Hysteresis and Memory Effects in Nanocrystal Embedded MOS Capacitors

by

Eralp Atmaca

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2000

June 2000

Author ..... Signature redacted

Department of Electrical Engineering and Computer Science

May 11, 2000

Certified by ..... Signature redacted

Jesus del Alamo

Professor of Electrical Engineering

Thesis Supervisor

Certified by ..... Signature redacted

Arvind Kumar

Research Staff Member, IBM Corporation

Thesis Supervisor

Accepted by ..... Signature redacted

Arthur C. Smith

Chairman, Department Committee on Graduate Students

Hysteresis and Memory Effects in Nanocrystal Embedded MOS Capacitors

by

Eralp Atmaca

Submitted to the Department of Electrical Engineering and Computer Science
on May 11, 2000, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Nanocrystal Memory is a promising new memory type which utilizes silicon nanocrystals and quantum mechanical direct tunneling current for charge storage. This thesis presents the work done to characterize the memory effect in nanocrystal embedded metal-oxide-semiconductor (NC-MOS) capacitors, the fundamental components of the nanocrystal memory. Various properties of the NC-MOS capacitors including gate stack composition, oxide charge storage and interface traps are studied by making high frequency and quasi-static capacitance voltage and current voltage measurements. High frequency and quasi-static capacitance characteristics reveal hysteresis which is evidence for the memory effect. A hysteresis of 2 V is demonstrated which is large enough to enable the use of nanocrystal embedded devices as memory devices. Measurement results suggest that the tunneling in the accumulation bias regime is mostly electron tunneling from the channel into the nanocrystals, and the tunneling in the inversion bias regime is hole tunneling from the channel into the nanocrystals. Charge is stored in the nanocrystals either in the discrete quantum dot states or in the interface traps that surround the nanocrystals. The oxide thickness is varied to control the tunneling rate and the retention time. A thinner tunnel oxide is necessary for achieving a higher tunneling rate which provides a faster write/erase. However, when the barrier thickness is lower, the charge confined in the nanocrystals can leak back into the channel more easily. Measurement conditions such as bias schemes, hold times, sweep rates and illumination can significantly influence the memory effect. It is demonstrated that the memory effect is enhanced by longer hold times, wider sweep regimes and light.

Thesis Supervisor: Jesus del Alamo
Title: Professor of Electrical Engineering

Thesis Supervisor: Arvind Kumar
Title: Research Staff Member
IBM Corporation

Acknowledgments

This thesis work is conducted at IBM T.J. Watson Research Center, Yorktown Heights, New York as part of the VI-A Internship Program of the Massachusetts Institute of Technology. I worked in the Exploratory Memory and Device Modeling Group with Arvind Kumar, Jeffrey Welser, Kern Rim and Sandip Tiwari during my stay at IBM Watson.

My most special thanks are for Arvind Kumar. Discussions with Arvind created the core of this thesis work. He also helped by reviewing the thesis. Jeffrey Welser and Kern Rim offered great help in the interpretation of the results.

Also at Watson, I want to thank Steve Cohen for sharing his expertise in electrical measurements. Kevin Chan's effort to fabricate the devices on-time for my thesis made this work possible. Discussions with Donelli DiMaria, James Tsang and Frank Stern were very insightful to understand the results.

At MIT, I want to thank my thesis supervisor, Prof. Jesus del Alamo, for his effort to review the thesis, and clarify the fine points. His guidance helped me to understand and appreciate the results of this thesis.

As a last word, I want to thank my parents Semra and Fevzi Atmaca, and my sister Hande Atmaca for their love and support. Their existence makes the life meaningful for me. I also want to commemorate my late aunt Sabriye Ilker. She will be there in my commencement, in my heart.

Contents

1 Introduction **10**

1.1 Overview of Semiconductor Memories 12

1.1.1 Performance Considerations 12

1.1.2 Non-volatile Semiconductor Memories 14

1.1.3 Volatile Semiconductor Memories 16

1.2 Nanocrystal Memories 18

1.2.1 Motivation for Nanocrystal Memories 18

1.2.2 Nanocrystal Memory Performance 19

1.2.3 Nanocrystal Memory Architecture Issues 21

1.3 Organization of this Thesis 23

2 Device Physics and Fabrication **25**

2.1 Device Fabrication 25

2.1.1 RCA Clean 27

2.1.2 Dry Oxidation 28

2.1.3 Silicon Nanocrystal Formation 29

2.1.4 High-Temperature Oxide Deposition 30

2.1.5 Hot Annealing 31

2.1.6 Metallization 31

2.2 Device Characteristics 32

2.2.1 NC-MOS Electrostatics 32

2.2.2 Charge Confinement in Nanocrystals 37

2.2.3 Tunneling Through a Thin Oxide 38

2.2.4 Energy Band Model 40
2.2.5 Interface Traps 46
2.2.6 Optical Properties 47

3 Measurements **49**
3.1 Measurement Techniques 49
3.1.1 High Frequency C-V Measurements 53
3.1.2 Quasi-Static C-V Measurements 53
3.1.3 Current-Voltage Measurements 54
3.2 Measurement Conditions 55
3.3 Measurement Setup 57
3.3.1 Instrumentation 57
3.3.2 Calibration 59

4 Results **60**
4.1 NC-MOS Wafer Splits 60
4.2 Memory Effect 61
4.2.1 FGA Dependence of the Memory Effect 65
4.2.2 Sweep Range Dependence of the Memory Effect 67
4.2.3 Hold Time Dependence of the Memory Effect 70
4.2.4 Light Dependence of the Memory Effect 71
4.2.5 Sweep Rate Dependence of the Memory Effect 73
4.3 QS C-V and I-V Characteristics 75
4.3.1 FGA Dependence of QS C-V and I-V characteristics 76
4.3.2 Light Dependence of QS C-V and I-V Characteristics 78
4.3.3 Sweep Rate Dependence of QS C-V Characteristics 81
4.3.4 Sweep Direction Dependence of QS C-V Characteristics 82
4.4 Summary of the Results 83

5 Discussion **85**
5.1 Electrical Thickness 86

5.2 Flat-band Voltage Shift 88
5.3 The Nanocrystal Charge 90

6 Conclusion **92**

A NC-MOS Capacitor Run-Sheet for the Fabrication Facility **95**

Bibliography **103**

List of Figures

1-1 A General Classification of Memories 11

1-2 256Mb DRAM Cell 17

1-3 Cross-section of a NC-MEM Cell 19

1-4 Nanocrystal Memory Architecture 22

2-1 NC-MOS Capacitor Illustrated Cross-Section and TEM image 26

2-2 A 2-Dimensional Problem Separated into Two 1-Dimensional Problems 33

2-3 Energy Band Diagram in Thermal Equilibrium 42

2-4 Energy Band Diagram under Flat-band Condition 43

2-5 Energy Band Diagram under Accumulation Condition 44

2-6 Energy Band Diagram under Inversion Condition 45

3-1 A Simplified RC Circuit Model of the NC-MOS Capacitor 50

3-2 The Block Diagram of the Measurement Set-up 57

4-1 HF C-V : Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF 62

4-2 QS C-V : Memory Effect: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF 65

4-3 HF C-V : FGA Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF 66

4-4 HF C-V : Voltage Sweep Range Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF 68

4-5 HF C-V : Hold Time Dependence of the Memory Effect: Frequency = 100 KHz; Light OFF 71
4-6 HF C-V : Light Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec 72
4-7 HF C-V : Sweep Rate Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF 74
4-8 QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF 75
4-9 I-V Characteristics: Hold Time = 10 sec; Light OFF 76
4-10 QS C-V : FGA Dependence of QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF 77
4-11 FGA Temperature Dependence of the Interface Trap Density 78
4-12 IV : FGA Dependence of I-V Characteristics: Hold Time = 10 sec; Light OFF 78
4-13 QS C-V : Light Dependence of QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec 79
4-14 I-V : Light Dependence of I-V Characteristics: Hold Time = 10 sec . 80
4-15 QS C-V : Sweep Rate Dependence of QS C-V Characteristics: Hold Time = 10 sec; Light OFF 81
4-16 QS C-V : Sweep Direction Dependence of QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF 82

List of Tables

1.1 Read/Write Non-volatile Semiconductor Memory Performance 16

1.2 Volatile Semiconductor Memory Performance 18

2.1 Summary of some non-standard NC-MOS Fabrication Steps 32

3.1 Outline of Measurement Conditions Used in the Experiments 56

4.1 NC-MOS Capacitor Run Wafer Splits 61

5.1 Electrical and Predicted Oxide Thickness 88

5.2 Calculated Flat-Band Capacitance 89

5.3 The Flat-Band Voltage Shifts 89

5.4 The Nanocrystal Charge 91

Chapter 1

Introduction

Memory is one of the most important components of a modern electronic system. There are several different memory types with grossly varying features, and the choice of the memory type depends on the application. Today's memory technology is based on two different kinds of memories, solid-state[1] and mechanical.

Solid-state memories are based on semiconductor technology. All components of a solid-state memory are integrated on a single silicon chip, and the information is processed by using mobile charge carriers. Semiconductor memories can be classified by using various methods, but a most general one divides them into two main categories, non-volatile and volatile. Memories discussed in this section can be classified as shown in Figure 1-1.

A non-volatile memory does not lose its information contents when the power is turned off. Non-volatile semiconductor memories can be either read-only or read/write. In a read-only memory, memory contents can be accessed only for read operations, and the data cannot be changed. Masked Read-Only Memories (Masked ROMs), which can be programmed only once during the manufacturing, PROMs (Programmable ROMs), which are one-time programmable (OTP), and EPROMs (Electrically Programmable ROMs) are the most widely used read-only non-volatile semiconductor memories [1, 2]. In a read/write memory, the memory contents can be changed on-

[1]Solid-state memories are also called semiconductor memories.

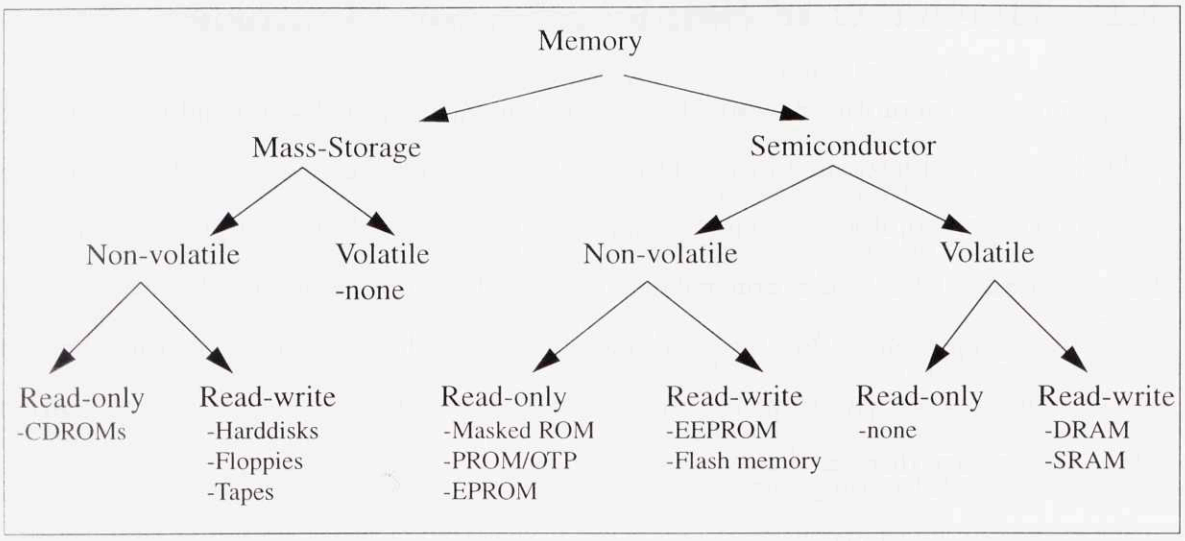


Figure 1-1: A General Classification of Memories

circuit as part of the routine memory operation. The most important read/write non-volatile semiconductor memories are EEPROMs (Electrically Erasable Programmable ROMs) and Flash Memories [1, 2].

A volatile memory loses its information contents when the power source is removed. All volatile semiconductor memories are read/write memories, and the most important members of this family are DRAMs (Dynamic Random Access Memories) and SRAMs (Static Random Access Memories). DRAM is the most abundant memory type in the world[2], and its most popular application is computer main memory.

Mechanical memories have moving parts; hence the name mechanical. The most common mechanical memories are CDROMs, hard-disks, floppy disks and tapes.[3] CDROMs use an optical storage medium, and hard-disks and floppy disks use a magnetic storage medium. The advantage of the mechanical memories is the capability to store very large amounts of data. Their handicap is low access speed which makes them unsuitable for high speed applications such as computer main memories.

[2]DRAM (1-T) cell is the most abundant man made object in the world [3].

[3]Hard-disks, floppy disks and tapes are read/write memories. CDROMs are read-only memories.

1.1 Overview of Semiconductor Memories

Semiconductor memories account for 50% of the entire world semiconductor market. Volatile semiconductor memories, DRAMs and SRAMs, have a total dollar share of 88% in the semiconductor memory market: 78% for DRAMs and 10% for SRAMs. The remaining 12% is the non-volatile semiconductor memory market where more than 50% is represented by the Flash memories [1]. There are many factors that can be considered when choosing the most convenient memory type for an application. These factors are discussed in the next section.

1.1.1 Performance Considerations

The desired features of an ideal semiconductor memory are high retention rate, low power consumption, fast random access (read, write/erase), high density (low cost) and high endurance.

A high retention rate refers to the ability to retain the data when the power is turned off. Therefore, non-volatile memories have high retention rates, and volatile memories have low retention rates. Some volatile memories can lose data even when the power is on, and this shortcoming is resolved by a special operation called refresh, which reads and restores the data during the routine memory operation. Therefore, a higher retention rate for a volatile memory that needs a refresh cycle means longer time intervals between subsequent refresh operations. A high retention rate is useful, because it reduces the power consumption, and it increases the CPU's access ability.

A high retention rate is desired when power consumption is a design constraint. Today's portable applications run on batteries that can provide a very limited power source for short periods of time. Minimizing the memory power consumption is a crucial issue in some electronic applications. Unlike volatile memories, non-volatile memories do not require stand-by power which means that the non-volatile memory cells do not consume power when they are not accessed. However, non-volatile memories consume operating power, the power needed for the access operations, in rates that are usually higher than those that are consumed by volatile memories.

As the CPU speeds approach the GHz regime, the memory should run fast enough to avoid becoming a bottleneck for the system operation. In read-only memories, where the only possible memory operation is read, the access speed is measured with the read cycle speed. On the other hand, in read/write memories the programming speed, i.e. write/erase speed, should also be considered when the memory speed is assessed.

Another important factor for the memory speed is the access scheme. Almost all of today's memories are random (direct) access memories which means that any data bit (or block) can be accessed individually without touching other bits (blocks). Some random access memories use bit-wise random access such that every individual bit can be accessed separately for any operation, and others use block-wise access where blocks (pages) of data are accessed together in one operation. An alternative access scheme is the sequential access which is used mostly in older memory technologies. In a sequential access system, data must be accessed in a sequential order, and this means that in order to access any bit, all of its preceding bits must also be touched. Sequential access memories prove to be very inefficient as online components for modern electronic systems, but they are still used in some very large scale back-up systems.[4]

It should be noted that the two terms RAM and ROM are not used carefully in the memory literature. RAM is used as a synonym for read/write volatile semiconductor memories, and ROM is used as a synonym for non-volatile semiconductor memories. Although this use is almost universal, it does not reveal what these terms really mean. In fact, some of the most important ROMs like EEPROMs and Flash memories are read/write memories that can make on-circuit write/erase operations. However since the write/erase operation is much slower than the read operation, a separate memory cycle is necessary for the write/erase operation. Moreover, these memories are also random access memories since individual bits can be accessed separately. For example, an EEPROM is a random access read/write memory, unlike its name suggests. All memory devices in Fig 1-1 are random access memories, except tapes

[4]Tape storage in AFS back-up systems.

which are sequential access memories.

The increasing complexity of the software applications demands more memory resources, and achieving more memory on a constant chip area means denser memory chips. The most important determinant factor of the memory density is the unit cell area which is defined as the area required to store one bit of information. Recent advancements in the silicon technology have enabled a major scaling in device dimensions, and consequently, smaller unit cell areas are now possible.

Finally, endurance is a measure of the memory lifetime. High endurance implies that the memory can be accessed many times without degrading the unit cell components. Some memory devices can provide very high endurance measured in billions of cycles [4]. In some applications, however, the memory becomes unusable after a very limited number of access operations.

1.1.2 Non-volatile Semiconductor Memories

The most important industry standard non-volatile memories are EEPROMs and Flash Memories. The relative ease of on-circuit electrical programming makes EEPROMs and Flash memories favorable over other non-volatile memories such as masked ROMs, PROMs and EPROMs.[5] An EEPROM cell is composed of two transistors whereas a Flash memory cell consists of a single transistor. The extra transistor in the EEPROM cell enables bitwise programmability at the cost of a lower chip density. In both the EEPROMs and Flash memories there is one transistor which has a charge-storing floating gate embedded in the oxide. The floating gate is a continuous thin-film conductor sandwiched between two insulating layers. The floating gate acts like a potential well where the carriers can be confined. The information is encoded with the electrical state of the floating gate.[6]

Two frequently used floating gate charging and discharging mechanisms are Fowler-

[5]EPROMs can be programmed off-line by first exposing to UV light to erase the data. Next, a specialized programmer instrument is used to electrically write the new data.

[6]The electrical state of the floating gate is defined by the charge level in the floating gate. For example, a charged floating gate can encode the bit 1, and a discharged floating gate can encode the bit 0.

Nordheim tunneling and hot electron injection. Fowler-Nordheim tunneling takes place when the potential barrier to tunneling is made low enough so that electrons can tunnel through it. The shape and the size of the barrier is changed by a high gate-to-substrate potential. Electrons that tunnel through this triangular barrier then relax into the floating gate or to the device channel. In hot electron injection, a few of the many electrons that flow from the device source terminal to drain terminal under the effect of a high enough drain-to-source potential can acquire enough speed in the vertical direction (towards the gate) to surmount the potential barrier imposed by the oxide. These few electrons are called the hot electrons, and they can be used to transport charge across the oxide.

The industry-popular high density Flash memories are programmed by hot electron injection, and erased by Fowler-Nordheim tunneling. These are slow mechanisms, and they also bring out endurance problems since high energy electrons degrade the oxide. The programming times are in the μs to ms range. The erase times are longer than 100 ms [1].

