Optimization of silicon nitride for SONOS flash memories by trap characterization

A dissertation

Submitted in the partial fulfillment of the requirements for the degree of

Master of Technology

by

VVSS SATYASURESH CH

08307R22

under the guidance of

Prof. Anil Kottantharayil

Department of Electrical Engineering

Indian Institute of Technology Bombay

June 2011

Abstract

This report describes the process optimization and trap characterization of the silicon nitride charge trap layer for SONOS flash memories by using Low Pressure Chemical Vapor Deposition (LPCVD) with SiH_4 and NH_3 as source gases. Silicon nitride in LPCVD can also be deposited using chloro-silanes typically dichloro-silane (DCS) and ammonia (NH_3) but requires extra bubbler and vaporizer arrangement to feed the reaction tube. Also the formation of NH_4Cl during deposition is detrimental to the pumping system due to the clogging of NH_4Cl particulates. On the other hand, the LPCVD system with SiH_4 and NH_3 is simple and easy to maintain. However the large non-uniformity in thickness and film composition are the main issues with silicon nitride films deposited by SiH_4 and NH_3 system. In this work the results of SiH_4 - NH_3 system with N_2 as diluent for depositing silicon nitride films are presented. A significant improvement in uniformity is observed when N_2 is used during deposition. Effect of process parameters on the film thickness and composition is studied. A breakdown field of 8.5 MV/cm is obtained for stoichiometric silicon nitride films.

Trap characterization results on SNS capacitors are described and effect of N_2 dilution on trap density and distribution is studied. Trap density and energy distribution of traps are determined from program transients and retention measurements. The trap density is of the order of 2.5E18 cm^{-3} and is distributed in the energy range 0.74 - 1.5 eV. Capture cross sections of both electron and hole traps are calculated.

Acknowledgements

First of all, I would like to thank Indian Institute of Technology Bombay for giving me the opportunity and providing facilities to do this project.

I sincerely thank my guide Prof. Anil Kottantharayil for his encouragement, direction, guidance and support during the project. His recommendations and suggestions have been invaluable for the project. I would like to thank him for sparing his valuable time to review manuscript of the thesis.

It is my pleasure to thank all the people who helped me directly or indirectly in the progress of this project. I am grateful to all the technical and non-technical staff members of the nanoelectronics fabrication facility for their support without whom this project would not have been possible.

I would like to thank all the members of flash memory group for their valuable suggestions in solving process related issues. I would also like to thank the RAs of characterization lab for their support.

Special thanks are due to my friends in IITB for their support and motivation during hard times.

Finally, I would like to thank my family members for their support and encouragement during my academics.

Contents

1 Introduction **1**

1.1 Organization of the Report 4

2 Literature Review **5**

2.1 Evolution of SONOS memory 5

2.2 SONOS charge injection mechanisms 6

2.2.1 Tunneling based charge injection 7

2.2.1.1 Direct band to band tunneling (DT) 7

2.2.1.2 Fowler-Nordheim Tunneling (FN) 8

2.2.1.3 Modified Fowler-Nordheim Tunneling (MFN) 8

2.2.1.4 Trap Assisted Tunneling (TAT) 8

2.2.2 Hot carrier injection (HCI) 8

2.3 Erasing 10

2.4 Issues with SONOS 10

2.5 Silicon nitride by LPCVD - Literature review 11

2.5.1 Silicon nitride in semiconductor industry 12

2.5.2 Silicon nitride in memory 14

2.6 Nature of the traps in silicon nitride 15

2.6.1 Amphoteric Trap model 15

2.6.2 Trap distribution 15

2.7 Trap characterization methods - Literature review 18

3 Silicon nitride process optimization and Results **20**

3.1 Experiments 20

3.1.1 Results of initial experiments 21

3.1.2 Solutions to reduce non-uniformity 23

3.2 Effect of N_2 dilution . . . 25
3.3 Change in SiH_4 flow . . . 26
3.4 Effect of pressure . . . 28
3.5 Effect of deposition temperature . . . 28

4 SONOS device fabrication and measurements **30**
4.1 Development of thin films for SONOS flash memory gate stack . . . 30
4.1.1 Tunnel oxide . . . 30
4.1.2 Blocking oxide . . . 31
4.1.3 N-poly gate . . . 32
4.2 SONOS structure fabrication and Characterization . . . 33
4.2.1 SONOS capacitor process flow . . . 34
4.2.2 SONOS memory stack characterization . . . 35

5 Trap characterization methods of MNS devices **37**
5.1 Trap density . . . 37
5.2 Trap capture cross section . . . 38
5.3 Retention Model for trap level extraction . . . 40

6 Extraction of trap parameters and Results **43**
6.1 Electrical Characterization Setup . . . 43
6.2 Results . . . 44
6.2.1 SiN_x film thickness and RI data . . . 44
6.2.2 Electrical characterization . . . 44
6.3 Trap Characterization Measurements . . . 45
6.3.1 Trap Density from Programming Transients . . . 46
6.3.1.1 C-V characteristics with program and erase . . . 47
6.3.2 Trap capture cross section . . . 49
6.3.3 Retention measurements for extracting the trap energy level . . . 50
6.4 Discussion . . . 51

7 Conclusions and Future Work **57**
7.1 Conclusions . . . 57
7.2 Future Work . . . 58

References 58

A Process recipes 63

A.1 Tunnel oxide 63

A.2 Silicon nitride 64

A.3 Low temperature oxide (LTO) 65

A.4 n-doped poly Si 65

A.5 Dopant activation anneal 65

A.6 Front side etching - Dry etching 65

A.7 Back side n-poly Si etch 66

A.8 Back side ONO stack etch 66

A.9 Back side Al contact 66

A.10 Forming gas anneal 66

List of Figures

1.1 Floating Gate flash. 2

1.2 Percolation path in thin and thick tunnel oxides [2]. 2

2.1 Evolution of SONOS NVSM device [6]. 5

2.2 The SONOS memory transistor. 6

2.3 I-V characteristics of tunnel oxide for memory [8]. 6

2.4 Band diagram during Direct tunneling. 7

2.5 Fowler-Nordheim tunneling. 8

2.6 Modified Fowler-Nordheim tunneling. 8

2.7 Trap assisted tunneling. 9

2.8 (a) Channel hot electron injection (b) Band diagram during CHEI. 9

2.9 Erasing in SONOS. 10

2.10 Band diagram of SONOS device with standard nitride [9]. 11

2.11 Erase saturation. 12

2.12 Trap distribution model by Ross and Wallmark [13]. 16

2.13 Energetical trap distribution model by Dorda and Pulver [14]. 16

2.14 Spatial and Energetical trap distribution model by Ferris-prabhu [15]. 17

2.15 Simple trap distribution model by White and Cricchi [16]. 17

2.16 Currents in n-channel SONOS device (a) Programming (b) Erasing [17]. 18

3.1 Typical three Zone LPCVD reactor. 21

3.2 Temperature vs time of a example recipe (not to scale). Standby temperature is 300^oC, process temperature is 780^oC and pressure is 300 mTorr. Base pressure is < 0.005 Torr. 22

3.3 Thickness nonuniformity in two cases (a) SiH_4- 15 sccm, NH_3-150 sccm at 780°C and 300 mTorr (for two wafers). (b) SiH_4- 15 sccm, NH_3-225 sccm at 780°C and 300 mTorr (measured in the center part of the wafer map in (a)). 23
3.4 (a) Bull's eye effect, SiH_4- 30 sccm, NH_3-75 sccm at 780°C and 300 mTorr. (b) Ring formation along the wafer circumference, SiH_4- 15 sccm, NH_3-45 sccm at 780°C and 300 mTorr. 23
3.5 Thickness variations after dummy wafers used during deposition for two recipes. SiH_4:NH_3 are 15:75 sccm and 30:75 sccm respectively at 780°C and 300 mTorr. . 24
3.6 Deposition rate and refractive index with change in NH_3 flow. SiH_4 - 15 sccm, Temperature - 780°C and Pressure - 300 mTorr. 24
3.7 Effect of N_2 dilution on (a) thickness (b) refractive index. SiH_4:NH_3 - 40:50 sccm, 780°C, 300 mTorr and 20 min. 25
3.8 Wafer map showing the measurement points on 4-inch wafer by ellipsometer. . . 25
3.9 Ring width vs N_2 flow. SiH_4:NH_3 - 40:50 sccm, 780°C, 300 mTorr and 20 min. . . 26
3.10 FTIR measurements on as-deposited SiN_x films. The legend in the figure indicates gas flow in sccm. Temperature-780°C, pressure-300 mTorr and deposition time-20 min. 26
3.11 Etch rate in 5:1 HF with different N_2 flow. SiH_4:NH_3 - 40:50 sccm, 780°C, 300 mTorr and 20 min. 27
3.12 Etch rate in 5:1 HF with different SiH_4 flow. SiH_4:N_2 ratio is 1:10, NH_3 - 50 sccm, 780°C, 300 mTorr and 20 min. 27
3.13 Effect of SiH_4 flow on thickness. SiH_4:N_2 ratio is 1:10, NH_3 - 50 sccm, 780°C, 300 mTorr and 20 min. 27
3.14 Effect of deposition pressure. SiH_4:NH_3:N_2 - 40:50:400 sccm, 780°C, 300 mTorr and 20 min. 27
3.15 Effect of deposition temperature and dilution on (a) thickness (b) refractive index. Numbers in the figure indicates temperature (^{o}C) - SiH_4 flow (sccm) - NH_3 flow (sccm) - N_2flow (sccm) - deposition time (minutes). 28

4.1 (a) C-V characteristics and (b) Hauser fit of the C-V of 4.2 nm tunnel oxide. . . 31
4.2 (a) I-E characteristics and (b) Weibull distribution plot of 4.2 nm tunnel oxide. . 32
4.3 (a) C-V characteristics and (b) Hauser fit of the C-V of 13.5 nm LTO after annealing at 850°C in N_2. 32

4.4 (a) J-E characteristics and (b) Weibull distribution plot of 13 nm LTO. 33
4.5 Sheet resistance map of n-poly film deposited at 630^0C with SiH_4:PH_3:N_2 - 90:20:210 sccm and 300 mTorr for 4500 seconds. 33
4.6 Programming transients. 35
4.7 P/E transients with improved blocking oxide. 35

5.1 (a) Schematic of SNS structure for trap density calculation (b) V_{FB} vs time during charge trapping. 37
5.2 Charge loss mechanisms in SONOS device [23]. 40

6.1 Characterization setup for for C-V, I-V, P/E and retention measurements. 44
6.2 (a) C-V characteristics and (b) J-E characteristics of SNS devices with different N_2 flow. The device area is 5.02E-5 cm^2. 45
6.3 Program transients of devices D1,D2 and D3. 46
6.4 C-V characteristics of D1 (a) Fresh device - after stressing with -14 V - after stressing with +14 V and light ON. (b) Fresh device - after stressing with +14 V and light ON - after stressing with -14 V -after stressing with +14 V and light ON. 48
6.5 C-V characteristics of D2 (a) Fresh device - after stressing with -14 V - after stressing with +14 V and light ON. (b) Fresh device - after stressing with +14 V and light ON - after stressing with -14 V -after stressing with +14 V and light ON. 48
6.6 C-V characteristics of D3 (a) Fresh device - after stressing with -14 V - after stressing with +14 V and light ON. (b) Fresh device - after stressing with +14 V and light ON - after stressing with -14 V -after stressing with +14 V and light ON. 49
6.7 Retention measurement on devices D1,D2 and D3 showing V_{FB} vs time. The devices are programmed for 1 s. 50
6.8 Electron trap distribution in devices D1,D2 and D3. 50

List of Tables

2.1 Charge injection mechanisms . . . 10

2.2 Trap distribution and memory performance . . . 15

4.1 Tunnel oxide parameters extracted from C-V by using hauser fit. Process conditions: T - 800^oC,O_2 - 5000 sccm, time - 10 minutes and annealed at 900^oC for 5 minutes in N_2 ambient. . . . 31

4.2 Blocking oxide parameters extracted from C-V by using hauser fit. Process conditions: SiH_4:O_2 - 40:100 sccm, N_2 - 500 sccm, temperature - 430^oC, pressure - 150 mTorr, Time - 5 minutes. . . . 31

4.3 SONOS device details . . . 34

6.1 Thickness and RI data of the SiN_x films deposited at 780^oC, 300 mTorr and with different N_2 flow. SiH_4-40 sccm and NH_3-50 sccm . . . 44

6.2 ΔV_{FB} values of samples D1,D2 and D3 with $\pm$14 V programming . . . 46

6.3 Calculated trap density values from program transients . . . 47

6.4 Current density values of D1,D2 and D3 in accumulation and inversion at 14 V . . . 50

6.5 Hole and Electron trap capture cross sections . . . 50

A.1 Process recipe for 4 nm tunnel oxide . . . 67

A.2 Process recipe for Silicon nitride . . . 68

A.3 Process recipe for LTO . . . 69

A.4 Process recipe for n-doped poly Si . . . 70

List of Abbreviations

BD	Blocking Dielectric
CHE	Channel Hot Electron
CHEI	Channel Hot Electron Injection
CMOS	Complementary Metal Oxide Semiconductor
CTF	Charge Trap Flash
C-V	Capacitance - Voltage
DCS	Di-chloro silane
DCTS	Discharge Current Transient Spectroscopy
DT	Direct Tunneling
EOT	Equivalent Oxide Thickness
FG	Floating Gate
FN	Fowler - Nordheim
FTIR	Fourier Transform Infrared spectroscopy
GCR	Gate Coupling Ratio
HF	Hydro Flouric Acid
HNA	HF/Nitric acid/Acetic acid
ITRS	International Technology Road map for Semiconductor industry
LPCVD	Low Pressure Chemical vapor Deposition
LTO	Low temperature Oxide
MFN	Modified FN
MNS	Metal/Silicon nitride/Silicon
MNOS	Metal/Silicon nitride/Oxide/Silicon
MOS	Metal/Oxide/Silicon
NC	Nano Crystal
ONO	Oxide/Silicon nitride/Oxide

P/E	Program/Erase
p-Si	p- type Silicon
RBS	Rutherford Back Scattering
RI	Refractive Index
RIE	Reactive Ion Etching
RTP	Rapid Thermal Processing
SNOS	poly Silicon/Silicon nitride/Oxide/Silicon
SNS	poly Silicon/Silicon nitride/Silicon
SONOS	poly Silicon/Oxide/Silicon nitride/Oxide/Silicon
TAT	Trap Assisted Tunneling
TOX	Tunnel Oxide
XPS	X-ray Photoelectron Spectroscopy

Constants

Electron charge $q = 1.602 \times 10^{-19}$ C

Free electron mass $m_0 = 9 \times 10^{-35}$ g.J/erg

Boltzmann's constant $k_B = 1.38 \times 10^{-23}$ J/K

Planck's constant $h = 6.62 \times 10^{-34}$ J.s

Permittivity of free space $\epsilon_0 = 8.854 \times 10^{-14}$ F/cm

List of symbols

X_{OX}	Thickness of oxide (nm)
X_N	Silicon nitride thickness (nm)
X_T	Thickness of the charge trapped region in Silicon nitride (nm)
ϕ_{OX}	Barrier height at Si-SiO_2 interface w.r.t Si conduction band (eV)
ϕ_N	Barrier height at SiO_2-Si_3N_4 interface (eV)
ϕ_t	Energy level of the trap in silicon nitride (eV)
E_{OX}	Electric field in the oxide (MV/cm)
E_N	Electric field in silicon nitride (MV/cm)
ϵ_{OX}	Dielectric constant of SiO_2
ϵ_N	Dielectric constant of silicon nitride
ρ	Spatial charge density (C/cm^3)
σ_t	Trap capture cross section (cm^2)
ϑ_d	Drift velocity of the carriers in silicon nitride (cm/s)
ξ	Total emission rate from traps (#/s)
J	Current density in the silicon nitride (A/cm^2)
E_{TA}	Trap energy level (eV)
$g(E_{TA})$	Density function of traps (# cm$^3 eV^{-1}$)
C_N	Capacitance of MNS/SNS structure (F/cm^2)
N_t	Total number of traps (#/cm^3)
n_t	Number of filled traps (#/cm^3)
n_c	Number of available carriers for trapping (#/cm^3)
V_G	Gate voltage
V_{FB}	Flat-band voltage
ΔV_{FB}	Change in flat-band voltage
ΔQ_P	Change in positive trapped charge

ΔQ_N	Change in negative trapped charge
V_{TH}	Threshold voltage
ΔV_{TH}	Change in threshold voltage

Chapter 1

Introduction

Non volatile semiconductor flash memory market is one of the driving forces of the present semiconductor industry. This storage technology simply consists of storing a few electrons surrounded by a dielectric well with the application of suitable voltage. Though the concept looks simple, producing reliable and high density memories is a challenging task. Floating gate flash is a mature technology in the semiconductor memory area that has been well understood and demonstrated from 0.8 μm to 45 nm technology nodes. The requirement of high density data storage in portable storage drives enables continuous floating gate (FG) flash scaling up to 32 nm node, but it may not be possible for further scaling.

Floating gate device stores charge in a polysilicon floating gate that is surrounded by insulators from all sides. The floating gate is thicker than both tunnel oxide and blocking dielectric. The floating gate device is similar to that of standard MOSFET except the gate stack. In standard MOSFET the gate stack consists of single oxide layer followed by poly-Si/metal gate whereas in floating gate devices the gate stack is made up of Tunnel Oxide/Poly-Si FG/Blocking Oxide/poly or metal gate (Figure 1.1).

The performance of the FG flash memory depends on the thickness of the layers used in the gate stack. Thick tunnel oxide is good for long term retention but higher programming and erase voltages are not compatible with the present CMOS technology. To be consistent with the scaling trend tunnel oxide thickness must be reduced below 4 nm. It is very difficult to grow such good quality thin tunnel oxides without any pin hole defects. Since the electrons stored in conductive floating gate, any single pin hole defect can cause tunneling of electrons back to substrate through Poole-Frenkel mechanism. The tunnel oxide must be thick enough to prevent charge tunneling through percolation path for long term retention. Defects in the oxide are created by electrical stress during P/E cycles. The defect formation continues until a chain of defects is formed called

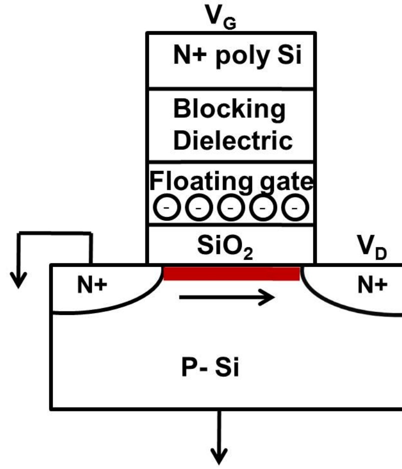

Figure 1.1: Floating Gate flash.

percolation path. The red circle in Fig. 1.2 shows the characteristic length of the electron while tunneling and the blue dot represents the defect in oxide [2]. If the thickness of the tunnel oxide is comparable to tunneling characteristic length of the electron (as is the case with Ultra Thin Tunnel oxides), a single defect can leak the charge through Poole-Frenkel conduction, whereas this is not a problem with thicker oxides.

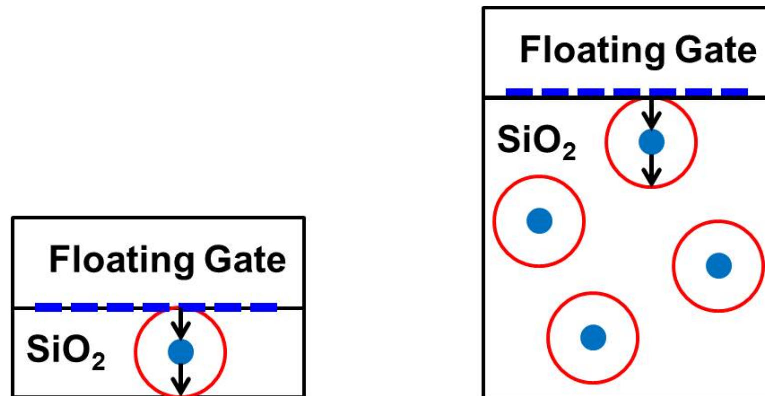

Figure 1.2: Percolation path in thin and thick tunnel oxides [2].

As per ITRS 2009, the tunnel oxide thickness for floating gate flash memory is 6-7 nm (up to 32 nm node) and the P/E voltages are 17-19 V and 15-17 V for 45 nm and 32 nm technology nodes respectively [1]. Tunnel oxide with thickness >8 nm can be relatively leak free because a single defect in the oxide is insufficient to provide a leakage path, but oxides with thickness less than 8 nm are vulnerable to leakage through defects. This limits the scaling of floating gate technology in vertical direction. Also higher P/E voltages causes reliability problems such as stress induced leakage currents and hot carrier degradation.

