HIGH PERFORMANCE FLOATING GATE MEMORIES USING GRAPHENE AS CHARGE STORAGE MEDIUM AND ATOMIC LAYER DEPOSITED HIGH-K DIELECTRIC LAYERS AS TUNNEL BARRIER

A THESIS
SUBMITTED TO THE MATERIALS SCIENCE AND
NANOTECHNOLOGY PROGRAM
OF THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Deniz Kocaay
January, 2013

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Aykutlu Dana (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Coşkun Kocabaş

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Necmi Bıyıklı

Approved for the Graduate School of Engineering and Science:

Prof. Dr. Levent Onural
Director of Graduate School

ABSTRACT

HIGH PERFORMANCE FLOATING GATE MEMORIES USING GRAPHENE AS CHARGE STORAGE MEDIUM AND ATOMIC LAYER DEPOSITED HIGH-K DIELECTRIC LAYERS AS TUNNEL BARRIER

Deniz Kocaay

M.S. in Materials Science and Nanotechnology

Supervisor: Assist. Prof. Dr. Aykutlu Dana

January, 2013

With the ongoing development in portable electronic devices, low power consumption, improved data retention rate and higher operation speed are the merits demanded by modern non-volatile memory technology. Flash memory devices with discrete charge-trapping media are regarded as an alternative solution to conventional floating gate technology. Flash memories utilizing Si-nitride as charge storage media dominate due to enhanced endurance, better scaling capability and simple fabrication. The use of high-k dielectrics as tunnel layer and control layer is also crucial in charge-trap flash memory devices since they allow further scaling and enhanced charge injection without data retention degradation. Atomic layer deposition (ALD) is a powerful technique for the growth of pinhole-free high-k dielectrics with precisely controlled thickness and high conformality. The application of graphene as charge trapping medium in flash memory devices is promising to obtain improved charge storage capability with miniaturization. Graphene acts as an effective charge storage medium due to high density of states in deep energy levels.

In this thesis, we fabricate graphene flash memory devices with ALD-grown HfO_2/AlN as tunnel layer and Al_2O_3 as control layer. Graphene oxide nanosheets are derived from the acid exfoliation of natural graphite by Hummers Method.

The graphene layer is obtained by spin-coating of water soluble graphene oxide suspension followed by a thermal annealing process. Memory performance including hysteresis window, data retention rate and program transient characteristics for both electron and hole storage mechanisms are determined by performing high frequency capacitance-voltage measurements. For comparing the memory effect of graphene on device performance, we also fabricate and characterize identical flash capacitors with Si-rich SiN layer as charge storage medium and HfO_2 as tunnel oxide layer. The Si-nitride films are deposited with high SiH_4/NH_3 gas flow ratio by plasma-enhanced chemical vapor deposition system.

Graphene flash memory devices exhibit superior memory performance. Compared with Si-nitride based cells, hysteresis window, retention performance and programming speed are both significantly enhanced with the use of graphene. For electron storage, graphene flash memory provides a saturated flat band shift of 1.2 V at a write-pulse duration of 100 ns with a voltage bias of 5 V. The high density of states and high work function of graphene improve the memory performance, leading to increased charge storage capability, enhanced retention rate and faster programming operation at low voltages.

The use of graphene as charge storage medium and ALD-grown high-k dielectrics as tunnel and control layers improves the existing flash technology and satisfies the requirements including scalability, at least 10-year retention, low voltage operation, faster write performance and CMOS-compatible fabrication**.**

Keywords: Flash memory, graphene, high-k dielectrics, ALD

ÖZET

YÜK DEPOLAMA ORTAMI OLARAK GRAFEN, TÜNELLEME BARİYERİ OLARAK ATOMİK KATMAN KAPLAMA TEKNİĞİYLE ÜRETİLMİŞ YÜKSEK-K DİELEKTRİK KULLANILARAK OLUŞTURULAN YÜKSEK PERFORMANSLI İKİNCİL KAPILI HAFIZA YAPILARI

Deniz Kocaay

Malzeme Bilimi ve Nanoteknoloji, Yüksek Lisans

Tez Yöneticisi: Yar. Doç. Dr. Aykutlu Dana

Ocak, 2013

Taşınabilir elektronik cihaz teknolojisinde süregelen gelişim ile birlikte düşük güç tüketimi, gelişmiş veri tutma oranı ve daha yüksek çalışma hızı, modern flaş bellek teknolojisi tarafından talep edilen özellikler haline gelmiştir. Ayrık yük yakalama ortamı taşıyan flaş bellek aygıtları, geleneksel ikincil kapılı hafıza teknolojisine alternatif bir çözüm olarak kabul edilmektedir. Yük depolama ortamı olarak Si-nitrür kullanarak geliştirilmiş flaş bellekler dayanıklılık, daha iyi ölçekleme yeteneği ve basit üretimi nedeniyle yaygın olarak kullanılmaktadır. Tünelleme ve kontrol bariyerleri olarak Atomik katman kaplama (ALD) tekniğiyle üretilmiş yüksek-k dielektrikli filmlerin kullanımı, daha küçük yapıların üretimine ve verinin tutulma zamanını etkilemeksizin daha fazla yük tutulumuna elverdiğinden flaş bellek yapıları için çok önemlidir. ALD tekniği tam olarak kontrol kalınlığı ve yüksek konformalite sağlayabildiğinden deliksiz yüksek-k dielektrikli film büyütülmesi için çok güçlü bir tekniktir. Flaş bellek aygıtları için yük yakalama aracı olarak grafen kullanılması, geliştirilmiş yük depolama kapasitesine sahip daha küçük ölçekli aygıtlar elde etmek açısından umut vericidir. Grafen, derin enerji seviyelerinde daha fazla yük tutulabilecek yer bulunması nedeniyle etkili bir yük depolama ortamı olarak görülmektedir.

Bu tez çalışmasında, ALD ile büyütülmüş tünel katmanı olarak HfO_2/AlN ve kontrol katmanı olarak Al_2O_3 malzemeler içeren grafen flaş bellek cihazları imal edilir. Grafen oksit nanotabakalar Hummers yöntemi ile doğal grafitin asit eksfoliyasyonu ile elde edilmektedir. Grafen tabakası suda çözünür grafen oksit süspansiyonunun döndürerek kaplanması ve ısıl tavlama işlemi ile elde edilir. Histerezis penceresi, veri saklama oranı ve elektron/deşik depolama mekanizmalarını içeren bellek performansı yüksek frekansta kapasitans-voltaj ölçümleri yapılarak belirlenir. Grafenin cihaz performansına etkisini karşılaştırmak için, yük saklama ortamı olarak Si-zengin SiN tabaka kullanan özdeş flaş kapasitörler imal edilmiş ve karakterizasyonu yapılmıştır. Si-nitrür filmler yüksek SiH_4/NH_3 gaz akış oranı ile plazma destekli kimyasal buhar biriktirme sisteminde kaplanmıştır.

Grafen flaş bellek cihazları üstün bellek performansı gösterir. Si-nitrür temelli hücrelerle ile karşılaştırıldığında, histerezis pencere, veri tutma performansı ve programlama hızı grafen kullanımı ile önemli ölçüde geliştirilmiştir. Elektron depolama için grafen flaş bellek 5 V gerilimde ve 100 ns darbe süresi ile 1.2 V düz bant kayması sağlar.

Grafenin yük depolama ortamı ve ALD ile büyütülen yüksek-k dielektriklerin tünel ve kontrol katmanları olarak kullanılması mevcut flaş teknolojisini geliştirir ve en az 10 yıllık saklama, düşük voltajda çalışma, hızlı yazma performansı ve CMOS-uyumlu fabrikasyon gibi gereksinimleri karşılar.

Anahtar kelimeler: Flaş bellek, grafen, yüksek-k dielektrikler, ALD

Acknowledgements

I would like to express my sincere gratitude, first and foremost, to my supervisor, Assist. Prof. Dr. Aykutlu Dana for his outstanding support, encouragement and guidance throughout my research. I have learned a lot of valuable knowledge and the way of thinking as a scientist but also as a person. I would like to thank the committee members, Assist. Prof. Dr. Coşkun Kocabaş and Assist. Prof. Dr. Necmi Bıyıklı for their generous guidance and review on this work.

I also would like to expand my thanks to my group members: Mustafa Ürel, Okan Öner Ekiz, Sencer Ayas, Hasan Güner, Serkan Karayalçın, Andi Çupallari and Ayşegül Abdelal. I also would like to thank UNAM engineers Fatih Büker, Semih Yaşar, Adem Saraç, Fikret Piri and Mustafa Kemal Ruhi.

I would like to thank my clean room family: Feyza Bozkurt-Oruç, İnci Dönmez, Çağla Özgit-Akgün, Engin Çağatay, Mehmet Alican Noyan, Fatih Bilge Atar, Levent Erdal Aygün and Oğuz Hanoğlu. It was a great experience that I worked and enjoyed my time in the clean room with all of you. I would like to thank my dear friends: Gözde Turan, Zeynep Seval, Hülya Budunoğlu, Gülesin Eren, Aylin Civan and Pınar Beyazkılıç.

My deepest gratitude goes to my parents Fundagül and Çetinkaya Kocaay for their encouragement, support, patience and endless love. Special thanks go to my dear brother Aziz Berkin Kocaay for his help for figures in the thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Flash memory technology has recently gained much attention with the growing demand of non-volatile memories for mobile electronic devices. The conventional flash memory technology is composed of a MOSFET structure with a continuous floating gate (FG) as the charge storage medium. The basic operation principle of FG flash cell is based on the threshold voltage shift as a result of injected electrons in the charge storage medium. High retention rate, low power consumption, high density and high endurance are the desired features of a flash memory device.

As the dimensions scaled down, the FG technology encounters serious problems such as lateral charge leakage into drain or source regions, and complete charge loss because of pinholes and defects in the tunnel layer. Charge-trapping flash memory structures are proposed to overcome the drawbacks of scaled FG cells. In charge-trapping flash memories, the charge storage medium consists of a large number of electrically isolated traps. Nanoclusters in Si-nitride, semiconductor nanocrystals (Si, Ge etc.) and metal nanocrystals (Co, Ni, Al, Au, Ag etc.) are utilized as discrete traps. Discrete charge storage nodes prevent lateral charge leakage and allow further scaling of the tunnel oxide resulting in direct tunnelling mechanism. Direct tunnelling mechanism is desirable in flash technology to avoid oxide degradation due to

Fowler-Nordheim tunnelling and hot electron injection. Moreover, it allows faster write/erase speed at low operating voltages. The use of discrete charge-trapping medium offers high endurance, low power consumption and fast operation in high density memory applications.

Leakage current increases exponentially as the gate oxide gets thinner, causing reliability problems in flash memories. Various materials with high dielectric constants are proposed as tunnel oxide and blocking oxide in flash memory technology. With the application of high-k dielectrics, the leakage current is significantly suppressed due to increased physical dielectric thickness. High-k dielectrics in flash cells enhance the write/erase speed without degrading the data retention performance. To further enhance the charge injection current and the retention rate, charge storage medium should provide great number of available states in deep energy levels. High-k dielectric flash devices with dense nanocrystals exhibit improved retention performance and faster operation.

This thesis investigates the charge storage property of graphene in flash memory applications and the effect of atomic layer deposited tunnel dielectrics on memory performance. For comparing graphene flash memories, identical flash cells with Si-rich nitride layer as the charge-trapping medium are fabricated and characterized. Our motivation is to improve the performance of the existing flash technology with the use of high-k dielectrics instead of SiO_2 as tunnel/control barrier and graphene as charge storage medium.

This thesis is organized as follows: Chapter 2 includes theoretical preliminaries and basic concepts in flash memory devices. Chapter 3 addresses the promising solutions to the scaling issue in modern flash technology, including the usage of nanocrystals, ALD-grown high-k dielectrics and graphene. Chapter 4 provides detailed explanation of fabrication process and characterization methods. In Chapter 5, performance of fabricated graphene flash memories and detailed comparison with Si-nitride based flash memories

are presented. Chapter 6 concludes the thesis by summarizing the results and provides future work in this direction.

Chapter 2

Device Physics

2.1 Overview of Non-Volatile Memory

Memory is an inherent crucial component of any information processing system. Various types of memories find application in different parts of modern computer systems. Solid state memories offer high density, low power and no mechanical components. Memory devices can be classified into two main types, volatile and non-volatile. A volatile memory loses the information when the power is off. Static Random Access memory (SRAM) and Dynamic Random Access Memory (DRAM) are the most widely-used types of a volatile memory family. SRAMs offer fast writing and reading operations and DRAMs have denser structure.

Unlike volatile memories, non-volatile memories retain the stored content even if the power is turned off. Non-volatile memories (NVM) can be random access and read only. Read-only Memory (ROM), Erasable programmable read-only memory **(**EPROM), Electrically Erasable Programmable Read-Only Memory (EEPROM) and Flash are the main types of NVMs. Among all, EEPROMs and Flash memories dominant in applications due to continuous improvement in performance and density. An EEPROM cell consists of two transistors while there is a single transistor in each Flash cell. Due to extra transistor, EEPROMs offer lower chip density compared to Flash memories.

Both EEPROMs and Flash memories require a floating gate layer for charge storage.

In recent years, the share of NVM in semiconductor market is increasing due to portable electronic equipments such as cellular phones, digital cameras, laptops etc. Low operating voltage, high density and high retention performance are the desired features of an ideal NVM. Compared to others, Flash seems to be more advantageous considering the cost and the ability to be programmed and erased many times. In the coming years, as mobile computing becomes more widely used, the demand for flash is expected to increase because of the need for low-cost, low-power, and high-density applications such as solid state hard drives, or expanding storage capacity in cell phones and tablets.

2.2 Performance Considerations of Flash Memories

High retention rate, high density, high endurance and low power consumption are the main features of a desired non-volatile semiconductor memory.

Retention rate is the ability to retain the stored information even without power supply. Non-volatility implies that the flash cell should retain the data for at least 10 years at temperatures near 80°C and after many write/erase cycles. In contrast, volatile memories are needed to be refreshed to prevent from losing the stored data. The refresh cycle is important for power consumption. Memories with high retention rate need lower refresh cycles, which decreases power consumption rate. Retention performance is determined by testing the reliability of a flash cell for nearly 10^4 - 10^5 s and extrapolated up to 10^8 s (approximately ten years). Moreover, different experiments are also conducted at high temperatures to determine the retention capability of the flash cell.

With the current trend in scaling device dimensions, today's electronic applications require high-density components. A unit cell that is used to store one bit of information determines the memory density. Flash memories are suitable for high density applications due to single transistor structure in each cell. However, serious problems such as difficulty in lateral charge confinement and neighboring cell disturbance may occur due to scaled dimensions. When lateral dimension of memory cell becomes less than 25 nm, confined charges on FG can leak out into drain or source regions; therefore, such cells may need to be refreshed frequently to retain the data, which is not desired for NVMs. In addition, because of lateral charge leakage, one cell can easily disturb neighboring cells. Therefore, development of new technologies to improve floating gate memories must take into account geometry related effects in the device.

Endurance is another criterion to determine the memory performance. Endurance of a memory cell is referred as the ability to be accessed many times without the degradation of the performance. After many write/erase cycles, oxide defects may occur due to charge injection, which results in degradation of endurance and retention, and also reliability problems due to narrowing of the memory window. For solid state hard drives, endurance values above 10^5 are desirable, where the data is frequently updated.

Low operating voltages are another desired feature of an ideal NVM. It is possible to reduce power consumption by decreasing oxide thickness between the FG and the channel. However, stored charges can leak into the channel easily when the tunnel oxide is scaled down. Moreover, the effect of hot electrons to thinner oxide becomes more serious. As a result, the application of thinner oxide in a Flash cell may reduce both endurance and retention performance.

As outlined above, there are several directions in improvement of floating gate non-volatile memories. A more detailed discussion as given below can help identify routes to improve various qualities of flash NVMs.

2.3 Floating Gate Flash Memories

Floating gate (FG) flash memory is basically a metal-oxide semiconductor field effect transistor (MOSFET) with a poly-silicon FG layer sandwiched between two insulating thin-films. Actually, the storage location is mainly a MOS capacitor with two dielectrics and a FG gate layer between them. The structure of conventional flash cell can be seen in Figure 2.1. FG is a continuous thin-film in which charges are stored. The state of the memory cell is determined by the electrical charge state of the FG. The layer separating FG from the device channel is called as tunnel barrier. The thickness of this layer is generally in the range of 2-10 nm. The FG layer is blocked from the control gate by a control oxide that is used to prevent from discharging through gate contact.

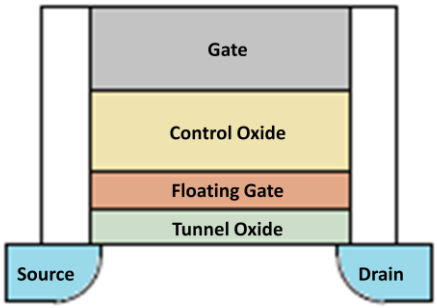

Figure 2.1: Device Structure of Conventional FG Flash Cell

Basic operation principle of a FG flash cell is as follows. When positive bias is applied to the gate, electrons are attracted from substrate through FG tunnel barrier layer via tunneling and trapped on the FG when the gate bias is removed. Negative bias causes the lowering of the barrier and facilitates escape

of electrons back to the substrate. The amount of charge stored on FG results in threshold voltage shift of the underlying transistor. Generally, neutral or positively charged state is defined as logical "1" and programmed or negatively charged state is referred as logical "0". Fowler-Nordheim (FN) and hot electron injection are the most frequently used mechanisms for charging and discharging. Compared to direct tunneling, these mechanisms are faster; however, they cause damage in the oxide layer due to high fields, which degrades endurance and retention performance of the flash cell. Thickness and quality of each oxide layer are also crucial to guarantee both non-volatility and program/erase repeatability. The information is preserved as long as the charges are trapped of FG. For a conventional flash cell, it is expected to retain the data for at least 10 years at temperature range of -40°C and 125°C [1]. Conductive paths along dielectrics may cause trapped charges on FG to tunnel to substrate or to gate; therefore, defect-free oxides are necessary for high performance flash applications.

D. Kahng and S. M. Sze suggested the use of FG for charge storage for the first time in 1967 [2]. The first proposed structure for non-volatile MOS memory is as follows. A metal conducting layer as FG is sandwiched between a thin insulator as tunnel oxide and a thick insulator that isolates the FG from the gate metal. This memory cell is called as MIMIS (metal-insulator-metal-insulator-semiconductor). Extremely thin oxide layer (< 5 nm) allows direct tunneling mechanism for programming. The main drawback of this structure is that all stored charges in metal layer can leak off in case of any pinhole in the tunneling oxide. Hence, there is a reliability problem with MIMIS cell. Increasing the thickness of tunneling oxide and replacing conducting layer with a dielectric without losing the capture probability would be the possible solutions to MIMIS structure.

The first improvement to reliability problem of MIMIS cell was first suggested by Wegener in 1967 [3]. Wegener proposed to replace conducting

layer and the insulator on top it with a nitride layer. This structure is referred as MNOS (metal-nitride-oxide-semiconductor). Nitride layer is composed of large number of traps for both electron and hole storage. Since individual trapping centers are separated from each other, any pinhole in thin tunnel oxide will not cause complete discharging. Applying high positive voltage to gate metal, electrons tunnel from silicon conduction band (CB) to nitride CB and are confined in nitride traps, causing positive shift in threshold voltage. Conversely, high negative bias is applied to erase the cell, resulting in negative threshold voltage shift.

Floating gate Avalanche injection MOS (FAMOS) was suggested by Frohman-Bentchkowsky in 1971 [4-7]. In FAMOS cell, a poly-silicon FG surrounded by a thick oxide is used as charge storage layer. The programming mechanism is mainly based on the creation of highly energetic electrons by drain avalanche plasma. On the other hand, UV or X-ray radiation can be utilized to erase the cell. FAMOS cell was the first structure that is compatible with mass production.

2.4 Charge-Trapping Flash Memories

The main distinguishing feature of a charge-trapping flash memory is a discontinuous charge storage medium that contains a large number of electrically isolated traps. The cross-section of charge-trapping cells can be seen in Figure 2.2. Charges are stored in discrete traps, preventing from lateral charge leakage. Nitride films, semiconductor nanocrystals and metal nanocrystals can be used as charge-trapping layers.

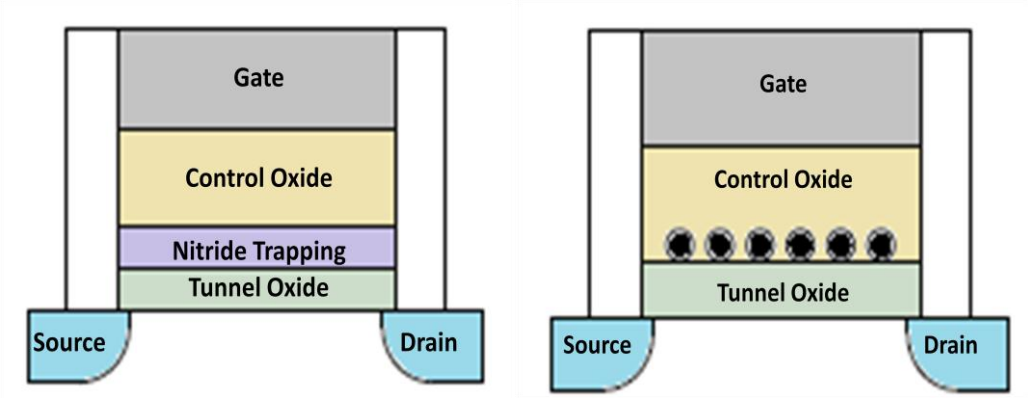

Figure 2.2: Device Structures of SONOS and Nanocrystal Flash Cells

The application of discontinuous layer as charge storage medium allows for the reduction of tunnel oxide thickness. In charge-trapping devices, direct tunneling mechanism can be used for charging and discharging because of ultra-thin tunnel oxide. Charge-trapping devices gain some advantages over floating gate memories due to direct tunneling mechanism. One of the main advantages is the life-time of the memory cell is increased since direct tunneling does not degrade the oxide as do Fowler-Nordheim or hot electron injection mechanisms. Another advantage of direct tunneling mechanism is to allow faster program/erase operations. In addition, retention rate of nanocrystal memories is less prone to oxide defects compared to a continuous floating gate memory since there is no lateral conduction between trap centers. Although charge-trapping memories can offer high endurance, low power consumption and fast program/erase speed at high-density memory applications, the distribution and the size of nanocrystals are both crucial to determine device performance.

2.5 MOS Physics

A nanocrystal flash cell can be modeled as a one-dimensional metal-oxide-semiconductor (MOS) capacitor since thick dielectric electric isolates the nanocrystals laterally and the electric field is applied along gate-to-substrate

direction. Size, shape and distribution of nanocrystals, interfacial traps, oxide defects should be carefully analyzed to determine the electrostatic characteristics of a nanocrystal flash memory. To examine the electrical properties, MOS capacitor physics should be studied.

MOS capacitor is the main structure of a single flash cell. A MOS capacitor is composed of a dielectric layer sandwiched between a doped silicon substrate and a metal contact. An ideal MOS capacitor has no charge-trapping centers in the oxide and at oxide/semiconductor interface. Moreover, for an ideal MOS structure, the metal work function should be equal to the semiconductor work function. Leakage current under all static conditions is assumed to be zero in an ideal MOS.

According to applied voltage, MOS capacitor has three biasing regions, accumulation, depletion and inversion; all of which will be studied over an ideal n-type MOS.

2.5.1 Accumulation

When positive voltage is applied, Fermi level (E_F) of the metal is lowered relative to E_F of semiconductor, causing a positive slope in the energy band diagram. Excessive positive charges placed in gate electrode and this positive bias should be balanced by the electrons at the interface. In accumulation region, the total capacitance is the result of an oxide capacitance.

2.5.2 Depletion

The application of small negative voltage to n-type MOS capacitor slightly raises E_F of metal with respect to E_F of semiconductor. The concentration of electrons on the gate contact increases, causing the repulsion of majority carriers from the interface. Electron concentration is becoming less than the doping

concentration of Si substrate as the bias voltage is decreasing. The total capacitance is the serial combination of two parallel-plate capacitors, oxide capacitance and depletion capacitance. The depletion capacitance decreases with the increase of depletion thickness.

2.5.3 Inversion

When large negative voltage applied, bands at the interface bend more and depletion width reaches the maximum value. Hole concentration at the surface gets larger than the intrinsic carrier concentration. At low frequencies, minority carriers can follow the ac signal and total capacitance approaches to oxide capacitance value as the negative bias becomes larger. On the other hand, holes in the inversion layer cannot follow the ac signal at high frequencies. Then, the total capacitance will be equal to serially connected oxide capacitance and depletion capacitance.

Band diagrams and charge distributions of each biasing regions including flat-band condition for an ideal n-type MOS capacitor are demonstrated in Figure 2.3.

