joo, B. J. Cho, C. C. Yeo, D. S. H. Chan, S. J. Whoang, S. Matthew, L. K. Bera, N. Bala and D. L. Kwong, "Formation of hafnium-aluminum-oxide

gate dielectric using single cocktail liquid source in MOCVD process", *IEEE Trans. Electron Devices*, vol. 50, pp. 2088-2094, 2003.

[9] M. H. White, L. Yang, A. Purwar and M. L. French, "A low voltage SONOS nonvolatile semiconductor memory technology", *IEEE Trans. Components, Packaging and Manufacturing Technology*, Part A, vol. 20, pp. 190-195, 1997.

[10] G. D. Wilk, R. M. Wallace, J. M. Anthony, "High-K gate dielectrics: Current status and materials properties considerations", *J. Appl. Phys.*, vol. 89, pp. 5243-5275, 2001.

Chapter 4

Hafnium Aluminum Oxide as the Charge Storage Layer in SONOS-type Nonvolatile Memory for High-Speed Operation with Improved Charge Retention and Endurance Performance

4.1 Introduction

SONOS (polysilicon-oxide-silicon nitride-oxide-silicon) Flash memory is one of the most attractive candidates to realize Flash vertical scaling [1]. Increase in programming speed of SONOS devices and lower voltage operation had been accomplished previously by reducing the tunnel oxide thickness [2-3]. However, this seriously degraded the charge retention capability of the device. To overcome this limitation, the so-called SOHOS (polysilicon-oxide-high-κ-oxide-silicon) Flash memory had been attempted by replacement of the silicon nitride layer with a high dielectric constant (high-κ) material [4-6]. In the previous chapter, SOHOS structure with hafnium oxide (HfO_2) as the charge storage layer demonstrated superior charge storage capability at low voltages, faster programming and less over-erase problem as compared to the conventional SONOS device. However, such a SOHOS device had poorer charge retention capability than SONOS. On the other hand, using aluminum oxide (Al_2O_3) as the charge storage layer resulted in a SOHOS structure with improved charge retention performance, but at the expense of a slower programming speed [5]. By adding a small amount of aluminum to HfO_2 to form hafnium aluminum oxide (HfAlO), we will demonstrate that the resultant SOHOS structure with HfAlO

as a charge storage layer can combine the advantages of both HfO_2 and Al_2O_3, such as fast programming speed, good charge retention and good program/erase endurance.

4.2 Sample Fabrication

A tunnel oxide layer of 25 Å to 34 Å was thermally grown at 800°C on 4-8 Ω-cm (100) p-type silicon substrates. For SONOS device, a Si_3N_4 layer (60 and 75 Å) was deposited by low pressure chemical-vapor-deposition (LPCVD). For SOHOS device, pure HfO_2 and Al_2O_3 films were deposited by atomic-layer-deposition (ALD) while HfAlO films were deposited by metal-organic-chemical-vapor-deposition (MOCVD) using a single cocktail source [6, 7]. The Al_2O_3 concentration in HfAlO was controlled to be 10% [6, 7]. The blocking oxide (55 and 65 Å thickness) was deposited using LPCVD TEOS ($Si(OC_2H_5)_4$). Lastly, HfN/TaN metal gate was formed by physical-vapor-deposition for the control gate [8-11]. The resulting structure is shown in Fig. 4.1. The transistors undergo source/drain implantation followed by activation annealing at 950°C for 30s. All capacitor structures have similar gate areas of 800 × 800 μm^2 while the transistor structures tested have gate width (W) to gate length (L) dimensions of W/L = 100 μm/20 μm.

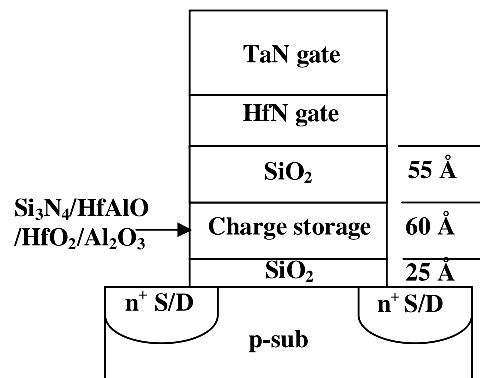


Figure 4.1: Fabricated SOHOS (with HfO_2 or HfAlO or Al_2O_3 charge storage layer) and SONOS (Si_3N_4) transistor structures with HfN/TaN gate electrode.

4.3 Results and Discussion

The charge retention performance of SONOS-type structures with four different charge storage layers was compared in Fig. 4.2 by measuring the flatband voltage (V_{fb}) shift of the programmed devices. The devices were programmed to an initial V_{fb} shift of 1.1 V before the retention measurements. The retention measurements were performed for durations of up to 10^4 s. After about 10^3 s, the V_{fb} shift showed a logarithmic decay in time. Hence, assuming that the V_{fb} decay followed a constant rate, the V_{fb} shift was extrapolated to 10 years. It is seen that the device with HfO_2 as the charge storage layer showed the worst charge retention characteristic. However, by adding 10% of Al_2O_3 into HfO_2 to form HfAlO, the charge retention performance of the resulting structure had been greatly improved. The HfAlO device had similar charge retention performance as Si_3N_4 and was only slightly worse than Al_2O_3.

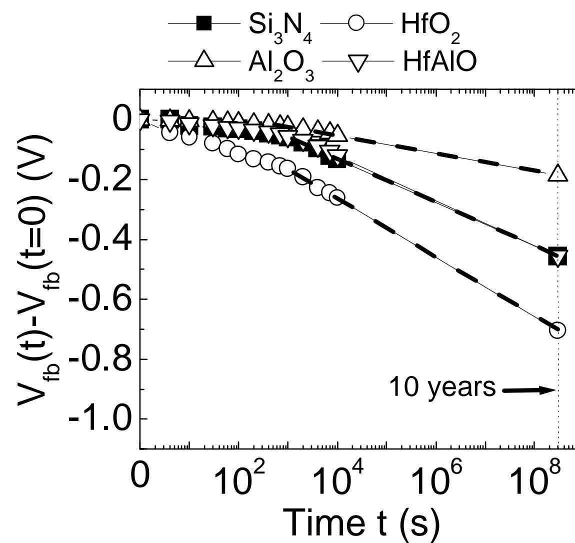


Figure 4.2: Flatband voltage shift during charge retention measurements versus time of SONOS-type memory devices with Si_3N_4, Al_2O_3, HfO_2 or HfAlO as the charge storage layer during discharging at a gate bias of -1.45 V below the initial flatband voltage of a charged device. The devices were programmed to an initial V_{fb} shift of 1.1 V before the retention measurements.

The dependence of charge loss with respect to temperature was further investigated for Al_2O_3 and HfO_2 charge storage layer devices in order to find the relative trap depth in Al_2O_3 as compared to HfO_2. Al_2O_3 devices had the best retention characteristics while HfO_2 devices had the worst retention characteristics. The relative trap depth of HfAlO was expected to be intermediate between HfO_2 and Al_2O_3. The rate of discharge was monitored by the difference of the drain current from its initial state (during the discharge process) at a particular read voltage after writing, conducted over a range of temperatures.

The linear drain current of a transistor can be approximated by [12]:

$$I_{D(lin)} = \mu \frac{W}{L} C_{ox} (V_G - V_{th}) V_D \tag{4.1}$$

where the $0.5V_D^2$ term is neglected for small V_D, μ is the mobility of the minority charge carrier in the inversion channel, W/L is the channel width to channel length ratio, C_{ox} is the oxide capacitance, V_G is the gate voltage, V_{th} is the threshold voltage and V_D is the drain voltage of the transistor device.

After some re-arrangements, Eq. (4.1) could be rewritten as:

$$C_{ox}(V_G - V_{th}) = \frac{I_D}{\mu \frac{W}{L} V_D} = K_1 I_D \tag{4.2}$$

where K_1 is a constant for a fixed V_D.

The threshold voltage of a transistor is given by [12]:

$$V_{th} = V_{FB} + \gamma \sqrt{\phi_S} + \phi_S \tag{4.3}$$

where γ is the body effect parameter and ϕ_S ($=2\phi_F$) is the surface potential. In a MOS system, the charge in the gate (Q_G) would be balanced by the oxide trap charge (Q_{ot}) as well as the charge in silicon (Q_s).

$$Q_G(t) = -Q_{ot}(t) - Q_s(t) \quad (4.4)$$

Since the gate voltage is fixed during the discharge experiment, $Q_G(t)$ is independent of time, that is,

$$Q_G(t) = Q_G \quad (4.5)$$

At time t = 0,

$$Q_G = -Q_{ot}(0) - Q_s(0) \quad (4.6)$$

At time t = *t*,

$$Q_G = -Q_{ot}(t) - Q_s(t) \quad (4.7)$$

By equating Eq. (4.6) and Eq. (4.7), the expression in Eq. (4.8) could be obtained.

$$Q_{ot}(0) - Q_{ot}(t) = Q_s(t) - Q_s(0) \quad (4.8)$$

where

$$Q_{ot}(0) = qn_o \quad (4.9a)$$

$$Q_{ot}(t) = qn_o\exp(-e_n t)[13] \quad (4.9b)$$

The time-dependent silicon charge $Q_s(t)$ can be obtained from Eq. (4.2) as

$$Q_s(t) = C_{ox}[(V_G - V_{th}(t)] = K_1 I_D(t) \quad (4.10)$$

where V_G is time independent because it is fixed during the experiment. The threshold voltage $V_{th}(t)$ changes with time due to the detrapping of charges from the charge storage layer.

Using Eqs. (4.9a), (4.9b) and (4.10), Eq. (4.8) can be expressed as

$$qn_o[1-\exp(-e_n t)] = K_1[I_D(t) - I_D(0)] \quad (4.11)$$

If the argument in the exponential term in Eq. (4.11) is small (i.e., $e_n t << 1$), the following

$$\exp(-e_n t) \approx 1 - e_n t \quad (4.12)$$

Hence Eq. (4.11) can be simplified to

$$K_1[I_D(t) - I_D(0)] \approx e_n t \tag{4.13}$$

where $e_n = AT^2 \exp(-\frac{E_{trap}}{kT})$ [13].

After some mathematical re-arrangements, Eq. (4.13) can be simplified to Eq. (4.14):

$$\ln[\frac{I_D(t) - I_D(0)}{T^2}] = \ln K_2 - \frac{E_{trap}}{kT} \tag{4.14}$$

Using this approach, the E_{trap} level could be extracted based on the analysis of the change in the drain current at time t with respect to the initial state.

Figure 4.3 shows the drain current transient plots of memory devices with Al_2O_3 and HfO_2 charge storage layers at different temperatures. The increase in drain current at higher temperatures could be explained by the higher intrinsic carrier concentration of the substrate. The increased intrinsic carrier concentration reduced the Fermi level (ϕ_F) of the substrate. This caused the voltage drop at the substrate (~$2\phi_F$) during inversion to reduce which would enhance the vertical electric field across the gate stack oxide.

The drain current difference during discharging divided by the squared temperature (T^2) versus the inverse of temperature is shown in Fig. 4.4. From Fig. 4.4, it could be seen that the charge loss mechanism of Al_2O_3 devices had a stronger temperature dependence compared to that of HfO_2. The extracted trap energy levels were approximately 0.12 eV and 0.64 eV below the charge storage layer conduction bands for HfO_2 and Al_2O_3, respectively. The shallow traps in HfO_2 may be attributed to grain boundary defects as a result of crystallization [14, 15]. As will be shown subsequently, as-deposited HfO_2 was already crystallized. The extracted trap energy

level of Al_2O_3 was very close to that observed by Jonnard *et al.* (E_c – 0.6 eV) by electron-induced x-ray emission which was attributed to oxygen vacancies [16]. Hence the good charge retention performance of Al_2O_3 devices was probably due to deeper trap levels, as also shown by Sugizaki *et al.* [5].

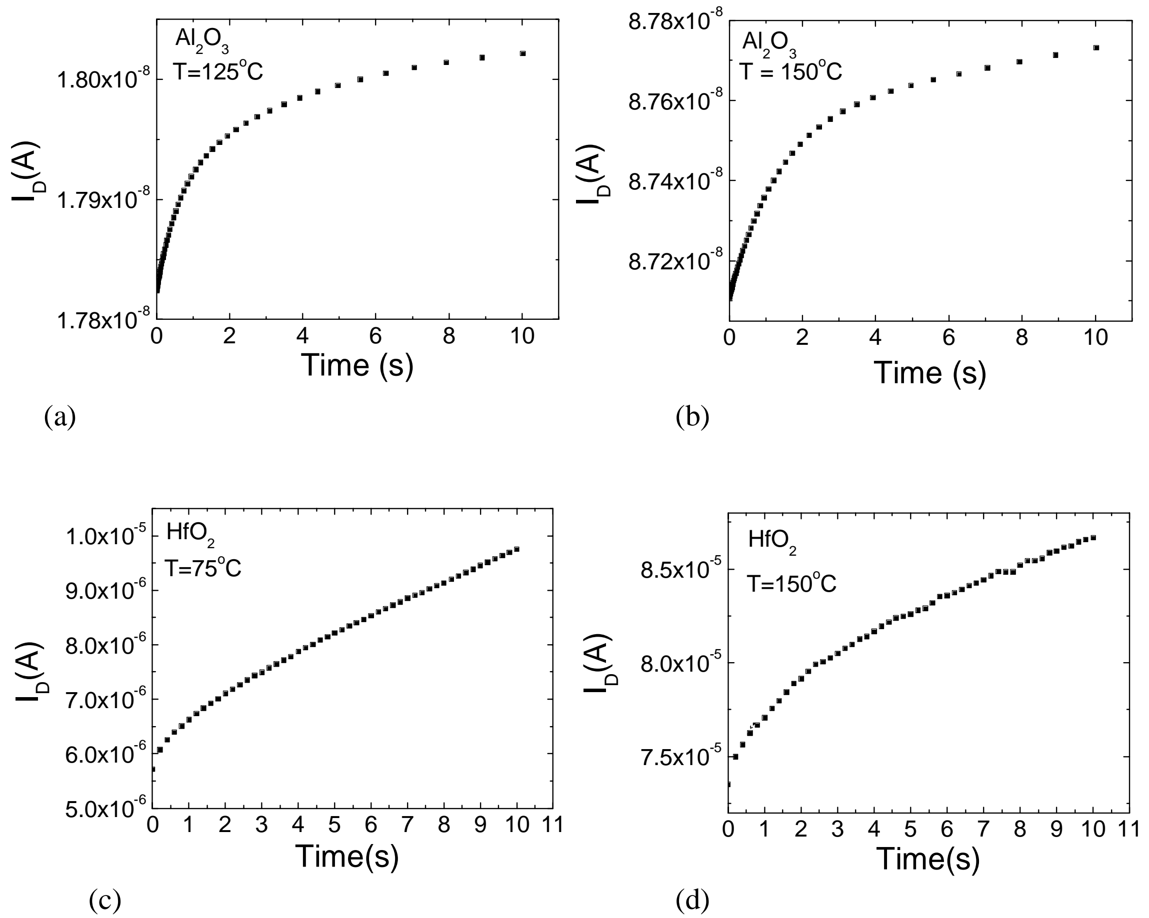


Figure 4.3: The drain current transients of (a), (b) Al_2O_3 memory devices and (c), (d) HfO_2 memory devices during the application of a read voltage after the application of a program voltage for 20s. The read and program voltages for Al_2O_3 devices were 3.3 V and 9 V, respectively. For HfO_2 devices the read and program voltages were 2.9 V and 7 V respectively.

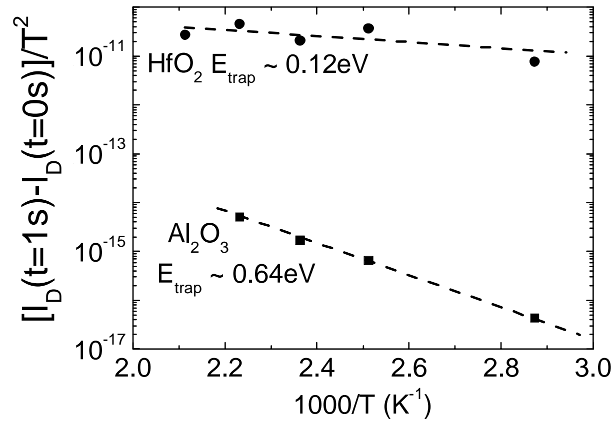


Figure 4.4: Drain current difference during discharging divided by squared temperature (T) versus the inverse of T for HfO_2 and Al_2O_3 memory devices.

The charge storage capability of the devices with good charge retention characteristics, namely HfAlO, Si_3N_4 and Al_2O_3 devices, were further investigated as shown in Fig. 4.5. The HfAlO structure showed a clear advantage over Al_2O_3 and was comparable to Si_3N_4. The charge storage capability was calculated from the C-V curves with counterclockwise hysteresis, assuming that the charge centroid was located at the charge storage layer/tunnel oxide interface. The relative concentration of the traps in the various materials approximated from the charge storage capability curves were $1.3\times10^{13}cm^{-2}$ for both Si_3N_4 and HfAlO and $2.5\times10^{12}cm^{-2}$ for Al_2O_3.

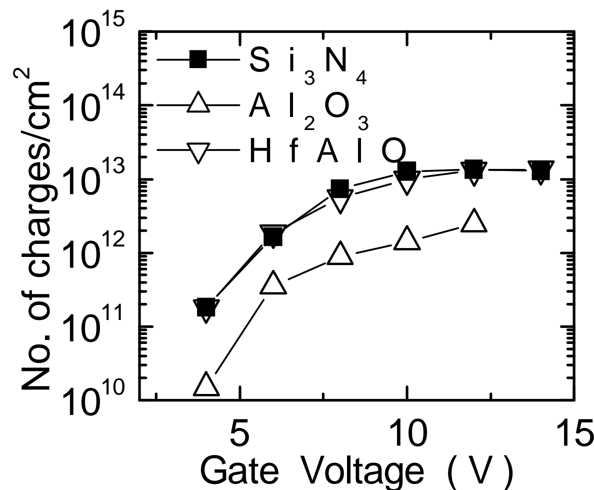


Figure 4.5: Density of stored charge, extracted from the hysteresis in the C-V curves, and plotted against the gate voltage sweep range for SONOS-type capacitor structures with Si_3N_4, Al_2O_3 or HfAlO as the charge storage layer.

Figure 4.6 shows the over-erase characteristics of the different devices. Both HfAlO and Al_2O_3 devices showed better over-erase performance than the Si_3N_4 device, with over-erase-free characteristics down to a negative gate voltage sweep of -8 V and -10 V for HfAlO and Al_2O_3 devices, respectively, as compared to -4 V for the Si_3N_4 device. From the programming and erasing characteristics, shown in Figs. 4.7(a) and (b) respectively, HfAlO devices showed the fastest programming and erase speed while the Al_2O_3 devices were the slowest. The program/erase (P/E) endurance characteristic of the HfAlO device showed no discernible difference from that of the Si_3N_4 device as shown in Fig. 4.8. Both Si_3N_4 and HfAlO devices showed negligible degradation in the threshold voltage (V_{th}) window after 10^4 P/E cycles.

HfAlO devices showed improved endurance characteristics as compared to HfO_2 devices, as discussed in the previous chapter. The high electric fields across the gate stack during program/erase operations could introduce some sort of permanent damage. Threshold window closing was usually associated with electron trapping in the tunnel oxide for floating gate Flash memory [17]. The trapped charges in the tunnel oxide of a floating gate Flash memory device would result in a decrease in the tunnel oxide electric field during programming and hence reduced the amount of charge transferred to the floating gate during programming. After erasing, the device threshold voltage would be higher than that of an unstressed device as the trapped electrons in the tunnel oxide of the floating gate Flash memory device would cause an increase in the threshold voltage. A similar mechanism may be proposed in the case of HfO_2 and HfAlO devices. In this case, the tunnel oxide quality was similar as the devices were fabricated at the same time. Hence, the only difference was the charge storage layer. Additional electron traps or negative defects may be generated at the tunnel oxide/charge storage layer interface or in the bulk of the charge storage layer

during programming. It is possible that electron trap or negative defect generation was reduced in the case of HfAlO devices as compared to the HfO_2 devices. Hence, threshold window closing occurred after 400 program/erase cycles for the HfO_2 devices while HfAlO devices showed negligible degradation in the threshold voltage window after 10^4 cycles.

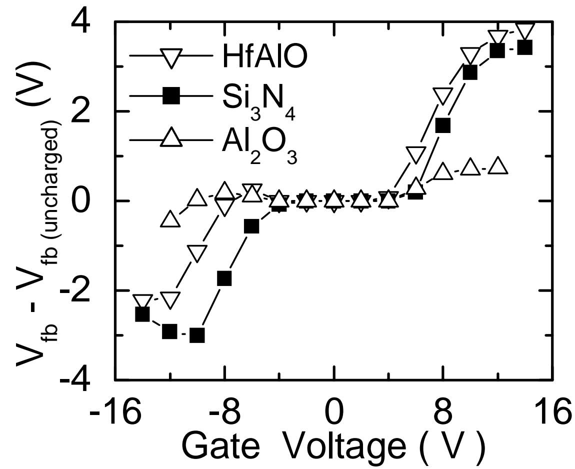


Figure 4.6: Flatband voltage shift plotted against the charging/discharging (program/erase) voltage extracted from the hysteresis in the C-V curves for memory capacitors with Si_3N_4, Al_2O_3 or HfAlO as the charge storage layer.

