Observation and characterization of near-interface oxide traps with C-V techniques

Neil Laurence Cohen
Lehigh University
OBSERVATION AND CHARACTERIZATION OF NEAR-INTERFACE OXIDE TRAPS WITH C-V TECHNIQUES

by

Neil Laurence Cohen

A Thesis
Presented to the Graduate Committee
of Lehigh University
in Candidacy for the Degree of
Master of Science
in
Electrical Engineering

Lehigh University
1995
This thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science.

April 21, 1995
(date)

Dr. Marvin H. White
Thesis Advisor

Dr. Alastair McAulay
Chairman of Department
Acknowledgements

This thesis is dedicated to my parents and brother for their continuous love and support.

I would like to thank my thesis advisor Dr. Marvin H. White for his academic guidance and support, without whom this work would not have been possible. The financial support of the Sherman Fairchild Foundation, National Science Foundation, Naval Research Laboratory through the Defense Nuclear Agency, and the Office of Naval Research is greatly appreciated.

I have had the privilege of working with many graduate students from our group, all of whom have patiently answered my numerous questions and made the research environment enjoyable. In particular, I have worked very closely with Ron Paulsen and Matt Martin the past two years. Their friendship and availability for technical discussions deserve special mentioning. The other members of our research group are much appreciated for creating a productive and stimulating working environment: Dr. Shexin Gong, Ansha Purwar, Bill Wagner, Dr. Zhiganag Ma, Dr. Chun-Yu Malcolm Chen, Dr. Margaret French, Harikaran Sathianathan, and Amit Banerjee.

In addition, I have benefitted from discussions with other members of the Sherman Fairchild Center: Dr. Miltiadis Hatalis, Rob Klamka, David Fleisher, Greg Sarcona, Su-Heng Lin, Mark Stewart, Dr. Fuyu Lin, Dr. Apostolis Voutsas, and Dr. Ji-Ho Kung. I would also like to thank Ray Filozof and Floyd Miller for helping me with my processing and keeping the fabrication laboratory functional. Special thanks to Mrs. Linda Dreisbach for taking care of all the administrative details. Finally, I wish to mention the summer students Abood Al-Quraini and Theerachet Soorapanth for assisting in our group’s research this past summer.
Contents

Acknowledgements iii
List of Figures vii
List of Symbols xi
Abstract 1

1 Introduction 2
1.1 Historical Background 2
1.2 Scope of Thesis 4

2 C-V Measurement and Techniques 5
2.1 Ideal Capacitance-Voltage Measurement 5
  2.1.1 Surface Potential 5
  2.1.2 Accumulation Region 7
  2.1.3 Depletion Region 7
  2.1.4 Inversion Region 8
2.2 Gate Oxide Charge 9
  2.2.1 Fixed Charge 11
  2.2.2 Mobile Charge 11
  2.2.3 Interface Trapped Charge 13
  2.2.4 Near-Interface Oxide Trapped Charge 13
2.3 C-V Techniques 14
  2.3.1 Hi-Lo C-V Technique 14
  2.3.2 Berglund Technique 15
  2.3.3 Linear Voltage Ramp 17
# Low Frequency C-V Model

## 3.1 Motivation

## 3.2 Observation of NIOT Generation
- 3.2.1 Vertical Field Injector
- 3.2.2 Variable Frequency Charge Pumping
- 3.2.3 Anomalous 'Hump'

## 3.3 Evidence of Tunneling

## 3.4 Low Frequency C-V Model
- 3.4.1 Continuum of Interface States
- 3.4.2 Acceptor Type Monoenergetic NIOTs
- 3.4.3 Communication of Interface Traps with NIOTs

## 3.5 Theoretical Results

## 3.6 Donor and Amphoteric NIOTs
- 3.6.1 Donor Type
- 3.6.2 Amphoteric Type

# Characterization of Near-Interface Oxide Traps

## 4.1 Characterization of NIOTs

## 4.2 Device Fabrication

## 4.3 Effective Areal Density

## 4.4 Spatial Distribution

## 4.5 Energy

## 4.6 Summary

# Conclusions

## 5.1 Accomplishments

## 5.2 Applications

## 5.3 Recommendation for Future Work

# References
A Solution to Equation 4.4 55

B Computer Simulations 57

C SONOS Nonvolatile Memory Research 63
  C.1 Oxynitride Films 63
  C.2 Surface Passivation 66

D Series Resistance 67

Vitae 70
List of Figures

2.1 Energy band diagram of a Metal Oxide Semiconductor (MOS) system illustrating the definition of surface potential. Application of a DC gate voltage causes band-bending in the p-type semiconductor and the surface (semiconductor/dielectric interface) intrinsic energy is not equal to the bulk intrinsic energy. ................................................. 6

2.2 Theoretical C-V curve for a p-type MOS capacitor with gate dielectric thickness of 100Å. Depending on the DC gate bias, the C-V measurement can be divided into three regions; accumulation, depletion, and inversion. ................................................................. 9

2.3 Depending on the potential of the semiconductor surface, $\phi_s$, the device is accumulated with majority carriers, depleted of carriers, or inverted with minority carriers. ............................ 10

2.4 Quasi-static C-V data obtained with the LVR technique at 300K. By measuring the triangular area of the LVR curve, the amount of mobile charge can be accurately determined. ................................................. 12

2.5 Energy band-diagram of an inverted MOS system. When the Fermi energy is approximately equivalent to the interface trap energy an electron can trap. Near-interface oxide traps are located at very small distances from the interface (<25Å). ................................................. 14

3.1 Substrate hot-carrier injection was performed with a vertical field injector structure. With this structure, the effects of hot-carriers can be observed more rapidly since the injection is uniform throughout the channel. In addition, independent control over the oxide field is possible. ................................................. 19
3.2 Using variable frequency charge pumping, it is demonstrated NIOT generation simultaneously occurs with interface trap generation at high hot carrier injection levels. The sharp increase in charge recombined per cycle at lower frequencies is attributed to trap-to-trap tunneling of carriers to NIOTs.

3.3 A MOSFET before and after substrate hot carrier injection of \(1 \times 10^{16}\) electrons/cm\(^2\). Similar results obtained for a NIOT-rich ONO device suggest NIOTs are responsible for the anomalous 'hump', rather than interface traps.

3.4 Energy band diagrams of the two-step process involving capture of free carriers by SRH recombination followed by a trap-to-trap tunneling mechanism. (i) When the AC measurement signal is positive, an electron is captured from the conduction band and tunnels to the monoenergetic near-interface oxide trap. (ii) The electron is able to back tunnel and then, unlike charge pumping, emit to the conduction band when the AC signal is negative.

3.5 A plot of trap-to-trap tunneling time vs. tunneling distance for various trap energies. The energy of the trap has minimal effect on the trap-to-trap tunneling time when spatially located within 20Å of the interface.

3.6 Theoretical simulations demonstrate communication from a continuum of interface states to monoenergetic acceptor type NIOTs via trap to trap tunneling introduces a frequency dependent anomalous 'hump' in the inversion regime of the C-V curve.

3.7 Theoretical simulations of a continuum of interface states to spatially located monoenergetic donor type NIOTs via trap to trap tunneling introduces a frequency dependent anomalous 'hump' in the depletion regime of the C-V curve. The curve begins left of ideal and shifts right once the donor traps capture an electron and become neutral.
3.8 Theoretical simulations demonstrate communication from a continuum of interface states to spatially located monoenergetic amphoteric type NIOTs via trap to trap tunneling introduces two frequency dependent anomalous 'humps'.

4.1 A cross section of the ONO transistor fabricated with a NIOT-rich nitride layer located 12Å from the interface.

4.2 Experimental low frequency C-V curves for a NIOT-rich ONO transistor. The NIOTs are located 12Å away from the Si-SiO₂ interface. The hump increases as frequency is decreased since charge can tunnel further into the oxide and fill more NIOTs.

4.3 Utilizing the Hi-Lo C-V technique, an effective areal density of both interface traps and NIOTs may be obtained. The u-shaped distribution is due to interface traps and the superimposed 'hump' is due to NIOTs. As the measurement frequency is decreased, more NIOTs can fill and the peak value $C_{TOT_{max}}(f)$ increases.

4.4 Variable frequency Hi-Lo C-V measurements have been taken on a trap-rich ONO transistor with $D_{it}=2\times10^{11}$ (cm$^{-2}$eV$^{-1}$) and a breakpoint at approximately 175kHz is observed. For frequencies below the breakpoint, charge has sufficient time to tunnel to the NIOTs placed 12Å away.

4.5 Plotting (4.5), the breakpoint in Fig. 4.4 corresponds to a tunneling distance of 12-13Å, consistent with the location of the NIOT-rich nitride layer as determined by ellipsometry. A higher interface trap density increases the distance charge can tunnel at a particular frequency.
4.6 A comparison of extracted areal densities between variable frequency Hi-Lo C-V and charge pumping. Close agreement between extracted trap densities from charge pumping and C-V measurements is observed. The constant offset is due to the difference in interface trap response between both techniques and the drop-off is due to non-equilibrium as the trap-to-trap tunneling time approaches the emission time. 43

4.7 By differentiating the effective areal density with respect to \( \log(f) \), a volume density of NIOTs is obtained. Excellent convergence between variable frequency Hi-Lo C-V and charge pumping exists provided equilibrium with tunneling electrons is maintained \((\tau_T > \tau_e)\). 45

4.8 A summary of the variable frequency Hi-Lo C-V extraction technique. With this technique, areal density, spatial distribution, and energy characterization of NIOTs is possible. 46

5.1 Low frequency C-V curves of the 12Å trap-rich ONO device under illumination. A larger anomalous ‘hump’ is observed due to the reduction in emission time. 49

D.1 Series resistance results in a lowering of the C-V curve in the strongly accumulated region. The series resistance problem is more evident at higher measurement frequencies. 69
List of Symbols

- $A_G$: area of the gate ($\text{cm}^2$)
- $A_{ion}$: area of C-V curve due to mobile ions
- $C_{HF}$: high frequency capacitance ($\text{F}$)
- $C_{it}$: interface trap capacitance ($\text{f}$)
- $C_{LF}$: low frequency capacitance ($\text{F}$)
- $C_m$: measured capacitance ($\text{F}$)
- $C_{N1OT}$: near-interface oxide trap capacitance ($\text{F}$)
- $C_{N1OT_{\text{max}}}$: frequency dependent maximum near-interface oxide trap capacitance ($\text{F}$)
- $C_{ox}$: oxide capacitance ($\text{F}$)
- $C_s$: semiconductor capacitance ($\text{F}$)
- $C_{TOT}$: effective trap capacitance ($\text{F}$)
- $\bar{D}_{it}$: average density of interface traps ($\text{cm}^{-2}\text{eV}^{-1}$)
- $E_{c0}$: conduction band energy at the Si-SiO$_2$ interface ($\text{eV}$)
- $E_F$: Fermi level ($\text{eV}$)
- $E_g$: semiconductor energy gap ($\text{eV}$)
- $E_i(x)$: intrinsic energy level at distance $x$ ($\text{eV}$)
- $E_{ib}$: bulk intrinsic energy level ($\text{eV}$)
- $E_{is}$: intrinsic energy level at the Si-SiO$_2$ interface ($\text{eV}$)
- $E_t$: interface trap energy ($\text{eV}$)
- $E_{ts}$: near-interface trap energy ($\text{eV}$)
- $E_{vs}$: valence band energy at the Si-SiO$_2$ interface ($\text{eV}$)
- $f$: AC measurement frequency ($\text{Hz}$)
- $f^0$: neutral charge state occupation probability
- $f^-$: negative charge state occupation probability
- $f^+$: positive charge state occupation probability
\( G_m \) real admittance component

\( h \) Planck’s constant (J-sec)

\( I_G \) gate current (A)

\( I_{NIOT} \) near-interface oxide trap current (A)