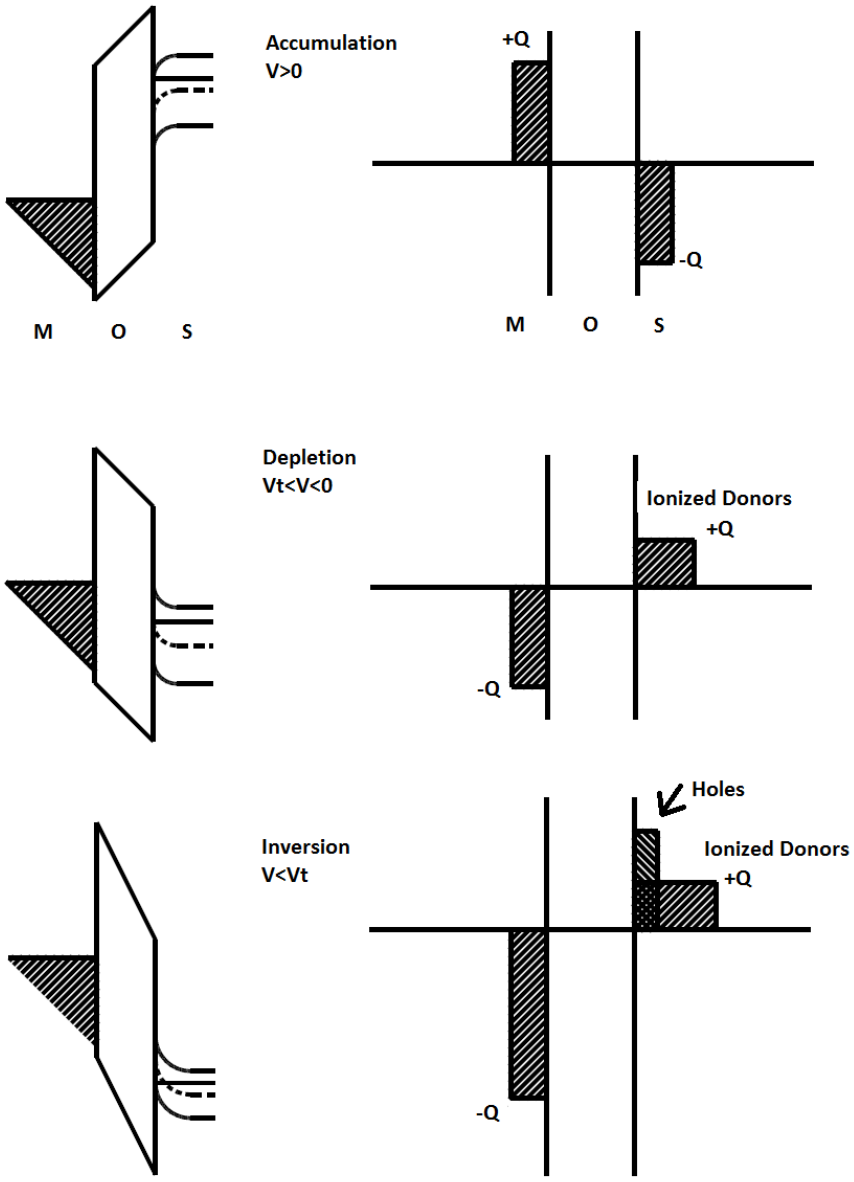

Figure 2.3: Energy band diagrams and corresponding block charge diagrams in an ideal n-type MOS capacitor.

2.6 Tunneling Mechanism

An ideal insulator does not allow any leakage current. But, when high electric field is applied on a thin dielectric layer, charges will pass through the dielectric. Tunneling mechanism is explained as the propagation of a particle through a

potential barrier that it could not pass according to classical view. Tunneling depends on the barrier height, availability of the states and applied electric field. Direct tunneling, Fowler-Nordheim tunneling and trap-assisted tunneling are the tunneling mechanisms to explain charging/discharging of a flash cell, shown in Figure 2.4.

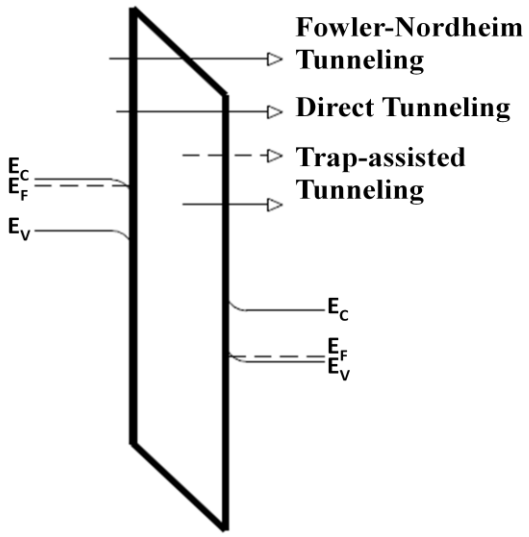

Figure 2.4: Tunneling Mechanisms

2.6.1 Direct Tunneling

Direct tunneling mechanism is referred as the propagation of electrons through the barrier without using the conduction band of the dielectric. It becomes dominant for the dielectrics with a thickness below 5 nm. Direct tunneling is temperature-dependent. It increases exponentially as the decrease of oxide thickness. It does not dependent on the electric field across the dielectric. Direct tunneling does not create oxide defects after repeated cycle operations [8].

2.6.2 Fowler-Nordheim Tunneling

Fowler-Nordheim (FN) tunneling is a quantum mechanical mechanism in which electrons can propagate through the barrier by passing the conduction band of the oxide. The probability of FN tunneling increases exponentially as the electric field in the oxide layer increases [9].

2.6.3 Trap-assisted Tunneling

Oxide defects and interface traps give rise to two-step tunneling mechanisms. Trap-assisted tunneling becomes significant after many write/erase cycles in EEPROMs and flash memories. Due to repeated high stress, the tunneling current increases at low voltages. This is regarded as stress induced leakage current (SILC). SILC is generally accepted as the main reason for the degradation of oxide quality and retention performance in NVMs [10]. SILC was examined for MOS capacitors and EEPROMs in recent studies [11-13]. Single defect in a thick oxide would not be sufficient to initiate a leakage current. However, the existence of several traps can create large SILC and results in gate oxide breakdown.

Chapter 3

Possible Solutions to the Scaling Issue

3.1 Nanocrystal Flash Memory

Floating gate (FG) flash memories encounter major problems as the flash gate stack is scaled down, explained in Chapter 2. In recent years, various materials as charge-trapping layer and high-k dielectrics as tunneling and blocking layers are proposed to overcome scaling issue of the conventional FG memory. One possible solution is to create a memory structure that stores charges in discrete-traps, such as SONOS type memory and nanocrystal memory [14, 15].

The structure of a nanocrystal cell resembles conventional FG memory. The main difference is that charges are not stored in a continuous FG, but in nanocrystals or natural traps of a nitride layer. Those discrete charge-trapping centers allow thinner tunnel oxide thickness without a drastic decrease in the retention performance.

In a FG cell, oxide defects in the tunneling layer may form a conductive path the channel and the FG, which may cause full discharge of stored charges. On the other hand, nanocrystal memories can preserve most of the charge and improve the retention capability since they are stored in discrete traps and such a conductive path may discharge only a small number of nanocrystals. Therefore,

charge-trapping flash memories offer better scalability of the tunnel oxide thickness without retention degradation.

Another major benefit of charge-trapping memories is to allow multi-bit storage due to localized charge storage mechanism [16]. By injecting the charge through channel hot electron (CHE) to the source, to the drain or to the both sides and reading on only one end, it is possible to obtain multi-bit storage mechanism in a single cell, which has a great importance to increase the data density.

The application of nanocrystals as charge-trapping medium seems to be advantageous. They require the use of well-known materials and CMOS-compatible fabrication steps since they have a similar structure to the conventional FG.

3.1.1 Semiconductor Nanocrystal Preparation Methods

The fabrication of semiconductor nanocrystals embedded in a MOS structure has been commonly studied. There are four major methods of nanocrystal preparation in non-volatile memory technology, which are the chemical vapor deposition (CVD) growth, the non-stoichiometric layer deposition, ion-beam synthesis and the layer-by-layer growth.

Si and Ge nanocrystals embedded in a dielectric matrix have been widely used in nanocrystal memory technology. In general, high temperature annealing process is required to complete nanocrystal formation. Ion implantation of Si and Ge in SiO_2 matrix is the most frequently used method. This method is based on implantation of Si or Ge into SiO_2 with low energies and post-implant annealing at high temperatures (>900ºC). At layer-by-layer growth technique, a thin amorphous Ge or Si layer is deposited by e-beam evaporation or thermal evaporation system, and then high-temperature annealing process is applied to

the sample. Si or Ge thin layer is either oxidized or covered by another dielectric layer to complete nanocrystal formation.

Another method of preparation is the deposition of Si-rich SiO_x or SiN_x layer by plasma enhanced chemical vapor deposition (PECVD) and subsequent high-temperature annealing. Iacona and coworkers showed that Si nanoclusters with a diameter between 1.4 nm and 4.2 nm are formed after high-temperature annealing of PECVD-deposited Si-rich SiO_2 films [17]. The study of Sung et al. studied the formation of Si nanocrystals embedded in amorphous SiN_x films grown by PECVD at 250ºC [18]. The size of Si nanocrystals varies from 2.1nm to 6.1 nm as altering deposition parameters [18, 19]. Further post-annealing is not applied to the Si-rich oxide and nitride films deposited by PECVD.

CVD is recently the most frequently used technique to form self-assembled nanocrystals in memory applications. The crucial point of this method is to control the initial nucleation of nanocrystals on top of a dielectric. The study of Ammendola et al. seems to be promising to obtain uniformly-distributed Si quantum dots incorporated in a thin SiO_2 layer with low-pressure CVD (LPCVD). The average size of Si quantum dots was determined as 4-6 nm [20].

3.1.2 Metal Nanocrystals

The application of metal nanocrystals as charge-storage layer is also promising because metal nanocrystals offer selectable work function and high density of states. In general, the fabrication of metal nanocrystals is based on the deposition of a thin metal layer by evaporation or sputtering techniques and post-deposition annealing.

Lee et al. demonstrated the charge storage property of Au, Ag and Pt nanocrystals in EEPROM devices [21]. A thin layer of Au, Ag and Pt was deposited o a direct tunneling oxide by e-beam evaporation and subsequently

annealed at 1000ºC. The size and the distribution of metal nanocrystals could be controlled with evaporation and annealing conditions. Besides the study of Lee et al., the charge-trapping property of various metal nanoparticles such as Co, Mo, Ni, TiN and Al nanocrystals were investigated [22-28].

Metal with high work functions are desirable for memory applications because higher effective potential well depth with respect to Si conduction band improves both retention performance and operation speed of the memory. Metal nanoparticles are crucial for engineering of effective potential well depth (d_{eff}). The energy band diagram of a flash cell with metal nanoparticles can be seen in Figure 3.1. High d_{eff} offers lower barrier for writing and longer barrier for retention. Thus, fast write/erase speed with improved retention can be achieved with metal nanocrystals as charge storage nodes.

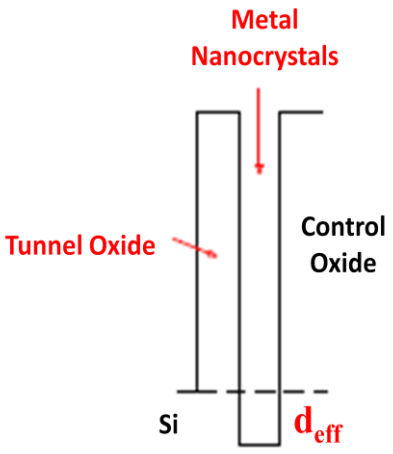

Figure 3.1: Energy Band Diagram of a Flash Cell with Metal Nanocrystals

Au and Pt have higher work function compared to Co, W, Ag, Al and Ni; as a result, they exhibit better memory performance. However, the application of Au and Pt nanoparticles does not seem to be a cost-effective solution in memory technology. There have been many reports on metal species with high work functions such as W_5Si_3, $NiSi_2$, $CoSi_2$ nanocrystals to enhance the memory performance and to allow further scaling [29-32]. The use of metal nanoparticles

also offer multi-bit storage mechanism. 2-bits per cell metal nanocrystal memory and 4-bits per cell with quad source/drain devices were reported [33, 34].

3.1.3 Basic Principle of Operation

In conventional flash memories, data storage is mainly based on the threshold shift of FETs due to stored charges. During programming, charges are tunnelled from the channel to the FG. Reading operation is dependent on the measurement of source-drain current. Tiwari et al. explained the charge storage mechanism of nanocrystal flash memories as follows [35]. Reverse bias applied to the gate causes the injection of electrons into the nanocrystals. Conductance of the inversion layer is reduced because of the screening effect of the trapped charges in nanocrystals, causing a shift in the threshold voltage of the FET.

Write/Erase speed and operation voltage are both based on the injection current between the channel and the charge-stored medium. The injection current is exponentially dependent on the electric field across the tunnel oxide for FN tunneling mechanism.

A nanocrystal flash cell can be modelled as a serially connected three parallel plate capacitors. The capacitance of a parallel-plate capacitor is dependent on the permittivity of the dielectric between two plates, the thickness of the dielectric and the area.

When there is no charge stored in the flash memory, the electric field in the bottom oxide can be found as

$$E_{ox} = \frac{Vg}{t_{ox}\left(1+\frac{\varepsilon_{ox}t_c}{\varepsilon_c t_{ox}}\right)} \tag{3.1}$$

where

Vg : Applied gate voltage

E_{ox}: Electric field in the tunnel oxide

ε_{ox} : Dielectric constant of the tunnel oxide

t_{ox} : Tunnel oxide thickness

ε_c : Dielectric constant of the control oxide

t_c : Control oxide thickness

The electric field in the tunnel oxide is a function of the thickness and the dielectric constant of both tunnel oxide and control oxide layers. The decrease in the ratio $\frac{\varepsilon_{ox}}{\varepsilon_c}$ enhances the electric field in the tunnel oxide. If the control oxide with a higher dielectric permittivity with respect to the tunnel oxide is used in the gate stack, the programming time duration or the operation voltage will decrease.

Taking into consideration that nanocrystals are the only charge storage nodes and exactly one electron is stored in each nanocrystal, flat band voltage shift is given as [35]

$$\nabla V_{FB} = \frac{q n_{nc}}{\varepsilon_{ox}} \left(t_c + \frac{1}{2} \frac{\varepsilon_{ox}}{\varepsilon_{Si}} t_{nc} \right) \tag{3.2}$$

The charge density, n_{nc}, can be directly calculated as

$$n_{nc} = \frac{\nabla V_{FB} \varepsilon_{ox}}{q} \frac{1}{\left(t_c + \frac{1}{2} \frac{\varepsilon_{ox}}{\varepsilon_{Si}} t_{nc} \right)} \tag{3.3}$$

where

n_{nc} : Charge density in nanocrystal medium

∇V_{FB}: Flat band voltage shift

ε_{ox} : Dielectric constant of the tunnel oxide

t_c : Control oxide thickness

t_{nc} : Nanocrystal medium thickness

ε_{Si} : Dielectric constant of Si

3.1.4 Challenges of Nanocrystal Memories

Although nanocrystal memories is regarded as a solution to scaling issue in conventional FG technology, there still exist major challenges that should be overcome. First of all, the memory performance of nanocrystal-based flash cells is directly affected by the size, the shape and the distribution of nanocrystals. The nanocrystal size must be around 3-4 nm due to carrier confinement effects. When more than one electron is stored in a single nanocrystal, the barrier potential for the electron occupying the highest energy level will reduce due to the increase in energy level separation according to the given formula assuming the nanocrystal shape is a perfectly symmetric sphere [36].

$$\nabla E_{n,n-1} = \frac{q^2}{2C_{nc}}(n^2 - (n-1)^2) \qquad C_{nc} = 2\pi\varepsilon d \tag{3.4}$$

where

$\nabla E_{n,n-1}$: Energy separation

C_{nc} : Capacitance of a spherical nanocrystal

d : Diameter of the nanocrystal

ε : Dielectric constant of the nanocrystal

In addition, the density of nanocrystals should be at least 10^{12}cm^{-2} to obtain an appropriate shift in threshold voltage. Moreover, the separation between two neighboring nanocrystals is also significant to prevent lateral current flow. Generally, nanocrystals are formed by thermal annealing process, rather than by patterning. Thus, it becomes challenging to produce nanocrystals with a certain size, shape and distribution in each individual memory cells.

The use of metal nanoparticles in charge-trapping memories has been also receiving attention since metal nanoparticles offer large density of states around the Fermi level and high work function. Metal nanoparticle memories show improved retention and greater charge storage capability compared to semiconductor nanocrystals due to the deeper effective potential well formed between the metal and the conduction band of the semiconductor. However, the formation of metal nanocrystals is usually required high-temperature process. They can easily diffuse into other layers during thermal annealing. Therefore, they create defects and conductive paths in the tunnel oxide and the control oxide. As a result, non-volatile memories with metal nanocrystals may have reliability problems. Moreover, flash cells with Au nanoparticles as charge storage nodes exhibit better performance due to its higher work function. However, the Au is considered as a non-CMOS compatible material.

3.2 High-k Dielectrics

Modern semiconductor technology requires high-density and low-power applications. In recent years, there have been many applications on possible solutions to overcome challenges arising from scaled dimensions. Silicon-based technology dominates modern electronic devices because of the excellent interface at Si / SiO_2. Although other semiconductors such as Ge, GaAs, InGaAs offer greater transport properties, low-trap density at the interface of Si and its native oxide, SiO_2 is considered as one reason that silicon-based FETs are the basis of modern electronics.

Further scaling of SiO_2 beyond 10 nm brings problems due to two main reasons. First of all, leakage current rises exponentially as the gate oxide gets thinner because of quantum mechanical tunneling of carriers through such a thin SiO_2 layer. The leakage current not only reduces the reliability of the device, but also creates oxide defects after repeated cycles of operation. The second reason is that the SiO_2 loses its bulk electronic properties when its thickness is scaled

down to 0.7 nm. Because of such problems arising from the scaling issue, semiconductor technology is looking for new materials that form low-trap density on top of Si and offer equivalent electrical properties of a scaled SiO_2.

High-k materials are regarded as an alternative solution to scaling issue to suppress the leakage current without reducing the oxide capacitance. The reduction of the leakage current is achieved by increasing the physical thickness of gate oxide, which is explained with equivalent oxide thickness (EOT). The EOT represents the electrical thickness of a dielectric that corresponds to the equivalent capacitance of a physically thinner SiO_2 layer. The EOT is given as

$$EOT = t_{high-k}\left(\frac{k_{SiO_2}}{k_{high-k}}\right) \tag{3.5}$$

where t_{high-k} is the thickness of the high-k layer, k_{SiO_2} and k_{high-k} are the relative dielectric constants of SiO_2 and high-k dielectric.

Band gaps and band offsets of various high-k materials can be seen in Figure 3.3 [37]. A high-k material with large band gap and high band offsets for both electrons and holes are required for an alternative gate dielectric. The inverse relation between the dielectric constants and the band gaps is shown in Figure 3.4 [38].

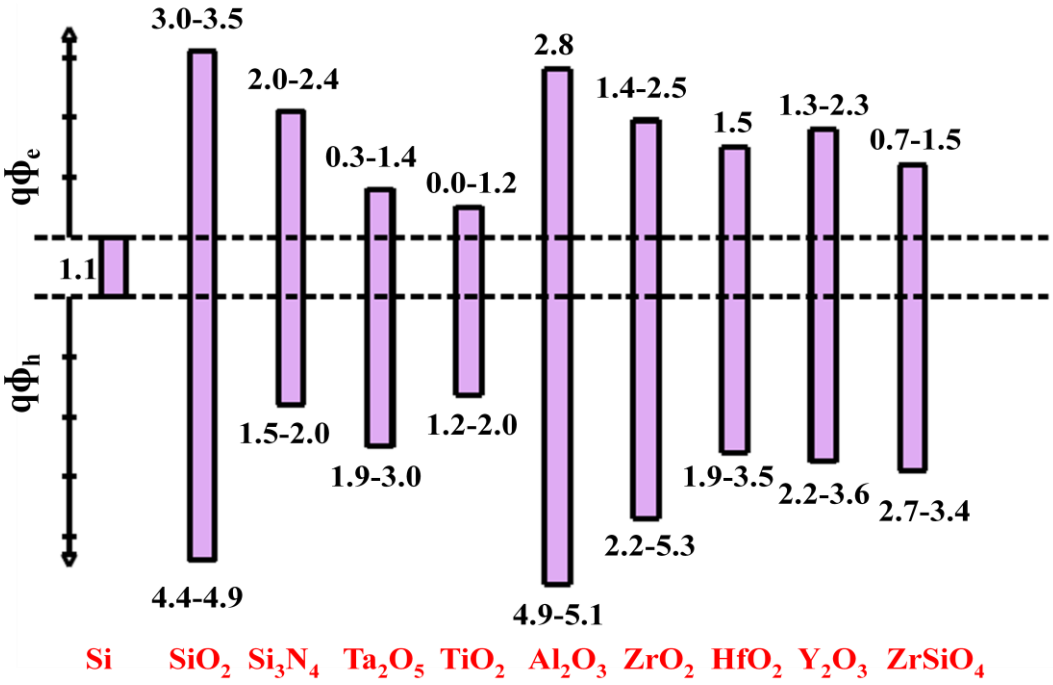

Figure 3.2: Band gaps and band offsets of various high-k materials [37].

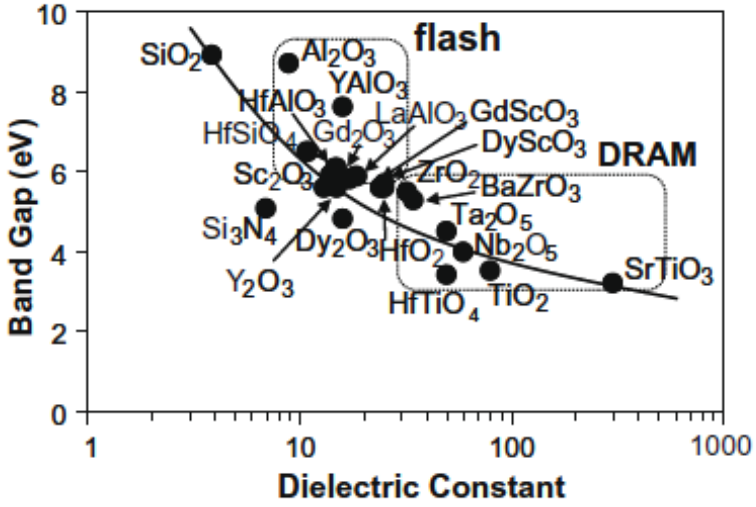

Figure 3.3: Inverse relation between band gap and dielectric constant of various high-k materials [38].

The first motivation for the replacement of SiO_2 with a high-k material was to reduce the leakage current. Much work has focused on the leakage current measurements of devices with various high-k gate dielectrics. A transistor with Al_2O_3 as a gate dielectric was reported in 2000 [39]. Further studies showed that the high-k dielectric layer as a gate oxide causes a flat band voltage shift in capacitance-voltage measurements, indicating the existence of oxide charges and traps. Studies on thermodynamic stability of the high-k materials onto Si have also reported [40-42]. BeO, ZrO_2 and HfO_2 are shown to be thermodynamically more stable than Al_2O_3, TiO_2 and Ta_2O_5 [43].

Gritsenko et al. compared the simulated write/erase characteristics of SONOS devices with SiO_2, Al_2O_3 and ZrO_2 as a blocking oxide. They concluded that the application of high-k as a control oxide reduces the operation voltage or the write/erase speed from 1ms to 10μs [44]. Moreover, interfacial properties of high-k materials with Si in flash applications were studied [45].

The use of high-k materials as a tunnel oxide in flash memories is advantageous over SiO_2 gate oxide. First, it improves the write/erase speed. The charge injection mechanism of most flash memories is based on FN tunneling and hot electron injection. These tunneling mechanisms are dependent on the potential barrier height between the Si substrate and the tunnel oxide. Actually, larger barrier height requires higher voltage operation or longer programming duration. Most high-k dielectrics form lower barrier height on Si, allowing faster operations or low-voltage applications. The second is the improved retention rate. Since high-k dielectrics are physically thicker than a SiO_2 layer to achieve the EOT, SONOS with a high-k dielectric as a tunnel oxide can exhibit better retention performance.

Dana et al. proposed a model that explains the charge/discharge dynamics of nanocrystal flash memories and studied the effect of dielectric properties on memory performance [46]. In this study, the figure of merit (FOM) is defined as

$$FOM = log_{10} \frac{\tau_{ret}}{\tau_{charge}} \tag{3.6}$$

where τ_{ret} and τ_{charge} denote the retention and the charging times. The FOM is calculated for various high-k dielectric materials with same EOT and other parameters that effect charge/discharge dynamics such as nanocrystal size and distribution, write voltage, and gate work function are kept unchanged, given in Figure 3.4.