Flash memories have high densities because of the single transistor unit cell structure. However, scaling of this transistor is more challenging than it is for an ordinary transistor. The most serious problem is the difficulty of lateral charge confinement within a continuous film. With device dimensions as small as 0.25 μm, floating gate charge can leak into the nearby drain and source regions; hence cells might require a frequent refresh operation like in DRAMs, something that cannot be tolerated for a non-volatile memory that does not offer a fast operating speed. Lateral charge leakage can also bring out some disturbance problems; an operation on one cell can affect the neighboring cells since the space between the adjacent cells also gets smaller with the scaling device dimensions.

Another scaling problem associated with a continuous-film floating gate is the charge leakage from the floating gate into the device channel , a problem that arises when these two layers are close enough to allow charge transfer. Even tiny contact regions between the two layers can result in device failure since the floating gate is highly conductive, and the charge stored there can easily flow into the channel

Table 1.1: Read/Write Non-volatile Semiconductor Memory Performance

	Read speed	Write/Erase speed	Density	Endurance	Standby Power	Operating Power	Access Scheme
EEPROM	very low	very low	high	low	none	high	bitwise
Flash memory	very low	very low	very high	low	none	high	blockwise

through the contact region. It is very hard to totally eliminate such defects even with the latest silicon processing techniques, and therefore a small-scale continuous floating gate device is prone to failure because of minor defects in the insulating barrier between the device channel and the floating gate.

EEPROMs and Flash memories consume high operating powers. The two programming mechanisms, Fowler-Nordheim tunneling and hot electron injection, require energetic electrons, and therefore the memory power consumption is high. One way of reducing the power consumption is lowering the thickness of the potential energy barrier between the floating gate and the device channel. This can be done by decreasing the thickness of the intermediate oxide layer. The disadvantage of using a thinner oxide is that a thinner oxide degrades much faster when subjected to high-energy charge carriers, and this reduces the memory endurance.

The high power operation of EEPROMs and Flash Memories has an effect on the support circuitry, too. Since most EEPROMs and Flash memories require high voltages, very large charge pumps are necessary to produce the voltage, and therefore support transistors must be very large, and this reduces the chip density. The performance issues for the read/write non-volatile semiconductor memories are summarized in Table 1.1.

1.1.3 Volatile Semiconductor Memories

The most important industry standard volatile memories are SRAMs and DRAMs. These are both bitwise random access read/write memories, and they have high endurance. An SRAM cell consists of multiple transistors in a flip-flop configuration. Once it is programmed, an SRAM cell preserves the data so long as the power is on. There is no need for a refresh cycle, and this simplifies the SRAM operation. The

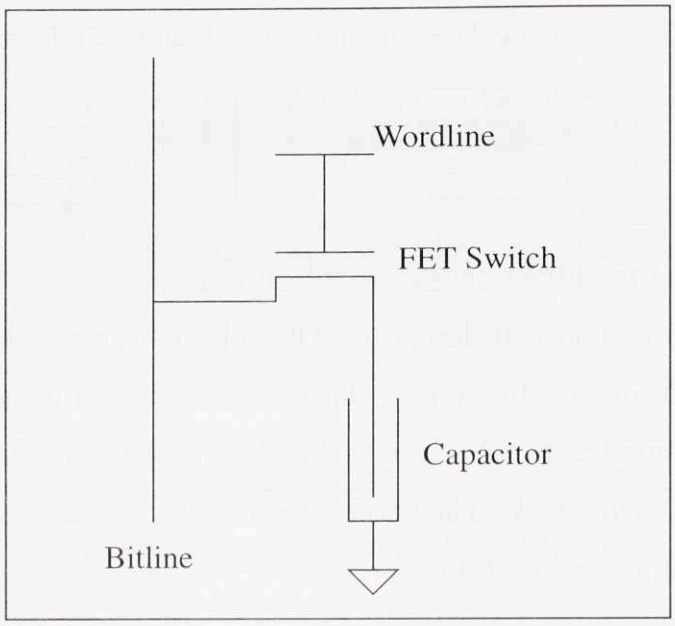


Figure 1-2: 256Mb DRAM Cell

standby power is low, but the operating power is becoming an issue as the SRAM densities get higher. SRAMs have very fast access times in the order of a few nanoseconds (5 ns–15 ns), but they suffer from low density, and they are used mostly as local fast memories [2].

A DRAM cell is composed of a pass transistor and a MOS capacitor as shown in Figure 1-2. Traditionally, in a DRAM cell about 30% of the cell area is occupied by the capacitor. The information is encoded with the charge stored in the capacitor, and the capacitor must be refreshed every few milliseconds to prevent information loss due to charge leakage into the substrate. DRAMs have fast access times in the 50 ns to 70 ns range. For DRAMs, the cell size is usually measured in terms of F, where F is the smallest feature that can be made in lithography at the time. For current technology (256 Mb DRAM), the cell size is $8F^2$, and F equals about 0.18 μm. The next generation of DRAM (1 Gb) is expected to have a cell size around $7F^2$, and the theoretical limit on the cell size is $4F^2$.[7] DRAM offers a higher density than SRAM, but as the memory chips are getting denser, there appear to be certain scaling and power consumption limitations for the DRAM technology. Pass

[7]It takes $4F^2$ just to hold the bitline and word-line wiring.

Table 1.2: Volatile Semiconductor Memory Performance

	Read speed	Write/Erase speed	Density	Endurance	Standby Power	Operating Power	Access Scheme	Retention
SRAM	very low	very low	low	low	none	high	bitwise	no refresh
DRAM	very low	very low	high	low	medium	high	bitwise	refresh

transistors with 0.1 μm dimensions show sub-threshold currents. Even if this problem is reduced by using a vertical design or SOI (silicon-on-insulator), there is still the problem of integrating gigabits into a chip area of 1 cm^2, a density which demands unit cell areas as small as 1 μm^2. Even with today's advanced lithography techniques, this is still a challenging task. The performance issues for the semiconductor volatile memories are summarized in Table 1.2.

1.2 Nanocrystal Memories

Nanocrystal Memory (NC-MEM) is a promising low-power, high density memory compliant with today's intensive computing needs [5]. An NC-MEM cell consists of a single floating gate transistor, and the most distinguishing feature of the NC-MEM is the discontinuous floating gate made up of an array of silicon nanocrystals. A discontinuous floating gate means that the nanocrystals are separate and electrically isolated. The fabrication of discontinuous films of very small size quantum dots[8] has been possible only recently with the advancements in chemical vapor deposition techniques [6]. The cross-section of a possible NC-MEM cell is shown in Figure 1-3.

1.2.1 Motivation for Nanocrystal Memories

As discussed in section 1.1.1 high retention, high density, fast access times, low power consumption and high endurance are the desired features for an ideal memory device. None of the currently known memory types has all of these desired features. DRAMs have scaling and associated power consumption and retention problems. SRAMs have very low densities making them unsuitable for most applications. EEPROMs

[8]The size of the floating gate charge storage units in NC-MEMs is about 5 nm.

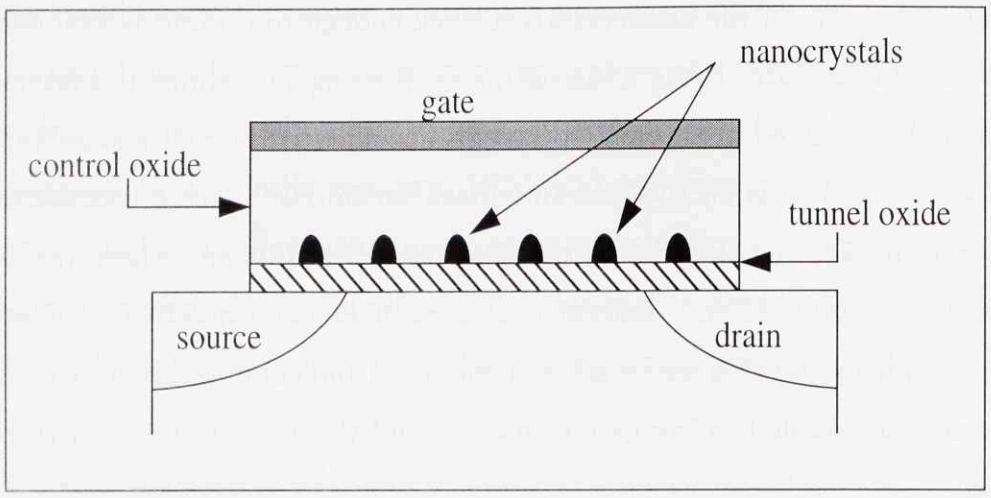


Figure 1-3: Cross-section of a NC-MEM Cell

and Flash memories have scaling and associated disturbance problems. They operate on high power, they have low endurance and long access times. There is substantial research to build a low-power memory with a long retention time, high endurance and a write/erase time in the sub-100 ns regime.

1.2.2 Nanocrystal Memory Performance

An NC-MEM cell consists of a single transistor with silicon nanocrystals of 5 nm size embedded in the oxide. This transistor is called the nanocrystal field-effect transistor (NC-FET). The charging and discharging of nanocrystals are made by direct quantum tunneling which happens at much smaller voltage regimes compared to Fowler-Nordheim tunneling and hot electron injection. Direct tunneling improves device endurance over the Fowler-Nordheim tunneling and hot electron injection since direct tunneling in principle does not degrade the oxide as do the higher voltage charge transport mechanisms. Another big advantage of using direct tunneling is the high write/erase speed. Since tunneling happens through a much thinner oxide, tunneling current density is higher than it is in the Fowler-Nordheim regime, and the charge/discharge is faster. NC-MEMs can offer faster write/erase times, lower power consumption and higher endurance than the industry-available non-volatile memories like EEPROMs and Flash Memories.

The oxide between the nanocrystals is thick enough to prevent lateral conduction inside the floating gate. Using a discontinuous floating gate eliminates lateral charge leakage and associated disturbance problems encountered in continuous-film floating gate memories such as EEPROMs and Flash memories. Access transistors, which are used to resolve the disturbance problems in EEPROMs and Flash memories, are not needed for NC-MEMs. Lateral charge confinement capability of nanocrystals saves the chip space taken up by access transistors and support circuitry used in conventional non-volatile floating-gate memories, and this in turn increases the memory density.

Another advantage of a discontinuous floating gate is that it is less prone to failure due to defects in the barrier oxide. Fabrication flaws can put some nanocrystals in direct electrical contact with the device channel, but since the nanocrystals are electrically isolated from each other, the operation of other nanocrystals is not disturbed, and the device can still store charge in the isolated nanocrystals. Therefore, a discontinuous floating gate device is less prone to failure due to defects in the oxide compared to a continuous floating gate device.

The retention rate of NC-MEMs can be engineered by changing the height and the width of the potential barrier between the discontinuous floating gate of silicon nanocrystals and the device channel. In fact, an NC-MEM can be designed both as a non-volatile or volatile read/write memory. As it is discussed above, when it is designed as a non-volatile read/write memory, NC-MEM can compete in the same market with EEPROMs and Flash memories. When it is designed as a volatile read/write memory, NC-MEM can compete in the same market with DRAMs.

NC-MEM can be designed as a volatile memory by reducing the thickness of the potential barrier to direct tunneling. A thinner barrier increases the operating speed but the trade-off is a lower retention time. As a volatile read/write memory, NC-MEM does not have write/erase times as fast as DRAMs. Although faster than the Fowler-Nordheim and hot electron injection based write/erase operation, the write/erase operation in NC-MEMs is still a tunneling process, and tunneling processes are in general slower than the capacitor charging in DRAMs. On the other hand, using a

tunneling based charge/discharge mechanism provides a higher retention rate since charge leakage from the nanocrystals into the channel can occur only via tunneling, and therefore it is also limited by the long time constants of the tunneling process. When designed as a volatile memory, NC-MEM can operate with a much longer refresh cycle, and this reduces the power consumption. In summary, NC-MEMs can offer higher retention, lower standby power and higher density than the industry-available DRAMs, but they have much lower operating speeds than DRAMs.[9]

1.2.3 Nanocrystal Memory Architecture Issues

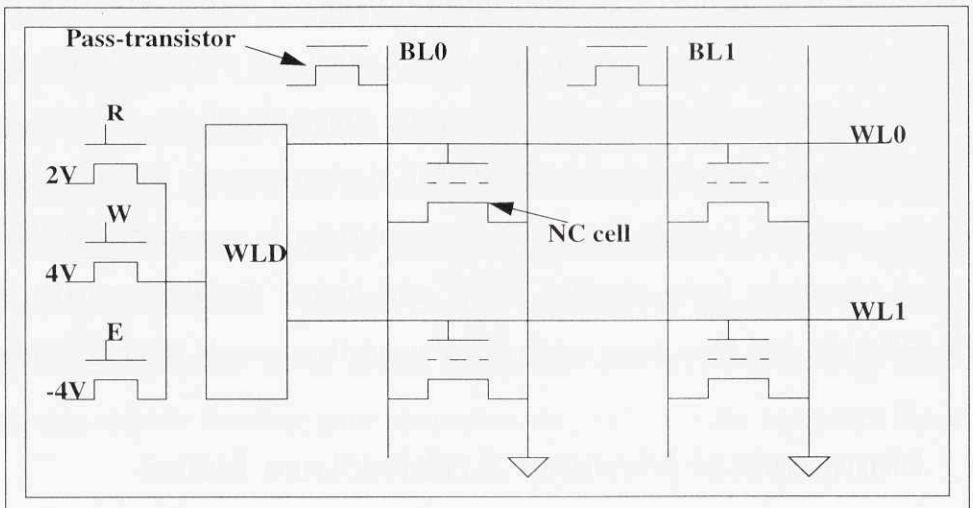
Nanocrystal memory can be designed as an array-structured module, and it resembles the DRAM chip in many aspects. The memory cells are connected by bitlines and wordlines. Bitlines are controlled by pass-transistors. Wordlines and bitlines are used together to change the state of the memory cell much like in conventional DRAM. However, in NC-MEM, bitlines and wordlines must be used together to change the cell contents, and the memory architecture is different from DRAMs. In this section, an NC-MEM architecture is suggested based on [4, 7].

NC-MEM operation requires four different voltage levels in contrast with the two level, high/low, operation of the DRAM. Hence, the algorithm used to access cell contents is also different from the DRAMs. In the suggested implementation, these voltage levels are -4 V, GND, 2 V, 4 V.

The read operation, ideally, does not disturb the contents of a cell, and hence it is a high speed, low power operation compared to the DRAM read operation. The write algorithm is composed of a block erase followed by a selected write which is required since there is no way to independently and simultaneously write 0's and 1's along a row without disturbing other cells. Figure 1-4 shows the array structure of nanocrystal memory and a possible algorithm for the read/write operations.

In its usual operation, each row in a nanocrystal memory array is accessed only when there is a read/write request for a cell in that row. When a row is accessed,

[9]There is intense research effort to increase the operating speeds of NC-MEMs.



Reading (WL1, BL1):

1. Set WL1 to 2V
2. Sense the current in BL1 // current means no charge, no current means charge
3. Set WL1 to 0V

Writing (WL1,BL1):

1. Set WL1 to -4V // block-erase all cells along WL1
2. Set WL1 to 2V // selectively write a 1 to the desired cell
3. Set BL0 -2V // by using pass-transistors
4. Set WL1 to 0V
5. Set BL0 to 0V

Figure 1-4: Nanocrystal Memory Architecture

the first operation is to sense the current in the cells in that row by using the sense amplifiers. There is one sense amplifier for each bitline, and hence, in a row-wise operation each sense amplifier senses the current from exactly one cell in that row. The sense amplifiers then latch the bits that they collect from the cells. There is one latch for each bitline. After the cell contents are latched, the external driver is checked to see what read/write operation is requested. If there is a read request then the bit in the corresponding latch is returned to the processor. If there is a write request, the contents of the corresponding latch is updated accordingly. As a final step, the latched bit is stored back into the memory cells in that row. The write operation consists of a block erase step which writes 0's into all cells in that row, followed by a selected write step which selectively writes 1's into the appropriate cells.

Even though the nanocrystals have much higher retention times than the ca-

pacitors in DRAMs, there is still some charge leakage if the nanocrystal memory is designed as a volatile memory. To prevent the information loss due to this leakage, there is a refresh cycle in the nanocrystal memory just like in the DRAMs. The refresh cycle is very similar to the read/write operation except that the memory does not look at the external driver to modify the latched information, and it simply re-writes the latch contents back into the cells. The advantage of the nanocrystal memory is that the refresh time is much longer than the DRAM refresh time. A standard DRAM cell needs to be refreshed 4 times in a second, whereas the nanocrystal memory cell needs to be refreshed almost once a day. This low leakage rate remarkably decreases the power consumption of the nanocrystal memories.

When the nanocrystal memory is designed as a non-volatile memory, no refresh cycle is necessary. However, the noted trade-off between the retention time and the block-erase/selected-write speed should be kept in mind as a design constraint.

1.3 Organization of this Thesis

The operation of the nanocrystal field-effect transistor (NC-FET), the core structure of the nanocrystal memory, must be well understood before the NC-FETs can be used to build a nanocrystal memory. Metal-oxide-semiconductor (MOS) capacitors are known as efficient probes for studying the capacitance and gate current characteristics of the field-effect transistors (FETs) since it is relatively easy to fabricate a MOS capacitor than fabricating a field-effect transistor[10] [8]. Similarly, nanocrystal metal-oxide semiconductor (NC-MOS) capacitors can be used for studying the capacitance and gate current characteristics of the NC-FETs. This thesis describes the results of research on nanocrystal metal-oxide-semiconductor capacitors.

The work done for this thesis is composed of three parts. The first part is the design of new NC-MOS capacitors based on earlier designs and results [9, 10]. In the next part, careful high frequency and quasi-static capacitance-voltage and current-

[10]The existence of source and drain regions in an FET makes the FET fabrication more complicated than the MOS fabrication.

voltage measurements are performed to study the device characteristics. In the last part, the measurement results are analyzed by using the results from the semiconductor and solid-state physics.

In Chapter 2 the design, fabrication and physics of the NC-MOS capacitors are discussed. Chapter 3 describes the measurement techniques. The results of the measurements are discussed in Chapter 4. Chapter 5 provides a quantitative background for the hysteresis effect which is the main focus of this thesis. Finally in Chapter 6, the overall results and the future research directions are discussed.