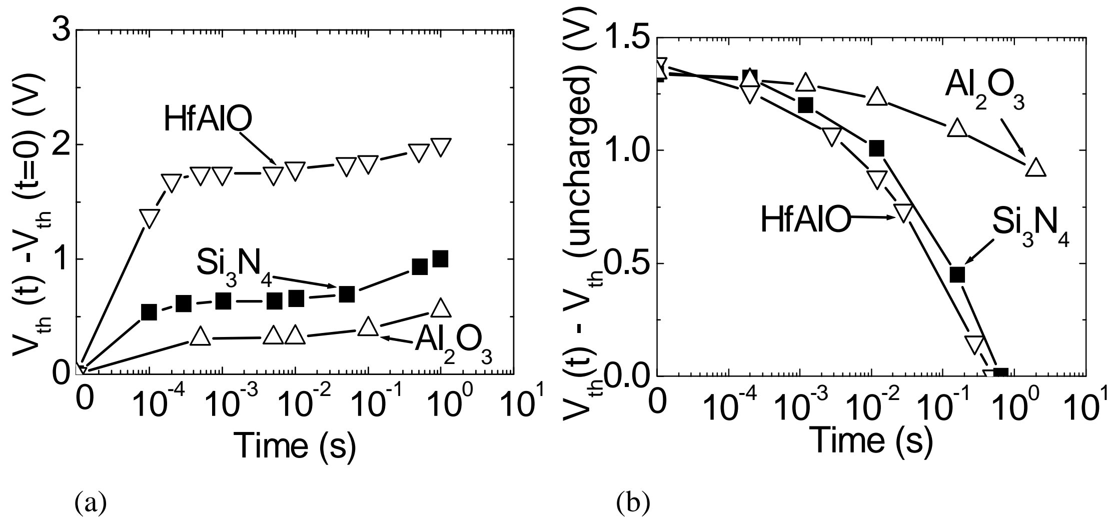


Figure 4.7: (a) Programming (V_g-V_{fb} = 6V) and (b) erasing (V_g-V_{fb} = –6V) characteristics of SONOS and SOHOS transistors with Si_3N_4, HfAlO and Al_2O_3 charge storage layers.

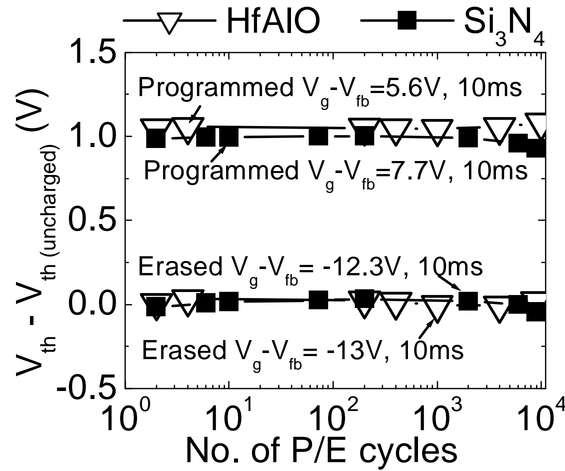


Figure 4.8: Program/Erase (P/E) endurance characteristics of SONOS and SOHOS transistors with Si_3N_4 and HfAlO charge storage layers.

The improvement in programming speed and over-erase characteristics of HfAlO is attributed to a suitable valence and conduction band offset with respect to silicon as illustrated in the schematic diagrams in Figs. 4.9 and 4.10. HfAlO with 10% Al_2O_3 has a similar band offset to pure HfO_2 [18, 19]. The conduction band offset between HfAlO with respect to Si is the smallest (1.63 eV), compared to 2 eV [20] for Si_3N_4 and 2.8 eV for Al_2O_3 [19]. Hence, at the same gate bias where modified Fowler-Nordheim tunneling dominates, the electron tunneling distance from the Si substrate to the conduction band of the storage dielectric is shortest in HfAlO and longest in Al_2O_3. This agrees well with the programming speed results in Fig. 4.7. During erase, both electron and hole tunneling are involved. However, there is a possibility that the over-erase problem is due to hole injection from Si substrate after all the stored electrons have tunneled back to the Si substrate. The valence band offset of Si_3N_4 with respect to Si is the smallest (2 eV) [20], compared to 3.3 eV for HfAlO [18] and 4.8 eV for Al_2O_3 [19]. Therefore, Si_3N_4 will start to experience hole tunneling and positive charge trapping at the lowest negative voltage, which causes

the negative shift in V_{fb}, while Al_2O_3 has hole tunneling at the highest negative voltage. This also agrees well with the over-erase result in Fig. 4.6.

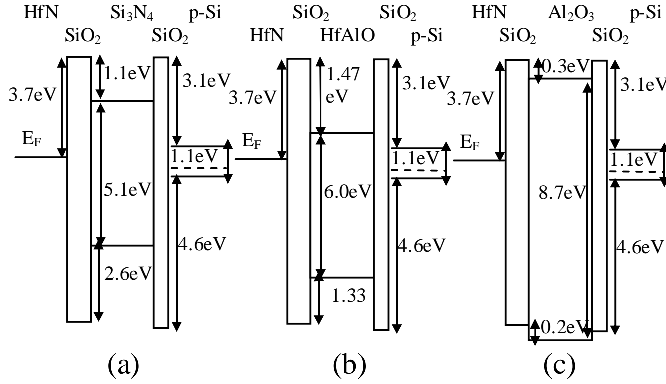


Figure 4.9: Ideal energy band diagrams of SONOS-type structures (HfN/TaN gate) with (a) Si_3N_4 (conventional SONOS), (b) HfAlO (10% Al_2O_3 concentration) and (c) Al_2O_3 as the charge storage layer.

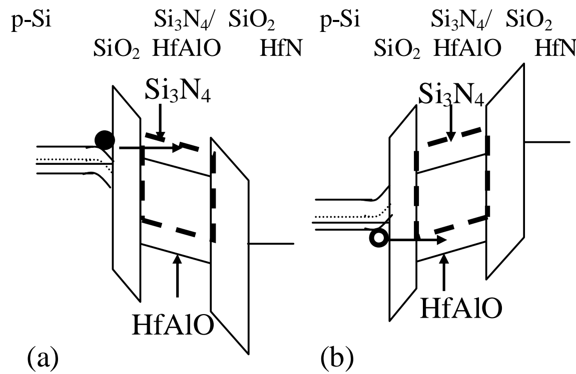


Figure 4.10: Energy band diagram schematic of SONOS-type structures with HfAlO (solid lines) or Si_3N_4 (dashed lines) as the charge storage layer during (a) write (program) and (b) erase operations.

The poor charge retention characteristics of pure HfO_2 can be attributed to the presence of conduction paths in the high-κ material [7]. As seen from the x-ray diffraction (XRD) results in Figs. 4.11(a) and (b), the pure HfO_2 film was fully crystallized, while HfAlO was still amorphous even after annealing at 800°C for 60 s and only slightly crystallized after 1000°C annealing. From Fig. 4.2, Si_3N_4, Al_2O_3 and

HfAlO SONOS-type structures showed significantly better charge retention performance than the HfO_2 device. This is because the Si_3N_4, Al_2O_3 and HfAlO charge storage layers were still relatively amorphous while the HfO_2 charge storage layer was already crystallized. Polycrystallization of thin films will generate grain boundaries which could act as current leakage paths. Hence there would be an increase in lateral conduction which could result in poorer charge retention [7]. The retention performance of the HfAlO device was only slightly worse than the Al_2O_3 device as the HfAlO film was partially crystallized due to the source/drain annealing at 950°C for 30 s.

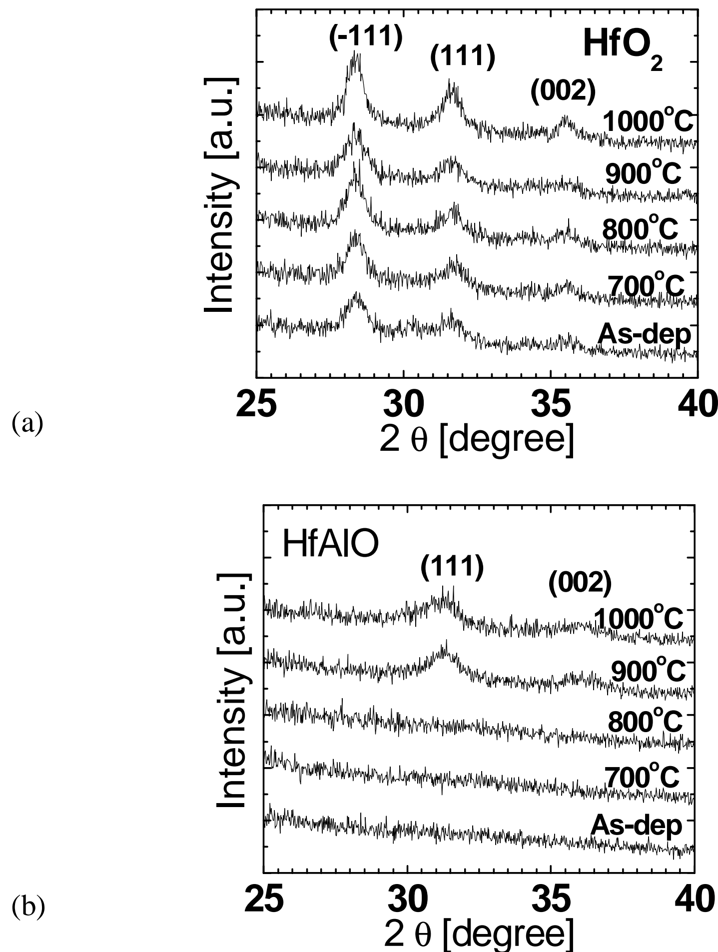


Figure 4.11: XRD spectra of (a) HfO_2 and (b) HfAlO. As-deposited HfO_2 was already crystallized while HfAlO remained amorphous up to 800°C.

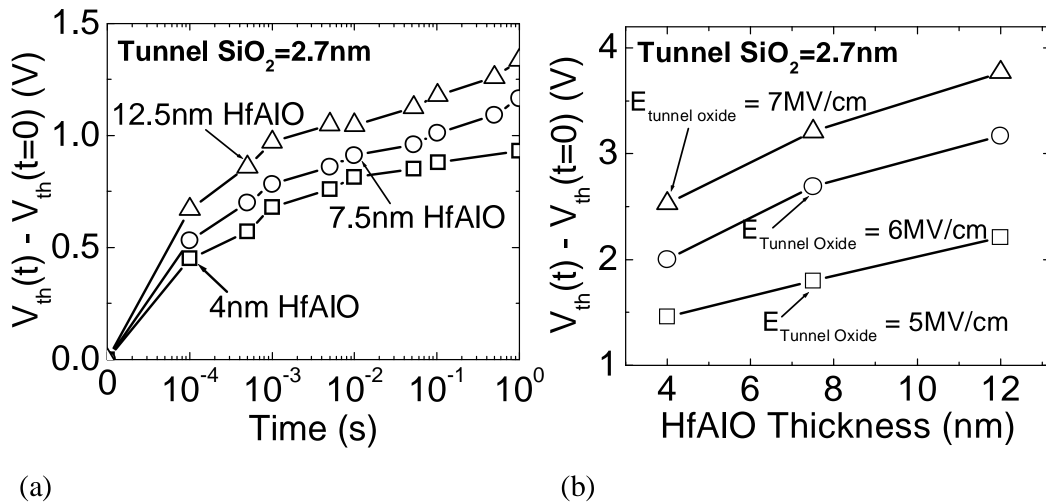
In order to further investigate the charging mechanism in the SOHOS HfAlO structure, SOHOS memory transistors with different thickness of HfAlO charge storage layer, as shown in Table 4.1, were evaluated. At a given tunnel oxide thickness, the threshold voltage shift increased with increasing HfAlO thickness, as shown in Figs. 4.12(a) and 4.12(b), indicating that the amount of negative trapped charge increased with increasing HfAlO thickness. This suggested that the dominant charge storage mechanism was due to electron trapping in the bulk of the HfAlO layer, rather than negative charge trapping at the tunnel oxide/high-κ interface which would be independent of the HfAlO thickness.

Table 4.1: The split conditions of samples with different HfAlO charge storage layer thicknesses, different tunnel oxide thickness and 65 Å blocking oxide. The cell structure is similar to Fig. 4.1.

Wafer No.	**SiO_2 tunnel oxide (Å)**	**Charge storage layer**
1	27	75 Å Si_3N_4
2	27	40 Å HfAlO
3	27	75 Å HfAlO
4	27	125 Å HfAlO
5	34	75 Å HfAlO

Figure 4.13 shows the charge retention characteristics of SOHOS devices with different HfAlO charge storage layer thicknesses. Charge retention measurements were carried out for durations of up to 10^4 s. After about 10^3 s, the V_{th} had a logarithmic decay in time. Hence, assuming that the V_{th} decay followed a constant rate, the V_{th} was extrapolated to 10 years. The charge retention performance of the SOHOS device also degraded with decreasing HfAlO thickness as shown in Fig. 4.13. This can be understood from the fact that for SOHOS devices with a thicker HfAlO layer, electrons that were trapped within the bulk may had to tunnel through a longer

distance through the HfAlO layer to the tunnel SiO_2 and silicon substrate. HfAlO SOHOS transistors with a thicker tunnel SiO_2 were compared to Si_3N_4 SONOS transistor with a thinner tunnel SiO_2 in Figs. 4.14(a) and (b) for the purpose of demonstrating better charge retention performance. From the threshold voltage shift ($V_{th}(t)$ - $V_{th}(t=0)$) with respect to programming time in Fig. 4.14(a), it is seen that the SOHOS devices still showed a faster programming speed than the SONOS device, even though the former had a thicker tunnel oxide than the latter. For the erasing characteristics in Fig. 4.14(b), the V_{th} (t=0) denoted the V_{th} of an uncharged device. Hence, the devices were programmed to a V_{th} shift of 0.8 V before erasing. However, the SOHOS devices had a much slower erasing speed even though they show better charge retention because of the thicker tunnel oxide. During erasing, electrons had to tunnel through the tunnel oxide layer to the silicon substrate by direct tunneling, which was greatly affected by the tunnel oxide thickness [2].



(a) (b)

Figure 4.12: (a) Programming characteristics (i.e., threshold voltage shift versus time at a tunnel oxide field of 5 MV/cm) of SOHOS transistors with 40 Å, 75 Å and 125 Å thick HfAlO charge storage layer and 27 Å thick tunnel oxide. (b) Threshold voltage shift of SOHOS transistors after 50 s programming versus thickness of the HfAlO charge storage layer for tunnel oxide fields of 5, 6 and 7 MV/cm during programming. The tunnel oxide is 27 Å thick. The traps are saturated after 50s programming.

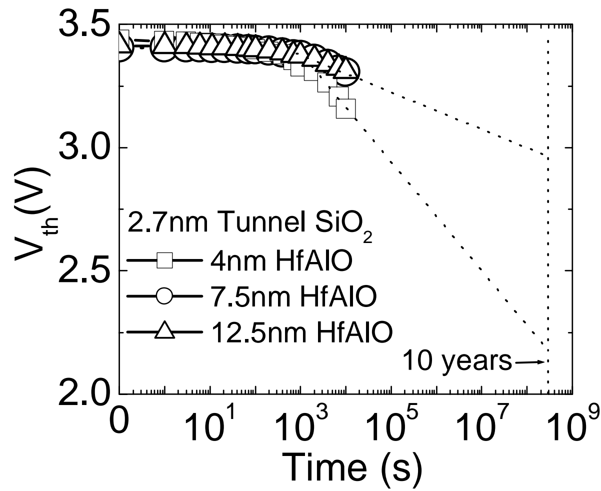


Figure 4.13: Charge retention characteristics (i.e., threshold voltage versus time) of SOHOS transistors with HfAlO charge storage layer of 40 Å, 75 Å and 125 Å thickness and 27 Å tunnel oxide performed at $V_g = 0V$ with source/drain and substrate grounded.

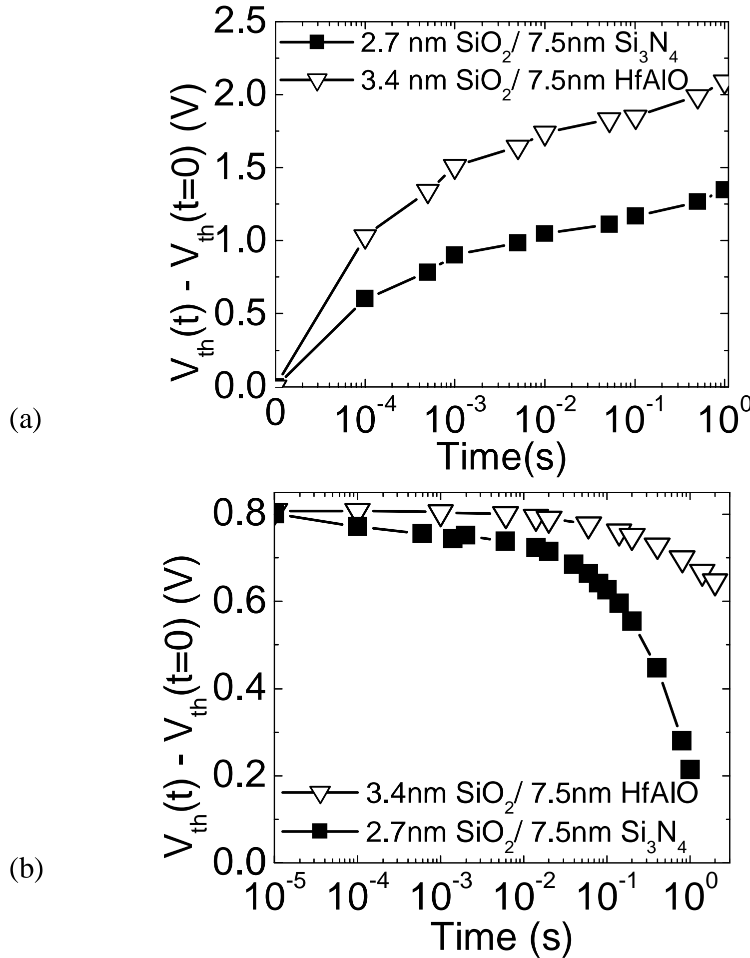


Figure 4.14: (a) Programming ($V_g - V_{fb} = 8.5V$) and (b) erasing ($V_g - V_{fb} = -15V$) characteristics of threshold voltage shift versus time of SONOS transistor with 27 Å tunnel SiO_2/75 Å Si_3N_4 charge storage layer and SOHOS transistor with 34 Å tunnel SiO_2/75Å HfAlO charge storage layer. $V_{th}(t=0)$ denoted the V_{th} of uncharged device.

As the programming speed and charge retention have a trade-off relationship in Flash memory, both parameters were evaluated together in Fig. 4.15 and Table 4.2. The V_{th} decay rate per decade is the V_{th} shift per decade of measurement time during charge retention measurement. This was taken at the later stages of retention measurement (from 10^3 to 10^4 s), when the V_{th} shift with respect to time followed a logarithmic decay. In comparison to other SONOS devices from bench-marked data, SOHOS with HfAlO charge storage layer showed a clear improvement in performance, considering both programming speed and charge retention.

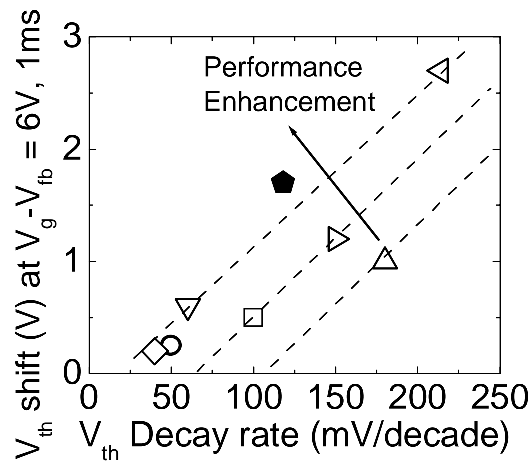


Figure 4.15: Graph of V_{th} shift after programming at V_g-V_{fb} = 6V, 1ms against the V_{th} decay rate per decade of retention measurement time. Comparison between this work (HfAlO device) and published data (refer to Table 4.2).

Table 4.2: Comparison between this work (HfAlO device) and published data. SRO is silicon rich oxide.

Symbol	Dielectric structure	V_{th} Decay rate in mV/decade	V_{th} shift at V_g-V_{fb} = 6V, 1ms (V)	V_{th} shift at V_g-V_{fb} = -6V, 1ms (V)	Tunnel oxide thickness (Å)	References
⬟	$SiO_2/HfAlO/SiO_2$	118	1.7V	- 0.5V	25	This work
○	$HfO_2/Ta_2O_5/HfO_2$	50	0.25V	-0.3V	48 (HfO_2)	[21]
△	SiO_2/SRO/ SiO_2	180	1V	-1V	25	[22]
▽	$SiO_2/Si_3N_4/$ SiO_2	60	0.6V	-1.4V	15	[23]
◇	$SiO_2/Si_3N_4/$ SiO_2	40	0.2V	-0.6V	23	[24]
◁	$SiO_2/Si_3N_4/$ SiO_2	214	2.7V	-3V	15	[25]
▷	$SiO_2/Si_3N_4/$ SiO_2	150	1.2V	-3V	18	[2]
□	$SiO_2/Si_3N_4/$ SiO_2	100	0.5V	-	22	[26]

4.4 Summary

SOHOS memory with HfAlO charge storage layer was demonstrated. HfAlO SOHOS showed faster programming speed than conventional SONOS, together with good charge retention and program/erase endurance characteristics. Therefore, the SiO_2/HfAlO/SiO_2 gate insulator stack structure has attractive advantages for Flash memory application. An alternative method to increase the program and erase speed is to use high-κ material as blocking oxide, which will be discussed in the next chapter.

References

[1] B. Jiankang and M. H. White, "Design considerations in scaled SONOS nonvolatile memory devices", *Solid State Electronics*, vol. 45, pp. 113-120, 2001.

[2] M. L. French and M. H. White, "Scaling of multidielectric nonvolatile SONOS memory structures", *Solid State Electronics*, vol. 37, pp. 1913-1923, 1994.

[3] M. L. French, C. Y. Chen, H. Sathianathan and M. H. White, "Design and scaling of a SONOS multidielectric device for nonvolatile memory applications", *IEEE Trans. Components, Packaging and Manufacturing Technology*-Part A, vol. 17, pp. 390-397, 1994.

[4] Y. N. Tan, W. K. Chim, B. J. Cho, W. K. Choi, "Over-erase phenomenon in SONOS-type Flash memory and its minimization using a hafnium oxide charge storage layer", *IEEE Trans. Electron Devices*, vol. 51, pp. 1143-1147, 2004.