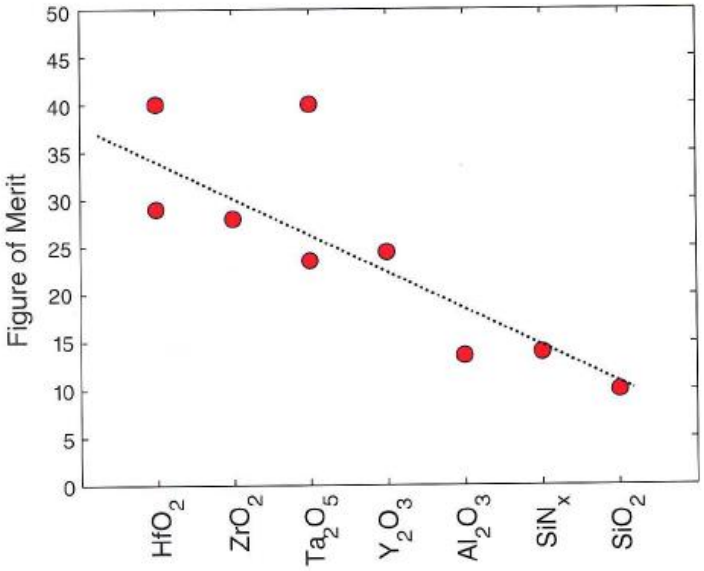

Figure 3.4: FOM for various high-k materials for nanocrystal diameter of 6 nm, write voltage of 10 V, and EOT of 4 nm [46].

The use of materials with higher dielectric constants improves the FOM, leading to enhanced retention rate and decreased charging time. Since high-k materials provide larger physical thickness, the escape tunnelling current is significantly decreased. Charging time is also improved as a result of smaller conduction band offset of the high-k material on Si.

3.2.1 Reliability Issues of High-k Dielectrics

Although high-k dielectrics are regarded as a solution to scaling issue of SiO_2 gate oxide, they have various disadvantages that directly affect the performance and the reliability of the device.

One major problem for the integration of high-k dielectrics to CMOS technology is the presence of dipoles and defects at Si/high-k interface. Capacitance-voltage measurements of a MOS capacitor with a high-k layer demonstrate that the flat band voltage is shifted from the ideal position, proving the existence of uncompensated charges and dipoles in the high-k oxide. Diffusion of oxygen through the film and formation of SiO_x layer at the interface was observed in most high-k oxides including Al_2O_3, ZrO_2 and HfO_2. To prevent oxygen diffusion from gate electrode, poly-Si gate is replaced with a metal electrode such as TiN [47]. At this point, the metal contacts should have appropriate work functions and thermal stability at temperatures required for CMOS technology.

Secondly, mobility degrades as the thickness of gate stack decreases. The reason for this is explained as the coupling of soft phonon modes in the dielectric with channel electrons [48]. It is probable that charge trapping and fixed charges in high-k films are responsible for lower carrier mobility. Moreover, it was reported that the mobility becomes higher when the thickness of the interfacial SiO_x becomes more than 1nm by thermal treatments. Forming gas annealing at high temperatures and O_3 surface treatments were carried out to reduce interfacial defects [49-51].

In addition, another challenge with the integration of high-k dielectrics is the presence of fast charge trapping and detrapping. Hysteresis phenomena and threshold voltage instability in MOSFETs with HfO_2 was reported [52]. It was

observed that threshold voltage shift depends on the HfO_2 thickness and post deposition annealing conditions.

The thermal stability of high-k materials is also crucial for device performance because conventional CMOS technology needs for high-temperature steps. At high temperatures, ultra thin high-k layers cannot prevent O_2 diffusion through Si surface and there occurs SiO_2 growth at the interface [53, 54].

3.3 Graphene

Graphene is a one-atom thick material made of carbon atoms that are arranged in honeycomb order. In 2010, graphene was worth of Nobel Prize since it is considered as one of the most promising material ever discovered. Graphene is regarded as a unique material because of its electrical, thermal, mechanical, optical and chemical properties. Graphene is one of the strongest and the most conductive material ever known. The unique properties of graphene make it the most promising material for electronic applications, including transistors, sensors, composite materials and flexible electronics. Especially for transistor technology, graphene is proposed as a new solution that can be used to further Moore's Law.

The main challenge is the growth of high quality graphene layers. Graphene films are desired to be uniform with minimal roughness and few defects. One method of graphene deposition is metal exfoliation or scotch-tape method [55]. In this method, the starting material is commercially available highly ordered pyrolytic graphite (HOPG). To mechanically exfoliate graphene on a substrate, a piece of sticky tape is used to peel graphene layers from HOPG and sticky tape is removed slowly from the substrate. Although high-quality graphene layers with few defects are obtained, this method would not seem to be applicable to mass production. In addition, graphene is deposited in a small area

and number of graphene layers varies. The number of graphene layers is also a crucial issue in device applications. In recent years, many studies have focused on the fabrication of wafer-sized graphene layers and the interaction of graphene with other materials.

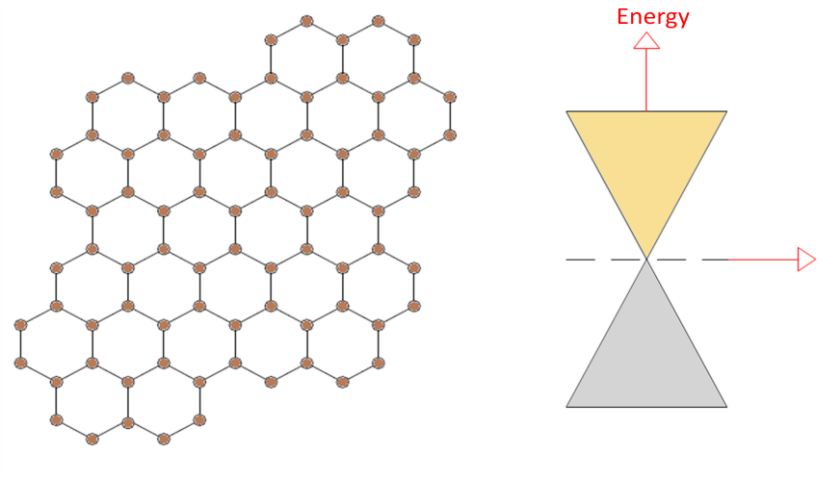

Figure 3.5: Graphene and its band structure

The high mobility and the doping property allow the graphene an alternative material to the silicon technology. Although the graphene has a zero band gap, recent studies showed that bilayer and few-layer graphene introduce a band gap [56]. Many researchers and industrial companies such as Intel and IBM are doing research on field effect transistor with graphene as a channel. The stability of graphene has already moved the attention into one-electron transistors and molecular-sized electronic devices.

3.3.1 Graphene Growth Techniques

The mechanical exfoliation is a simple method to produce graphene flakes with high purity. But, it is not useful for mass production and large-area applications. Besides mechanical exfoliation method, recent techniques are chemical vapor

deposition (CVD), epitaxy deposition and chemical decomposition of graphite oxide.

3.3.1.1 CVD Growth

Chemical Vapor Deposition (CVD) is a technique to deposit solid materials from a gaseous phase at low or vacuum pressure. In CVD, precursor gases are delivered into the reaction chamber at ambient temperatures. As they come into contact with a heated substrate, they react or decompose forming a solid phase and are deposited onto the substrate. The most commonly used material that catalyzes a CVD process is metal. In CVD growth of graphene, organic gases are used to initiate the growth of carbon monolayer. Growth condition can be optimized by changing the gas flow rates, the metal film thickness, the cooling rate and the temperature.

Li and coworkers have demonstrated graphene growth on copper using a mixture of hydrogen and methane gases at temperatures up to 1000°C [57]. One of the major benefits of this method is that it can be integrated to CMOS fabrication technology because 300 nm copper films on silicon wafer is a standard substrate in CMOS technology. They used SEM, TEM and Raman Spectroscopy to determine the uniformity and quality of graphene flakes. The process yielded graphene that was 95% monolayer over an area of 1cmx1cm. Mobility was reported as 4.050 $cm^2V^{-1}s^{-1}$, confirming the high concentration of monolayer graphene. The grown graphene can also be easily transferred to other substrates such as Si/SiO_2. This work seems to be promising for graphene-based applications; however, maintaining chamber pressure during the procedure is a crucial issue due to the gaseous methane. They also reported that graphene growth on copper is self-limited. Growth that proceeded for more than 60 min yielded is similar to growth performed for 10 min. For less than 10 min, SEM images show that copper film is not completely covered. The growth of graphene on Cu foils of varying thicknesses also yields similar structures. Based

on these observations, they reiterated that graphene is growing by a surface-catalyzed process rather than a precipitation process.

A method for the direct CVD of monolayer or few-layer graphene film on dielectric surfaces via a sacrificial copper film was reported [58]. Based on the observation of evaporation of Cu during CVD process, they proposed a new method for graphene growth on insulators by a controlled metal evaporation during or immediately after the catalytic growth. According to metal thickness and duration of CVD process, the areas between metal fingers change in size and shape. It was reported that almost 20 μm^2 size areas filled with continuous graphene layers. The continuity of the metal film on the surface depends on its thickness, the metal-dielectric wetting properties, the temperature and the time. This method was reported to be applicable to various insulating surfaces including single-crystal quartz, sapphire, fused silica, and silicon oxide. This process could suppress the need for a transfer to a new substrate. Additionally, further improvements on the control of dewetting and evaporation process could result in direct deposition of graphene for large-area electronics.

Rather than copper films, graphene can also be grown on nickel films. The film thickness should be less than 300 nm, preventing from graphite deposition. Graphene is produced on nickel surface using a diluted hydrocarbon gas at ambient pressure and 1000°C [59]. The film is then evaporated by e-beam deposition onto Si/SiO_2 substrate. Thermal annealing facilitates the growth of a nickel film. Nickel film can be etched with chemicals to transfer the graphene layers to another substrate. Direct patterning of graphene film was also demonstrated and this method eliminates the need for post-processing.

The study of Lee and coworkers is mainly based on the idea of graphene growth at low temperatures by plasma enhanced chemical vapor deposition (PECVD) [60]. A methane/hydrogen mixture is utilized to form plasma. Plasma is applied at a pressure of 2 mbar at a process temperature of 450°C. The

substrate with graphene film is then annealed at 900°C and rapidly cooled. This method produces at both top of nickel film and at interface between Si and nickel film. They also observed that annealing does not seem to enhance the graphene film quality; however, they proposed that annealing could be applied to increase the quality of graphene grown at Si/nickel film interface at low temperature processes.

Lee and coworkers further studied on production of wafer scale, high-quality graphene films on Ni and Cu films under ambient pressure and transfer them onto arbitrary substrates by chemical etching of metal layers and polymer-supported transfer onto arbitrary substrates [61]. They used 3-inch SiO_2/Si wafers with 300 nm-thick Ni or 700 nm-thick Cu. They applied a mixture of hydrogen/helium/methane for Cu and a mixture of hydrogen/argon/methane for Ni for 5 min. The samples were then rapidly cooled down to room temperature. They observed that the average number of graphene layers grown on a Ni film ranged from 3 to 8. On the other hand, the monolayer and bilayer graphene grew on a Cu film. In their study, the polymer supports such as PDMS and thermal-release tapes were used to transfer graphene layers.

In case of flexible electronics, it seems to be difficult to grow graphene directly on plastic surfaces due to low temperature melting point of plastic. In a recent process, mechanical rollers are used to transfer graphene film from a thermal release tape to a PET film at 120°C [62]. With roll-to-roll production of 8-inch wafer scale graphene films, commercial production of graphene-based transparent electrodes could be released in the near future.

3.3.1.2 Epitaxial Growth

In epitaxial growth, the substrate acts as a seed crystal and the deposited film will have the similar crystallographic orientations with respect to the substrate. The most popular epitaxial process makes use of a silicon carbide (SiC)

substrate. For large-scale few-layer graphene production, epitaxial growth on SiC by thermal decomposition is receiving much interest. When SiC substrates are annealed at high temperatures, Si atoms selectively desorbs from the surface and the C atoms are left behind naturally form FLG. Because SiC is a wide-band-gap semiconductor, FLG on SiC can serve as a graphene substrate for electronics applications [63]. The two fundamental problems with thermal decomposition are thickness distribution and understanding how the SiC substrate affects physical properties of FLG. In epitaxial growth method, the substrate increases the possibility of achieving a large area of graphene due to the uniform structure of the substrate crystal.

Emtsev and coworkers proposed a near atmospheric pressure method to grow graphene epitaxially from SiC and compare their results with high-vacuum experiments [64]. The authors grew graphene at 676.4 Torr, increasing the temperature, which is necessary for decomposition of the SiC sample. They observed that growth rate increases; however, the surface of the grown graphene improved. To decrease the growth rate, they applied argon to the sample to slow down the Si atoms leaving the surface. Their near-atmospheric method results in the increase of the mobility by a factor of 2 compared to high-vacuum experiments. This method seems to be promising due to importance of growth rate control in epitaxial methods.

3.3.1.3 Chemical Techniques

Chemical exfoliation of graphite to form graphene monolayer has been a promising method for mass production. The starting point for the fabrication of reduced graphene oxide (rGO) thin films is the oxidation of graphite. For electronic and optoelectronic applications, GO should be reduced to become electrically conductive. Synthesis of graphite oxide can be achieved by placing graphite in concentrated acid. Hummers proposed a less dangerous process of graphite oxidation with a mixture of sodium nitrate, potassium permanganate

and sulfuric acid [65]. Individual sheets of GO can be viewed as graphene with oxygen functional groups on both sides of the flakes. Exfoliation of graphite oxide into individual sheets can be obtained by ultrasonic agitation [66]. Thermal reduction of GO is typically achieved above 200°C in inert atmosphere and becomes more efficient at higher temperatures. GO sheets can be deposited on any surface using techniques such as drop-casting, dip-coating, spraying, spin coating and electrophoresis [67]. Uniformity, thickness, surface coverage and number of layers of GO film are dependent on the deposition technique and the deposition parameters. For instance, dip-coating, drop-casting and spraying result in non-uniform film deposition and thickness may not be controllable in these methods. On the other hand, for spin-coating technique, GO suspensions with high concentrations should be utilized to obtain uniform films. N_2 gun should be applied during spin-coating procedure to facilitate solvent evaporation [68].

3.3.2 Graphene in Flash Memory Applications

Graphene as charge storage medium has potential applications in flash memory technology because of the fascinating intrinsic properties such as high work function, high density of states and low dimensionality. The use of graphene in charge-trapping memories is regarded as a solution to increase charge storage capability and reduce device dimensions.

The charge trapping property of graphene oxide (GO) was reported [69]. The solution-processable GO monolayers are sandwiched between a SiO_2 layer as tunnel barrier and Al_2O_3 as control barrier in TANOS ($TaN/Al_2O_3/GO/SiO_2/p\text{-}Si$) structure. The production of GO sheets is based on modified Hummer's method and the solution is spin-coated on top of the tunnel layer. The TANOS structure with GO layer exhibits a wide memory window up o 7.5 V at the sweep range of -5V/14V. In addition, the memory window is reduced to 1.4 V after thermal treatment.

Non-volatile devices with gold nanoparticles and reduced GO were reported [70]. Reduced GO layer is used as a channel and gold nanoparticles are covalently bounded to the channel with a molecular linker. This molecular linker forms an energy barrier between the reduced GO and the channel. Au NPs with reduced GO memory device shows nonlinear hysteresis behavior, stable write/multiple read/erase/multiple read cycles and a retention rate of 700 s.

Hang et al. demonstrated the advantages of graphene as a charge-trapping medium in flash memories [71]. Single layer and multilayer graphene sheets are grown by CVD method and transferred onto SiO_2 tunnel oxide layer. Al_2O_3 layer grown by atomic layer deposition (ALD) is used for blocking oxide. Single layer graphene device exhibits a memory window of ~2 V at sweep range of $\pm$7 V; on the other hand, memory window width is around 6 V for multilayer graphene flash. In addition, the graphene flash memory provides a long retention rate of 8% charge loss after 10 years.

3.4 ALD Basics

Atomic layer deposition (ALD) is a thin film deposition method based on surface reactions of subsequently pulsed source vapors. In ALD, vapor sources are separated by evacuation or purge periods. Chemical reactions are similar to CVD method, but the separation of precursor materials during the reaction is the main difference from CVD technique. ALD achieves atomic-scale deposition control due to self-limited growth property.

ALD was first introduced in the late 1970s by Suntab et al. with the motivation to produce thin film electroluminescent (TFEL) flat panel display [73-76]. Since mid 1990s, the interest towards ALD has been increasing due to the need for thin and conformal films, which is a direct consequence of scaled device dimensions of IC technology. Recently, ALD has been a widely-used technique for the growth of high-k materials. In this study, HfO_2 as tunnel oxide

and Al_2O_3 as control oxide are deposited by Cambridge Nanotech Savannah S100.

ALD deposition cycle for an ideal growth is demonstrated in Figure 3.6. Film growth by ALD appears in a cyclic manner with four main steps per cycle, which are

1. exposure of the first precursor into the process chamber
2. purge or evacuation
3. exposure of the second precursor into the process chamber
4. purge or evacuation

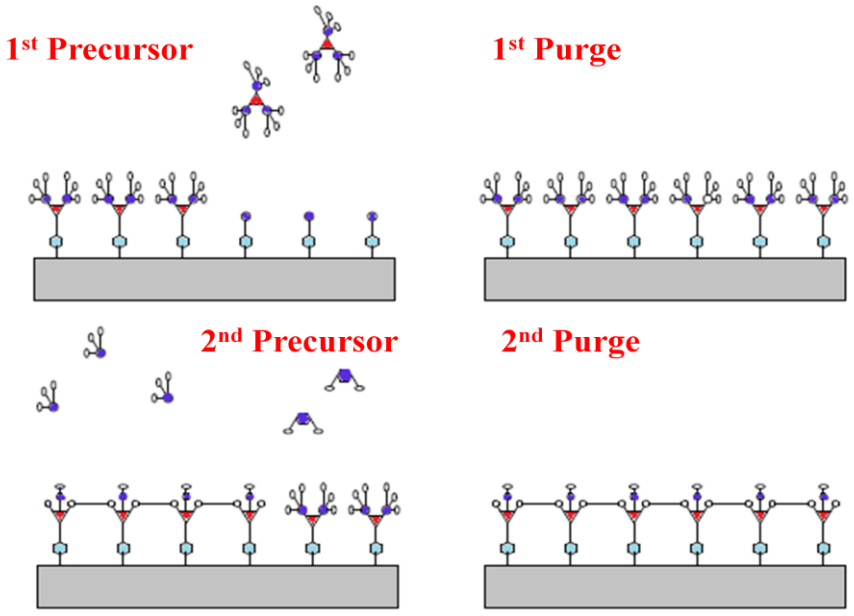

Figure 3.6: ALD deposition cycle for an ideal growth

The thickness of deposited films can be precisely controlled with the number of cycles since each cycle grows exactly one monolayer of film due to self-limiting property of ALD. In addition, self-limiting growth mechanism guarantees the perfect conformality because it is not necessary for the precursor

flux to be uniformly diffused over the sample. The flux should be large enough to saturate the chemisorption layer. Moreover, it is also possible to deposit films subsequently and create multilayer structures. Low-temperature deposition is another advantage of ALD. Plasma-enhanced ALD can grow materials at lower temperatures than thermal ALD. Low-temperature deposition is significant for especially for polymers. For instance, Al_2O_3 can be grown with TMA precursor and O_2 plasma by plasma-enhanced ALD at room temperature [77]. In addition, ALD-grown thin films are very continuous and pinhole-free. This property has great significance to reduce the leakage current through scaled gate dielectrics.

One major disadvantage of ALD is the slow deposition rates. For example, the growth rate of HfO_2 with Savannah reactor is closed to 0.1Å /cycle. The second is the variety of materials that can be deposited by ALD is still less compared other widely-used methods such as MBE and CVD. The nucleation of ALD has great importance to grow defect-free thin films. If the precursor cannot react with the initial substrate, the film may not nucleate uniformly. This lack of nucleation is a very serious problem for ultra-thin gate oxides. Recently, much research has been focusing on the nucleation issue of high-k dielectrics on hydrogen-passivated Si substrate. The nucleation of ZrO_2, Al_2O_3 and HfO_2 was reported [78]. The nucleation difficulty was also observed during the deposition of Al_2O_3 on carbon nanotubes and graphene sheets [79-81]. Al_2O_3 was grown only along the step edges of graphene sheets. Chemical treatment with perylenetetracarboxylic acid and ozone treatment was tried to functionalize the graphene surface [82, 83].

Chapter 4

Device Fabrication and Characterization Methods

4.1 Device Fabrication

The fabrication of flash memory devices with SiN and graphene as charge storage nodes is compatible with the modern semiconductor manufacturing. The fabrication process is composed of five main parts, which are wafer cleaning, back contact deposition, active area definition, gate stack formation and gate contact definition. All fabrication processes are conducted in Class100 clean room facility at UNAM. Characterization methods include spectroscopic ellipsometry, atomic force microscopy (AFM), scanning electron microscopy (SEM) and semiconductor parameter analyzer (SPA). Figure 4.1 illustrates the schematics of the fabricated flash memory capacitors.

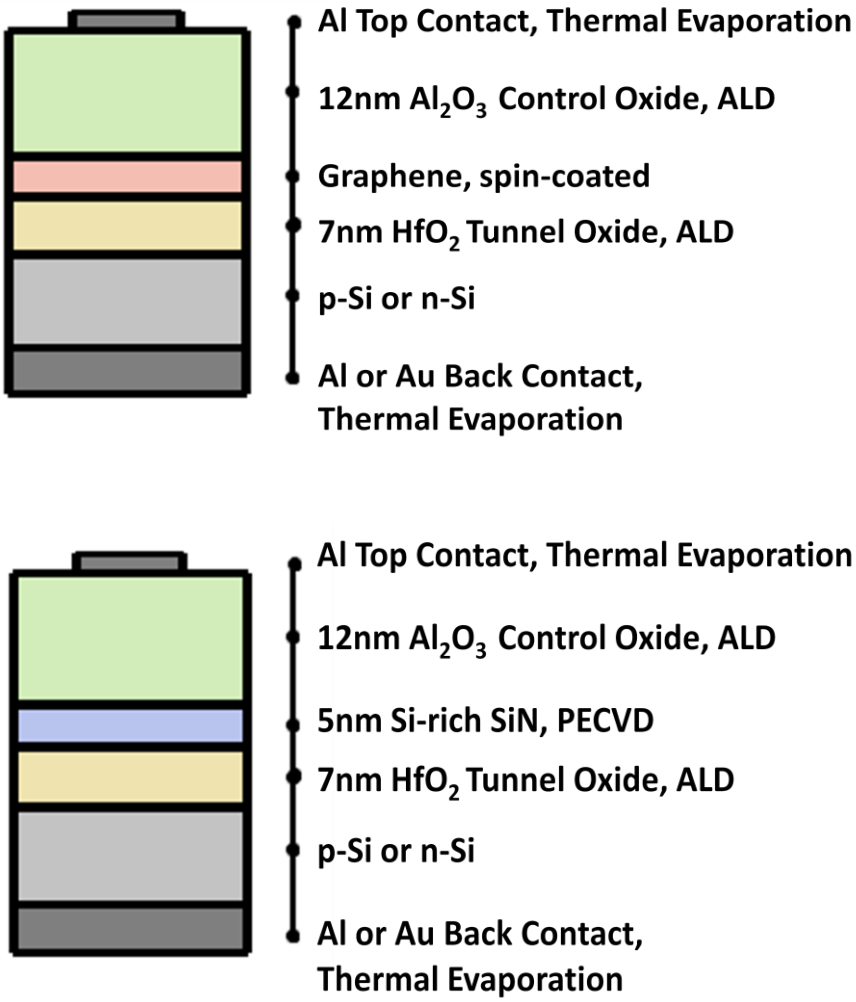

Figure 4.1: Schematics of the fabricated flash memory capacitors.

4.1.1 Wafer Cleaning

Before device fabrication, organic and metal contaminations should be removed off the silicon wafer since contaminants may affect electrical properties of fabricated devices. Both p-type and n-type (100) Si wafers with a resistivity of 1-10 Ω.cm are used for the fabrication. This resistivity range corresponds to a non-degenerate semiconductor with a doping level of 10^{15} - 10^{16} cm^{-3}.

The wafer cleaning consists of three major steps:

1. The wafer is agitated in acetone, methanol and isopropyl alcohol solutions for about 10 minutes to remove dusts and organic particles. Ultrasonic agitation is beneficial to improve efficiency of removal.

2. A mixture of sulfuric acid and hydrogen peroxide, piranha solution, is used to remove off organic contaminants. The wafers are cleaned in H_2SO_4:H_2O_2 (4:1) solution for 20 minutes. Then, the samples are rinsed with DI water and dried with N_2 gun. Piranha clean requires great care since it is highly reactive and easily damages the skin.