\( k \) Boltzmann’s constant (eV/K)

\( K_{ox} \) dielectric constant

\( L_D \) extrinsic Debye length (cm)

\( N_B \) substrate doping (cm\(^{-3}\))

\( N_{NIOT}(f) \) frequency dependent areal density of near-interface oxide traps (traps/cm\(^2\))

\( N_{NIOT}^- \) areal density of filled near-interface oxide traps (traps/cm\(^2\))

\( n_i \) intrinsic carrier density (cm\(^{-3}\))

\( n_e(t) \) carrier density (electrons) in the traps (cm\(^{-2}\))

\( n_h(t) \) carrier density (holes) in the traps (cm\(^{-2}\))

\( N_T(x) \) volume density of near-interface oxide traps (cm\(^{-3}\))

\( P(U_s) \) normal probability distribution function

\( Q_f \) total fixed charge density (C/cm\(^2\))

\( Q_G \) total gate charge density (C/cm\(^2\))

\( Q_{it} \) total interface trapped charge density (C/cm\(^2\))

\( Q_m \) total mobile charge density (C/cm\(^2\))

\( Q_{NIOT} \) total near-interface oxide trap charge density (C/cm\(^2\))

\( Q_{NIOT}^- \) average near-interface oxide trap charge per cycle

\( Q_s \) total semiconductor charge density (C/cm\(^2\))

\( q \) electronic charge (C)

\( R_s \) series resistance (ohms)

\( T \) temperature (K)

\( t \) time (sec)

\( U_F \) normalized Fermi potential

\( U_s \) normalized surface potential

\( U_t \) normalized trap potential

\( U_s \) mean-normalized surface potential
\[ V_{FB} \] flatband voltage (V)
\[ V_G \] gate voltage (V)
\[ V_{GB} \] gate to bulk voltage (V)
\[ V_{TH} \] threshold voltage (V)
\[ v_{th} \] thermal velocity (cm/sec)
\[ \bar{x} \] charge centroid (cm)
\[ x_m(f) \] maximum tunneling distance (cm)
\[ x_{min} \] minimum near-interface oxide trap distance (cm)
\[ x_{ob} \] blocking oxide thickness (cm)
\[ x_{ot} \] tunnel oxide thickness (cm)
\[ x_{ox} \] oxide thickness (cm)
\[ x_n \] nitride thickness (cm)
\[ Y_m \] measured admittance (S)
\[ \alpha \] ramp rate (V/sec)
\[ \alpha_1 \] attenuation coefficient in SiO_2
\[ \alpha_2 \] attenuation coefficient in Si
\[ \epsilon_0 \] permittivity of free space (F/cm)
\[ \omega \] angular frequency (radians/sec)
\[ \phi(x) \] semiconductor potential (V)
\[ \phi_B \] Si and SiO_2 conduction band barrier height (eV)
\[ \phi_F \] Fermi potential (V)
\[ \phi_s \] surface potential (V)
\[ \phi_{so} \] additive constant (V)
\[ \rho_{NIO} \] volume density of monoenergetic near-interface oxide traps (cm\(^{-3}\)eV\(^{-1}\))
\[ \rho_{NIO} \] volume density of occupided monoenergetic near-interface oxide traps (cm\(^{-3}\)eV\(^{-1}\))
\[ \sigma_n \] trap capture cross section for electrons (cm\(^2\))
\[ \sigma_p \] trap capture cross section for holes (cm\(^2\))
\[ \sigma_s^2 \] normalized variance of the distribution
\[ \tau_T(E_t, x) \] trap-to-trap tunneling time constant (sec)
\[ \tau_e \] SRH trapping center emission time (sec)
\( \tau_0 \)  
trap-to-trap tunneling parameter (sec)

\( \tau_p \)
majority carrier capture time constant (sec)

\( \tau_{po} \)
intrinsic majority carrier time constant (sec)

\( \tau_{no} \)
midgap trapping center emission time (sec)
Abstract

Characterization of near-interface oxide traps is essential to evaluate the dielectric reliability of today’s thin gate CMOS devices. Presently, variable frequency charge pumping is a viable approach for determining the spatial distribution of near-interface oxide traps. Charge pumping, however, requires a transistor structure and the measurement is not widely adopted by industry.

The capacitance-voltage (C-V) measurement is a useful characterization tool due to its simplicity, suitability for extracting a variety of parameters, and necessitates only a capacitor structure. In addition, familiarity with the measurement is ubiquitous. For these reasons, a method of characterizing near-interface oxide traps with conventional C-V techniques would be very useful to evaluate device reliability and for process monitoring.

In this work, a low frequency C-V model has been developed to interpret the effect of near-interface oxide traps which communicate with the Si-SiO₂ interface traps and influence C-V characteristics. The model is based on trap-to-trap tunneling of carriers from a continuum of interface traps to acceptor type monoenergetic near-interface oxide traps (NIOTs). At low measurement frequencies, sufficient time exists for the trap-to-trap tunneling mechanism to occur. We have employed a trap-rich thin oxide ONO device (SONOS) to provide evidence of tunneling and suggest strongly the anomalous ‘hump’ repeatably observed in low frequency C-V curves of stressed devices is attributed to near-interface oxide traps, rather than so-called ‘interface’ traps.

From the low frequency C-V model, an extraction procedure based solely on C-V techniques permits both spatial and energetic characterization of near-interface oxide traps. To support the proposed model, a spatial and energetic distribution of near-interface oxide traps has been extracted for a trap-rich ONO transistor with variable frequency Hi-Lo C-V measurements and is consistent with conventional bi-level charge pumping results.
Chapter 1

Introduction

1.1 Historical Background

Over the past four decades, research toward understanding the fundamental device physics of the MOS system has allowed the microelectronics industry to progress at an astonishing rate. Over this time, a large emphasis has been placed on the scaling of semiconductor devices in order to satisfy customer demand for faster, smaller, and low power products. As a result, today's state of the art CMOS technology necessitates submicron channel lengths and gate oxides less than 100Å thick. Unfortunately, scaling of the MOS system has amplified the effects of undesirable device phenomenon such as tunneling of carriers and hot-carrier injection[1]. Thus, the reliability of small geometry devices is a major concern for the semiconductor industry.

Semiconductor device reliability testing is performed by application of a known stress condition for a given period of time. With electrical characterization techniques such as capacitance-voltage, current-voltage, and charge pumping, it is possible to evaluate the withstandability to the applied stress. This is accomplished by monitoring the change in device parameters (ie. interface trap density, transconductance, mobility, and threshold voltage) after the applied stress. Some common stress conditions to simulate actual device operating environments include hot-carrier injection, Fowler-Nordheim[2]/high-field stressing, irradiation exposure, and snapback[3].

The electrical properties of the semiconductor-dielectric interface play a key role in device reliability. Experimental conductance techniques and analysis methods by Nicollian and Goetzberger[4] provide the basis for interface modeling. Due to
CHAPTER 1. INTRODUCTION

the scaling of gate dielectric thicknesses, near-interface oxide trapping has recently become an increasingly important reliability concern. Near-interface oxide traps are spatially located within a tunneling distance of the semiconductor-dielectric interface. The near-interface oxide traps can communicate with interface traps by exchanging charge provided the appropriate conditions are available for the tunneling mechanism to occur. Previous work by Heiman and Warfield[5] extended the existing model of the interface to incorporate the effect of near-interface oxide traps on the MOS capacitance. However, a detailed expression of the tunneling mechanism was not available at the time. In this work, a trap-to-trap tunneling mechanism is proposed as the means of communication between interface and near-interface oxide traps.

Verification of the trap-to-trap tunneling mechanism has been provided by performing C-V measurements on devices with a trap-rich silicon nitride layer located near (12-20Å) from the Si-SiO₂ interface. These poly-Silicon-Oxide-Nitride-Oxide-Silicon (SONOS) devices act as nonvolatile memory elements since the silicon nitride layer can trap and store charge tunnelled from the interface.

A consistent occurrence in highly stressed devices has been the presence of an anomalous 'hump' in the inversion region of the C-V curve. The anomalous 'hump' has been observed when evaluating device reliability via hot carrier injection[6], Fowler-Nordheim/high-field stressing[7], and radiation exposure[8] and has been attributed to so-called 'interface' state generation. The anomalous 'hump' is more apparent at higher stressing conditions and is observable only at low AC measurement frequencies. We have proposed that the origin of the anomalous 'hump' is attributed to trap-to-trap tunneling of carriers from interface traps to near-interface oxide traps. At low AC measurement frequencies, sufficient time exists for the trap-to-trap tunneling mechanism to occur. The trap-rich SONOS transistor has been employed to provide experimental evidence and suggest strongly the anomalous 'hump' repeatably observed in low frequency C-V curves is attributed to near-interface oxide traps, rather than so-called 'interface' traps.
CHAPTER 1. INTRODUCTION

1.2 Scope of Thesis

This thesis is devoted to modeling the effects of near-interface oxide trapping on the C-V measurement and the use of conventional C-V techniques for near-interface oxide trap characterization.

In the next chapter, a discussion of the ideal C-V measurement is reviewed. The different regions of the C-V measurement are illustrated and analyzed in terms of semiconductor surface potential. The effect of gate oxide charges on the C-V measurement is presented and conventional C-V techniques are described for gate oxide charge characterization. Chapter 3 begins by providing experimental evidence of tunneling in low frequency C-V curves on stressed gated-diode MOS structures. This evidence of tunneling provides a basis for the low frequency C-V model developed in the remainder of the chapter. The model is based on trap-to-trap tunneling from a continuum of interface traps to spatially located monoenergetic near-interface oxide traps. From the low frequency C-V model, an extraction technique using the aforementioned C-V techniques is developed in Chapter 4 permitting spatial and energetic characterization of near-interface oxide traps. To support the low frequency C-V model and extraction procedure, extracted near-interface oxide trap densities for a trap-rich SONOS transistor is compared with well established bi-level charge pumping techniques. Finally, concluding remarks, applications of the extraction technique, and recommendations for further investigation are discussed in the fifth chapter.
Chapter 2

C-V Measurement and Techniques

2.1 Ideal Capacitance-Voltage Measurement

The metal-oxide-semiconductor (MOS) capacitor serves as a useful test structure to characterize devices due to its ease of fabrication and suitability for extracting a variety of parameters. Processing of MOS capacitors can be completed in a small amount of time when compared to transistors and parameters extracted from MOS capacitors are useful for MOSFET transistor development. One common electrical measurement performed on a MOS capacitor is capacitance-voltage (C-V). The C-V measurement is a well accepted method for characterizing the MOS system. Some parameters obtained from C-V measurements include oxide thickness, doping profile, threshold voltage, and interface trap density[9].

To perform a C-V measurement, the semiconductor bulk is grounded while a linearly varying DC bias and small AC sinusoid test signal (amplitude< $kt/q$) are applied simultaneously to the gate. The small AC signal causes the amount of charge to fluctuate on the gate resulting in semiconductor surface charge fluctuations. A differential capacitance arises due to the small variations in charge. The DC bias changes the magnitude and amount of charge available at the semiconductor surface and will be considered in three different regions: accumulation, depletion, and inversion. The case of a p-type semiconductor will be examined.

2.1.1 Surface Potential

To be consistent, all potentials in the semiconductor will be referenced to the intrinsic Fermi level in the bulk. Defining $x=0$ at the Si-SiO$_2$ interface, the potential
for any distance $x$ in the semiconductor is defined as

$$\phi(x) = (E_{ib} - E_i(x))/q$$

(2.1)

where $E_{ib}$ is the intrinsic energy in the semiconductor bulk and $E_i(x)$ is the intrinsic energy for a given distance $x$ away from the interface. The potential at the semiconductor surface, $\phi_s$, is determined by applying (2.1) at the interface ($x=0$) such that

$$\phi_s = (E_{ib} - E_{is})/q$$

(2.2)

where $E_{is}$ is the intrinsic Fermi level at the surface. An increase in surface potential occurs as the DC bias is increased in the positive direction since more 'band-bending' is required to maintain charge neutrality in the MOS system (Fig. 2.1).