[5] T. Sugizaki, M. Kobayashi, M. Ishidao, H. Minakata, M. Yamaguchi, Y. Tamura, Y. Sugiyama, T. Nakanishi and H. Tanaka, "Novel multibit SONOS type Flash memory using a high-κ charge trapping layer", in *VLSI Tech. Symp.*, 2003, p. 27.

[6] Y. N. Tan, W. K. Chim, W. K. Choi, M. S. Joo, T. H. Ng and B. J. Cho, "High-k HfAlO charge trapping layer in SONOS-type nonvolatile memory device for high speed operation", in *IEEE IEDM Tech. Dig*., 2004, p. 889.

[7] M. S. Joo, B. J. Cho, C. C. Yeo, D. S. H. Chan, S. J. Whoang, S. Matthew, L. K. Bera, N. Bala and D. L. Kwong, "Formation of hafnium-aluminum-oxide

gate dielectric using single cocktail liquid source in MOCVD process", *IEEE Trans. Electron Devices*, vol. 50, pp. 2088-2094, 2003.

[8] H. Y. Yu, H. F. Lim, J. H. Chen, M. F. Li, C. X. Zhu, D. L. Kwong, C.H. Tung, K. L. Bera and C. J. Leo, "Robust HfN metal gate electrode for advanced MOS devices application", in *VLSI Tech. Symp.*, 2003, p.151.

[9] H. Y. Yu, M. F. Li and D. L. Kwong, "Thermally robust HfN metal as a promising gate electrode for advanced MOS device applications", *IEEE Trans. Electron Devices*, vol. 51, pp. 609-615, 2004.

[10] H. Y. Yu, J. F. Kang, Ren C., J. D. Chen, Y. T. Hou, Shen, C., M.F. Li, D. S. H. Chan, K. L. Bera, C. H. Tung, D. L. Kwong, "Robust high-quality HfN-HfO_2 gate stack for advanced MOS device applications", *IEEE Electron Device Lett.*, vol. 25, pp. 70-72, 2004.

[11] H. Y. Yu, H. F. Lim, J. H. Chen, M. F. Li, Zhu Chunxiang, C. H. Tung, A. Y. Du, W. D. Wang, D. Z. Chi, D. L. Kwong, "Physical and electrical characteristics of HfN gate electrode for advanced MOS devices," *IEEE Electron Device Lett.*, vol. 24, pp. 230-232, 2003.

[12] N. Arora, "MOSFET Models for VLSI Circuit Simulation, Theory and Practice", *Springer-Verlag*, chapter 6, 1992.

[13] Y. L. Yang and M. H. White, "Charge retention of scaled SONOS nonvolatile memory devices at elevated temperatures", *Solid State Electronics*, vol. 44, pp. 949-958, 2000.

[14] N. Zhan, K. L. Ng, H. Wong, M. C. Poon, C. W. Kok, "Effects of rapid thermal annealing on the interface and oxide trap distributions in hafnium oxide films", in *IEEE Conf. Electron Devices and Solid State Circuits*, 2003, p. 431.

[15] H. Wong, "Material and interface instabilities of hafnium gate oxide", in *Proc. International Conf. Solid State and Integrated Circuits Technology*, 2004, p. 378.

[16] P. Jonnard, F. Vergand, M. Kefi and C. Bonnelle, "Electron trapping in α-alumina observed by electron-induced x-ray emission", *J. Appl. Phys.*, vol. 79, pp. 2909-2912, 1996.

[17] "Nonvolatile Semiconductor Memory Technology, A Comprehensive Guide to Understanding and Using NVSM Devices", Edited by William D. Brown and Joe E. Brewer, IEEE Press, pp. 69-73, 1998.

[18] H. Y. Yu, M. F. Li, B. J. Cho, C. C. Yeo, M. S. Joo, D. L Kwong, J. S. Pan, C. H. Ang, J. Z. Zheng, S. Ramanathan, "Energy gap and band alignment for $(HfO_2)_x(Al_2O_3)_{1-x}$ on (100) Si," *Appl. Phys. Lett.*, vol. 81, pp. 376-378, 2002.

[19] G. D. Wilk, R. M. Wallace, J. M. Anthony, "High-κ gate dielectrics: Current status and materials properties considerations", *J. Appl. Phys.*, vol. 89, pp. 5243-5275, 2001.

[20] M. H. White, L. Yang, A. Purwar and M. L. French, "A low voltage SONOS nonvolatile semiconductor memory technology", *IEEE Trans. Components, Packaging and Manufacturing Technology*, Part A, vol. 20, pp. 190-195, 1997.

[21] X. Wang, J. Liu, W. Bai and D. L. Kwong, "A novel MONOS-type nonvolatile memory using high-κ dielectrics for improved data retention and programming speed", *IEEE Trans. Electron Devices*, vol. 51, pp. 597-602, 2004.

[22] M. Rosmeulen, E. Sleeckx and K. De Meyer, "Silicon-rich-oxides as an alternative charge-trapping medium in Fowler-Nordheim and hot carrier type non-volatile-memory cells", in *IEEE IEDM Tech. Dig*., 2002, p. 189.

[23] Y. K. Lee, S. K. Sung, J. S. Sim, C. J. Lee, T. H. Kim, S. H. Lee, J. D. Lee, B. G. Park, D. H. Lee and Y. W. Kim, "Multi-level vertical channel SONOS nonvolatile memory on SOI", in *VLSI Tech. Symp*., 2002, p. 208.

[24] M. Tanaka, S. Saida, Y. Mitani, I. Mizushima and Y. Tsunashima, "Highly reliable MONOS devices with optimized silicon nitride film having deuterium terminated charge traps", in *IEEE IEDM Tech. Dig.*, 2002, p. 237.

[25] M. H. White, D. A. Adams, B. Jiankang, "On the go with SONOS", *IEEE Circuits and Devices*, pp. 22-31, 2000.

[26] P. Olivo, Z. A. Weinberg, K. J. Stein and D. S. Wen, "Charge trapping and retention in ultra-thin oxide-nitride-oxide structures", *Solid State Electronics*, vol. 34, pp. 609-611, 1991.

Chapter 5

Development of High-κ Blocking Oxide Layer in SONOS-type Nonvolatile Memory

5.1 Introduction

The applications of digital electronics have resulted in a strong demand for nonvolatile memories that are densely integrated, fast and consume little power. Charge trapping memories such as SONOS (polysilicon-oxide-silicon nitride-oxide-silicon) device is an attractive candidate to realize Flash memory vertical scaling. In the previous chapters, we have shown that programming speed can be increased without reducing the tunnel oxide thickness through appropriate choice of the material for the charge storage layer. An alternative method to increase program/erase speed without decreasing the tunnel oxide thickness is by using a high-κ material as the blocking oxide [1-4] which will be demonstrated in this section. From electrostatics consideration, the use of a high dielectric constant blocking oxide layer will cause a smaller voltage drop across the blocking oxide and greater voltage drop across the tunnel oxide. This will result in a simultaneous increase of the electric field across the tunnel oxide and reduction of the electric field across the blocking oxide, leading to more efficient program and erase processes [1-4].

The purpose of the blocking oxide is to prevent charge transfer between the charge storage layer and the gate electrode during program/erase processes. Hence, during programming, the blocking oxide prevents both electron out-tunneling from the charge storage layer to the gate electrode and hole injection from the gate electrode to the charge storage layer. During erasing, the blocking oxide prevents both

electron injection from the gate electrode to the charge storage layer and hole tunneling from the charge storage layer to the gate electrode. Ideally, the blocking oxide should have a high κ value and large conduction and valence band offsets with respect to the charge storage layer. In this chapter, various materials with suitable properties to be used as blocking oxides for SONOS structures were evaluated. These are hafnium aluminum oxide, lanthanum aluminum oxide and lanthanum yttrium aluminum oxide.

5.2 Hafnium Aluminum Oxide Blocking Oxide Layer in SONOS-type Nonvolatile Memory for High Speed Operation

5.2.1 Introduction

Al_2O_3 has a large energy band gap value of 8.9 eV, a large conduction band offset with respect to silicon of 2.8 eV and a dielectric constant (κ) value of 9 making it an attractive candidate as a blocking oxide [5]. On the other hand, HfO_2 has a relatively smaller band gap of 5.7 eV, a smaller conduction band offset with respect to silicon of 1.5 eV but a much higher κ value of 25 [5]. Although HfO_2 has a smaller band gap, the electric field across the blocking oxide is much reduced due to its higher κ value while the electric fields across the tunnel SiO_2 and charge storage layers during program and erase are increased. Hence there is a trade-off between energy gap and κ value for the blocking oxide. It has been shown previously that the κ value, band gap energy and hence band offset with respect to silicon of HfAlO are proportional to the relative concentration of HfO_2 and Al_2O_3 [6]. It would be interesting to investigate the effect of the κ value and band gap energy of the blocking oxide layer on the program/erase speed and charge retention of SONOS devices. The

relative concentration of HfO_2 and Al_2O_3 in HfAlO can be selected by varying the number of deposition cycles in the atomic-layer-deposition (ALD) system.

5.2.2 Sample Fabrication

25 or 40 Å thick tunnel oxide was thermally grown at 800°C on 4-8 Ω-cm (100) p-type silicon substrates. Subsequently, 50 or 70 Å Si_3N_4 was deposited by low pressure chemical-vapor-deposition (LPCVD). 75 or 120 Å thick blocking oxide layer, consisting of either pure HfO_2, Al_2O_3 or HfAlO film, was deposited by ALD. For the control devices, 75 Å blocking oxide was deposited by LPCVD TEOS ($Si(OC_2H_5)_4$). Lastly, either HfN/TaN or TaN metal gate was formed by physical vapor deposition for the control gate [7-10]. The resulting structures are shown in Fig. 5.1. HfN gate is one of the processes developed in our laboratory. The purpose of HfN is to block oxidation of the high-κ/Si interface. Since the tunnel oxide is SiO_2 in this case, the HfN blocking layer is not necessary. In addition, comparisons and analysis were made among devices using the same gate electrode material. Hence, the use of different gate electrode materials was not expected to affect the results significantly. The transistors undergo source/drain implantation followed by activation annealing at 950°C for 30 s. The transistor structures tested have gate width (W) to gate length (L) dimensions of W/L = 100 μm/20 μm.

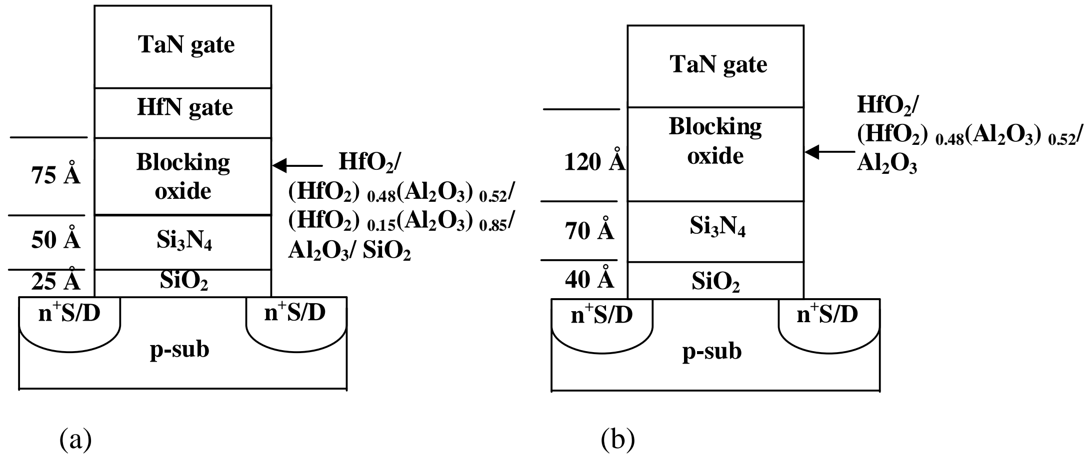


Figure 5.1: (a) Fabricated SONOS Flash transistor structures with HfN/TaN gate electrode. The blocking oxide layer is either SiO_2 or high-κ dielectric. (b) Fabricated SONOS Flash transistor structures with TaN gate electrode. The blocking oxide layer is high-κ dielectric.

5.2.3 Results and Discussion

In order to evaluate the use of HfAlO with different HfO_2 and Al_2O_3 compositions as a blocking oxide layer, SONOS devices with different blocking oxide layers were fabricated according to the device structure shown in Fig. 5.1(a). X-ray photoelectron spectroscopy (XPS) was performed on two samples with $(HfO_2)_x(Al_2O_3)_{1-x}$ (abbreviated as HfAlO) blocking layers of different HfO_2 concentration. The XPS spectra for Al 2*p*, O 1*s* and Hf 4*f* were shown in Figs. 5.2(a), (b) and (c), respectively. The Hf atomic percentage = x/(5-2x) and the Al atomic percentage = 2(1-x)/(5-2x) were determined from the intensities of the XPS lines [6]. The HfO_2 concentration was calculated to be 0.15 and 0.48 for the two samples with HfAlO blocking layer.

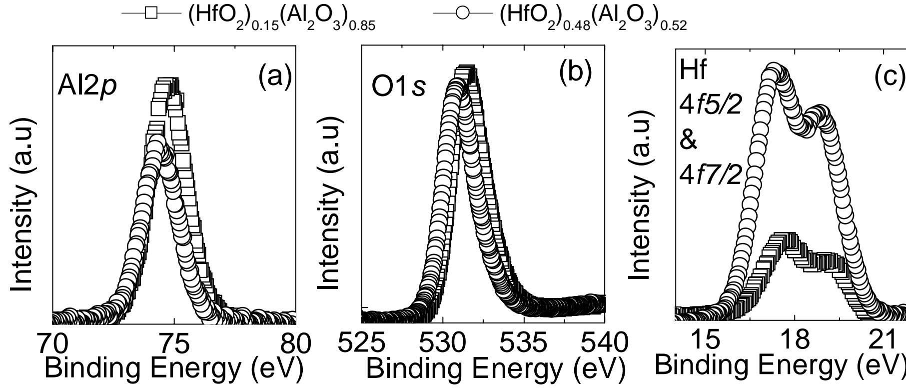


Figure 5.2: XPS spectra for (a) Al 2*p* core levels, (b) O 1*s* core levels and (c) Hf 4*f* core levels taken from $(HfO_2)_x(Al_2O_3)_{1-x}$ samples (used in the blocking oxide layer), with x values determined to be 0.15 and 0.48.

The programming transients of SONOS devices with high-κ and SiO_2 blocking oxide layers were shown in Figs. 5.3(a), (b) and (c) for V_g-V_{fb} = 6V, 7V and 9V, respectively. V_g and V_{fb} were the gate voltage and flatband voltage, respectively, while V_g-V_{fb} was the programming gate voltage after accounting for the flatband voltage. It could be observed that using a high-κ blocking oxide instead of SiO_2 increased the programming speed of the SONOS device significantly as seen from the faster increase in the threshold voltage, $V_{th}(t)$-$V_{th}(t=0)$, with time. However, the relationship between improvement in programming speed and HfO_2 concentration varied with the programming gate voltage. At low programming gate voltage (i.e., 6V), the programming speed increased with increasing HfO_2 concentration. However, at higher programming gate voltage such as 9V, the programming speed actually decreased with increasing HfO_2 concentration instead. The results were summarized in Fig. 5.4, which showed the threshold voltage shift after programming at V_g-V_{fb} = 6V, 7V, 8V and 9V for 100 μs for SONOS devices with HfAlO blocking oxide layer with different HfO_2 mole fraction x.

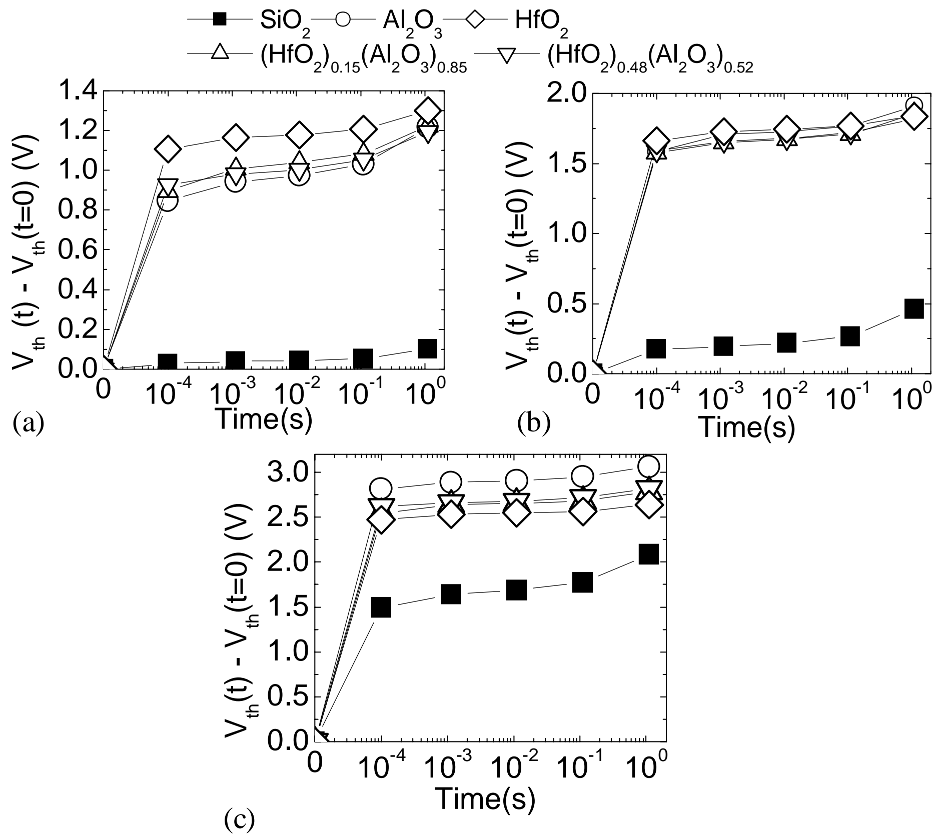


Figure 5.3: Programming transient for (a) V_g - V_{fb} = 6V (b) V_g - V_{fb} = 7V and (c) V_g - V_{fb} = 9V for SONOS devices with SiO_2 (solid symbol) or high-κ (open symbols) blocking oxide layers. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a).

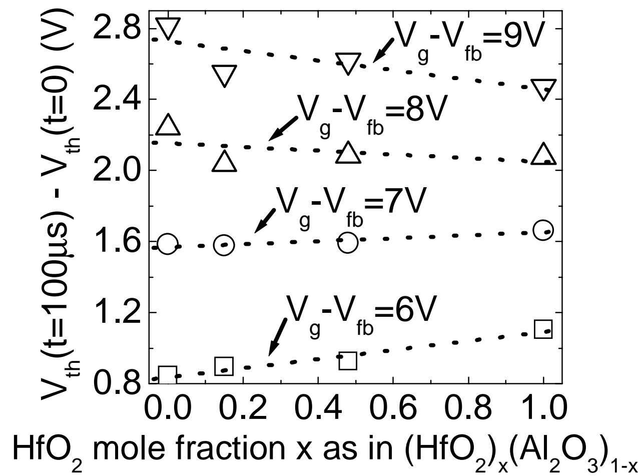


Figure 5.4: Threshold voltage shift after programming at V_g - V_{fb} = 6V, 7V, 8V and 9V for 100 μs for SONOS devices with HfAlO blocking oxide layer with different HfO_2 mole fraction x. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a).

The use of a high-κ blocking oxide increased the electric field across the tunnel oxide and charge storage layers and at the same time decreased the electric field across the blocking oxide layer [1-4]. Hence this increased the programming speed for SONOS devices with high-κ blocking oxide layer as compared to that with SiO_2 blocking oxide, the results of which were shown in Fig. 5.3. Schematic energy band diagrams for SONOS devices with Al_2O_3 and HfO_2 blocking oxide layers in the program mode for the low gate voltage case (e.g., 6V) were shown in Figs. 5.5(a) and (b), respectively. For simplicity and ease of explanation, only HfO_2 and Al_2O_3 cases are illustrated. The schematic energy band diagram for devices with HfAlO blocking oxide layers will be intermediate between that with Al_2O_3 and HfO_2 blocking oxide layers. Similar schematic energy band diagrams for SONOS devices with Al_2O_3 and HfO_2 blocking oxide layers for high program voltage (> 7V) situations were shown in Figs. 5.5(c) and (d), respectively. Increasing the κ-value of the blocking oxide (i.e., increasing HfO_2 content in this case) resulted in an increase in electric field across the tunnel oxide and charge storage layers. This would result in an increase in the programming speed especially at low program voltages. However, as illustrated in Figs. 5.5(c) and (d), for high program gate voltage, some of the electrons injected into the Si_3N_4 charge storage layer may tunnel out through the blocking oxide into the gate electrode. Increasing the HfO_2 percentage of the $(HfO_2)_x(Al_2O_3)_{1-x}$ layer resulted in a decrease of the band gap value and conduction band offset with respect to silicon, which decreased the effectiveness in preventing electron out-tunneling at higher positive gate voltages.