3. Since Piranha solution is a strong oxidizing agent, H_2O:HF (95:5) mixture is applied to remove the native oxide layer on silicon surface. The wafers are cleaned in HF solution for 3 min. Then, they are rinsed with DI water and dried with N_2 gun.

4.1.2 Back Contact Deposition

Back contacts are formed by depositing metal layers by VAKSIS PVD Vapor-3S Thermal Evaporation System. Operation principle of thermal evaporator is as follows. The source material is heated to several hundreds of degrees in vacuum. The vapor particles from the source move towards the sample without colliding air molecules and condense to a solid state on the surface. Chamber pressure, substrate temperature, power applied to source material and the distance between the source and the substrate are the main parameters of thermal evaporation systems. Thermal evaporators are widely used for metal contact and masking layer deposition in microfabrication. Compared to sputtering systems, the density and the adhesion of deposited films are lower.

For n-type wafers, the back surface is covered with 5 nm-thick Cr layer and 100-nm-thick Au layer by thermal evaporator. Thin Cr layer is applied to improve Au adhesion on Si surface. For p-type devices, 100nm-thick Al is deposited in a chamber pressure of around 10^{-6} Torr. Material coating rate is kept constant at 0.5-0.6 Å/s to achieve high density deposition.

Tungsten and molybdenum boats are used to place pallets of coating materials. Samples are rotated during deposition to allow uniformity. Cooling is not provided by VAKSIS, thus the temperature of the sample is observed to increase up to 80ºC during coating. Chamber cleaning prior to pumping is significant to avoid the particles remaining on the inner walls of the chamber; otherwise, it requires extended pumping time before reaching the desired chamber pressure.

After metal layer deposition, the samples are annealed by ATV SRO-704 rapid thermal annealing system (RTA) to form ohmic contacts. RTA is a common process in semiconductor technology in which silicon wafers are heated to high temperatures in a few seconds. RTA has many applications on semiconductor manufacturing such as activation of dopants, crystallization, densification of thin film and ohmic contact formation.

Metal-semiconductor contacts with low impedance are essential in modern electronics. Ohmic contacts are formed by doping the Si substrate heavily enough to allow tunneling. Ohmic contact formation includes a high temperature process at which metal atoms diffuse into the substrate and make an alloy with the semiconductor. P-type devices with Al as back contact are heated to 450ºC for 2 min. in inert atmosphere. Similarly, n-type devices with Au layer are annealed at 400ºC for 2 min. annealing process is applied under atmospheric pressure and takes approximately 10 min.

4.1.3 Active Region Definition

This fabrication step includes the deposition of an insulating layer on Si surface and patterning this layer to form an active region for each flash cell.

SiO_x insulating layer is deposited on Si surface by VAKSIS CVD-Handy Plasma-Enhanced Chemical Vapor Deposition (PECVD) System to isolate flash cells from each other. PECVD is a technique in which various thin films are deposited from gas state to solid state on a substrate at lower temperatures compared to chemical vapor deposition (CVD). Plasma is created due to the capacitive coupling between a ground electrode and a RF electrode. Chemical reactions between the gases are initialized after the creation of the plasma, forming the desired material film on the sample. In PECVD technique, the substrate is heated to 250ºC - 350ºC. The lower deposition temperature makes PECVD more advantageous than CVD in many applications. PECVD system has limited capacity and allow individual wafer loading. In addition, it can be easily contaminated and requires frequent cleaning. However, most PECVD-coated films such as Si, SiO_x and SiN_x can be easily cleaned from the electrodes and the chamber walls with a plasma containing fluorine. Frequent cleaning of chamber has great importance to prevent contamination from deposited films. Compared to physical deposition systems, PECVD reaches faster deposition rates while maintaining conformality and uniformity.

Argon is used to dilute the plasma, which is crucial to achieve uniformity. Low deposition rate is achieved by lowering gas flow rates. The thickness at the center is measured as 204 nm by spectroscopic ellipsometry and it gradually increases through the edge, reaching around 212 nm.

Active regions are created by photolithography and wet etching with buffered oxide solution (BOE). The photomask with a minimum feature size of 150 μm is designed by Layout Editor Software and produced by Heidelberg

Instruments DWL-66 Mask Writer. Prior to photolithography, samples are heated at 110ºC on a hot plate for 5 min. to remove humidity from the surface. Photoresist adhesion is improved by spin-coated HMDS (hexamethyldisilazane) at 5000 rpm for 40 s. Then, AZ5214E photoresist is spin coated at 4000 rpm for 50 s., achieving a thickness of 1.4 μm. The samples are baked at 110ºC for 60 s to decrease the solvent concentration of the photoresist. After exposure, a mixture of AZ400K:H_2O (1:4) is used to pattern the active regions on the surface. Photolithography process is finished with post-exposure bake. The post-exposure bake is applied at 120ºC for 20 s. to increase the physical stability of the resist for chemical etching process. After photolithography, active regions are formed with wet etching of SiO_x layer in BOE solution for approximately 3 min. During etching process, the color of the sample is frequently observed with an optical microscopy to prevent over-etching.

4.1.4 Gate Stack Formation

Gate stack is composed of a charge storage medium sandwiched between HfO_2 as tunnel barrier and Al_2O_3 as control barrier. For graphene flash memories, chemically-synthesized graphene sheets are used as charge storage medium. For ONO-type flash memories, PECVD-grown Si-rich nitride layer is utilized for charge trapping. Gate stack formation includes deposition of graphene flakes, growth of nitride layer and deposition of tunnel layer and control layer by ALD.

Water-soluble graphene oxide is synthesized from natural graphite by Hummers Method [65]. Based on Hummers Method, graphene oxide sheets are derived from the acid exfoliation of natural graphite (SP-1, Bay Carbon). A mixture of sodium nitrate, potassium permanganete and sulfuric acid is applied to achieve acid oxidation. The graphite oxide is exfoliated into individual sheets by ultrasonic agitation for 60 min. Unexfoliated graphene oxide particles are completely eliminated by centrifuging the suspension sequentially at 8000 rpm

and 14000 rpm for 20 min each. Graphene flakes are obtained by thermally annealing the spin-coated graphene oxide layer.

The first process for gate stack is to grow 7 nm-thick HfO_2 as tunnel oxide. 69 cycles of HfO_2 layer is deposited at 200ºC by Cambridge Nanotech Savannah thermal ALD system. For graphene flash memories, graphene oxide solution is spin-coated onto tunnel oxide layer at 300 rpm for 50 s. Thermal reduction of graphene oxide sheets is achieved at 300ºC for 30 min in ALD chamber. Prior to Al_2O_3 deposition, 50 cycles of TMA is deposited at 300ºC for functionalization of graphene surface. It was reported that metal oxides cannot be grown directly on graphene sheets by ALD since graphene surface is lack of dangling bonds and functional groups [84]. Surface functionalization is required to deposit defect-free thin films on graphene layers. After surface functionalization, 110 cycles of Al_2O_3 as blocking oxide is grown at 300ºC by ALD. N_2 with a flow rate of 20 sccm is applied to ALD chamber as a carrier gas. Deposition conditions of HfO_2 and Al_2O_3 are summarized in Table 4.1.

Table 4.1: ALD Deposition Conditions

Film	Precursor #1	Precursor #2	Growth Temp.	Growth Rate	# of cycles
HfO_2	$Hf(NMe_2)_4$	H_2O	200°C	1.08 Å/cycle	69
Al_2O_3	TMA	H_2O	300°C	1.01 Å/cycle	110

SiN flash memories contain Si-rich nitride layer as charge-trapping medium. SiN layer is deposited by VAKSIS CVD-Handy PECVD System. Uniformity is achieved by applying helium into the chamber to dilute the atmosphere. It is difficult to obtain stoichiometric ratio in PECVD-deposited films because of the complexity of the chemical reactions in the plasma. Therefore, the control of gas flow ratio is crucial on nitride film properties. The SiH_4/NH_3 ratio is kept constant as (4sccm/200sccm) at 250°C under 0.6 Torr chamber pressure.

Thickness of nitride layer is measured as 5.4 nm by ellipsometry. In ONO-type flash cells, Si-rich nitride layer is sandwiched between HfO_2 and Al_2O_3, which are deposited by ALD in similar growth conditions as do in graphene flash cells.

4.1.5 Gate Contact Definition

Gate contacts are formed by lift-off process after deposition of Al layer. Before Al evaporation, photolithography process is applied to pattern the gate electrodes. Details of photolithography steps are given in Section 4.1.3. However, post-deposition bake is not utilized not to harden the resist. After patterning, 100 nm-thick Al layer is deposited by thermal evaporation. Then, samples are left in acetone for 20 min to complete the lift-off process.

4.2 Characterization Methods

4.2.1 Spectroscopic Ellipsometry

Spectroscopic ellipsometer is a machine to measure the refractive index and thickness of films. The working principle is based on the change of polarization of light due to the reflection at the surface and the change of phase of the incoming light when the light transmits through the transparent parts of the sample. The polarization change is represented as an amplitude component, Ψ and the phase difference, Δ.

The spectroscopic ellipsometry allows film characterization with a thickness ranging from a few angstroms to several micrometers with superior precision. It is widely used in many different fields such as microelectronics, materials science and biology. It can yield information about various properties of thin films, including morphology, electrical conductivity, crystal quality and chemical composition.

V-Vase Ellipsometry is used to determine the thickness of ALD-grown HfO_2 and Al_2O_3, and PECVD-deposited SiN layer. Optical constants of Si-rich nitride are also measured to determine the charge-trapping property of nitride films in ONO stack.

4.2.2 Atomic Force Microscopy

Atomic force microscopy (AFM) is one of the most widely used instruments for imaging, manipulation and measuring the sample at nano-scale. Information is gathered by scanning the surface with a sharp tip and measuring the forces between the tip and the sample.

There are many advantages of AFM over conventional microscopy techniques. First, AFM provides a three-dimensional image of a sample with atomic resolution. AFM does not require a vacuum condition. Moreover, it can be used in a liquid environment and sample preparation is not necessary before characterization.

AFM finds great number of applications in the field of materials science, semiconductor manufacturing, chemistry and biology. Scanning speed and scan image size are regarded as the main drawbacks of AFM compared with scanning electron microscopy (SEM).

Asylum MFP-30 AFM at UNAM is used to obtain topographical and surface potential data of graphene flakes in graphene flash memory. A conventional NA-NC polysilicon cantilever with a force constant of 9.5 N/m is driven in the repulsive mode.

4.2.3 Semiconductor Parameter Analyzer

Capacitance-voltage measurements of SONOS-type and graphene flash memories are conducted with Keithley Model 4200 Semiconductor Characterization System. The measurement setup also includes a manual probe station, Cascade PM-5, where the sample is mounted o a vacuum chuck and microneedles are gently placed on bottom and top electrodes to achieve electrical connection. Coaxial cables are used for the connection to avoid environmental noise.

The semiconductor parameter analyzer (SPA) includes Model 4200-CVU card, which is an impedance measurement card at frequency range from 10 kHz to 10 MHz. The impedance measurement is conducted by sourcing AC signal to device terminals, then measuring the resulting AC current and the phase difference. Capacitance and conductance are extracted from measured impedance and phase difference.

Before the measurements, connection compensation data should be generated for open correction, short correction and load correction to avoid gain errors and parasitic effects caused by the electrical connection between the device under test (DUT) and the SPA.

The performance of the memory devices is determined with high frequency capacitance-voltage (C-V) measurement. All measurements are conducted in an ambient environment without light illumination to eliminate minority carrier generation by light.

Chapter 5

Results and Discussion

The fabrication of flash cells with Si-nitride and graphene sheets as charge trapping layer is given in the previous chapter. In this chapter, electrical characterization of graphene and SiN flash cells will be presented. Charge storage property of graphene is compared with ONO-type flash cells with Si-nitride as storage medium. C-V measurements are performed for both electron and hole storage. This chapter also includes the structural characterization of graphene flakes by SEM and AFM.

5.1 Characterization of Graphene

5.1.1 SEM

Characterization of graphene flakes is performed by SEM. Graphene oxide suspension is spin-coated onto SiO_2/Si wafer at 300 rpm for 50 s. SEM images of graphene oxide layer can be observed in Figure 5.1. It is clearly shown that the surface is not uniformly covered with graphene flakes. Graphene layer is composed of monolayer, bi-layer and few-layer graphene nanosheets. In addition, SEM images confirm that individual sheets are in the range of hundreds of nanometers. Chemical exfoliation of graphene is a simple and

inexpensive method to deposit graphene flakes in large dimensions. Uniform and continuous films can be achieved by altering the density of graphene oxide suspension and spin rate.

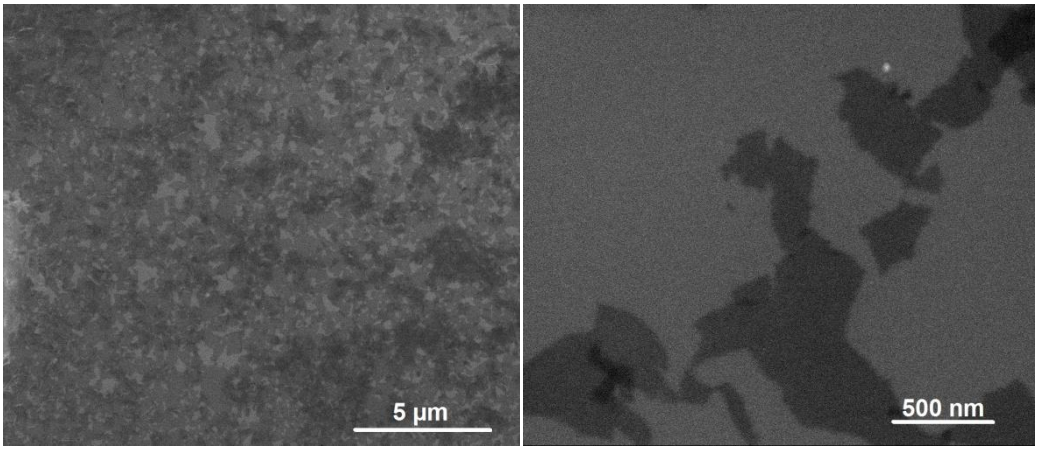

Figure 5.1: SEM images of graphene flakes on a Si substrate

5.1.2 AFM

Topographical and surface potential data of graphene flakes are obtained simultaneously by Asylum MFP-30 AFM, given in Figure 5.2. The used nanoprobe is commercial HA_NC polysilicon cantilever of NT MDT company which is of force constant 9.5 N/m. The probe is coated with 10 nm of Pt for conductance necessary for surface potential data. The measured cantilever resonance frequency is 235.9 kHz. Back gate contact of the sample was DC biased, and AC bias was applied to the cantilever. DC bias was provided with AFM electronics, whereas AC bias of frequency 25 kHz was provided with the Stanford Research Systems DS345 Function Generator. The cantilever was driven in the repulsive mode. To measure surface potential data, the deflection signal connected to the SR830 lock-in amplifier which is locked to 25 kHz. The topographic height of graphene layer is determined as 2 nm, approving the existence of monolayer and few-layer sheets.

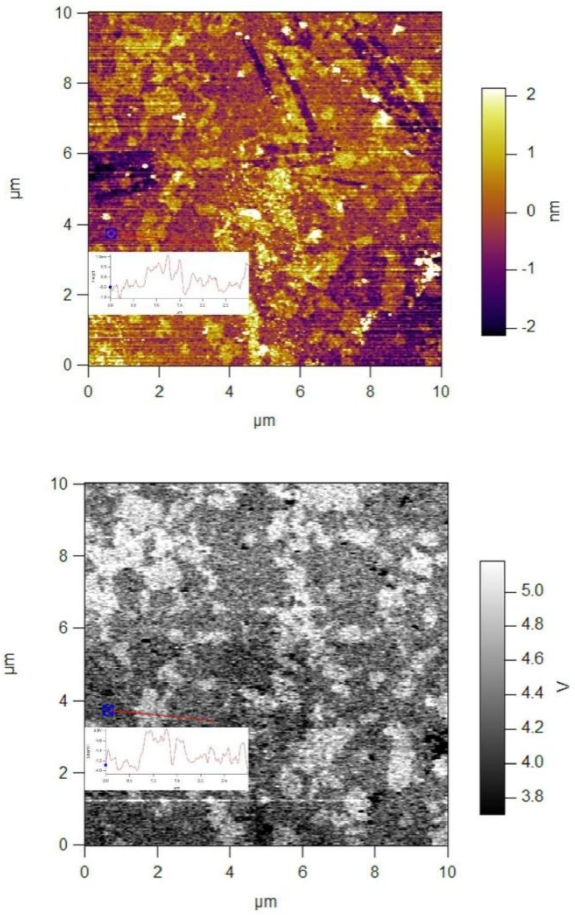

Figure 5.2: AFM images of graphene flakes.

5.1.3 Raman

The Raman spectrum of exfoliated graphene oxide is illustrated in Figure 5.3. The sample for Raman spectroscopy is prepared by spin-coating the graphene oxide suspension on a Si substrate with a 120nm-thick Al layer on top of it. The measurement is performed at room temperature with Witec Raman Module. The raman spectrum of graphene oxide shows a D peak at 1347 cm^{-1} and a G peak at 1597 cm^{-1} [84].

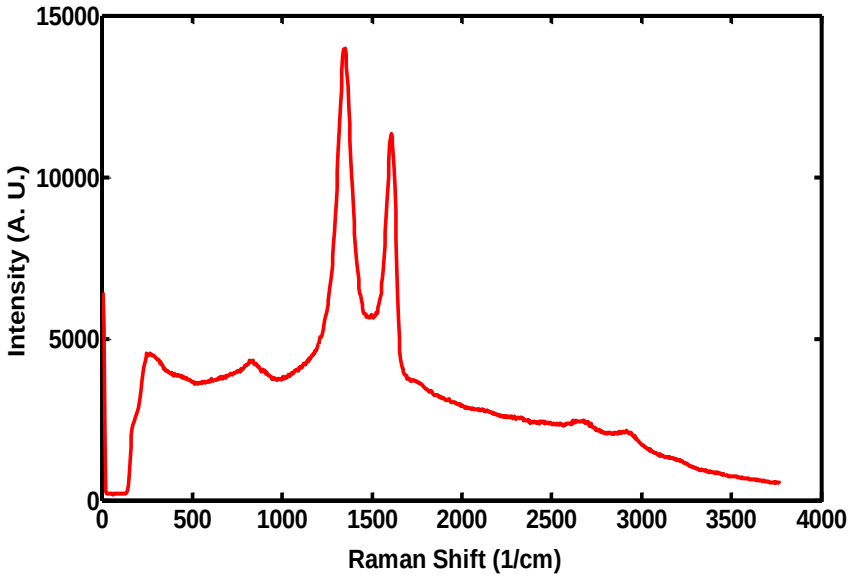

Figure 5.3: Raman spectrum of exfoliated graphene oxide

5.2 Electrical Characterization of Control Samples

To understand the storage property of both graphene and Si-nitride layers, p-type and n-type control samples with HfO_2 as the tunnel oxide and Al_2O_3 as the control oxide are fabricated. C-V measurements are conducted at frequency of 1 MHz and AC signal amplitude of 25 mV_{rms}. Large hysteresis in C-V curve is a strong indicative of charging/discharging in MOS capacitors. As observed in Figure 5.4 and Figure 5.5, C-V characterization of control samples exhibit negligible hysteresis at a voltage range of ± 6 V, approving that high-k dielectric layers and interface traps have negligible effect on charge storage in flash cells. In other words, graphene and Si-nitride layers will be responsible for charge storage in this study.

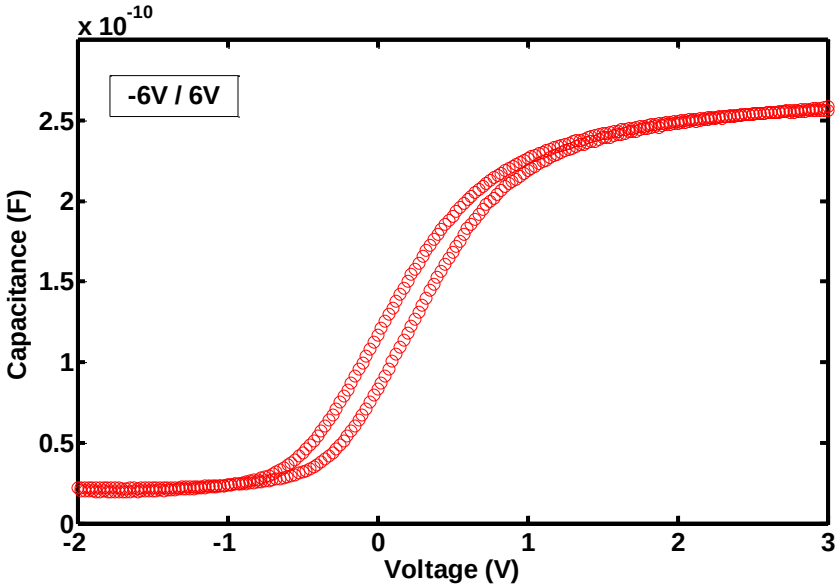

Figure 5.4: C-V characteristics of n-type control sample

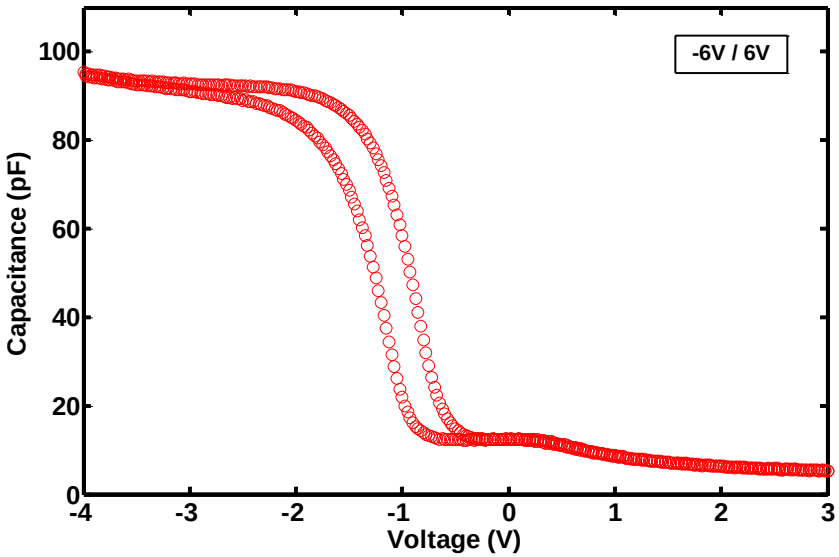

Figure 5.5: C-V characteristics of p-type control sample

5.3 Surface Functionalization of Graphene

Recently there has been much research conducted on the deposition of high-k dielectrics on top of graphene flakes. Dielectric films deposited on graphene surface should have high uniformity with a pinhole-free structure, high breakdown voltage and low leakage current at nano-scale thicknesses. ALD system seems to be a possible solution to deposit ultrathin homogenous films on top of graphene layer. Wang et al. tried to grow a 2 nm-thick Al_2O_3 layer on mechanically exfoliated graphene sheets using TMA and H_2O as ALD precursors at 100ºC [83]. They concluded that direct deposition of metal oxides on graphene surface is not possible due to nonexistence of dangling bonds. Oxide growth can be initiated on the edge and defect sites, where dangling bonds and functional groups enhance the possibility of nucleation in ALD deposition. Surface functionalization of graphene is crucial before ALD deposition to nucleate continuous and uniform growth of thin high-k dielectrics. Many efforts have been carried out in the field of high-k/graphene interaction including deposition of seed layer with Al and polymer, pretreatment with ozone and fluorine [85-89]. These attempts are based on to create hydrophilic characteristics on graphene surface with reactive species such as hydroxyl groups. These functional groups have great importance to improve chemical reaction between graphene surface and ALD precursors.

In this study, graphene surface is functionalized with TMA precursor before Al_2O_3 growth. A graphene flash memory sample without surface functionalization is fabricated to observe the importance of surface pretreatment before ALD deposition. In this sample, Al_2O_3 layer is directly deposited on top of graphene surface with TMA and H_2O as precursors at 300ºC. For the sample with surface functionalization, 50 cycles of TMA precursor is deposited before Al_2O_3 deposition to create dangling bonds and functional groups on the graphene surface. Figure 5.6 shows the C-V curves of the samples with/without surface functionalization at voltage sweep range of ± 4V. The C-V characteristics of the

sample without surface functionalization exhibit a narrow memory hysteresis with a significant corruption in accumulation region. This corruption indicates non-uniform and discontinuous growth of Al_2O_3 layer on graphene surface. The graphene layer is partially adhered to the control oxide due to hydrophobic characteristics of graphene surface; therefore, stored charges can easily tunnel through the gate metal. On the other hand, the sample with surface functionalization exhibits a large hysteresis window, confirming that Al_2O_3 layer is continuous and can avoid charge leakage through the gate contact.