Figure 2.1: Energy band diagram of a Metal Oxide Semiconductor (MOS) system illustrating the definition of surface potential. Application of a DC gate voltage causes band-bending in the p-type semiconductor and the surface (semiconductor/dielectric interface) intrinsic energy is not equal to the bulk intrinsic energy.
CHAPTER 2. C-V MEASUREMENT AND TECHNIQUES

2.1.2 Accumulation Region

When a negative DC bias voltage is applied to the gate, holes (majority carriers) are attracted to the surface to balance the negative charge on the gate. The surface is said to be *accumulated* when the surface potential condition $-(E_g/2q - \phi_F) \leq \phi_s \leq 0$ is satisfied where $\phi_F$ is the Fermi potential and $E_g$ is the energy-gap of the semiconductor. The Fermi potential is also referenced to the intrinsic energy in the bulk

$$\phi_F = \frac{E_{ib} - E_F}{q} = \frac{kT}{q} \ln\left(\frac{N_B}{n_i}\right)$$  \hspace{1cm} (2.3)

where $N_B$ is the bulk doping density and $n_i$ the intrinsic carrier concentration. The Fermi potential remains constant regardless of the change in surface potential. The accumulated layer of holes at the Si-SiO$_2$ interface act as a capacitor plate equal to the gate area. Therefore, the MOS capacitance in accumulation consists of two parallel plates separated by the gate dielectric. The measured capacitance is simply the oxide capacitance, $C_{ox}$. The capacitance for a parallel plate capacitor is

$$C_{ox} = \frac{K_{ox} \epsilon_o}{x_{ox}}$$  \hspace{1cm} (2.4)

where $x_{ox}$ is the dielectric thickness, $K_{ox}$ is the dielectric constant, and $\epsilon_o$ is the permittivity of free space.

A special condition known as flatband arises when the intrinsic bulk energy equals the intrinsic surface energy ($\phi_s = 0$).

2.1.3 Depletion Region

As the gate bias increases, the resultant electric field causes the semiconductor surface to *deplete* of carriers and act as an insulator. The depletion width increases as more carriers are repelled from the surface. Depletion occurs when the surface potential is between $0 \leq \phi_s \leq \phi_F$. When this condition occurs, the insulating depletion region causes the effective parallel capacitor plates to move further apart and the overall capacitance consists of oxide and semiconductor capacitances, $C_s$, 


in series. A reduction in measured capacitance is observed as the depletion region widens.

### 2.1.4 Inversion Region

A surface potential of $\phi_s \leq \phi_s \leq (E_g/2q + \phi_F)$ results in an abundance of electrons (minority carriers) at the semiconductor surface. The surface is said to be inverted due to the high minority carrier concentration. The gate voltage inducing a surface potential of $\phi_s = 2\phi_F$ is known as the threshold voltage and marks the onset of strong inversion. The threshold voltage is generally a good reference point for sufficient channel conduction in MOSFETs. Similar to accumulation, a parallel plate capacitor exists between the positive gate charge and the inversion layer of negative surface charge separated by the gate dielectric. The measured capacitance is once again simply the oxide capacitance, $C_{ox}$.

In a capacitor structure, the lack of source/drain minority carrier regions can make it difficult for surface inversion to occur. Therefore, it is often desirable to generate excess minority carriers by illuminating the device or ion implanting a minority carrier grid during wafer fabrication.

It should be noted that the above analysis assumes the AC measurement signal is at a sufficiently low frequency so the generation-recombination of electrons have time to respond to the small signal voltage fluctuations. At higher frequencies, the inversion layer can no longer respond to the quickly changing charge variations and the C-V curve will not rise back to $C_{ox}$.

It is interesting to note that the total energy band-bending from accumulation to inversion is $E_g$. A theoretical C-V curve has been computed and plotted in Figure 2.2. In addition, the energy band-bending and equivalent circuits are illustrated for all three regions of the C-V measurement in Fig. 2.3.
CHAPTER 2. C-V MEASUREMENT AND TECHNIQUES

Figure 2.2: Theoretical C-V curve for a p-type MOS capacitor with gate dielectric thickness of 100Å. Depending on the DC gate bias, the C-V measurement can be divided into three regions; accumulation, depletion, and inversion.

2.2 Gate Oxide Charge

The previous discussion concerning the C-V measurement was idealized in order to introduce fundamental theory. Experimentally, however, charges present in the gate oxide introduce deviations to the ideal C-V curve and can alter predicted device performance considerably. Various types of charge can arise in the gate oxide of a MOS system such as fixed, mobile, interface, and near-interface oxide trapped charge. Each will be discussed in terms of their subsequent effect on the C-V measurement.
Figure 2.3: Depending on the potential of the semiconductor surface, \( \phi_s \), the device is accumulated with majority carriers, depleted of carriers, or inverted with minority carriers.
CHAPTER 2. C-V MEASUREMENT AND TECHNIQUES

2.2.1 Fixed Charge

Charge will trap in the gate oxide during the device fabrication process. Trapped charge spatially located far from the Si-SiO₂ interface is unable to free itself from the gate oxide, hence the name fixed charge. Since the amount of charge remains constant regardless of the surface potential, the fixed charge will not contribute additional capacitance to the theoretical C-V curve. However, the fixed charge present in the oxide will cause the entire C-V curve to shift parallel by a voltage \( \Delta V \). A net positive charge shifts the experimental curve left of ideal and a net negative charge causes a parallel shift to the right. The amount of fixed charge can be determined by determining the voltage shift between experimental and theoretical C-V curves

\[
Q_f = -\frac{K_{ox} \varepsilon_o \Delta V}{(\bar{x}_{ox} - \bar{x})}
\]  

(2.5)

where \( Q_f \) is the amount of fixed charge and \( \bar{x} \) is the centroid of the fixed charge.

The amount of fixed charge has been correlated to the oxidation temperature[10]. A higher oxidation temperature reduces the amount of fixed charge trapped in the oxide. Fixed charge can also be reduced by annealing in an unreactive environment (ie. nitrogen, argon) immediately after oxidation.

2.2.2 Mobile Charge

Another class of gate oxide charge known as mobile charge is typically introduced from alkali-metal ions such as sodium. These ions can drift into the oxide at low voltages and cause a shift to the measured C-V curve. For this reason, only a small amount of mobile charge (\( \leq 5 \times 10^{11} / \text{cm}^2 \)) can be tolerated to maintain the important reliability issue of flatband voltage stability. As temperature increases, mobile charge effects become more detectable since the ionic mobility increases. One method of quantifying mobile charge is by performing a quasi-static C-V (linear voltage ramp) measurement on a device at elevated temperatures, commonly referred to as the Kuhn technique[9]. From this technique, the amount of mobile charge can be
CHAPTER 2. C-V MEASUREMENT AND TECHNIQUES

Figure 2.4: Quasi-static C-V data obtained with the LVR technique at 300K. By measuring the triangular area of the LVR curve, the amount of mobile charge can be accurately determined.

determined by

\[
Q_m = \frac{1}{\alpha A_g} \int_{-V_o}^{+V_o} dV GB I_g = \frac{A_{\text{ion}}}{\alpha A_g}
\]

where \(A_{\text{ion}}\) is the C-V area due to mobile ions (Figure 2.4), \(A_g\) is the gate area, \(\alpha\) is the measurement ramp rate, and \(I_g\) is the gate current.

To keep mobile charge at tolerable levels, frequent monitoring of the easily transmittable alkali-metal ions is necessary. Alkali-metal ions are introduced by certain processing steps such as the metalization, polysilicon deposition, and handling processes. In addition, chemicals (including water) are also sources of ionic contamination. If unacceptable levels exist, furnaces must be recleaned. An alkali-metal removal bath of hydrochloric acid and hydrogen peroxide is part of the standard
RCA pre-oxidation wafer cleaning procedure.

2.2.3 Interface Trapped Charge

Historically, interface traps have been treated as existing at the semiconductor-insulator interface with distributed energy levels within the forbidden energy band-gap. In practice, however, the interface will be spread out spatially over some layers of the lattice to relieve the strain[11]. As gate voltage is swept, the occupancy of the interface traps can change depending on the surface potential. Because of band-bending, an interface trap is capable of changing charge states when the Fermi energy approximately equals the interface trap energy as illustrated by Fig. 2.5. The changing of charge states introduces an additional capacitance to the MOS system. The interface trap capacitance, $C_{it}$, is modeled in parallel with the semiconductor capacitance, $C_s$. An observed C-V stretch-out is observed due to the interface trapped charge[10]. A detailed modeling of the interface will be discussed with greater detail in subsequent chapters. A technique known as Hi-Lo C-V can be used to measure interface trap densities, as well as charge pumping[12].

Commercial CMOS devices have interface trap densities on the order of $10^{10}$ traps/cm²eV[13]. This is achieved by oxidizing in a ultra-clean dry environment, rather than steam. A post-metal-anneal (PMA) reduces interface trap densities by removing processing induced damage from sputtering.

2.2.4 Near-Interface Oxide Trapped Charge

The final category of gate oxide charge to be discussed is near-interface oxide trapped charge (NIOT, or border trapped charge[14]). The term near-interface is used to classify traps located at very small distances into the gate oxide, typically within 25Å of the Si-SiO₂ interface (Fig. 2.5). Near-interface oxide trapping effects are more evident in today's low power, small-geometry devices since their effects become more prominent as gate dielectrics scale. To this day, however, a comprehensive investigation regarding the effect of NIOTs on the C-V measurement has
CHAPTER 2. C-V MEASUREMENT AND TECHNIQUES

Near-Interface Oxide Trap

E_{FM}  

E_{IS}  

Interface Traps

semiconductor/dielectric interface (x=0)

Figure 2.5: Energy band-diagram of an inverted MOS system. When the Fermi energy is approximately equivalent to the interface trap energy an electron can trap. Near-interface oxide traps are located at very small distances from the interface (<25Å).

not been undertaken.

2.3 C-V Techniques

Numerous techniques based on simple high and low frequency C-V measurements have been developed to characterize charge trapping in the gate dielectric of the MOS system. The following sections describe the C-V techniques used in the experimental work of this thesis.

2.3.1 Hi-Lo C-V Technique

The Hi-Lo C-V technique is performed by taking two separate C-V measurements. One is taken at high frequency to prevent interface trap response. In doing
so, the equivalent high frequency capacitance is simply the oxide capacitance and semiconductor capacitance in series. The other C-V measurement is taken under identical conditions, however, at a low frequency to permit trap response. At low frequency, the oxide capacitance is in series with the parallel branch of trap and semiconductor capacitances. Therefore, the normalized high and low frequency capacitances are expressed by the following equations

\[
\frac{C_{HF}}{C_{ox}} = \frac{C_s}{C_{ox} + C_s} \quad (2.7)
\]

and

\[
\frac{C_{LF}}{C_{ox}} = \frac{C_s + C_{it}}{C_{ox} + C_s + C_{it}} \quad (2.8)
\]

where \(C_{it}\) is an effective trap capacitance. By solving (2.7) for \(C_s\) and then substituting into (2.8), an expression is obtained for the normalized effective trap capacitance.

\[
\frac{C_{it}}{C_{ox}} = \frac{C_{LF} - C_{HF}}{C_{ox} - C_{ox} - C_{HF}} \quad (2.9)
\]

As previously mentioned, the principal application of the Hi-Lo C-V technique is determination of the average density of interface traps, \(\bar{D}_{it}\) (traps/cm\(^2\)eV\(^{-1}\)). This is achieved by evaluating the trap capacitance at midgap and dividing by electronic charge.

\[
\bar{D}_{it} = \frac{C_{it}}{q} \quad (2.10)
\]

### 2.3.2 Berglund Technique

It is often necessary to determine the surface potential, \(\phi_s\), corresponding to a particular gate voltage. For example, in order to determine the threshold of a device the voltage corresponding to surface potential \(2\phi_F\) must be determined. A popular method to determine surface potential is the Berglund Technique[15].