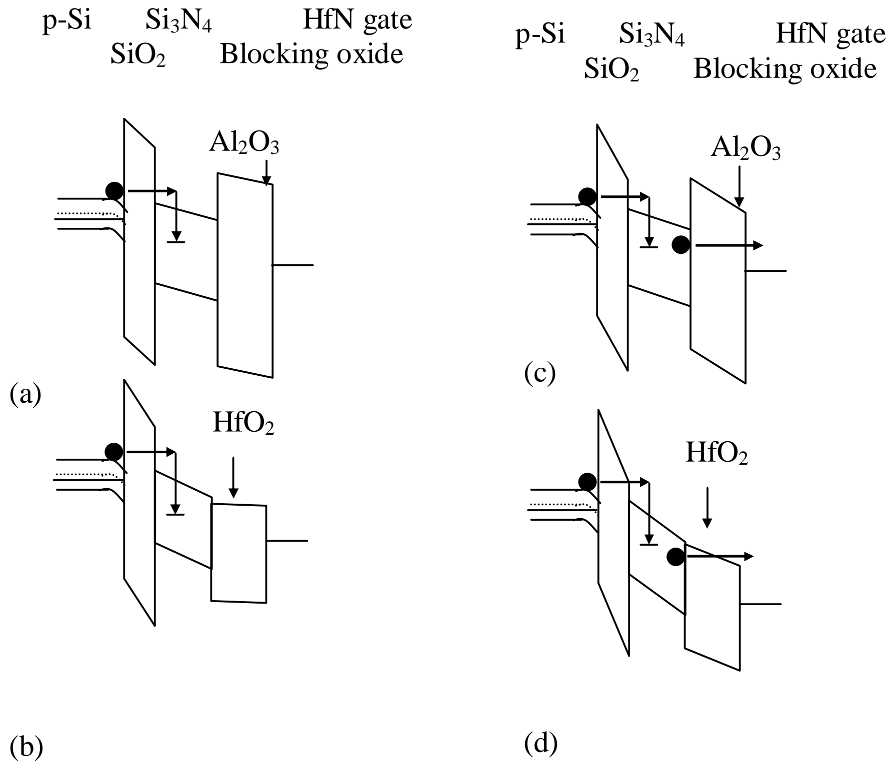


Figure 5.5: Schematic energy band diagrams for SONOS devices with Al_2O_3 [(a) and (c)] and HfO_2 [(b) and (d)] blocking oxide layers in the program mode for low [(a) and (b)] and high [(c) and (d)] gate voltage situations.

Figures 5.6(a), (b) and (c) showed the erasing transients of the SONOS devices with high-κ and SiO_2 blocking oxide layers for erase voltages (after accounting for the flatband voltage) V_g-V_{fb} = -6V, -7V, and -8V, respectively. It can be seen that SONOS devices with high-κ blocking oxide layers had faster erase speed than that with SiO_2 blocking oxide, especially at low gate erase voltages. The erase speed generally increased with increasing dielectric constant of the blocking oxide layer. However, at high erase voltages, the erase threshold voltage shift, V_{th}(t)-V_{th}(uncharged), would saturate for SONOS devices with high-κ blocking oxide layers. The onset of saturation occurred at lower erase gate voltages for devices with higher κ value or smaller band gap blocking layer (i.e., devices with blocking layers containing more HfO_2 content). SONOS devices with SiO_2 blocking oxide layer

showed no erase threshold voltage shift saturation for the range of erase gate voltages investigated.

(a)

(b)

(c)

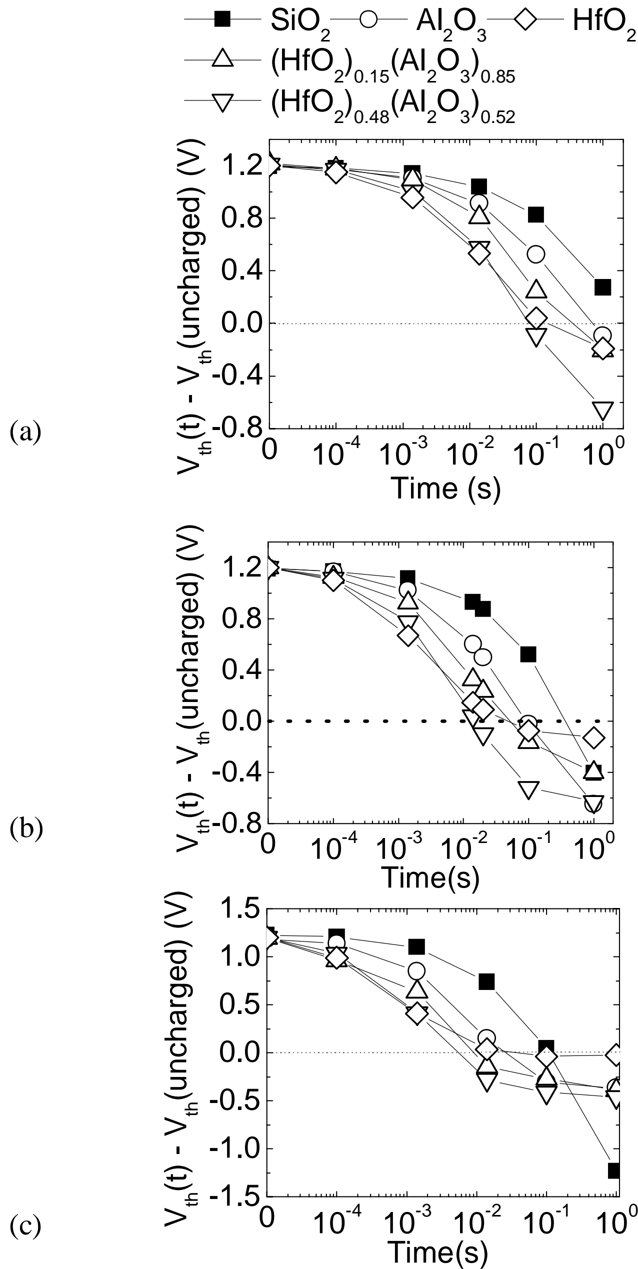


Figure 5.6: Erasing transient for (a) $V_g - V_{fb} = -6V$ (b) $V_g - V_{fb} = -7V$ and (c) $V_g - V_{fb} = -8V$ for SONOS devices with SiO_2 (solid symbol) or high-κ (open symbols) blocking oxide layers. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a).

Schematic energy band diagrams comparing SONOS devices with SiO_2 (solid lines) and high-κ (e.g., Al_2O_3) (dashed lines) blocking oxide layers in the erase mode were shown in Fig. 5.7(a). Similar to the program mode, the use of a high-κ blocking oxide layer increased the electric field across the tunnel SiO_2 and Si_3N_4 charge storage layers and decreased the electric field across the blocking oxide layer during erase operation [1-4]. Hence SONOS devices with high-κ blocking oxide had higher erase speed as compared to that with SiO_2, especially at low erase voltages as shown in Fig. 5.6(a). However, at high erase gate voltages, SONOS devices with high-κ blocking oxide experienced erase threshold voltage shift saturation, especially for devices with blocking layers containing more HfO_2 content as seen in Fig. 5.6(c). This was probably due to electron injection from the negatively biased gate electrode during the erase operation [2], as illustrated in Fig. 5.7(b) for the device with HfO_2 as the blocking oxide. Increasing HfO_2 concentration would result in an increase in the dielectric constant and a decrease in the band gap value of the $(HfO_2)_x(Al_2O_3)_{1-x}$ layer. Hence, this decreased its blocking capability against gate electron injection during erasing at high negative gate voltages.

Even though $(HfO_2)_x(Al_2O_3)_{1-x}$ blocking oxide showed some degree of erase saturation, the endurance test result of $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ blocking oxide device showed reasonably good endurance characteristics lasting to more than 100,000 program/erase cycles (Fig. 5.8).

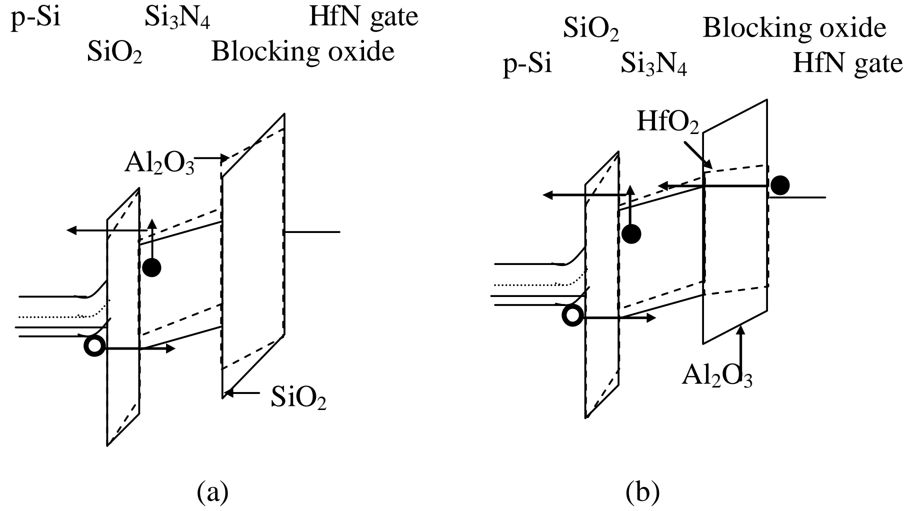


Figure 5.7: Schematic energy band diagrams for SONOS devices in the erase mode: (a) Comparing SiO_2 (solid lines) and high-κ (e.g., Al_2O_3) (dashed lines) blocking oxide layers, and (b) Comparing Al_2O_3 (solid lines) and HfO_2 (dashed lines) blocking oxide layers.

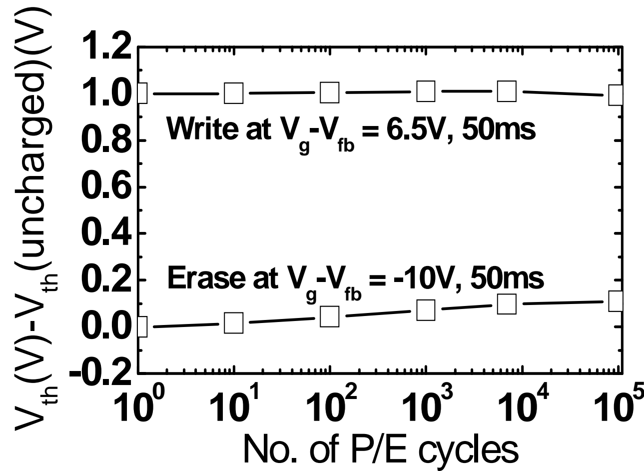


Figure 5.8: Program/Erase (P/E) endurance characteristics of SONOS device with $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ (48% HfO_2) blocking oxide. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a).

The charge retention characteristics of SONOS devices with SiO_2 and high-κ blocking oxide layers were shown in Fig. 5.9. Referring to Fig. 5.9(a), the charge

retention performance of SONOS devices with high-κ blocking oxide layer improved with increasing Al_2O_3 concentration. Interestingly, the memory devices with Al_2O_3, $(HfO_2)_{0.15}(Al_2O_3)_{0.85}$ and $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ blocking oxide layers showed better charge retention performance as compared to SONOS devices with a SiO_2 blocking oxide layer. SONOS devices with HfO_2 blocking oxide layer showed the worst retention performance. It could be seen from Fig. 5.9(b) that SONOS devices with HfAlO blocking oxide layers showed good charge retention lasting up to 10 years.

The charge retention performance of SONOS devices with HfAlO blocking oxide layer was also related to the band gap value of the $(HfO_2)_x(Al_2O_3)_{1-x}$ film. Reduction in the band gap value reduced the effectiveness of the blocking oxide layer in preventing electron out-tunneling from the Si_3N_4 charge storage layer to the gate electrode during charge retention as illustrated in Fig. 5.10.

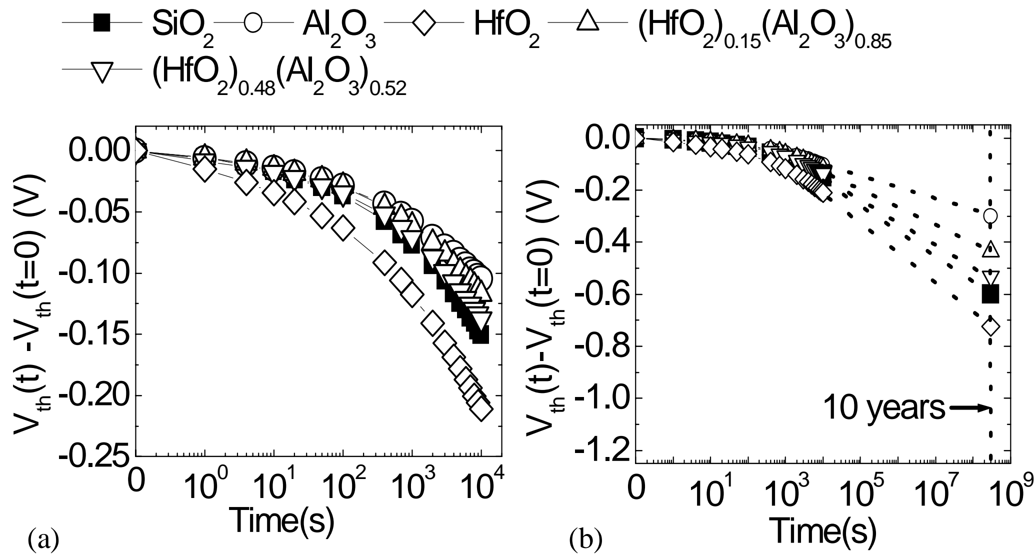


Figure 5.9: (a) Charge retention characteristics of SONOS devices with SiO_2 (solid symbol) or high-κ (open symbols) blocking oxide layers performed at $V_g = 0V$ with source/drain and substrate grounded. The devices were programmed to an initial V_{th} shift of 1.25V before the retention measurements. (b) The same result as in (a) but with the time scale plotted up to 10^9 seconds. The gate stacks of the SONOS devices are 25 Å SiO_2/ 50 Å Si_3N_4/ 75 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (a).

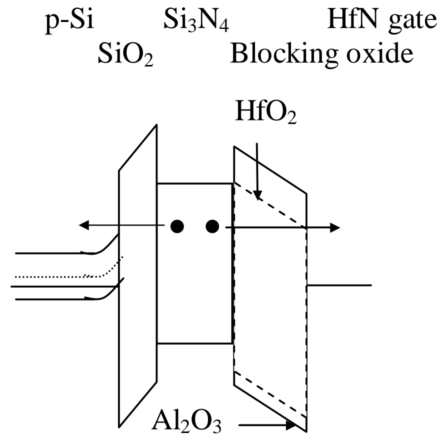


Figure 5.10: Schematic energy band diagrams for SONOS devices with Al_2O_3 (solid lines) and HfO_2 (dashed lines) blocking oxide layer during charge retention measurement.

Since there is a trade-off between programming speed and charge retention in Flash memory device design, both parameters were evaluated together in Fig. 5.11 and Table 5.1. The V_{th} decay rate per decade is the V_{th} shift per decade of measurement time during retention. This was obtained at the later stages (from 10^3 to 10^4 s) of charge retention measurement, when the V_{th} shift with respect to time followed a logarithmic decay. Compared to other SONOS-type devices from bench marked data, SONOS devices with high-κ blocking oxide in this work showed clear improvement in performance considering both programming speed and charge retention as shown in Fig. 5.11.

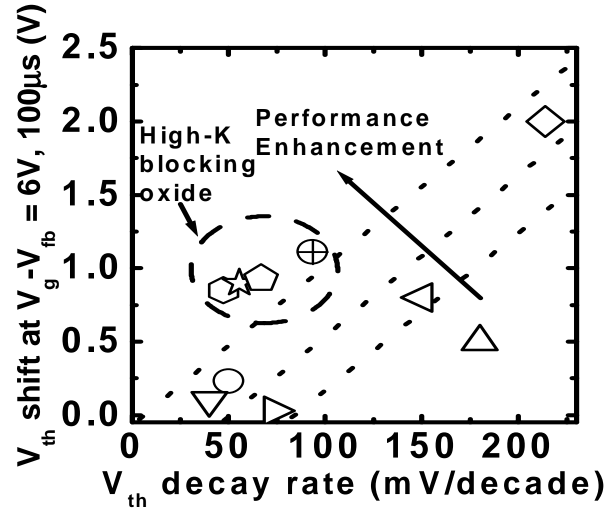


Figure 5.11: Graph of V_{th} shift after programming at V_g-V_{fb} = 6V, 100µs against the V_{th} decay rate per decade of retention measurement time. Comparison between this work and published data (refer to Table 5.1).

Table 5.1: Comparison between this work and published data. SRO is silicon rich oxide.

Sign	Dielectric structure	V_{th} Decay rate (mV/decade)	V_{th} shift at V_g-V_{fb} = 6V, 100µs (V)	V_{th} shift at V_g-V_{fb} = -6V, 1ms (V)	Tunnel oxide thickness (Å)	References
▷	$SiO_2/Si_3N_4/SiO_2$	73	0.03	-0.1551	25	This work
⬡	$SiO_2/Si_3N_4/Al_2O_3$	47	0.85	-0.2787	25	This work
☆	$SiO_2/Si_3N_4/(HfO_2)_{0.15}(Al_2O_3)_{0.85}$	56	0.90	-0.2857	25	This work
⬠	$SiO_2/Si_3N_4/(HfO_2)_{0.48}(Al_2O_3)_{0.52}$	67	0.93	-0.3868	25	This work
⊕	$SiO_2/Si_3N_4/HfO_2$	93	1.11	-0.4579	25	This work
○	$HfO_2/Ta_2O_5/HfO_2$	50	0.23	-0.3V	48 (HfO_2)	[11]
△	SiO_2/SRO/ SiO_2	180	0.5	-1V	25	[12]
▽	$SiO_2/Si_3N_4/SiO_2$	40	0.1	-0.6V	23	[13]
◇	$SiO_2/Si_3N_4/SiO_2$	214	2.0	-3V	15	[14]
◁	$SiO_2/Si_3N_4/SiO_2$	150	0.8	-3V	18	[15]

In order to improve the charge retention performance of the devices, the tunnel oxide thickness was increased to 40 Å and the blocking oxide thickness increased to 120 Å as shown in Fig. 5.1 (b). The programming transients of SONOS devices with HfO_2, $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ or Al_2O_3 blocking oxide layers were shown in Figs. 5.12(a), (b) and (c) for V_g-V_{fb} = 9V, 11V and 13.5V, respectively. For lower programming voltages (i.e., V_g-V_{fb} = 9V), programming speed increased with increasing HfO_2 concentration, similar to the results of devices illustrated in Fig. 5.1 (a). However, as the programming progressed, electrons trapped in the charge storage layer decreased the electric field across the tunnel oxide and increased the electric field across the blocking oxide. Some of the electrons trapped in the charge storage layer may tunnel out through the blocking oxide to the gate electrode. Since Al_2O_3 had a larger conduction band offset with respect to silicon, it could more effectively prevent electron out-tunneling from the charge storage layer to the gate electrode. Hence, Al_2O_3 devices would show a larger V_{th} shift compared to HfO_2 and HfAlO devices as programming time increased. As programming voltage increased (i.e., V_g-V_{fb} = 11V), Al_2O_3 devices showed even faster programming speed compared to $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ devices and the programming time taken for the V_{th} shift to equal to that of HfO_2 devices decreased. Increasing programming voltage increased both the electrons trapped in the charge storage layer and the electric field across the blocking oxide. Hence, the onset of electron out-tunneling through the blocking oxide layer occured earlier especially for dielectrics with smaller conduction band offsets. As the programming voltage increased even further (i.e., V_g-V_{fb} = 13.5V), the HfO_2 devices showed the slowest programming speed due to electron out-tunneling from the Si_3N_4 charge storage layer through the blocking oxide into the gate electrode (refer to Fig. 5.5(c) and (d)). The $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ devices showed faster programming speed as

increasing the Al_2O_3 concentration increased the band gap value and conduction band offset with respect to silicon. Hence, this increased the effectiveness in preventing electron out-tunneling at higher positive gate voltages. The Al_2O_3 devices showed a slower initial programming speed compared to the $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ devices due to the lower tunnel oxide coupling ratio. The tunnel oxide coupling ratio refers to the fraction of the applied gate voltage that is capacitively-coupled to the tunnel oxide. However, as the programming time increased, the Al_2O_3 devices showed the largest amount of V_{th} shift compared to HfO_2 and $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ devices as Al_2O_3 was effective in preventing electron out-tunneling at high positive gate voltage.

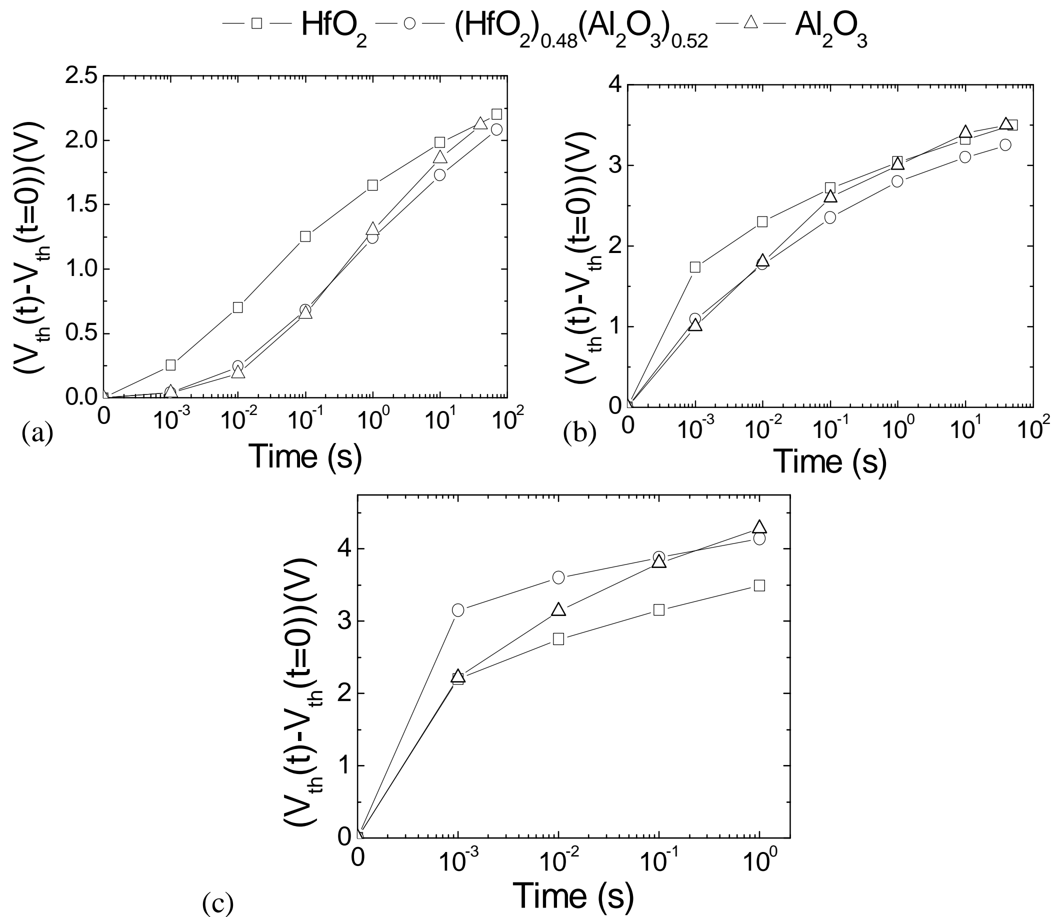


Figure 5.12: Programming transient for (a) $V_g - V_{fb} = 9V$ (b) $V_g - V_{fb} = 11V$ and (c) $V_g - V_{fb} = 13.5V$ for SONOS devices with HfO_2, $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ or Al_2O_3 blocking oxide layers. The gate stacks of the SONOS devices are 40 Å SiO_2/ 70 Å Si_3N_4/ 120 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (b).