Gate coupling ratio (GCR) is another concern with the floating gate devices with respect

to scaling. The GCR indicates the amount of control by the control gate over channel. GCR is defined as C (control gate to floating gate capacitance)/C (total floating gate capacitance). The GCR must be >0.60 for the control gate to be effective. But the scaling limitation on the blocking oxide to reduce leakage does not permit to achieve GCR > 0.5 [2]. To achieve high GCR, floating gate is wrapped around by the control gate in the width direction. With continuous scaling there is no space left for the control gate to wrap around the floating gate after the blocking dielectric filling. This limits the scaling in lateral direction.

Another problem in FG flash is inter-cell interference. As the scaling continues, cell to cell interference increases due to the physical proximity of the devices. To reduce inter-cell interference with continuous scaling the floating gate must be thin. But recent results demonstrated that there is a limit to thinning of the floating gate up to ~10 nm below which further thinning is not possible. This is due to the ballistic transport of those electrons which are not scattered in the thin poly floating gate during programming [3]. These ballistic currents in scaled floating gate devices causes reliability problems, reduces P/E speeds and limits the FG scaling. Overall further scaling of FG flash does not seem to be effective and alternative technologies must be considered to drive the scaling of the semiconductor non-volatile memory.

The main issue with FG flash is the charge storage in a conductive poly-Si layer which is continuous. If the charge storage is discrete in nature then the fundamental problem of scaling the tunnel oxide can be solved (to the extent that direct tunneling limits the retention). Charge Trap Flash (CTF) is the best alternative to the floating gate flash in which charge is stored in the discrete form. Two types of CTF technologies are under research. One is SONOS (Si/Oxide/Nitride/Oxide/Poly-Si) technology in which the charge is stored in the traps in the nitride[0] layer. The other CTF technology is the Nanocrystal (NC) memory in which the storage node is either Si or metal nanocrystals. The nanocrystal memory has its own problems such as poor programming efficiency due to smaller density and area coverage, poor size distribution etc.

The main motivation for SONOS technology is that it allows scaling in both directions and can be operated with lower P/E voltages compatible with the CMOS technology. Also the SONOS technology is immune to radiation (radiation hard) and can be used in space applications.

The main focus of this work is to develop thin silicon nitride films for the application of charge trap layer in SONOS flash memories. LPCVD is the best method for depositing such memory quality nitride films. The performance of memory stack depends on the composition of charge trap layer. Si dangling bonds act as the trap centers which are responsible for memory

[0]Here nitride means Silicon nitride. The same convention is followed elsewhere in this report.

action [4]. Si-rich nitride is needed for faster programming where as for long term retention the traps in the nitride should be deep [23]. Therefore proper (spatial and energetic) distribution of traps is very important and this depends on the processing conditions. The deposition conditions should be free from oxygen and hydrogen impurities as well as metallic impurities. The metallic impurities cause charge loss whereas the O and H atoms reduce the trap density [4].

In this work, development of good quality uniform silicon nitride films from SiH_4-NH_3-N_2 system with varying stoichiometry and thickness is presented. The memory properties of as-deposited films were investigated and trap parameters are extracted.

1.1 Organization of the Report

1. Chapter 1 The present chapter served as introduction.
2. Chapter 2 Literature review, describes the basic principles of SONOS technology and nature of traps.
3. Chapter 3 Silicon nitride process optimization, issues and results.
4. Chapter 4 SONOS device fabrication and P/E measurements.
5. Chapter 5 Trap density analysis and trap characterization methods.
6. Chapter 6 Trap characterization experiments, results and discussion.
7. Chapter 7 Conclusion and Future work to be done.

Chapter 2

Literature Review

2.1 Evolution of SONOS memory

Early charge trapping devices in 1960s were MNOS (Metal/Nitride/Oxide/Si) type. The first MNOS device was a p-channel device with thick silicon nitride (~45 nm) as charge trap layer and thin tunnel oxide (2 nm). The gate was Al metal gate (Figure 2.1). Because of thicker charge trap layer, the programming and erasing voltages were relatively high (~30 V). These memory devices suffered from low speed and less density.

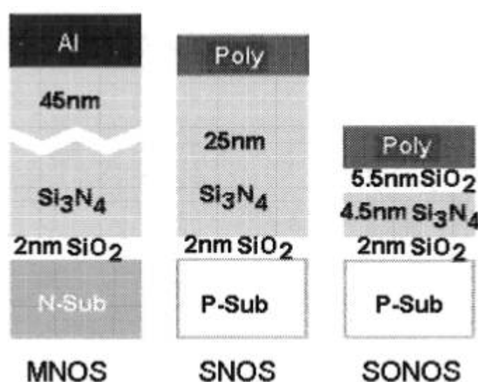

Figure 2.1: Evolution of SONOS NVSM device [6].

With the advent of LPCVD, researchers started depositing high quality silicon nitride and poly silicon films. Slowly the MNOS device transformed into SNOS (Poly-Si/Nitride/Oxide/Si) device with thinner dielectric layers. The nitride layer was reduced to 25 nm whereas the tunnel oxide was 2 nm. This resulted in lower programming and erase voltages compared to MNOS devices. Programming and erasing of these devices were done by Fowler-Nordheim tunneling. Later many researchers proposed scaled SNOS devices in which the nitride layer was scaled to 20 nm and the tunnel oxide of 1.6 nm with a low programming voltage of 10 V. The main problem

with the scaled SNOS devices was hole injection from the gate that limits the memory window.

The charge injection from gate can be reduced by sandwiching an oxide layer between the poly-Si gate and nitride layer [7]. This idea led to the present form of SONOS transistor. The advantages of triple dielectric structure are:

- Reduced programming voltages due to reduction in nitride thickness
- Because of the top oxide, charge injection from and to the gate electrode is minimized.
- Since there is minimal loss of charge through the gate, the memory window is improved.

The SONOS device is triple dielectric device in which a silicon nitride layer is sandwiched between two oxide layers as shown in Fig. 2.2. Charge storage takes place in the traps distributed in the nitride layer. The bottom oxide called as tunnel oxide permits the charge transfer from silicon substrate to the charge trap layer through quantum mechanical tunneling.

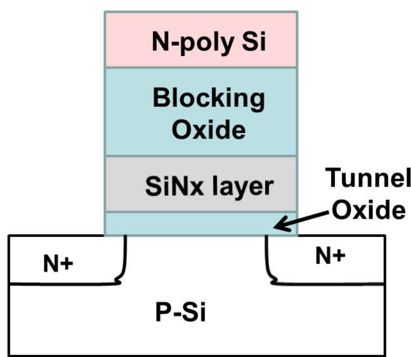

Figure 2.2: The SONOS memory transistor.

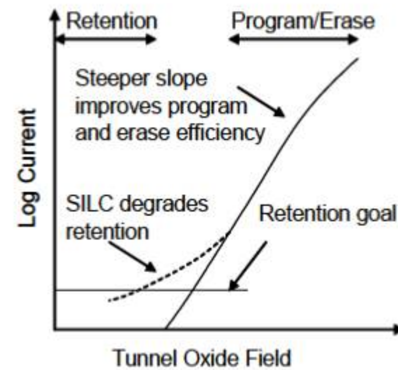

Figure 2.3: I-V characteristics of tunnel oxide for memory [8].

The thickness of the tunnel oxide decides the programming efficiency and retention characteristics. Figure 2.3 shows the requirements of the tunnel oxide in the memory devices. Leakage at low fields should be very less to improve the retention whereas higher tunneling currents are needed at moderate fields for faster and low voltage programming.

The top oxide layer between poly-Si gate and charge trap layer is called as blocking oxide. As the name indicates its function is to block the charge transfer to and from the gate. The blocking oxide thickness is decided in such a way that it should not allow any charge injection from the gate.

2.2 SONOS charge injection mechanisms

The charge injection mechanisms in SONOS devices are broadly classified into two categories:

- Tunneling based charge injection and
- Hot carrier injection

The type of charge injection depends on how the SONOS device is connected in memory configuration, either NAND or NOR.

2.2.1 Tunneling based charge injection

This type of charge injection is based on the quantum mechanical tunneling of carriers through oxide. In these mechanisms the required voltage is applied between the gate and substrate of the memory transistor. SONOS devices in NAND configuration use tunneling based charge injection. The different tunneling mechanisms are :

1. Direct band to band tunneling (DT)
2. Fowler-Nordheim tunneling (FN)
3. Modified Fowler-Nordheim tunneling (MFN)
4. Trap Assisted Tunneling (TAT)

The type of tunneling depends on the thickness of the dielectric layers and applied field.

2.2.1.1 Direct band to band tunneling (DT)

In this mechanism, the carriers tunnel from the silicon conduction band to nitride traps through the oxide directly. The band diagram for this type of tunneling is shown in Fig. 2.4. Direct tunneling depends on the thickness of the oxide and occurs when the fields are not very high. In the SONOS devices with ultra-thin oxides (< 3 nm), the dominant injection mechanism is Direct tunneling.

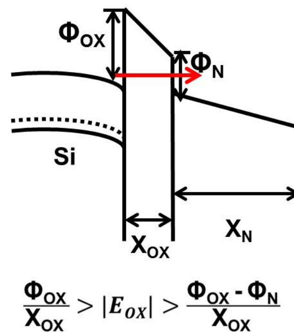

Figure 2.4: Band diagram during Direct tunneling.

2.2.1.2 Fowler-Nordheim Tunneling (FN)

This is the most prominent charge injection mechanism in flash memory devices. In fact this is field assisted tunneling. Under the application of large electric fields (~10 MV/cm), the electrons in the silicon conduction band see a triangular energy barrier whose width depends on the applied field. The width of the barrier becomes so thin that electrons in the silicon conduction band tunnel through the barrier (Figure 2.5). Since the effective barrier width is reduced, the FN tunneling currents are higher than DT currents.

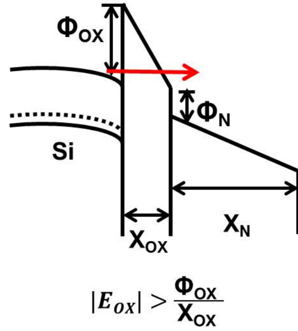

Figure 2.5: Fowler-Nordheim tunneling.

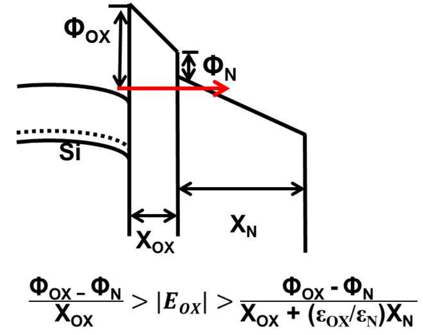

Figure 2.6: Modified Fowler-Nordheim tunneling.

2.2.1.3 Modified Fowler-Nordheim Tunneling (MFN)

In this mechanism, the carriers first tunnel through the oxide directly into the nitride band gap and then through FN tunneling to the nitride conduction band (Figure 2.6). This depends on the barrier height and the barrier height on the nitride side should be less than the barrier on gate or substrate.

2.2.1.4 Trap Assisted Tunneling (TAT)

This is similar to MFN. In this case the carriers are captured by the traps in nitride close to nitride conduction band by direct tunneling. Then the trapped carriers tunnel from the traps to the conduction band (Figure 2.7). Since the trap is closer than the conduction band edge in the tunneling path, TAT currents are higher compared to that of MFN.

2.2.2 Hot carrier injection (HCI)

In tunneling based charge injection mechanisms the electric field in the channel is constant and depends on the voltage applied between gate and substrate. Whereas in hot carrier injection, voltage is applied to both gate and drain terminals such that the carriers in the channel see a

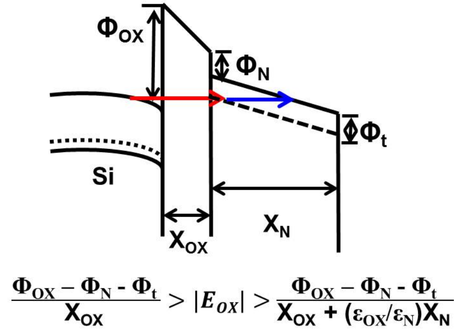

Figure 2.7: Trap assisted tunneling.

gradual increase in lateral electric field towards drain and becomes energetic (Figure 2.8). Those 'hot'energetic carriers that are having sufficient energy to cross the oxide barrier will be trapped in the nitride . For the SONOS devices made on p-Si substrate, Channel hot electron (CHE) injection is the hot carrier injection mechanism. SONOS devices in NOR configuration use HCI technique for programming.

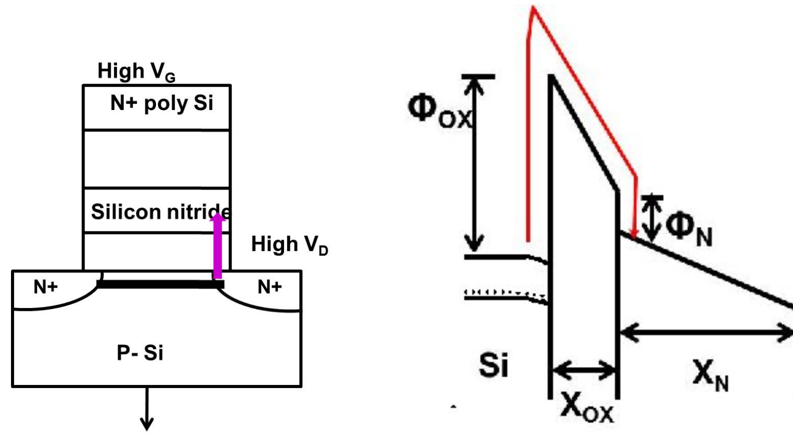

Figure 2.8: (a) Channel hot electron injection (b) Band diagram during CHEI.

With the tunneling mechanisms used for programming, the charge distribution in the nitride layer is uniform whereas with CHE programming the charge is localized near drain. The trapping efficiency is better in the tunneling charge injection compared to hot carrier injection. Since, the carriers are accelerated by lateral field, programming time with HCI is small. Also, CHE programming needs higher voltages and results in higher charge pump voltages. Table 2.1 summarizes the charge trapping mechanisms.

[1]Since the electrons overcome the barrier at Si-SiO_2 interface, CHEI can be used with the oxides of any thickness. But the thickness decided by leakage during retention (Section 2.1).

[2]Both vertical and lateral fields are needed for carrier injection. Lateral field ensures that carriers get enough energy, vertical field favors the charge injection.

Injection Mechanism	Tunnel oxide thickness	Electric Field
FN Tunneling	Thick (∼5nm)	Very High (10MV/cm)
Direct Tunneling	Thin (< 3nm)	High (6-7 MV/cm)
MFN Tunneling	Thin	Medium
TAT	Thin	Low
CHE Injection	Any [1]	depends on the channel length [2]

Table 2.1: Charge injection mechanisms

2.3 Erasing

The stored charge in a SONOS device is removed by tunneling in both configurations NAND and NOR. By applying a voltage of opposite polarity (to that of programming) between gate and substrate, the fields are opposite and the energy bands are in such a way that hole tunneling from substrate takes place. This neutralizes the electrons stored in the nitride layer. At the same time FN tunneling of electrons from nitride to Si conduction band also takes place. But the main component for erasing is the hole direct tunneling from Si valance band to nitride Fig. 2.9. Therefore, the thickness of tunnel oxide should be minimum for faster erase but this increases the chance of low field leakage during retention and memory window is reduced.

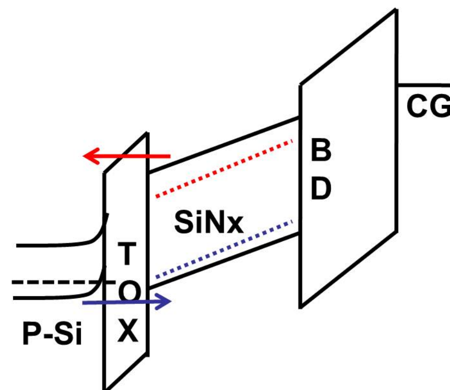

Figure 2.9: Erasing in SONOS.

2.4 Issues with SONOS

The localized charge trapping nature of the SONOS devices allows further scaling and overcomes some of the disadvantages of floating gate technology. It allows higher density due to reduced cell size in both directions, lower programming and erasing voltages, excellent scalability, compatibility with existing CMOS process technology and superior resistance to radiation. In spite of many advantages there are some issues with SONOS technology such as poor trapping efficiency, short retention times and erase saturation. These problems slowed down the growth of the SONOS

technology and needs to be resolved before replacing the existing floating gate technology.

Because of smaller barrier heights between nitride and surrounding oxides (~1.3 eV) (Figure 2.10) the trapped charges in the nitride escape easily and leads to poor trapping efficiency [9]. If the traps are located at deeper energy levels, this small barrier height is not an issue but in silicon nitride the traps are distributed every where. Therefore, the electrons trapped in those shallow traps distributed near the band edge can escape easily.

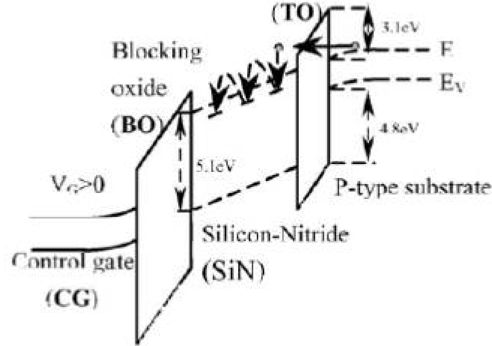

Figure 2.10: Band diagram of SONOS device with standard nitride [9].

The deep trap levels in the standard nitride is another concern for the SONOS memories. Detrapping is very slow and high gate voltages (*i.e*, High electric fields) are not much helpful to detrap the electrons from these deep traps. Moreover injection from the gate is a problem at high gate voltages. One solution to this is to use ultra thin oxides which allow hole tunneling from the substrate. But with ultra thin oxides low field leakage is a problem and this leads to retention loss even at room temperature thus reduces memory window. If thicker oxides are to be used to reduce low field leakage then erase speed will be reduced. Thus there is a trade-off between retention and erase speed [10].

Another problem with SONOS is the erase saturation. During erasing by FN tunneling, electron injection from gate (through FN tunneling) is faster than hole tunneling from the substrate to nitride layer. This is shown in Fig. 2.11. The memory device is not completely erased and the threshold voltage is clamped at single value. The solution to this problem is to use p-poly gate instead of n-poly gate or use high work function metal gate such as TaN [11].

2.5 Silicon nitride by LPCVD - Literature review

This section highlights some of the research done on silicon nitride from LPCVD. The main sources of Si reacting species in LPCVD silicon nitride are either chlorosilanes ($SiCl_4$ or SiH_2Cl_2)

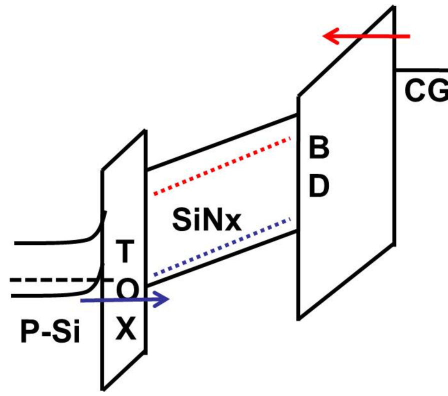

Figure 2.11: Erase saturation.

or monosilane (SiH_4) and the nitrogen source is NH_3. The other Si source under research is Si_2H_6 (Disilane). The first LPCVD nitride was produced by SiH_4 and NH_3. Very few research groups have experimented with SiH_4-NH_3 system. Due to the non-uniformity issues the SiH_4-NH_3 process was quickly abandoned. Majority research groups across the globe are currently using SiH_2Cl_2-NH_3 process.

2.5.1 Silicon nitride in semiconductor industry

After the introduction of LPCVD in the semiconductor industry, synthesis of silicon nitride films by using LPCVD was first reported by R.S.Rosler et al., [29]. They first used SiH_4, NH_3 gases with N_2 under different pressures and gas flows in a tubular hot wall reactor. The deposition temperature ranges from 780^oC-850^oC. The thicknesses obtained were in the range of 50 nm-200 nm and the refractive index from 1.98 to 2.02 for different gas flows and temperature profiles. They reported a large thickness non-uniformity of $\pm 20\%$ within wafer. They observed the bull′s eye effect of being much thicker deposition at the wafer edges compared to the center part of the wafer. With the modification of quartz boat they reported better within wafer uniformity of ± 2 to 5%. They also reported that the deposition rate and thickness uniformity were sensitive to wafer spacing. Etch rates (in 48% HF) of 1.2-1.5 nm/min were reported.

The same group reported the silicon nitride films from SiH_2Cl_2 and NH_3. They mentioned that excellent uniformity of $\pm 3.2\%$ has been obtained without any quartz modification. The refractive index obtained was 2 and the reported etch rates in 48% HF were 1.8-2.1 nm/min. They also have reported that the SiH_2Cl_2-NH_3 system is very insensitive to wafer spacing. This may be due to difference in diffusivity between SiH_2Cl_2 and SiH_4.

Chramova et al., [43] deposited amorphous Si_3N_4 films to study the oxidation effects of silicon nitride films. They used 2.5% SiH_4 (diluted in Ar) and NH_3 to deposit 100-300 nm thick films.

No details about thickness and compositional uniformity were mentioned. The IR measurement results indicate strong Si-N peak at 900 cm^{-1} - 1200 cm^{-1} and N-H peak at 3000 cm^{-1} - 3500 cm^{-1}.