Due to charge neutrality in the MOS system, the sum of all charges must equal zero. Using this fact, and by differentiation with respect to gate-to-bulk voltage, the following expression is obtained:

\[
\frac{\partial Q_G}{\partial V_{GB}} + \frac{\partial Q_s}{\partial V_{GB}} + \frac{\partial Q_{it}}{\partial V_{GB}} = 0 \quad (2.11)
\]
CHAPTER 2. C-V MEASUREMENT AND TECHNIQUES

Capacitances are defined as follows:

\[ C_{LF} = \frac{\partial Q_G}{\partial V_{GB}} \tag{2.12} \]

\[ C_s = -\frac{\partial Q_s}{\partial \phi_s} \tag{2.13} \]

and

\[ C_{it} = -\frac{\partial Q_{it}}{\partial \phi_s} \tag{2.14} \]

Substitution of (2.12) (2.13) and (2.14) into (2.11) yields,

\[ C_{LF} - C_s \frac{\partial \phi_s}{\partial V_{GB}} - C_{it} \frac{\partial \phi_s}{\partial V_{GB}} = 0 \tag{2.15} \]

Solving (2.15) for \( C_{LF} \),

\[ C_{LF} = \frac{1}{\frac{1}{C_{LF}} - \frac{1}{C_{ox}}} \frac{\partial \phi_s}{\partial V_{GB}} \tag{2.16} \]

Cross multiplying and integrating both sides results in the following equation:

\[ \int_{V_{GB_0}}^{V_{GB}} dV_{GB} (1 - \frac{C_{LF}}{C_{ox}}) = \int_{\phi_s}^{\phi_s} d\phi_s \tag{2.17} \]

By performing the integration and solving for \( \phi_s \), a final expression is obtained for surface potential

\[ \phi_s = \phi_{s_0} + \int_{V_{GB_0}}^{V_{GB}} dV_{GB} (1 - \frac{C_{LF}}{C_{ox}}) \tag{2.18} \]

where \( \phi_{s_0} \) is an additive constant. Equation 2.18 indicates the surface potential within an additive constant is simply the area of the normalized low frequency C-V curve.

There are numerous procedures to determine the additive constant \( \phi_{s_0} \). For example, since theoretical and experimental C-V curves deviate only slightly in strong accumulation, the surface potential is known to be \(-\frac{E_g}{2q} - \phi_F\). Experimental methods to determine the \( 2\phi_F \) point, or threshold voltage, can also be used to evaluate the additive constant[16, 17].
2.3.3 Linear Voltage Ramp

The linear voltage ramp (LVR) technique [9, 18] is commonly referred to as quasi-static C-V. A LVR measurement is performed by applying a slowly varying (quasi-static) triangular voltage to the gate of a device while an electrometer measures the resultant bulk current. In practice, however, the terminals are often reversed since ramping the semiconductor bulk results in less noise. The ramp rate, $\alpha$ (V/sec), is characteristic of the applied triangular wavefunction. Using definitions

$$\alpha = \pm \frac{\partial V_{GB}}{\partial t} \quad (2.19)$$

and

$$C_{LF} = \frac{\partial Q_G}{\partial V_{GB}} \quad (2.20)$$

the gate current can be expressed with use of the chain rule

$$I_G = \frac{\partial Q_G}{\partial t} = \pm \frac{\partial Q_G}{\partial V_{GB}} \cdot \frac{\partial V_{GB}}{\partial t} \quad (2.21)$$

Therefore, the quasi-static capacitance is proportional to the gate current normalized by the ramp rate.

$$C_{LF} = \pm \frac{I_G}{\alpha} \quad (2.22)$$
Chapter 3
Low Frequency C-V Model

3.1 Motivation

As gate dielectrics continue to scale in order to satisfy low power requirements and faster operating speeds, the influence of interface and near-interface oxide traps on device performance becomes more prominent. Threshold voltage shifts\[19\], mobility and transconductance degradation\[17, 20\], premature dielectric breakdown\[21\], and 1/f noise\[22\] are reliability concerns for scaled CMOS devices which are associated with interface traps and NIOTs. In addition, near-interface oxide trapping may be observed as dielectric relaxation effects\[23\] in sampled-data analog circuitry and DRAM storage capacitors.

For these reasons, the ability to characterize NIOTs electrically is necessary for the process development of more-reliable gate dielectrics. Since the C-V measurement is commonly used to monitor interface trap generation, it would be very useful if the same could be done for NIOTs.

3.2 Observation of NIOT Generation

First, it is necessary to determine the appropriate stressing conditions for NIOT generation to occur. A vertical field injection structure has been used to degrade the gate oxide by substrate hot carrier injection. The variable frequency charge pumping technique\[24\] will be used to correlate a relationship between substrate hot-carrier injection levels and NIOT generation.
3.2.1 Vertical Field Injector

In order to stress devices via substrate hot-carrier injection, a vertical field injection structure [17] has been used (Fig. 3.1). A positive gate bias and negative substrate bias induce an electric field such that minority carriers (electrons) are directed toward the channel. By biasing a nearby minority carrier junction (injector) more negative than the bulk, supplied electrons are attracted towards the bulk and then accelerated by the electric field into the gate oxide. This structure allows independent control over the oxide field and uniform injection throughout the entire channel.

![Diagram of Vertical Field Injector](image)

Figure 3.1: Substrate hot-carrier injection was performed with a vertical field injector structure. With this structure, the effects of hot-carriers can be observed more rapidly since the injection is uniform throughout the channel. In addition, independent control over the oxide field is possible.
CHAPTER 3. LOW FREQUENCY C-V MODEL

3.2.2 Variable Frequency Charge Pumping

Interface state generation due to hot carrier stressing is typically monitored with bi-level charge pumping[25] and low frequency C-V measurements[6]. However, the variable frequency charge pumping technique has the desirable capability of monitoring both interface and NIOT generation simultaneously. NIOT generation is observed as a rapid increase of charge pumped per cycle at low frequencies and has been attributed to trap-to-trap tunneling of carriers to NIOTs[24].

With the vertical field injection structure, substrate hot carrier injection was performed on pure oxide MOSFETs at various injection levels. The variable frequency charge pumping technique was performed at each injection level (Fig. 3.2) and indicates NIOT generation occurs concurrent with interface trap generation only at high injection levels of $N_{inj} > 1 \times 10^{15}$ electrons[26, 27].

3.2.3 Anomalous ‘Hump’

A consistent occurrence in stressed thin gate devices has been the presence of an anomalous ‘hump’ in the inversion region of the quasi-static C-V curve. The anomalous ‘hump’ has been observed when evaluating device reliability via hot carrier injection[6], Fowler-Nordheim/high-field stressing[7], and radiation exposure[8] and has been attributed to so-called ‘interface’ state generation. At higher stressing conditions, the anomalous ‘hump’ becomes more observable.

Since the anomalous ‘hump’ becomes more evident at higher stressing conditions and NIOT generation occurs only at high injection levels, it is suspected that NIOT generation is accountable for its origin, rather than interface trap generation.

3.3 Evidence of Tunneling

In an attempt to understand the origin of the anomalous ‘hump’, ONO devices have been fabricated with a NIOT-rich nitride layer located 12Å from the Si-SiO₂ interface as determined by ellipsometry and charge pumping experiments[28]. The
Figure 3.2: Using variable frequency charge pumping, it is demonstrated NIOT generation simultaneously occurs with interface trap generation at high hot carrier injection levels. The sharp increase in charge recombined per cycle at lower frequencies is attributed to trap-to-trap tunneling of carriers to NIOTs.
CHAPTER 3. LOW FREQUENCY C-V MODEL

quasi-static C-V data obtained for the NIOT-rich ONO device appears remarkably similar to the highly stressed devices at which NIOT generation has occurred (Fig. 3.3). As the NIOT-rich nitride layer is placed farther from the interface, the anomalous 'hump' becomes less evident. This suggests a communication between interface and NIOTs via a tunneling process, since the tunneling probability is an exponential function of distance.

These preliminary experiments have been performed to suggest strongly NIOTs rather than so-called 'interface' traps are responsible for the anomalous 'hump' observed in quasi-static C-V measurements on gated-diode MOS structures. In addition, evidence of a tunneling mechanism provides a basis for the low frequency C-V model based on trap-to-trap tunneling from interface traps to monoenergetic NIOTs[27] developed in the following section.

3.4 Low Frequency C-V Model

A low frequency C-V model is presented to interpret the effect of near-interface oxide traps which communicate with the Si-SiO₂ interface and influence the C-V characteristics. The model assumes (1) a continuum of interface states, (2) acceptor type monoenergetic NIOTs, and (3) communication of free carriers at the Si-SiO₂ interface with NIOTs via a two-step process[29] involving (a) capture of free carriers by SRH recombination[30, 31] followed by (b) a trap-to-trap tunneling mechanism[32]. The two-step process is illustrated by Fig. 3.4.

3.4.1 Continuum of Interface States

The continuum of interface states has been modeled as constant-density, uniformly distributed energy states across the bandgap of the semiconductor[33]. Surface potential fluctuations associated with random surface charge have been introduced with a normal probability distribution function \( P(U_s) \)[4]. The capacitance
CHAPTER 3. LOW FREQUENCY C-V MODEL

Figure 3.3: A MOSFET before and after substrate hot carrier injection of $1 \times 10^{16}$ electrons/cm$^2$. Similar results obtained for a NIOT-rich ONO device suggest NIOTs are responsible for the anomalous 'hump', rather than interface traps.
CHAPTER 3. LOW FREQUENCY C-V MODEL

Figure 3.4: Energy band diagrams of the two-step process involving capture of free carriers by SRH recombination followed by a trap-to-trap tunneling mechanism. (i) When the AC measurement signal is positive, an electron is captured from the conduction band and tunnels to the monoenergetic near-interface oxide trap. (ii) The electron is able to back tunnel and then, unlike charge pumping, emit to the conduction band when the AC signal is negative.
CHAPTER 3. LOW FREQUENCY C-V MODEL

due to interface traps is expressed as

\[ C_{it} = qD_{it} \int_{-\infty}^{\infty} \frac{\arctan(2\pi f \tau_p)}{(2\pi f \tau_p)} P(U_s) dU_s \] (3.1)

and

\[ P(U_s) = \frac{(2\pi \sigma_s^2)^{-1/2}}{\exp\left[-\frac{(U_s - U_s)^2}{2\sigma_s^2}\right]} \] (3.2)

where \( \tau_p = \tau_{p0} e^{U_s - U_F} \) is the majority carrier capture time constant for a p-type substrate, \( \tau_{p0} = (\sigma_p v_{th} n_i)^{-1} \) is the intrinsic majority carrier time constant, \( U_s \) and \( \bar{U}_s \) are the normalized and mean-normalized surface potential respectively, and \( \sigma_s^2 \) is the normalized variance \((kT/q)\) of the distribution. \( \bar{D}_{it} \) is the average density of interface traps, and \( f \) is the AC measurement frequency.

In order to determine the interface trapped charge, \( Q_{it} \), a uniform interface trap density \( \bar{D}_{it} \) of acceptor traps in the upper half of the energy-gap and donor traps in the lower half of the energy-gap has been assumed[34].

\[ Q_{it} = \frac{q}{C_{ox}} [kT(U_s - U_F)\bar{D}_{it}] \] (3.3)

When normalized surface potential, \( U_s \), equals normalized Fermi potential, \( U_F \), or midgap, the net charge due to interface traps is zero since all donor and acceptor traps are in neutral charge states.