Figure 5.13 showed the erasing transient at V_g - V_{fb} = -12.5V for SONOS devices with HfO_2, HfAlO or Al_2O_3 blocking oxide layers. Erasing speed generally increased with increasing HfO_2 concentration due to higher tunnel oxide coupling ratio. The charge retention characteristic of the SONOS devices was shown as Fig. 5.14. The Al_2O_3 and $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ devices showed comparable charge retention characteristics while the HfO_2 device showed the worst retention performance. This can be explained by the difference in band gap values of the films, similar to the devices fabricated with thinner dielectric layers shown in Fig. 5.1 (a).

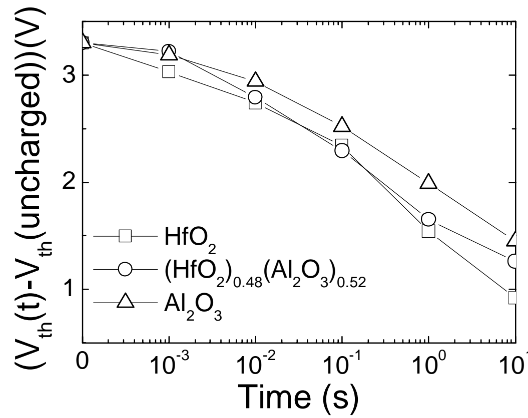


Figure 5.13: Erasing transient at V_g - V_{fb} = -12.5V for SONOS devices with HfO_2, $(HfO_2)_{0.48}(Al_2O_3)_{0.52}$ or Al_2O_3 blocking oxide layers. The gate stacks of the SONOS devices are 40 Å SiO_2/ 70 Å Si_3N_4/ 120 Å high-κ or SiO_2 blocking oxide, as illustrated in Fig. 5.1 (b).

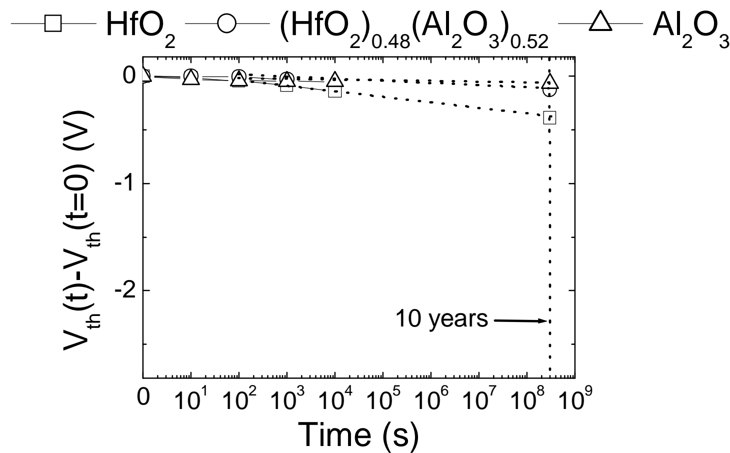


Figure 5.14: Charge retention characteristics of SONOS devices with HfO_2, HfAlO or Al_2O_3 blocking oxide layers performed at V_g = 0V and source/drain and substrate grounded. The devices were programmed to an initial V_{th} shift of 2.9V before retention measurements.

5.3 Evaluation of Lanthanum Aluminum Oxide and Lanthanum Yttrium Aluminum Oxide as the Blocking Oxide Layer in SONOS-type Nonvolatile Memory

5.3.1 Introduction

$LaAlO_3$ has a high-κ of 21 to 25 [16] with high conduction and valence band offsets with respect to silicon of 1.56 eV and 3.2 eV respectively [17]. These properties make it an attractive candidate as a blocking oxide layer in the SONOS structure. In this section, the feasibility of integrating $(La_2O_3)_x(Al_2O_3)_{1-x}$ with different composition ratios as a blocking oxide in the SONOS-type memory structure was investigated.

From calculations based on density-functional theory in the virtual crystal approximation, it was shown that aluminates $(LaAl)_xY_{1-x}O_3$ alloys derived by mixing aluminum oxide with lanthanum and yttrium oxides have unique physical attributes for possible application as gate dielectrics when stabilized in the rhombohedral perovskite structure [16]. However, these attributes are lost in the orthorhombic modification. The room-temperature structure of $LaAlO_3$ is the rhombohedral perovskite (PV or LAP) with a κ value of ~ 21-25. It was shown (by calculation) that the substitution of La in the aluminate with a lighter atom such as Y will result in an increase in the dielectric constant (31.7) for the rhombohedral perovskite structure. However, $YAlO_3$ is not stable at room temperature as a rhombohedral PV, but rather assumes the orthorhombic structure (YAP) in which the dielectric constant is dramatically lower (~16). The addition of La to form $(LaAl)_xY_{1-x}O_3$ was shown to stabilize the rhombohedral modification of $YalO_3$. Stability arguments locate this interesting composition range as $0.2<x<0.4$ [16]. The calculated average κ value in this composition range was between 22 and 25. In addition, the band gap was

calculated to be ~ 1 eV higher than $LaAlO_3$, which may lead to an increase in the conduction band offset with silicon [16]. If $(LaAl)_xY_{1-x}O_3$ is used as a blocking oxide in the SONOS structure, the increase in conduction band offset may lead to less charge transfer between the gate electrode and the charge storage layer, hence minimizing erase saturation. In this section, the feasibility of integrating $(LaAl)_xY_{1-x}O_3$ into the SONOS structure was also investigated.

5.3.2 Sample Fabrication

SONOS devices with $(La_2O_3)_x(Al_2O_3)_{1-x}$ of different composition ratios as blocking oxide were fabricated. 36 Å thick tunnel oxide was thermally grown at 800°C on 4-8 Ω-cm (100) p-type silicon substrates. Subsequently, 65 Å Si_3N_4 was deposited by low pressure chemical-vapor-deposition (LPCVD). $(La_2O_3)_x(Al_2O_3)_{1-x}$ of different composition ratios were deposited by co-sputtering La and Al metals with different power ratios followed by oxidation at 500°C for 60s in O_2 ambient. Lastly, TaN metal gate was formed by physical-vapor-deposition for the control gate. The devices undergo source/drain implantation followed by activation annealing at 950°C for 30s. Both transistor and capacitor structures were fabricated together on the same wafer. The resulting structure is shown in Fig. 5.15.

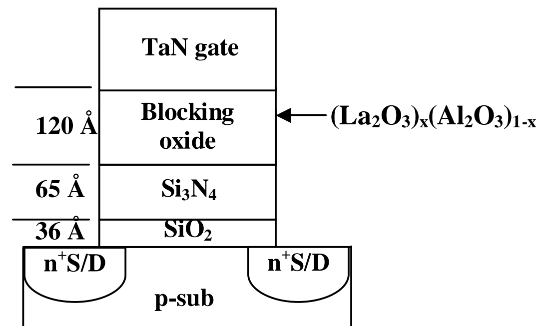


Figure 5.15: Fabricated SONOS structures with TaN gate electrode. The blocking oxide layer is $(La_2O_3)_x(Al_2O_3)_{1-x}$ with different composition ratios.

For $(LaAl)_xY_{1-x}O_3$ evaluation, 80 Å thick $(LaAl)_xY_{1-x}O_3$ was deposited by co-sputtering LaAl and Y metals at different power ratios followed by oxidation at 500°C for 300s in O_2 ambient. The substrate used in this case was 4-8 Ω-cm (100) n-type silicon. Thermal stability test was conducted by annealing some of the devices at 900°C, N_2 for 60s after TaN gate electrode deposition. The resulting device structure is illustrated in Fig. 5.16.

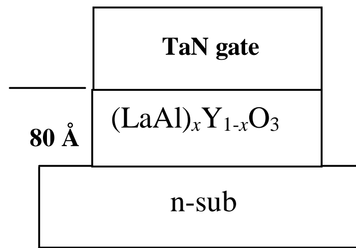


Figure 5.16: Fabricated $(LaAl)_xY_{1-x}O_3$ capacitor structures with TaN gate electrode.

5.3.3 Results and Discussion

(A) Evaluation of $(La_2O_3)_x(Al_2O_3)_{1-x}$ with different composition ratios as blocking oxide

Figure 5.17 shows the high-frequency capacitance-voltage (HFCV) graph of SONOS capacitors with $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide. The capacitors have dimensions of 200 μm × 200 μm**.** Quantum-mechanically corrected equivalent oxide thickness values of the devices with $La_{0.78}Al_{0.22}O_3$, $La_{0.68}Al_{0.32}O_3$ and $La_{0.47}Al_{0.53}O_3$ blocking oxides were 88, 96 and 104 Å respectively. Quantum-mechanical CV correction was done by fitting the theoretical CV curves generated by the CV simulator developed by UC Berkeley Device Group [18] to the experimental CV curves. In the simulation model [18], electrons or holes were confined in the narrow potential well existing at the insulator-silicon interface and quantized in the direction

normal to the insulator-silicon interface. The quantization effect became significant with higher substrate doping and a larger electric field that resulted from a thinner gate insulator. Within the potential well, the carriers were quantum-mechanically confined as a two-dimensional charge sheet. The behavior of the carriers in the potential well deviated substantially from classical theory and required more rigorous quantum-mechanical calculations to describe them. The one-dimensional Poisson and Schrödinger equations were solved self-consistently to find the bounded solutions for energy states and potentials [18]. X-ray photoelectron spectroscopy was used to quantify the film composition. It could be seen that the equivalent oxide thickness decreased, and hence the κ value increased with increasing La concentration. The corresponding gate current density (J_g) versus gate voltage (V_g) graph of the capacitors is illustrated in Fig. 5.18. As the breakdown voltages of the SONOS memory devices were less than 8 V in magnitude, $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide was considered unsuitable for SONOS-type devices. For programming using Fowler-Nordheim tunneling, appreciable electron injection through the tunnel oxide occurs for tunnel oxide electric fields in excess of 10MV/cm. Hence, a suitable minimum breakdown voltage would be around 20 V to allow efficient program and erase by Fowler-Nordheim tunneling. In addition, transistors with high La_2O_3 content showed delamination problem after 950°C anneal. The low breakdown voltage and delamination problem may be due to stress caused by different thermal coefficient of expansion of $(La_2O_3)_x(Al_2O_3)_{1-x}$ compared to the underlying dielectric.

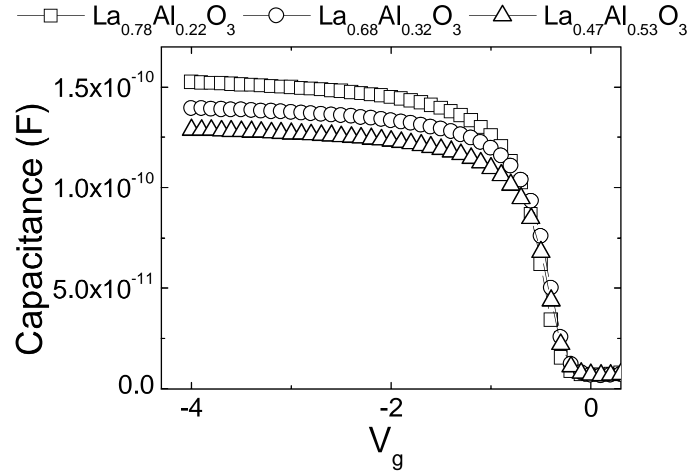


Figure 5.17: High-Frequency Capacitance-Voltage (HFCV) measurements of SONOS capacitors with $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide. The capacitors have dimensions of 200 μm × 200 μm.

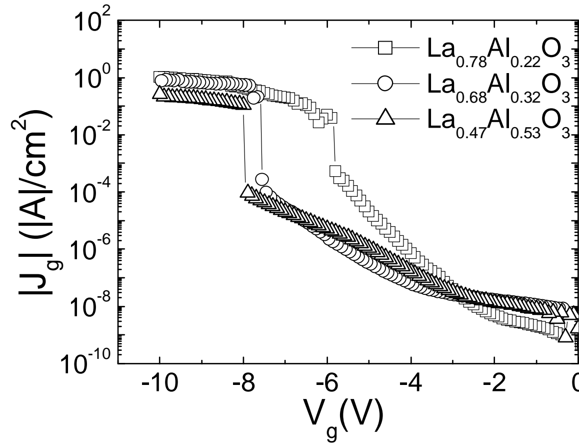


Figure 5.18: Gate current density versus gate voltage (J_g-V_g) measurements of SONOS capacitors with $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide. The capacitors have dimensions of 200 μm × 200 μm.

(B) Feasibility study of $(LaAl)_xY_{1-x}O_3$ with different composition ratios as blocking oxide for SONOS memory

The high-frequency capacitance-voltage (HFCV) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions are shown in Fig. 5.19. The

different composition ratios are calculated by XPS. The capacitors have dimensions of 200 μm × 200 μm. Quantum-mechanically corrected equivalent oxide thicknesses (EOTs) obtained varied between 23 Å and 34 Å. Silicate formation at the dielectric-silicon interface is expected as both Y_2O_3 [19, 20] and $LaAlO_3$ [21, 22] tend to form silicates with lower dielectric constants when in direct contact with silicon. The corresponding gate current versus gate voltage (J_g-V_g) results were shown in Fig. 5.20. As can be seen from Fig. 5.20, all the dielectrics showed good J_g-V_g characteristics before the 900°C, 60s anneal. The gate current density at an excess gate voltage of 3V above the flatband voltage was plotted against EOT in Fig. 5.21. Generally, the addition of $LaAlO_3$ to Y_2O_3 resulted in a decrease in EOT without significant increase in gate current density.

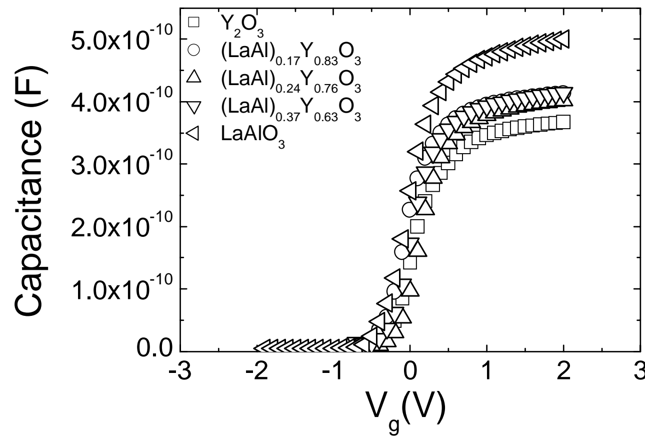


Figure 5.19: High-frequency capacitance-voltage (HFCV) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions. The capacitors have dimensions of 200 μm × 200 μm.

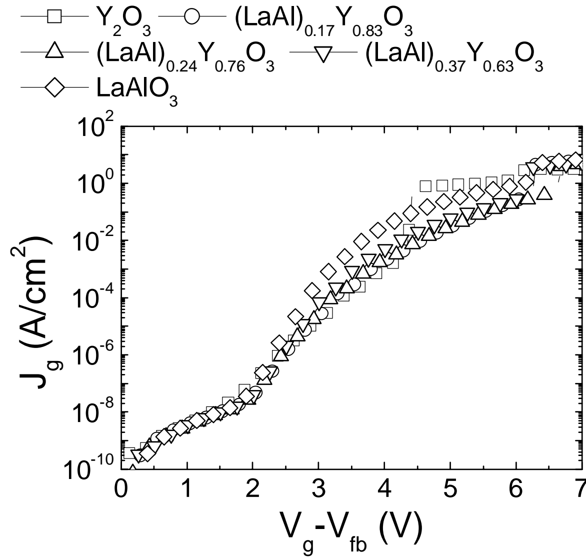


Figure 5.20: Gate-current versus gate voltage (J_g-V_g) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions. The capacitors have dimensions of 200 μm × 200 μm.

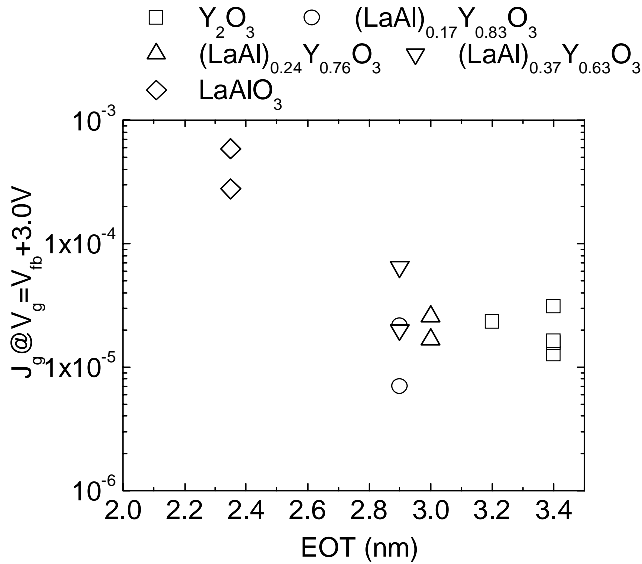


Figure 5.21: Gate-current density at gate voltage of 3V above the flatband voltage against EOT of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions. The capacitors have dimensions of 200 μm × 200 μm.

The barrier height at the TaN/$(LaAl)_xY_{1-x}O_3$ dielectric interface was calculated using the method proposed by Zafar *et al.*[23]. The quantity ΔJ_V, as defined in Eq. (5.1), is greatest when the applied voltage is equivalent to the anode barrier height.

$$\Delta J_V = d(\ln J)/dV \tag{5.1}$$

Figure 5.22 showed the plot of ΔJ_V against V_g. The peak maximum (the barrier height of the TaN/$(LaAl)_xY_{1-x}O_3$ interface) and the $(LaAl)_xY_{1-x}O_3$ conduction band offsets with respect to silicon (ΔE_c) were estimated and tabulated in Table 5.2. The ΔE_c values of $(LaAl)_xY_{1-x}O_3$ were calculated using the value of 4.4 eV for the TaN workfunction [24] and 4.03 eV for the electron affinity of silicon. The calculated value for ΔE_c of $LaAlO_3$ (1.48 ± 0.05 eV) was similar to the value obtained by X-ray photoelectron spectroscopy; Edge *et al.* obtained a value of 1.56 eV [17]. Although the value obtained for ΔE_c of Y_2O_3 (1.63 ± 0.05 eV) was lower than that in literature (2.3 eV as calculated by Robertson [25]), ΔE_c of $(LaAl)_xY_{1-x}O_3$, with $0.17 < x < 0.37$, was higher than that of $LaAlO_3$, as predicted by Shevlin *et al.*[16]. The calculated values of dielectric/metal interface barrier height and ΔE_c by this method may be affected by the presence of traps in the dielectric. Charged traps would cause a distortion in the internal dielectric electric field and trap-assisted tunneling may cause a shift in the location of the peak maximum.

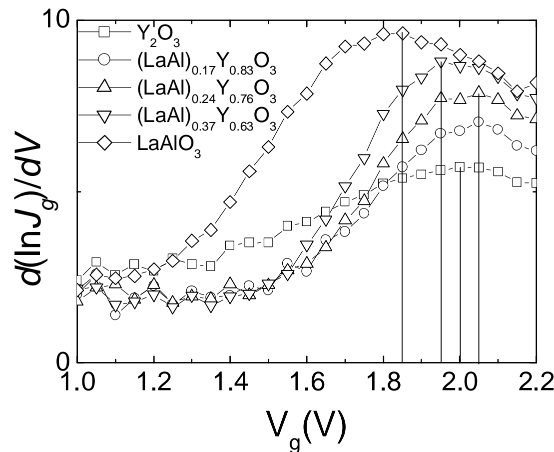


Figure 5.22: $d(\ln J)/dV$ plotted against V_g for TaN/$(LaAl)_xY_{1-x}O_3$ / n-Si devices.

Table 5.2: Estimated barrier heights of the TaN/$(LaAl)_xY_{1-x}O_3$ interface and conduction band offsets of $(LaAl)_xY_{1-x}O_3$ with respect to silicon.

Dielectric type	Barrier height of the TaN/ $(LaAl)_xY_{1-x}O_3$ interface (eV)	Conduction band offset of $(LaAl)_xY_{1-x}O_3$ with respect to silicon (eV).
$LaAlO_3$	1.85 ± 0.05	1.48 ± 0.05
$(LaAl)_{0.37}Y_{0.63}O_3$	1.95 ± 0.05	1.58 ± 0.05
$(LaAl)_{0.24}Y_{0.76}O_3$	2.05 ± 0.05	1.68 ± 0.05
$(LaAl)_{0.17}Y_{0.83}O_3$	2.05 ± 0.05	1.68 ± 0.05
Y_2O_3	2.0 ± 0.05	1.63 ± 0.05

HFCV and J_g-V_g characteristics of $(LaAl)_xY_{1-x}O_3$ capacitors after 900°C, 60s, N_2 anneal were shown in Figs. 5.23 and 5.24, respectively. The high temperature annealing resulted in further increase in EOT, due possibly to interfacial oxide growth or silicate formation [21, 22]. The quantum mechanically corrected EOTs obtained varied between 26 to 34 Å. Flatband voltage shifts in the negative direction, indicating positive fixed charges, were observed in films with higher Y_2O_3 content. The origin of positive charges in Y_2O_3 dielectric was attributed to oxygen vacancies [26]. Dielectrics with a higher Y_2O_3 content showed better thermal stability and lower leakage current densities. A higher $LaAlO_3$ content (≥ 37% in our case) led to higher leakage and early breakdown after high temperature anneal. Hence, $La_xY_{1-x}AlO_3$ with higher Y_2O_3 (> 63%) content may be considered as blocking oxide. The $LaAlO_3$ concentration was limited as it may lead to higher leakage current after high temperature source/drain annealing.