Chapter 2

Device Physics and Fabrication

Nanocrystal metal-oxide-semiconductor (NC-MOS) capacitor is a MOS capacitor with silicon nanocrystals embedded in the device oxide. The most prominent characteristic of this device is the quantum mechanical tunneling current between the device channel[1] and a discrete floating gate media of silicon nanocrystals. The NC-MOS capacitor can be used as a memory device by utilizing the electrical state of the nanocrystals[2] to encode information. However, as it is discussed in the following sections, it is difficult to control the charge level in the nanocrystals, and therefore designing a reliable encoding scheme for NC-MEMs is a challenging task.[3]. A cross-section of the NC-MOS capacitor is shown on the left side of Figure 2-1, and a Tunneling Electron Microscope (TEM) image of the NC-MOS gate stack is shown on the right side of Figure 2-1.

2.1 Device Fabrication

The fabrication of the NC-MOS capacitor is 100% compatible with the modern silicon technology. The fabrication can be done on standard silicon wafers in a moderate

[1]The term channel accounts for the sub-surface charge layer in both the accumulation and the inversion bias conditions.

[2]The electrical state of the nanocrystals can be defined by the level of the charge stored in the nanocrystals.

[3]Similar problems such as over-erase also exist for Flash memories, and techniques are known to get around these problems.

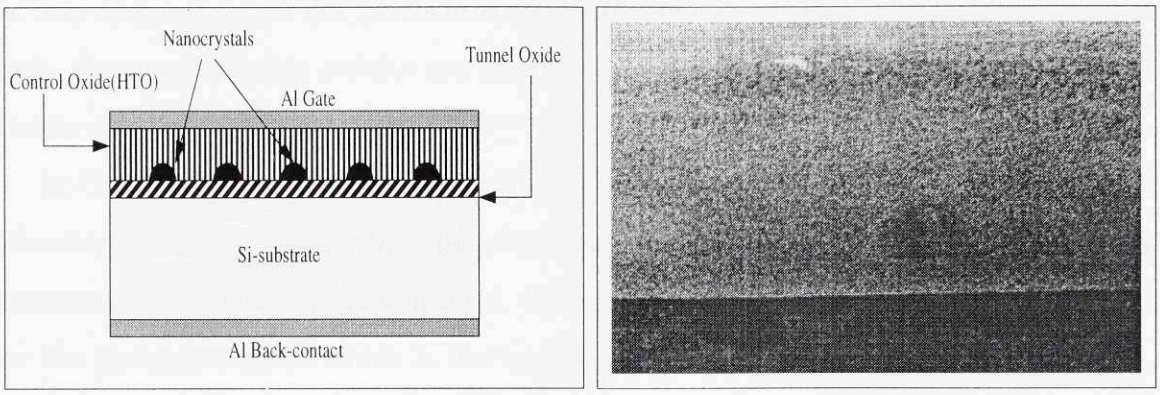


Figure 2-1: NC-MOS Capacitor Illustrated Cross-Section and TEM image

silicon processing facility. Although both types of wafers (n or p) can be used, the devices reported in this thesis work are always fabricated on n-type wafers. This wafer type is chosen because the electron tunneling rate is higher than the hole tunneling rate, and as it is discussed in Section 2.2, the tunneling rate through an oxide under accumulation bias conditions plays an important role in the device speed and the retention rate. In FETs (and also NC-FETs) p-type wafers are preferred because of the higher conductivity of the inversion channel in a p-type wafer. However, NC-MOS capacitor is essentially a one-dimensional device[4], and the most crucial device operation is in the gate-to-substrate direction. The resistivity of the wafers is about 1 Ω-cm which corresponds to a non-degenerate semiconductor with a doping level of approximately 10^{16} donors/cm^3.[5] The wafers are 8 inch in diameter and 0.5 mm in thickness.

The fabrication process has six parts: RCA clean, dry oxidation, silicon nanocrystal formation, high-temperature oxide deposition, hot annealing, and metallization. The fabrication of the gate stack encompasses three of these processes: dry oxidation, silicon nanocrystal formation, and high-temperature oxide deposition. These three processes are done subsequently in the UHV-CVD cluster tool[6], a specialized

[4]A MOS capacitor does not have drain and source regions, and therefore the device can be modeled in one-dimension which is oriented in the gate-to-substrate direction. A more detailed analysis can be found in Section 2.2.1.

[5]The donor type is phosphorus.

[6]UHV-CVD stands for Ultra-High Vacuum Chemical Vapor Deposition Tool.

tool which is capable of growing thermal oxides, forming nanocrystals and depositing high temperature oxides sequentially without outside intervention. Minimizing the outside intervention during the gate stack fabrication is very important because the quality of the interfaces between the gate stack layers can dramatically change the NC-MOS characteristics, and small amounts of contamination which will normally not disturb the regular MOS operation can affect the NC-MOS operation. Therefore, it is advantageous to fabricate the entire gate stack in a single session, and this is done using the UHV-CVD cluster tool.

2.1.1 RCA Clean

The electrical quality of the silicon surface is related to the contaminants present on the surface of silicon wafers. Contaminants can have very different electrical properties than the silicon crystal, and they can disturb the electrical characteristics of the device. Therefore, in order to obtain high performance and high reliability semiconductor devices, contaminants must be removed off the silicon surface. RCA clean is a cleaning procedure that has become an industry standard shortly after its development in 1970[7]. RCA is also the method of choice for cleaning the NC-MOS wafers prior to fabrication.

The RCA cleaning procedure has three major steps:

1. A thin silicon dioxide layer that forms as a result of atmospheric contact is removed using a diluted 50:1 H_2O:HF solution.

2. Insoluble organic contaminants are removed by placing the wafer in a 5:1:1 H_2O:H_2O_2:NH_4OH solution.

3. Ionic and heavy metal atomic contaminants are removed using a solution of 6:1:1 H_2O:H_2O_2:HCl.

In each step, the wafer sits in the solution for about 5 minutes at a temperature of 45 C. After the final step, the wafer is dried in a centrifuge drier. It should also be

[7]IBM calls the RCA clean the Huang clean.

noted that the water is rinsed after each major processing step.

2.1.2 Dry Oxidation

After the RCA clean, the wafer is placed in the UHV-CVD cluster tool. The UHV-CVD cluster tool is comprised of 5 interconnected chambers which are used for different processes. They are the nanocrystal chamber where the nanocrystals are formed, the oxidation chamber where the thermal oxides are grown, the HTO chamber where high-temperature oxides are deposited, the transfer chamber, and the load chamber. The tool can provide a very low base pressure (10^{-9} Torr), and this enables ultra-high vacuum processing.

The first process in the UHV-CVD tool is to grow a thermal oxide on the silicon wafer surface by using dry oxidation. During the dry oxidation process, the silicon wafer is placed in a quartz tube inside the oxidation chamber, which is a resistance-heated furnace. When the temperature becomes stable at 900 C, oxygen gas is allowed in the chamber, and silicon reacts with the oxygen gas. O_2 diffuses through the existing oxide[8], and reacts with silicon at the Si-SiO_2 interface to form more SiO_2.[9] The reaction takes place at low pressure, and it lasts about an hour.

The thickness of the oxide can be controlled by controlling the furnace temperature, the oxidation time or the gas pressure. Since the oxide is thin, and it is grown at a relatively lower temperature[10], the growth rate of the tunnel oxide is limited by the surface reaction rate, and the thickness of the oxide layer is a linear function of the oxidation time.

This layer of thermal oxide is referred to as the tunnel oxide because all quantum mechanical tunneling that occurs during the device operation takes place across this oxide. The devices reported in this work have 1.5 nm and 3 nm oxide thicknesses.

---

[8]The word oxide is used interchangeably with SiO_2.

[9]Another thermal oxidation method is wet oxidation. In this method, silicon is reacted with water vapor (H_2O) instead of oxygen gas. Dry oxidation gives a superior Si-SiO_2 interface, and it is used to form the most critical insulator regions in a device structure. Wet oxidation proceeds at a faster rate, and is therefore preferred in forming the thicker barrier oxides.

[10]Oxidation temperatures in the range 900 C to 1200 C are typical, the reaction proceeding more rapidly at higher temperatures.

2.1.3 Silicon Nanocrystal Formation

The next process is the deposition of silicon nanocrystals using the NC chamber of the UHV-CVD tool. The wafer is heated in the existence of silane gas (SiH_4) under 1 mTorr gas pressure. When a gas molecule hits the oxide surface at a high temperature (about 600 C), the silicon atom in the silane molecule can stick to the surface releasing the hydrogen atoms in the form of H_2 gas. Empirical observations suggest that on a clean oxide surface, the first silicon atom is more likely to stick at some defect or micro-rough area of the oxide, while the more smooth areas of the oxide remain clear. After a single silicon atom has stuck to some point, more silicon atoms can come in and stick to it, increasing the size of the silicon clump, eventually producing a nanocrystal. So the initial defect, or the initial silicon atom that sticks to it, acts as a nucleation center upon which the nanocrystal can build up. If the process continues for a long enough period of time, the various nanocrystals that have formed on the surface will eventually merge into one continuous film, but by stopping the process earlier (by turning off the silane gas), the average size of the nanocrystals can be controlled and the nanocrystals can be kept separate.[11] The devices used in this work have an average nanocrystal size of 6 nm and a sheet density of 5×10^{11} cm^{-2}. The distance between adjacent nanocrystals is 8 nm, and the nanocrystals occupy approximately 18% of the total capacitor area. It is predicted [6] that the nanocrystals resemble hemispheres, but for most purposes it is reasonable to approximate the nanocrystal geometry with a cube of 6 nm edge length. It should be noted that some devices are fabricated without depositing any nanocrystals. These devices are called the control devices, and all other fabrication processes except the nanocrystal formation are identical with those of the nanocrystal devices. Control devices can be used as effective tools to pinpoint the effects of nanocrystals on device operation.

[11] This assumes that most of them nucleate at the same initial time and grow at the same rate, which is only partially accurate.

2.1.4 High-Temperature Oxide Deposition

Although most oxides are formed by a thermal oxidation process, silicon oxides can also be directly deposited on the silicon surface without using any silicon from the wafer. Deposited oxides are not as high in quality as the thermally grown oxides. Thermally grown oxides are more commonly used since they are better passivators[12], but deposited oxides may be preferred in some occasions. The last process in the UHV-CVD cluster tool is high-temperature oxide (HTO) deposition[13] which is done at 750 C under 400-500 mTorr pressure. Growing a thermal oxide on top of the nanocrystals is avoided since the nanocrystals themselves can react with the oxygen gas. Moreover, a thermally grown oxide builds up at the Si-SiO_2 interface, and this means that growing a thermal oxide will change the thickness of the tunnel oxide. On the other hand, a deposited oxide deposits itself all over the surface of the wafer, regardless of what is on the surface. The purpose of this second layer of oxide is to shield the nanocrystals from the top and the sides without disturbing the tunnel oxide and the nanocrystals. The deposited oxide layer formed by this process is called the control oxide, and its thickness is about 9 nm. The top of the control oxide is not flat, but this does not disturb the electrical behavior since the aluminum deposited by the metallization process is much thicker than the oxide layers. The deposition of the control oxide takes place at a relatively higher temperature compared to the typical temperatures for oxide deposition,[14] and that is why the oxide deposited by this process is called the high-temperature oxide.

As it can be seen from Figure 2-1, the control oxide acts as the upper and lateral potential wall to confine charge carriers in the nanocrystals. It is thick enough to prevent any direct tunneling from the metal gate into the nanocrystals.

---

[12]The oxide-silicon interface is electrically active because of the existence of dangling Si bonds which can capture electrons, and a material that is known to electrically deactivate this interface is called a passivator.

[13]Chemical vapor deposition (CVD) technique is used to form the deposited oxide layer of the NC-MOS capacitors reported in this work.

[14]A higher temperature is used to get a higher quality oxide.

2.1.5 Hot Annealing

Interfacial trap states are allowed energy states in which electrons are localized in the vicinity of a material's surface. Interface traps introduce energy levels in the forbidden gap at the Si-SiO_2 interface. Experimental evidence suggests that the interfacial traps primarily arise from unsatisfied chemical bonds at the surface of the semiconductor. When the silicon lattice is abruptly terminated along a given plane to form a surface, one of the four surface-atom bonds is left dangling. The formation of SiO_2 ties up some of these bonds, and the remaining bonds become the interfacial traps.

Annealing - that is, heating - in the presence of hydrogen at relatively low temperatures minimizes the interface trap density. The annealing is accomplished through hydrogen (H_2) ambient annealing. The wafer is heated in the presence of a H_2 and N_2 gas mixture[15], a process known as forming gas annealing (FGA). In this process, a hydrogen species[16] migrates through the SiO_2 layer to the nanocrystal-oxide interface or the oxide-silicon interface where it can attach itself to a dangling silicon bond, making the bond electrically inactive.

Four different FGA temperatures are tried: 400 C, 450 C, 550 C, and 600 C. Annealing is done under 1 atm pressure, and the process takes about 30 minutes. In order to further investigate the effects of annealing, some wafers are not annealed.

2.1.6 Metallization

Once the HTO is formed, the wafer is taken out of the UHV-CVD cluster tool, and it is metallized to form the gate and the substrate contacts. The choice of metal is aluminum, a common choice in research-level MOS devices. The substrate contact is made by covering the back surface entirely with aluminum. The gate contacts are formed by using a special mask set that provides circular aluminum dots of four different diameters: 0.4064 mm, 0.8128 mm, 1.6256 mm, 3.2512 mm. Almost all of the measurements are made by using the smallest size dots, but other dots are also

[15] N_2 is not active in the annealing process, but it is used since 100% H_2 gas can be explosive.
[16] The species is thought to be atomic hydrogen.

Table 2.1: Summary of some non-standard NC-MOS Fabrication Steps

	Temperature	Pressure	Duration	Reactants
Dry oxidation	900 C	low pressure	1 hour	O_2
Nanocrystal formation	$\sim$ 600 C	1 mTorr	few mins	SiH_4
HTO deposition	750 C	400-500 mTorr	depends on t_{ox}	SiH_4 and N_2O
Hot anneal	400-600 C	1 atm	30 min	H_2 (and N_2)

measured to check if the capacitance scaling is correct. The smallest dots are the choice because they have lower capacitance, and hence higher impedance. Using a high impedance device improves the signal to noise ratio, and provides more accurate capacitance measurements.

The fabrication processes discussed in this section are summarized in Table 2.1.

2.2 Device Characteristics

As it is noted before, an NC-MOS capacitor can be modeled as a one-dimensional device since there are no source-drain regions, and the nanocrystals are laterally isolated. However, since the floating gate is not uniform, the one-dimensional problem should be studied in two parts as it is discussed in Section 2.2.1. The tunnel oxide, nanocrystals, and interface traps have important effects on the device behavior. NC-MOS capacitors also have optical properties which will be discussed in Section 2.2.6.

2.2.1 NC-MOS Electrostatics

Although it resembles an ordinary MOS capacitor, an NC-MOS capacitor has distinct electrostatic characteristics. MOS capacitors (including continuous-film floating gate devices) can be analyzed in one dimension in the gate-to-substrate direction since they are laterally uniform. However, in NC-MOS capacitors, the nanocrystals form a non-uniform discrete floating gate, and as a result of this there are two different types of regions between the aluminum gate and the back contact. One of them includes nanocrystals, and the other is free of nanocrystals. This is illustrated in Figure 2-2.

The distance between the nanocrystals varies depending on the nanocrystal sheet density. When the sheet density is 10^{11} cm^{-2}, the distance between the nanocrystals is approximately 25 nm, and when the sheet density is 5×10^{11} cm^{-2}, the distance is approximately 8 nm.

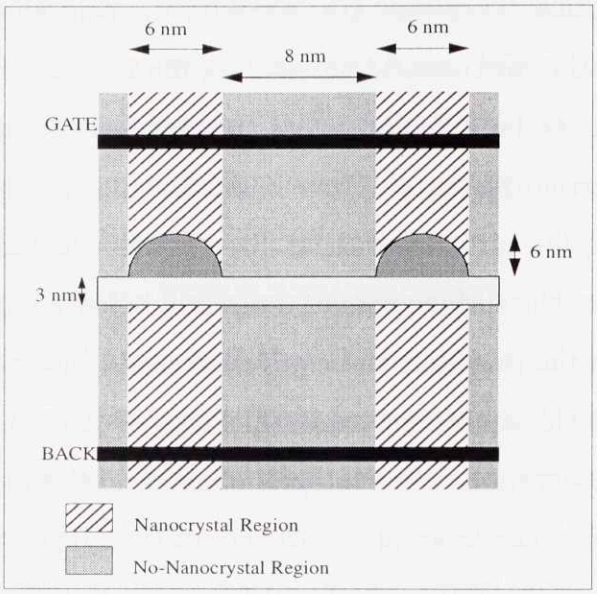


Figure 2-2: A 2-Dimensional Problem Separated into Two 1-Dimensional Problems

The exact flat-band condition at the surface can be calculated by solving the Poisson's equation. However, because of the quantum mechanical effects such as quantum confinement that are distinct to NC-MOS devices, the charge density in the Poisson's equation should account for the quantum mechanical nature of the problem. A rigorous method for solving the electrostatics of the NC-MOS capacitor is the self-consistent approximation [11]. This method assumes an initial charge distribution, and then iterates the solution of the effective mass theorem with the solution of the Poisson's equation to obtain the electrostatic potential across the NC-MOS capacitor. By this way, the surface potential conditions can be accurately calculated.

In this thesis, an approximation method is used to calculate the important quantities such as the flat-band voltage. This method approximates the charge in the nanocrystals with a continuous sheet charge located at the mid-point of the nanocrystals. More should be said about the theoretical correctness of this approximation.

The nanocrystals have a different dielectric constant from the surrounding oxide,

and there is charge stored in the nanocrystals. Therefore, the electric field varies around the NC-MOS capacitor. In the absence of the nanocrystals, in the case of a conventional MOS capacitor, the field lines are straight, and the potential is uniform in the channel. When there are nanocrystals, the electric field lines will fringe around the nanocrystals. Therefore, the length of the field lines from the gate to the channel will vary. Since electrostatic potential is the line integral of the electric field, the potential is going to be different across the channel. Another important consideration is the dielectric difference. Even if the fringing is ignored, which is a very unrealistic approach, the surface potential directly below a nanocrystal-free region is approximately 85% of the surface potential directly below a nanocrystal region. This can be seen easily by noting that the equivalent oxide thickness of the nanocrystal region is 14 nm and the equivalent oxide thickness of the nanocrystal-free region is 12 nm. With this observation, a line integral of the field from the gate to the channel yields this result. Therefore, even in the absence of the fringing fields, it is shown that there is still some spatial variation in the surface potential. These observations make it difficult to justify the uniform charge approximation (UCA) because it requires to show that the surface potential does not vary much around the channel.