3.4.2 Acceptor Type Monoenergetic NIOTs

The anomalous ‘hump’ has been observed typically in the upper half of the bandgap by previous investigators[6, 7, 8] as well as our own studies[27]. In our work, the experimental C-V curve has been observed to shift to the right of the ideal C-V curve. Thus, the NIOTs have been modeled as a monoenergetic, spatially distributed acceptor-like trap. An acceptor trap can reside in either neutral or negative charge states such that

\[ f^0 + f^- = 1 \] (3.4)
CHAPTER 3. LOW FREQUENCY C-V MODEL

where $f^o$ and $f^-$ are the probabilities of being in neutral and negative charge states respectively. The probability of an acceptor trap being in its $f^-$ state is

$$f^- = \frac{1}{1 + e^{U_F - U_t - U_\text{S}}}, \quad (3.5)$$

The total charge contributed by NIOTs is, therefore,

$$Q_{\text{NIOT}} = -qN_{\text{NIOT}}(f)f^- = -qN_{\text{NIOT}}(f)\frac{1}{1 + e^{U_F - U_t - U_\text{S}}}, \quad (3.6)$$

where $N_{\text{NIOT}}(f)$ (traps/cm$^2$) is the frequency dependent, areal density of NIOTs at energy $E_{\text{tr}}$. $U_F$ and $U_\text{S}$ are normalized Fermi and surface potentials referenced to the bulk intrinsic level respectively. The normalized trap potential $U_t = (E_{\text{tr}} - E_{\text{i}})/kT$ is referenced to the intrinsic level at the surface such that decreasing $U_t$ places the trap energy closer to the conduction band edge. The charge is differentiated with respect to normalized surface potential to obtain an expression for the capacitance

$$C_{\text{NIOT}} = -\frac{q}{kT} \cdot \frac{\partial Q_{\text{NIOT}}}{\partial U_\text{S}} \quad (3.7)$$

Substitution of (3.6) into (3.7) yields an expression for the NIOT capacitance

$$C_{\text{NIOT}} = \frac{q^2N_{\text{NIOT}}(f)e^{U_F - U_t - U_\text{S}}}{kT[1 + e^{U_F - U_t - U_\text{S}}]^2} \quad (3.8)$$

3.4.3 Communication of Interface Traps with NIOTs

When an interface state traps an electron, it is possible for the charge to communicate with, or tunnel to, a NIOT provided a state is available and enough time is allowed for the transition to occur. The filling of the NIOTs introduce additional capacitance several $kT$ about the trap energy, thus influencing the C-V characteristics by introducing a 'hump'[5]. As the measurement frequency is decreased, for a given trap level, electrons can tunnel further into the dielectric and fill more traps, thus, adding to the amplitude of the 'hump'. The trap-to-trap tunneling time constant used to model the communication of interface traps and NIOTs is[32, 35]

$$\tau_T(E, x) = \frac{m^*_1 x (1 + \frac{1}{\alpha_1 x})}{\pi^2 \hbar^3 \alpha_2 D_{it}} e^{\alpha_1 x} \approx \tau_0 e^{\alpha_1 x} \quad (3.9)$$
CHAPTER 3. LOW FREQUENCY C-V MODEL

where $\alpha_1$ and $\alpha_2$ are the attenuation coefficients in SiO$_2$ and Si, respectively, $m^*_1$ is the effective mass in the oxide, and $\tau_o$ is assumed to have minimal spatial dependence. The attenuation coefficients $\alpha_1$ and $\alpha_2$ are expressed as follows

$$\alpha_1^2 = \frac{8m^*_1}{h^2}(\phi_B + E_{cs} - E_{ts})$$  \hspace{1cm} (3.10)

and

$$\alpha_2^2 = \frac{8m^*_2}{h^2}(E_{cs} - E_{ts})$$  \hspace{1cm} (3.11)

where

$$\phi_B = \frac{\chi_s - \chi_o}{q}$$  \hspace{1cm} (3.12)

is the semiconductor and oxide conduction band barrier height (Fig. 3.4). $\chi_s$ and $\chi_o$ are the electron affinities of the semiconductor and oxide respectively, $m^*_1$ is the effective mass in the semiconductor, and $E_{cs}$ and $E_{ts}$ are conduction and NIOT energies at the surface respectively. Theoretical calculations indicate the energy of the NIOTs has minimal dependence on the trap-to-trap tunneling time when close ($< 20 \text{Å}$) to the interface (Fig. 3.5).

As $D_i$ increases, a larger ‘hump’ is observed due to the decreased trap-to-trap tunneling time constant. The observation of an anomalous ‘hump’ depends both on the aforementioned tunneling time constant, $\tau_T$, and the SRH emission time, $\tau_e = \tau_{no}e^{Ut}$ of the trapping center, where $\tau_{no} = (\sigma_n\nu_hn_i)^{-1}$ is the emission time associated with a trapping center at the center of the energy gap. If the reciprocal of the measurement frequency is considerably larger than the tunneling time constant, which in turn is larger than the emission time ($1/\tau_T > \tau_e$), then the conditions are favorable for the two-step process and the presence of an anomalous ‘hump’ in the low frequency C-V characteristics.
Figure 3.5: A plot of trap-to-trap tunneling time vs. tunneling distance for various trap energies. The energy of the trap has minimal effect on the trap-to-trap tunneling time when spatially located within 20Å of the interface.
3.5 Theoretical Results

The capacitance contribution from interface and near-interface oxide traps is additive. The equivalent normalized, low frequency capacitance is expressed as

$$\frac{C}{C_{ox}} = [1 + \left( \frac{C_{ox}}{C_s + C_{it} + C_{NIOT}} \right)]^{-1} \quad (3.13)$$

The semiconductor charge can be derived with a straightforward analysis beginning with Poisson's equation and applying Gauss' law at the Si-SiO_2 interface[10]. By differentiation of the semiconductor charge with respect to surface potential the semiconductor capacitance, $C_s$, can be obtained. For p-type semiconductor,

$$C_s = \frac{U_s K_s \epsilon_o [1 - e^{-U_s} + e^{-2UF(U_s, U_F)}]}{\sqrt{2} L_D F(U_s, U_F)} \quad (3.14)$$

where $L_D$ is the extrinsic Debye length and

$$F(U_s, U_F) = \sqrt{e^{-U_s} + U_s - 1 + e^{-2UF(U_s, U_F)}} \quad (3.15)$$

Substituting (3.1), (3.8), and (3.14) into (3.13) an expression is obtained for the overall normalized low frequency capacitance. In addition, utilizing Kirchoff's voltage law to sum voltage drops from gate to bulk yields

$$V_{GB} = V_{FB} + \frac{kT}{q} U_s + \frac{U_s \sqrt{2} K_s \epsilon_o kTF(U_s, U_F)}{L_D q C_{ox}} +$$

$$\frac{q}{C_{ox}} \left[ kT(U_s - U_F) D_{it} + \frac{(1 - \bar{x}/X_{ox}) N_{NIOT}(f)}{1 + e^{U_F - U_s - U_s - 1}} \right] \quad (3.16)$$

where $X_{ox}$ is the oxide thickness, and $\bar{x}$ is the centroid of the trapped charge in the oxide as measured from the SiSiO_3 interface into the oxide. The centroid is defined as

$$\bar{x} = \frac{\int_{0}^{X_{ox}} dx N_T(x) x}{\int_{0}^{X_{ox}} dx N_T(x)} \quad (3.17)$$

where $N_T(x)$ is the volume density of NIOTs (traps/cm^3). The weighting factor of $\bar{x}/X_{ox}$ has been introduced to accurately express the voltage drop due to NIOTs and becomes increasingly important as $X_{ox}$ scales.
CHAPTER 3. LOW FREQUENCY C-V MODEL

Computer simulations of (3.13) and (3.16) demonstrate the inclusion of trap-to-trap tunneling from interface traps to monoenergetic NIOTs indeed introduces a frequency dependent anomalous 'hump' in the C-V curve (Fig. 3.6). In addition, the curve will shift to the right by \( q_{\text{NIOT}}(f)/C_\infty \) once the acceptor type NIOTs trap electrons and change charge states from neutral to negative. The variations observed in the depletion region of the C-V curve is due to the frequency dependence of the interface traps\([33, 9]\) as indicated by, Equation 3.1.

3.6 Donor and Amphoteric NIOTs

The NIOTs have been modeled as acceptor type since the experimental C-V curve shifts right of theoretical and the anomalous 'hump' is typically observed in the upper half in the bandgap. However, the low frequency C-V model can easily be extended for donor and amphoteric type NIOTs.

3.6.1 Donor Type

Similar to the acceptor type trap, a donor trap has two charge states. Upon trapping an electron, a donor trap changes from positive to neutral charge states. The occupancy probabilities are written as

\[
f^+ + f^0 = 1 \tag{3.18}
\]

where \( f^+ \) is the probability of the trap residing in its positive charge state. The charge due to donor type NIOTs is

\[
Q_{\text{NIOT}} = q_{\text{NIOT}}(f)(1 - f^0) = \frac{q_{\text{NIOT}}(f)}{1 + e^{-(U_F - U_i - U_s)}} \tag{3.19}
\]

and

\[
C_{\text{NIOT}} = \frac{q^2_{\text{NIOT}}(f)e^{-(U_F - U_i - U_s)}}{kT[1 + e^{-(U_F - U_i - U_s)}]^2} \tag{3.20}
\]

is found by differentiating \( Q_{\text{NIOT}} \) with respect to surface potential. For donor type NIOTs, the C-V curve starts left of theoretical and approaches ideal as donor traps change charge states from positive to neutral (Fig. 3.7).
CHAPTER 3. LOW FREQUENCY C-V MODEL

Figure 3.6: Theoretical simulations demonstrate communication from a continuum of interface states to monoenergetic acceptor type NIOTs via trap to trap tunneling introduces a frequency dependent anomalous 'hump' in the inversion regime of the C-V curve.
CHAPTER 3. LOW FREQUENCY C-V MODEL

Figure 3.7: Theoretical simulations of a continuum of interface states to spatially located monoenergetic donor type NIOTs via trap to trap tunneling introduces a frequency dependent anomalous 'hump' in the depletion regime of the C-V curve. The curve begins left of ideal and shifts right once the donor traps capture an electron and become neutral.

3.6.2 Amphoteric Type

Amphoteric traps are essentially a combination of both acceptor and donor type traps. With two discrete energy levels, \( E_{tA} \) and \( E_{tD} \) an amphoteric trap can change charge states twice.

\[
f^+ + f'^+ + f^- = 1 \tag{3.21}
\]

For the amphoteric trap, the following expressions are written[36]:

\[
\frac{f^-}{f^- + f'^+} = \frac{1}{1 + \frac{2e(E_{tA} - E_F)}{kT}} \tag{3.22}
\]

and

\[
\frac{f'^+}{f'^+ + f^+} = \frac{1}{1 + \frac{2e(E_{tD} - E_F)}{kT}} \tag{3.23}
\]
Solving for $f^+$ and $f^-$ yields

$$f^+ = \frac{\frac{1}{2} e V - U_s + U_{tD}}{1 + \frac{1}{2} e V - U_s - U_{tD} + \frac{1}{2} e U_s - U_P + U_{tA}}$$  \hspace{1cm} (3.24)$$

$$f^- = \frac{\frac{1}{2} e U_s - U_P + U_{tA}}{1 + \frac{1}{2} e U_s - U_P - U_{tD} + \frac{1}{2} e U_s - U_P + U_{tA}}$$  \hspace{1cm} (3.25)$$

The following definitions are made:

$$\Delta U = U_{tA} - U_{tD}$$  \hspace{1cm} (3.26)$$

and

$$U_o = \frac{U_{tA} + U_{tD}}{2}$$  \hspace{1cm} (3.27)$$

thus, the charge due to NIOTs is expressed as:

$$Q_{NIOT} = f^+ - f^- = q N_{NIOT}(f) e^{\Delta U/2} \frac{\sinh(U_F - U_s - U_o)}{1 + e^{\Delta U/2} \cosh(U_F - U_s - U_o)}$$  \hspace{1cm} (3.28)$$

For $\Delta U/2 \gg 1$, (3.28) simplifies to

$$q N_{NIOT}(f) \tanh(U_F - U_s - U_o)$$  \hspace{1cm} (3.29)$$

and the capacitance is obtained by differentiating (3.29) with respect to normalized surface potential.

$$C_{NIOT} = \frac{q^2 N_{NIOT}(f)}{kT} \left[ \cosh^2(U_F - U_s - U_o) \right]^{-1}$$  \hspace{1cm} (3.30)$$

An amphoteric trap introduces two anomalous 'humps' in the C-V curve. Initially, the C-V curve is left of ideal and shifts to ideal once the Fermi energy approaches the donor energy $E_{tD}$. The C-V curve will shift right as the Fermi energy approaches $E_{tA}$.  