The effect of change in precursor gas flow on the composition and structure of silicon nitride was reported by T.Makino [30] and correlated the refractive index with N/Si atomic ratio. The gas flow ratio (R = NH_3 flow/ SiH_2Cl_2 flow) was varied from 0.1 to 10 and the deposition temperature was 770^oC. The RBS measurements show that the silicon concentration increases with the decrease in R. As R increases the refractive index and etch rate in 49% HF decreases.

Popova et al., [31] reported the electrical characteristics (I-V and C-V measurements) on thin silicon nitride films deposited from SiH_2Cl_2 and NH_3 at 800^oC with 500 mTorr pressure. The deposition rate was around 6 nm/min with gas flow ratio (SiH_2Cl_2:NH_3) of 1:3. They concluded, from the RBS measurements, that LPCVD process yields stoichiometric films. Their conduction model shows that both FN tunneling and Poole-Frenkel conduction are the dominant mechanisms.

Many researchers modeled the LPCVD silicon nitride process from SiH_2Cl_2-NH_3 ([32], [33], [34]). Carlos et al., [35] studied the thermophysical properties of low stress Si-rich silicon nitride films for sensor and actuator applications prepared from SiH_2Cl_2-NH_3.

Schied et al., produced the silicon nitride films from Si_2H_6 and NH_3 [20]. They compared all the three processes (SiH_4-NH_3, SiH_2Cl_2-NH_3 and Si_2H_6-NH_3). The main advantage of Si_2H_6 process is the low deposition temperature (600^oC) compared to other processes. They also reported large non-uniformity in thickness in the silicon nitride films produced from the SiH_4-NH_3 process compared to other two processes. As-deposited films from Si_2H_6-NH_3 show chemical resistance to buffer HF is comparable to that of standard nitride. They reported a breakdown field of 7 MV/cm for stoichiometric nitride (with RI of 2).

Liu et al., [36] reported silicon nitride films from SiH_4-NH_3-N_2 system with varying composition from stoichiometric to Si-rich. They studied the deposition kinetics and concluded that the deposition mechanism follows Langmuir-Hinhselwood theory where the surface adsorption and desorption of the SiH_4 limit the reaction rate. They also investigated the effects of temperature, pressure and gas flow ratio. The FTIR measurements showing strong peak at 837 cm^{-1} suggests the Si-N bond and no other peaks relating to hydrogen were observed.

The detailed analysis of LPCVD silicon nitride from SiH_4 and NH_3 is given by Yacoubi et al. [37]. They reported 26 gas phase reactions are involved in the final film deposition on the silicon surface. They also observed a significant radial non-uniformity in thickness (bull's eye).

They proposed complete modeling of the silicon nitride from SiH_4-NH_3.

2.5.2 Silicon nitride in memory

The early memory transistors were MNOS and the silicon nitride in those structures obtained by normal pressure CVD at a temperature of 900°C [16]. Fujita et al., [4] used the LPCVD silicon nitride at 700°C with SiH_4, NH_3 and Ar as carrier gas to study the behavior of trap states by changing the SiH_4/NH_3 ratio. They concluded that Si dangling bonds create deep trap states and the trap density increased from 2E19 cm^{-3} to 8E19 cm^{-3} with increase in SiH_4/NH_3 ratio from 0.001 to 0.1.

The variation in stoichiometry of the silicon nitride and its relation to memory traps was studied by Bailey and Kapoor [38]. They used single wafer cold wall reactor with SiH_4 and NH_3 gases at temperatures ranging from 700°C to 950°C. They reported an increase in Si concentration with decreasing gas ratio NH_3 : SiH_4 from 300:1 to 25:1 and corresponding trapped electron density from 9.5E17 cm^{-3} to 2.1E18 cm^{-3}.

Brown et al., [39] have used 45 nm LPCVD silicon nitride in the MONOS memory transistor deposited at 750°C. They studied the effect of N_2 anneal at 900°C after the formation of blocking oxide and concluded that the memory window is increased by 40%. They attributed this to annihilation of Si-H bonds with N_2 anneal.

Kamigaki et al., [40] calculated the electron and hole traps in amorphous silicon nitride films from LPCVD from SiH_2Cl_2 and NH_3. The deposition temperature was 790°C and other parameters were not mentioned. They used MNOS structures to extract the trap densities and reported electron and hole trap densities were 7E18 cm^{-3} and 1.2E19 cm^{-3} respectively.

Libsch and White [41] studied the charge transport in SONOS devices and proposed a low voltage SONOS device programmable with 5 V. The reported deposition conditions were: SiH_2Cl_2 : NH_3 - 100:30 (sccm), pressure 300 mTorr and temperature 725°C. The deposition time was varied from 4 minutes to 8 minutes to get thicknesses from 5 nm to 8 nm. Almost all publications from the research group led by M. H. White in the SONOS area use the LPCVD silicon nitride from SiH_2Cl_2 and NH_3.

Wong et al., [42] reported high quality dielectric films with small hydrogen content for non-volatile memory applications by using $SiCl_4$ and NH_3 at 820°C. Yun et al., [21] investigated the formation of Si-rich silicon nitride for nanoscale nonvolatile memory applications. They have used SiH_2Cl_2 and NH_3 at different pressures (100-600 mTorr) and temperatures (685°C-785°C) with varying flow ratios. They reported that increase in leakage current and large flat band

voltage shift for the Si-rich (RI-2.255) nitride films compared to the films with RI 1.96.

Thus majority of the research groups uses the SiH_2Cl_2 - NH_3 process for depositing the silicon nitride films. Only very few groups worked with SiH_4 and NH_3 precursors in horizontal reaction tubes for the development of silicon nitride films. The aim is to develop uniform silicon nitride films by using SiH_4 and NH_3 gases in LPCVD system.

2.6 Nature of the traps in silicon nitride

The traps in the silicon nitride layer are responsible for the memory action in the SONOS device. Therefore the performance of the SONOS device depends on the distribution and density of the traps in the silicon nitride layer. Table 2.2 compares the erase speed and retention time for different energy distribution of traps. The carriers trapped in shallow traps have smaller emission times and therefore shorter retention time.

Density of Traps	**Energy Distribution**	**Erase speed**	**Retention time**
High	Shallow	Fast	Short
High	Deep	Slow	Good

Table 2.2: Trap distribution and memory performance

2.6.1 Amphoteric Trap model

The origin of traps is attributed to the Si dangling bonds during silicon nitride film formation [12]. A saturated Si-N bond has four N atoms covalent bonded with Si atom whereas in a dangling Si-N bond the Si atom is short of one N atom. This incomplete covalent bonding creates an energy level in the band gap of Si and acts as trap center. When a single electron is attached to this dangling bond, the charge state is neutral (D^0). When two electrons are attached the charge state is negative (D^-) and the state is positive (D^+) when no electron is attached. The negatively charged D^- state is a hole trap and the D^+ is an electron trap.

2.6.2 Trap distribution

Since the earlier nitride based devices were MNOS type, the studies on trap properties and their distribution were entirely on MNOS devices. Researchers have proposed different trap distributions to study the charge trapping and detrapping mechanisms. Ross and Wallmark [13] assumed that the traps located at single energy level and distributed uniformly (Figure 2.12). They used direct tunneling theory to explain trapping and detrapping. During programming

electrons tunnel from the valance band of silicon to the traps in the nitride. During detrapping, the trapped electrons in the nitride tunnel through the oxide to the conduction band of Si.

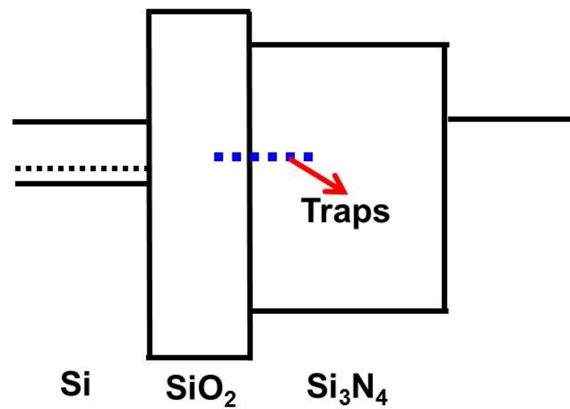

Figure 2.12: Trap distribution model by Ross and Wallmark [13].

They considered rectangular barrier for tunneling calculations and found that the flat band voltage shift ΔV_{FB} is proportional to logarithm of the applied pulse width.

$$\Delta V_{FB} \propto (0.577 + ln(\frac{t}{t_0})) \tag{2.1}$$

where t_0 is the time required for one transition from conduction band (valance band) edge to trap center.

Dorda and Pulver [14] considered spatially fixed but energetically distributed traps (Figure 2.13). They used direct tunneling model to explain trapping and detrapping. They neglected the conduction current in the nitride and considered the effect of transferred charge on the tunneling probability. They proposed that the traps are located at the nitride-oxide interface.

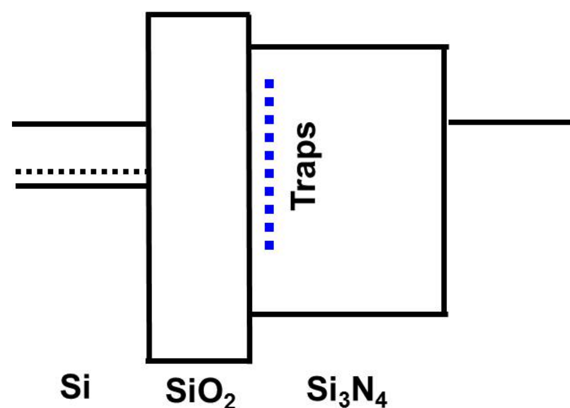

Figure 2.13: Energetical trap distribution model by Dorda and Pulver [14].

As per their theory the shift in flat band voltage given by

$$\Delta V_{FB} \propto 1 - e^{[-C_1(1-e^{(-C_2 t)})]} \tag{2.2}$$

where C_1 and C_2 are positive constants determined by the device structure and material properties. From the tunneling kinetics they found the trap distribution is continuous and trap density of $10^{13} cm^{-2} eV^{-1}$.

Ferris-Prabhu [15] considered traps distributed spatially and energetically (Figure 2.14). By using direct tunneling and charge transport equation they have shown that the charge transfer with respect to time is initially linear and then varied logarithmically.

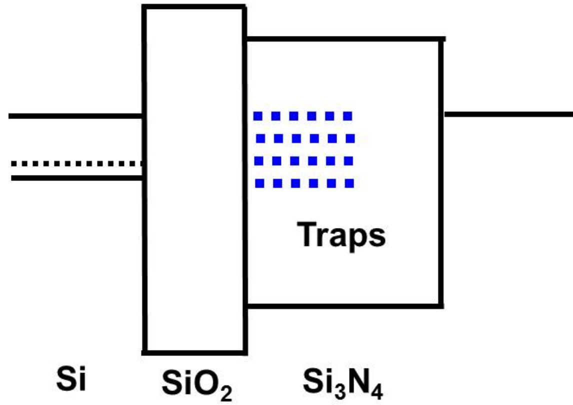

Figure 2.14: Spatial and Energetical trap distribution model by Ferris-prabhu [15].

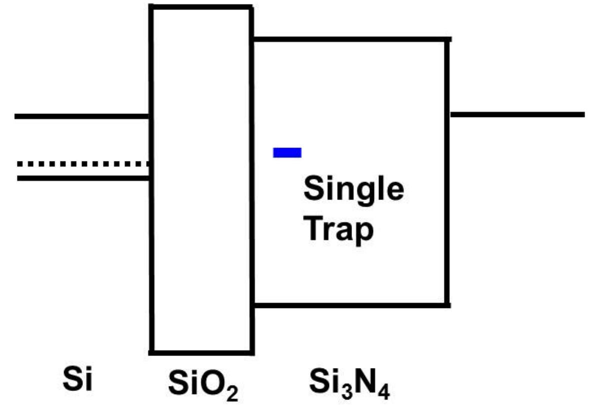

Figure 2.15: Simple trap distribution model by White and Cricchi [16].

All the above models assumed rectangular barriers for tunneling approximation and no effect of transferred charge on the tunneling. White and Cricchi [16] assumed a simple trap distribution with single trap located near the oxide-nitride interface (Figure 2.15).

They considered the effect of transferred charge on further tunneling of carriers and used direct tunneling theory for charge transfer from deep traps to Si conduction band. Their results indicate that the shift in V_{TH} is proportional to logarithm of pulse duration.

The above models used thin oxide MNOS devices and the direct tunneling theory is applicable. Similar theory can also be applied to SONOS devices. But the tunneling mechanism depends on thickness of the ONO stack. Bachhofer et al., [17] studied the transient conduction in the SONOS devices and suggested one of the tunneling mechanisms (depends on the field and thickness) is responsible for charge transfer. The various conduction mechanisms in a n-channel SONOS device under both biases is shown in Fig. 2.16.

When the gate bias is positive, during charge trapping, electron tunneling from substrate to

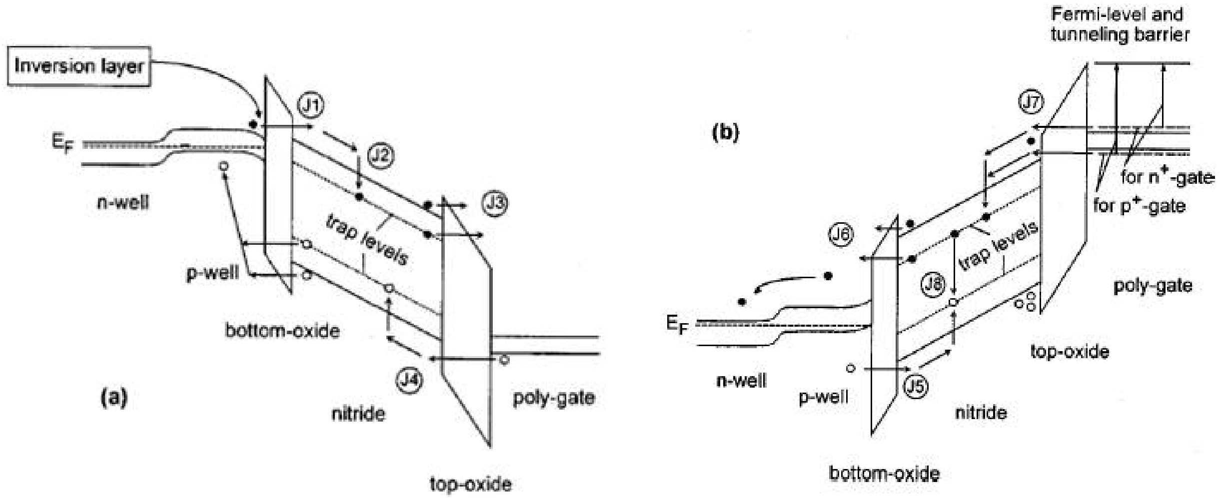

Figure 2.16: Currents in n-channel SONOS device (a) Programming (b) Erasing [17].

nitride traps is the dominant component (J1). These electrons are then redistributed by Poole-Frenkel conduction (J2). Those electrons that are reaching the blocking oxide interface may tunnel through the blocking oxide and collected by the gate (J3). This component is negligible due to the blocking action of the top oxide. The component J4 is the hole injection from the gate and is negligible. During erase (with negative gate voltage) hole tunneling from the substrate is the main component (J5). The electric field is such that electrons tunnel from traps to substrate (J6). Electron injection from gate (J7) is the other component and is the main reason for erase saturation. Electron-hole recombination (J8) in the nitride is another component during charge transfer.

2.7 Trap characterization methods - Literature review

Discharge current measurement, change in flat band voltage ΔV_{FB} and retention measurement *i.e.*, decay in ΔV_{FB} w.r.t time are the important characterization methods to evaluate the traps in MNOS/SONOS devices. Discharge current measurement and ΔV_{FB} are used to calculate the trap density whereas the retention measurement is used to estimate the energy levels of the traps.

Matsuura et al., [24] used DCTS technique for evaluating the trap density and their distribution. In this method, sufficiently high voltage pulse of suitable polarity is applied to the MNS device for complete charge trapping. After the pulse is applied, the discharge current from the MNS capacitor is measured and analyzed to evaluate the trap density and trap distribution. MNS devices made on p-Si with 76 nm SiN_x layer and Al gate were used for DCTS measurements. A voltage pulse of 50 V and width 600 s was applied. They considered the hole trapping

in the nitride. The discharge currents were measured up to 10^3 seconds. The calculated trap density was 2E17 cm^{-3} in the energy level 0.83 eV $<$ E-E$_V$ $<$ 0.94 eV.

Roy and White [25] used retention model to extract the spatial distribution and density of traps. During the retention measurement the MONOS transistor is biased such that the Si surface is in either weak inversion or depletion. Therefore the back tunneling of the captured charge is the dominant component for the threshold voltage decay. They used the ΔV_{TH} decay *vs* time data to calculate the trap distribution. The extracted trap density was 1E19 cm^{-3} and located from 2-3 nm from the tunnel oxide-nitride interface.

Martin and Aymhich [26] characterized the charge distribution in the nitride layer by using ΔV_{FB} decay with field assisted discharge. Since the charge stored in the deep traps, it takes very long time for detrapping. In order to accelerate the discharge process they applied voltage such that decay rate is increased. They then used the basic trap kinetics equation with Pool-Frenkel emission to extract the spatial distribution. The trap density calculated by ΔV_{FB} was 2.4E19 cm^{-3} and located at oxide-nitride interface with exponential tail distribution up to 10 nm in the nitride.

Chao and White [27] used linear voltage ramp technique and basic trap kinetic equation with amphoteric trap model to extract the trap density and capture cross section. The linear ramp voltage technique separates the charges at the interface, minimizes the back tunneling of trapped charge and simultaneously measures ΔV_{FB} and injected charges ΔQ_n, ΔQ_p. By considering charge centroid, they calculated the trap density and trap capture cross section 2E19 cm^{-3} and 2E-14 cm^2 respectively.

Kim et al., [28] used charge decay model to determine the energy distribution of traps in SONOS devices. This is similar to the retention model proposed by White [25] but considered the effect of internal electric field at the interface between tunnel oxide and nitride. The calculated trap density was 9E18 cm^{-3} in the energy level 1.15 - 1.55 eV from the conduction band edge.

This chapter discussed the evolution and basic principle of SONOS memory device. Issues with SONOS technology, nature of traps in the nitride layer, various trap distributions and charge transfer mechanisms were described. Literature review of process details of silicon nitride by LPCVD and trap characterization methods were presented. The next chapter describes the process optimization of silicon nitride charge trap layer.

Chapter 3

Silicon nitride process optimization and Results

As a charge trapping node for the scaled SONOS memories, the film thickness and composition of silicon nitride layer greatly determines the nonvolatile memory properties. As per ITRS 2009, charge trap layer thickness in current technology is 5 nm and expected to scale down to 4nm by 2014 [1]. Moreover the trapping efficiency depends on the thickness of the charge trap layer and the film composition. Therefore the deposition conditions must be chosen to yield good deposition rate and proper composition.

The target thickness for optimizing the deposition conditions is 20 nm with stoichiometric nitride. The refractive index of stoichiometric nitride is 1.99 - 2.0. We started the process optimization by selecting the process conditions based on literature review. The following are generally used process parameters for LPCVD silicon nitride:

1. Temperature range: 750°C -850°C
2. Pressure: 50 mTorr - 1 Torr (depends on pumping system)
3. Gas flow ratio: depends on the required film composition, thickness and the type of precursor gases.
4. Deposition time: depends on the target thickness

3.1 Experiments

The LPCVD system used for deposition consists of a tubular hot wall reactor with three zone heater control. Figure 3.1 shows the schematic of typical LPCVD system. P-type 100 mm Si

<100>, 5-10 ohm-cm resistivity wafers were used for SiN_x deposition. After the standard RCA cleaning with HF last, the wafers were loaded into the LPCVD system and kept normal to the gas flow.

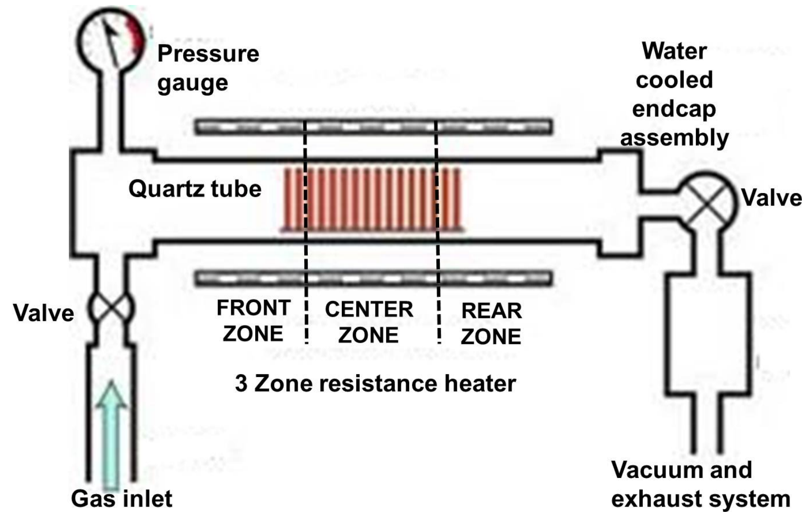

Figure 3.1: Typical three Zone LPCVD reactor.