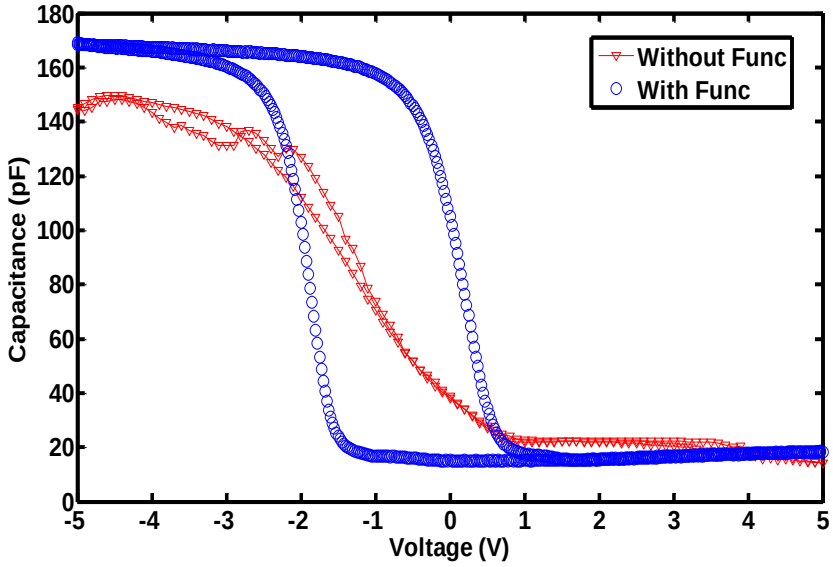

Figure 5.6: C-V characteristics of the samples with/without surface functionalization

5.4 Electrical Characterization of p-type Graphene Flash Memories

Charging/discharging performance of p-type flash cells with graphene as charge storage medium is obtained by taking high frequency C-V measurements with successively increasing voltage sweep range. As illustrated in Figure 5.7, the memory cell exhibits large hysteresis window of 5 V in sweep range of ±8 V. The hysteresis window widens toward negative direction, proving that hole

trapping is the leading mechanism. Both electron and hole charging begins at ±3 V. Compared to n-type graphene flash devices, charge trapping/detrapping begins at larger voltage range because of the fact that holes face higher band offset while tunneling through HfO_2 as tunneling oxide.

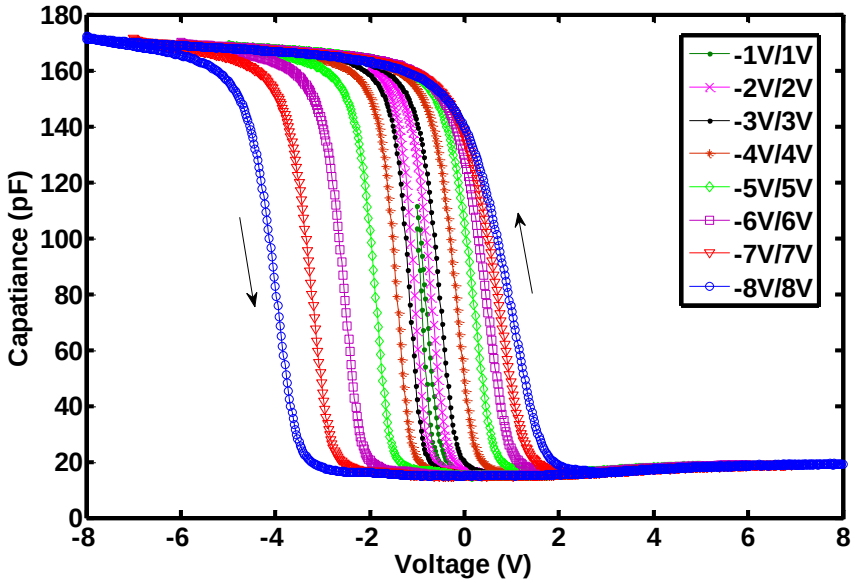

Figure 5.7: Hysteresis window of p-type graphene flash memory device

Hole storage performance of p-type graphene flash cells for varying gate voltage amplitudes is given in Figure 5.8. Flat band voltage shift is determined while pulse duration of the gate bias is kept constant at 0.5s and its amplitude is varied from 0V to -7 V. Shift in flat band voltage increases and hole storage become significant after applied voltage exceeds -2.5 V. A flat band voltage shift of -0.77 V can be achieved at programming voltage of -5 V. This result confirms that p-type graphene flash cells exhibit great memory performance at low programming voltages.

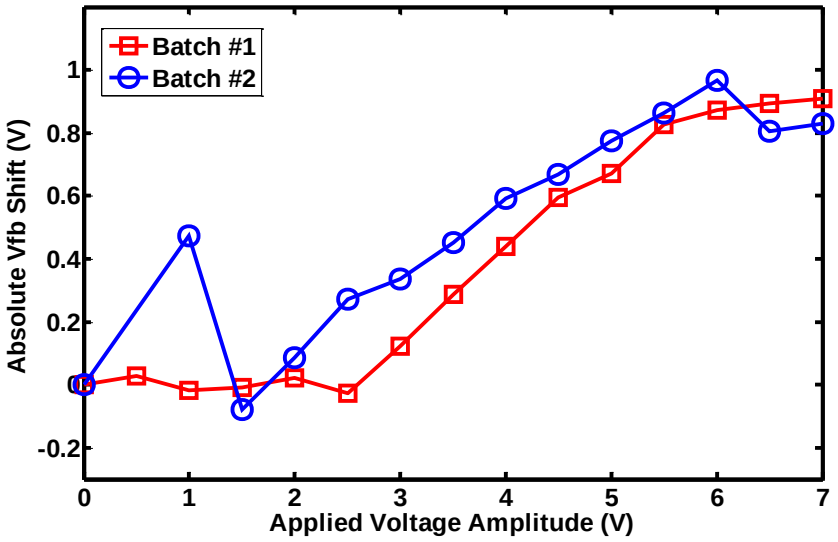

Figure 5.8: Hole storage performance of graphene flash devices for varying bias voltages

Retention rate of p-type graphene flash structures for hole storage is given in Figure 5.9. Uncharged p-type graphene cells are programmed at -5 V for 0.5 s. Retention characteristics is determined by observing flat band voltage shift. The measurements take more than 12 hours and the data is extrapolated up to 10^8 s (10 years). After programming two identical cells in a single sample, a flat band shift is found as around 0.6 V. According to the results, initial charge emission seems to be faster because of discharging of holes in shallow traps and interface states. After a certain period of time, charge loss rate has decreased significantly. Extrapolated data shows that 50% of trapped charge will be lost after ten years, approving that p-type graphene flash devices for hole storage exhibit enhanced retention rate.

Memory performance of graphene flash memories based on programming pulse duration is illustrated in Figure 5.10. The cells are programmed at -5 V and flat band voltage shift is determined as a function of pulse duration from 100 ns to 0.5 s. It is clearly shown that four separated cells exhibit similar memory performance for varying pulse duration. Flat band shift of 0.2 V - 0.3 V is obtained at a pulse duration of 100 ns and hole trapping increases with the

increase of pulse duration. Non-uniform deposition of graphene nanosheets causes a slight difference in the maximum shift obtained at 0.5 s.

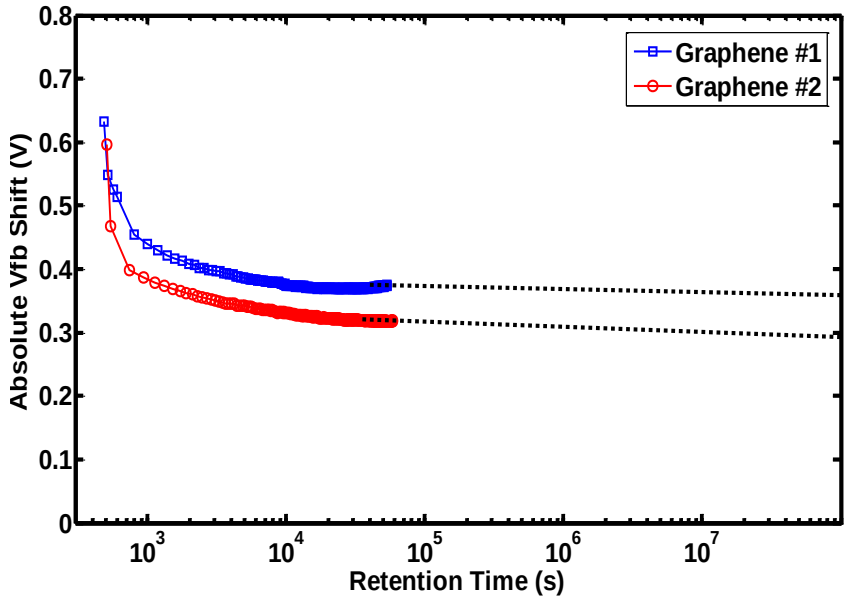

Figure 5.9: Retention characteristics of p-type graphene flash devices

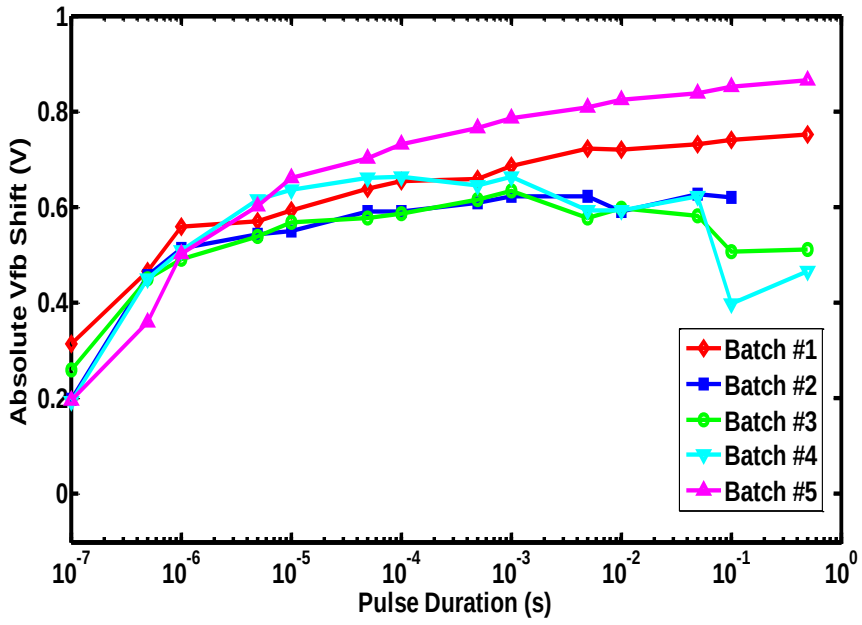

Figure 5.10: Program transient characteristics of p-type graphene flash devices

Memory window of p-type graphene flash cells is obtained after many write/erase cycles, shown in Figure 5.11. The cell is programmed/erased at -5V/5V for 0.1 s and C-V measurements are repeated for 30 times. The memory window of 1.1V is obtained after the first cycle. The memory window is observed to decrease to 0.8 V after 30 cycles. Oxide defects may be the reason of memory window degradation. After many write/erase cycles, charge injection through tunnel oxide may cause forming oxide defects and narrowing of the memory window. This result has a great importance to confirm that p-type graphene flash cells can be written/erased with low bias voltages and short pulses.

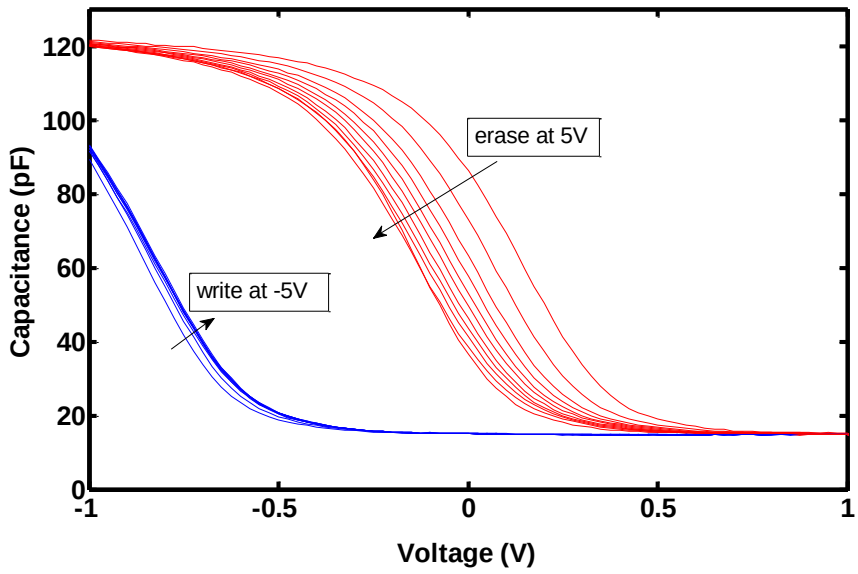

Figure 5.11: Memory transfer characteristics of p-type graphene falsh device under written and erased conditions.

5.5 Electrical Characterization of n-type Graphene Flash Memories

Voltage sweep range has a great importance on memory effect since tunneling rate is exponentially depended on applied bias. C-V measurements of n-type

graphene flash memory cells are taken to determine the charging/discharging ability and the memory window. C-V curves on n-type graphene cells with successively increasing range from ±1 V to ±6 V are shown in Figure 5.12. All measurements are taken at a frequency of 1 MHz and it is concluded that CV hysteresis is strongly dependent on the maximum voltage limit. Large memory window is obtained at low voltage ranges, indicating that the usage of graphene as charge trapping medium improves the charge storage property of flash memories at low operation voltages.

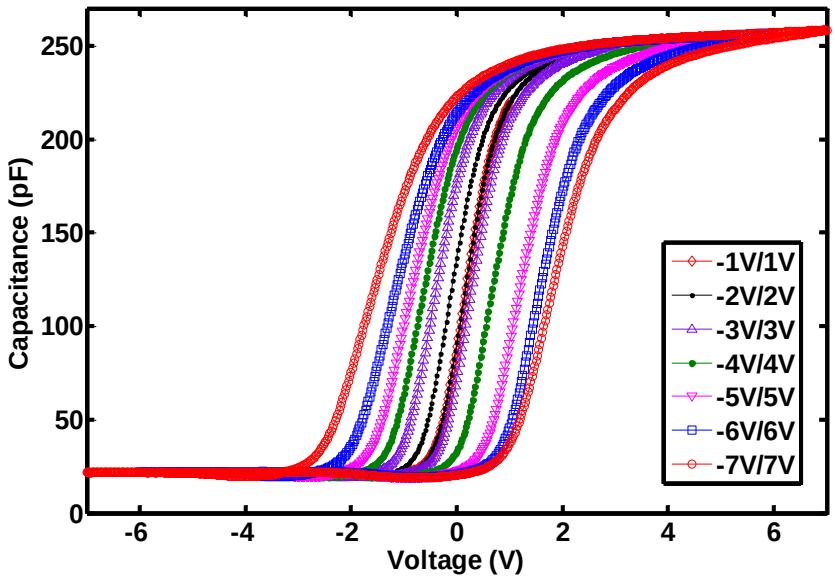

Figure 5.12: Hysteresis window of n-type graphene flash memory device

Retention performance of graphene flash cells for electron storage is given in Figure 5.13. Retention performance is determined by observing change in flat band voltage after programming the sample. The flat band voltage shift is estimated with respect to the C-V curve of the uncharged sample. The flash cell is programmed at 5 V for 0.5 s and subsequent C-V measurements are taken for more than 10 hours at room temperature. It is estimated that the cell Batch #1 will lose 34% of the total stored charge after 10 years. Charge loss at the beginning of retention seems to be faster because of trapped charges at interface states. Moreover, maximum flat band shifts obtained from two separated cells

are different due to non-uniform deposition of graphene sheets. The retention rate of the graphene flash cell exhibits that graphene layer as charge storage medium has great capability in retaining the stored charges over long periods of time.

The dependency of memory effects as a function of pulse duration is shown in Figure 5.14. Bias amplitude is kept constant at 5 V for electron storage and pulse width is varied from 100 ns to 0.5 s. It is clearly shown that electron injection is significant even when a gate bias with a pulse width of 100 ns is applied to the n-type graphene flash structure. High charge storage capability of graphene is the main reason for such a fast programming with a low voltage bias. Reverse flat band shift is observed for Batch #1 in program transient characteristics. A probable reason for this effect could be the transfer of oxygen atoms between two graphene flakes.

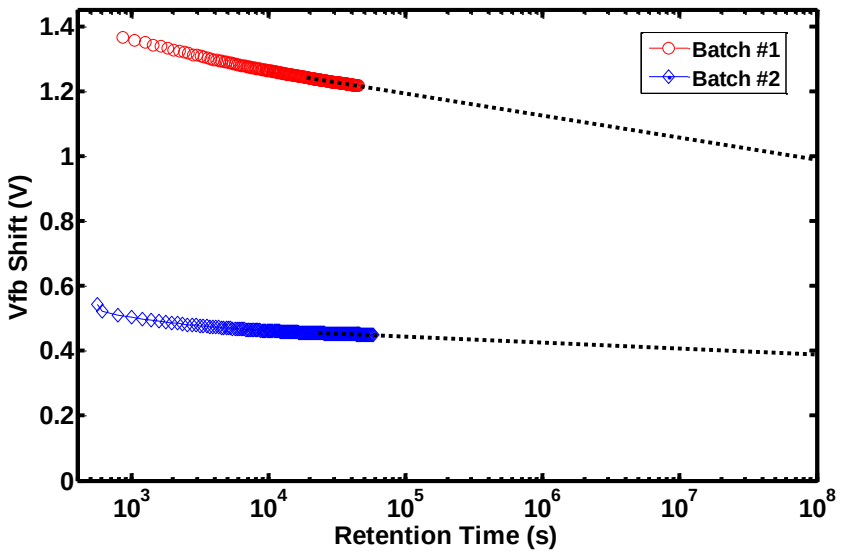

Figure 5.13: Retention characteristics of n-type graphene flash devices.

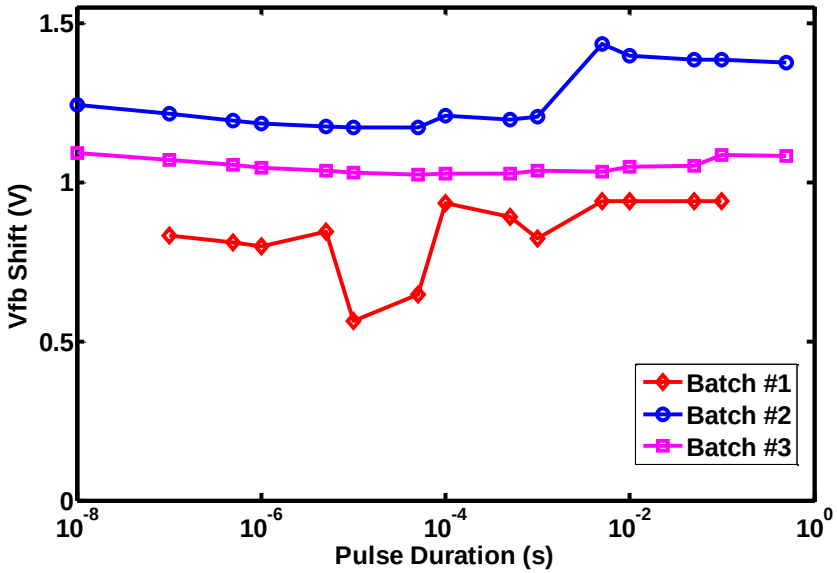

Figure 5.14: Program transient characteristics of n-type graphene flash devices.

5.6 Electrical Characterization of p-type SiN Flash Memories

Figure 5.15 illustrates the hysteresis window of p-type SiN flash memories. High frequency C-V measurements are taken from -3V/3V to -10V/10V. The hysteresis window widens toward negative direction, confirming that hole storage is the leading mechanism. In addition, it is also observed that charge trapping/detrapping is not significant at low voltages. For a similar p-type memory cell with graphene as charge trapping layer, a hysteresis window of 2 V can be obtained for a voltage range of -5V/5V. On the other hand, this value is less than 1 V for p-type SiN flash structures. This results shows that graphene flash cells exhibit better charge storage property at low operating voltages compared to SiN flash cells.

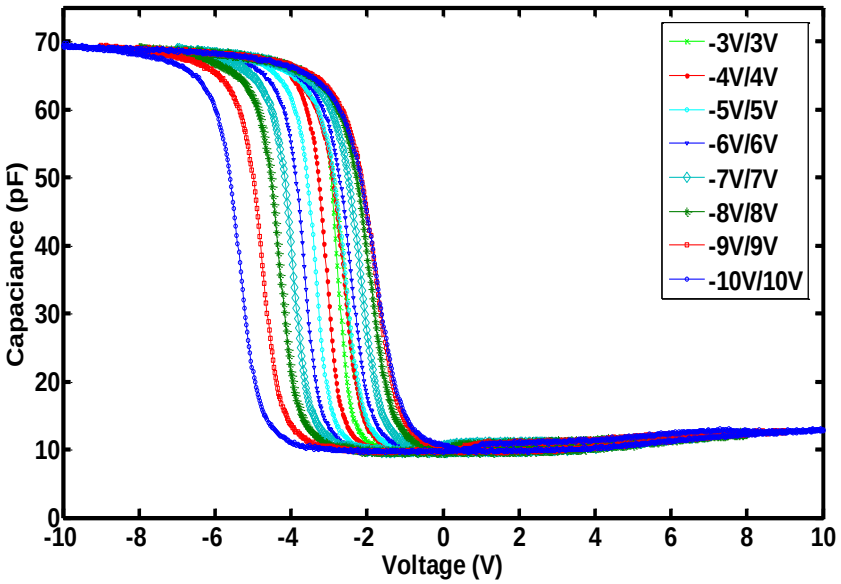

Figure 5.15: Hysteresis window of p-type SiN flash memory device.

The impact of programming voltage on p-type SiN flash memories is given in Figure 5.16. Pulse duration of gate voltage is kept constant as 0.5 s and its amplitude is increased up to 9 V. It is found that hole storage becomes significant when the programming voltage exceeds -5 V. This result is also consistent with the hysteresis window experiment of p-type SiN devices. Flat band voltage shift starts to increase with the applied voltage of -5 V and reaches a maximum shift of 0.92 V when the bias is -8 V. A reverse flat band shift is observed when the voltage bias is equal to -4 V. It is predicted that this reverse shift is not related with either back tunneling of holes from the blocking oxide or discharging through the tunnel oxide. In fact, drift of charge carriers within SiN layer will cause a flat band shift in opposite direction. Based on the flat band shift with programming voltage bias, p-type SiN flash memories are not appropriate for low voltage operations.

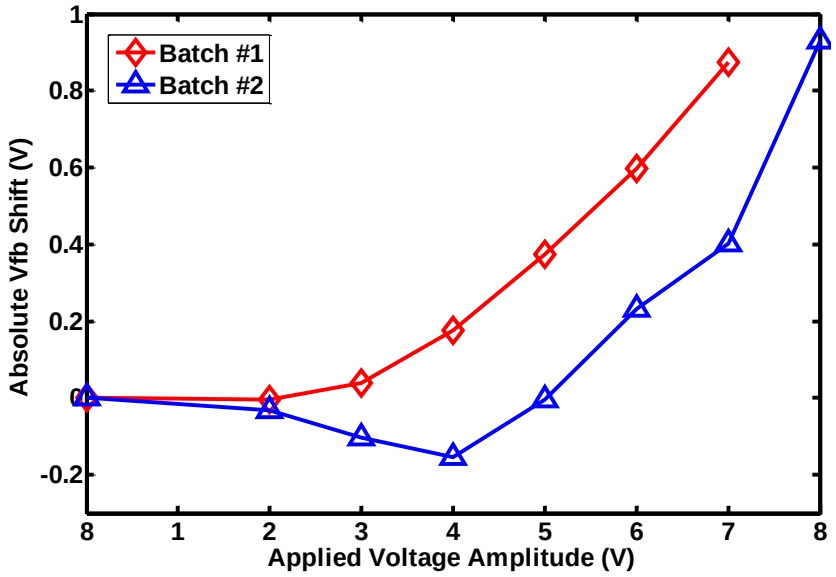

Figure 5.16: Hole storage performance of SiN flash devices for varying bias voltages.

The retention characteristics are shown in Figure 5.17 for a p-type sample with SiN layer as charge trapping medium. High frequency C-V measurements are conducted at zero retention voltage after programming the sample at -7 V for 0.5 s. As mentioned before, bias voltage of -5 V for hole storage is not sufficient for these memory cells. Therefore, -7 V as gate voltage is utilized to make a comparison of retention rate with graphene flash cells for hole storage. A flat band voltage becomes less than -0.6 V after programming the sample. According to the retention characteristics, the cell exhibits fast charge loss. The reason is that shallow traps are the main storage nodes in SiN layer and emission of holes from these traps becomes faster.