33
CHAPTER 3. LOW FREQUENCY C-V MODEL

Figure 3.8: Theoretical simulations demonstrate communication from a continuum of interface states to spatially located monoenergetic amphoteric type NIOTs via trap to trap tunneling introduces two frequency dependent anomalous 'humps'.
Chapter 4
Characterization of Near-Interface Oxide Traps

4.1 Characterization of NIOTs

A set of experiments has been performed to verify the communication of interface and NIOTs, via trap-to-trap tunneling, is indeed accountable for the anomalous ‘hump’. With use of a NIOT-rich ONO device (SONOS), the spatial location of the NIOTs are known and can be utilized to experimentally justify the proposed tunneling model. Based on the model, an experimental procedure requiring only C-V techniques permits effective areal density, spatial distribution, and energetic characterization of NIOTs. The variable frequency Hi-Lo C-V extraction technique has been applied to a NIOT-rich ONO memory transistor and compared with the well established charge pumping technique[24] to test its validity.

4.2 Device Fabrication

N-channel SONOS transistors have been fabricated with a NIOT-rich nitride layer located 12Å from the Si-SiO₂ interface (Fig. 4.1) as determined by ellipsometry and charge pumping experiments[28, 24]. This is accomplished by immediately loading the wafer into the chemical vapor deposition (CVD) reactor after cleaning. As a result, an ultra-thin native oxide of 12Å is grown between the semiconductor and deposited silicon nitride film. Since the native oxide is grown at the reactor temperature of 725°C, a high interface trap density of \( D_{it} = 2 \times 10^{11} \text{ cm}^{-2}\text{eV}^{-1} \) is measured. This, however, is beneficial with regards to enhanced trap-to-trap tunneling.
CHAPTER 4. CHARACTERIZATION OF NEAR-INTERFACE OXIDE TRAPS

of carriers. Experimental low frequency C-V and linear voltage ramp (quasi-static) measurements have been performed on the NIOT-rich ONO transistor and a frequency dependent anomalous 'hump' is observed (Fig. 4.2). The 'hump' size increases as frequency is decreased since charge can tunnel further into the oxide and fill more NIOTs.

Figure 4.1: A cross section of the ONO transistor fabricated with a NIOT-rich nitride layer located 12Å from the interface.
Figure 4.2: Experimental low frequency C-V curves for a NIOT-rich ONO transistor. The NIOTs are located 12Å away from the Si-SiO₂ interface. The hump increases as frequency is decreased since charge can tunnel further into the oxide and fill more NIOTs.
4.3 Effective Areal Density

By differentiation of (4.2) and equating to zero, an expression is obtained for the frequency dependent maximum NIOT capacitance, $C_{NIOT_{max}}(f)$.

\[
\frac{\partial C_{NIOT}}{\partial U_s} = \frac{q^2 N_{NIOT}}{kT} \cdot \frac{\partial}{\partial U_s} \frac{e^{U_p - U_t - U_s}}{[1 + e^{U_p - U_t - U_s}]^2} = 0
\]

(4.1)

From this expression, an effective areal density of interface and NIOTs (traps/cm$^2$) is obtained

\[
N_{TOT} = \frac{4kT}{A_G q^2} C_{NIOT_{max}}(f) + \frac{C_{it}}{A_G q}
\]

(4.2)

at a normalized trap potential

\[
U_t = U_p - U_s
\]

(4.3)

where $A_G$ is the device area. As the measurement frequency is lowered, the maximum capacitance is increased since electrons have more time to tunnel deeper into the oxide and fill additional traps.

The Hi-Lo C-V technique\[9\] has been used to obtain an experimental effective trap capacitance, $C_{TOT}$, a culmination of both interface and NIOT capacitance. By subtracting the u-shaped interface trap distribution from the superimposed hump attributed to NIOTs, a value for $C_{NIOT_{max}}(f)$ is obtained. Keeping the high frequency constant while varying the low frequency, the peak capacitance of the hump is used to extract areal densities at different frequencies.

Application of the Hi-Lo technique on a NIOT-rich ONO transistor at various frequencies is illustrated by Fig. 4.3. The measurements were taken by varying the low frequency from 750Hz-300kHz and keeping the high frequency constant at 1MHz.

Applying (4.2) to obtain an effective areal density at each frequency, a sharp increase, or breakpoint, is observed at approximately 175kHz (Fig. 4.4). At the breakpoint frequency electrons are able to tunnel 12-13Å into the oxide as determined from trap-to-trap tunneling theory (Fig. 4.5), which is consistent with the
CHAPTER 4. CHARACTERIZATION OF NEAR-INTERFACE OXIDE TRAPS

Figure 4.3: Utilizing the Hi-Lo C-V technique, an effective areal density of both interface traps and NIOTs may be obtained. The u-shaped distribution is due to interface traps and the superimposed 'hump' is due to NIOTs. As the measurement frequency is decreased, more NIOTs can fill and the peak value $C_{TOT_{\text{max}}}(f)$ increases.
CHAPTER 4. CHARACTERIZATION OF NEAR-INTERFACE OXIDE TRAPS

location of the NIOT-rich nitride layer as determined by ellipsometry and charge pumping. This suggests strongly that NIOTs rather than interface states are accountable for the hump, since the additional capacitance is observed only when charge has sufficient time to tunnel to NIOTs at least 12Å away.

When compared with charge pumping, the effective areal densities obtained at low frequencies (large tunneling distances) are consistent (Fig. 4.6). A drop off is noticed at high measurement frequencies (small tunneling distances) which suggests a non-equilibrium condition exists as the trap-to-trap tunneling time approaches the emission time. The constant offset is due to the difference in interface trap response between both techniques (ie. interface traps respond at 1MHz with C-V).

4.4 Spatial Distribution

In a manner analogous to the bi-level charge pumping treatment, the volume density of monoenergetic NIOTs (traps/cm³) is related to the areal density of oxide traps by

\[ N_{NIOT} = \int_{x_{\text{min}}}^{x_{\text{ox}}} dx N_T(x) [1 - e^{-\frac{x}{10\tau}}] \]  

as derived in Appendix A where \( x_{\text{ox}} \) is the oxide thickness and \( x_{\text{min}} \) is the minimum distance a NIOT can be distinguished from an interface trap[24]. In this instance, \( f \) is a the small-signal AC frequency associated with the differential C-V measurement.

With the use of (3.9) and (4.4) a maximum tunneling distance, \( x_m(f) \), can be defined for a particular measurement frequency [37, 28]

\[ x_m(f) = \frac{1}{\alpha_1} \ln \left( \frac{1}{f\tau_o} \right) \]  

where \( \tau_o \) and \( \alpha_1 \) are trap-to-trap tunneling parameters from (3.9). Therefore, all NIOTs located closer than \( x_m(f) \) are filled and those further than \( x_m(f) \) are empty. Thus, we may simplify (4.4) to

\[ N_{NIOT} = \int_{x_{\text{min}}}^{x_m(f)} dx N_T(x) \]  

40
CHAPTER 4. CHARACTERIZATION OF NEAR-INTERFACE OXIDE TRAPS

Figure 4.4: Variable frequency Hi-Lo C-V measurements have been taken on a trap-rich ONO transistor with $D_{it}=2\times10^{11}$ (cm$^{-2}$eV$^{-1}$) and a breakpoint at approximately 175kHz is observed. For frequencies below the breakpoint, charge has sufficient time to tunnel to the NIOTs placed 12Å away.
Figure 4.5: Plotting (4.5), the breakpoint in Fig. 4.4 corresponds to a tunneling distance of 12-13 Å, consistent with the location of the NIOT-rich nitride layer as determined by ellipsometry. A higher interface trap density increases the distance charge can tunnel at a particular frequency.
Figure 4.6: A comparison of extracted areal densities between variable frequency Hi-Lo C-V and charge pumping. Close agreement between extracted trap densities from charge pumping and C-V measurements is observed. The constant offset is due to the difference in interface trap response between both techniques and the drop-off is due to non-equilibrium as the trap-to-trap tunneling time approaches the emission time.
By differentiating (4.6) with respect to \( \log(f) \), the density of NIOTs at the maximum tunneling distance \( N_T(x_m) \) is obtained

\[
N_T(x_m) = \frac{-\alpha_1}{2.3} \frac{\partial N_{NIOT}}{\partial \log f}
\]  

(4.7)

With (4.7) and (4.5), a profile of trapped charge due solely to NIOTs can be obtained. These equations have been applied to the trap-rich ONO transistor to obtain a spatial distribution of NIOTs. Excellent agreement between variable frequency Hi-Lo C-V and charge pumping is evident (Fig. 4.7) provided equilibrium is maintained \((\tau_T > \tau_e)\). Theoretical calculations indicate electrons must tunnel approximately 16-17 Å into the oxide to satisfy the equilibrium requirement for the NIOT-rich ONO device. Experimentally, the onset of convergence between Hi-Lo C-V and charge pumping is observed at a maximum tunneling distance of 17 Å. At this distance, tunneling limited trap emission allows equilibrium to be maintained.

### 4.5 Energy

The energy of the NIOTs can be determined using (11) since \( U_F = \ln(N_B/n_i) \) is fixed and \( U_e \) can be evaluated by applying the Berglund technique.

\[
U_e = U_F - \frac{q}{kT} \left[ \phi_{so} + \int_{V_{GB}}^{V_{GB}} dV_{GB} \left( 1 - \frac{C_{LF}}{C_{OX}} \right) \right]
\]

(4.8)

By determining the normalized surface potential at \( C_{NIOT_{max}}(f) \), it has been determined that the NIOTs are energetically located approximately 0.275 eV below the conduction band, which agrees with reported nitride characteristics[38].

### 4.6 Summary

A summary of the variable frequency Hi-Lo C-V extraction technique has been illustrated by Figure 4.8. Aside from Hi-Lo and Berglund C-V techniques, no additional measurements are required in determining the areal density, spatial distribution, and energy of NIOTs. However, trap-to-trap tunneling parameters \( \tau_0 \) and \( \alpha_1 \)
CHAPTER 4. CHARACTERIZATION OF NEAR-INTERFACE OXIDE TRAPS

Figure 4.7: By differentiating the effective areal density with respect to $\log(f)$, a volume density of NIOTs is obtained. Excellent convergence between variable frequency Hi-Lo C-V and charge pumping exists provided equilibrium with tunneling electrons is maintained ($\tau_T > \tau_e$).
must be determined theoretically. Computer programs have been developed to aid in the evaluation of these parameters and are included in Appendix B.

\[
N_t = \frac{4C_{t_{\text{max}}}(f)kT}{Aq^2}
\]

effective areal density

\[
N_{\text{vol}} = -\frac{\alpha_t \partial N_t}{2.3 \partial \log f}
\]

volume density

\[
X_m = \frac{\ln(1/ft_0)}{\alpha_1}
\]

maximum tunneling distance

Figure 4.8: A summary of the variable frequency Hi-Lo C-V extraction technique. With this technique, areal density, spatial distribution, and energy characterization of NIOTs is possible.
Chapter 5

Conclusions

5.1 Accomplishments

The anomalous hump observed in stressed devices has been attributed to NIOTs rather than interface traps. A new low frequency C-V model based on trap-to-trap tunneling from a continuum of interface states to monoenergetic NIOTs has been developed. From the model, an extraction procedure based on variable frequency Hi-Lo C-V measurements permits spatial distribution and energetic characterization of NIOTs.