It is very hard to obtain an exact analytic solution to this electrostatics problem because of the size and the shape of the nanocrystals. One approach is to use the method of images by approximating the nanocrystals with spherical conductors. The analytic solution to this problem is known, however there is an important difficulty with applying this method to this problem. The nanocrystal size is larger than the oxide thickness, and when the nanocrystal is completed to a full sphere, the sphere merges with the conducting channel. Therefore, the nanocrystals are too close to the channel to obtain correct results with the method of images. Another approach is to use a numerical solution based on electrostatic simulation. One such solution is suggested by Kohno [12]. A two dimensional simulation of the potential in the quantum dot floating gate MOS structures was carried out, and the flat-band shift for the quantum dot floating gate is compared with the flat-band shift for the uniform floating gate. It is found that the flat-band shifts are different. It is argued that the

reason for this difference is that the shape of the quantum dot greatly influences the potential distribution.

The electrostatics of this problem also depends on the nanocrystal density. It has been observed that the nanocrystal density can vary significantly across different runs. TEM measurements reveal that the nanocrystal density is usually in the range $1-10\times 10^{11}$ cm^{-2} [10, 11]. The nanocrystal size is about 6 nm. When the nanocrystal density is 10^{11} cm^{-2}, the nanocrystals occupy 3.6% of the total capacitor area. When the nanocrystal density is 10^{12} cm^{-2}, the nanocrystals occupy 36% of the total capacitor area. When the nanocrystal density is low, the amount of the charge stored in the nanocrystals is small. Since the flat-band voltage shifts are due to the charge stored in the nanocrystals, it can be expected that the hysteresis will be small. The devices studied in this thesis have shown significant hysteresis which is in the order of a few volts. From this, it can be concluded that the devices studied in this thesis have large nanocrystal densities. It is predicted that the nanocrystal sheet density for the NC-MOS capacitors used in this work is around 5×10^{11} cm^{-2}.

These observations make it difficult to provide an analytic justification for the uniform layer approximation. In this thesis, the uniform layer approximation will be used to provide a quantitative background for the flat-band shift and the nanocrystal charge. It is understood that this is a first-order approximation, and a more correct solution can be obtained by numerically solving the problem by iterating the Poisson's equation with the Schrodinger's equation as it is discussed before in this section.

To study the electrical properties of the NC-MOS capacitors, a potential is applied between the two device terminals: the gate and the back contact. Just like an ordinary MOS capacitor, an NC-MOS capacitor can be biased in three different conditions depending on the magnitude and polarity of the gate-to-substrate potential and the wafer type. These bias conditions are accumulation, depletion, and inversion. The devices used in this work are all n-type devices, and the discussion will be made for this wafer type.

In accumulation, majority carriers pile up right at the oxide-semiconductor interface, and the state of the system can be changed very rapidly since the majority

carriers, the only carriers involved in the operation of the accumulated device, can equilibrate very fast. Consequently, at standard probing frequencies of 1MHz or less it is reasonable to assume the device can follow the applied AC signal, with the small AC signal adding or subtracting a small amount of charge on the two sides of the oxide. Since the AC signal merely adds or subtracts charge close to the edges of an insulator, the charge configuration inside the accumulated NC-MOS capacitor is essentially that of an ordinary parallel-plate capacitor.

Under depletion biasing, the DC state of the NC-MOS capacitor is characterized by a negative charge on the gate and an equal amount of positive depletion layer charge in the semiconductor. The majority carriers are withdrawn from an effective width adjacent to the oxide-semiconductor interface. As in the case of an accumulated device, only majority carriers are involved in the operation of the device and the charge state inside the system can be changed very rapidly. When the AC signal places an increased negative charge on the NC-MOS gate, the depletion layer inside the semiconductor widens instantaneously, and the depletion width fluctuates about its DC value in response to the applied AC signal. This AC behavior is analogous to two parallel plate capacitors in series, the oxide and the depletion region.

Under inversion biasing, a large number of minority carriers pile up near the oxide-semiconductor interface in response to the applied DC bias. The DC width of the depletion layer maximizes at a certain value which will be referred to as W_T. The AC fluctuations of the charge in the gate can be compensated by variations either in the inversion layer charge or in the depletion layer width. A combination of the two extremes is also possible. The problem is to determine which alternative describes the AC fluctuation inside the device.

If the measurement frequency is very low, minority carriers can be generated or annihilated in response to the applied AC signal, and the time-varying AC state is essentially a succession of DC states. If the measurement frequency is very high, the relatively sluggish generation-recombination process will not be able to supply or eliminate minority carriers in response to the applied AC signal. Therefore, the number of minority carriers in the inversion layer remains fixed at its DC value and

the depletion width simply fluctuates about the W_T DC value.

2.2.2 Charge Confinement in Nanocrystals

A nanocrystal, as its name suggests, is a crystal of nano-scale dimension. As it is discussed in Section 2.1.3, nanocrystals are formed by building up a silicon mass around the so-called nucleation centers. It is predicted that when new silicon atoms are added to the silicon clump during the nucleation process, the silicon atoms arrange themselves in a crystalline order just like in the bulk silicon.[17]

Crystalline solids are generally described by using energy band models. An energy band model assumes that the crystal is large enough to form continuous bands of electron (or hole) states. Therefore, the characteristics of the nanocrystals can be described by using an energy band model. The nanocrystal is much like the silicon crystal but with a larger bandgap which is due to the quantum confinement effect. Nanocrystals act like nano-scale islands of lower bandgap in a sea of oxide. A lower bandgap results in better conductivity, and therefore nanocrystals can be thought of as quantum dots that are isolated in a region of poor conductivity, the oxide.

Solid-state theory states that the charge carriers in a crystal behave like free particles with an effective mass which is different from that of the rest mass of the free particles. The effective mass of a charge carrier is determined by the crystal properties. A consequence of this theorem is that the charge carriers in a nanocrystal can behave like free particles confined in a box which is defined by the walls of the nanocrystal. A quantum physical calculation of the eigen-energies in a confined system reveals a discrete spectrum of energies. Therefore, the nanocrystals have discrete energy levels, and the charge carriers distribute themselves in these discrete levels in agreement with the Fermi-Dirac statistics.[18]

A rough estimation of the nanocrystal bandgap can be obtained by considering the particle-in-a-box problem. If the nanocrystals are approximated by a cube with

[17] Atwater group at Caltech reports verifying the crystal structure by an interference-based measurement [13].

[18] Holes and electrons are both fermions which means that they obey the Pauli exclusion principle when they fill up the discrete energy levels of a nanocrystal.

a 6 nm edge length then the ground state energy of the electrons in the conduction band of the quantum dot, $E^0_{nc,e}$, is given by:

$$E^0_{nc,e} = E_{con,si} + \frac{2\hbar^2\pi^2}{2m^*_t L^2} + \frac{\hbar^2\pi^2}{2m^*_l L^2} \tag{2.1}$$

where $E_{con,si}$ is the energy of the conduction band edge of bulk silicon, $\hbar$ is Planck's constant, m^*_t and m^*_l are respectively the transverse and longitudinal effective masses of the electrons in the silicon conduction band, and L is size of the quantum dot which corresponds to the edge length of the cubic box in this approximation. Similarly, the ground state energy for the holes in the valence band of the quantum dot, $E^0_{nc,h}$, is given by:

$$E^0_{nc,h} = E_{val,si} - \frac{3\hbar^2\pi^2}{2m^*_{hh} L^2} \tag{2.2}$$

where $E_{val,si}$ is the energy of the valence band edge of bulk silicon, and m^*_{hh} is the effective mass of the heavy holes in the silicon valence band. The heavy hole mass is used since the light holes yield a higher energy ground state. These equations yield that the ground state energy of the electrons in the quantum dot is approximately 0.12 eV above the conduction band edge of bulk silicon, and the ground state energy of the holes in the quantum dot is approximately 0.06 eV below the valence band edge of the bulk silicon. Therefore, the bandgap of a silicon nanocrystal of 6 nm size is approximately 0.18 eV larger than the bandgap of bulk silicon. Since this energy is much greater than the thermal energy of the charge carriers at room temperature, 0.026 eV, it is reasonable to expect that quantum effects will play a role in the device operation.

2.2.3 Tunneling Through a Thin Oxide

Tunneling through thin oxides has been extensively studied in the past years. Semi-classical and self-consistent models that can account for the experimental results have been suggested. In the self-consistent model, the tunneling current from the accumu-

lation layer into a polysilicon gate through a thin oxide is studied by iterating the Schrodinger's equation with the Poisson's equation to get a self-consistent approximation. This model can be generalized to other bias conditions, and furthermore it can be extended to explain charge tunneling from the channel into the nanocrystals where the device channel refers to the surface charge layer of mobile carriers in both the accumulation and inversion bias conditions. It follows from the WKB approximation that the tunneling rate is exponentially dependent on the barrier width. Therefore, tunneling rate through thinner oxides is much higher than the tunneling rate through thicker oxides. This result is used to explain many of the NC-MOS characteristics in Chapter 4.

In general, there are two important parameters associated with the tunnel oxide: the oxide thickness and the barrier height. The thickness of the tunnel oxide is important in NC-MOS device characteristics. Two different thicknesses of tunnel oxide are used: 1.5 nm and 3 nm. Larger tunneling current densities are obtained with the thinner tunnel oxide since reducing the thickness of the tunneling barrier increases the tunneling rate. The advantage of a larger tunneling current density is that the charge/discharge and therefore write/erase times are small compared to the EEPROMs and the Flash memories. However, a thin oxide can also cause the charge stored in the nanocrystals to tunnel more easily into the channel, and this induces a low retention rate. Therefore, when the tunnel oxide thickness is chosen, the trade-off between the write/erase speed and the retention rate should be kept in mind. The barrier height for an oxide-silicon system is defined as the difference in the work functions of the oxide and the silicon. The tunneling rate gets lower with an increasing barrier height.

Although some results of the self-consistent model can be used to explain the tunneling in NC-MOS capacitors, it should be noted that the tunneling process in an NC-MOS device has different characteristics from the tunneling process in the devices theorized in the self-consistent model. The most important difference is that tunneling in an NC-MOS capacitor occurs between the nanocrystals and the device channel. The polysilicon gate offers a continuum of energy levels since it is a conductor, whereas

the nanocrystals are quantum dots that have quantized energy levels.

For tunneling to occur, there should be available hole or electron states on both sides of the oxide to accept or donate the tunneling mobile charge carriers. This means that the tunneling process is strongly dependent on the bias conditions discussed in Section 2.2.1. In the next section, an energy band model is used in order to better understand the type and the nature of the tunneling process in NC-MOS capacitors.

2.2.4 Energy Band Model

In this section, an energy band model is used to explain the characteristics of the NC-MOS capacitors. The energy band model is developed for the region which includes a nanocrystal in Figure 2-2. For the nanocrystal-free region in the same figure, conventional energy band models for MOS structures should apply. It is very important to note that within the continuous sheet charge approximation that is discussed in Section 2.2.1, the energy band diagrams which are developed in this section apply to the entire NC-MOS structure.

The major assumption of this model is that the chemical vapor deposition process produces silicon nano-crystals which are actual crystals of intrinsic silicon [13]. Once a crystal structure is assumed, the mobile charge carriers inside the nanocrystals can be modeled as free particles with an effective mass determined by the crystal properties. Since the nanocrystals are surrounded by oxide layers that have a much wider energy gap, the mobile charge carriers inside a nanocrystal are confined within a volume defined by the boundaries of the nanocrystal. Therefore, the problem is reduced to finding the eigen-energies of mobile charge carriers confined in the nanocrystal quantum dots.

The exact solution to the problem defined in the previous paragraph can be found by solving the Schrodinger's equation for a free electron (hole) with an effective mass confined in a potential dot which has the shape and the size of a nanocrystal. Since this is a problem of a confined particle, the boundary conditions impose a discrete set of eigen-energies. Therefore, the mobile charge carriers inside a nanocrystal occupy discrete energy levels in accordance with the Pauli exclusion principle.

A practical problem with this method is that the size and the shape of the nanocrystals can vary on the same wafer. A more important problem is caused by the extra set of energy levels introduced by the interface traps surrounding the nanocrystals. Although it is hard to theoretically estimate the interface trap energy spectrum, it is expected that there is significant overlap between the energy spectrums of the quantum dot and the interface traps. Therefore, the actual energy levels inside a nanocrystal are likely to be very different from the results obtained by solving the Schrodinger's equation.

The energy gap for a nanocrystal is larger than the energy gap of bulk silicon due to the quantum confinement effect. Moreover, the confined nature of the nanocrystals causes other important differences. The nanocrystal conduction and valence bands are composed of a discrete set of states rather than a continuum because of the quantum mechanical effects. The interface traps surrounding the nanocrystals also form a discrete set of energy levels. The trap states are mostly concentrated around the conduction and valence band edges since the forming gas anneal process drastically reduces the mid-gap trap density as it will be discussed in Chapter 4.

It should also be noted that although energy band models normally exist for crystalline structures, it is common practice to use energy band models also for amorphous solids such as silicon oxide. Although it is conceptually wrong, this approach enables the construction of an energy band diagram for the whole device, and the physics of the device operation can be studied conveniently by an energy band diagram. High temperature oxide and the thermal oxide are produced by different processes, but they have the same band structure since they are chemically equivalent.

The energy band diagrams change with bias conditions, and the energy band diagrams at thermal equilibrium and under accumulation, inversion and flat-band conditions will be studied.

1. **Thermal Equilibrium:** Thermal equilibrium is defined as the condition at which there is no external potential applied to the device, and the system is in detailed balance. The NC-MOS capacitor is brought to thermal equilibrium by connecting the back contact and the aluminum gate by a wire that will

allow charge transfer from one material to the other. The work-function of aluminum is smaller than the work-function of the substrate silicon, and hence, at thermal equilibrium the gate seems to be positively charged with respect to the substrate. Figure 2-3 shows the energy band diagram of an NC-MOS capacitor at thermal equilibrium.

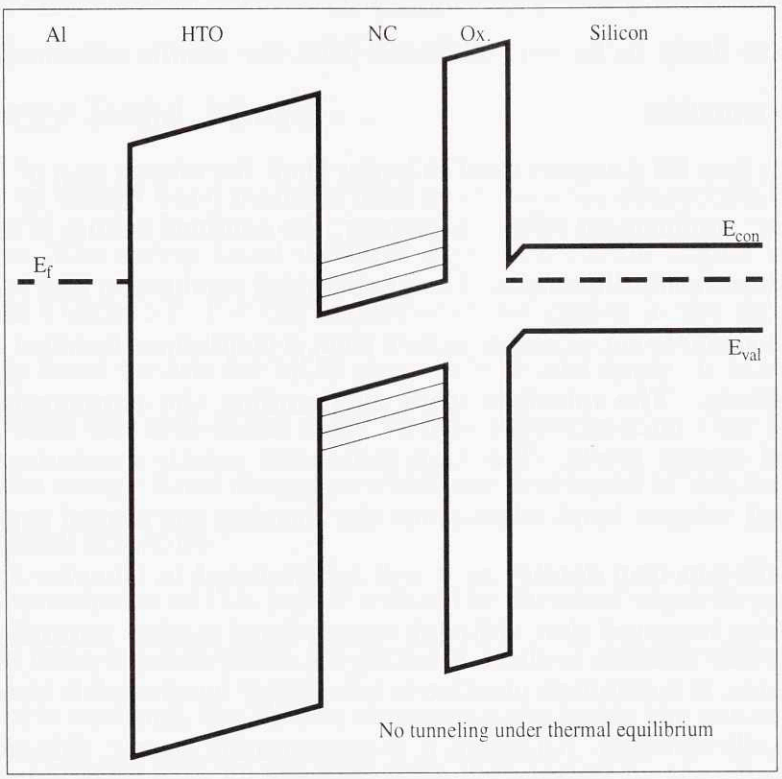


Figure 2-3: Energy Band Diagram in Thermal Equilibrium

At thermal equilibrium, the Fermi level is constant throughout the whole device. As a consequence of this, the bands are bent at the silicon surface. Under the thermal equilibrium conditions, there is no significant tunneling process since the thermal energy of the electrons (and holes) is not big enough to surmount the potential energy barrier.

2. **Flat-Band Condition:** The metal-semiconductor work-function difference can be compensated by applying a negative voltage to the gate. The magnitude of this voltage has to be equal to the built-in potential. The band diagram of the NC-MOS capacitor under flat-band condition is shown in Figure 2-4.

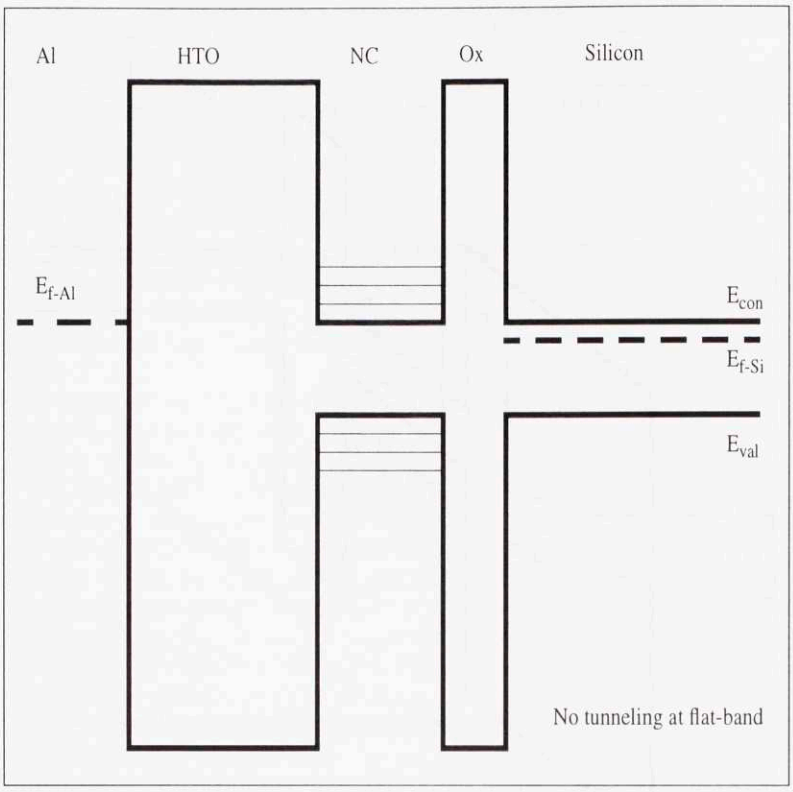


Figure 2-4: Energy Band Diagram under Flat-band Condition

Under the flat-band conditions, the silicon surface is depleted of charge carriers. Hence, there are no holes or electrons below the tunnel oxide which means that tunneling cannot occur in this bias regime. As it will be discussed in Chapter 4, the flat-band voltage varies with the amount of charge stored in the nanocrystals, and hence, in Figure 2-4 the Fermi level of the aluminum gate will move up or down in agreement with the flat-band conditions imposed by the charge stored in the nanocrystals.