Figure 5.18 illustrates the transient program characteristics of p-type SiN flash devices for hole storage. The C-V curves are taken from three separated cells in a single sample at a constant write voltage of -5 V. The measurements are conducted without illumination to avoid generation of charge carriers with light-assisted voltage pulses. Based on the results, it is clear that a continuous increase in flat band shift cannot be obtained with increasing pulse duration. A

bias voltage of -5 V seems not to be sufficient to inject holes into the storage medium. The program transient characteristics of Batch #2 show a continuous increase in flat band shift up to 100 μs; however, it returns its uncharged condition at the following voltage pulse. Based on this behavior, we predict that charges are mainly stored in shallow traps and can be easily detrapped in a few seconds. The magnitude of voltage bias should be increased to store charges in deep traps.

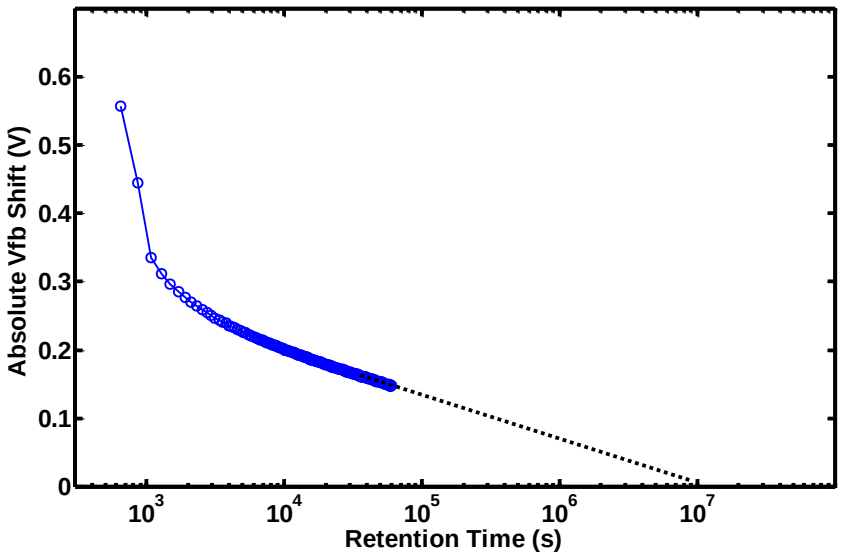

Figure 5.17: Retention characteristics of p-type SiN flash device.

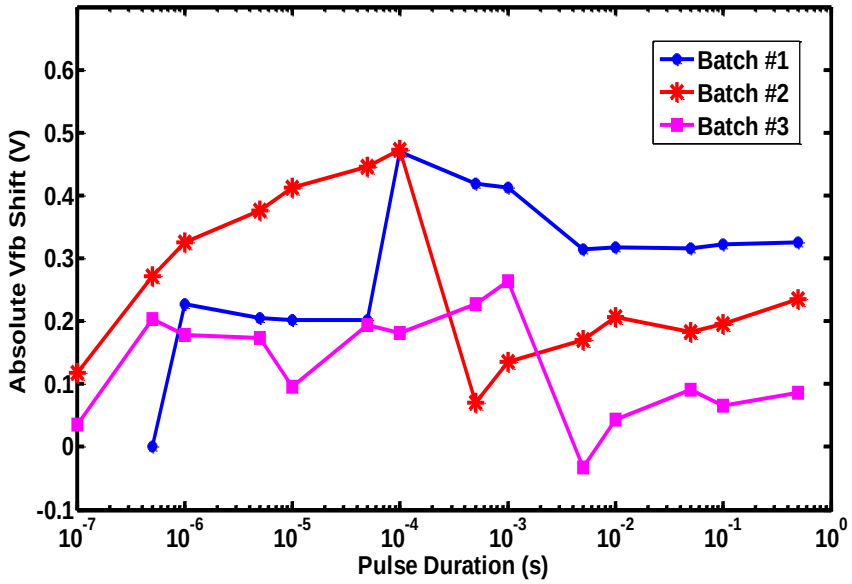

Figure 5.18: Program transient characteristics of p-type SiN flash devices.

5.7 Electrical Characterization of n-type SiN Flash Memories

The high frequency (1 MHz) C-V curves of an n-type SiN flash cell under successively increasing voltage ranges are illustrated in Figure 5.19. According to the results, hysteresis width gets larger as dual-directional sweep range increases. For n-type SiN structure, the C-V characteristics exhibit clockwise hysteresis window around 4 V under a range of -8V/8V. The hysteresis window expands toward positive voltages, confirming that electron trapping is the dominant mechanism. In addition, hole trapping at the inversion state is also observed in the experiment.

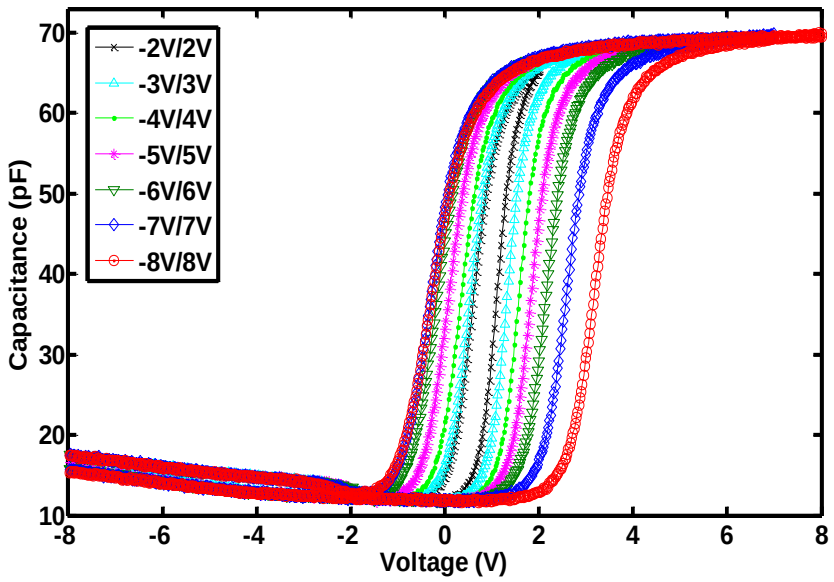

Figure 5.19: Hysteresis window of n-type SiN flash memory device.

Figure 5.20 shows the retention characteristics of n-type flash cell with SiN as charge storage medium. The cell is programmed 5 V for 0.5 s and flat band shift is determined as a function of waiting time at room temperature. Retention rate is obtained under zero gate voltage and extrapolated to 10 years. After programming the device, a flat band shift of 0.82 V is observed. The n-type SiN flash structure exhibits a charge loss of 62% according to extrapolated data. This

significant amount of charge loss proves that electrons mostly stored in shallow traps of SiN layer. These shallow traps result in large emission rates from the beginning of the retention.

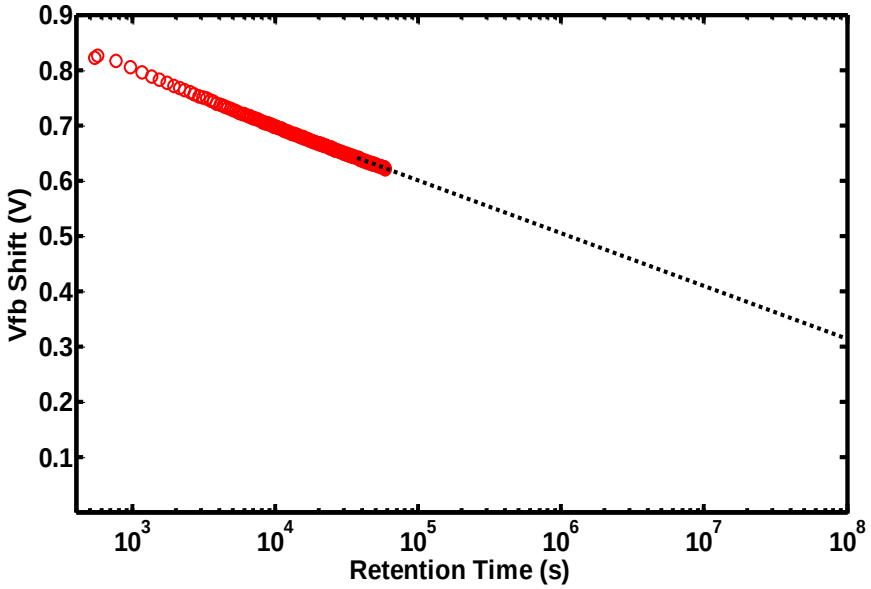

Figure 5.20: Retention characteristics of n-type SiN flash device.

Figure 5.21 illustrates the program transient characteristics of n-type capacitors with SiN as charge trapping medium. High frequency C-V measurements are taken under a constant gate bias of 5 V. The SiN flash cells cannot provide significant electron storage until pulse duration exceeds 100 μs. Flat band shift reaches a maximum value of 0.48 V for Batch #1 and 0.68 V for Batch #2. The difference between maximum values of flat band shifts is a consequence of non-uniform deposition of SiN layer by PECVD. In addition, a flat band shift of 0.5 V will not be sufficient for read operation of flash memories. Operation voltage should be increased to improve injection of charges and to reach a larger memory window. For Batch #2, program transfer characteristics make jumps at 100 μs and 10 ms, confirming that electrons are stored in traps with multiple levels in SiN layer.

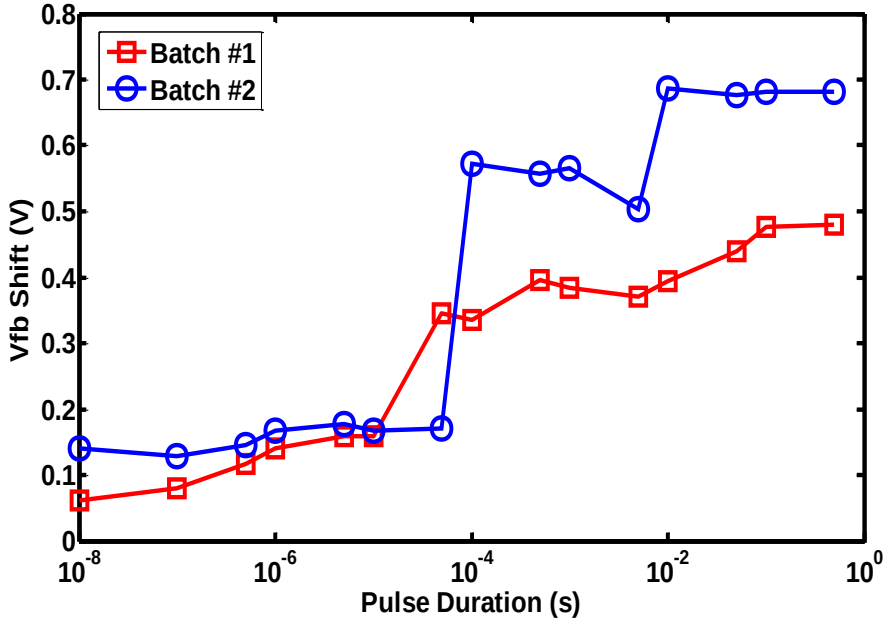

Figure 5.21: Program transient characteristics of n-type SiN flash devices.

5.8 Comparison of Graphene and SiN as Charge Storage Medium for Electron Storage

In this section, electron storage property of graphene nanosheets in flash applications is compared with ONO-type flash cell with Si-rich nitride layer as the charge-trapping medium. The band diagrams of two structures are given in Figure 5.22. Defect free graphene has zero band gap and its band lies within the forbidden gap of Si [90]. The band diagram of Si-rich SiN trapping layer is shown in accordance with the existing literature [91]. Compared to stoichiometric SiN, Si-rich SiN provides a smaller band gap with increased both conduction and valence band offsets. It was reported that physical trap densities of both stoichiometric SiN and Si-rich SiN are very similar; however, traps in Si-rich SiN are mostly located in shallow energy level [92].

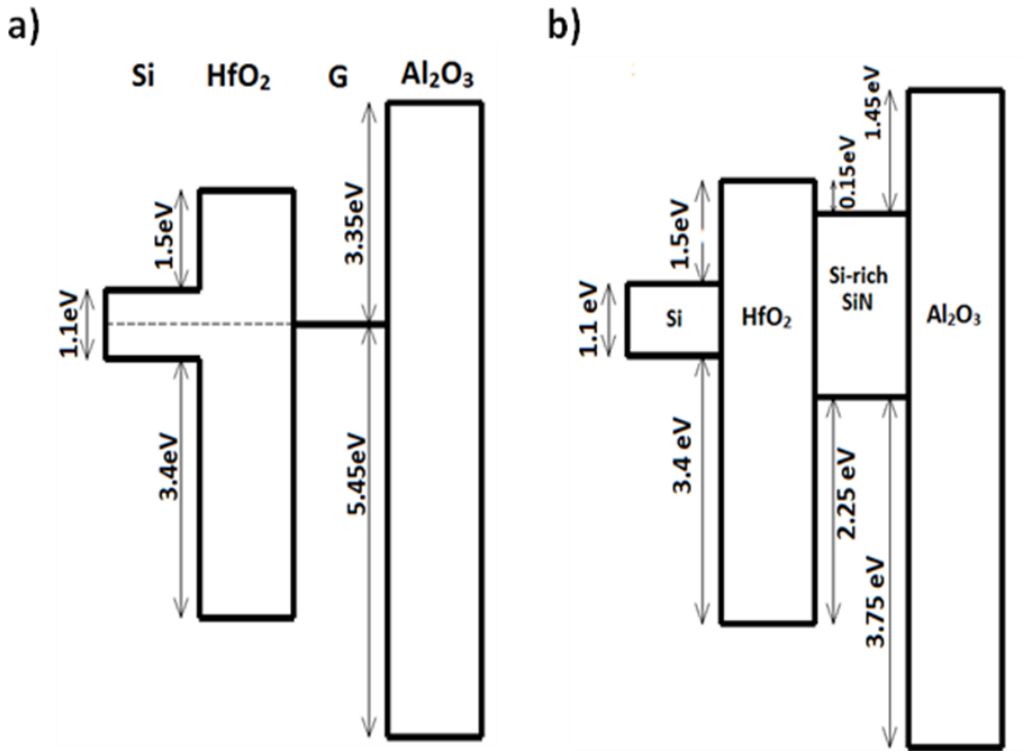

Figure 5.22: Band diagrams for samples with (a) graphene nanosheets and (b) Si-rich nitride as charge-trapping mediums.

The high frequency C-V curves of flash cells with graphene and Si-rich nitride under different sweep ranges are illustrated in Figure 5.23. For graphene flash cell, the C-V characteristics exhibit larger hysteresis window (about 2.03 V) under voltage sweep of ±5 V whereas the hysteresis window for SiN flash cell is 1.75 V. The enhanced memory window confirms that the use of graphene improves the charge storage capability and reduces the operation voltage.

The stored charge density can be calculated with the following equation:

$$n = \frac{-\Delta V_{fb} C_{ox}}{qA} \tag{5.1}$$

where n is the stored charge density, A is the contact area and q is the charge. ΔV_{fb} and C_{ox} are determined from the measured data. The trapped charge density for graphene flash is calculated as 8.17×10^{12} cm^{-2} and for SiN flash

memory device; it is 1.90x10^{12} cm^{-2} under ±5 V dual sweep range. Considering the trapped charge density, graphene layer provides better charge storage capability.

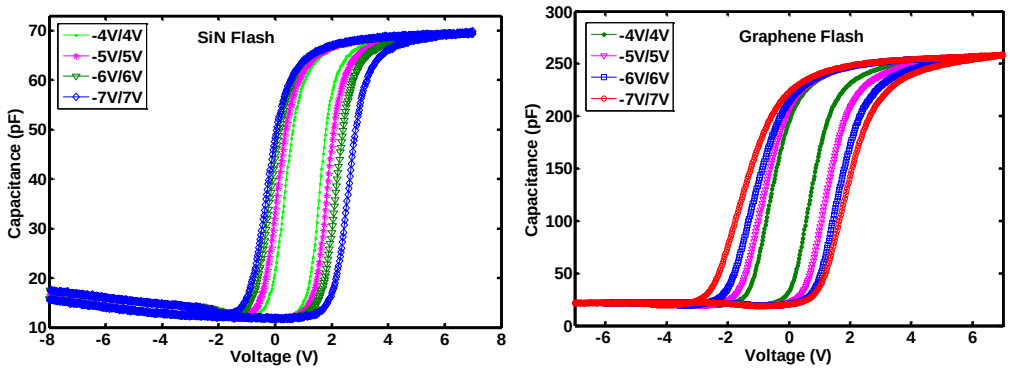

Figure 5.23: Comparision of hysteresis window for electron storage.

Retention rates for electron storage are compared in Figure 5.24. Graphene flash devices exhibit improved retention characteristics at room temperature. Based on the extrapolated data, graphene flash structure will lose 34% of trapped charges with an initial flat band shift of 1.37 V after 10 years whereas it is 62% with a 0.82 V initial flat band shift for a memory cell with Si-rich nitride medium. Graphene provides a deeper potential well for electrons; however, for Si-rich SiN, charges are stored in shallow traps and trapped electrons face a lower barrier height during retention. Since graphene flash memory structure has traps in deep energy level, it shows superior retention rate.

The program transient characteristics of graphene flash and SiN flash structures under a constant bias voltage of 5 V are given in Figure 5.25. Based on the results, it is obvious that programming efficiency of graphene-based flash cells is significantly enhanced compared to SiN-based flash cells. At the same programming voltage, the graphene flash exhibits more than 1 V flat band shift at 100 ns pulse duration while for SiN flash cells, the pulse duration should be at least 100 μs to obtain 0.57 V flat band shift. Moreover, SiN-based structure

cannot achieve the same saturated flat ban shift even at 0.5 s. The write voltage should be increased or longer pulse durations should be applied for SiN flash devices. The improvement in programming speed is attributed to the higher density of states in graphene nanosheets. A higher saturation flat band shift is observed for graphene flash, confirming that the application of graphene as the trapping layer enhances the charge storage capability. Moreover, graphene offers deep trapping levels and increases the barrier height with the control oxide. Then, back tunneling of trap charges is decreased. While, for Si-rich nitride, the traps are located in shallow energy levels and the probability of back tunneling is enhanced.

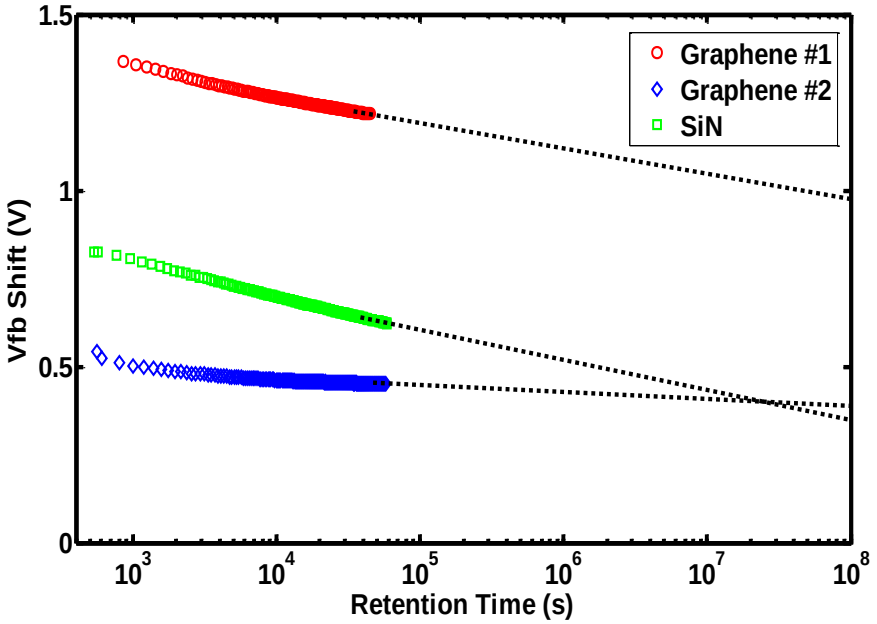

Figure 5.24: Retention characteristics of samples with graphene trapping layer and SiN trapping layer for electron storage.

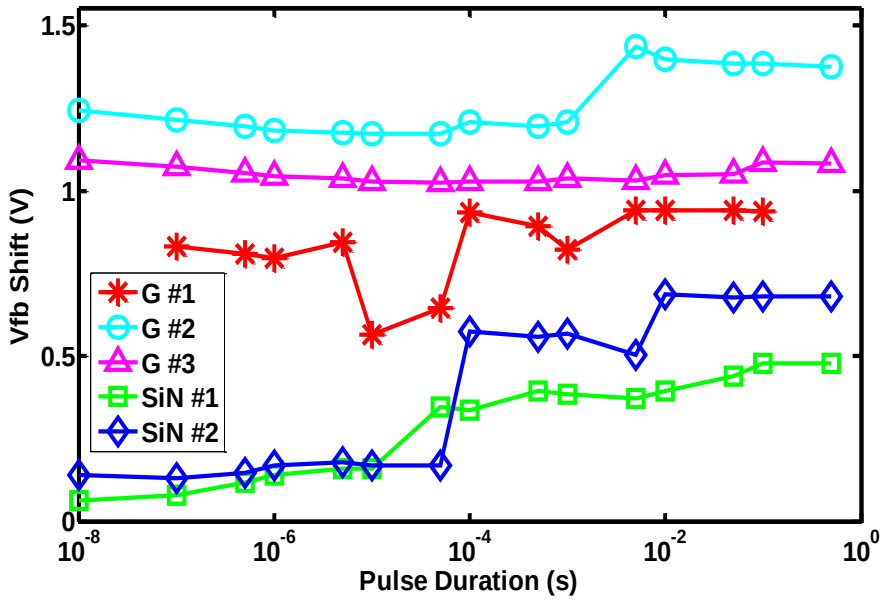

Figure 5.25: Program Transient characteristics of samples with graphene trapping layer and SiN trapping layer for electron storage.

5.9 Comparison of Graphene and SiN as Charge Storage Medium for Hole Storage

In this section, hole storage capability of graphene flash cell are compared with Si-rich nitride flash cell. Figure 5.26 represents the clockwise C-V hysteresis of the memory structures with different sweep voltages. For p-type graphene flash memory, an obvious memory window of 2.03 V is observed at ±5 V sweep range whereas it is 0.75 V for SiN flash cell. It is clear that hysteresis window of graphene-based device is larger compared to SiN-based device, indicating that charge storage capability is enhanced with the use of graphene as the charge storage layer.

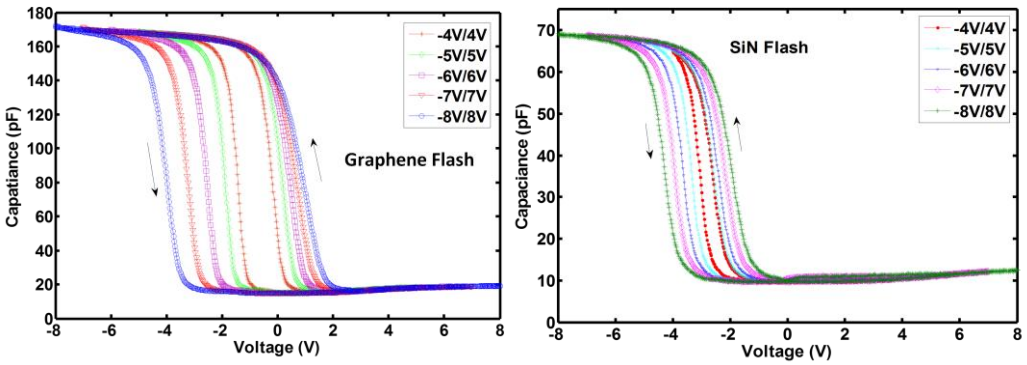

Figure 5.26: Comparision of hysteresis window for hole storage.

Hole storage performance of SiN flash and graphene flash cells for varying gate voltage at a pulse duration of 0.5 s is presented in Figure 5.27. Graphene flash cells exhibit significant flat band shifts at low voltages and reach the saturated flat band shift at -6 V. For Si-rich nitride flash devices, flat band shift is negligible with the applied voltage of -5 V and nitride-based cells do not reach a saturated flat band shift. Since graphene provides higher density of states, hole storage capability is better and larger flat band shifts are observed at low programming voltages in comparison to p-type SiN flash memories.