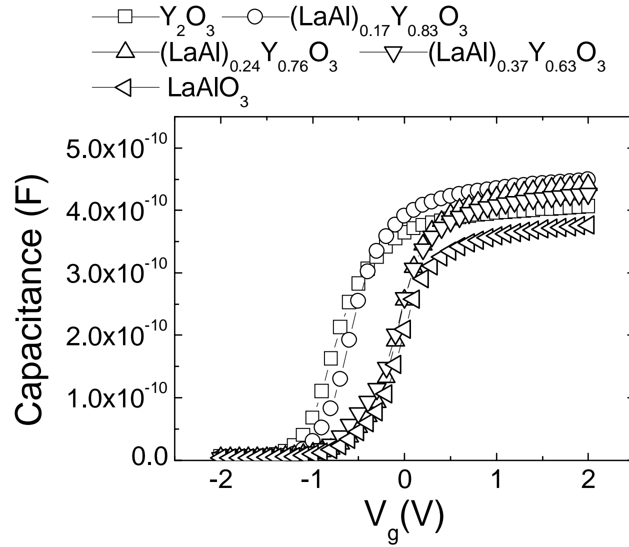


Figure 5.23: High-Frequency Capacitance-Voltage (HFCV) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions after 900°C, 60s, N_2 anneal. The capacitors have dimensions of 200 μm × 200 μm.

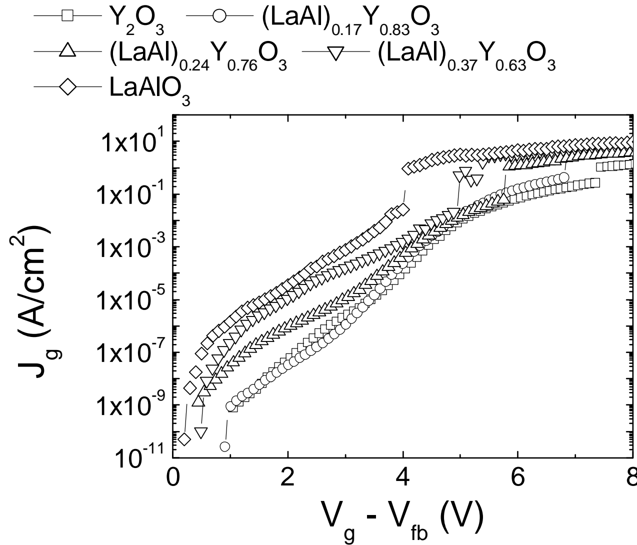


Figure 5.24: Gate current versus gate voltage (J_g-V_g) results of capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions after 900°C, 60s, N_2 anneal. The capacitors have dimensions of 200 μm × 200 μm.

5.4 Summary

The use of HfAlO with different HfO_2 and Al_2O_3 compositions as a blocking oxide in SONOS structure was investigated. The use of high-κ blocking oxide instead

of the conventional SiO_2 in SONOS memory devices resulted in an increase in program and erase speeds, especially at low gate voltages. At high gate voltages, the effectiveness of the high-κ blocking oxide layer in preventing electron tunneling to and from the gate electrode was related to the band-gap value of the blocking oxide, which was inversely proportional to its κ-value. SONOS devices with high-κ blocking oxide layers also showed good charge retention performance. The charge retention performance of SONOS devices improved with increasing Al_2O_3 concentration. Hence the use of a high-κ HfAlO blocking oxide resulted in improvement in program and erase speeds without compromising charge retention capability.

The feasibility of integrating $(La_2O_3)_x(Al_2O_3)_{1-x}$ into the SONOS structure was investigated. SONOS transistors with $(La_2O_3)_x(Al_2O_3)_{1-x}$ of different composition ratios as blocking oxide were fabricated. The low breakdown voltage made $(La_2O_3)_x(Al_2O_3)_{1-x}$ blocking oxide unsuitable for SONOS structure. Transistors with high La_2O_3 content showed delamination problem after 950^oC anneal.

Lastly, capacitors with $(LaAl)_xY_{1-x}O_3$ dielectric with different compositions had also been fabricated to investigate the feasibility of integrating $(LaAl)_xY_{1-x}O_3$ into the SONOS structure. Dielectrics with higher Y_2O_3 content showed better thermal stability and lower leakage current densities. Higher $LaAlO_3$ content led to higher leakage and early breakdown after high temperature anneal. Hence, $La_xY_{1-x}AlO_3$ with higher Y_2O_3 content may be considered as blocking oxide. In the next chapter, the integration of high-κ tunnel and blocking oxides with ultra-high-κ charge storage layer to further improve device performance would be discussed.

References

[1] V. A. Gritsenko, “Design of SONOS memory transistor for terabit scale EEPROM”, in *IEEE Conf. Electron Device and Solid State Circuits,* 2003, p. 345.

[2] C. H. Lee, K. I. Choi, M. K. Cho, Y. H. Song, K. C. Park and K. Kim, “A novel SONOS structure of SiO_2/SiN/Al_2O_3 with TaN metal gate for multi-giga bit flash memories”, in *IEEE IEDM Tech. Dig*., 2003, p. 613.

[3] S. Choi, M. Cho, H. Hwang and J. W. Kim, “Improved metal-oxide-nitride-oxide-silicon-type flash device with high-κ dielectrics for blocking layer”, *J. Appl. Phys.*, vol. 94, pp. 5408-5410, 2003.

[4] C. H. Lee, S. H. Hur, Y. C. Shin, J. H. Choi, D. G. Park and K. Kim, “Charge-trapping device structure of SiO_2/SiN/high-κ dielectric Al_2O_3 for high-density flash memory,” *Appl. Phys. Lett.*, vol. 86, pp. 152908 (1-3), 2005.

[5] G. D. Wilk, R. M. Wallace, J. M. Anthony, “High-κ gate dielectrics: Current status and materials properties considerations”, *J. Appl. Phys.*, vol. 89, pp. 5243-5275, 2001.

[6] H. Y. Yu, M. F. Li, B. J. Cho, C. C. Yeo, M. S. Joo, D. L Kwong, J. S. Pan, C. H. Ang, J. Z. Zheng, S. Ramanathan, “Energy gap and band alignment for $(HfO_2)_x(Al_2O_3)_{1-x}$ on (100) Si”, *Appl. Phys. Lett.*, vol. 81, pp. 376-378, 2002.

[7] H. Y. Yu, H. F. Lim, J. H. Chen, M. F. Li, C. X. Zhu, D. L. Kwong, C.H. Tung, K. L. Bera and C. J. Leo, “Robust HfN metal gate electrode for advanced MOS devices application”, in *VLSI Tech. Symp.*, 2003, p.151.

[8] H. Y. Yu, M. F. Li and D. L. Kwong, “Thermally robust HfN metal as a promising gate electrode for advanced MOS device applications”, *IEEE Trans. Electron Devices*, vol. 51, pp. 609-615, 2004.

[9] H. Y. Yu, J. F. Kang, Ren C., J. D. Chen, Y. T. Hou, Shen C., M. F. Li, D. S. H. Chan, K. L. Bera, C. H. Tung, D. L. Kwong, "Robust high-quality HfN-HfO_2 gate stack for advanced MOS device applications", *IEEE Electron Device Lett.*, vol. 25, pp. 70-72, 2004.

[10] H. Y. Yu, H. F. Lim, J. H. Chen, M. F. Li, Zhu Chunxiang, C. H. Tung, A. Y. Du, W. D. Wang, D. Z. Chi, D. L. Kwong, "Physical and electrical characteristics of HfN gate electrode for advanced MOS devices," *IEEE Electron Device Lett.*, vol. 24, pp. 230-232, 2003.

[11] X. Wang, J. Liu, W. Bai and D. L. Kwong, "A novel MONOS-type nonvolatile memory using high-κ dielectrics for improved data retention and programming speed", *IEEE Trans. Electron Devices*, vol. 51, pp. 597-602, 2004.

[12] M. Rosmeulen, E. Sleeckx and K. De Meyer, "Silicon-rich-oxides as an alternative charge-trapping medium in Fowler-Nordheim and hot carrier type non-volatile-memory cells", in *IEEE IEDM Tech. Dig*., 2002, p. 189.

[13] M. Tanaka, S. Saida, Y. Mitani, I. Mizushima and Y. Tsunashima, "Highly reliable MONOS devices with optimized silicon nitride film having deuterium terminated charge traps", in *IEEE IEDM Tech. Dig.*, 2002, p. 237.

[14] M. H. White, D. A. Adams, B. Jiankang, "On the go with SONOS", *IEEE Circuits and Devices*, pp. 22-31, 2000.

[15] M. L. French and M. H. White, "Scaling of multidielectric nonvolatile SONOS memory structures", *Solid State Electronics*, vol. 37, pp. 1913-1923, 1994.

[16] S. A. Shevlin, A. Curioni and W. Andreoni, "*Ab Initio* Design of High-κ Dielectrics: La_xY_{1-x} AlO_3", *Phys. Rev. Lett.*, vol. 94, pp. 146401(1)-146401(4), 2005.

[17] L. F. Edge, D. G. Schlom, S. A. Chambers, E. Cicerella, J. L. Freeouf, B. Hollander and J. Schubert, "Measurement of the band offsets between amorphous $LaAlO_3$ and silicon", *Appl. Phys. Lett.*, vol. 84, pp. 726-728, 2004.

[18] Quantum-mechanical C-V program from the UC Berkeley Device Group. Available online at http://www-device.eecs.berkeley.edu/qmcv.html.

[19] F. Zhu, S. J. Rhee, C. Y. Kang, C. H. Choi, M. S. Akbar, S. A. Krishan, M. Zhang, H. S. Kim, T. Lee, I. Ok and J. C. Lee, "Improving channel carrier mobility and immunity to charge trapping of high-κ NMOSFET by using stacked Y_2O_3/ HfO_2 gate dielectric", *IEEE Electron Device Letters.*, vol. 26, pp. 876-878, 2005.

[20] J. J. Chambers and G. N. Parsons, "Physical and electrical characterization of ultrathin yttrium silicate insulators on silicon," *J. Appl. Phys.*, vol. 90, pp. 918-933, 2001.

[21] H. Watanabe, N. Ikarashi and F. Ito, "La-silicate gate dielectrics fabricated by solid phase reaction between La metal and SiO_2 underlayers", *Appl. Phys. Lett.*, vol. 83, pp. 3546-3548, 2003.

[22] P. Sivasubramani, M. J. Kim, B. E. Gnade, R. M. Wallace, L .F. Edge, D. G. Schlom, H. S. Craft and J. P. Maria, "Outdiffusion of La and Al from amorphous $LaAlO_3$ in direct contact with Si (001)", *Appl. Phys. Lett.*, vol. 86, pp. 201901(1)-201901(3), 2005.

[23] S. Zafar, C. Cabral, J. R. Amos and A. Callegari, “A method for measuring barrier heights, metal work functions and fixed charge densities in metal/SiO_2/Si capacitors”, *Appl. Phys. Lett.*, vol. 80, pp. 4858-4860, 2002.

[24] C. Ren, H. Y. Yu, J. F. Kang, Y. T. Hou, M.-F. Li, W. D. Wang, D. S. H. Chan and D.-L. Kwong, “Fermi-Level Pinning Induced Thermal Instability in the Effective Work Function of TaN in TaN/SiO_2 Gate Stack”, *IEEE Electron Device Lett.*, vol. 25, pp. 123-125, 2004.

[25] J. Robertson and B. Falabretti, “Band offsets of high-κ gate oxides on III-IV semiconductors”, *J. Appl. Phys.*, vol. 100, pp. 014111(1)-014111(8), 2006.

[26] E. K. Evangelou, C. Wiemer, M. Fanciulli, M. Sethu and W. Cranton, “Electrical and structural characteristics of yttrium oxide films deposited by rf-magnetron sputtering on n-Si”, *J. Appl. Phys.*, vol. 94, pp. 318-325, 2003.

Chapter 6

SONOS-type Nonvolatile Memory with Ultra-high-κ Charge Storage Layer and High-κ Tunnel and Blocking Oxide Layers

6.1 Introduction

SONOS (polysilicon-oxide-silicon nitride-oxide-silicon) Flash memory is one of the most attractive candidates to realize FLASH vertical scaling. In the previous chapters, it has been shown that the programming speed can be increased without reducing the tunnel oxide thickness through appropriate choice of the material for the charge storage layer and also by using a high-κ material as the blocking oxide. The integration of high-κ charge storage and blocking oxide layers will further improve the device performance. The advantages of using a high-κ material, instead of the conventional Si_3N_4 and SiO_2, would be reduction in total equivalent-oxide-thickness (EOT) of the gate stack and increase in the tunnel oxide coupling ratio.

Titanium dioxide (TiO_2) is an attractive material for use as a charge trapping (storage) layer due to its small band gap energy of 3.5 eV [1] and small conduction band offset with respect to silicon of 1.2 eV [1] which would lead to faster programming speed. It has a high relative dielectric constant (κ) value of 80 [1], [2] which would potentially lead to low programming and erasing voltages. However, for TiO_2 to be considered a suitable material for use as a charge storage layer in a gate-first process, it has to be thermally stable, which means that it should not react with the tunnel and blocking oxide layers during the high temperature source/drain

annealing. TiO_2 crystallizes at temperatures above 400^oC [3]. In addition, TiO_2 has been shown to intermix with SiO_2 after high temperature annealing [3]. Si_3N_4 has been shown to be a good barrier layer for the TiO_2 transistor [2], [4]. The leakage current of the TiO_2/Si_3N_4 stack is only degraded slightly after a 900^oC, 10 seconds anneal in N_2 ambient but increased sharply after 1050^oC annealing [2].

The reaction between TiO_2 and Si to form $TiSi_2$ and SiO_2 has a negative Gibbs free energy change value at 1000 K (ΔG^o_{1000}) [5] of -23.014 kcal/mol. This indicates that the reaction is thermodynamically favorable. Hence the extremely high leakage shown by transistors with a TiO_2/Si_3N_4 gate stack after high temperature annealing may be caused by $TiSi_2$ formation. At high temperature, TiO_2 crystallizes. TiO_2 crystallites may penetrate through the Si_3N_4 barrier to reach the Si substrate resulting in $TiSi_2$ formation. In a separate study, a 20 Å thick TiN layer on top of 35 Å SiO_2 was annealed at 850^oC for 10s in vacuum to form TiN nanocrystals [6]. Figure 6.1 shows the Transmission Electron Microscopy (TEM) picture of the SiO_2/TiN stack after 850^oC annealing [6]. Energy Dispersive X-Ray (EDX) analysis of the sample revealed the formation of $TiSi_2$ film. Hence SiO_2 is not an effective barrier against $TiSi_2$ formation. On the other hand, for a 17nm TiO_2/4nm SiO_2 stack after undergoing low temperature forming gas anneal at 420^oC for 30 minutes in this work, the two dielectric layers showed a very smooth interface indicating negligible intermixing, as illustrated in the TEM micrograph in Fig. 6.2. A schematic diagram of the device structure is shown in Fig. 6.3.

Nitridation has been shown to improve the thermal stability of HfO_2 [7]-[10], HfAlO [11] and HfSiO [12]. It would be interesting to investigate the effect of nitridation on the thermal stability of TiO_2.

HfO_2 has been reported to be a good barrier layer for TiO_2 [13] . The HfO_2/TiO_2 gate stack was reported to be thermally stable up to 900°C with negligible intermixing [13]. HfAlO has been shown to be more thermally stable than HfO_2 as it remained relatively amorphous after 900°C annealing [14]. Hence, it is also of interest to evaluate the feasibility of the HfAlO/TiO_2/HfAlO SOHOS structure.

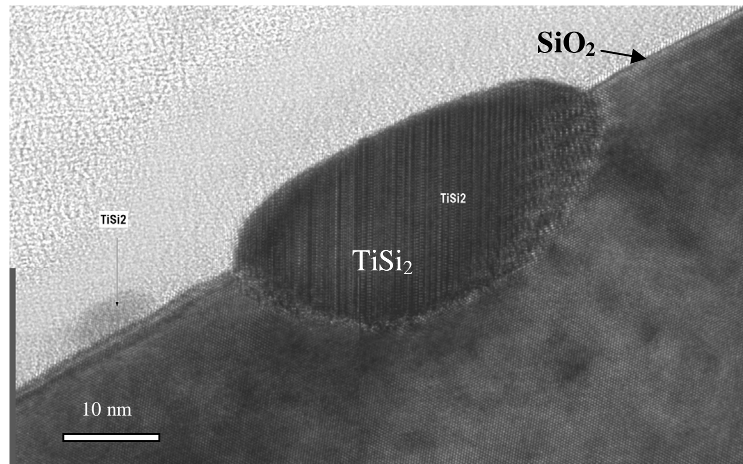


Figure 6.1: TEM micrograph of TiN film on SiO_2 underlayer after 850°C, 10 s anneal in vacuum. EDX analysis revealed formation of $TiSi_2$ after annealing [6].

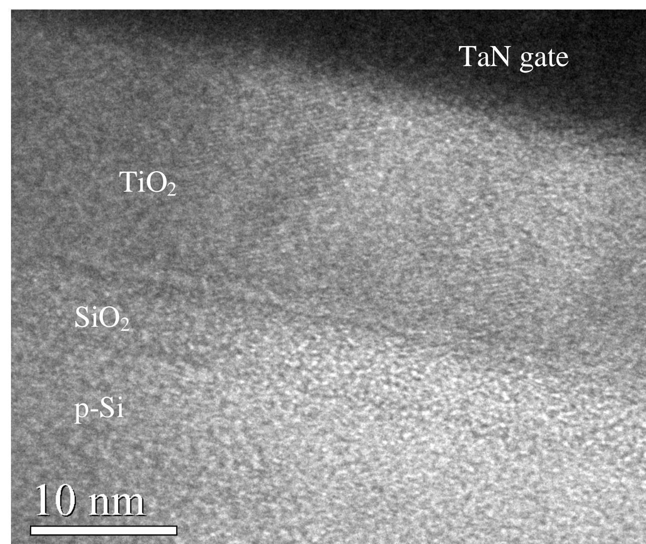


Figure 6.2: TEM micrograph of 4nm SiO_2/17nm TiO_2 layers after forming gas annealing at 420°C for 30 minutes.

6.2 Sample Fabrication

40 Å thick tunnel oxide was thermally grown at 800°C on 4-8 Ω-cm (100) p-type silicon substrates. Subsequently, 170 Å TiO_2 was deposited by reactive sputtering of Ti in oxygen or in a mixture of oxygen and nitrogen with different O_2/N_2 ratios, followed by post-deposition-annealing (PDA) at 700°C for 30s in an O_2 ambient. The devices were annealed in O_2 ambient to ensure that the TiO_2 layer is fully oxidized. Lastly, TaN metal gate was formed by physical vapor deposition for the control gate. Some devices were annealed at 950°C for 30s in N_2 (to simulate the source/drain anneal condition during transistor fabrication) to evaluate the thermal stability of the TiO_2 film. All devices were annealed in forming gas at 420°C for 30 minutes. A schematic diagram of the fabricated structure is shown in Fig. 6.3.

For the HfAlO/TiO_2/HfAlO structure, 60 Å thick HfAlO (with 10% Al_2O_3 concentration) was deposited by metal organic chemical vapor deposition (MOCVD) on 4-8 Ω-cm (100) p-type silicon substrates followed by PDA at 700°C for 60s in N_2 ambient [14]. The devices were annealed in N_2 ambient to minimize interfacial SiO_2 formation at the p-Si/HfAlO interface. 20 Å aluminum nitride (AlN) was deposited by reactive sputtering on some devices to act as both a barrier layer and a nitrogen source for the nitridation of the surrounding dielectrics [15]. Subsequently, 60 Å TiO_2 was deposited by reactive sputtering of Ti in oxygen followed by PDA at 600°C or 700°C for 30s in O_2 ambient. 20 Å Aluminum nitride (AlN) was again deposited on some devices to act as both a barrier layer and a nitrogen source for the nitridation of the surrounding dielectrics [15]. 120 Å HfAlO film was deposited by atomic-layer-deposition (ALD) for blocking oxide. Lastly, TaN metal gate was formed by physical-vapor-deposition for the control gate. Some devices were annealed at 800°C or 900°C for 30 s in N_2 to evaluate the thermal stability of the deposited films. All devices were

annealed in forming gas at 420°C for 30 minutes. The resulting device structures are illustrated in Figs. 6.4(a) and (b) and the ideal energy band diagram of the HfAlO/TiO_2/HfAlO device is shown in Fig. 6.4(c) [1]. The gate area of the fabricated capacitor test structures is 200µm × 200 µm.

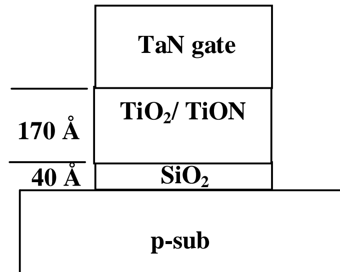


Figure 6.3: Fabricated SiO_2/TiO_2 capacitor structures with TaN gate electrode.