3. **Accumulation Condition:** When a large enough positive bias is applied to the NC-MOS gate, accumulation bias conditions prevail. Under the accumulation bias conditions, the Fermi level is much closer to the conduction band edge of the silicon substrate, and there is a pile-up of electrons at the surface. Therefore, there are available electron states on the surface side to provide electrons that can tunnel into the nanocrystals under the effect of the electric field. The energy band diagram under the accumulation bias conditions is shown in

Figure 2-5.

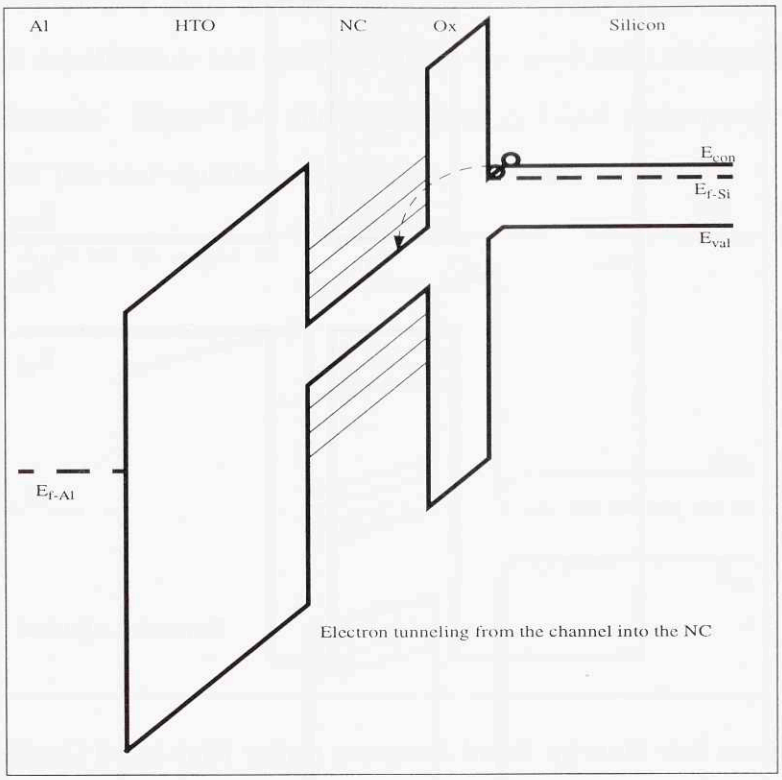


Figure 2-5: Energy Band Diagram under Accumulation Condition

4. **Inversion Condition:** When a large enough negative bias is applied to the NC-MOS gate, inversion bias conditions prevail. Under the inversion bias conditions, the Fermi level is much closer to the valence band edge of the silicon substrate, and there is a pile-up of holes at the surface. Therefore, there are available hole states on the surface side to provide holes that can tunnel into the nanocrystals under the effect of the electric field. The barrier to hole tunneling is 4.7 eV which is larger than the 3.2 eV barrier to electron tunneling. Therefore, the hole tunneling current density is lower than the electron tunneling current density. This is in agreement with the hysteresis that will be discussed in Chapter 4. It will be shown that the shift for the right-branch of the HF C-V data which corresponds to accumulation regime charging is always larger than the left-branch shift which corresponds to inversion regime charging. The energy band diagram under inversion condition is shown in Figure 2-6.

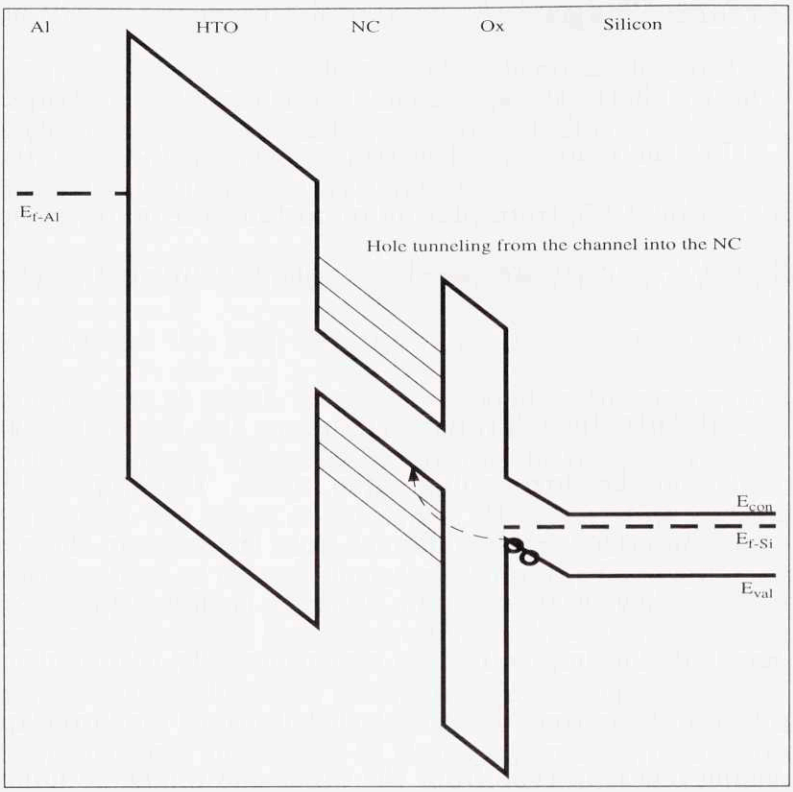


Figure 2-6: Energy Band Diagram under Inversion Condition

In summary, there is no tunneling under depletion bias conditions because there are no available holes or electrons on the substrate side of the tunnel oxide. The tunneling in accumulation bias conditions is mostly electron tunneling from the accumulated device into the nanocrystals. It can be argued that hole tunneling from the nanocrystals into the device channel can also take place since the direction of the electric field favors this transport mechanism. When the channel is accumulated there are unoccupied hole states in the channel that can accommodate the holes that tunnel down from the nanocrystals. As it will be discussed in Chapter 4, the charge transport under inversion bias condition is primarily hole tunneling from the channel into the nanocrystals.

As a conclusion, the characteristics of the tunneling process depend on the density of the available states as well as the barrier height and width. One of the goals of this thesis is to better understand the type and source of the tunneling current under various bias conditions.

2.2.5 Interface Traps

The other alternative charge storage mechanism is the interface traps. It can be seen from Figure 2-1 that there are several interfaces between the gate stack components. As discussed in Section 2.1.5, traps play an important role in device operation. There are several different traps that are associated with the different interfaces in the NC-MOS device.

1. **Nanocrystal Interface Traps:** These are the traps that are between the nanocrystals and the surrounding oxide layers. The traps in the nanocrystal-thermal oxide interface can be different from the traps in the nanocrystal-HTO interface since they originate from different chemical processes. However, it can be hard, if not impossible, to experimentally distinguish between these two trap types. Moreover, the energy distributions may overlap, and therefore distinguishing one trap type from the other one might be unnecessary for understanding the effects of nanocrystal interface traps on device characteristics. In this thesis work, the nanocrystal-HTO and the nanocrystal-thermal oxide interface traps are not treated separately, and they are considered together in one as the nanocrystal interface traps.

2. **Thermal Oxide-HTO Interface Traps:** The thermal oxide is in direct contact with the HTO in the regions that are free from nanocrystals. Although the two oxides are chemically equivalent, there can still be interface traps in between since the two materials are formed by different processes. However, it can be expected that the thermal oxide-HTO interface trap density will be lower compared to the nanocrystal interface trap density.

3. **Silicon-Thermal Oxide Interface Traps:** This type is common in all MOS capacitors. It is predicted that depositing the nanocrystals on top of the thermal oxide does not change the oxide-silicon interface states. The origin of these states is the dangling bonds of the silicon crystal which act like charge traps. For a standard MOS capacitor, forming gas annealing greatly reduces

the silicon-thermal oxide interface traps. Although having a multi-component gate structure can have some effect on the silicon-oxide interface, it is expected that this effect is minimal, and it is believed that FGA greatly reduces the silicon-thermal oxide interface trap density.

It is hard to theoretically calculate the trap energy levels, but empirical methods can be used to estimate the trap density. One method is to compare the results obtained in a bias regime at which traps can respond to the change in the bias with a regime at which they cannot. Therefore by varying the frequency of the gate potential it is possible to extract the effect of the traps. However, this method requires that there is no hysteresis effect so that all deviations from the ideal are due to the interface traps. Since there is a significant hysteresis effect in NC-MOS capacitors, this method or its derivatives do not yield correct results for these devices.

Since the traps are very important in device operation, it is important to study the variations in the device behavior for different interface trap densities. Forming gas anneal (FGA) is a processing technique that is used to change and reduce the trap density. Applied before the metallization process, FGA reduces the trap density by heating the wafer in the presence of excess hydrogen gas. The FGA process is described in Section 2.1.5, and Chapter 4 discusses the effects of FGA on device characteristics.

2.2.6 Optical Properties

The response of trap states depends also on the availability of free carriers. This is observed by contrasting the measurements made in the presence and the absence of light. Light can have two different effects on the device operation. The first one is on the device channel. In the inversion regime, a MOS capacitor can exhibit a phenomenon called deep depletion which occurs when there are not enough minority carriers available in the substrate to supply the demand for the minority carriers. Light speeds up the generation process, and the deep depletion is not observed when the measurement is made in the presence of light. Deep depletion is not an issue

for FETs since the drain and source doped regions can act as a source of excess minority carriers in these devices. The second effect of the light is on the nanocrystals themselves. A nanocrystal is a quantum dot, and therefore it has both electrical and optical properties. The aim of the work in this thesis is to explore the electrical properties of the nanocrystals, but as the light dependency measurements reveal, the photonic properties of the quantum dots have a certain effect on the electronic device behavior. Chapter 4 includes a discussion about the effect of light on the quantum dots.

Chapter 3

Measurements

There are various measurement techniques to study NC-MOS capacitors. In this thesis, electrical and optical properties of NC-MOS capacitors are explored by using capacitance and current measurements. Other conventional MOS measurement techniques such as photo-conductivity, electron probe microscopy, and corona-based capacitance spectroscopy are also known, but they are not used within the work done for this thesis. Capacitance and current measurements alone can provide sufficient information to understand and model the NC-MOS capacitors.

The next section discusses the measurement techniques used in this thesis work, and elaborates the reasons for using different techniques. Section 3.2 describes the measurement conditions including the voltage bias regimes, sweep rate, illumination and frequency. In Section 3.3, the measurement set-up, instrument specifications, and the calibration procedures are explained.

3.1 Measurement Techniques

Standard electrical measurement technique for an NC-MOS capacitor is to apply a potential difference between the two device terminals and to measure the gate current. Maxwell's equations reveal that there can be two components of the gate current in an NC-MOS capacitor: the conduction current and the displacement current. Distinguishing between these two current types can be crucial in understanding various

electrical and optical properties of NC-MOS capacitors.

In order to understand these two current components, the device under test can be modeled by a capacitor and a resistor connected in parallel where the current through the capacitor is the displacement current, and the current through the resistor is the conduction current. This simple RC model is shown in Figure 3-1.

The current that flows through the resistor is the conduction current. It can be either DC or AC current, and it occurs when an electric field exerts a Coulomb force on an electrical charge, and the charge responds by moving in the direction (or anti-direction) of the electric field. The movement of the electrical charge creates a current, and this current is known as the conduction current.

The other current component is the displacement current. When an AC electric field is applied across the NC-MOS capacitor, charge on both sides of the insulating layer varies by time, and creates an electro-motive force that can drive current into the circuit. The difference from the conduction current is that there is no charge transport across the insulator. The capacitor in Figure 3-1 accounts for the displacement current which is well explained by Maxwell's equations.

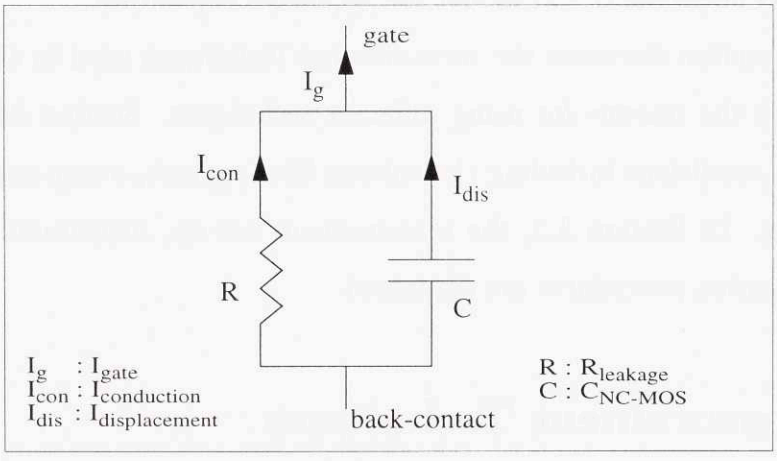


Figure 3-1: A Simplified RC Circuit Model of the NC-MOS Capacitor

The gate current is the sum of the conduction current and the displacement cur-

rent:

$$I_{gate} = I_{conduction} + I_{displacement} \tag{3.1}$$

In the case of an ideal NC-MOS (or MOS) capacitor, the conduction current is zero because of the insulating oxide layer. In fact, the most important advantage of the field-effect devices over bipolar junction devices is the almost negligible gate current of the field-effect devices. A small gate current reduces power consumption, and prevents excess heating of the chip. The heat removal problem hinders the high scale integration of junction devices such as bipolar junction transistors. Therefore, field-effect devices are the choice for VLSI systems even though they operate at lower frequencies than the bipolar devices.

Although the gate conduction current is negligible in an ideal NC-MOS capacitor, a real NC-MOS device can have current pathways that originate from imperfect chemistry of the device oxide, and this gives rise to a small current which is referred to as the leakage current[1]. It should be noted that the voltage regime used for these devices (generally between -4 V and 4 V) does not allow Fowler-Nordheim tunneling. A straightforward application of the Fowler-Nordheim tunneling equation shows that the gate-to-substrate voltage should be more than 7 V to turn on the Fowler-Nordheim tunneling current. Similarly, current due to hot electron injection is not an issue since the NC-MOS capacitors do not have source and drain regions. Therefore, the leakage current is the only possible source of a conduction current, and so the conduction current component of the NC-MOS capacitor gate current can be written as:

$$I_{conduction} = I_{leakage} \tag{3.2}$$

The displacement current component of the NC-MOS capacitor gate current can be

[1]The word leakage refers to the leakage of the oxide.

derived from the generic charge-voltage relationship for a capacitor:

$$Q_{gate} = C_{NC-MOS}V \tag{3.3}$$

where Q_{gate} is the charge stored in the NC-MOS gate, C_{NC-MOS} is the NC-MOS capacitance, and V is the external voltage applied between the gate and the back-contact. Taking the total time derivative of both sides of this equation yields:

$$\frac{dQ_{gate}}{dt} = C_{NC-MOS}\frac{dV}{dt} + \frac{dC_{NC-MOS}}{dt}V \tag{3.4}$$

It should be noted that both the voltage and the NC-MOS capacitance can be time-dependent, a situation significantly different from that of an ordinary capacitor where the capacitance value is not time-dependent. The charge/discharge centers such as the nanocrystals and the interface traps are responsible for the time-varying nature of the NC-MOS capacitance. As an example of this, electrons in the channel can tunnel into the nanocrystals under the accumulation bias conditions, and when this happens the capacitance of the device changes even though the measurement frequency and the bias is kept the same.

The displacement current, as it is defined, is the current that flows through the capacitor by the effect of time-varying fields, and it can be written as:

$$I_{displacement} = \frac{dQ_{gate}}{dt} = C_{NC-MOS}\frac{dV}{dt} + \frac{dC_{NC-MOS}}{dt}V \tag{3.5}$$

The total gate current can now be written by inserting Eq. 3.2 and Eq. 3.5 into Eq. 3.1:

$$I_{gate} = I_{leakage} + C_{NC-MOS}\frac{dV}{dt} + \frac{dC_{NC-MOS}}{dt}V \tag{3.6}$$

The conduction current and the displacement current have different characteristics which reveal different aspects of the NC-MOS capacitor. The three current components in Eq. 3.6 can be distinguished by using special measurement techniques. In

this study, three different measurement techniques are used to study the NC-MOS capacitors: high-frequency capacitance-voltage measurements, quasi-static capacitance-voltage measurements, and current-voltage measurements.

3.1.1 High Frequency C-V Measurements

High frequency measurements are accurate measurements of the device capacitance. A small amplitude time-varying signal is applied across the terminals of a voltage-biased capacitor. As the small signal input varies, the charge carriers in the NC-MOS capacitor respond, and the motion of the charge carriers reveals information about the device capacitance. The measurement is based on sensing the change in the gate current when the small signal input is changed. The impedance is then calculated by using this change. When the high frequency C-V measurement technique is used, the first term on the right hand side of Eq. 3.6 does not contribute to the measurement results since the capacitor looks like a short-circuit, i.e. the impedance of the capacitor is low, for high frequencies, and therefore negligible current flows through the resistor. Therefore, only the displacement current component can be measured in a HF C-V measurement, and the HF C-V NC-MOS capacitance can be written as:

$$C_{HF}(w) = \frac{C_{NC-MOS}(w)\frac{dV}{dt} + \frac{dC_{NC-MOS}(w)}{dt}V}{V}\frac{1}{\omega \sin\theta} \tag{3.7}$$

where ω is the measurement frequency and θ is the phase angle of the impedance that is being measured by the LC meter. It should also be noted that for large enough frequencies (like 100 kHz used in these measurements) the time dependent part of the MOS capacitance is not responsive since the carriers can not respond to very fast-changing signals.