On each side of the process wafer(s) 10 dummy wafers were placed. The wafer spacing was 0.5 cm unless otherwise specified. Every time one of the process parameters was varied systematically while keeping the other parameters constant to study the effect of process parameters on the film deposition and quality. The deposition process was automated such that all the gas lines and process tube were evacuated and temperature has been stabilized before the actual deposition step. Figure 3.2 shows temperature vs time of a example silicon nitride recipe with other process parameters.

3.1.1 Results of initial experiments

The initial experiments were done with one dummy wafer on each side with temperature fixed at 780°C. To avoid gas depletion along the tube length and maintain the wafer to wafer uniformity along the boat, temperature gradient is maintained inside the process tube. The front zone is operating at 750°C, the center zone at 780°C and the rear zone at 790°C.

As-deposited films with pure silane, ammonia and without nitrogen dilution were very non-uniform in thickness and composition. The deposited film at the wafer periphery was very thick and of different composition when compared with the middle part of the wafer. This thick film along the circumference of the wafer forms a ring with a width Δw. Figure 3.3 shows the thickness non-uniformity of two wafers measured by the ellipsometer.

It is observed that the ring width (measured from wafer periphery) is increasing with the

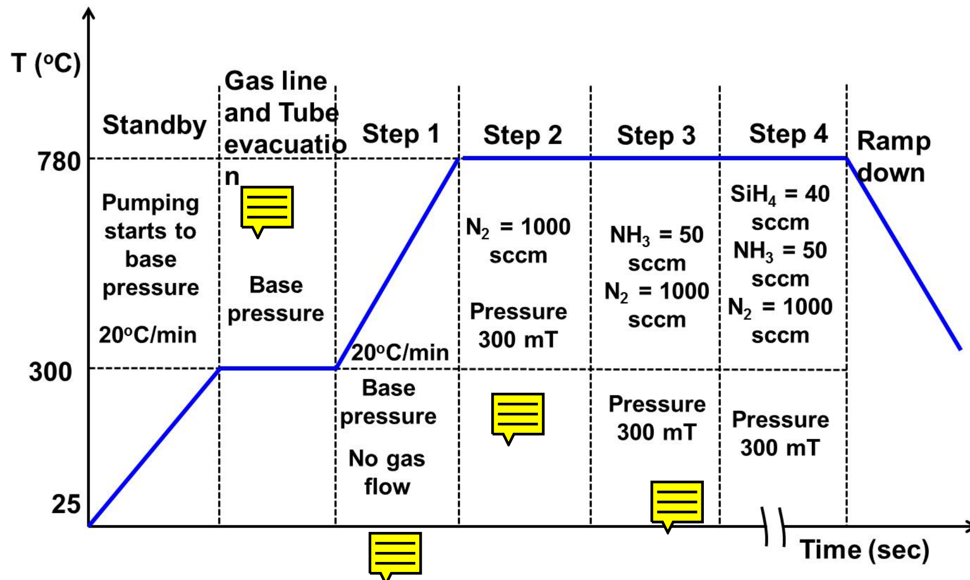

Figure 3.2: Temperature vs time of a example recipe (not to scale). Standby temperature is 300°C, process temperature is 780°C and pressure is 300 mTorr. Base pressure is < 0.005 Torr.

film thickness (Bull's eye effect). This may be due to insufficient supply of intermediate gas phase reacting species to the center part of the wafer. Bull's eye effect is significant in depositing thicker nitride films. Figure 3.4 shows the images of bull's eye and ring formation along the wafer periphery.

The following gas phase reactions are important for the successful deposition [18].

- Diffusion of initial reactants into the gap between wafers ($W_{D,in}$)
- Reaction in the gas phase (W_{hom})
- Diffusion of intermediates to the surface ($W_{D,P}$)
- Heterogeneous reaction generating final product (W_{het})

The non-uniformity is due to different values of $W_{D,P}$ and W_{het} at the wafer edge and at the center part of the wafer. Especially, $W_{D,P}$ is very less than W_{het} in the center part of the wafer. The ellipsometry measurements confirmed the non-uniformity in thickness and composition. The measurement on the ring gives high thickness with silicon rich nitride compared to the middle part of the wafer. The reason for this non-uniformity could be the gas flow inside the process tube is not laminar enough. In order to streamline the gas flow inside the process tube, 10 dummy wafers were kept on both sides of the process wafer. Both the wafer-to-wafer and within wafer thickness variations were reduced by small amount. Figure 3.5 shows the thickness variations on 4-inch wafer for two recipes with dummy wafers used during deposition.

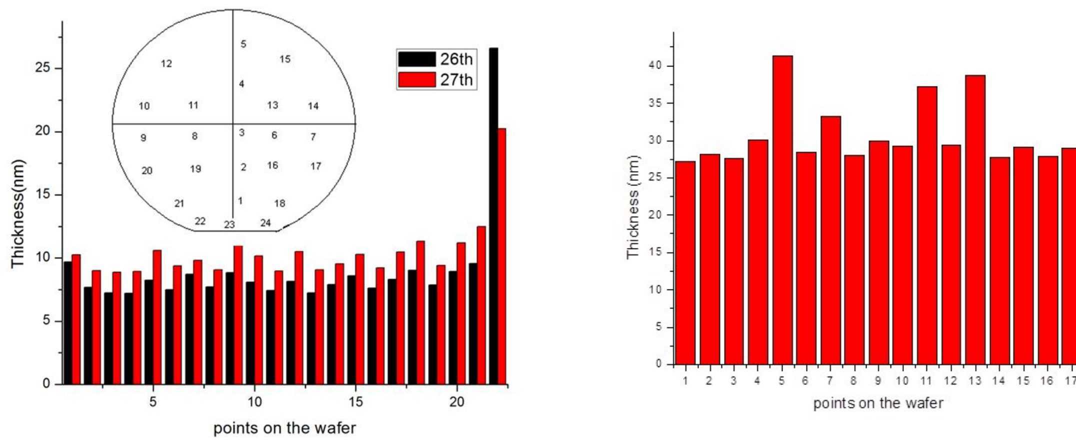

Figure 3.3: Thickness nonuniformity in two cases (a) SiH_4- 15 sccm, NH_3-150 sccm at 780°C and 300 mTorr (for two wafers). (b) SiH_4- 15 sccm, NH_3-225 sccm at 780°C and 300 mTorr (measured in the center part of the wafer map in (a)).

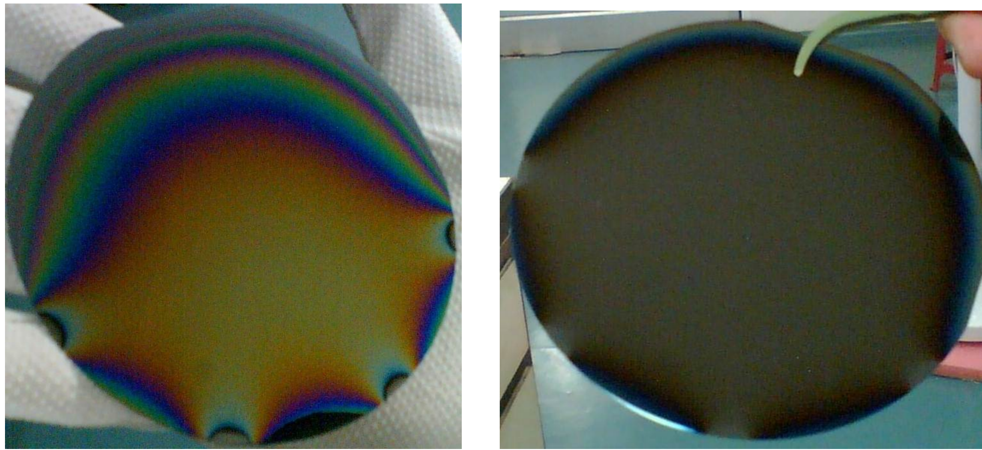

Figure 3.4: (a) Bull's eye effect, SiH_4- 30 sccm, NH_3-75 sccm at 780°C and 300 mTorr. (b) Ring formation along the wafer circumference, SiH_4- 15 sccm, NH_3-45 sccm at 780°C and 300 mTorr.

For a given NH_3 flow rate, the deposition rate increases with the SiH_4 flow rate and also the non-uniformity. It is also observed that deposition rate drops with the increase in NH_3 flow suggesting that NH_3 inhibits the adsorption of Si reacting species onto the substrate. The deposition rate and refractive index with varying NH_3 flow is shown in Fig 3.6.

Following no.of trials, it is observed that the flow ratio (R = SiH_4:NH_3) 40:50 (sccm) giving optimum results and further optimization is carried out by using this gas flow ratio.

3.1.2 Solutions to reduce non-uniformity

The main reason for this non-uniformity in thickness and composition across the wafer could be insufficient diffusion of the intermediate reaction products (that are responsible for the final

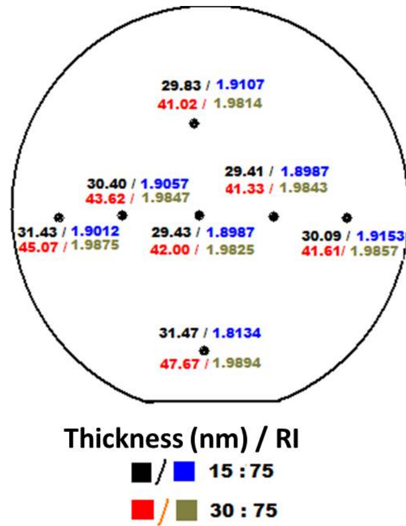

Figure 3.5: Thickness variations after dummy wafers used during deposition for two recipes. SiH_4:NH_3 are 15:75 sccm and 30:75 sccm respectively at 780°C and 300 mTorr.

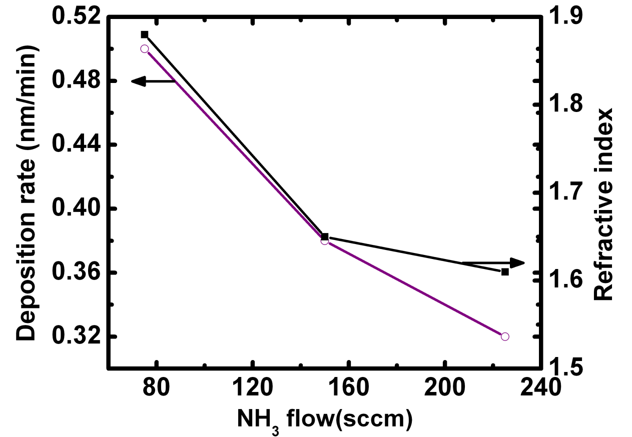

Figure 3.6: Deposition rate and refractive index with change in NH_3 flow. SiH_4 - 15 sccm, Temperature - 780°C and Pressure - 300 mTorr.

film formation) to the center part of the wafer. From section 3.1.1, the non- uniformity can be reduced by making $W_{D,P}$ is equal to W_{het}. The following could be the probable solutions to improve the uniformity.

1. Use diluent gas

   The diluent gas helps in reducing the concentration of the reactive gases and enables the mixing of gases without any spontaneous reaction. The diluent gas also helps in diffusing enough mixture to the surface of the heated substrate. This condition is necessary for film deposition [19].

2. Modification of wafer boat

   The significant radial non-uniformity in both thickness and composition can also reduced by modification of the wafer boat[1]. A specially designed wafer boat having small holes to allow the gases in which the wafers are enclosed can reduce the non- uniformity. However, the design and maintenance of such wafer boats are difficult and expensive.

We implemented the solution 1 by adding extra N_2 line to the existing system. The following sections describe the results of the modified process.

[1] Unpublished work by Rosler *et al.*(1977)

3.2 Effect of N_2 dilution

A significant improvement in uniformity is observed when N_2 is used during deposition. The diluent gas optimizes the turbulence and gas flow conditions for establishing good film uniformity. Therefore, adding N_2 to the reaction chamber reduces desorption of SiH_2 intermediate from the surface of the wafer and hence the deposition rate. Too high flow rate of diluent gas reduces the deposition rate drastically. The SiH_4 to N_2 ratio is the key to obtain optimum deposition rate. Figure 3.7 shows the effect of N_2 dilution on deposition rate and refractive index respectively. The decrease in refractive index is attributed to the the desorption of SiH_2.

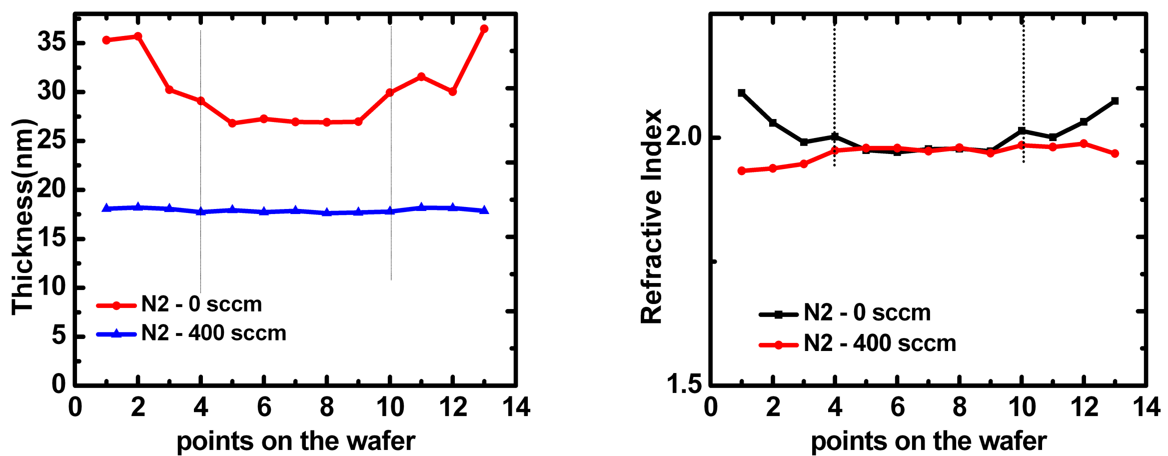

Figure 3.7: Effect of N_2 dilution on (a) thickness (b) refractive index. SiH_4:NH_3 - 40:50 sccm, 780°C, 300 mTorr and 20 min.

Points 1-4 and 11-13 are measured along the wafer periphery. Points 5-10 are measured in the central part of the wafer. This convention is followed in the figures with x-axis named "points on the wafer". This is shown in Fig. 3.8.

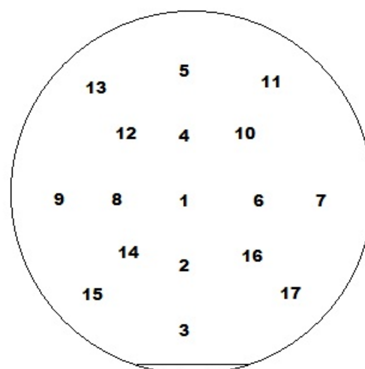

Figure 3.8: Wafer map showing the measurement points on 4-inch wafer by ellipsometer.

With N_2 dilution the uniform area on the wafer increases. This is shown in Fig. 3.9. On a 4-inch wafer, around 95% area with uniform film thickness was obtained when 1000 sccm of N_2 is used in the reaction chamber. In terms of yield, around 13% increment was obtained with dilution for the same deposition time.

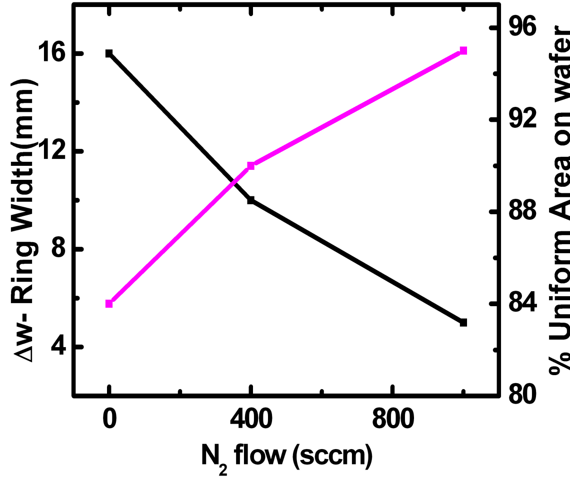

Figure 3.9: Ring width vs N_2 flow. SiH_4:NH_3 - 40:50 sccm, 780°C, 300 mTorr and 20 min.

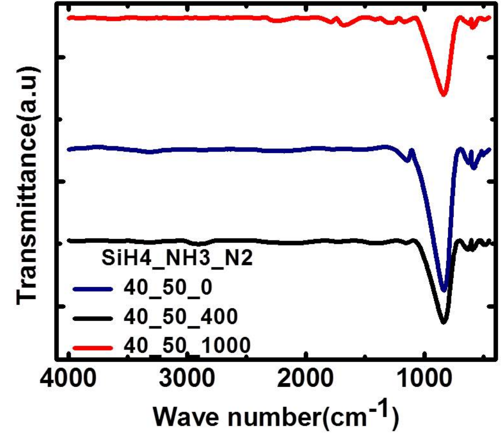

Figure 3.10: FTIR measurements on as-deposited SiN_x films. The legend in the figure indicates gas flow in sccm. Temperature-780°C, pressure-300 mTorr and deposition time-20 min.

In order to study the presence of other impurities such as oxygen and hydrogen, FTIR measurements were done. Figure 3.10 shows FTIR transmission spectra of as-deposited SiN_x films. The main absorption band located at 833 cm^{-1} to 840 cm^{-1}. This belongs to Si-N stretching bond. Some weak bonds were also observed in the vicinity of 580 cm^{-1}. These were identified as Si-N bending bonds. No significant Si-H and N-H bond was observed in the as-deposited SiNx films. With no dilution the main absorption peak was observed at 833 cm^{-1} whereas the same absorption peak was shifted to 837 cm^{-1} when N_2 was used as diluent.

The quality of the as-deposited films was verified by chemical etch test. Figure 3.11 shows the effect of N_2 dilution on etch rate and refractive index.The deposition time is same for these experiments. The etch rate of the SiN_x films in 5:1 HF increases with the N_2 dilution. This suggests that the films with more silicon content were very dense in nature. In order to confirm this, Si-rich SiN_x films were deposited by changing the SiH_4 flow and tested (Figure 3.12).

3.3 Change in SiH_4 flow

For a fixed NH_3 flow of 50 sccm, the thickness increases with SiH_4 flow. For comparison, the deposition time is same for all experiments and also SiH_4:N_2 is 1:10. The ring width Δw along

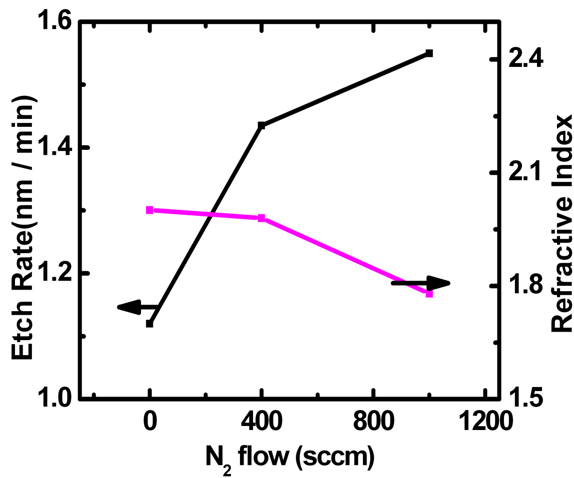

Figure 3.11: Etch rate in 5:1 HF with different N_2 flow. SiH_4:NH_3 - 40:50 sccm, 780°C, 300 mTorr and 20 min.

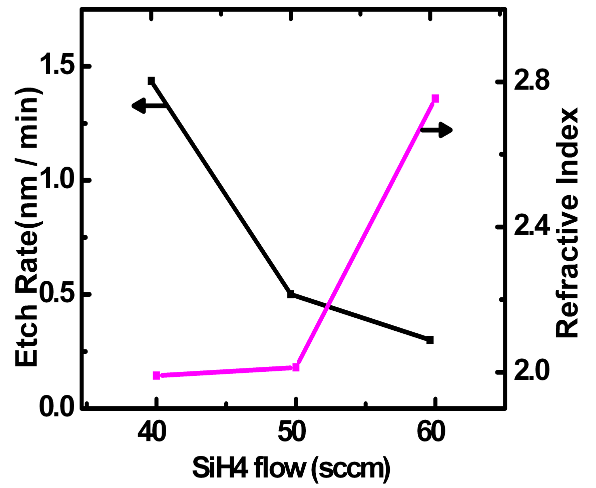

Figure 3.12: Etch rate in 5:1 HF with different SiH_4 flow. SiH_4:N_2 ratio is 1:10, NH_3 - 50 sccm, 780°C, 300 mTorr and 20 min.

the wafer periphery increases with the SiH_4 flow. For the flow ratio (R = SiH_4 flow/ NH_3 flow) 0.8, the film is stoichiometric. With R = 1.2 the film becomes Si rich as expected and also the deposition rate increases. Thickness measurements by ellipsometer are shown in Fig. 3.13. As the films become Si rich the etch rate is reduced (Figure 3.12).