The variable frequency Hi-Lo C-V technique has been used to extract trap densities for a trap-rich ONO transistor. By knowing the distance of NIOTs from the semiconductor-dielectric interface, correlation between a breakpoint frequency and the location of the NIOTs demonstrates trap-to-trap tunneling from interface traps to NIOTs is indeed occurring. The collected data displays excellent agreement with the charge pumping technique indicating the variable frequency Hi-Lo C-V extraction technique provides an effective means of evaluating device reliability by monitoring interface and NIOT generation. The technique will prove useful in evaluating hot carrier degradation, high field/Fowler-Nordheim stressing, radiation damage, and nonvolatile memory films in thin gate devices.

5.2 Applications

The ability to characterize NIOTs with simple C-V measurement techniques is highly promising for industrial use. The C-V measurement is widely used to determine other device parameters, and familiarity with the measurement is ubiquitous.
CHAPTER 5. CONCLUSIONS

The capability of testing capacitors dramatically reduces the process development time to evaluate gate dielectric reliability. Unlike the charge pumping technique, a MOS capacitor with implanted minority carrier grid can be used to characterize NIOTs. The variable frequency Hi-Lo C-V technique permits NIOT characterization throughout the entire bandgap, whereas, other techniques such as conductance[39] measurements are restricted to certain energies.

The introduction of processing induced defects can alter predicted device performance considerably. Obtaining the spatial distribution and energy of NIOTs at various stages of fabrication enables process monitoring by identifying and quantifying contaminants. The variable frequency Hi-Lo C-V technique is suitable as a process monitoring tool due to its accuracy and simplicity.

For a device to be reliable, its performance must not alter the intended circuit operation over the expected lifetime. By monitoring NIOT generation with the variable frequency Hi-Lo C-V technique, lifetime expectancies can be estimated through appropriate stressing conditions. Another salient feature of the technique is the capability to quickly characterize new, promising materials, such as wide bandgap semiconductors, where charge trapping effects near the interface are critical to device performance.

5.3 Recommendation for Future Work

A limitation of the Hi-Lo C-V technique is the difficulty of maintaining equilibrium at small tunneling distances as the trap-to-trap tunneling time decreases and is comparable to the emission time. A plausible solution to this problem is to reduce the emission time by means of illumination. Preliminary results appear promising since a larger anomalous 'hump' is observed at a given frequency (Figure 5.1). Another possibility is heating the sample to reduce the emission time since tunneling probability is independent of temperature.
Figure 5.1: Low frequency C-V curves of the 12Å trap-rich ONO device under illumination. A larger anomalous 'hump' is observed due to the reduction in emission time.
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Appendix A

Solution to Equation 4.4

The volume density of occupied monoenergetic NIOTs (cm\(^{-3}\)eV\(^{-1}\)) is expressed as follows

\[
\rho_{\text{NIOT}}(E_t, x, t) = \rho_{\text{NIOT}}(E_t, x) e^{-\frac{x}{\tau_T}}
\]  \hspace{1cm} (A.1)

where \(\rho_{\text{NIOT}}(E_t, x)\) is the volume density of oxide traps (cm\(^{-3}\)eV\(^{-1}\)). The areal density of filled oxide traps \(N_{\text{NIOT}}^{-}\) is obtained by integrating the occupied volume density over the oxide thickness and energy bandgap.

\[
N_{\text{NIOT}}^{-} = \int_{x_{\text{min}}}^{x_{\text{max}}} dx \int_{E_{\text{v}}}^{E_{\text{c}}} dE_t \rho_{\text{NIOT}}(E_t, x) e^{-\frac{x}{\tau_T}}
\]  \hspace{1cm} (A.2)

The near-interface oxide trap current, \(I_{\text{NIOT}}\), is determined by differentiation of \(N_{\text{NIOT}}^{-}\) with respect to time.

\[
I_{\text{NIOT}} = -qA_G \frac{\partial N_{\text{NIOT}}^{-}}{\partial t}
\]  \hspace{1cm} (A.3)

or

\[
I_{\text{NIOT}} = qA_G \int_{x_{\text{min}}}^{x_{\text{max}}} dx \int_{E_{\text{v}}}^{E_{\text{c}}} dE_t \rho_{\text{NIOT}}(E_t, x) e^{-\frac{x}{\tau_T}} \frac{1}{\tau_T(E_t, x)}
\]  \hspace{1cm} (A.4)

Averaging over the period of the applied AC waveform, the average NIOT charge per cycle is

\[
\overline{Q_{\text{NIOT}}} = \int_0^T \frac{dt}{T} \frac{I_{\text{NIOT}}}{T} = qN_{\text{NIOT}}A_G
\]  \hspace{1cm} (A.5)

Therefore,

\[
N_{\text{NIOT}} = \int_0^T dt \frac{e^{-\frac{x}{\tau_T(E_t, x)}}}{\tau_T(E_t, x)} \int_{x_{\text{min}}}^{x_{\text{max}}} dx \int_{E_{\text{v}}}^{E_{\text{c}}} dE_t \rho_{\text{NIOT}}(E_t, x)
\]  \hspace{1cm} (A.6)

Performing the integration over time,

\[
N_{\text{NIOT}} = \int_{x_{\text{min}}}^{x_{\text{max}}} dx \int_{E_{\text{v}}}^{E_{\text{c}}} dE_t (1 - e^{-\frac{x}{\tau_T(E_t, x)}}) \rho_{\text{NIOT}}(E_t, x)
\]  \hspace{1cm} (A.7)
APPENDIX A. SOLUTION TO EQUATION 4.440

\[ N_{NIOT} = \int_{x_{\text{min}}}^{x_{\text{max}}} dx \int_{E_{t^*}}^{E_{t^*}} dE_t (1 - e^{-\frac{1}{\tau_T(E_t, x)}}) \rho_{NIOT}(E_t, x) \]  \hspace{1cm} (A.8)

Since the trap energy has minimal influence on trap-to-trap tunneling time at small tunneling distances, \( \tau_T \) is assumed to be energy independent.

\[ \tau_T(E_t, x) \approx \tau_T(x) \]  \hspace{1cm} (A.9)

A volume density of NIOTs, \( N_T(x) \) (traps/cm\(^3\)), is defined

\[ N_T(x) = \int_{E_{t^*}}^{E_{t^*}} dE_t \rho_{NIOT}(E_t, x) \]  \hspace{1cm} (A.10)

By substitution of (A.10) into (A.8), a final expression is obtained relating the areal density of NIOTs to the volume density.

\[ N_{NIOT} = \int_{x_{\text{min}}}^{x_{\text{max}}} dx N_T(x) [1 - e^{-\frac{1}{\tau_T(x)}}] \]  \hspace{1cm} (A.11)
Appendix B

Computer Simulations

/* The following program will generate a theoretical C-V curve */
/* incorporating trap-to-trap tunneling from interface to a */
/* monoenergetic spatially located acceptor type NIOT. */

#include <stdio.h>
#include <stdlib.h>
#include <ctype.h>
#include <math.h>

FILE *iofile;
main() {
    double q,k,t,pi,eo,ks,kox,tx,ni,plow,phigh,pstep,pf,ec,ei,ut,fm,
    omega,not,tmo,cox,ld,uf,x,mo,m1,m2,pb,dit,mb,al,a2,tt,xm,nt,ps,
    us,f,qniot,v,k1,cs,tm,sliga,lo,hi,h,ux,pu,csitns,temp,cit,cniot,cnorm;
    int i,numit,j,iter;
    char *cvdata;

    cvdata="data";
    q=1.60e-19;
    k=1.38e-23;
    t=293;
    pi=acos(-1.0);
    eo=8.854e-12;
APPENDIX B. COMPUTER SIMULATIONS

ks=11.7;
kox=3.9;
tox=100e-10;
nb=8e21;
ni=1.45e16;
plow=-.3;
phigh=.95;
pstep=0.025;

pf=((k*t)/q)*log(nb/ni);
ec=1.1*1.6e-19;
ei=.55*1.6e-19;

ut=11;
fm=1000;
omega=2*pi*fm;
not=2e24;
tmo=.01;

cox=(kox*eo)/tox;
ld=sqrt((k*t*ks*eo)/(q*q*nb));
uf=(q*pf)/(k*t);
numit=(phigh-plow)/pstep;

x=12e-10;
mo=9.1e-31;
m1=.5*mo;
m2=.98*mo;
pb=3.1*1.6e-19;
dit=3e15/1.6e-19;
hb=1.06e-34;
APPENDIX B. COMPUTER SIMULATIONS

\[ a_1 = 2\sqrt{2m_1(pb + ec - (ei - utk^t))/(hbhb)}; \]
\[ a_2 = 2\sqrt{2m_2(ec - (ei - utk^t))/(hbhb)}; \]
\[ t_0 = (m_1m_1x(1 + 1/(a_1x)))/(pi^2a_2hbhb^2dit); \]
\[ t_t = t_0\exp(a_1x); \]
\[ xm = (1/a_1)\log(1/(fm^2t_0)); \]
if \(xm < x\) \(xm = x;\) /* no NIOTs exist before trapping layer*/
\[ nt = not\*(xm-x); \]
\[ iofile = fopen(cvdata,"w"); \]
/* sweep surface potential to determine Vgs*/
for \((i=0;i<=numit;i++)\{
\[ ps = plow + i*pstep; \]
\[ us = (q*ps)/(k^t); \]
\[ f = \sqrt{\exp(-us) + us - 1 + \exp(-2uf)(\exp(us) - us - 1)}; \]
\[ qniot = (-q*nt)/(1.0 + \exp(uf - ut - us)); \]
\[ v = ps + (\sqrt{2.0}, ks*eo*us*k^t*f)/(ld*cox*fabs(us)*q) - (qniot/cox); \]
\[ k_1 = 1 - \exp(-us) + \exp(-2uf)(\exp(us) - 1); \]
\[ cs = (ks*eo*us*k_1)/(\sqrt{2.0}*ld*fabs(us)*f); \]
\[ tm = tmo*exp(us - uf); \]
\[ sigma = 1.5; \]
\[ iter = 750; \]
\[ lo = -9.07; \]
\[ hi = 34.93; \]
\[ h = (hi - lo)/iter; \]
\[ fx = 0; \]
for \((j=0;j<=iter;j++)\{
\[ pus = (1/\sqrt{2\pi*sigma*sigma})*\exp(-((j-us)(j-us))/(2*sigma*sigma)); \]
\[ citns = (q*dit*1.6e-19*atan(omega*tmo*exp(j - uf)))/(omega*tmo*exp(j - uf)); \]
\[ temp = pus*citns; \]
\]
APPENDIX B. COMPUTER SIMULATIONS

```c
if (j==0 || j==numit)
    fx=fx+temp;
if (j>0 || j<numit)
    fx=fx+2*temp;
}
fx=fx*h/2;
citns=(q*dit*1.6e-19*atan(omega*tmo*exp(us-uf)))/(omega*tmo*exp(us-uf));
cit=citns+fx;
cniot=(q*q*nt*exp(uf-ut-us))/(k*t*(1+exp(uf-ut-us))*(1+exp(uf-ut-us)));

/* cnorm is the normalized capacitance C/Cox */
cnorm=1.0/(1.0+cox/(cs+cit+cniot));

printf("v=%e C/Cox=%e \n",v,cnorm);
fprintf(iofile,"%e %e \n",v,cnorm);
}
fclose(iofile);
}

/* The following program will generate the maximum trap-to-trap tunneling */
/* distance for various frequencies. Trap-to-trap parameters to and al are */
/* also calculated. */

#include <stdio.h>
#include <stdlib.h>
#include <ctype.h>
#include <math.h>

FILE *iofile;
```
APPENDIX B. COMPUTER SIMULATIONS