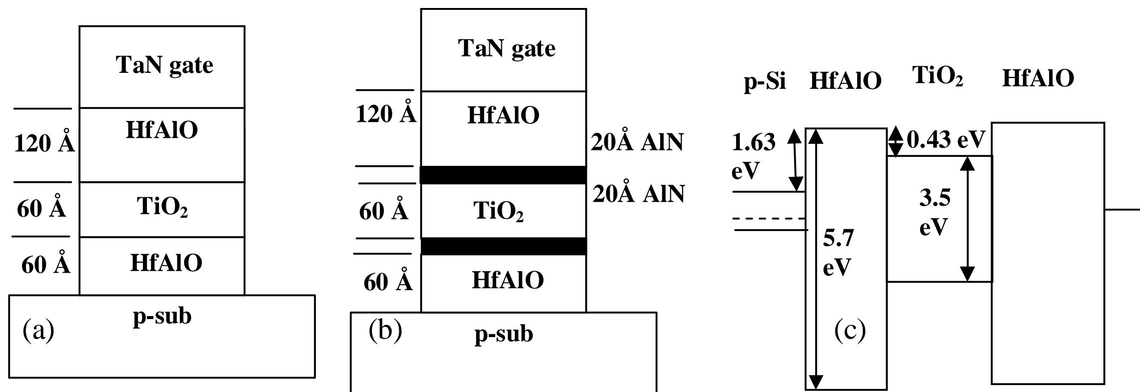


Figure 6.4: (a) Device structure of fabricated HfAlO/TiO_2/HfAlO capacitor structures with TaN gate electrode (b) Device structure of fabricated HfAlO/AlN/TiO_2/AlN/HfAlO capacitor structures with TaN gate electrode. (c) Ideal energy band diagram of HfAlO/TiO_2/HfAlO capacitor.

6.3 Results and Discussion

Figures 6.5(a), (c) and (e) show the high-frequency capacitance-voltage (HFCV) plots of SiO_2/TiO_2 capacitors (Fig. 6.3) after forming gas anneal only, after PDA at 700°C for 30s in an O_2 ambient and after N_2 anneal at 950°C for 30s,

respectively. The corresponding J_g-V_g data are shown in Figs. 6.5(b), (d) and (f), respectively. The gas flow ratios used during reactive sputtering are indicated in the figure legend. As seen from the figures, the leakage current of all the devices increased after high temperature annealing (both 700°C and 950°C anneals). However, the addition of a small amount of N_2 during TiO_2 reactive sputtering reduced the leakage current after high temperature annealing. Devices fabricated with the lowest N_2/O_2 ratio of 2/10 showed the best thermal stability. High N_2 concentration may result in the TiN metal formation, instead of TiON dielectric formation, which will result in leakage current increase as TiN is conductive.

As the SiO_2/TiO_2 structure is unstable after high temperature (700°C and above) annealing, another dielectric must be used to realize SOHOS transistors with TiO_2 charge storage layer. By using high-κ dielectrics as tunnel oxide and blocking oxide layers, the physical thicknesses of both layers can be increased while achieving smaller EOT. Hence, lower program and erasing voltages can potentially be used. Increasing physical thickness will result in lower leakage currents through the tunnel and blocking oxide as compared to SiO_2 layers with the same EOT. This will result in better charge retention. In addition, thicker tunnel and blocking oxide layers can act as more effective barrier layers to prevent TiO_2 interaction with the p-Si substrate. HfO_2 has been reported to be a good barrier layer for TiO_2 [13]. As HfAlO has been shown to be more thermally stable than HfO_2 [14], the feasibility of HfAlO/TiO_2/HfAlO SOHOS structure (Fig. 6.4) was next evaluated.

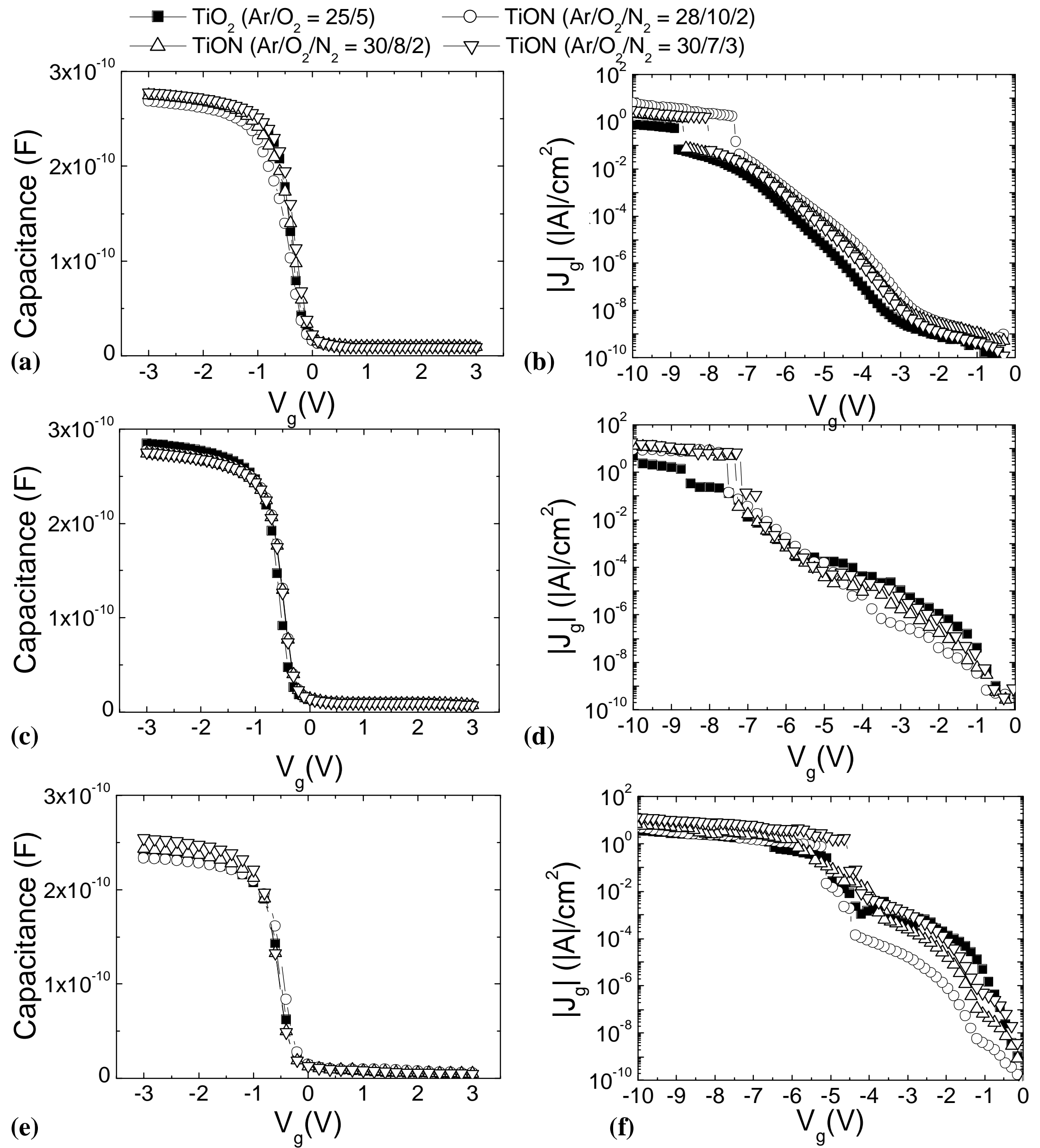


Figure 6.5: (a), (c) and (e) HFCV and (b), (d) and (f) J_g-V_g graphs of SiO_2/TiO_2 capacitors; (a) and (b) after forming gas anneal only, (c) and (d) after 700^oC, 30 s, O_2 PDA and (e) and (f) after 950^oC, 30 s, N_2 anneal. The devices have gate areas of 200 μm × 200 μm.

Figures 6.6(a) and (c) show the HFCV plots of HfAlO/TiO_2/HfAlO SOHOS capacitors after undergoing 700°C, 30 s, O_2 PDA of the TiO_2 layer only and after 900°C N_2 anneal for 30s, respectively. The corresponding J_g-V_g plots are shown as Figs. 6.6(b) and (d), respectively. HfAlO/TiO_2/HfAlO capacitors with only 700°C PDA showed reasonable insulator characteristics. However, some intermixing might have occurred due to the 700°C PDA. On the other hand, 900°C annealed devices showed extremely high leakage. Figure 6.7 shows the TEM micrograph of the HfAlO/TiO_2/HfAlO SOHOS capacitor after the 900°C anneal. The TiO_2/HfAlO layers had intermixed after 900°C annealing as evidenced by the rough interface between the different dielectric layers. From the phase diagram proposed by Ruh and Hollenberg [16], HfO_2 and TiO_2 showed significant intermixing at temperatures higher than 600°C. Hence, further optimization of the process must be done to realize HfAlO/TiO_2/HfAlO SOHOS structures.

Figure 6.8 shows the HFCV curves of HfAlO/TiO_2/HfAlO memory capacitors after undergoing the 700°C PDA, showing counter-clockwise hysteresis for various gate voltage (V_g) sweep ranges as indicated. The capacitance was measured at 100 kHz, with a gate voltage sweep rate of 0.1 V/s. Flatband voltage shifts plotted against the charging (positive) and discharging (negative) gate voltages for 60 Å HfAlO/60 Å TiO_2/120 Å HfAlO and 25 Å SiO_2/60 Å Si_3N_4/60 Å SiO_2 (conventional SONOS) memory devices are shown in Fig. 6.9. The TiO_2 memory device showed much greater flatband voltage shift at lower program/erase voltages compared to the conventional SONOS device. This is an important advantage of using a TiO_2 charge storage layer.

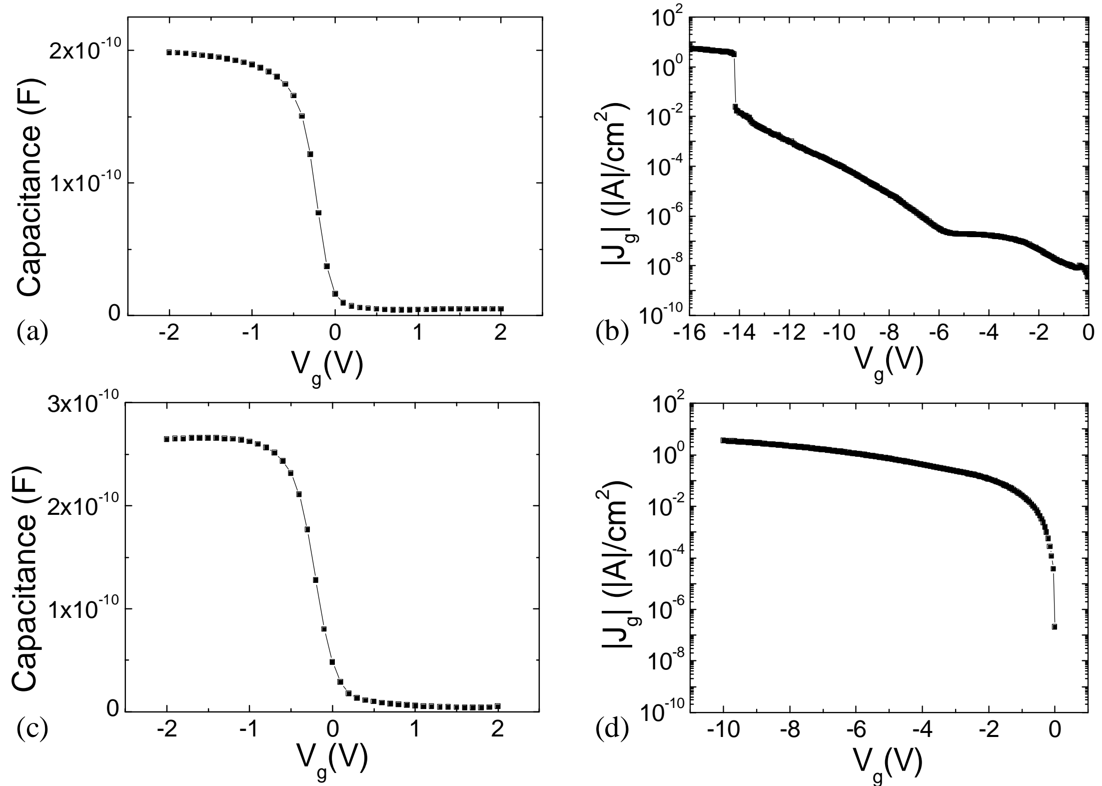


Figure 6.6: (a) and (c) HFCV and (b) and (d) J_g-V_g graphs of HfAlO/TiO_2/HfAlO capacitors; (a) and (b) after 700°C, 30 s, O_2 PDA of the TiO_2 layer and (c) and (d) after 900°C, 30 s, N_2 anneal. The devices have gate areas of 200 μm × 200 μm.

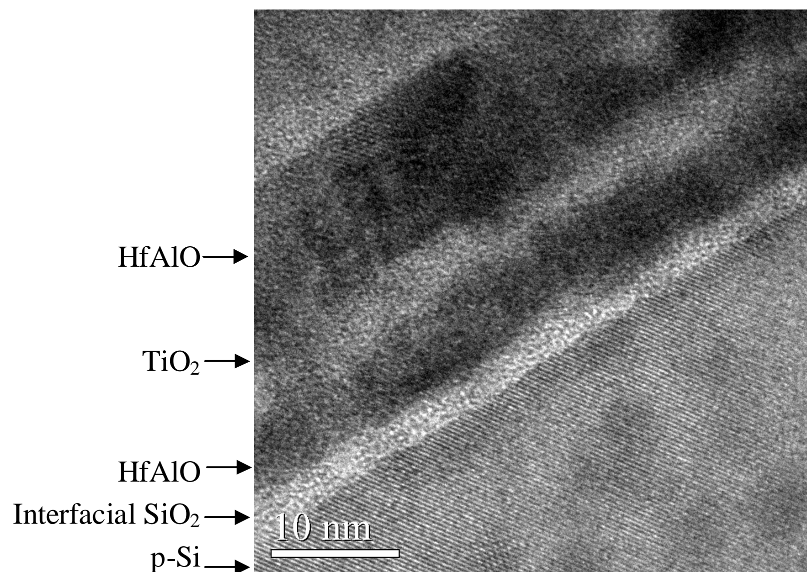


Figure 6.7: TEM micrograph of HfAlO/TiO_2/HfAlO capacitors after 900°C N_2 anneal for 30s.

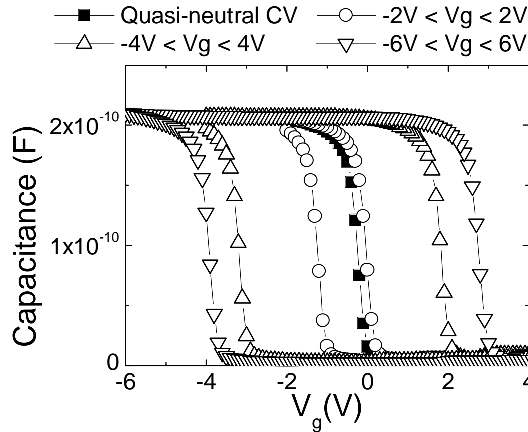


Figure 6.8 C-V curves of HfAlO/TiO_2/HfAlO memory capacitors after PDA at 700°C for 30s in O_2 showing counter-clockwise hysteresis for various gate voltage (V_g) sweep ranges as indicated. The capacitance was measured at 100 kHz, with a gate voltage sweep rate of 0.1 V/s. Gate area is 200 μm × 200 μm.

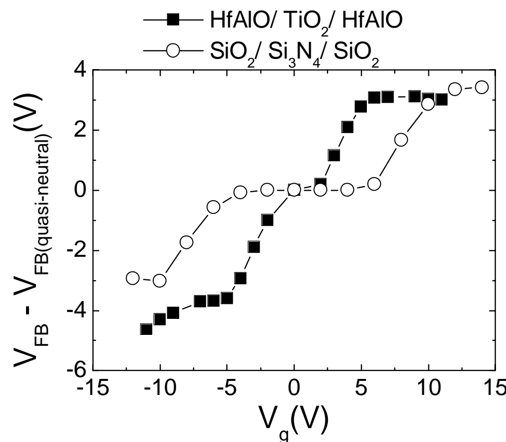


Figure 6.9: Flatband voltage shift extracted from the hysteresis C-V curves plotted against the charging (positive) and discharging (negative) gate voltage for 60 Å HfAlO/60 Å TiO_2/120 Å HfAlO and 25 Å SiO_2/60 Å Si_3N_4/60 Å SiO_2 memory devices. Gate area is 200 μm × 200 μm.

The charge retention characteristic (i.e., V_{fb} shift during retention versus retention time t) of the HfAlO/TiO_2/HfAlO capacitors is shown in Fig. 6.10. The poor retention characteristics of the device (the device has lost more than 50% of its' initial charge after just 100s) may be due to intermixing between TiO_2 and HfAlO during the TiO_2 post-deposition anneal at 700°C for 30s in O_2. Hence, the effective tunnel oxide thickness may be less than that of the deposited value of 60 Å.

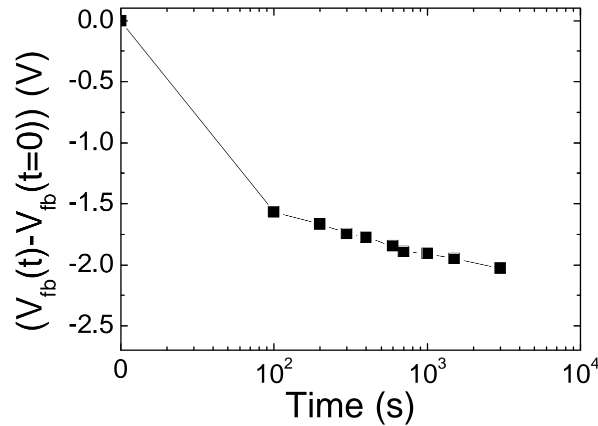


Figure 6.10: Charge retention characteristics of HfAlO/TiO_2/HfAlO memory devices measured with V_g = 0V. The devices were programmed to a V_{fb} shift of 2.7V before retention measurement.

In order to improve the thermal stability of the HfAlO/TiO_2/HfAlO gate stack, 20 Å AlN was deposited as both a barrier layer and as a nitrogen source for the nitridation of HfAlO and TiO_2 [15] as illustrated in Fig. 6.4(b). PDA of the TiO_2 layer was conducted at 600°C for 30s in O_2 ambient. The PDA was done at 600°C instead of 700°C in order to minimize intermixing between TiO_2 and the adjacent dielectric layers. Some of the devices were annealed at 800°C or 900°C for 30s in N_2 after TaN gate deposition to investigate the thermal stability of the gate stack. Figures 6.11(a), (c) and (e) show the HFCV plots of the HfAlO/AlN/TiO_2/AlN/HfAlO gate stack with only 600°C/30s/O_2 PDA, after 800°C/30s/N_2 and after 900°C/30s/N_2 anneals, respectively. The corresponding J_g-V_g plots are shown in Figs. 6.11(b), (d) and (f), respectively. Leakage current increased significantly after 800°C and 900°C anneals. Hence, the AlN barrier layer is not very effective in improving the thermal stability of the gate stack.

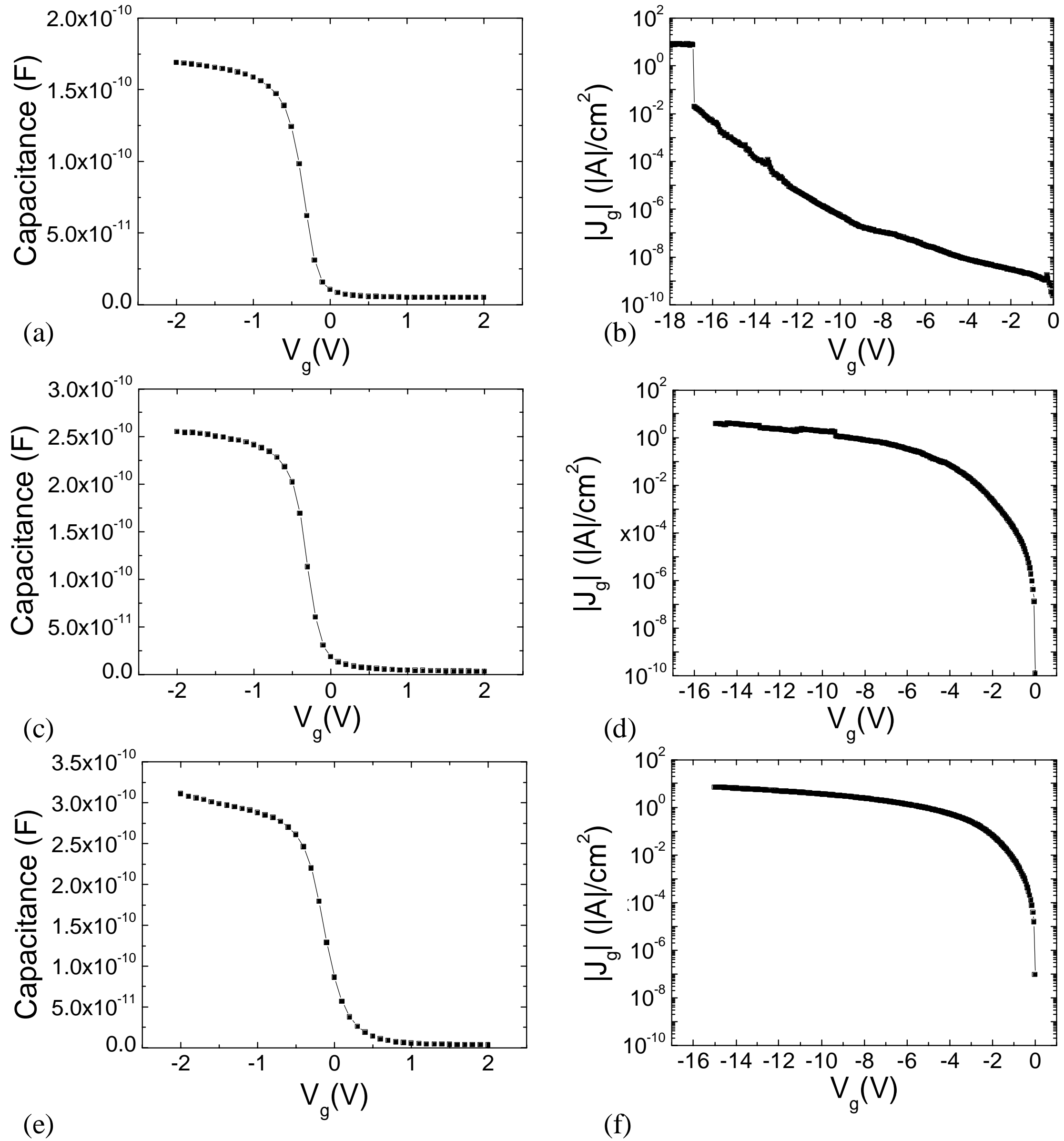


Figure 6.11: (a), (c) and (e) are HFCV while (b), (d) and (f) are J_g-V_g graphs of HfAlO/AlN/TiO_2/AlN/HfAlO capacitors; (a) and (b) with only 600°C, 30s, O_2 PDA of the TiO_2 layer, (c) and (d) after 800°C, 30 s, N_2 anneal while (e) and (f) after 900°C, 30 s, N_2 anneal. The devices have gate areas of 200 μm × 200 μm.