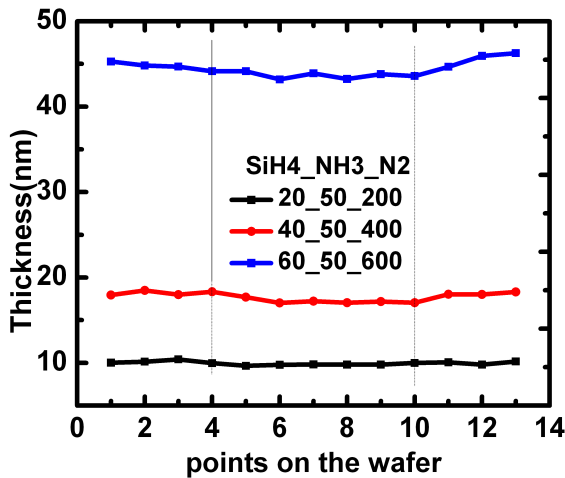

Figure 3.13: Effect of SiH_4 flow on thickness. SiH_4:N_2 ratio is 1:10, NH_3 - 50 sccm, 780°C, 300 mTorr and 20 min.

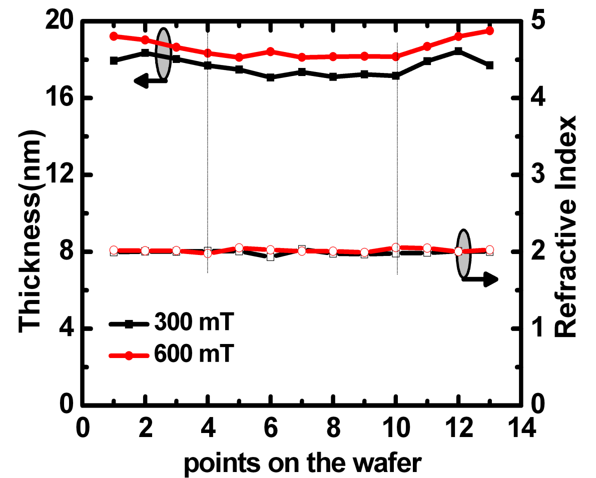

Figure 3.14: Effect of deposition pressure. SiH_4:NH_3:N_2 - 40:50:400 sccm, 780°C, 300 mTorr and 20 min.

3.4 Effect of pressure

The deposition pressure has very little effect on the deposition rate and film composition as shown in Fig. 3.14. The thickness and refractive index increases by very small value with the deposition pressure. But for reduced pressure and gas flow, the deposition rate increases and also the uniformity. This may be due to higher diffusion at lower pressures still maintaining the surface reaction limited deposition.

3.5 Effect of deposition temperature

The effect of temperature on deposition rate and refractive index is shown in Fig. 3.15. For a given gas flow ratio, the deposition rate increases with the temperature. The refractive index measurement shows that Si content in the SiN_x film is increasing with the temperature. The increase in thickness is attributed to increase of surface diffusion of Si reacting species at high temperatures. This is similar to the results of LPCVD SiN_x thin films from DCS and NH_3 reported by Yun et al., [21].

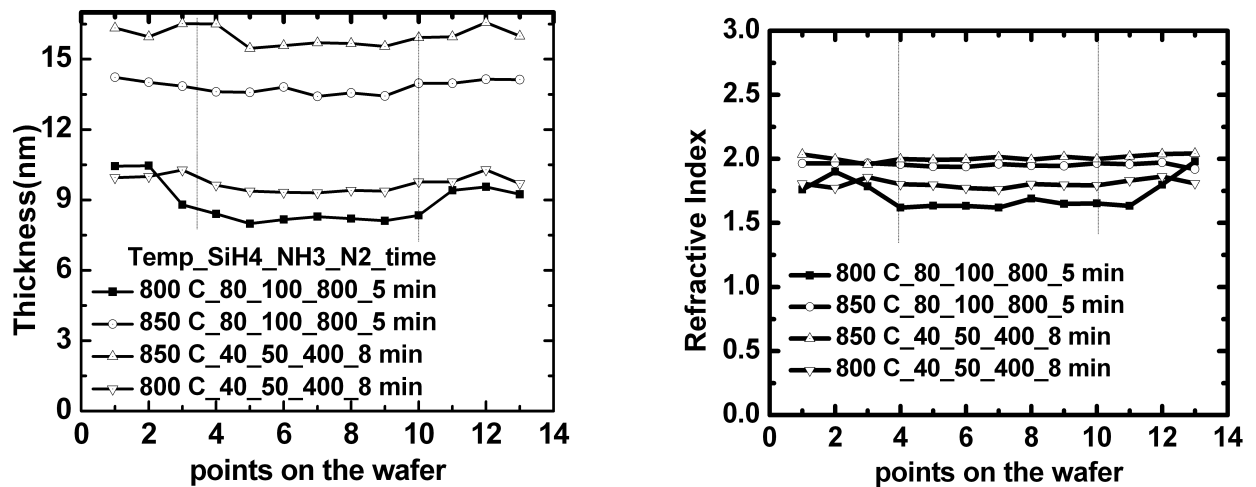

Figure 3.15: Effect of deposition temperature and dilution on (a) thickness (b) refractive index. Numbers in the figure indicates temperature (^{o}C) - SiH_4 flow (sccm) - NH_3 flow (sccm) - N_2flow (sccm) - deposition time (minutes).

The effect of process parameters on the silicon nitride film deposition is summarized as follows:

- The dummy wafers on both sides of the process wafer/s makes the gas flow smoother and laminar.

- Adding N_2 as diluent to process chamber during deposition improves film uniformity by

optimizing the gas flow conditions and diffusing enough gas mixture to the wafer surface.

- NH_3 is limiting the absorption of SiH_2 intermediate gas phase reactant and reducing the deposition rate.
- For a given NH_3 flow, deposition rate and RI increases with SiH_4 flow making the silicon nitride films denser.
- For a given flow ratio of SiH_4:NH_3, deposition rate and film density decreases with increase in N_2 flow.
- Deposition pressure has a very little effect of increasing the thickness and RI for a given gas flow.
- Increase in deposition temperature causes increment in deposition rate and RI by increasing the surface diffusion of gas phase reactants.

Thus proper choice and combination of the process parameters is needed to obtain good quality silicon nitride films.

This chapter discussed the process optimization and effect of different process parameters on the silicon nitride film formation. The results of this chapter were used in the fabrication of SONOS/SNS devices. Next chapter describes the process flow of SONOS flash capacitor fabrication.

Chapter 4

SONOS device fabrication and measurements

This chapter describes the process details and electrical characteristics of individual layers (tunnel oxide, blocking oxide and n-doped poly-Si) of SONOS device using optimized recipes.The process flow of making SONOS device and its characterization is presented. The recipe details are given in Appendix A.

4.1 Development of thin films for SONOS flash memory gate stack

4.1.1 Tunnel oxide

Tunnel oxide is grown on RCA cleaned p-type silicon substrates in horizontal tube furnace in O_2 ambient at various temperatures (750°C-850°C). As-grown films are annealed in N_2 ambient at 900°C. Following a number of trials, process recipes to grow 3.5 to 4.5 nm thick oxide are developed. Detailed process recipe is given in Appendix A (Table A.1). Ellipsometry measurements of as-grown oxide indicates good thickness uniformity. Electrical characteristics of p-Si/SiO_2/N-poly MOS capacitors with the optimum oxide recipe are shown in Fig. 4.1. Thickness measured by ellipsometer is 4.2 nm. The thickness data extracted from C-V measurement is given in the Table. 4.1.

I vs E characteristics and weibull distribution plot are shown in Fig. 4.2. The breakdown field is 15 MV/cm and is par with the values reported elsewhere. The gate is n- doped poly and the device area is 7.85E-5 cm^2.

Device No.	TOX from C-V (nm)	V_{FB} (V)	D_{it} (# $cm^{-2}eV^{-1}$)
1.	3.9	-0.834	4.9E11
2.	4.2	-0.833	4.46E11
3.	4.05	-0.831	4.7E11

Table 4.1: Tunnel oxide parameters extracted from C-V by using hauser fit. Process conditions: T - 800^oC,O_2 - 5000 sccm, time - 10 minutes and annealed at 900^oC for 5 minutes in N_2 ambient.

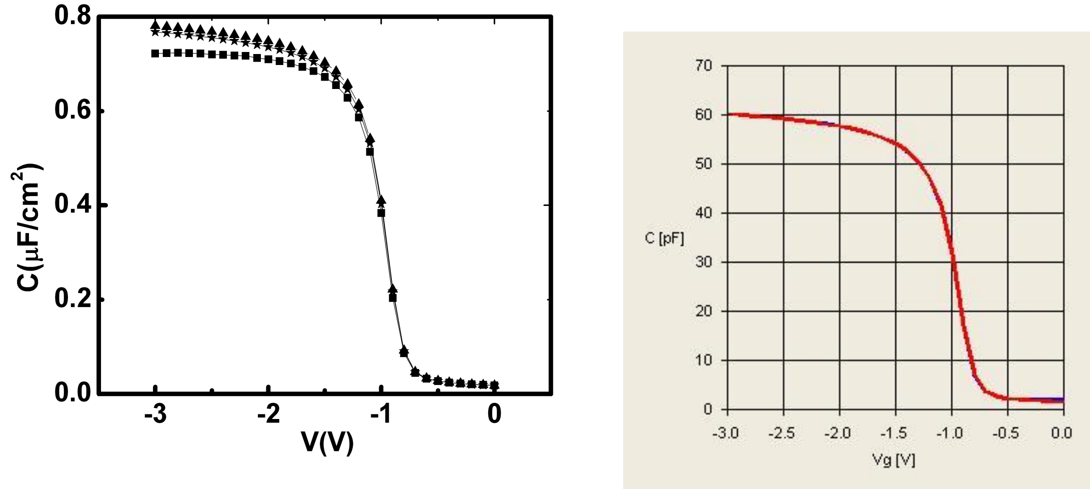

Figure 4.1: (a) C-V characteristics and (b) Hauser fit of the C-V of 4.2 nm tunnel oxide.

4.1.2 Blocking oxide

Blocking dielectric should be free of traps and should not allow electron injection from the gate during erase process. Back injection would reduce the program window of flash memory devices. The blocking oxide is deposited in LPCVD tube at low temperature (430^oC) with SiH_4 and O_2 as precursors and N_2 as diluent. The deposition pressure was 150 mTorr. Thickness measured by ellipsometer is 13.5 nm. Process details are given in Appendix A (Table A.3). Table 4.2 gives the extracted parameters by using hauser program. Thickness obtained from C-V is in good agreement with ellipsometry data. Figure 4.3 shows the C-V characteristics and corresponding hauser fit. The red curve in Fig. 4.3 (b) indicates the C-V plot obtained from measurement and the blue curve is the hauser fit.

Device No.	TOX from C-V (nm)	V_{FB} (V)	D_{it} (# $cm^{-2}eV^{-1}$)
1.	12.72	-1.09	2.65E11
2.	12.7	-1.037	1.56E11

Table 4.2: Blocking oxide parameters extracted from C-V by using hauser fit. Process conditions: SiH_4:O_2 - 40:100 sccm, N_2 - 500 sccm, temperature - 430^oC, pressure - 150 mTorr, Time - 5 minutes.

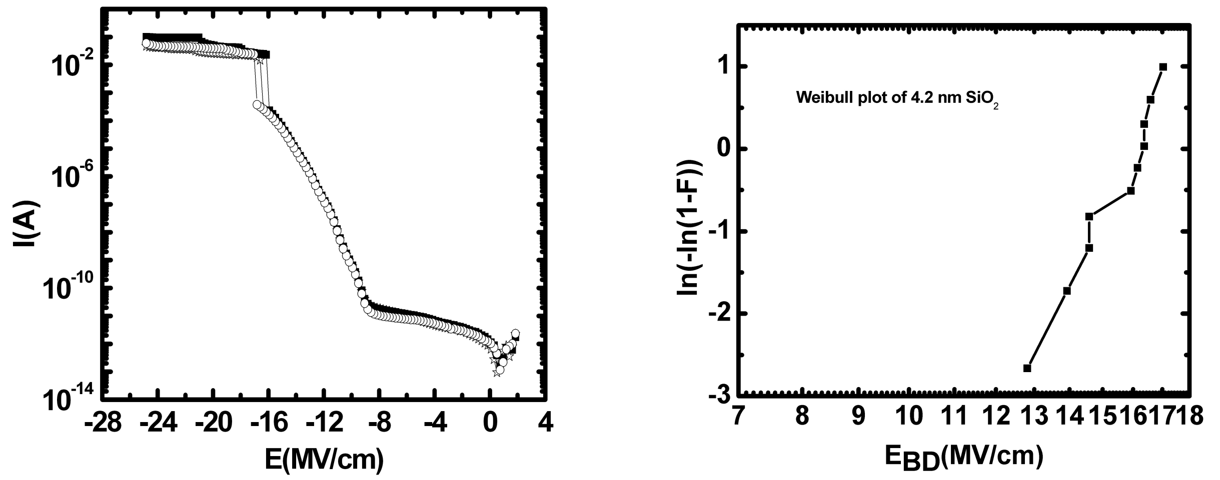

Figure 4.2: (a) I-E characteristics and (b) Weibull distribution plot of 4.2 nm tunnel oxide.

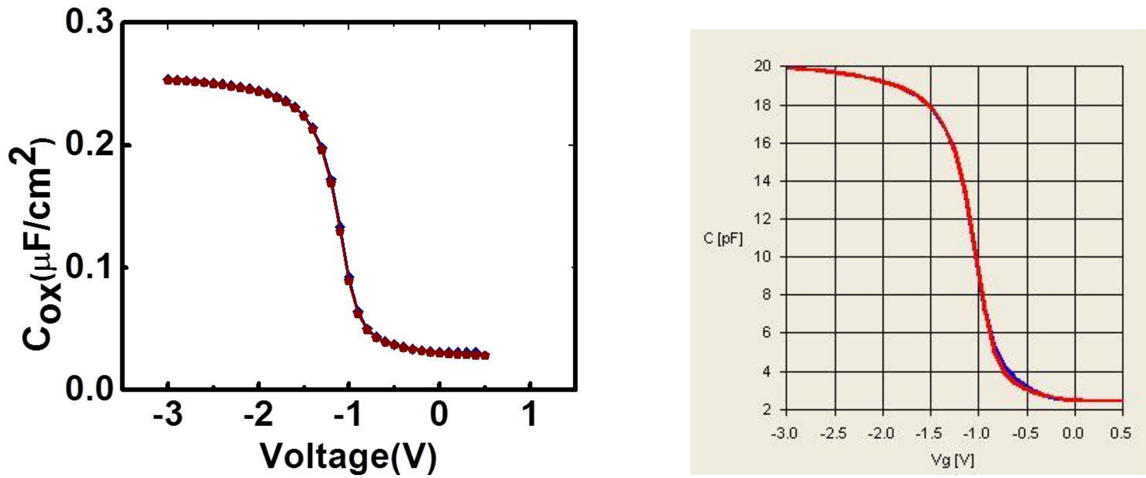

Figure 4.3: (a) C-V characteristics and (b) Hauser fit of the C-V of 13.5 nm LTO after annealing at 850°C in N_2.

A breakdown field of 14 MV/cm is obtained. This is shown in Fig. 4.4(a). The weibull plot in Fig. 4.4 (b) shows that the mean breakdown field is around 14.5 MV/cm. It is found that the leakage has been reduced by annealing in N_2 ambient at 850°C for 5 minutes.

4.1.3 N-poly gate

Polysilicon film with $in-situ$ phosphorous doping is used as the gate material in the SONOS devices. The deposition was done at 630°C and 300 mTorr with SiH_4 and PH_3 as source gases and N_2 as diluent. The deposition kinetics were similar to the silicon nitride. The only difference is the PH_3 in this case. Process details are given in Appendix A (Table A.4). Uniform deposition

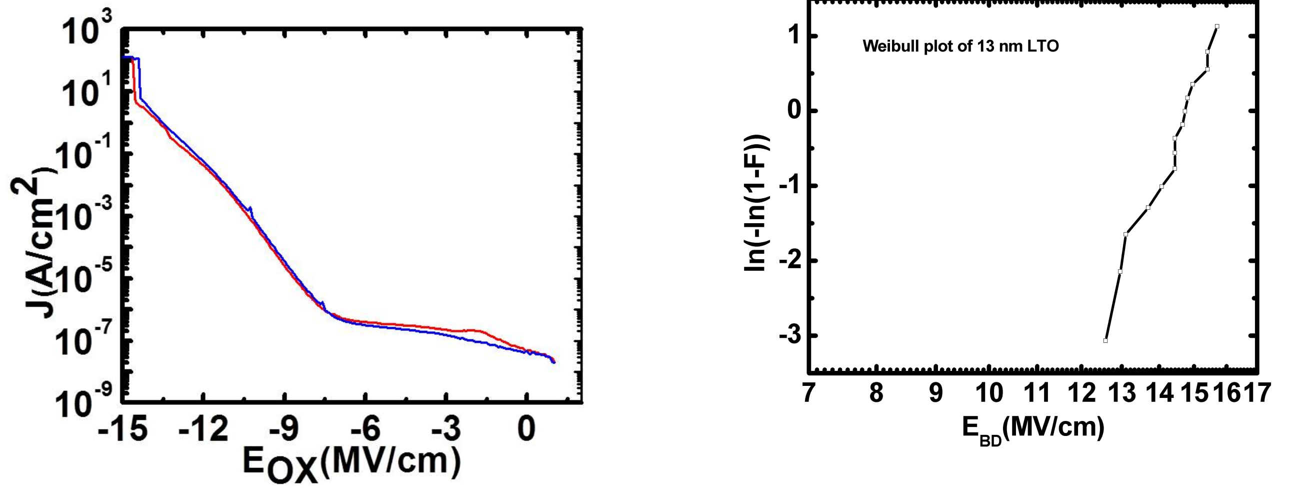

Figure 4.4: (a) J-E characteristics and (b) Weibull distribution plot of 13 nm LTO.

rate of 1.15 nm/min is obtained. Dopant activation is done in Rapid Thermal Process (RTP) chamber at 1000°C for 30 seconds in N_2+O_2 ambient. The sheet resistance measured by four probe system is 65 -120 $\Omega/\Box$ (Figure 4.5).

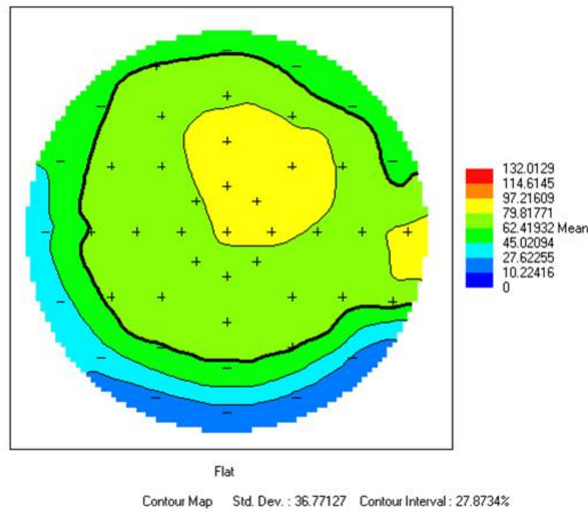

Figure 4.5: Sheet resistance map of n-poly film deposited at 630^{0}C with SiH_4:PH_3:N_2 - 90:20:210 sccm and 300 mTorr for 4500 seconds.

4.2 SONOS structure fabrication and Characterization

The films developed in earlier process steps were integrated in the process flow mentioned below to fabricate SONOS memory capacitors. For detailed process recipes please refer to Appendix A.

4.2.1 SONOS capacitor process flow

1. P-type Si of resistivity 5-10 $\Omega - cm$.
2. Standard RCA clean with HF last
3. Tunnel oxide growth
4. SiNx deposition in LPCVD
5. Blocking oxide deposition
6. Anneal at 850^oC in N_2
7. In-situ doped n-poly deposition
8. RTP annealing at 1000^oC
9. Resist coating and Mask pattern transfer onto wafer
10. Etching of n-poly in RIE
11. Back side n-poly etch with HNA
12. Back side ONO stack etch
13. Back side Al-metal contact
14. Forming gas anneal at 420^oC for 20 minutes.

Capacitors with 100 μm diameter were made and characterized. Two different compositions were used for the charge trap layer. The split conditions of the devices are given in the table 4.3. The differences between S1 and S2 are the gas flow ratio and deposition time during SiN_x

Layer	**S1**	**S2**
Substrate	p type <100>	p type <100>
Tunnel Oxide	3.25 nm	3.25 nm
SiN_x layer, RI	3.5 nm, 1.98	3.5 nm, 1.9
Blocking Oxide	7.5 nm	7.5 nm

Table 4.3: SONOS device details

layer deposition. Precursor gas flow SiH_4 : NH_3 for S1 and S2 are 40:50 sccm and 20:50 sccm respectively. Deposition times are 5 minutes and 8 minutes for S1 and S2 respectively.