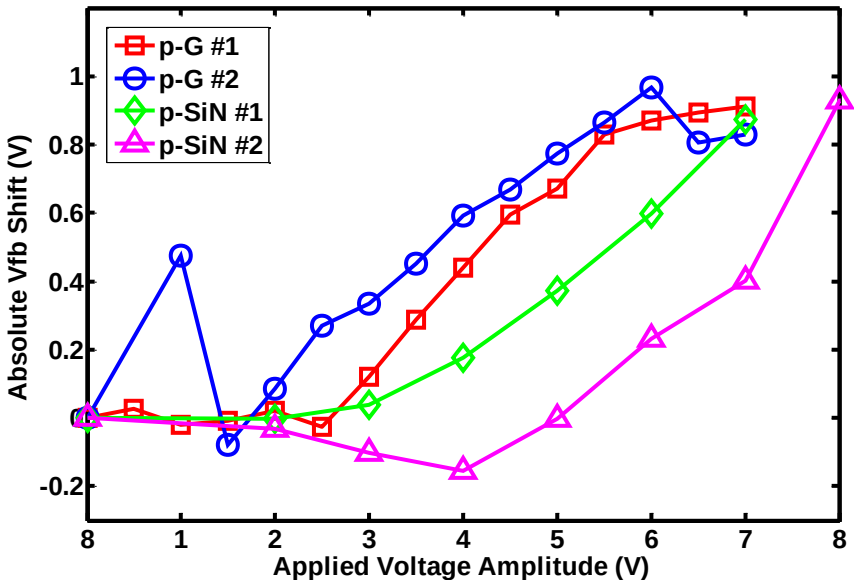

Figure 5.27: Hole storage performance of samples with graphene trapping layer and SiN trapping layer for varying bias voltages.

The long-term retention characteristics are compared in Figure 5.28. The graphene flash devices are programmed at -5 V for 0.5 s. For SiN memory structure, a gate voltage of -5 V is not sufficient to obtain an obvious flat band shift; therefore, it is charged at -7 V for 0.5 s to compare the retention rate with the graphene-based memory cell. It is observed that the nitride flash cell exhibits faster charge loss with a lower initial flat band shift. For graphene flash cells, a quick charge loss is observed in the early stage of retention, but charge loss becomes negligible after waiting time of 10^5 s. The good retention rate of graphene flash memories is the result of charge storage nodes located in deep energy level. The leakage currents towards the tunnel oxide and the control oxide are suppressed by storing charges in deep traps. P-type flash memory with Si-rich nitride shows worse retention rate in the early stage of retention time and better retention in the later stages because of the reason that Si-rich nitride have higher trap density in shallow levels and less charge traps in deep levels.

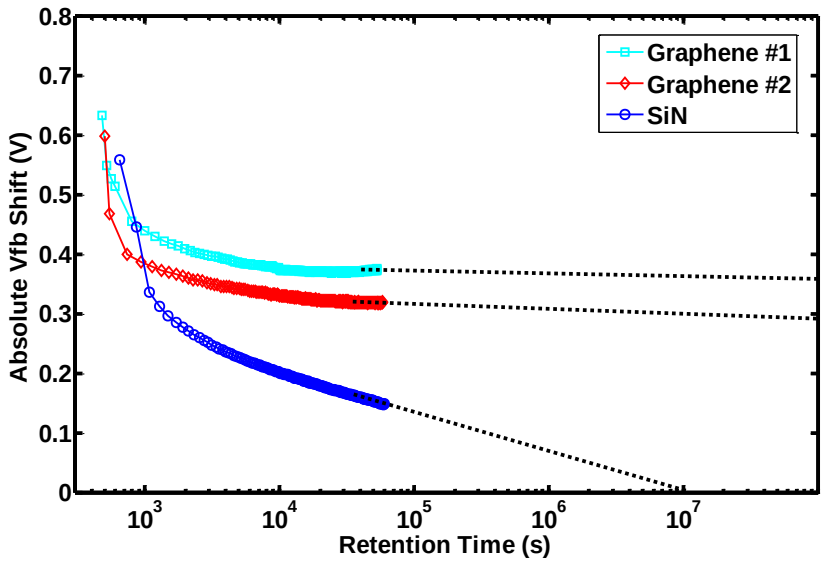

Figure 5.28: Retention characteristics of samples with graphene trapping layer and SiN trapping layer for hole storage.

The programming performance of graphene and Si-nitride for hole storage is compared in Figure 5.29. A voltage bias of -5 V is utilized to charge the

memory cells. The devices with Si-rich nitride as charge-trapping layer cannot exhibit a continuous increase in flat band shift, confirming that injected holes are stored in shallow traps with the application of -5 V and can easily detrap in a few seconds. Whereas graphene flash cells shows superior program transient characteristics with 0.3 V flat band shift at 100 ns. Faster programming speed of graphene-based devices is related to the greater number of trapping nodes. In addition, holes encounter higher barrier when tunneling thorough the control oxide, confirming that the use of graphene flakes decreases the back-tunneling probability. For comparison with electron storage, graphene flash cells achieved quick saturation at 100 ns when electrons are injected in the trapping medium. The faster behavior is attributed that HfO_2 tunnel dielectric offers asymmetric band offsets with a lower barrier for electrons. In addition, hole injection reaches the saturated flat band shift at longer pulse durations due to their larger effective mass.

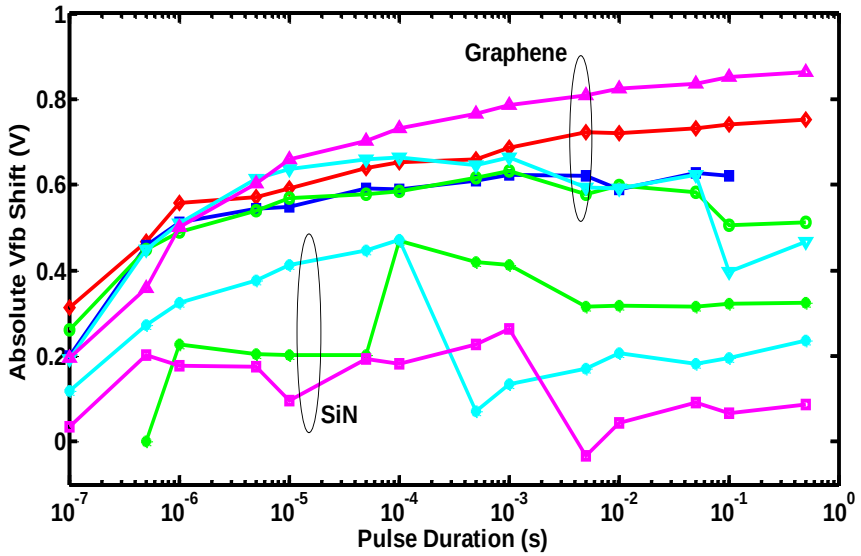

Figure 5.29: Program transient characteristics of samples with graphene trapping layer and SiN trapping layer for hole storage.

5.10 Graphene Flash Memory with AlN as Tunnel Layer

In this study, we have fabricated graphene flash structure with AlN as a tunnel layer. We compare the retention, operation voltage and program transient characteristics of AlN-based graphene flash memories with the structure using HfO_2 as the tunnel oxide. The results presented in this study are collaboration efforts with Cagla Ozgit, Inci Donmez and Dr. Necmi Biyikli.

5.10.1 Device Fabrication

The graphene flash structure with AlN as the tunnel layer is fabricated on p-type Si wafers. The fabrication process is similar to the one explained in Chapter 4. The main difference is that the HfO_2 tunnel layer is replaced with a 7-nm-thick AlN film deposited by plasma-enhanced atomic layer deposition (PEALD) using TMA and NH_3 as precursors at 185ºC. Ozgit et al. reported the deposition conditions and characterization of PEALD-grown AlN thin films [93]. The structure of AlN-based graphene flash cells is given in Figure 5.30.

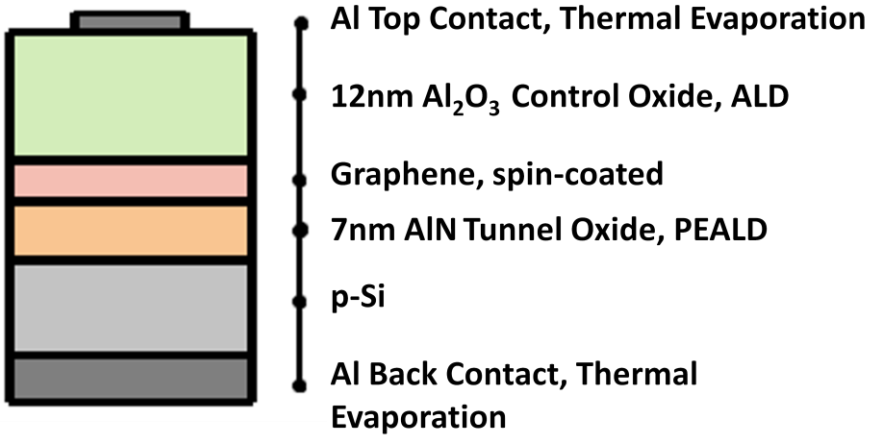

Figure 5.30: Device structure for graphene flash device with AlN as charge storage medium.

5.10.2 Electrical Characterization of AlN Thin Film

The AlN thin film can be seen a prospective candidate for tunnel oxide in flash applications due to its appropriate band offsets and high dielectric constant. However, the AlN thin film should provide minimal charge trapping in order to utilize as a tunnel layer. Tunnel oxides in flash devices should provide low charge trapping to avoid trap-assisted tunneling of stored charges.

The electrical performance of PEALD-grown AlN thin film is determined by fabricating MIS capacitors and characterizing the charge trapping property. For this study, the MIS capacitors are fabricated on a p-type Si substrate with 7 nm AlN layer as an insulator. Thermally-evaporated Al layers are utilized as a top and bottom contact. The charge trapping property is determined by carrying out C-V measurements over a large voltage range under dark. As given in Figure 5.31. The hysteresis window slightly increases towards negative direction with the increase of a dual-directional voltage range. Based on the shift of flat band voltage in negative direction, it is concluded that the AlN layer has hole traps. C-V curves of the MIS capacitor confirm that the dielectric has an excellent interface with the Si for a film thickness less than 10 nm. Although hole traps exist in the AlN layer, it can be still utilized as a tunnel layer due to the fact that charge trapping behavior will become insignificant at low operating voltages according to the hysteresis curve.

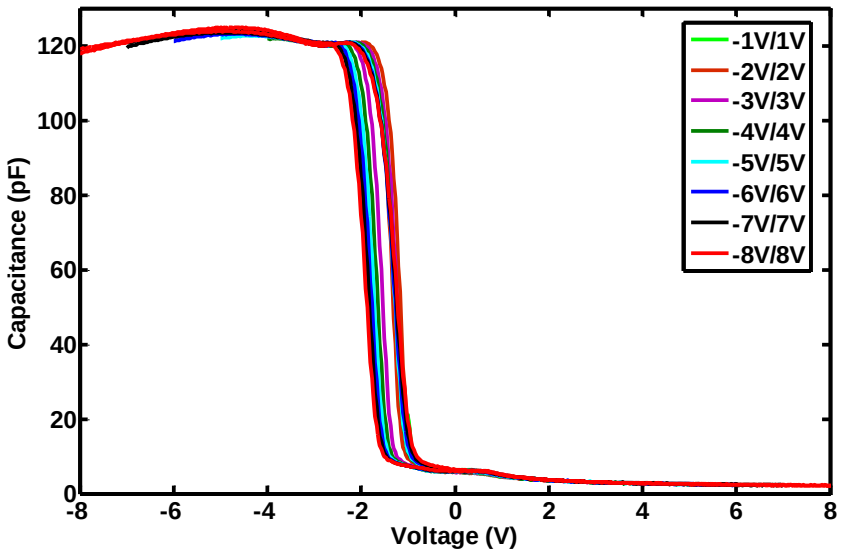

Figure 5.31: C-V crahateristics of MIS structure with 7 nm-thick AlN layer.

5.10.3 Electrical Characterization of the Memory Structure

Figure 5.32 illustrates the band diagram of the memory capacitor under zero bias [94, 95]. AlN provides a higher conduction band offset with Si compared to HfO_2. Therefore, the use of AlN is expected to improve the retention performance for electron storage. Moreover, AlN-based flash cells can exhibit faster write/erase speed at lower operating voltage since the dielectric constant of AlN is much lower than HfO_2. The electric field across the tunnel oxide will increase when the tunnel oxide with a lower dielectric permittivity compared to the control oxide is utilized according to the Eq. 5.2, explained in Chapter 3.

$$E_{ox} = \frac{Vg}{t_{ox}\left(1 + \frac{\varepsilon_{ox} t_c}{\varepsilon_c t_{ox}}\right)} \tag{5.2}$$

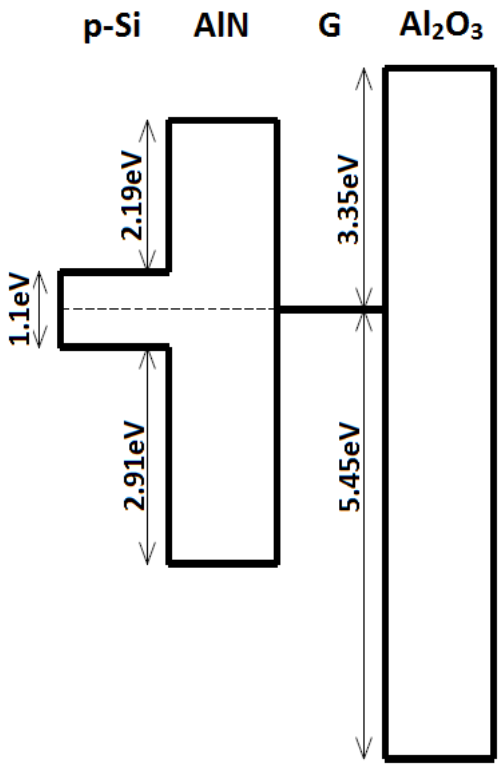

Figure 5.32: Band diagram for samples with AlN as the tunnel layer [94, 95].

The high frequency (1 MHz) C-V measurements of graphene flash memory with AlN as the tunnel layer are given in Figure 5.33. The hysteresis window gets larger with the increase of the dual-sweep range. The structure exhibits a counterclockwise hysteresis of 6 V under sweep range of ±8 V. The hysteresis window expands towards both negative and positive directions; however, electron trapping becomes dominant after the applied voltage range of ±3 V. Based on the large hysteresis width, it is concluded that the structure can exhibit great memory performance at low operating voltages.

Electron storage performance of the flash cell for varying the bias voltage is illustrated in Figure 5.34. Pulse duration is kept constant during the measurements. Based on the results, electron storage begins when the device is biased at 1 V and flat band shift becomes significant at 3 V biasing. Programming voltage characteristics approves that the use of AlN as the tunnel

oxide decreases the operating voltage. Since AlN provides smaller conduction band offset, the charge injection efficiency at low voltages is improved.

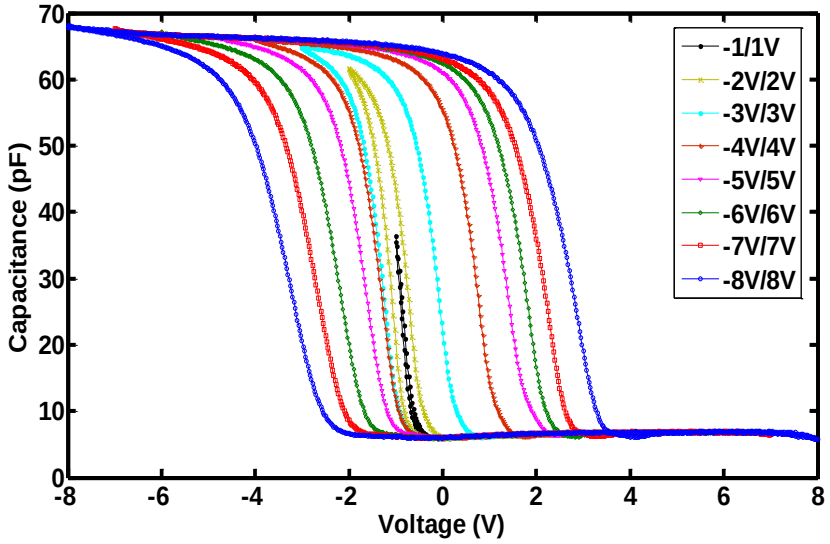

Figure 5.33: Hysteresis window of graphene flash memory device with AlN as the tunnel layer.

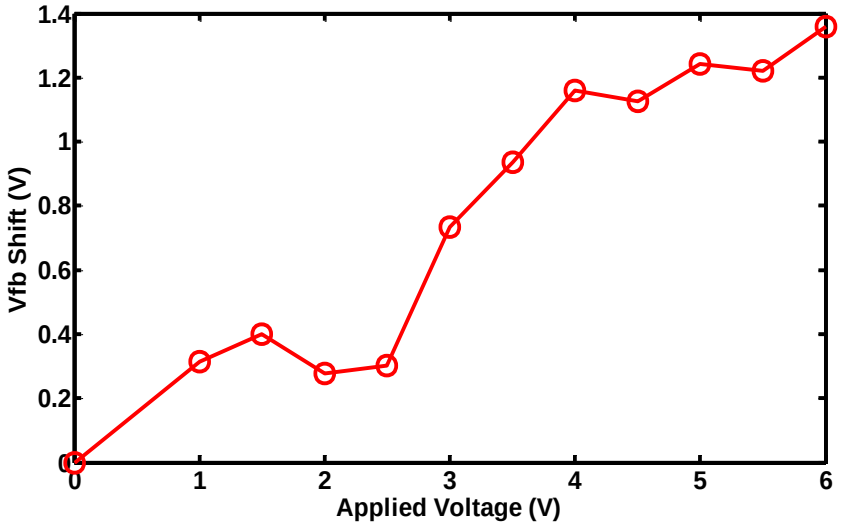

Figure 5.34: Electron storage performance of samples with graphene trapping layer and AlN tunneling layer for varying bias voltages.

Figure 5.35 shows the retention characteristics of AlN-based flash structure for electron storage. C-V measurements are conducted at zero retention voltage after programming the cell at 5 V for 0.5 s. A flat band shift of 1.36 V is obtained after programming and the flat band shift is observed as a function of retention time. The charge loss rate for the structure is worse than expected.

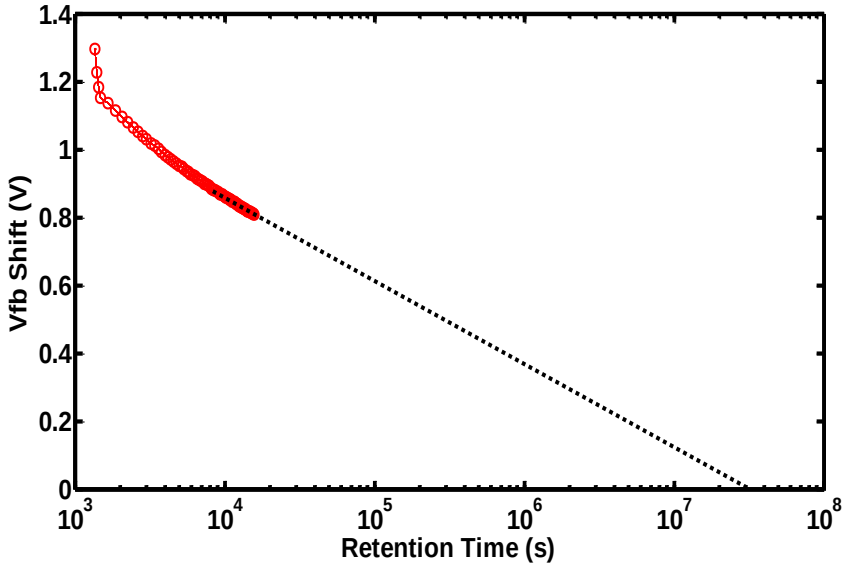

Figure 5.35: Retention characteristics of graphene flash device with AlN tunneling layer

To investigate the reason for fast charge loss, 7nm-thick AlN film deposited on Si(100) is examined with X-Ray photoelectron spectroscopy (XPS). Elemental composition as a function of ion beam etching is given in Figure 5.36. According to the compositional depth profile, C concentration rapidly decays and O concentration is high in the film. High concentration of oxygen impurities may create a conductive path between the Si substrate and the charge-trapping medium, resulting in a faster charge loss. Oxygen concentration in the film should be decreased to obtain improved retention capability for graphene flash cells with AlN as the tunnel layer.

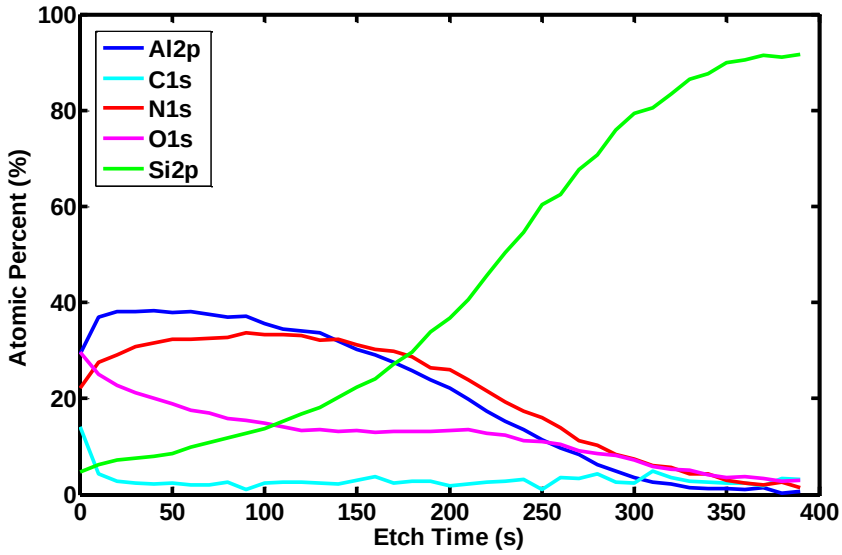

Figure 5.36: XPS results of 7nm-thick AlN film deposited on Si(100).

The transient program characteristics are presented in Figure 5.37. The device is programmed at 5 V and the flat band shift is determined as a function of pulse duration. Although the structure provides a significant flat band shift of 0.4 V at pulse duration of 100 ns, it was expected that the charge injection rate would be improved when the tunnel oxide was replaced with AlN layer. Since the oxygen impurities in the AlN thin film increases the charge leakage, the program transient characteristics become worse than expected. Lower oxygen concentration could be beneficial for the leakage and the operation speed. But, the study still shows promising results to enhance the sub-microsecond programming speed.

The transfer characteristics of the graphene flash with AlN tunnel layer under programmed and erased states is given in Figure 5.38. The write/erase conditions are +5V/-5V for 0.5 s. A memory window of 1.43 V is clearly achieved for the first write/erase cycle. The memory window narrows with the repeated write/erase cycles as the result of defects formed in the tunnel dielectric.

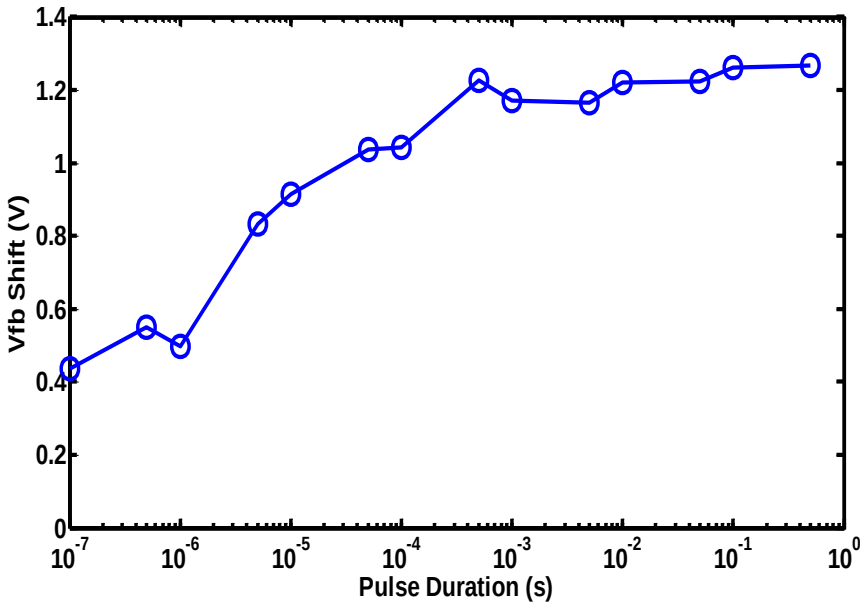

Figure 5.37: Program transient characteristics of graphene flash device with AlN tunneling layer.