The charge retention characteristics of the HfAlO/AlN/TiO_2/AlN/HfAlO device with only 600°C PDA for 30s in O_2 ambient is illustrated in Fig. 6.12. There is a significant improvement in the charge retention of the device (it has lost only 23% of its' initial charge after 1000s) compared to that of the HfAlO/TiO_2/HfAlO device with 700°C PDA for 30s in O_2 ambient (Fig. 6.10). The retention improvement could

be due to the lower PDA temperature which reduced the intermixing between TiO_2 and the surrounding dielectrics (HfAlO or AlN). In addition, the addition of the AlN barrier layer increased the total physical thickness of the tunnel oxide and thus possibly resulting in better charge retention.

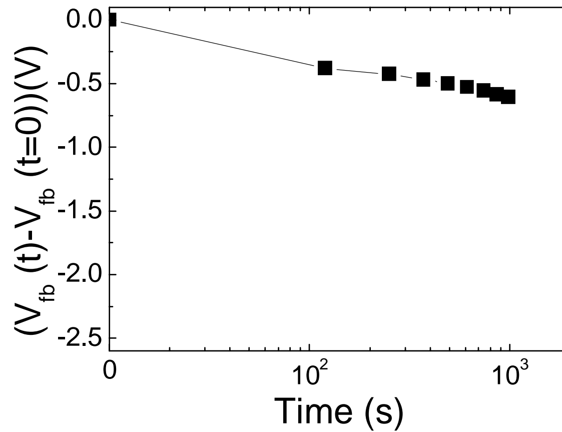


Figure 6.12: Retention characteristics of HfAlO/AlN/TiO_2/AlN/HfAlO memory devices measured with V_g = 0V. The devices were programmed to a V_{fb} shift of 2.6V before retention measurement.

Figure 6.13 shows the flatband voltage shifts extracted from the hysteresis of C-V curves plotted against the charging (positive) and discharging (negative) gate voltages for the 60 Å HfAlO/60 Å TiO_2/120 Å HfAlO and 60 Å HfAlO/20 Å AlN/60 Å TiO_2/20 Å AlN/120 Å HfAlO memory devices. The addition of AlN resulted in a slight reduction of V_{fb} shift with program/erase voltages as the charge carriers (electrons and holes) had to tunnel through a thicker tunnel oxide compared to the HfAlO/TiO_2/HfAlO devices. Hence, there is a trade-off in V_{fb} shift during program/erase operations and charge retention. Nevertheless, the HfAlO/AlN/TiO_2/AlN/HfAlO memory devices are promising structures for gate-last processes due to the large V_{fb} shift during program/erase operations and good charge retention.

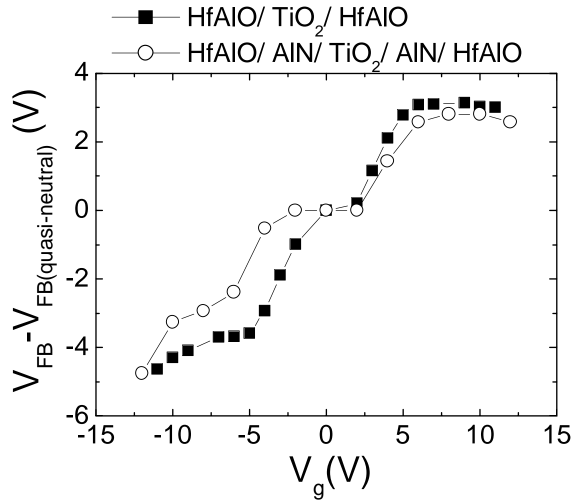


Figure 6.13: Flatband voltage shift extracted from the hysteresis C-V curves plotted against the charging (positive) and discharging (negative) gate voltage for 60 Å HfAlO/60 Å TiO_2/ 120 Å HfAlO and 60 Å HfAlO/20 Å AlN/60 Å TiO_2/20 Å AlN/120 Å HfAlO memory devices. Gate area is 200 μm × 200 μm.

6.4 Summary

The feasibility of using TiO_2 as a charge storage layer in SONOS memory was investigated. Dielectric intermixing in TiO_2/SiO_2 devices was suspected after high temperature annealing resulting in leakage current increase. TiO_2 nitridation possibly reduced dielectric intermixing and lower leakage current was obtained. HfAlO/TiO_2/HfAlO SOHOS capacitors showed much greater flatband voltage shift at lower program/erase voltages compared to the conventional SONOS device after PDA and forming gas anneal. The poor charge retention characteristics observed may be due to HfAlO/TiO_2 intermixing during the high temperature PDA (700°C, 30s, O_2) process. The charge retention performance of the devices was much improved by the addition of AlN to result in a HfAlO/AlN/TiO_2/AlN/HfAlO gate stack and reduction in PDA temperature from 700°C to 600°C. Therefore, if the intermixing problem is solved, TiO_2 charge storage layer can be a very promising candidate for next generation SONOS type memory device.

References

[1] G. D. Wilk, R. M. Wallace and J. M. Anthony, "High-κ Gate Dielectrics: Current Status and Materials Properties Considerations", *J. Appl. Phys.*, vol. 89, pp. 5243-5275, 2001.

[2] X. Guo, X. Wang, Z. Luo, T. P. Ma and T. Tamagawa, "High quality ultra-thin (1.5 nm) TiO_2/Si_3N_4 gate dielectric for deep sub-micron CMOS technology", in *IEEE IEDM Tech. Dig.*, 1999, p.137.

[3] N. S. Gluck, H. Sankur, J. Heuer, J. Denatale and W. J. Gunning, "Microstructure and composition of composite SiO_2/TiO_2 thin films", *J. Appl. Phys.*, vol. 69, pp. 3037-3045, 1991.

[4] Y. Jeon, B. H. Lee, K. Zawadzki, W. J. Qi, A. Lucas, R. Nieh, J. C. Lee, "Effect of barrier layer on the electrical and reliability characteristics of high-κ gate dielectric films", in *IEEE IEDM Tech. Dig.*, 1998, p.797.

[5] K. J. Hubbard and D. G. Schlom, "Thermodynamic stability of binary oxides in contact with silicon", *J. Mater. Res.*, vol. 11, pp. 2757-2776, 1996.

[6] S. K. Samanta, "TiN nanocrystal memory devices", unpublished.

[7] C. S. Kang, H. –J. Cho, K. Onishi, R. Choi, Y. H. Kim, R. Nieh, J. Han, S. Krishnan, S. Krishnan, A. Shahriar and J. C. Lee, "Nitrogen concentration effects and performance improvement of MOSFETs using thermally stable HfO_xN_y Gate dielectrics", in *IEEE IEDM Tech. Dig.*, 2002, p.865.

[8] H. J. Cho, C. S. Kang, K. Onishi, S. Gopalan, R. Nieh, R. Choi, S. Krishnan and J. C. Lee, "Structural and electrical properties of HfO_2 with top nitrogen incorporated layer", *IEEE Electron Device Lett.*, vol. 23, pp. 249-251, 2002.

[9] C. S. Kang, H. –J. Cho, K. Onishi, R. Choi, R. Nieh, S. Goplan, S. Krishnan and J. C. Lee, "Improved thermal stability and device performance of ultra-

thin (EOT < 10 Å) gate dielectric MOSFETs by using Hafnium Oxynitride (HfO_xN_y)", in *VLSI Tech. Symp.*, 2002, p.146.

[10] C. H. Choi, S. J. Rhee, T. S. Jeon, N. Lu, J. H. Sim, R. Clark, M. Niwa and D. L. Kwong, "Thermally stable CVD HfO_xN_y advanced gate dielectrics with poly-Si gate electrode", in *IEEE IEDM Tech. Dig.*, 2002, p.857.

[11] T. Nishimura, K. Iwamoto, K. Tominaga, T. Yasuda, W. Mizubayashi, S. Fujii, T. Nabatame and A. Toriumi, "Effects of Nitrogen incorporation into $HfAlO_x$ films on gate leakage current – from XPS study of Hf bonding states", in *Int. Workshop on Gate Insulator*, 2003, p.180.

[12] M. R. Visokay, J. J. Chambers, A. L. P. Rotondaro, A. Shanware and L. Colombo, "Application of HfSiON as a gate dielectric material", *Appl. Phys. Lett.*, vol. 80, pp. 3183-3185, 2002.

[13] S. J. Rhee, C. S. Kang, C. H. Choi, C. Y. Kang, S. Krishnan, M. Zhang, M. S. Akbar and J. C. Lee, "Improved electrical and material characteristics of hafnium titanate multi-metal oxide n-MOSFETs with ultra-thin EOT (8 Å) gate dielectric application", in *IEEE IEDM Tech. Dig.*, 2004, p. 837.

[14] M. S. Joo, B. J. Cho, C. C. Yeo, D. S. H. Chan, S. J. Whoang, S. Matthew, L. K. Bera, N. Balasubramaniam and D. L. Kwong, "Formation of Hafnium-Aluminum-Oxide gate dielectric using single cocktail liquid source in MOCVD process", *IEEE Trans. Electron Dev.*, vol. 50, pp. 2088-2094, 2003.

[15] C. S. Park, B. J. Cho and D. L. Kwong, "MOS characteristics of synthesized HfAlON-HfO_2 stack using AlN-HfO_2", *IEEE Electron Device Lett.*, vol. 25, pp. 619-621, 2004.

[16] R. Ruh and G. W. Hollenberg, “Phase Relations and Thermal Expansion in the System HfO_2-TiO_2”, *J. American Ceramics Society*, vol. 59, pp. 495-499, 1976.

Chapter 7

Conclusion

7.1 Summary of Findings

According to the International Technology Roadmap for Semiconductors, the difficult challenge for Flash scaling to 32 nm technology and beyond is the non-scalability of the tunnel and interpoly dielectrics of the floating-gate memory structure [1]. SONOS (polysilicon-oxide-silicon nitride-oxide-silicon) Flash memory is considered to be one of the most attractive candidates to replace the conventional floating-gate structure.

One of the more effective methods for improving the programming speed of the SONOS memory device is to reduce the tunnel oxide thickness. However, such a method has the inevitable disadvantage of degradation in the charge retention. Another method to improve the Flash device performance is by using alternative materials such as high-κ dielectrics as part of the gate stack, and this was investigated in the work presented in this dissertation. Basically, due to the higher dielectric constant or κ value, the equivalent oxide thickness is reduced for the same physical thickness of the film. Hence, the effect on device performance is expected to be similar to that of ONO stack scaling without the disadvantages that come with smaller physical thicknesses [2].

In the first part of the project, the effect of replacing the silicon nitride charge storage layer with a higher κ HfO_2 layer was investigated. The resulting device was referred to as the SOHOS (polysilicon-oxide-high-κ-oxide-silicon) Flash memory. The SOHOS structure with hafnium oxide (HfO_2) as the charge storage layer

demonstrated superior charge storage capability at low voltages, faster programming and less over-erase problem as compared to the conventional SONOS device. These were attributed to differences in the band offsets of the charge storage layer.

However, SOHOS devices with HfO_2 charge storage layer had poorer charge retention capability than SONOS devices and also poor endurance characteristics. On the other hand, using Al_2O_3 as the charge storage layer resulted in a SOHOS structure with improved charge retention performance, but with slower programming speed. The charge loss in devices with Al_2O_3 as the charge storage layer showed stronger temperature dependence compared to devices with HfO_2 as the charge storage layer. Hence, the good charge retention performance of Al_2O_3 devices was probably due to deeper trap levels. Therefore, by adding a small amount of aluminum to HfO_2 to form hafnium aluminum oxide (HfAlO), the resultant SOHOS structure with HfAlO as a charge storage layer can combine the advantages of both HfO_2 and Al_2O_3, such as fast programming speed, good charge retention and good program/erase endurance. The charge storage mechanism in SOHOS devices with HfAlO charge storage layer was attributed to electron traps within the bulk.

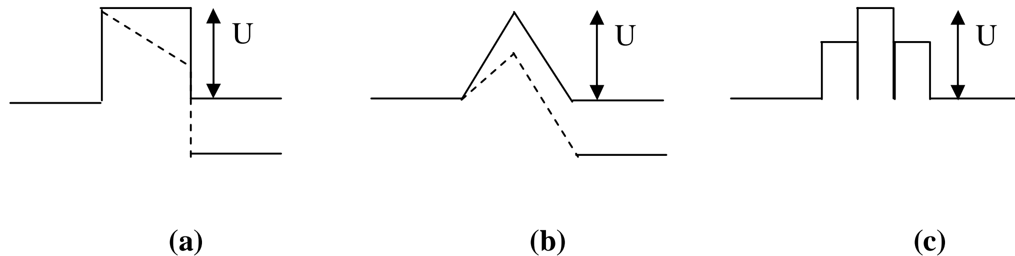
The use of a high-κ material as the blocking oxide was investigated as an alternative method to increase program/erase speed without decreasing the tunnel oxide thickness. From electrostatics consideration, the use of a high dielectric constant blocking oxide layer will cause a smaller voltage drop across the blocking oxide and greater voltage drop across the tunnel oxide. This will result in a simultaneous increase of the electric field across the tunnel oxide and reduction of the electric field across the blocking oxide, leading to more efficient program and erase processes. The effect of the κ value and band gap energy of the blocking oxide layer on the program/erase speed and charge retention of SONOS devices was investigated

by using HfAlO or $(HfO_2)_x(Al_2O_3)_{1-x}$ with different HfO_2 concentration ratios (x) as the blocking oxide. The use of the HfAlO high-κ blocking oxide instead of the conventional SiO_2 blocking oxide in SONOS memory devices resulted in an increase in program and erase speeds, especially at low gate voltages. At high gate voltages, the effectiveness of the high-κ blocking oxide layer in preventing electron tunneling to and from the gate electrode was related to the band-gap value of the blocking oxide, which was inversely related to its κ-value. SONOS devices with HfAlO high-κ blocking oxide layers also showed good charge retention performance. The charge retention performance of SONOS devices improved with increasing Al_2O_3 concentration. Hence the use of a high-κ HfAlO blocking oxide resulted in improvement in program and erase speeds without compromising the charge retention capability. Other high-κ materials with suitable conduction and valence band offsets were also evaluated.

Finally, the integration of high-κ tunnel and blocking oxides and an ultra-high-κ TiO_2 charge storage layer was also demonstrated in this project. HfAlO/TiO_2/HfAlO SOHOS capacitors showed much greater flatband voltage shift at lower program/erase voltages compared to the conventional SONOS device after post-deposition and forming gas anneals. The poor charge retention of the devices was attributed to dielectric intermixing between the TiO_2 and HfAlO layers during the post-deposition annealing. The charge retention performance of the devices was significantly improved by decreasing the post-deposition annealing temperature and by the addition of AlN to result in an HfAlO/AlN/TiO_2/AlN/HfAlO gate stack.

7.2 Recommendations for Future Work

Simultaneous improvements in both program/erase speeds and charge retention performance may be achieved by using layered tunnel barriers as the tunnel dielectrics [3]. Fowler-Nordheim tunneling of electrons through crested energy barriers (with the height peak in the middle) had been shown to be much more sensitive to applied voltage than that through barriers of uniform height [3], [4]. Figure 7.1 illustrates the conduction band edge diagrams of uniform and crested symmetric barriers.



(a) **(b)** **(c)**

Figure 7.1: Conduction band edge diagrams of various tunnel barriers: (a) a typical uniform barrier; (b) idealized crested symmetric barrier; (c) crested, symmetric layered barrier. U is the maximum barrier height, expressed in units of energy.

The conventional uniform barrier, illustrated in Fig. 7.1(a), has relatively low sensitivity to the applied electric field, as shown by Likharev [3]. This was attributed to the fact that the highest part of the barrier, closest to the electron source, was only weakly affected by the applied voltage V, that is $U_{max}(V) \approx U_{max}(0)$. On the other hand, the current through a crested barrier changes much faster with respect to the applied electric field [3]. The reason for this dramatic improvement was that in the crested barrier the highest part (in the middle) was pulled down by the electric field very quickly, that is $U_{max}(V) \approx U_{max}(0) - eV/2$, where e is the electron charge. This was illustrated in Fig. 7.1(b) for the idealized crested symmetric barrier case. The

crested symmetric barrier may be implemented by using dielectric layers with different band gaps and conduction and valence band offsets, as illustrated in Fig. 7.1(c). Some feasible combinations are $Si_3N_4/Al_2O_3/Si_3N_4$, $HfO_2/Al_2O_3/HfO_2$ [4], $HfO_2/La_2O_3/HfO_2$ and $Ta_2O_5/Al_2O_3/Ta_2O_5$.

For the blocking oxide, the most suitable barrier structure is still the conventional uniform barrier. The function of the blocking oxide is to prevent charge transfer from the charge storage layer to the gate electrode during programming and from the gate electrode to the charge storage layer during erasing. The conventional uniform barrier has the lowest sensitivity to applied voltage [3], hence it will effectively prevent charge transfer during program/erase processes.

The integration of p-type metals with high work functions into the SONOS memory structure will lead to a larger threshold voltage window due to less erase saturation. High work function metal gate increases the energy barrier for electron tunneling from the gate electrode to the charge storage layer during erase. Hence, electron tunneling from the gate is minimized. This will lead to a more effective erasing process and prevents erase saturation. Some p-type metal gate candidates that can be integrated into the CMOS process are Ruthenium (Ru) [5] and Molybdenum (Mo) [6].

References

[1] International Technology Roadmap for Semiconductors, 2005.

[2] M. Specht, H. Reisinger, M. Stadele, F. Hofmann, A. Gschwandtner, E. Landgraf, R. J. Luyken, T. Schulz, J. Hartwich, L. Dreeskornfeld, W. Rosner, J. Kretz and L. Risch, "Retention times of novel charge trapping memories using Al_2O_3 dielectrics", in *33rd Conf. on European Solid-State Device Research*, 2003, p. 155.

[3] K. K. Likharev, "Layered tunnel barriers for nonvolatile memory devices", *Appl. Phys. Lett.*, vol. 73, pp. 2137-2139, 1998.

[4] B. H. Koh, W. K. Chim, T. H. Ng, J. X. Zheng and W. K. Choi, "Quantum mechanical modeling of gate capacitance and gate current in tunnel dielectric stack structures for nonvolatile memory application", *J. Appl. Phys.*, vol. 95, pp. 5094-5103, 2004.

[5] Z. B. Zhang, S. C. Song, C. Huffman1, J. Barnett, N. Moumen, H. Alshareef, P. Majhi, M. Hussain, M. S. Akbar, J. H. Sim, S. H. Bae, B. Sassman and B. H. Lee, "Integration of dual metal gate CMOS with TaSiN (NMOS) and Ru (PMOS) gate electrodes on HfO_2 Gate Dielectric", in *VLSI Tech. Symp. Proc.*, 2005, p. 50.

[6] Y. C. Yeo, Q. Lu, P. Ranade, H. Takeuchi, K. J. Yang, I. Polishchuk, T. J. King, C. Hu, S. C. Song, H. F. Luan and D. L. Kwong, "Dual metal gate CMOS technology with silicon nitride gate dielectric", *IEEE Electron Device Lett.*, vol. 22, pp. 227-229, 2001.

List of Publications

[1] Y. N. Tan, W. K. Chim, B. J. Cho and W. K. Choi, "Over-erase phenomenon in SONOS-type Flash memory and its minimization using a hafnium oxide charge storage layer", *IEEE Trans. Electron Devices*, vol. 51, no. 7, pp. 1143-1147, 2004.

[2] Y. N. Tan, W. K. Chim, B. J. Cho and W. K. Choi, "A Novel MONOS-type Flash memory using a high-κ HfAlO charge trapping layer", *Electrochemical and Solid-State Letters*, vol. 7, no. 9, pp. G198-G200, 2004.

[3] Y. N. Tan, W. K. Chim, W. K. Choi, M. S. Joo, T. H. Ng and B. J. Cho, "High-κ HfAlO Charge Trapping Layer in SONOS-type Nonvolatile Memory Device for High Speed Operation", in *IEEE IEDM Tech. Dig.*, 2004, p. 889.

[4] Y. N. Tan, W. K. Chim, W. K. Choi, M. S. Joo and B. J. Cho, "Hafnium Aluminum Oxide as Charge Storage and Blocking-Oxide Layers in SONOS-type Nonvolatile Memory for High Speed Operation", *IEEE Trans. Electron Devices*, vol. 53, pp. 654-662, 2006.

[5] Y. N. Tan, W. K. Chim, B. J. Cho and W. K. Choi, "Memory Gate Stack Structure", International Patent (Application No.: PCT/SG2004/000050) filed on 11 March 2004 (ETPL Ref: SRC/P//1760/PCT).

[6] Y. N. Tan, W. K. Chim, W. K. Choi, M. S. Joo, T. H. Ng and B. J. Cho, "Evaluation of SOHOS (polysilicon-oxide-high κ-oxide-silicon) structure for Flash memory device application", Poster presented at the 3rd International Conference on Materials for Advanced Technologies (ICMAT 2005) and 9th International Conference on Advanced Materials, 3 - 8 July 2005, Suntec Singapore International Convention and Exhibition Centre, Singapore.