4.2.2 SONOS memory stack characterization

Program/Erase transients were measured on the as-fabricated SONOS devices by using Keithley semiconductor analyzer. Breakdown voltages of S1 and S2 are 16 V and 17 V respectively. Programming is done at 10 V and 12 V for varying program pulse widths from 1 μsec to 10 sec. During programming light is ON to generate the carriers in the substrate. A program window of 2.8 V is observed at 14 V for the devices as shown in Fig. 4.6.

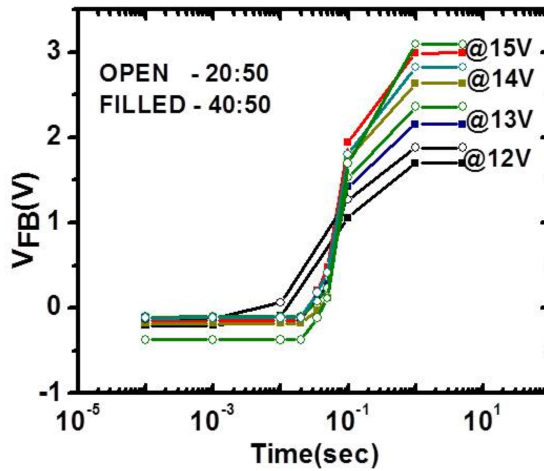

Figure 4.6: Programming transients.

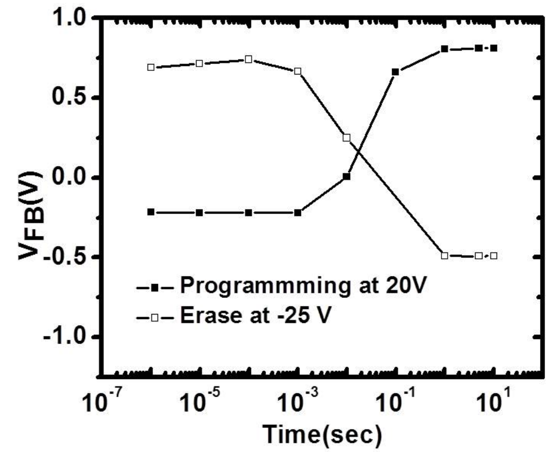

Figure 4.7: P/E transients with improved blocking oxide.

In order to erase the programmed device, negative pulse with 12 V with different widths is applied to the gate. But no erasing is observed. This may be due to weak blocking action from the top dielectric. To improve the quality of the blocking oxide, wet oxidation is done at 900°C on top of the 3 nm deposited Low Temperature Oxide. During this process a small fraction of SiN_x is converted into SiO_2 and also any remaining Si species during LTO deposition gets oxidized. Figure. 4.7 shows the P/E transients after the improved blocking oxide. The data shown in the Fig. 4.7 belongs to relatively thick ONO stack. Tunnel oxide thickness is 5 nm, silicon nitride thickness is 4.5 nm with refractive index 1.98 and the blocking oxide is 9 nm.

The programming and erasing was very slow and very high voltages were used (+20 V/-25 V). Erase saturation is observed and this may be due to relatively thick tunnel oxide in which the hole tunneling is small. Also, because of high voltages, FN tunneling of electrons from n-poly gate may be another reason for erase saturation.

Our idea in this part of work is to make a working SONOS device and characterize it. But the SONOS device has many interfaces compared to the simple MNS device. The main objective of this work is to analyze the traps in Silicon nitride film. Because of many interfaces (Si - SiO_2,

SiO_2 - SiN_x, SiN_x - SiO_2 and SiO_2 - n-poly), it is very difficult to separate the trap contribution from each. Therefore, to extract the trap parameters in the Silicon nitride film, SNS devices are made and characterized. The next chapter describes the trap characterization methods used for SNS capacitors. Chapter 6 describes the trap characterization experiments and trap parameter extraction.

Chapter 5

Trap characterization methods of MNS devices

This chapter describes the techniques used to extract the trap parameters (trap density, trap capture cross section and energetical distribution of the traps).

5.1 Trap density

The flat band voltage shift i.e, the change in V_{FB} (ΔV_{FB}) for a fresh device to that of trap filled device is used for determining the trap density. SNS structures are used to evaluate the trap density (Figure 5.1). Trap filling is done by applying voltage of proper polarity and magnitude for certain amount of time. As the traps started filling the V_{FB} changes and after certain time the V_{FB} value saturates. This means almost all the traps have been filled.

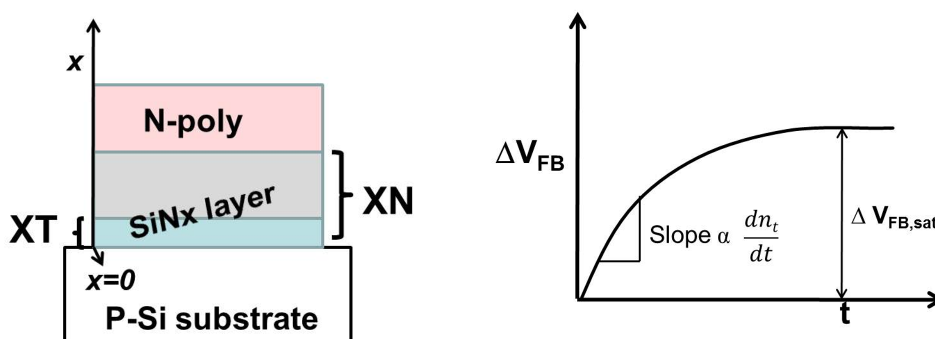

Figure 5.1: (a) Schematic of SNS structure for trap density calculation (b) V_{FB} vs time during charge trapping.

The shift in flat band voltage can be expressed as

$$\Delta V_{FB} = -\frac{1}{C_N}\int_0^{X_T} \frac{x}{X_N}\rho(x)\ dx \tag{5.1}$$

where,

ρ is the spatial charge density (C/cm^3)

X_N is the nitride thickness (nm)

X_T is the thickness of the nitride up to which charges have been trapped (nm)

C_N is the capacitance of the structure (F/cm^2)

It is assumed that the traps are distributed uniformly in the charge trapped region (X_T). Therefore,

$$\Delta V_{FB} = -\frac{\rho.{X_T}^2}{2\epsilon_N} \tag{5.2}$$

But, $\rho = q.N_t$. Therefore,

$$N_t = \frac{2\epsilon_N \Delta V_{FB}}{q{X_T}^2} \tag{5.3}$$

5.2 Trap capture cross section

The transient behaviour of the traps during charge trapping is governed by the basic rate equation [22]. The number of trapped charges at any point in the nitride is proportional to the available carriers for trapping (n_c) and number of free traps ($N_t - n_t$).

$$\frac{dn_t}{dt} \propto n_c(N_t - n_t) \tag{5.4}$$

$$\frac{dn_t}{dt} = \sigma_t \vartheta_d n_c(N_t - n_t) \tag{5.5}$$

where,

N_t is the total number of traps ($\#/cm^3$)

n_t is the number of filled traps ($\#/cm^3$)

σ_t is the trap capture cross section (cm^2)

ϑ_d is the drift velocity of the carriers in the nitride (cm/s)

Once the charges are trapped, there exists an internal electric field in the nitride. However, the trapped carriers gradually escapes from the traps due to any of the tunneling mechanisms or emission process. Two important emission mechanisms are Thermal emission and Pool-Frenkel emission. After taking these charge loss mechanisms into account, Eq. (5.5) is modified as

$$\frac{dn_t}{dt} = \sigma_t \vartheta_d n_c (N_t - n_t) - n_t \xi \tag{5.6}$$

where, ξ is the total emission rate.

But the emission rate is proportional to the number of filled traps n_t which are very small initially. Therefore, emission rate is very small and neglected during initial conditions. While calculating the trap capture cross section we are considering the initial rate of change of filled traps, this assumption is valid and Eq. (5.5) can be used.

The other unknown in Eq. (5.5) is n_c, the number of carriers available for trapping. This can be determined by current measurements. Initially, prior to charge injection, the insulator is neutral. The net charge in the nitride is $q(n_c + n_t)$. But n_t is zero initially, the net charge in the nitride becomes $q.n_c$. Therefore, during initial conditions the current through insulator is decided by n_c and the current density J is given by

$$J \cong q n_c \vartheta_d \tag{5.7}$$

Now, Eq. 5.5 becomes

$$\frac{dn_t}{dt} = \sigma_t \frac{J}{q} (N_t - n_t) \tag{5.8}$$

But, initially the number of filled traps n_c is very small compared to N_t. Therefore, Eq. (5.8) is modified as

$$\frac{dn_t}{dt} = \sigma_t \frac{J}{q} N_t \tag{5.9}$$

The initial slope of ΔV_{FB} vs t of Fig. 5.1(b) gives the information about $\frac{dn_t}{dt}$, the rate of change of filled traps. In Eq. 5.2, the time varying quantities are ΔV_{FB} and ρ (N_t is determined by the saturation value of ΔV_{FB}). Therefore,

$$\Delta V_{FB}(t) = -\frac{\rho(t).{X_T}^2}{2\epsilon_N} \tag{5.10}$$

Then

$$\frac{dn_t}{dt} = \frac{1}{q}\frac{d\rho}{dt} \tag{5.11}$$

But, from Eq. 5.10

$$\frac{d\rho}{dt} = \frac{2\epsilon_N}{{X_T}^2}\frac{d\Delta V_{FB}}{dt} \tag{5.12}$$

Upon substitution of Eqs. 5.11 and 5.12 in Eq. 5.8, the trap capture cross section can be calculated.

$$\sigma_t = \frac{1}{J.N_t}\frac{2\epsilon_N}{{X_T}^2}\frac{d\Delta V_{FB}}{dt} \tag{5.13}$$

Where, $\frac{d\Delta V_{FB}}{dt}$ is the initial slope of the Fig. 5.1(b), the program transients curve.

5.3 Retention Model for trap level extraction

The energy levels of traps are determined by retention measurements. By the application of energy, electric field or temperature, the trapped charge can be escaped. The energy needed to detrap a trapped charge depends on the energy level of the trap. Therefore by observing the detrapping rate of the captured charge (charge decay profile), energy level of the traps can be determined.

The retention model developed by Yang and White [23] considered the amphoteric nature of the traps in Si_3N_4. They have used SONOS devices to identify the charge loss mechanisms and the charge decay model at 175°C. The various charge loss mechanisms is shown in Fig. 5.2. The charge loss *i.e.*, threshold voltage decay can be due to:

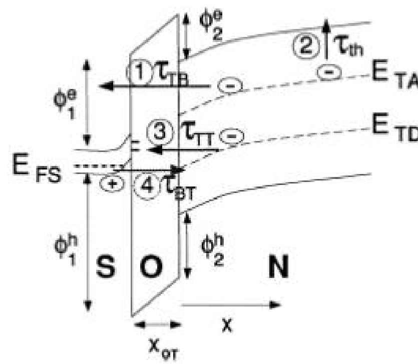

Figure 5.2: Charge loss mechanisms in SONOS device [23].

1. Trap to band tunneling (τ_{T-B})

2. Thermal excitation from the trap to nitride conduction band (τ_{th})

3. Trap to trap tunneling from nitride to $Si - SiO_2$ interface (τ_{T-T})

4. Hole tunneling from Si valance band to nitride traps (τ_{B-T})

The terms in the bracket are the time constants associated with respective charge loss mechanism shown in Fig. 5.2. They used the threshold voltage decay *vs* time (log scale) to extract the trap energy distribution. The SONOS device was programmed first and then observed the decay in threshold voltage at 175°C from time t_i (initial time) to t_f (final time). During retention measurements the internal electric field is small such that there is negligible amount of hole tunneling from Si valance band to traps and from nitride traps to $Si - Si_3N_4$ interface traps. The two processes trap to band tunneling and thermal excitation from nitride traps are dominant.

Based on the analysis [23], the threshold voltage decay in the case of MNS/SNS capacitor is given by:

$$\frac{\partial \Delta V_{TH}(t)}{\partial \log(t)} \cong -2.3qk_BT\frac{X_N^2}{2\epsilon_N}\,g(E_{TA}) \tag{5.14}$$

where,

$\Delta V_{TH}(t)$ - Change in threshold voltage with time

k_B - Boltzmann constant

T - Temperature in K during measurement

X_N - Nitride thickness

ϵ_N - Silicon nitride dielectric constant

$g(E_{TA})$ - Density function of traps in nitride ($\#/cm^3eV$)

E_{TA} - Trap energy level (eV)

The trap energy level E_{TA} depends on the start and end time of the measurement and is given by:

$$E_{TA} = k_BT\ln(AT^2t) \tag{5.15}$$

The constant A in Eq. 5.15 is temperature independent and is given by [23]:

$$A = 2\sigma_t\sqrt{\frac{3k_B}{m_{en}^*}}\left(\frac{2\pi k_B m_{en}^*}{h^2}\right)^{\frac{3}{2}} \tag{5.16}$$

Thus by using Eqs. 5.14, 5.15 and 5.16 the trap distribution can be determined.

In this chapter the methods for extracting the trap density, trap capture cross section and energetical trap distribution are discussed. These methods will be used in the next chapter to evaluate the trap parameters. Chapter 6 describes the relevant measurements to extract the trap parameters, extraction of the trap parameters and discusses the results.

Chapter 6

Extraction of trap parameters and Results

This chapter describes the characterization of devices such as C-V, I-V, program/erase and retention measurements. The data from these measurements is used in extraction of the trap parameters.

6.1 Electrical Characterization Setup

The measurement setup for characterizing as-fabricated SNS devices is shown in Fig. 6.1. The setup consists of low noise probe station with thermo chuck, Keithly 4200 semiconductor parameter analyzer and Agilent 4284 LCR meter. These are connected through Keithley 708A switch matrix. P/E measurements are done by connecting the gate of SNS device to the PGU/SMU of Keithley 4200. The C-V measurements are done by using Agilent 4284 LCR meter. During P/E measurements, switching between LCR and SMU is done by GPIB interface that controls the switching of all the units in the setup. P/E and read operations requires calculation of flat-band voltage (V_{FB}) from C-V measurement. For retention measurement, SMUs are kept connected and switched to LCR meter at certain time instants to measure the V_{FB}. The temperature of the stage is controlled by Thermax temperature controller during retention measurement at elevated temperatures. The I-V measurement is done by SMUs of the Keithley 4200.

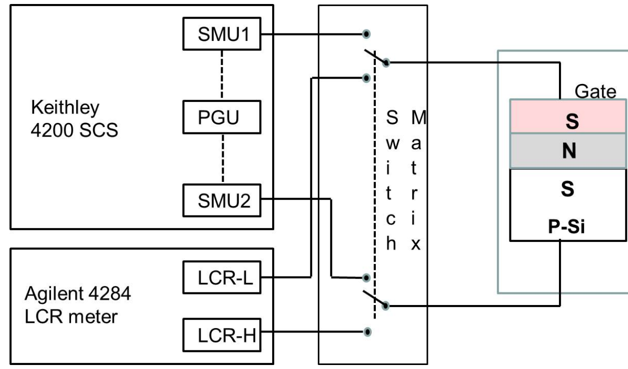

Figure 6.1: Characterization setup for for C-V, I-V, P/E and retention measurements.

6.2 Results

6.2.1 SiN_x film thickness and RI data

Thickness and refractive index of as-deposited films on Si wafer are measured by using spectroscopic ellipsometer. Table 6.1 shows the process conditions and ellipsometry data of as-deposited SiN_x films.

Sample description	N_2 Flow (sccm)	Deposition time (min)	Thickness (nm)	RI
D1	0	20	22.2	2.001
D2	400	25	22.5	1.999
D3	1000	40	25.5	1.950

Table 6.1: Thickness and RI data of the SiN_x films deposited at 780^oC, 300 mTorr and with different N_2 flow. SiH_4-40 sccm and NH_3-50 sccm

From the ellipsometry data it is clear that SiN_x films deposited with no N_2 or 400 sccm N_2 are stoichiometric[2]. The thickness uniformity is improved with N_2 dilution but film composition is changed. The reduction in refractive index with N_2 dilution is the indication of decrease in Si component in the film.

6.2.2 Electrical characterization

The C-V characteristics of as-fabricated SNS devices are shown in Fig. 6.2 (a). The difference in C_{max} is due to the difference in the SiN_x film thickness. The C-V characteristics of samples D1 and D2 are almost same except C_{max}. This suggests the SiN_x film composition is not varied drastically with reasonable amount of N_2 dilution. The refractive index from the ellipsometer also confirms the same. The C-V of sample D3 shifted towards left indicates some positive

[2] The refractive index (RI) for stoichiometric silicon nitride is $1.99 \sim 2.0$.

trapped charge compared to D1 and D2. The EOT of samples D1,D2 and D3 extracted from C-V measurement are 12.5 nm, 13.5 nm and 14.8 nm respectively. By using the thickness values from Table 6.1 as the physical thickness of nitride and EOT from C-V, dielectric constants were calculated and the values were 7.05 (for D1), 6.73 (for D2) and 6.58 (for D3).

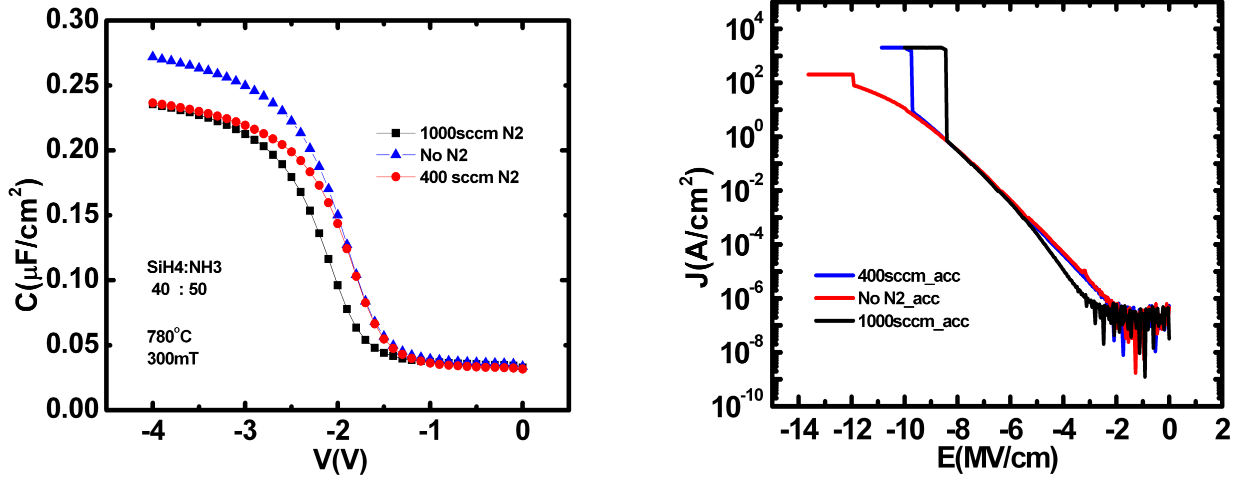

Figure 6.2: (a) C-V characteristics and (b) J-E characteristics of SNS devices with different N_2 flow. The device area is 5.02E-5 cm^2.

The breakdown voltage is determined from the Current - Voltage (I-V) measurements. This measurement is necessary (before the P/E measurements) to decide the maximum programming/erase voltage. The I-V measurement data is converted into Current density (J) - Electric field (E) and is shown in Fig. 6.2 (b). Although the breakdown fields are different (due to variation in film thickness) the current levels of samples D1 and D2 are same. This again confirms RI measurement from ellipsometer that the film composition is not changed drastically with 400 sccm of N_2 during deposition. The leakage in insulators is mainly due to traps present in the insulator. At low fields ($<$ 6 MV/cm), the leakage current of D3 is small compared to D1 and D2. This is some kind of indication of traps present in the SiN_x film.

6.3 Trap Characterization Measurements

Trap parameters are determined by P/E transients,C-V measurement and retention measurements. The program transients along with C-V data used to determine the trap density (N_t). This method is based on the assumption of spatially uniformly distributed traps. Data from retention measurement can also be used to extract the trap density. This technique gives the distribution of traps in energy.

6.3.1 Trap Density from Programming Transients

In this method, a ±14 V pulse with varying widths from 1 μsec to 1 sec is applied to the gate of SNS device. Between each pulse, C-V is taken to measure V_{FB}. When +14 V is applied, Si surface is in inversion and light is shining on the device to generate carriers. Electrons are now captured into the traps by FN tunneling ($|E_N| > \frac{\phi_N}{X_N}$). When -14 V is applied to the gate, Si surface is in accumulation and hole tunneling from the valance band occurs and the V_{FB} shifted in other direction. Measured program transients are shown in Fig. 6.3. Calculated ΔV_{FB} for both polarities is given in the Table. 6.2.