3.1.2 Quasi-Static C-V Measurements

Another technique for the C-V measurements is the quasi-static C-V technique. A linear ramp voltage is applied across the device terminals of an NC-MOS capacitor,

and the gate current is monitored as the voltage varies. In the case of a QS C-V measurement, all three components of the gate current in Eq. 3.6 are measured. The ramp rate is slow enough (50 mV/sec) to ensure that the trap states will have enough time to respond to the change in the bias voltage. The capacitance measured by the quasi-static technique can be written as

$$C_{QS} = \frac{I_{leakage} + C_{NC-MOS}\frac{dV}{dt} + \frac{dC_{NC-MOS}}{dt}V}{\frac{dV}{dt}} \tag{3.8}$$

where the numerator shows the total measured current value and the denominator is the ramp rate. It is necessary to compare this equation with Eq. 3.7 to understand the significance of the capacitance terms in the two equations. As it is discussed in Section 2.2.1, the capacitance of a MOS capacitor is frequency dependent, and therefore the capacitance terms in these equations have an implicit frequency dependence. Therefore, the measured current is different because of the frequency-dependent nature of the MOS capacitance.

3.1.3 Current-Voltage Measurements

Current voltage measurements are very similar to quasi-static C-V measurements in origin, and they can give more explicit information about the trap states. This time, a staircase ramp is applied to the gate, and the gate current is measured by using a current meter. With this technique, the first and the last terms on the right hand side of Eq. 3.6 contribute to the measured current while the middle term does not since the bias potential is constant at each step, and the step delay time[2] is adjusted to exclude the effects of the abrupt change between subsequent voltage steps. The current measured in the I-V measurements can therefore be written as:

$$I = I_{leakage} + \frac{dC_{NC-MOS}}{dt}V \tag{3.9}$$

[2]Step delay time is the time that the current meter waits after each increment in the bias voltage. It is used to avoid the undesired effects of the abrupt change in the gate voltage as discussed in the text.

Current voltage measurements are hard to make since the gate current is very small. The current voltage measurements are used to study the trap effects, and since the trap effects reveal themselves mostly in the time-varying capacitance term, $\frac{dC_{NC-MOS}}{dt}$, in equations 3.8 and 3.9, the quasi-static C-V measurements and I-V measurements should reveal similar information about the trap effects. This is supported by the results in Chapter 4.

3.2 Measurement Conditions

Measurement conditions have an important effect on the NC-MOS test results. Many aspects of the NC-MOS characteristics can be learned by using different measurement conditions. Bias scheme, hold time, sweep rate, and illumination are the measurement conditions that are studied in this thesis. The effects of these conditions are extensively investigated.

The voltage sweep scheme plays an important role in NC-MOS device behavior. The amount of the charge stored in the nanocrystals affects the state of the device, and in response, the flat-band voltage changes across repeated tests. Therefore, it is important to establish a reference state for the nanocrystals in order to get a relevant measure of the flat-band shift. This is a particularly difficult task since the nanocrystal charging and discharging do not seem to saturate. The flat-band voltage of an NC-MOS capacitor depends upon the initial state of the nanocrystals, and it is hard to control the initial state since it can change with carrier tunneling in and out of the nanocrystals. This behavior is explained in Section 4.2.

However, some sweep techniques such as hold time are shown to improve the control over the nanocrystal charge level. Since the devices are n-type, the sweep is made from a negative voltage to a positive voltage to ensure that the inversion layer is formed before the sweep starts. The usual start voltage is -4 V, and the usual stop voltage is 4 V. In step voltage sweeps, HF C-V and I-V, the step voltage is chosen as 0.05 V. The measurements are made in a double ramp scheme such that there is a 4 V to -4 V sweep in the reverse direction following the forward sweep. Three different

Table 3.1: Outline of Measurement Conditions Used in the Experiments

	Sweep	Hold Time	Sweep Rate	Frequency	Light
HFCV	-3V→3V→-3V -4V→4V→-4V -5V→5V→-5V	5 sec 10 sec 30 sec	0.05 sec 0.1 sec 0.5 sec	100 KHz	ON OFF
QSCV	-4V→4V→-4V	10 sec	0.05 V/s 0.1 V/s 0.5 V/s	N/A	ON OFF
IV	-4V→4V→-4V	10 sec	0.1 sec	N/A	ON OFF

sweep schemes are used: -3V→3V→-3V, -4V→4V→-4V, and -5V→5V→-5V. Before the start of a sweep in either direction, the device is held at the start or stop voltage for a certain time called the hold time. For the high frequency C-V measurements three different hold times are used: 5 sec, 10 sec and 30 sec. For the quasi-static C-V and current measurements the hold time is always 10 sec.

The effects of light on the C-V and I-V characteristics are studied by illuminating the wafer with visible light during the measurement. Since the NC-MOS capacitors do not have source-drain regions that can supply minority carriers, the carrier generation/recombination stimulated by illumination plays a crucial role on device characteristics. A comparison of the device characteristics under light and in dark reveals information about the tunneling process, interface traps, and charge storage.

The high frequency C-V measurements are made at a frequency of 100 kHz. The small signal amplitude is 50 mV. The step size for the DC bias voltage is chosen as 50 mV. Three different step delay times are used to study the effects of the sweep rate: 0.05 s, 0.1 s and 0.5 s.

The effect of the sweep rate on the QS C-V characteristics is also studied. As the sweep rate changes, it is believed that the charging and discharging of the nanocrystals is altered, and the flat-band conditions change. The quasi-static C-V measurements are made at three different sweep rates to verify this prediction: 0.05 V/s, 0.1 V/s, and 0.5 V/s.

3.3 Measurement Setup

It is convenient to test NC-MOS capacitors using a probe station, and all measurements are made on a standard probe station. The probe station has light insulation, and the controllers can be moved either manually or automatically by the computer. Careful compensation and calibration eliminate the contact and cable impedance from the measured impedance. In QS C-V measurements, leakage current is monitored before the start of the measurement to make sure that it is negligible compared to the displacement current component. The block diagram of the measurement set-up is shown in Figure 3-2.

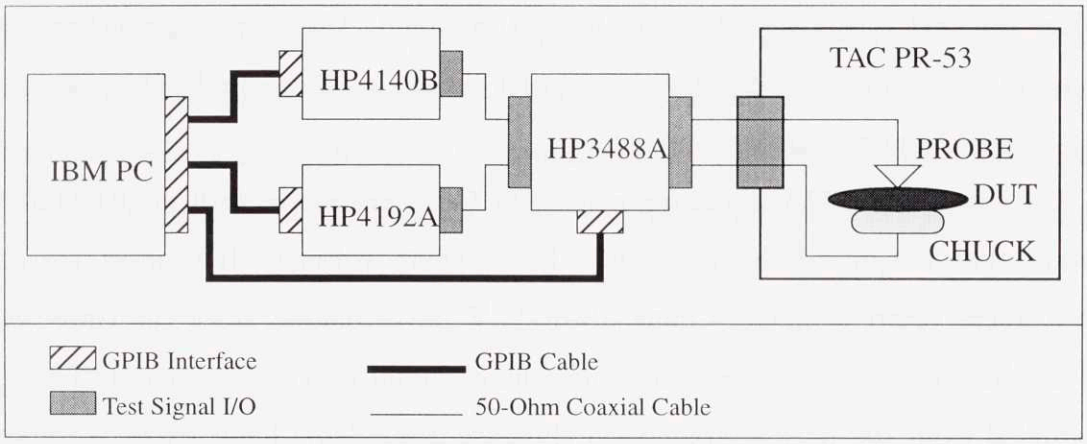


Figure 3-2: The Block Diagram of the Measurement Set-up

3.3.1 Instrumentation

The measurements are made on a Teledyne Tac PR-53 probe station. The micro-controller is magnetically mounted on the probe stand, and there is a coaxial tungsten tip of 5 mil diameter attached to the micro-controller. The tip is used to make the gate contact to the device. A coaxial tip is preferred because it minimizes the signal loss and the noise. The tip is connected to the test instruments with standard 50 Ω coaxial cables attached to the other end of the tip. The probe station has a tight lid to provide light insulation if desired. The wafers are placed on a gold-plated chuck made of molybdenum. The back of the wafers are aluminum coated as discussed

in Section 2.2, and the gold-plated chuck provides a good electrical contact for the wafers.

The coaxial cables which carry the applied and measured signals from the probe station are connected to the test instruments through an HP3488A VHF-compatible switch matrix. The switch matrix is used to reduce the amount of wiring and to run the system more efficiently.

The high frequency C-V measurements (HF C-V) are made with an HP4192A LF meter. The LF meter uses a staircase voltage output to bias the NC-MOS capacitors. The signal applied across the NC-MOS device terminals is comprised of a DC signal and an AC signal of 50 mV amplitude and 100 KHz frequency. The DC signal is used to bias the device, and the AC signal is the test signal. The source signals are applied from the back contact, and the probe contact is used as the ground. The gate current is measured from the probe contact.

The quasi-static C-V measurements (QS C-V) are made with an HP4140B pA meter. The output of the meter is a linear ramp voltage. It is most correct to use a linear ramp to make a quasi-static C-V measurement since the time-varying voltage based displacement current component should also be considered. The signal is applied from the back-contact with the gate grounded. The current is measured by using a current meter connected in series with the gate.

The current voltage measurements (I-V) are made with the same HP4140B pA meter. The scheme is similar with the quasi-static C-V measurements, except that a staircase voltage output is used instead of a linear ramp. As in the C-V measurements, signal is applied from the back-contact, and the gate is grounded. The current is measured from the gate. In all three techniques, the noise is reduced by measuring the current from the gate since the back contact is a chuck which can act like an antenna which can amplify the noise.

The entire set-up is controlled by an IBM PC running an HTBasic program. The inter-communication is done using the GPIB protocol.

3.3.2 Calibration

Direct tunneling currents play an important role in NC-MOS device operation. Since the tunneling current is small in magnitude, very sensitive calibration is necessary in order to obtain correct measurement results. During the calibration, the switch matrix between the test instruments and the probe station is set such that the circuit connecting the line from the instrument to the probe stand is closed, and so the whole assembly including the probe itself and the parasitic capacitance in air is compensated.

Standard open and short compensations are made. In the open compensation, first the probe is lifted up so that it is in the air without touching anything, and then the automatic instrument calibration is performed. In the short compensation, the probe is set in contact with the probe chuck during the calibration process. During the actual measurements, the wafer is placed on the probe chuck, and the aluminum coated back side of the wafer is in electrical contact with the chuck, and therefore short compensating from the chuck is as realistic as it can get to the real situation.

Chapter 4

Results

The measurement results are discussed in two parts. In the first part, the memory effect is demonstrated, and data obtained under various measurement conditions is used to explain the characteristics of the memory effect. In the next part, other features of the NC-MOS capacitors are described by using QS C-V and I-V characteristics.

4.1 NC-MOS Wafer Splits

NC-MOS capacitors are fabricated with different gate stack compositions and various forming gas anneal conditions in order to compare and study the dependence of NC-MOS characteristics on these variables. Because of fabrication limitations, it is not yet possible to fabricate NC-MOS capacitors with different gate stacks on the same wafer, and hence each different gate stack is fabricated on a different wafer. Furthermore, there are five different FGA conditions as discussed in Section 2.1.5. It is possible to cut a wafer into five pieces, and perform a different FGA on each piece, but in this thesis a whole wafer is used for each FGA condition.

Three gate stack splits and five FGA conditions make a total of fifteen wafers in the whole run. Ten of these wafers are NC-MOS capacitor wafers meaning that the devices on these wafers have nanocrystals embedded in the oxide. The other five wafers are the control wafers which have no nanocrystals. Five of the NC-MOS wafers have gate stack compositions of 3 nm thermal oxide, 6 nm nanocrystals and

Table 4.1: NC-MOS Capacitor Run Wafer Splits

	no FGA	FGA: 400 C	FGA: 450 C	FGA: 550 C	FGA: 600 C
30A ox / 60A NC / 90A HTO	wafer 1	wafer 2	wafer 3	wafer 4	wafer 5
15A ox / 60A NC / 90A HTO	wafer 6	wafer 7	wafer 8	wafer 9	wafer 10
30A ox / 90A HTO	wafer 11	wafer 12	wafer 13	wafer 14	wafer 15

9 nm HTO. The other five have gate stack compositions of 1.5 nm thermal oxide, 6 nm nanocrystals and 9 nm HTO. The control wafers do not have nanocrystals, and their gate stack is composed of 3 nm thermal oxide and 9 nm HTO. Each gate stack is annealed at five different conditions. There is one non-annealed wafer, and the remaining four wafers are annealed at four different temperatures: 400 C, 450 C, 550 C, 600 C.

In summary, there are three different gate stacks, and for each gate stack there are five different FGA conditions making a total of fifteen wafers in the whole run. The wafer splits are shown in Table 4.1. Table cells show the wafer numbers that are used in the fabrication facility. More information about the wafer splits can be found at Appendix A.

4.2 Memory Effect

Memory effect, the motive for researching nanocrystal devices, is the most important common characteristic of the NC-MOS capacitors. Hysteresis between C-V curves corresponding to forward and reverse voltage sweeps is evidence for variable charge storage in NC-MOS capacitors. Figure 4-1 shows the HF C-V data which demonstrates hysteresis. The measurement frequency is 100 KHz, and there is a step delay time of 0.1 s. The voltage sweep scheme is -4V→4V→-4V, and the hold time is 10 s. The light is off during the measurements. The plots in Figure 4-1 correspond to three different gate stack compositions with the same FGA condition which is 400 C.[1]

The first plot on the left shows the data taken from a control capacitor.[2] There

[1]The reason for choosing this FGA condition is explained in Section 4.2.1.

[2]As a reminder, control capacitor refers to the capacitors which have no nanocrystals embedded in, and have a gate stack composition of 30 A ox / 90 A HTO.

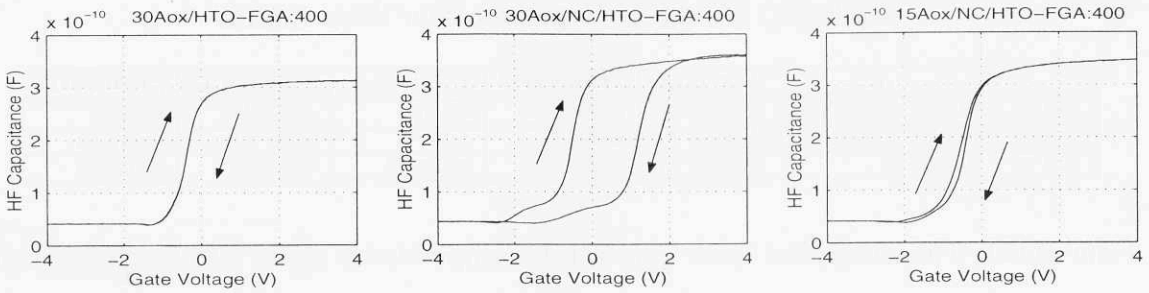


Figure 4-1: HF C-V : Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF

is no hysteresis between the two sweep directions. Therefore, a double layer oxide structure (thermal oxide and HTO) alone does not exhibit memory effect. This eliminates the possibility that charge trapped in the HTO-thermal oxide interface traps is large enough to induce memory effect.

The second plot in Figure 4-1 shows HF C-V data taken from a 30 A NC-MOS capacitor[3], and the data clearly exhibits memory effect. The left-hand branch of the C-V data corresponds to a forward sweep, and the right-hand branch corresponds to a reverse sweep. The arrows indicate the sweep direction. The data is better analyzed with the measurement conditions in mind to understand the physical origin of hysteresis.

Since the devices are n-type, measurements start with forming an inversion layer by applying a negative bias voltage to the gate. The energy band diagram under the inversion condition is as shown in Figure 2-6. During this stage, the electric field is directed upward (from the device channel towards the gate), and there are two possible charge transport mechanisms across the tunnel oxide. The first one is electron tunneling from the nanocrystals into the channel, and the second one is hole tunneling from the channel into the nanocrystals. Besides the direction and the magnitude of the electric field, charge transport also depends on the availability of electron and hole states on both sides of the tunnel oxide. The devices are held at a gate voltage of -4 V for 10 s to form the inversion layer. Once the inversion layer

[3]30 A NC-MOS capacitor refers to a device with a gate stack composition of 30 A ox / 60 A NC / 90 A HTO.

is formed, holes in the inversion layer can tunnel into the nanocrystals because the direction of the electric field is upward. For the second possible charge transport mechanism, it can be argued that the electrons in the nanocrystals can tunnel down into the channel since the inverted device channel has plenty of unoccupied electron states available to accommodate the tunneling electrons. This can be seen in Figure 2-6 by noting that the Fermi level at the substrate surface is far below the conduction band edge. This means that the occupation probability of the electronic states in the conduction band at the surface of the silicon substrate is very low, and hence, the electrons in the nanocrystal can easily find an empty state in the channel to tunnel into. Besides this, silicon/oxide interface traps are also available for electrons, and therefore there is a high chance for electron tunneling. In the following sections, the type of the tunneling charge under inversion bias conditions is investigated by using different measurement conditions.

The net effect of holding the gate voltage at -4 V is that positive charge is stored in the nanocrystals. Since the nanocrystals are embedded in the oxide, it is plausible to think of this charge as oxide charge.[4] A positive oxide charge causes a negative shift in the flat-band voltage since it increases the surface potential, and therefore enables the accumulation layer to form at lower voltages.[5]

After the 10 s hold time elapses, the gate bias is increased in the positive direction with some pre-determined rate. As the gate bias moves from -4 V towards 0 V, the NC-MOS capacitor changes its state from strong inversion to weak inversion, and it eventually reaches the depletion state where the substrate under the tunnel oxide is depleted of charge carriers. The energy bands at this flat-band condition are shown in Figure 2-4. When the depletion bias conditions prevail, there is no significant tunneling between the device channel and the nanocrystals because the electric field across the oxide is weak. This can be seen in Figure 2-4 where the bands are flat, and therefore there is no charge accumulated at the surface. The interface trap states can also cause some tunneling, but since the electric field is weak, the rate of tunneling

[4]It should not be confused with fixed oxide charge common in MOS devices.

[5]It should be kept in mind that the whole discussion in this thesis is based on n-type devices.

is too low to induce any significant charge transport. As the gate voltage keeps increasing towards 4 V, first the flat-band condition is attained at about -0.52 V. By increasing the gate potential beyond the flat-band condition, the accumulation layer is formed. As the magnitude of the downward-directed electric field increases, the device goes into the accumulation regime, and the electrons in the channel start to tunnel up into the nanocrystals. This is shown in Figure 2-5. Another possible charge transport mechanism under accumulation biasing is hole tunneling from the nanocrystals into the interface traps on the silicon surface. However, since the barrier to tunneling for holes is higher than for electrons, it can be argued that the dominant charge transport mechanism under accumulation bias conditions is electron tunneling from the channel into the nanocrystals.