main()
{
    double q,k,t,eo,ks,kox,tox,nb,ni,plow,phigh,pstep,pf,ec,ei,
        ut,x,cox,ld,uf,f,mo,m1,m2,pb,dit,hb,pi,a1,a2,to,tt,xm;
    int i,numit,j;
    char *cvdata;

    cvdata="data";
    q=1.60e-19;
    k=1.38e-23;
    t=293;
    eo=8.854e-12;
    ks=11.7;
    kox=6.5;
    tox=100e-10;
    nb=6e20;
    ni=1.45e16;
    plow=-0.3;
    phigh=0.9;
    pstep=0.025;
    pf=((k*t)/q)*log(nb/ni);
    ec=1.1*1.6e-19;
    ei=.55*1.6e-19;
    ut=-11;
    x=12e-10;

    cox=(kox*eo)/tox;
    ld=sqrt((k*t*ks*eo)/(q*q*nb));
    uf=(q*pf)/(k*t);
    numit=100;
APPENDIX B. COMPUTER SIMULATIONS

\[ f = 0; \]
\[ m_0 = 9.1 \times 10^{-31}; \]
\[ m_1 = 0.5 \times m_0; \]
\[ m_2 = 0.98 \times m_0; \]
\[ p_b = 3.1 \times 1.6 \times 10^{-19}; \]
\[ \delta t = 2 \times 10^5 / 1.6 \times 10^{-19}; \]
\[ h_b = 1.06 \times 10^{-34}; \]
\[ \pi = 3.141; \]

\[
\text{iofile} = \text{fopen(cvdata, "w");}
\]
/* sweep frequency to determine maximum tunneling distance */
for (i=0; i<numit; i++){
    \[ f = f + 100; \]
    \[ a_1 = 2 \times \sqrt{2 \times m_1 \times (p_b + e_c - (e_i - u_t \times k \times t)) / (h_b \times h_b)}; \]
    \[ a_2 = 2 \times \sqrt{2 \times m_2 \times (e_c - (e_i - u_t \times k \times t)) / (h_b \times h_b)}; \]
    \[ t_0 = (m_1 \times m_1 \times x \times (1 + 1 / (a_1 \times x))) / (\pi \times \pi \times a_2 \times h_b \times h_b \times \delta t); \]
    \[ t_t = t_0 \times \exp(a_1 \times x); \]
    \[ x_m = (1 / a_1) \times \log(1 / (f \times t_o)); \]

    printf("f=\%e x_m=\%e t_o=\%e a_1=\%e\n", f, x_m, t_o, a_1);
    fprintf(iofile, "\%e \%e \%e \%e\n", f, x_m / 1e-10, t_o, a_1);
}
fclose(iofile);
Appendix C

SONOS Nonvolatile Memory Research

C.1 Oxynitride Films

Nonvolatile semiconductor memory devices have the unique ability of retaining a programmed voltage without an external power supply. The majority of today’s nonvolatile applications employ the floating gate technology. An alternative structure, the SONOS (polySilicon-blockingOxide-Nitride-tunnelOxide-Silicon) memory transistor stores charge in the bulk traps of a nitride dielectric and exhibits low power dissipation, small cell size, and good memory retention. In order for the SONOS structure to become widely used, industry must view the SONOS structure as a highly reliable, low power, and cost effective method of obtaining nonvolatility.

Presently, conventional silicon nitride (Si₃N₄) films are used as the charge storing medium in SONOS devices. However, recent focus has shifted to oxynitride (SiOₓ₂Nᵧ) films. The addition of oxygen is believed to improve the trap distribution by passivating dangling silicon bonds and moving the charge centroid farther away from the tunnel oxide[40]. As a result, oxynitride films can be incorporated with existing SONOS technology to improve endurance and data retention.

Our LPCVD system has been successfully modified to flow ammonia gas in the center reactor. A 2:4:1 flow of N₂O:NH₃:SiCl₂H₂ has been used to deposit the oxynitride films for an oxygen concentration of 8%. The variable frequency Hi-Lo C-V technique will prove useful in comparing the oxynitride trapping layer with conventional silicon nitride films. The process sequence for oxynitride SONOS capacitors is described by the following text:
Oxynitride Capacitor Fabrication Sequence:

1. n+ Grid

   (a) Furnace Clean

      • Organic removal bath: 5 minutes in solution of 5:1:1 DI water, ammonium hydroxide, hydrogen peroxide at 75°C.
      • Tweezer bath: 5:1:1 solution of DI water, ammonium hydroxide, hydrogen peroxide; room temperature.
      • rinse: 5 rinses in DI water.
      • Metal removal bath: 5 minutes in 5:1:1 solution of DI water, hydrochloric acid, and hydrogen peroxide at 75°C.
      • rinse: 5 rinses in DI water.
      • HF dip: Add small quantity of HF (3 rinse to etch pad oxide. Rinse in DI water 5 more times.

      • Spin Dry

   (b) 300Å wet pad oxide (950°C, 12min)

   (c) Photo 1 (n+ grid)

   (d) Ion Implantation (Phosphorous, 5E15, 100Kev)

   (e) Plasma Etch (O₂, 5min, 300mT)

   (f) Strip (Photo (PRS-2000)

2. Active Area

   (a) Furnace Clean

   (b) 2800Å oxide (Steam Oxidation, 1100°C, 60 minutes.)

   (c) Photo 2 (FOX)

   (d) Etch (2800Å oxide, Buffered HF 10:1, 12 minutes)

      Check back as indicator, etch until back is hydrophobic.
APPENDIX C. SONOS NONVOLATILE MEMORY RESEARCH

3. Triple Dielectric

(a) Furnace Clean
(b) Tunnel Oxide
   i. 20Å Dry tunnel oxide (TWO, 700°C, 12 min)
   ii. 12Å Native tunnel oxide
(c) Ar anneal (700°C, 30 min.)
(d) 50Å Oxynitride (LPCVD: 2:4:1 flow N\textsubscript{2}O:NH\textsubscript{3}: SiCl\textsubscript{2}H\textsubscript{2} (80,40,20 sccm, 725°C, 8.5 min, 300mT)
(e) 40Å Blocking oxide (LPCVD: 10:1 flow N\textsubscript{2}O:SiCl\textsubscript{2}H\textsubscript{2} (100:10 sccm), 725°C, 15 min, 800mT)
(f) Densify blocking oxide (Wet Oxidation, 900°C, 30 min)

4. Metal

(a) 7000Å Al, sputtered
(b) Photo 3 (Metal)
(c) Etch (7000Å A, PAN etch, 45°C, 2 min)
(d) Strip (PRS-2000)
(e) Plasma Etch backside (CF\textsubscript{4}, 1 min, 300mT)
(f) 7KÅ Al backside, sputtered
(g) Organic Clean (Acetone, Methanol)
(h) Post Metal Anneal (H\textsubscript{2}/N\textsubscript{2}, 400°C, 60 minutes)
C.2 Surface Passivation

Nonvolatile applications such as laptop computers necessitate low power and operating voltages. In order to reduce power consumption in a SONOS device, the ONO dielectric must be slimmed down so a smaller voltage can be used to program the device. To obtain 5V programmability, SONOS tunnel oxides must be less than 20Å thick. However, when silicon is cleaned for furnace preparation, a native oxide of approximately 12Å remains due to the ambient oxygen. The native oxide is of very poor quality since it is grown at room temperature. Therefore, a need exists to develop a process sequence to inhibit the growth of native oxide.

One way of passivating native oxide growth with low metallic contamination is by incorporating hydrofluoric acid (HF) into the standard RCA cleaning procedure[41]. Room temperature studies have been performed using hydrofluoric acid to control undesirable oxide growth. It has been determined an unrinised 3% HF dip will minimize native oxide growth to only 7Å after several hours. However, the use of HF for surface passivation has not yet been studied at oxidation temperatures.
Appendix D
Series Resistance

Although this thesis has dealt primarily with the effects of gate oxide charge on the C-V measurement, series resistance can also introduce significant non-idealities and should be addressed. Series resistance can arise from various sources:

1. Contact made by the probe wire to the gate
2. The back contact to the silicon
3. Dirt or particulate matter between back contact and wafer chuck
4. Extremely nonuniform doping distribution underneath the gate

At a given frequency, series resistance effects become most apparent in the strong accumulation region of the C-V curve. When the semiconductor surface is accumulated with minority carriers the semiconductor capacitance is much larger than the oxide capacitance so the overall effective capacitance is approximately $C_{ox}$. The equivalent circuit is simply $C_{ox}$ in series with resistance $R_s$. A capacitance meter is used to measure the admittance across the gate and substrate terminals and will be expressed as $Y_m = G_m + j\omega C_m$. Since admittances in series act like resistors in parallel,

$$G_m + j\omega C_m = \frac{j\omega C_{ox}}{j\omega C_{ox} R_s + 1}$$  \hspace{1cm} (D.1)

If we multiply by the complex conjugate both the real and imaginary components of the measured admittance can be separated.

$$G_m + j\omega C_m = \frac{\omega^2 C_{ox}^2 R_s}{\omega^2 C_{ox}^2 R_s^2 + 1} + j\omega \frac{C_{ox}}{\omega^2 C_{ox}^2 R_s^2 + 1}$$  \hspace{1cm} (D.2)
APPENDIX D. SERIES RESISTANCE

Defining \( Q \) as the 'quality factor' of the MOS structure,

\[
Q = \frac{1}{\omega R_s C_{ox}} \quad (D.3)
\]

Eq. D.3 can be simplified.

\[
G_m + j \omega C_m = \frac{1}{R_s} \frac{1}{1 + Q^2} + j \omega \frac{C_{ox}}{1 + \frac{1}{Q^2}} \quad (D.4)
\]

From Eq. D.4 it is clear the meter actually measures a capacitance of

\[
C_m = C_{ox}(1 + \frac{1}{Q^2}) \quad (D.5)
\]

This demonstrates that when the surface is accumulated, the measured capacitance will be smaller than the actual capacitance due to series resistance. As the frequency is increased, the \( Q \) decreases and an exact measurement of oxide capacitance is difficult. Series resistance problems, however, can be corrected by measuring both real and imaginary parts. In order to reduce series resistance effects, smaller devices can be used to reduce the sample capacitance and the RC product. In addition, the wafer chuck should be cleaned with acetone before measurements are performed to remove particulate matter and provide a good back substrate contact.
Figure D.1: Series resistance results in a lowering of the C-V curve in the strongly accumulated region. The series resistance problem is more evident at higher measurement frequencies.
Vitae

Neil Laurence Cohen was born February 10, 1971 in Edison, New Jersey to Marcia and Barry Cohen. He attended Lehigh University from September 1989 to May 1993 and obtained a Bachelor of Science in Electrical Engineering with honors. During his undergraduate years, he was awarded a Sherman Fairchild Summer Fellowship for solid-state studies.

Since May 1993 he has been enrolled as a full time graduate student at Lehigh University under a Sherman Fairchild Fellowship. During this time, he has served as a research assistant under numerous contract supported projects. His studies have focused primarily on interface and near-interface trapping characterization in MOSFETs, low-power nonvolatile memory devices, and gate dielectric reliability issues.

He is a student member of the IEEE and Electron Devices Society and has been an officer in Tau Beta Pi and Eta Kappa Nu. He has two publications and conference presentations.
END OF TITLE