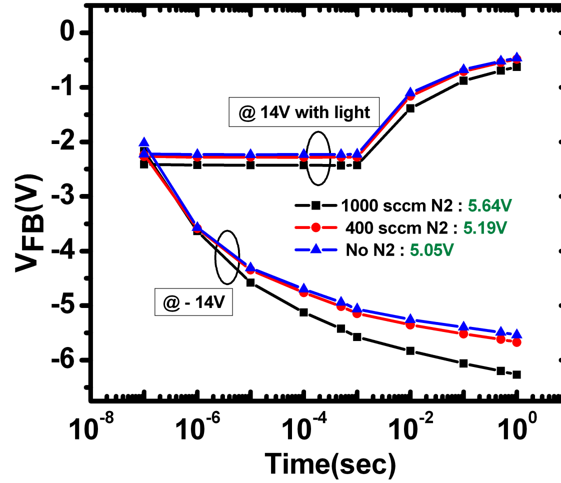

Figure 6.3: Program transients of devices D1,D2 and D3.

Sample Description	ΔV_{FB} for +14 V	ΔV_{FB} for -14 V
D1 (No N_2)	1.76 V	3.51 V
D2 (400 sccm)	1.78 V	3.45 V
D3 (1000 sccm)	1.81 V	4.09 V

Table 6.2: ΔV_{FB} values of samples D1,D2 and D3 with ±14 V programming

It is observed that the hole trapping is faster than the electron trapping. When programmed with +14 V (light ON), initially there is hardly any electron trapping. There is no significant shift in V_{FB} due to electron trapping till 1 ms. This may be due to the smaller stressing time (programming pulse width) than the minority carrier generation time.

During programming with -14 V, the substrate is p-type and Si surface is in accumulation. Therefore lot of holes are available for tunneling even for smaller programming pulse widths.

From section 5.1, the trap density can be calculated by using 5.2. ΔV_{FB} can be obtained from program transients, ϵ_N from C-V data and X_T by assuming uniform distribution of traps.

Actually X_T is the charge centroid up to which the traps have been filled and is assumed to be $X_T = \frac{X_N}{2}$. Table 6.3 shows the calculated electron and hole trap density values of the three samples.

Sample Description	**Thickness (nm)**	**EOT (nm)**	ϵ_N	**Electron Trap Density** N_{Te} $(\#/cm^3)$	**Hole Trap Density** N_{Th} $(\#/cm^3)$
D1 (No N_2)	22.2	12.5	7.05	2.7E18	5.4E18
D2 (400 sccm)	22.5	13.5	6.73	2.61E18	5.05E18
D3 (1000 sccm)	25.5	14.8	6.58	2.02E18	4.57E18

Table 6.3: Calculated trap density values from program transients

6.3.1.1 C-V characteristics with program and erase

This section describes the C-V measurements on the devices D1, D2 and D3 with programming and erasing. The following notation is used to describe the C-V characteristics. These are the sequence of operations performed on the SNS device.

- **Program** : Application of +14 V pulse with varying widths from 1 μs to 1 s and light ON. Between each pulse V_{FB} is calculated.
- **Erase** : Application of -14 V pulse with varying widths from 1 μs to 1 s. Between each pulse V_{FB} is calculated.
- **F-P-E** : C-V on fresh device - C-V on the same device after programming - C-V on the same device after erasing.
- **F-E-P** : C-V on fresh device - C-V on the same device after erasing - C-V on the same device after programming.
- **F-P-E-P** : C-V on fresh device - C-V on the same device after programming - C-V on the same device after erasing - C-V on the same device after programming.
- **F-E-P-E** : C-V on fresh device - C-V on the same device after erasing - C-V on the same device after programming C-V on the same device after erasing.

Device D1

Fig. 6.4 (a) shows the C-V characteristics of D1 with F-E-P. The C-V is shifted to left by 3.3 V during erasing. This suggests hole trapping in the nitride. After programming the C-V is

shifted to right by 0.9 V from that of fresh device. The memory window is 4.2 V. Now F-P-E-P measurements are performed on another device and corresponding C-Vs are shown in Fig. 6.4 (b). In this case the memory window is 3.7 V.

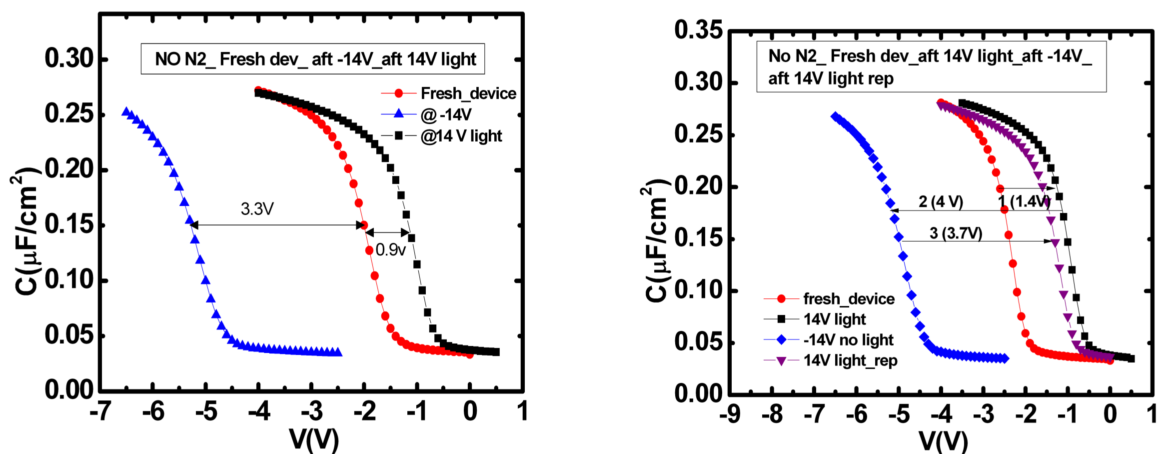

Figure 6.4: C-V characteristics of D1 (a) Fresh device - after stressing with -14 V - after stressing with +14 V and light ON. (b) Fresh device - after stressing with +14 V and light ON - after stressing with -14 V -after stressing with +14 V and light ON.

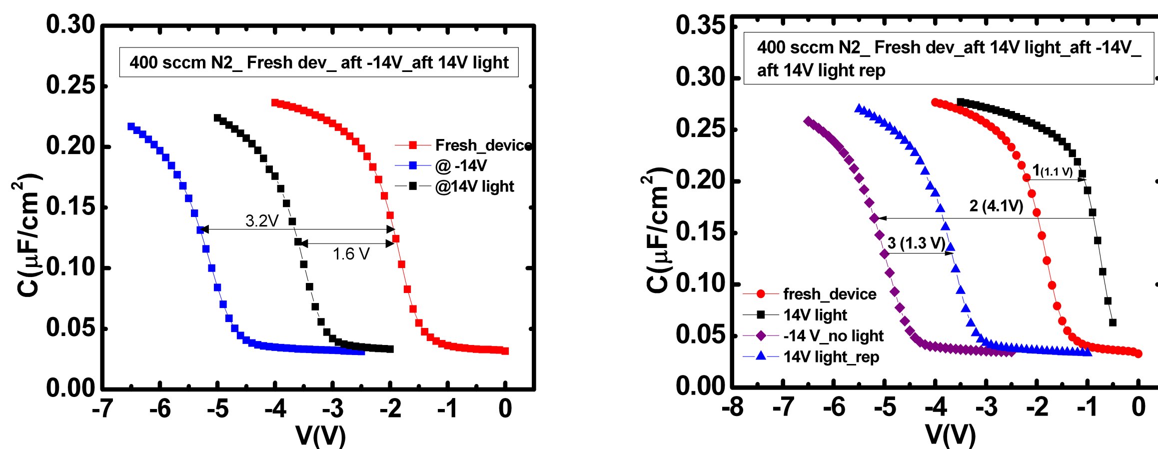

Figure 6.5: C-V characteristics of D2 (a) Fresh device - after stressing with -14 V - after stressing with +14 V and light ON. (b) Fresh device - after stressing with +14 V and light ON - after stressing with -14 V -after stressing with +14 V and light ON.

Device D2

The same set of measurements of D1 described above are repeated on D2. After erasing the ΔV_{FB} is -3.2 V. Then C-V is taken after programming. Now the V_{FB} is shifted to right by

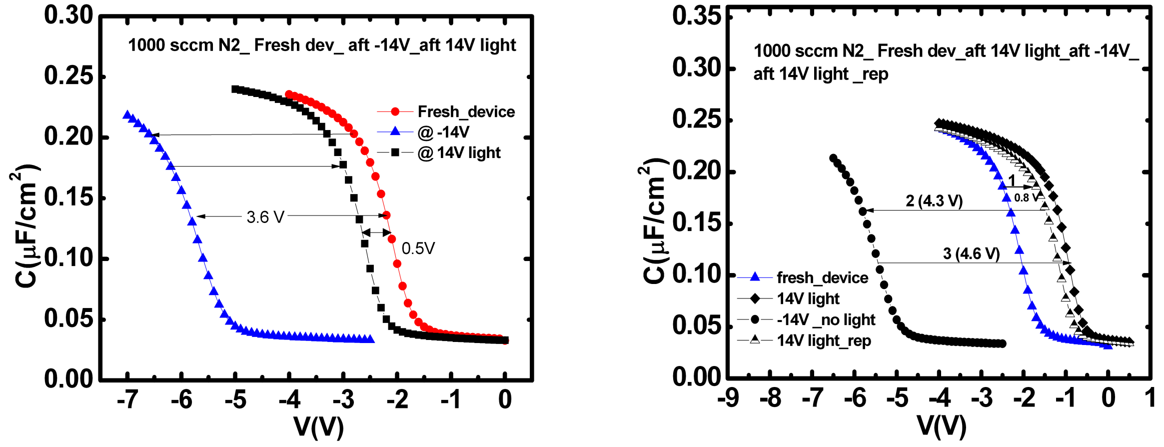

Figure 6.6: C-V characteristics of D3 (a) Fresh device - after stressing with -14 V - after stressing with +14 V and light ON. (b) Fresh device - after stressing with +14 V and light ON - after stressing with -14 V -after stressing with +14 V and light ON.

1.6 V (Figure 6.5 (a)). This suggests that all the trapped holes are not neutralized. F-P-E-P measurements on another device also shows similar result. This is shown in Fig. 6.5 (b).

Device D3

The F-E-P measurements on D3 shows that similar trend as D2. All the trapped positive charge is not completely neutralized. The memory window is 3.1 V. Whereas the F-P-E-P measurements on another D3 device shows complete erase of trapped positive charge and the memory window is 4.3 V. This is shown in Fig. 6.6.

6.3.2 Trap capture cross section

By using Eq. 5.12, the trap capture cross section can be determined. $\frac{d\Delta V_{FB}}{dt}$ is the initial slope of the program transient curves from section 6.4.1. The other unknown in Eq. 5.12 is the current density J and is obtained as follows. The SNS device is stressed by applying ± 14 V and the current is measured by switching the SMU to sampling mode. The current is measured at 1 μs intervals for 1000 samples. When +14 V is applied, the sample is illuminated. The measured current values are divided by the device area to get J. Table 6.4 gives the values of current density in accumulation and inversion.

Trap density, N_t, is determined in the previous section. The calculated values of σ_t by Eq. 5.12 are given in the Table 6.5.

Sample description	J -14 V (A/cm^2)	J at +14 V with light (A/cm^2)
D1	9E-3	1.69E-3
D2	4.98E-3	1.29E-3
D3	1.03E-3	6.37E-4

Table 6.4: Current density values of D1,D2 and D3 in accumulation and inversion at 14 V

Sample description	Hole trap capture cross section (cm^2)	Electron trap capture cross section (cm^2
D1	1.7E-11	1.34e-14
D2	2.6E-11	1.76E-14
D3	1.24E-10	3.2E-14

Table 6.5: Hole and Electron trap capture cross sections

6.3.3 Retention measurements for extracting the trap energy level

Retention measurements are done at 150°C. Trap filling (electron traps) is done by applying +15 V with light ON. The change in V$_{FB}$ is noted and the sample is heated to 150°C. Now C-V is taken at different time instants starting from 100 μsec to 10^5 seconds to obtain V$_{FB}$. Because of the heat energy, the electron detrapping is fast. This charge decay data i.e., V$_{FB}$ decay with time is analyzed with the model developed in section 5.2 to extract the energy distribution of the traps. Since ΔV_{FB} *vs* time data is used, the trap density can also be extracted. Fig. 6.7 shows the V$_{FB}$ *vs* time data from retention measurements at 150°C.

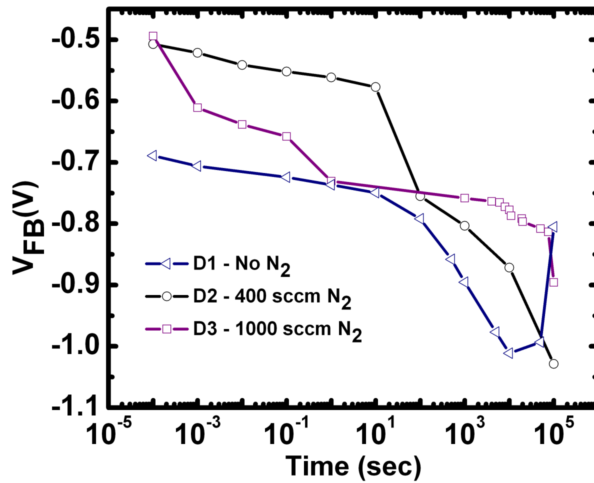

Figure 6.7: Retention measurement on devices D1,D2 and D3 showing V_{FB} vs time. The devices are programmed for 1 s.

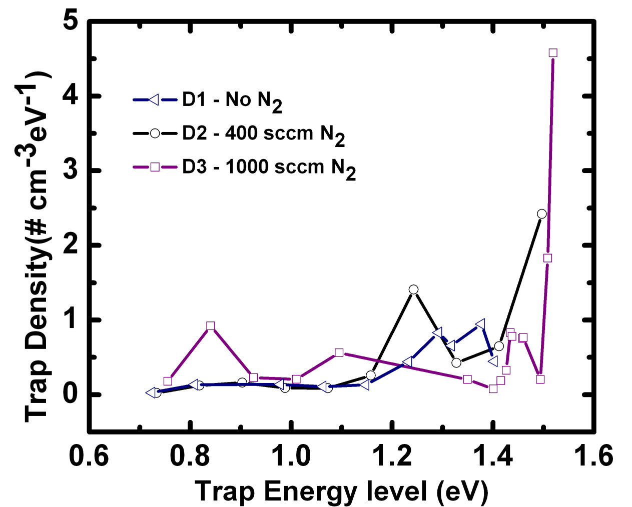

Figure 6.8: Electron trap distribution in devices D1,D2 and D3.

From the retention measurement data, $\frac{\partial \Delta V_{TH}(t)}{\partial \log(t)}$ is obtained for each t. Then E_{TA} by using

5.15 is obtained for corresponding t calculated earlier. By using Eq. 5.14 the density of trap states function is evaluated. The time range for which retention measurement is done decides the trap energy level to be observed. The temperature independent constant A in Eq. 5.15 is calculated for different σ_t (corresponds to different samples). Therefore, the time rage 100 μsec to 10^5 s corresponds to 0.74 - 1.55 eV (approximately) at 150°C. The trap density extracted in this way is comparable to the value obtained from program transient method. Figure. 6.8 shows the energetical distribution of traps.

6.4 Discussion

By observing all the results described above, it is clear that dilution of SiH_4 with N_2 affects the trap properties. Ellipsometry measurements indicate that the refractive index is decreased from 1.999 to 1.95 suggesting the reduction in Si component in the film. But the Si dangling bonds are the trap centers in nitride film which are responsible for memory action. Therefore the observation based on the ellipsometer data is that Si-rich nitride[3] films have higher trap density compared to stoichiometric and N-rich nitride films.

The leakage current is another indication of presence of traps in the insulators. From the J-E characteristics (Figure 6.2 (b)), it is observed that leakage current is decreasing with increase in the N_2 flow. From the above discussion and ellipsometry data it can be concluded that Si-rich films are more leaky than stoichiometric and N-rich films because of more number of traps confirming the results by Yun et al., [21].

The program transients for both polarities show that hole trapping rate is higher than electron trapping rate. Electron trapping is a minority carrier tunneling process (because the substrate is p-type) and hence depends on the surface inversion and minority carrier life time. During electron trapping, +14 V pulse with varying pulse width from 1 μsec to 1 sec (with light ON) is applied to the gate. Since the pulse widths are small (1 μsec - 10 μsec) compared to the generation time, there is hardly any initial V_{FB} shift with electron trapping.

During hole trapping the Si surface is in accumulation mode and plenty of majority carriers are available for tunneling. Faster initial hole trapping suggests that some hole traps are located near Si-Si_3N_4 interface.

The calculation of trap density from ΔV_{FB} obtained from program transients shows decrease in both electron and hole traps with increasing dilution. Since the traps are the result of Si

[3]RI - 1.999$\sim$2.000 classifies the type of silicon nitride film. If RI $>$ 1.999, the film is Si-rich and RI $<$ 1.999, film is N-rich.

dangling bonds this confirms the initial observation that Si component is decreasing with the dilution.

From the retention measurements (Figure 6.7) it is observed that the initial charge decay is faster with dilution. This suggests location of shallow traps in the nitride film deposited with 1000 sccm N_2. When the dilution gas flow is reduced, as in the case of 400 sccm, initial charge decay is very slow. This suggests that the traps are relatively deep. The analytical model to extract the energy distribution of traps (Section 5.3) also supports above observation. The trap distribution in Fig. 6.8 shows that there is a peak in the trap distribution at 0.85 eV in the case of nitride film deposited with 1000 sccm N_2 dilution. This is a shallow trap level with respect to the silicon nitride conduction band. There is another small peak with wider distribution at around 1.1 eV. The retention, whatever may be, comes from these small number of distributed traps. The peak height is another measure of trap density distributed in particular energy range. This is comparable to the trap density obtained from the program transients (based on the assumption of uniform trap distribution).

In the case of the nitride film deposited with 400 sccm N_2, there is a broad peak centered at 1.25 eV which is relatively deep. Since the distribution is broader and relatively deep energy level, the detrapping rate is small. It is observed from Fig. 6.8 that the trap density shows peaking trend at 1.45 eV. This energy level is very deep and hence the charge decay is slower in sample D2 compared to D3.

The energy distribution of traps in sample D1 shows two small peaks with wider distribution in the energy range of 1.3 - 1.4 eV. In the energy range 0.73 - 1.1 eV, the distribution of the traps in D2 and D3 are almost same. This is due to equal initial decay rates in D1 and D2. It is observed that the small distributed peak of D1 is shifting to the left with increasing N_2. The main peak in D1 is at 1.35 eV, at 1.25 eV in D2 and at 0.85 eV in D3. This suggests that the trap distribution can be tailored with proper amount of diluent.

Overall, the observation is that the trap density is decreasing with increase in the amount of diluent N_2. But the memory property in nitride based devices is the result of Si-N dangling bonds. Since the number of traps is directly proportional to the Si dangling bonds, it is clear that the diluent gas affects the Si dangling bonds during film formation. As described in section 2.6.1, Si-N dangling bond is a incomplete covalent bond. With one electron attached to this bond, the charge state becomes neutral (D^0). With two electrons bonded, the charge state becomes negative (D^-) and the charge state is positive (D^+) when no electron is attached to the dangling bond. The D^- state is a hole trap and D^+ state is a electron trap. Therefore, it is

clear that deposition of SiN_x by LPCVD from SiH_4 and NH_3 forms more number of D^- charge states. This may be the reason for high hole trap density in these nitride films.

Thus, effect of adding diluent N_2 gas during deposition of SiN_x is a two edged sword. It improves the uniformity of the film and at the same time decreases the trap density. The amount of diluent gas to be chosen carefully and proper SiH_4:N_2 ratio is the key for getting optimum results.