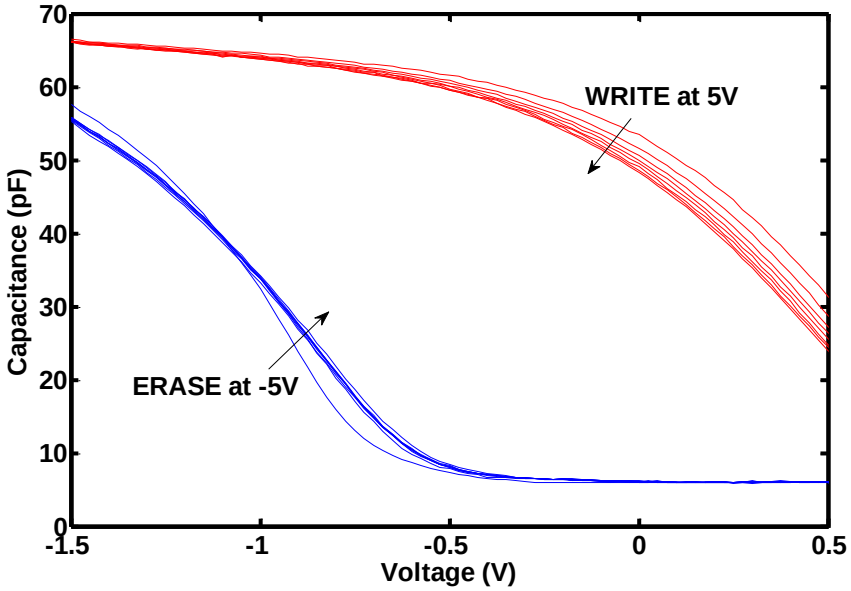

Figure 5.38: Memory transfer characteristics of graphene flash device with AlN tunneling layer under written and erased conditions.

Chapter 6

Conclusions

In recent years, non-volatile memory has dominated the semiconductor market due to increased demand for portable electronic equipments such as cellular phones, digital cameras, laptops etc. Conventional non-volatile memory structure is based on a MOSFET with a poly-silicon floating gate layer as the charge storage medium. Non-volatile flash memory devices with discrete charge-trap mediums are recently considered as an alternative solution to scaling issue in conventional floating gate technology. Currently, silicon-oxide-nitride-oxide-silicon (SONOS) type flash memory structures utilizing Si-nitride as the charge-trapping layer have been investigated because of its better scalability, simpler fabrication and improved endurance. Metal nanoparticles have been used as discrete charge storage nodes since they offer high density of states and selectable work function. However, diffusion of metal nanoparticles into other layers during high-temperature processes may cause reliability problems in flash memory applications.

High-k dielectrics in flash technology allow further scaling without increasing the leakage mechanism. In flash applications, dielectrics with wide band gaps, high band offsets and minimal charge trap densities are crucial to obtain improved retention characteristics with high charge injection currents.

Atomic layer deposition (ALD) is a powerful technique for the growth of pinhole-free high-k dielectrics with precisely controlled thickness and high conformality.

Graphene is a promising candidate as a charge storage medium for non-volatile technology since it provides high work function and advanced charge storage capability, which enable further scaling down of the tunnel and control dielectric thicknesses without degradation of retention characteristics. Therefore, replacing SiN layer in SONOS structure with graphene sheets is crucial to obtain low voltage operations, faster programming speed and enhanced retention.

In this thesis, our aim is to compare the charge storage performance of graphene sheets with those of SiN layer in flash memory applications. We fabricate graphene flash memories and characterize the memory performances of Al/Al_2O_3/Graphene sheets/HfO_2/(p-Si or n-Si) structures performing C-V measurements for both electron and hole storage. Based on Hummers Method, graphene oxide sheets are derived from the acid exfoliation of graphite and spin-coated onto tunnel oxide layer. Graphene sheets are obtained by thermal annealing and examined with AFM and SEM. ALD-grown HfO_2 and Al_2O_3 high-k dielectric layers are used as tunnel oxide and control oxide, respectively. For comparing the effect of graphene as the charge-trapping layer on the memory device performance, we also fabricate identical memory cells with Si-rich SiN layer replacing graphene. SiN films are deposited with high SiH_4/NH_3 gas flow ratio by PECVD and characterized by ellipsometer.

The flash memory structure with graphene nanosheets as the charge-trapping medium exhibits superior memory performance for both electron and hole storage. Compared to Si-rich nitride based devices, the retention rate and the programming speed are significantly enhanced with the application of graphene. Improved retention performance is a consequence of high density of charge nodes in deep energy level. The leakage mechanism through tunnel oxide

and control oxide is effectively suppressed with the charge storage in deep traps. Graphene trapping layers also improve the charge injection efficiency as a result of high density of states. For electron storage, the saturated flat band shift is obtained even with the pulse duration of 100 ns at applied bias of 5 V.

In order to investigate the memory performance of chemically-exfoliated graphene charge-trapping flash memory devices, hysteresis window, retention rate, operation voltage and programming speed characteristics are analyzed both for electron and hole storage mechanisms. The unique properties of graphene flakes offer promising solution for the requirements of high-density flash memory technology such as scalability, low operating voltage, faster write performance, at least 10-year retention and CMOS-compatible fabrication.

Bibliography

[1] P. J. McWhorter, S. L. Miller, and T. A. Dellin, "Modeling the memory retention characteristics of silicon-nitride-oxide-silicon nonvolatile transistors in a varying thermal environment," *Journal of Applied Physics*, vol. 68, pp. 1902 - 1909, 1990.

[2] D. Kahng and S. M. Sze, "A floating gate and its application to memory devices," *Bell System Technical Journal*, vol. 46, no. 6, pp. 1288 - 1295, 1967.

[3] H. A. R. Wegener, et al., "The variable threshold transistor, a new electrically alterable,
non-destructive read-only storage device," *IEEE Transactions on Electron Devices*, vol. 15, no. 6, pp. 420 - 421, 1967.

[4] D. Frohman-Bentchkowsky, "A fully decoded 2048-bit electrically programmable MOS-ROM," *Solid-State Circuits Conference. Digest of Technical Papers 1971 IEEE International*, vol. XIV, pp. 80 - 81, 1971.

[5] D. Frohman-Bentchkowsky, "Memory behaviour in a floating gate avalanche injection MOS (FAMOS) structure," *Applied Physics Letters*, vol. 18, no. 8, pp. 332 - 334, 1971.

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

[7] D. Frohman-Bentchkowsky, "FAMOS - A new semiconductor charge storage device," *Solid-State Electronics*, vol. 17, no. 6, pp. 517 - 528, 1974.

[8] J. Maserjian, "Tunneling in thin MOS structures," *Journal of Vacuum Science and Technology*, vol. 11, pp. 996 - 1003, 1974.

[9] S. Bharadwaj, "Investigation of oxide thickness dependence of Fowler-Nordheim parameter B," Graduate School Theses and Dissertations, University of South Florida, 2004.

[10] S. Aritome, et al., "Reliability Issues of Flash Memory Cells," *Proceedings of IEEE*, vol. 81, no. 5, pp. 776 - 788, 1993.

[11] B. Ricco, G. Gozzi, and M. Lanzoni, "Modeling and Simulation of Stress-Induced Leakage Current in Ultrathin SiO2 Films," *IEEE Transactions on Electron Devices*, vol. 45, no. 7, pp. 1554 - 1560, 1998.

[12] C.-M. Yih, Z.-H. Ho, M.-S. Liang, and S. S. Chung, "Characterization of Hot-Hole Injection Induced SILC and Related Disturbs in Flash Memories," *IEEE Transactions on Electron Devices*, vol. 48, no. 2, pp. 300 - 306, 2001.

[13] K. Sakakibara, N. Ajika, et al., "A Quantitative Analysis of Time-Decay Reproducible Stress-Induced Leakage Current in SiO2 Films," *IEEE Transactions on Electron Devices*, vol. 44, no. 6, pp. 1002 - 1008, 1997.

[14] B. De Salvo, C. Gerardi, et al., "How far will silicon nanocrystals push the scaling limits of NVMs technologies?," *IEEE - International Electron Devices Meeting*, pp. 26.1.1 - 26.1.4, 2003.

[15] B. De Salvo, C. Gerardi, et al., "Performance and Reliability features of advanced nonvolatile memories based on discrete traps (Silicon nanocrystal, SONOS)," *IEEE Transactions On Device and Materials Reliability*, Vol. 4, No. 3, pp. 377 – 389, 2004.

[16] D. Corso, I. Crupi, V. Ancarani, et al., "Localized Charge storage in nancrystal memories: feasibility of a multi-bit cell," *European solid-state device research, ESSDERC 2003*, pp. 91 - 94, 2003.

[17] F. Iacona, G. Franzo and C. Spinella, "Correlation between luminescence and structural properties of Si nanocrystals," *Journal of Applied Physics*, vol. 87, pp. 1295 - 1303, 2000.

[18] T.-Y. Kim, N.-M. Park, et al., "Quantum confinement effect of silicon nanocrystals in situ grown in silicon nitride films," *Applied Phyisics Letters*, vol. 85, pp. 5355 - 5357, 2004.

[19] K. S. Cho, N.-M. Park, et al., "High efficiency visible electroluminescence from silicon nanocrystals embedded in silicon nitride using a transparent doping layer," *Applied Physics Letters*, vol. 86, pp. 071909, 2005.

[20] G. Ammendola, M. Vulpio, et al., "Nanocrystal metal-oxide-semiconductor memories obtained by chemical vapor deposition of Si nanocrystals," *Journal of Vacuum Science and Technology B*, vol. 20, 2075 - 2079, 2002.

[21] C. Lee et al., "Operational and Reliability Comparison of Discrete-Storage Nonvolatile Memories: Advantages of Single- and Double-Layer Metal Nanocrystals," *Electron Devices Meeting, IEDM '03 Technical Digest*, pp. 22.6.1 - 22.6.4, 2003.

[22] D. Shahrjerdi, D. I. Garcia-Gutierrez and Sanjay K. Banerjee, "Fabrication of Ni Nanocrystal Flash Memories Using a Polymeric Self-Assembly Approach," *IEEE Electron Device Letters*, Vol. 28, No. 9, 2007.

[23] Y.-S. Jang, J.-Hw. Yoon and R. G. Elliman, "Formation of nickel-based nanocrystal monolayers for nonvolatile memory applications," *Applied Physics Letters*, vol. 92, pp. 253108, 2008.

[24] S. Maikap, P. J. Tzeng, et al., "Physical and electrical characteristics of atomic layer deposited TiN nanocrystal memory capacitors," *Applied Physics Letters* vol. 91, pp. 043114, 2007.

[25] F. M. Yang, T. C. Chang, et al., "Memory characteristics of Co nanocrystal memory device with HfO_2 as blocking oxide," *Applied Physics Letters*, vol. 90, pp. 132102, 2007.

[26] J. Kim, et al., "Memory characteristics of cobalt-silicide nanocrystals embedded in HfO_2 gate oxide for nonvolatile nanocrystal flash devices," *Applied Physics Letters*, vol. 92, pp. 013512, 2008.

[27] C.-C. Lin, T.-C. Chang, et al., "Charge storage characteristics of Mo nanocrystal dependence on Mo oxide reduction," *Applied Physics Letters*, vol. 93, pp. 222101, 2008.

[28] B. Park, K. Cho, et al., "Memory characteristics of Al nanocrystals embedded in Al_2O_3 layers," *Microelectronic Engineering*, vol. 84, pp. 1627 - 1630, 2007.

[29] F.M. Yang, T.C. Chang, et al.," Nickel silicide nanocrystals embedded in SiO_2 and HfO_2 for nonvolatile memory application," *Thin Solid Films*, vol. 516, no. 2 - 4, pp. 360 - 363, 2007.

[30] P. H. Yeh, L. J. Chen, et al., "Nonvolatile Memory Devices with $NiSi_2$/$CoSi_2$ Nanocrystals," *Journal of Nanoscience and Nanotechnology*, vol. 7, pp. 339 - 343, 2007.

[31] F. M. Yang, T. C. Chang, et al., "Using Double layer $CoSi_2$ Nanocrystals to Improve the Memory Effects of Nonvolatile Memory Devices," *Applied Physics Letters*, vol. 90, pp. 212108, 2007.

[32] P. H. Yeh, H. H. Wu, et al., "Fabrication of $NiSi_2$ Nanocrystals Embedded in SiO_2 with Memory Effect by Oxidation of the Amorphous $Si/Ni/SiO_2$ Structure," *Journal of Vacuum Science and Technology A*, vol. 23, pp. 851 - 855, 2005.

[33] Z. Liu, C. Lee, et al., "Metal Nanocrystal Memories Part II: Electrical Characteristics," *IEEE Transactions on Electron Devices*, vol. 49, pp. 1614 - 1622, 2002.

[34] Z. Liu, C. Lee, et al., "A Novel Quad Source/Drain Metal Nanocrystal Memory Device for Multibit-Per-Cell Storage," IEEE Electron Device Letters, vol. 24, no. 5, pp. 345 - 347, 2003.

[35] S. Tiwari, F. Rana, et al.,"A silicon nanocrystals based memory," *Applied Physics Letters*, vol. 68, no. 10, pp. 1377 - 1379, 1996.

[36] H. E. Maes and R. J. Van Overstraeten, "Low-field transient behavior of MNOS devices," *Journal of Applied Physics*, vol. 47, pp. 664 - 666, 1976.

[37] J. Robertson, "Band structures and band offsets of high K dielectrics on Si," *Applied Surfance Science*, vol. 190, pp. 2 - 10, 2002.

[38] J.A. Kittl, K. Opsomer, et al., "High-k dielectrics for future generation memory devices," *Microelectronic Engineering*, vol. 86, pp. 1789 - 1795, 2009.

[39] E.P. Gusev, D.A. Buchanan, et al., "Ultrathin high-K gate stacks for advanced CMOS devices," *IEDM Technical Digest 200,* pp. 20.1.1 - 20.1.4, 2001.

[40] K.J. Hubbard and D.G. Schlom, "Thermodynamic stability of binary oxides in contact with silicon," *Journal of Materials Research*, vol. 11, no. 11, pp. 2757 - 2776, 1996.

[41] D.G. Schlom and J.H. Haeni, "A Thermodynamic Approach to Selecting Alternative Gate Dielectrics," *MRS Bulletin*, vol. 27, pp. 198 - 204, 2002.

[42] P.S. Lysaght, P.J. Chen et al., "Experimental observations of the thermal stability of high-k gate dielectric materials on silicon," *Journal of Non-Crystalline Solids* vol. 303, pp. 54 - 63, 2002.

[43] S. K. Dixit, "Radiation-induced charge trapping studies of advanced Si and SiC based MOS devices," PhD thesis, Interdisciplinary Material Science, Vanderbilt University, 2008.

[44] V.A. Gritsenko, K.A. Nasyrov, et al., "A new low voltage fast SONOS memory with high-k dielectric," *Solid-State Electronics*, vol. 47, pp. 1651 - 1656, 2003.

[45] B. Govoreanu, D. Wellekens, et al., "Performance and reliability of $HfAlO_x$-based interpoly dielectrics for floating-gate Flash memory," *Solid-State Electronics*, vol. 52, pp. 557 - 563, 2008.

[46] A. Dana, Imran Akca, et al., "A Figure of Merit for Optimization of Nanocrystal Flash Memory Design," *Journal of Nanoscience and Nanotechnology*, vol. 8, pp. 510 - 517, 2008.

[47] R. Chau, S. Datta, et al., "High-k/Metal–Gate Stack and Its MOSFET Characteristics," *IEEE Electron Device Letters*, vol. 25, no. 6, pp. 408 - 410, 2004.

[48] M. V. Fischetti, D. A. Neumayer and E. A. Cartier, "Effective electron mobility in Si inversion layers in metal–oxide–semiconductor systems with a high-k insulator: The role of remote phonon scattering," Journal of *Applied Physics*, vol. 90, pp. 4587 - 4608, 2001.

[49] R. Choi, K. Onishi, et al., "Fabrication of high quality ultra-thin HfO_2 gate dielectric MOSFETs using deuterium anneal," *Technical Digest International Electron Devices Meeting*, pp. 613 - 616, 2002.

[50] K. Onishi, R. Choi, et al. "Effects of high-temperature forming gas anneal on HfO_2 MOSFET performance," *Technical Digest International Symposium VLSI*, pp. 22 - 23, 2002.

[51] HR. Huff, A. Hou, et al., "High-k gate stacks into planar, scaled CMOS integrated circuits," *Microelectronic Engineering*, vol. 69, pp. 152 - 167, 2003.

[52] C. Leroux, J. Mitard, et al., "Characterization and modeling of hysteresis phenomena in high K dielectrics", *IEDM Technical Digest*, pp. 737 - 740, 2004.

[53] M. Copel, M. Gribelyuk and E. Gusev, "Structure and stability of ultrathin zirconium oxide layers on Si (001)," *Applied Physics Letters*, vol. 76, pp. 436 - 438, 2000.

[54] M. Copel, E. Cartier, et al., "Robustness of ultrathin aluminum oxide dielectrics on Si(001)," *Applied Physics Letters*, vol. 78, pp. 2670 - 2672, 2001.

[55] A. Geim and K. Novoselov, "The rise of graphene," *Nature Materials*, vol. 6, no. 3, pp. 183 - 191, 2007.

[56] Y. Zhang, T. Tang, et al., "Direct observation of a widely tunable bandgap in bilayer graphene," *Nature*, vol. 459, no. 7248, pp. 820 - 823, 2009.

[57] X. Li, W. Cai, et al., "Large-area synthesis of high-quality and uniform graphene films on copper foils," *Science*, vol. 324, no. 5932, pp.1312 - 1314, 2009.

[58] A. Ismach, C. Druzgalski, et al., "Direct Chemical Vapor Deposition of Graphene on Dielectric Surfaces" *Nano letters*, vol. 10, no. 5, pp. 1542 - 1548, 2010.

[59] A. Reina, X. Jia, et al., "Large area, few-layer graphene films on arbitrary substrates by chemical vapor deposition," *Nano Letters*, vol. 9, no. 1, pp. 30 - 35, 2008.

[60] C. Lee, L. Baraton, et al., "Graphene growth directly on functional substrate," 2010.

[61] Y. Lee, S. Bae, et al., "Wafer-scale synthesis and transfer of graphene films," *Nano letters*, vol. 10, no. 2, pp. 490 - 493, 2010.

[62] S. Bae, H. Kim, et al., "Roll-to-roll production of 30-inch graphene films for transparent electrodes," *Nature Nanotechnology*, vol. 5, pp. 574 - 578, 2010.

[63] H. Hibino, H. Kageshima, and M. Nagase, "Graphene Growth on Silicon Carbide", *NTT Basic Research Laboratories* Atsugi-shi, 243-0198 Japan

[64] K. Emtsev, A. Bostwick, et al., "Towards wafer-size graphene layers by atmospheric pressure graphitization of silicon carbide," *Nature Materials*, vol. 8, no. 3, pp. 203–207, 2009.

[65] J. William, S. Hummers, R. E. Offeman, "Preparation of Graphitic Oxide**,**" *Journal of American Chemical Society*, vol. 80, pp. 1339 - 1339, 1958.

[66] S. Stankovich, R. D. Piner, et al., "Stable aqueous dispersions of graphitic nanoplatelets *via* the reduction of exfoliated graphite oxide in the presence of poly(sodium 4-styrenesulfonate)," *Journal of Materials Chemistry*, vol. 16, pp. 155 - 158, 2006.

[67] G. Eda and M. Chhowalla, "Chemically Derived Graphene Oxide: Towards Large-Area Thin-Film Electronics and Optoelectronics," *Advanced Materials*, vol. 22, pp. 2392 - 2415, 2010. 2010.

[68] J. T. Robinson, et al., "Wafer-scale Reduced Graphene Oxide Films for Nanomechanical Devices," *Nano Letters*, vol. 8, pp. 3441 - 3445, 2008.

[69] S. Wang, J. Pu, et al., "Wide memory window in graphene oxide charge storage nodes," *Applied Physics Letters*, vol. 96, pp. 143109, 2010.

[70] P. Cui, S. Seo, et al., "Nonvolatile memory device using gold nanoparticles covalently bound to reduced graphene oxide," *ACS Nano*, vol. 5 (9), pp. 6826 - 6833, 2011.

[71] A. J. Hong, E. B. Song, et al., "Graphene Flash Memory," *ACS Nano*, vol. 5 (10), pp 7812 - 7817, 2011.

[72] S. M. Kim, E. B. Song, et al., "Transparent and Flexible Graphene Charge-Trap Memory," *ACS Nano*, vol. 6 (9), pp. 7879 - 7884, 2012.

[73] T. Suntola and J. Antson, U.S. Patent 4.058.430, 1977.

[74] T. Suntola, J. Antson, A. Pakkala and S. Lindfors, "Thin Film Electroluminescent device," *SID 80 Digest*, vol. 11, pp. 108, 1980.

[75] T. S. Suntola, A. J. Pakkala and S. G. Lindfors, U.S. Patent 4,389,973, 1983.

[76] T. S. Suntola, A. J. Pakkala and S. G. Lindfors, U.S. Patent 4,413,022, 1983

[77] S. B. S. Heil, P. Kudlacek, et al., "In situ reaction mechanism studies of plasma-assisted atomic layer deposition of Al_2O_3," *Applied Physics Letters*, vol. 89, pp. 131505, 2006.

[78] W.F.A. Beslinga, E. Young, et al., "Characterization of ALCVD Al_2O_3–ZrO_2 nanolaminates, link between electrical and structural properties," *Journal of Non-Crystalline Solids*, vol. 303, pp. 123 - 133, 2002.

[79] A. S. Cavanagh, C. A. Wilson et al., "Atomic layer deposition on gram quantities of multi-walled carbon nanotubes," *Nanotechnology*, vol. 20, no. 25, pp. 255602, 2009.

[80] D. B. Farmer and R. G. Gordon, "ALD of High-κ Dielectrics on Suspended Functionalized SWNTs," *Electrochemical and Solid-State Letters*, vol. 8, pp. G89 - G91, 2005.

[81] Y. Xuan, Y. Q. Wu, et al., "Atomic-layer-deposited nanostructures for graphene-based nanoelectronics," *Applied Physics Letters*, vol. 92, pp. 013101, 2008.

[82] B. Lee, S.-Y, Park, et al., "Conformal Al_2O_3 dielectric layer deposited by atomic layer deposition for graphene-based nanoelectronics," *Applied Physics Letters*, vol. 92, pp. 203102, 2008.

[83] X. Wang, S. M. Tabakman, and H. Dai, "Atomic Layer Deposition of Metal Oxides on Pristine and Functionalized Graphene," *Journal of The American Chemical Society*, vol. 130, pp. 8152 - 8153, 2008.

[84] S. Stankovich, D. A. Dikin, et al., "Synthesis of graphene-based nanosheets via chemical reduction of exfoliated graphite oxide," *Carbon*, vol. 45, pp. 1558 - 1565, 2007.

[85] S. Kim, J. Nah, et al., "Realization of a High Mobility Dual-gated Graphene Field Effect Transistor with Al_2O_3 Dielectric," *Applied Physics Letters*, vol. 96, pp. 062107, 2009.

[86] B. Lee, S.-Y Park, et al., "Conformal Al_2O_3 Dielectric Layer Deposited by Atomic Layer Deposition for Graphene-Based Nanoelectronics," *Applied Physics Letters*, vol. 92, pp. 203102, 2008.

[87] B. Lee, G. Mordi, et al., "Characteristics of High-k Al2O3 Dielectric Using Ozone-Based
Atomic Layer Deposition for Dual-Gated Graphene Devices," *Applied Physics Letters*, vol. 97, pp. 043107, 2008.

[88] D. B. Farmer, H.-Y. Chiu, et al., "Utilization of a Buffered Dielectric to Achieve High Field-Effect Carrier Mobility in Graphene Transistors," *Nano Letters*, vol. 9, pp. 4474 - 4478, 2009.

[89] I. Meric, C. R. Dean, et al., "Channel Length Scaling in Graphene Field-Effect Transistors Studied with Pulsed Current-Voltage Measurements," *Nano Letters*, vol. 11, pp. 1093 - 1097, 2011.

[90] R. Yan, Q. Zhang, et al., "Determination of graphene work function and graphene-insulator semiconductor band alignment by internal photoemission spectroscopy," *Applied Physics Letters*, vol. 101, pp. 022105, 2012.

[91] C.-Y. Chen, K.-S. Chang-Liao, et al., "Improved programming/erasing speed of charge-trapping flash device with tunneling layer formed by low temperature nitrogen-rich SiN/SiO_2 stack," *Solid-State Electronics*, vol. 78, pp. 22 - 27, 2012.

[92] T. H. Kim, H. Park, et al., "Electron trap density distribution of Si-rich silicon nitride extracted using the modified negative charge decay model of silicon-oxide-nitrideoxide-silicon structure at elevated temperatures," *Applied Physics Letters*, vol. 89, pp. 063508, 2006.

[93] Cagla Ozgit, Inci Donmez, et al., "Self-limiting low-temperature growth of crystalline AlN thin films by plasma-enhanced atomic layer deposition," *Thin Solid Films*, vol. 520 pp. 2750 - 2755, 2012.

[94] P. D. C. King, T. D. Veal, et al., "Valence band offset of InN/AlN heterojunctions measured by x-ray photoelectron spectroscopy," *Applied Physics Letters*, vol. 90, pp. 132105, 2007.

[95] T. N. Bhat, M. Kumar, et al., "Band alignment studies in InN/p-Si(100) heterojunctions by x-ray photoelectron spectroscopy," Journal of Applied Physics, vol. 109, pp. 123707, 2011.