This latter tunneling process attains its maximum rate when the gate voltage reaches its maximum at 4 V. The device is held at this voltage for 10 s before the reverse sweep begins. The net effect of the positive gate bias is the negative charge stored in the nanocrystals. When the reverse sweep begins, the device can reach flat-band conditions at a higher gate voltage since the negative charge in the nanocrystals screens the gate potential. The flat-band voltage for the reverse sweep is about 1.18 V.

The flat-band voltage for an ideal aluminum gate MOS capacitor with no oxide charge is equal to the metal-semiconductor work-function difference which is -0.25 V. The forward sweep flat-band voltage is about 0.27 V less than the ideal value, and the reverse sweep flat-band voltage is 1.43 V more than the ideal value. Therefore, the negative and positive flat-band shifts seen in the HF C-V characteristics are in qualitative agreement with the predictions.

The last plot in Figure 4-1 shows the HF C-V characteristics of a 15 A NC-MOS capacitor.[6] It is clear that there is still hysteresis, but it is not as large as the hysteresis observed in the 30 A NC-MOS capacitors. The reason for a much smaller hysteresis is the lower retention rate of the 15 A NC-MOS capacitors. The charge that tunnels into the nanocrystals can leak back into the channel more easily with a thinner tunnel

[6] 15 A NC-MOS capacitor refers to a device with a gate stack composition of 15 A ox / 60 A NC / 90 A HTO.

oxide. It is predicted that the tunneling rate is exponentially dependent on the oxide thickness, and therefore a 15 A oxide permits a higher tunneling rate than does a 30 A oxide. When there is more leakage, less charge is retained in the nanocrystals, and the flat-band shifts are drastically smaller.

Figure 4-2 shows the quasi-static C-V data taken from the same set of wafers. The hysteresis trends in the QS C-V characteristics closely follow the hysteresis trends in the HF C-V characteristics. The vertical shift in the capacitance levels for the forward and reverse sweeps in the second plot which corresponds to a 30 A NC-MOS capacitor is due to the DC current component that is measured in a QS C-V measurement. The details of the QS C-V characteristics will be explained in Section 4.3.

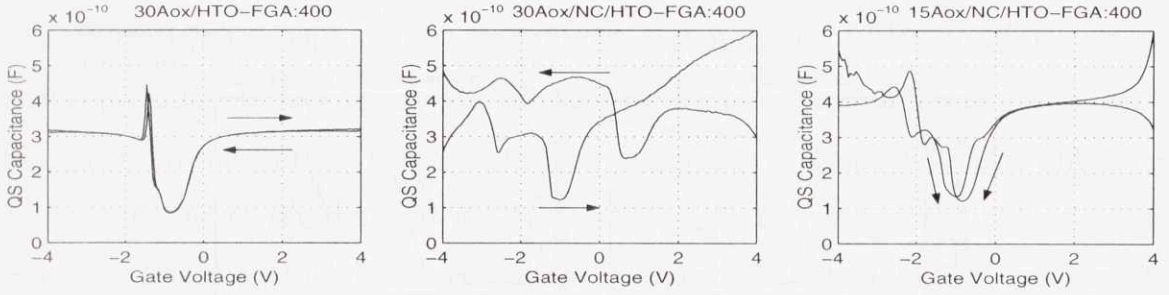


Figure 4-2: QS C-V : Memory Effect: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF

The rest of this section is devoted to analyzing the memory effect by studying the effects of FGA and measurement conditions. The discussions are based on HF C-V characteristics which are not affected by DC current components, and therefore are more easy to analyze.

4.2.1 FGA Dependence of the Memory Effect

Forming Gas Anneal is used to reduce the density of various types of interface traps in NC-MOS capacitors. The effect of interface traps on the device characteristics is studied by varying the interface trap density, and this is done by using different FGA conditions. An important goal is to compare the charge stored in the discrete states of the nanocrystal quantum dots with the charge stored in the nanocrystal interface

traps, and to understand the effects of the nanocrystal and Si/SiO_2 interface traps on the C-V characteristics.

Figure 4-3 shows the HF C-V characteristics of 30 A NC-MOS capacitors with five different FGA conditions. The measurements are made at 100 KHz with a step delay time of 0.1 s. The voltage sweep scheme is -4V→4V→-4V, and the hold time is 10 s. The light is off during the measurements.

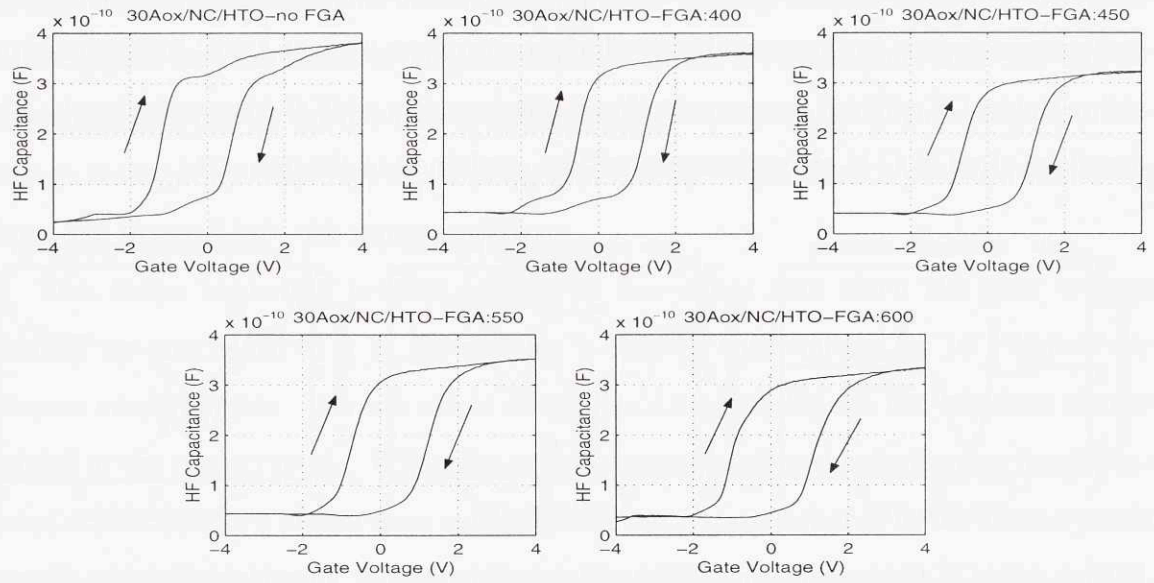


Figure 4-3: HF C-V : FGA Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF

By comparing the five plots it can be seen that there is no significant variation in the hysteresis for different FGA conditions. This can be explained since the magnitude of the hysteresis primarily depends on the tunneling rate, and the tunneling rate is not affected by the interface trap density. Tunneling rate is mostly controlled by the barrier height and width, and FGA does not change the barrier shape. On the other hand, if the tunneling is between the trap states on the opposite sides of the tunnel oxide then the trap density will determine the number of states that are available to accept or donate the charge that tunnels across the oxide. Therefore, the amount of charge that is stored in the nanocrystals may still depend on the trap density. This contradicts the FGA-independent nature of the hysteresis.

FGA seems to affect the shape of the C-V curves. The first plot on the first row

of Figure 4-3 corresponds to a non-annealed capacitor. An important observation is the vertical gap at 2 V between the forward and reverse sweep branches. The gap is largest for the non-annealed capacitor, it gets smaller for 400 C and 450 C FGA capacitors, and is large again for 550 C and 600 C FGA capacitors. A more careful examination of the C-V characteristics of the non-annealed capacitor (the first plot in the first row) reveals a hump at 0 V in the forward sweep branch. This hump in the capacitance may be due to the charging of the Si/SiO_2 interface traps with electrons. Similarly, the reverse sweep branch does not start with a flat accumulation capacitance as it normally should, but the accumulation capacitance becomes smaller as the gate voltage drops down from 4 V. The slope of the accumulation part of the reverse C-V curve may be due to the discharging of the same interface traps. The area between the two branches in the accumulation region can be due to the charge stored in silicon/oxide interface traps. The discharging starts between 3 V and 4 V which is different from the starting point of the charging which is about 0 V. This difference is primarily due to the different flat-band voltage of the forward and reverse sweep branches. Based on this discussion it can be argued that the interface trap density tends to decrease and then again increase with rising temperature. The reason is that the hydrogen atom that bonds with the dangling silicon bonds is released at high temperatures.

The discussion in the rest of this chapter, unless noted otherwise, is based on 400 C annealed capacitors since this FGA condition seems to give a lower trap density than do the other FGA conditions. It is desirable to study the memory effect with a low interface trap density since this weakens the possibility that most of the memory effect is due to the silicon/oxide interface traps.

4.2.2 Sweep Range Dependence of the Memory Effect

As discussed before, it is predicted that the tunneling rate is exponentially dependent on the gate voltage. Therefore, the voltage sweep scheme might have a major effect on the hysteresis. Three different sweep ranges are tested to understand the dependence of the tunneling rate on the gate voltage. Figure 4-4 shows the sweep

range dependence of the HF C-V characteristics. FGA condition is 400 C for all plots. The plots in the first row show C-V data collected by using a -3V→3V→-3V sweep regime, the second row of plots corresponds to a -4V→4V→-4V sweep, and the last row corresponds to a -5V→5V→-5V sweep. All measurements are made at a frequency of 100 KHz with a step delay time of 0.1 s. The hold time is 10 s, and the light is off during the measurements.

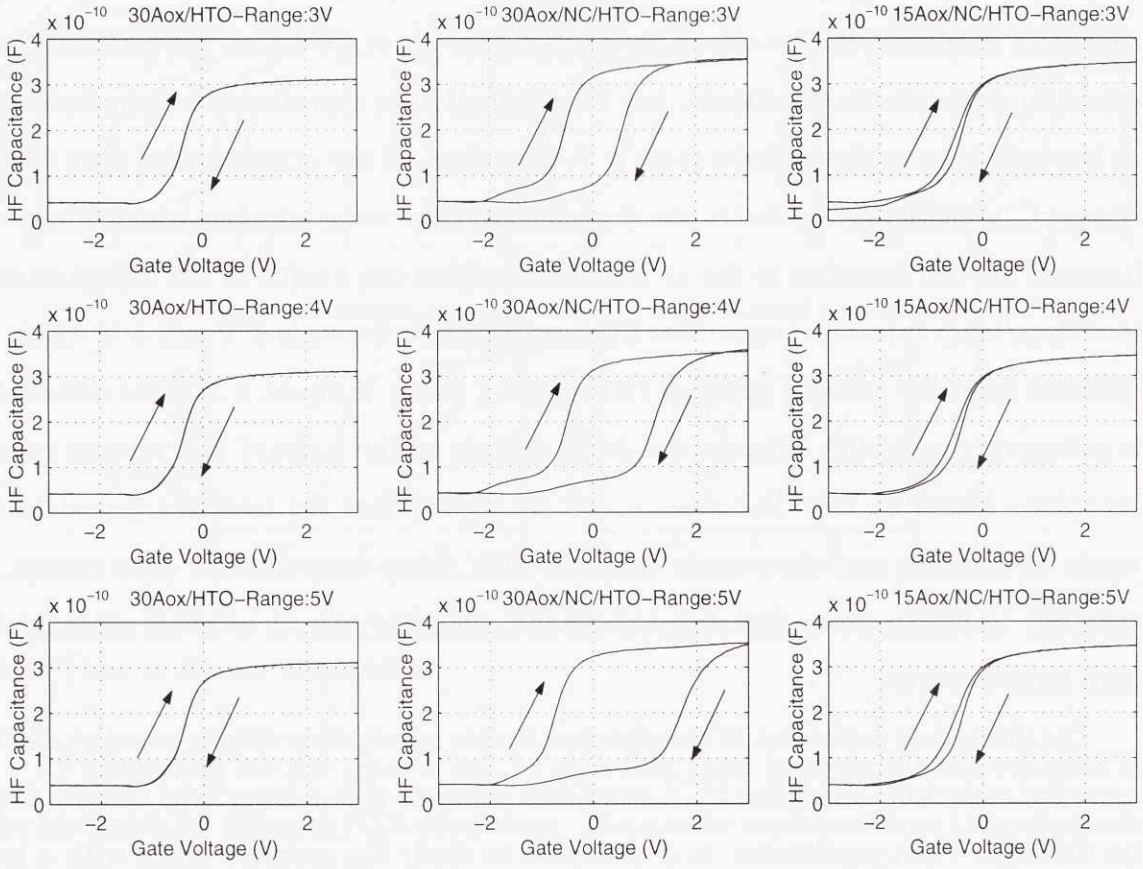


Figure 4-4: HF C-V : Voltage Sweep Range Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF

It can be seen that the memory effect is dependent on the sweep range only for 30 A NC-MOS capacitors. The hysteresis is larger for wider sweep ranges. The hysteresis in the -5V→5V→-5V sweep is approximately 1 V more than the hysteresis in the -3V→3V→-3V sweep. A close analysis of the two C-V branches shows that for wider sweep ranges, the left branch shifts more to the left, and the right branch shifts more to the right. This behavior is important, and it should be discussed in detail.

When the device is biased with a negative voltage, inversion layer forms, and charge can tunnel between the inversion layer and the nanocrystals with the help of the electric field. The amount of the total charge transport depends on the tunneling rate, the density and energy levels of the nanocrystals and the energy band conditions at the silicon surface. If the device is held at a larger negative gate bias, more charge will make it through the oxide. Therefore, it can be expected that holding the gate bias at -5 V will transport more charge than holding it at -3 V.

As it is discussed in Section 4.2, there can be two charge transport mechanisms under the inversion bias conditions: hole tunneling from the channel into the nanocrystals and electron tunneling from the nanocrystals into the channel. The rate of tunneling in both cases increases with a larger negative voltage, and therefore when the negative gate bias is larger in magnitude, there is a net positive charge transport into the nanocrystals which causes a negative shift in the flat-band voltage. This explains why the left branch of the 5 V sweep is shifted more to the left compared to the left branch of the 3 V and 4 V sweeps. However, the type of the actual charge transport mechanism, i.e. hole transport or electron transport, cannot be identified by studying the the effect of the sweep range. In the next section, the hold time dependence of the memory effect is discussed, and better insight is gained about the origin of the charge transport.

At the other end of the sweep range, the accumulation bias conditions prevail. Similarly, the right branch of the 5 V sweep is shifted more to the right of the same branches of the lower sweep ranges since the tunneling rate is higher for higher voltages. As it is discussed in Section 4.2, the charge transport mechanism under accumulation bias conditions is predominantly electron tunneling from the channel into the nanocrystals. Therefore, by holding the gate voltage at a higher positive voltage, more electrons are injected into the nanocrystals.

The HF-CV data for the 15 A NC-MOS capacitors is plotted in the last column. As it is seen, the hysteresis does not depend on the sweep range. The main reason for this is the higher charge leakage of the 15 A oxide. When 15 A capacitors are used, more charge can tunnel during the hold times, but once the sweep starts, this

charge can leak out very quickly since the oxide is thinner. With little charge left in the nanocrystals, the flat-band voltage shifts are much smaller, and the hysteresis is low in magnitude.

4.2.3 Hold Time Dependence of the Memory Effect

The amount of charge in the nanocrystals also depends on the time that the device is held at a bias voltage at which the tunneling can occur. Since the tunneling rate is higher at both positive and negative higher gate voltages, where strong inversion or accumulation bias conditions prevail, it is most convenient to vary the hold times to study the time dependence of the tunneling process.[7]

Hold time measurements are designed to study the tunneling rate and tunneling time constants of NC-MOS capacitors. Devices used in this experiment are annealed at 400 C. Measurements are made at 100 KHz with a step delay time of 0.1 s. The sweep regime is -4V→4V→-4V, and the light is off during the measurements. Three different hold times are used: 5 s, 10 s and 30 s. Figure 4-5 shows the results of the hold time measurements.

The effect of changing the hold time is more obvious with 30 A NC-MOS capacitors which are shown in the middle column. Hysteresis is larger for longer hold times. As the hold time increases, the right branch shifts more to the right, and the left branch shifts more to the left. The shift for the right branch is larger than it is for the left branch. This observation is useful to distinguish between electron or hole tunneling in inversion biasing. As it is discussed, the charge transport under accumulation biasing is primarily electron tunneling from the channel into the nanocrystals. Since the electron tunneling rate is much higher than the hole tunneling rate, it can be argued that the charge transport mechanism in inversion biasing is mostly hole tunneling from the nanocrystals into the channel since the magnitude of the flat-band shift is much lower on the left side than it is on the right side, and it is known that hysteresis primarily depends on the tunneling rate. Therefore, the hold time measurements

[7]For a discussion of what is meant by hold time, see Section 3.2.

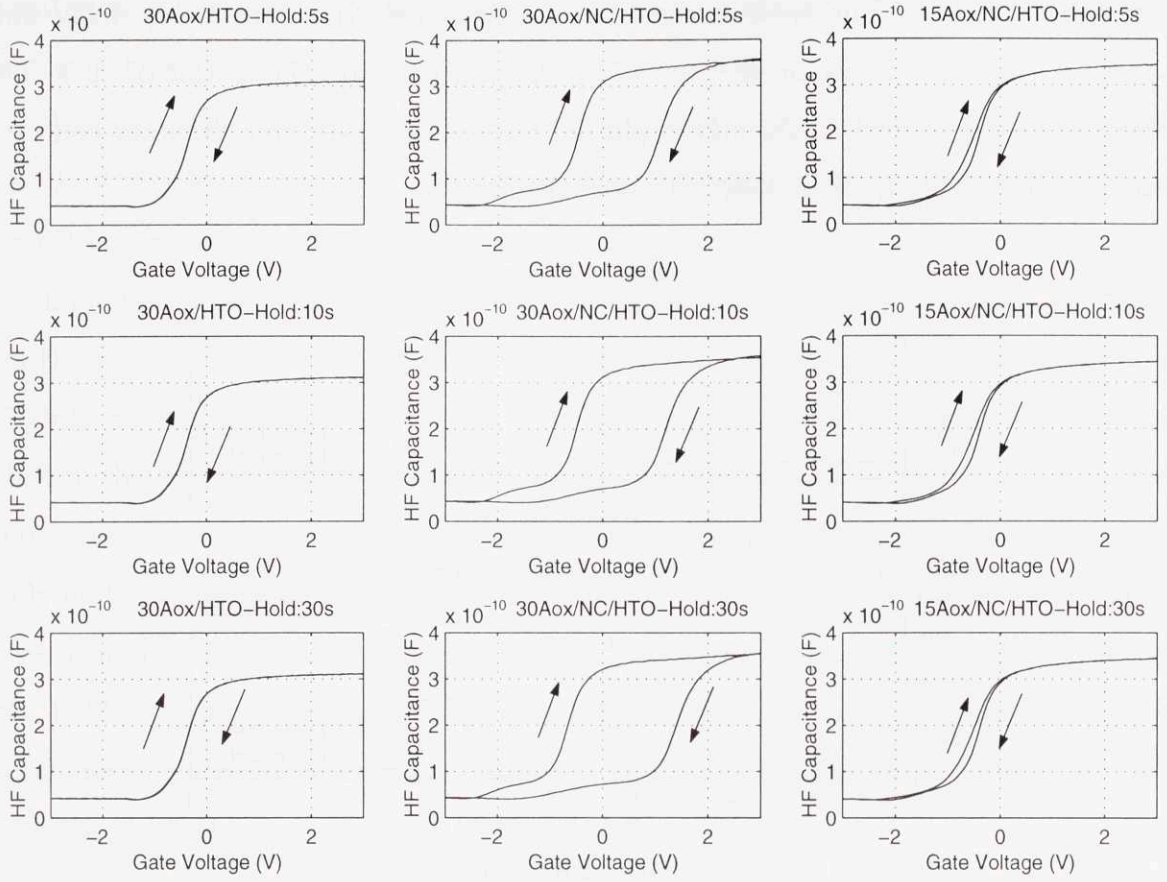


Figure 4-5: HF C-V : Hold Time Dependence of the Memory Effect: Frequency = 100 KHz; Light OFF

reveal the type of the charge transport mechanism in inversion biasing.