Summary of trap characterization techniques by various authors

S.No	Author [References]	Device used	Process parameters	Technique	Trap parameters
1.	Hsia et al., [12]	MNOS	• CVD • $SiH_4/NH_3/N_2$ as carrier gas • Temp. 770^oC	ΔV_{TH}	ΔV_{TH} increases with $SiH_4/NH_3/$ ratio
2.	Manzini. S. [44]	MONOS	• LPCVD nitride with DCS:NH_3 (1:3) • Temp. 800^oC and pressure 280 mTorr	ΔV_{FB} vs time data modeled with TAT and Poole-Frenkel emission	• N_t = 6E18 cm^{-3} • σ_t = 5E-13 cm^2
3.	Fujita et al., [4]	MNS	• CVD • $SiH_4/NH_3/Ar$ as carrier gas • R = SiH_4/NH_3 from 10 to 1000 • Temp. 700^oC	ΔV_{FB} and tunneling approximation by Ross and Wallmark [13]	N_t = 2 - 8 E19 cm^{-3}

4.	Matsuura et al., [24]	MNS	• Photo CVD • SiH_4/NH_3 (1.5 : 45 sccm) • Substrate Temp. 500°C	Trap filling by + ve voltage and measuring discharge current	• Hole traps and N_t = 2E17 cm^{-3} • Energy range: 0.83 eV $<$ E-E_V $<$ 0.94 eV
5.	Roy and White [25]	SONOS	No process details	• Retention measurement at different V_G • Electron back tunneling	N_t = 1E19 cm^{-3}
6.	Endo [45]	MNOS	CVD with Trichloro silane and NH_3	Obtained +ve charge profile	N_t = 4E17 cm^{-3}
7.	Martin and Aymhich [26]	MNOS	• LPCVD with DCS:NH_3 (3:13) • Temp. 800°C and Pressure 400mTorr	Accelerated ΔV_{FB} decay and Poole-Frenkel emission	N_t = 1.5E19 cm^{-3}

8.	Chao and White [27]	SONOS	• LPCVD with DCS:NH_3 (30:100 sccm) • Temp. 730°C and Pressure 300mTorr	• Linear voltage ramp for ΔV_{FB}, ΔQ_P and ΔQ_N • charge centroid	$N_t = 1.5E19$ cm^{-3}
9.	Gu et al., [46]	SONOS	No process details	Measured SILC and correlated with Poole -Frenkel emission	7E12 cm^{-2} eV $^{-1}$
10.	Kim et al., [28]	SONOS	LPCVD with DCS : NH_3 (10:5 and 75 sccm)	Charge decay model from retention measurements	$N_t = 9.5E19$ cm^{-3} in the range 1.15 eV $<$ E-E$_V$ $<$ 1.55 eV
11.	This work	SNS	• LPCVD with SiH_4:NH_3 (40:50 sccm)with N_2 as diluent • Temp. 780°C • Pressure: 300mTorr	• Program transients (ΔV_{FB} vs time) for N_t • Retention measurements at 150°C for trap level • trap rate equation and current measurements for σ_t	• Electron traps : 2 - 2.5E18 cm^{-3} • Hole traps : 4 - 5.3E18 cm^{-3} • Energy range: 0.75 - 1.55 eV w.r.t conduction band • Electron σ_t = 1.3 - 2.5E-14 cm^2 and Hole σ_t = 1.7 - 2.6E-11 cm^2

Summary of trap characterization techniques by various authors

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In this work we studied the effect of different process parameters on the silicon nitride film deposition. Effect of N_2 as a diluent is studied in detail. Process conditions were developed to obtain the silicon nitride films with optimum deposition rate, good uniformity and varying stoichiometry. A breakdown field of 8.5 MV/cm was reported. Trap density and energy level of the traps were determined from program transients and retention measurements at 150^oC. Electron and hole trap densities were calculated from program transients and the values were 2.5E18 cm^{-3} and 5E18 cm^{-3} respectively. Trap capture cross sections were calculated by using trap kinetic equation and current measurements. Energy distribution of electron traps was obtained from the retention measurements and the traps were distributed in the range 0.73 - 1.5 eV.

We concluded the following points from this work:

- Adding diluent N_2 to the reaction chamber during deposition of silicon nitride improves the film uniformity.
- Hole trap density is higher compared to electron trap density in the silicon nitride films deposited with SiH_4-NH_3-N_2 system.
- Increasing the amount of dilution reduces the Si content in the nitride film. Reduction of Si content affects the number of Si-N dangling bonds. Therefore, the trap density decreases with increasing the diluent N_2 flow.
- Proper choice of diluent flow and combination of other process parameters are the key for getting optimum results.

7.2 Future Work

The work done so far and the results produced were on SNS devices. This helped in characterizing the silicon nitride charge trap layer. But the actual device is SONOS and these results should be compared with SONOS devices. The following work would help in characterization of the silicon nitride charge trap layer for SONOS flash memories:

1. Making SNS and SONOS devices simultaneously and characterization for trap parameters.

2. Spatial distribution of traps needs to be determined.

3. Effect of different annealing ambients (N_2 and NH_3) on trap properties to be studied.

References

[1] ITRS, Process Integration, Devices, and Structures, 2009 Edition,pp.13-17.

[2] C. Y. Lu et al., "Future challenges of flash memory technologies," *Microelectron. Eng.*, 86, 3 (March 2009), 283-286.

[3] Shyam Raghunathan, T. Krishnamohan, K. Parat and K. C. Saraswat, "Investigation of Ballistic Current in Scaled Floating-gate NAND FLASH and a solution," *Int. Electron Dev. Meet.(IEDM) Technical Digest*, pp. 819-822, Baltimore, MD, Dec. 2009.

[4] Fujita, S., et al., "Variations of trap states and dangling bonds in CVD Si_3N_4 on Si substrate by NH_3/SiH_4 ratio," *J. Elect. Mat.*, vol. 11, No. 4,p. 795, 1982.

[5] Yang, et al., "Reliability considerations in scaled SONOS nonvolatile memory devices," *Solid State Elect.*, vol. 43,pp. 2025-2032, 1999.

[6] White,M.H, Adams,D.A., Bu,J., "On the go with SONOS," *Circuits and Devices Magazine, IEEE*, vol.16, no.4, pp.22-31, Jul 2000.

[7] Chen,P.C.Y., "Threshold-alterable Si- gate MOS devices," *Electron Devices, IEEE Transactions on*, vol.24, no.5, pp. 584- 586, May 1977.

[8] Stephen N. Keeney, "Dielectric scaling challenges and approaches in floating gate non volatile memories," *Electrochemical society proceedings*, vol. 4, pp.151-158,2004.

[9] Wu K. H., Chien H. C., Tsai T. K., Chang J. W., Chan C. C., Chen T. S., et al., "Phenomenal SONOS performance for next-generation flash memories, " *Proc of symposium on nano device technology*, 2004. p. 35 - 40.

[10] Szu-Yu Wang et al., "Reliability and Processing Effects of Bandgap-Engineered SONOS (BE-SONOS) Flash Memory and Study of the Gate-Stack Scaling Capability," *IEEE transactions on device and materials reliability*, vol. 8,no. 2, June 2008.

[11] C. H. Lee, K. I. Choi, M. K. Cho, Y. H. Song, K. C. Park, and K. Kim, "A Novel SONOS Structure of Si02/SiN/A1203 with TaN metal gate for multi-giga bit flash memories,"*Proceeding of IEDM Technical Digest* (IEEE, Washington, 2003), p. 613.

[12] Yukun Hsia and K. L. Ngai, "MNOS Traps and Tailored Trap Distribution Gate Dielectric MNOS," *Jpn. J. Appl. Phys.*, vol.19, Supplement 19-1, 245-248, 1980.

[13] E. C. Ross and J. T. Wallmark, "Theory of the Switching Behavior of MIS Memory Transistors," *RCA Review*, vol. 30, 366 (1969).

[14] G. Dorda and M. Pulver, "Tunnel Mechanism in MNOS Structures," *Phys. Stat. Solidi*, (a) 1, 71 (1970).

[15] A.V.Ferris-Prabhu., "Theory of MNOS memory device behavior," *IBM J. Res. Dev.*, 17, 2 (March 1973), 125-134.

[16] Marvin H. White and J. Donald cricchi, "Characterization of thin-oxide MNOS memory transistors," *IEEE Trans. Elect. Dev.*, vol. ED-19, no. 12,p. 1280, 1972.

[17] H. Bachhofer, H. Reisinger, E. Bertagnolli, and H. von Philipsborn, "Transient conduction in multidielectric silicon–oxide–nitride–oxide semiconductor structures," *J. Appl. Phys.*, 89, 2791 (2001).

[18] Beyli, V. I., et al., "Silicon nitride in electronics," *Material Science monographs*, vol. 34, pp.55-57, 1988.

[19] Kern, W., Rosler, R. S., "Advances in deposition processes for passivation films," *J. Vac. Sci. Technol.*, vol. 14, No.5, Sept/Oct. 1977.

[20] Schied, E., et al., "Silicon nitride elaborated by low pressure chemical vapour deposition from Si_2H_6 and NH_3 at low temperature," *Material Science and Engineering*, vol.B 17, p. 185, 1993.

[21] Jang-Gn Yun et al., "Formation of Si-Rich Silicon Nitride with Low Deposition Rate by Using LPCVD for Nanoscale Non-Volatile-Memory Application," *Journal of the Korean Physical Society*, vol. 51, pp. S 229-S 233, December 2007.

[22] Arnett,P. C., "Transient conduction in insulators at high fields," *Journal of Applied Physics*, vol. 46, Issue 12, pp. 5236-5243 (1975).

[23] Yang (Larr) Yang, Marvin H. White, "Charge retention of scaled SONOS nonvolatile memory devices at elevated temperatures," *Solid-State Electronics*, vol. 44, Issue 6, 1 June 2000, Pages 949-958.

[24] Hideharu Matsuura et al., "Discharging Current Transient Spectroscopy for Evaluating Traps in Insulators," *Jpn. J. Appl. Phys.*, vol. 34 (1995) L185.

[25] Anirban Roy, Marvin H. White, "Determination of the trapped charge distribution in scaled silicon nitride MONOS nonvolatile memory devices by tunneling spectroscopy," *Solid-State Electronics*, vol. 34, Issue 10, October 1991, Pages 1083-1089.

[26] F. Martin, X. Aymerich, "Characterization of the spatial distribution of traps in Si_3N_4 by field-assisted discharge of metal-nitride-oxide-semiconductor devices," *Thin Solid Films*, vol. 221, Issues 1-2, 10 December 1992, Pages 147-153.

[27] Chen-Chung Chao, Marvin H. White, "Characterization of charge injection and trapping in scaled SONOS/MONOS memory devices," *Solid-State Electronics*, vol. 30, Issue 3, March 1987, Pages 307-319.

[28] Tae Hun Kim et al., "Electron trap density distribution of Si-rich silicon nitride extracted using the modified negative charge decay model of silicon-oxide-nitride-oxide-silicon structure at elevated temperatures," *Appl. Phys. Lett.*, vol. 89,pp. 063508 (2006).

[29] Rosler,R.S., "Low pressure CVD production processes for poly,nitride and oxide," *Solid State Technol.* 20(4), 63(1977).

[30] Makino, T., "Composition and Structure Control by Source Gas Ratio in LPCVD SiN_x," *J. Electrochem. Soc. Solid-state science and Tech.*, vol. 130, No. 2, p. 450, 1983.

[31] Popova, L. I. et al., "Characterization of low pressure chemically vapour deposited thin silicon nitride films," *Thin solid films*, vol.122, p 153,1984.

[32] Karl F. Roenigk and Klavs F. Jensen, "Low pressure CVD of silicon nitride," *J.Electrochem. Soc. Solid-state science and Tech.*, vol. 134, No. 7, p. 1777, 1987.

[33] Ji-Tao Wang et al., "Modeling of LPCVD silicon nitride process," *Journal De Physique*, vol. 50, C5-67, 1989.

[34] Peev, G., et al.,"Modeling of low pressure chemical vapour deposition of Si_3N_4 thin films fron dichlorosilane and ammonia," *Thin solid films*, vol.190, p 341, 1990.

[35] Carlos, H. M., et al., "Thermophysical Properties of Low-residual Stress,Silicon-rich, LPCVD Silicon Nitride Films," *Sensors and Actuators*, A21-A23, p. 856, 1990.

[36] Liu, X. J., et al., "Silicon nitride films deposited by low pressure chemical vapour deposition from SiH_4-NH_3-N_2 system," *Key Engineering Materials*, vols. 264-268, pp. 643-648, 2004.

[37] Yacoubi, K., et al., "Analysis and Modeling of Low Pressure CVD of Silicon Nitride from Silane-Ammonia Mixture," *J. Electrochem. Soc.*, vol. 146, p.3009, 1999.

[38] Bailey, R. S. and Kapoor, V. J., "Variation in the stoichiometry of thin silicon nitride insulating films and its correlation with memory traps," *J. Vac. Sci. Technol.*, Vol. 20, No.3, p. 484, March 1982.

[39] Brown, W. D., et al., "The MONOS memory transistor: Application in a radiation-hard nonvolatile RAM, memory devices," *Solid State Elect.*, vol. 28, No. 9, pp. 877884, 1985.

[40] Kamigaki, Y., et al., "A new portrayal of electron and hole traps in amorphous silicon nitride," *J. Appl. Phys.*, Vol. 68, No.5, p. 2211, September 1990.

[41] Frank R. Libsch and Marvin H. White, "Charge transport and storage of low programming voltage SONOS/MONOS memory devices," *Solid State Elect.*, vol. 33,No. 1, pp. 105-126, 1990.

[42] Wong Hei et al., "Preparation of Thin Dielectric Film for Nonvolatile Memory by Thermal Oxidation of Si-Rich LPCVD Nitride," *J. Electrochem. Soc.*, vol. 148, G275-G278, 2001.

[43] Chramova, L. V. , et al.,"The influence of the chemical composition of silicon nitride films on their thermal oxidation parameters," *Thin solid films*, vol.78, p 303, 1981.

[44] S. Manzini, "Electronic processes in silicon nitride,"*J. Appl. Phys.*, 62, 3278 (1987).

[45] N. Endo, "Charge distributions in silicon nitride of MNOS devices,"*Solid-State Electronics*, Volume 21, Issue 9, September 1978, Pages 1153-1156.

[46] Shaw-Hung Gu, Tahui Wang, Wen-Pin Lu, Yen-Hui Joseph Ku, and Chih-Yuan Lu, "Extraction of nitride trap density from stress induced leakage current in silicon-oxide-nitride-oxide-silicon flash memory,"*Appl. Phys. Lett.*, 89, 163514 (2006).

Appendix A

Process recipes

This appendix describes the tools and process recipes used for fabricating the SONOS/SNS devices. Ultech furnaces are used for process optimization of the SiO_2 (growth and deposition), silicon nitride and n-poly Si. Activation anneal of in-situ doped n-poly Si is done in RTP (Rapid Thermal Process) chamber. Process recipes used in the fabrication of SONOS/SNS devices is given below.

A.1 Tunnel oxide

Tool: Ultech furnace stack # 1 - Tube # 1 (Thin dry oxide tube)
The recipe description given here is applicable to all the Ultech process tubes. This notation used in the recipe is according to the operating manual of Ultech furnace.

The process recipe consists of 1- 7 subprocesses indicated by "step name". Each sub process is identified by "Process X", where X is 1-7. Process parameters during each sub process is given in the rows of the Table A.1.

- **Step time** - Duration of subprocess in seconds.
- **Front (^{o}C)** - Temperature of the front zone.
- **Center (^{o}C)** - Temperature of the center zone.
- **Rear (^{o}C)** - Temperature of the rear zone.
- **Ramp rate (^{o}C/min)** - Ramp up rate of all the three zones.
- **Control thermocouple** - There are two sets of thermocouples installed in each zone of the tube. One is inside the process tube and the other is on the circumference of the stainless

steel tube that encloses process tube. The thermocouple inside process tube is called "PROFILE" and the thermocouple on the outer periphery of the tube is called "SPIKE". The SPIKE thermocouple controls the temperature of the system till Standby step. Once Process 1 started, PROFILE thermocouple takes the control. During actual growth (or deposition) step, the control thermocouple should be set to PROFILE.

- **O_2 flow (sccm)** - Process gas controlled by MFC(Mass Flow Controller).
- **N_2 flow** - N_2 is used during boat-in and boat-out for purging and controlled by flow meter.
- **Boat control** - Automatic boat-in if set to "IN".
- **Step mode** - if YES, subprocess is part of the recipe. if NO, corresponding subprocess will be skipped and control goes to next subprocess.

Parameters of tunnel oxide recipe given here are: Growth temperature - 800°C, time - 15 minutes and O_2 flow - 5000 sccm (maximum MFC limit). Process 3 is the actual growth step and Process 5 is the annealing step. Annealing parameters are: temperature - 900°C, time - 10 minutes and N_2 ambient. This recipe gives oxide thickness about 4.2$\pm$0.4 nm.

A.2 Silicon nitride

Tool: Ultech furnace stack # 2 - Tube # 3 (Silicon nitride)

This is a LPCVD process and the extra process parameter is "Pressure" in addition to those described in previous subsection. Process recipe is given in Table A.2.

The variable parameters are Deposition time (Process 4), gas flows during Process 4, temperature and pressure.

Optimized recipes for this work

S. No.	N_2 flow (sccm)	Deposition time	Thickness (nm)	RI
1.	0	20 min	22.5	2.001
2.	400	25 min	22.3	1.999
3.	1000	40 min	25	1.955

A.3 Low temperature oxide (LTO)

Tool: Ultech furnace stack # 2 - Tube # 2 (Low temperature oxide)
LTO process recipe is same as silicon nitride described in section A.2. except the temperature range, gas flows. In this case SiH_4, O_2 and N_2 are used as process gases. For the recipe given the Table A.3 deposition rate is 1.25 ± 0.15 nm with refractive index 1.43. Therefore, deposition time needs to be changed as per the required thickness.

A.4 n-doped poly Si

Tool: Ultech furnace stack # 3 - Tube # 3 (n-doped poly silicon)
This is similar to silicon nitride except the process gas PH_3 instead of NH_3. The temperature in this case is 630^oC . The recipe in Table A.4 gives n-poly film with thickness 65-70 nm and the sheet resistance of 65 - 120 $\Omega/\square$ (after dopant activation annealing).

A.5 Dopant activation anneal

Tool: Rapid thermal process chamber

- Temperature : 1000^oC
- Ambient : O_2 (900 sccm) + N_2 (450 sccm)
- Time : 30 seconds

A.6 Front side etching - Dry etching

Tool: STS Reactive ion etching

- Gases: CF_4 : O_2 - 40 : 4 (sccm)
- Pressure : 110 mTorr
- RF power : 50 watts.

The etch rate is 30 - 35 nm/min, almost same for both n-poly Si and silicon nitride.

A.7 Back side n-poly Si etch

This is done by wet etching.

Chemical used: HNA (HF :Nitric acid : DI water) - 1: 10 : 19.

A.8 Back side ONO stack etch

This is done by wet etching.

Chemical used: 5:1 HF.

A.9 Back side Al contact

Tool: Thermal evaporator

A.10 Forming gas anneal

Tool: FGA furnace

- Gases: Forming gas (N_2 + H_2)
- Temperature: 420°C
- Time: 20 minutes

Step name	**Prestandby**	**Process 1**	**Process 2**	**Process 3**	**Process 4**	**Process 5**	**Process 6**	**Process 7**	**Post standby**
Step time (sec)		600	120	900	900	600	300	1	
Front (^oC)		800	800	800	800	900	500	300	
Center (^oC)		800	800	800	800	900	500	300	
Rear (^oC)		800	800	800	800	900	500	300	
Ramp rate (^oC/min)		20	10	5	5	10	20	20	
Control thermocouple		PROFILE	PROFILE	PROFILE	PROFILE	PROFILE	SPIKE	SPIKE	
O_2 flow (sccm)		0	50	5000	0	0	0	0	
N_2 flow (sccm)		ON	ON	OFF	ON	ON	ON	ON	
Boat control	IN								
Step mode	YES	YES	YES	YES	YES	YES	YES	YES	KEEP

Table A.1: Process recipe for 4 nm tunnel oxide

Step name	**Prestandby**	**Gas evacuate**	**Tube evacuate**	**Process 1**	**Process 2**	**Process 3**	**Process 4**	**Post standby**
Step time (sec)				600	600	15	1200	
Front (^{o}C)				780	780	780	780	
Center (^{o}C)				780	780	780	780	
Rear (^{o}C)				780	780	780	780	
Ramp rate (^{o}C/min)				20	10	5	5	
Control thermocouple				PROFILE	PROFILE	PROFILE	PROFILE	
SiH_4 flow (sccm)				0	0	0	40	
NH_3 flow (sccm)				0	0	50	50	
N_2 flow (sccm)				0	1000	1000	1000	
Pressure (Torr)				0.001	0.300	0.300	0.300	
Boat control	IN							
Step mode	YES	YES	YES	YES	YES	YES	YES	NO

Table A.2: Process recipe for Silicon nitride

Step name	**Prestandby**	**Gas evacuate**	**Tube evacuate**	**Process 1**	**Process 2**	**Process 3**	**Process 4**	**Post standby**
Step time (sec)				600	600	15	300	
Front (^oC)				430	430	430	430	
Center (^oC)				430	430	430	430	
Rear (^oC)				430	430	430	430	
Ramp rate (^oC/min)				20	10	5	5	
Control thermocouple				PROFILE	PROFILE	PROFILE	PROFILE	
SiH_4 flow (sccm)				0	0	0	40	
O_2 flow (sccm)				0	0	100	100	
N_2 flow (sccm)				0	500	500	500	
Pressure (Torr)				0.001	0.150	0.150	0.150	
Boat control	IN							
Step mode	YES	YES	YES	YES	YES	YES	YES	NO

Table A.3: Process recipe for LTO

Step name	Prestandby	Gas evacuate	Tube evacuate	Process 1	Process 2	Process 3	Process 4	Post standby
Step time (sec)				600	600	15	4500	
Front (^oC)				630	630	630	630	
Center (^oC)				630	630	630	630	
Rear (^oC)				630	630	630	630	
Ramp rate (^oC/min)				20	10	5	5	
Control thermocouple				PROFILE	PROFILE	PROFILE	PROFILE	
SiH_4 flow (sccm)				0	0	90	90	
PH_3 flow (sccm)				0	0		20	
N_2 flow (sccm)				0	210	210	210	
Pressure (Torr)				0.001	0.300	0.300	0.300	
Boat control	IN							
Step mode	YES	YES	YES	YES	YES	YES	YES	NO

Table A.4: Process recipe for n-doped poly Si