Changing the hold time does not effect the hysteresis for the 15 A NC-MOS capacitors because of the lower retention time as discussed in the previous section.

4.2.4 Light Dependence of the Memory Effect

Light stimulates carrier generation/recombination in a semiconductor device. Carrier generation by light has a small effect on the majority carrier concentration, which is electron concentration in this case, but it can change the minority carrier concentration by orders of magnitude. Therefore, by studying the effect of light on the HF C-V characteristics, hole and electron transport mechanisms can be better understood.

Figure 4-6 shows the light dependence of the HF-CV characteristics. Devices are

annealed at 400 C. The measurements are made at 100 KHz with a step delay time of 0.1 s. The sweep scheme is -4V→4V→-4V, and there is a hold time of 10 s. The data collected in dark (light-off) is plotted with solid lines, and the data collected under light (light-on) is plotted with dotted lines.

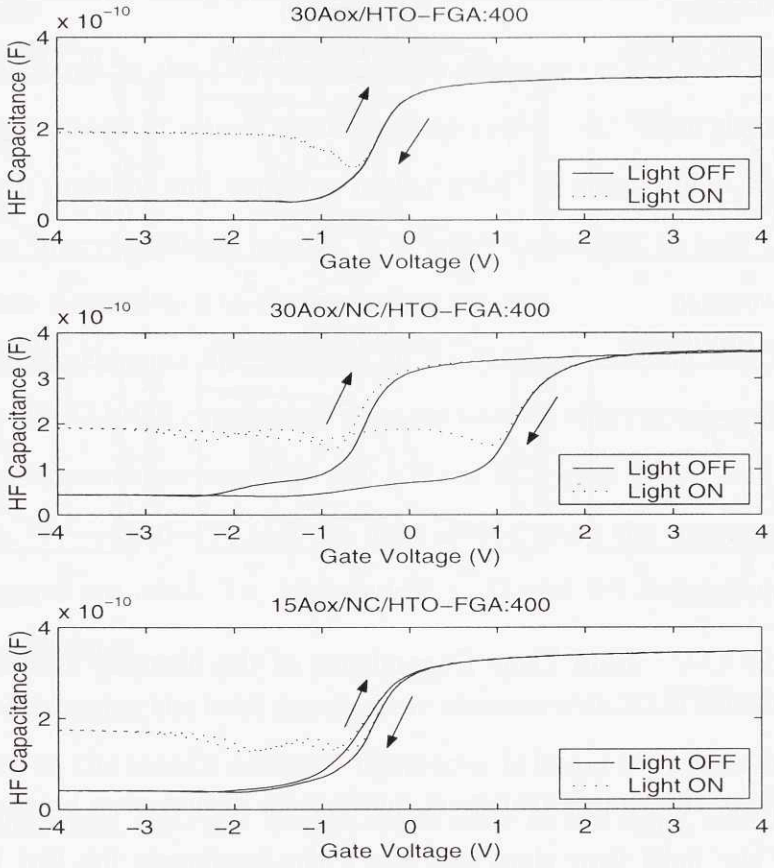


Figure 4-6: HF C-V : Light Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec

The first major point is the difference in the inversion layer capacitances of the light-on and light-off curves. The inversion capacitance is higher when the measurement is made under light. NC-MOS capacitors do not have source-drain regions which can provide minority carriers, and hence making the measurement under light significantly changes the density of holes in the inversion layer. When the hole density is higher, holes in the inversion layer can more easily respond to the AC test signal. When this happens, the inversion capacitance approaches the oxide capacitance just like in the accumulation biasing. Since the oxide capacitance is higher than the ca-

pacitance of the depletion region, the inversion capacitance measured under light is higher than the inversion capacitance measured in dark.

Another observation is the hump at -2 V in the forward sweep light-off curve of the 30 A NC-MOS capacitor. The reason for this hump is the deep depletion effect, and it is common in MOS capacitors.

Looking at the hysteresis behavior, it can be seen that light has a small effect on the forward sweep branch of the 30 A NC-MOS capacitor. The light-on curve is slightly more to the left of the light-off curve. As it is discussed above, the flat-band shift in the forward sweep direction is caused by the net positive charge stored in the nanocrystals under the inversion bias conditions. Since light has a significant effect only on the minority carrier concentration, the origin of the small shift in the forward sweep branch must be the hole tunneling. This observation supports the results of the hold time measurements which also reveal that the dominant charge transport mechanism under inversion bias conditions is hole tunneling from the nanocrystals into the channel.

4.2.5 Sweep Rate Dependence of the Memory Effect

It is plausible to think that the rate at which the voltage is swept in a HF C-V measurement can change the memory effect. Sweep rate measurements are made to understand if the hysteresis depends on the step delay time which is the time that the LC meter waits at each voltage step before it measures the capacitance. It is convenient to think of the step delay time as a parameter which sets the overall sweep rate for a HF C-V measurement. Figure 4-7 shows the sweep rate dependence of the HF C-V characteristics.

Three different step delays times are used: 0.05 s, 0.1 s and 0.5 s. All devices are annealed at 400 C. The measurements are made at a frequency of 100 KHz with a hold time of 10s. The light is off during the measurements, and the sweep scheme is -4V→4V→-4V.

There is one important observation: The hysteresis does not depend on the sweep rate. It is clearly seen that with all three step delay times, the hysteresis is the same

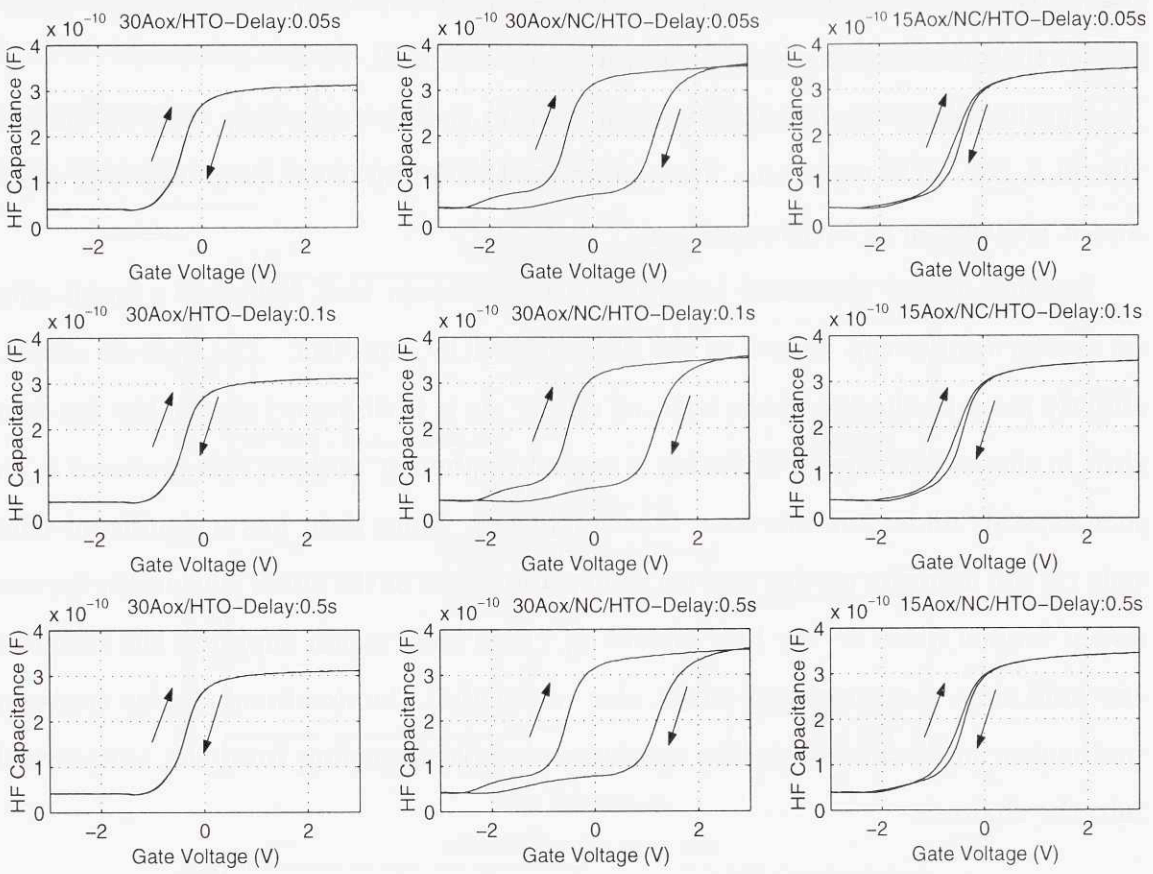


Figure 4-7: HF C-V : Sweep Rate Dependence of the Memory Effect: Frequency = 100 KHz; Hold Time = 10 sec; Light OFF

for the same device. Since the hysteresis is primarily controlled by the tunneling rate, it follows that sweep rate does not change the tunneling rate. This is an important result, and it will be used in Section 4.3.3 to analyze some QS C-V characteristics.

This result supports another remark from the discussion in the previous sections. Most of the charge transport takes place during the hold times when the gate bias is highest in magnitude. Increasing the step delay time only changes the time waited in the intermediate voltages, and gives more time to tunneling at these voltages. However, since the hysteresis does not seem to be affected by this, it can be argued that most of the charge transport, which causes the hysteresis, is taking place during the hold times, and this result supports the discussions in the previous sections.

In the next section, QS C-V and I-V characteristics are analyzed to better understand the tunneling process and the nature of the interface traps.

4.3 QS C-V and I-V Characteristics

In the previous section, the memory effect is demonstrated, and the type and the direction of charge transport under different bias conditions are explained. In this section, quasi-static C-V and I-V characteristics of the NC-MOS devices are used to explain the interface traps and other characteristics.

Figure 4-8 shows the QS C-V characteristics of three different NC-MOS capacitors with different gate stack compositions. The FGA condition is 400 C for all three plots. The voltage sweep scheme is -4V→4V and the hold time is 10 s. The sweep rate is 0.1 V/s and the light is off during the measurements.

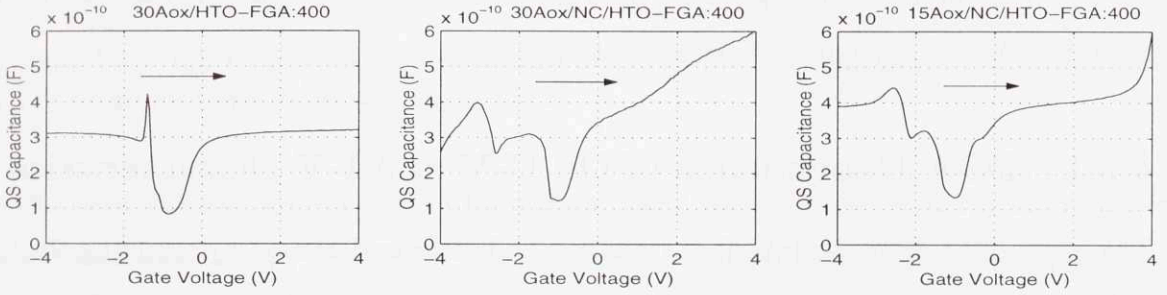


Figure 4-8: QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF

The first plot on the left side of Figure 4-8 shows the QS C-V characteristics of a control device. There are two important features on this plot. The first one is the narrow spike at about -1 V which is around the beginning of the inversion region. The second one is the small hump to the left of the bottom of the lowest depletion region capacitance. The second plot shows the QS C-V characteristics of a 30 A NC-MOS capacitor. This time, there are three important features. Starting from the left, the first one is an upward hill at about -3 V. To the right of this hill, there is a dip. Finally the accumulation capacitance has a positive slope. The last plot shows the QS C-V characteristics of a 15 A NC-MOS capacitor. The two peaks are still there, but the accumulation capacitance curve is less steep. The origins of these anomalies will be investigated by a series of measurements.

In Figure 4-9, I-V characteristics of the same set of devices are plotted with the

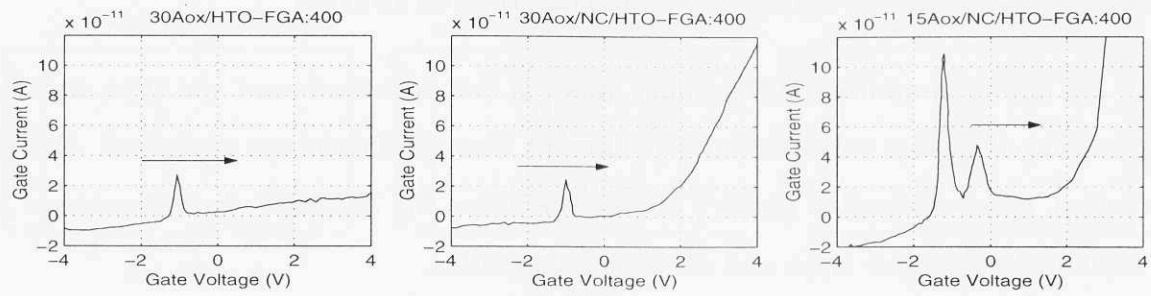


Figure 4-9: I-V Characteristics: Hold Time = 10 sec; Light OFF

same order. The control capacitor and the 30 A NC-MOS capacitor both have a narrow peak at about -1 V. The 15 A capacitor has two peaks, one higher peak on the left, and another lower peak on the right. As it will be discussed, the features of the QS C-V and I-V curves have similar origins.

4.3.1 FGA Dependence of QS C-V and I-V characteristics

The effect of FGA on the interface traps can be better understood by studying the QS C-V and I-V characteristics. Figure 4-10 shows the QS C-V data for 30 A NC-MOS capacitors with 5 different FGA conditions. The voltage sweep scheme is -4V→4V and the hold time is 10 s. The sweep rate is 0.1 V/s, and the light is off during the measurements.

The first thing to note is the variation in the depletion capacitance. The depletion capacitance, which is the capacitance at about -1 V, is lowest for 400 C and 450 C annealed capacitors, and it is highest for the non-annealed capacitor. Theoretically, depletion capacitance should be lowest when the surface is depleted of charge carriers so that the gate charge is accounted for entirely by the charge at the bottom of the depletion region. When there is a high density of silicon/oxide interface traps, the Fermi level gets pinned at the oxide, and some of the gate charge is accounted by the charge at the interface traps. This causes a higher depletion capacitance since the effective capacitance for the charge in the interface traps is the oxide capacitance.

It is known that FGA is much more effective in annealing the mid-gap traps than the band-edge traps. FGA reduces the density of the traps that have energy

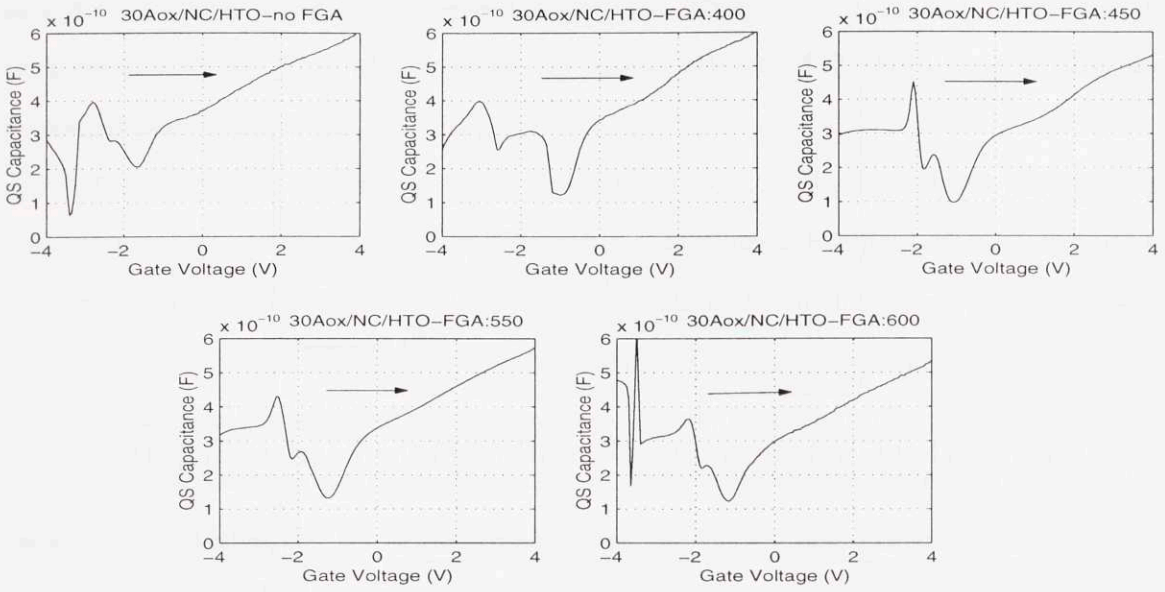


Figure 4-10: QS C-V : FGA Dependence of QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF

levels near the mid-gap energy of the silicon more than it reduces the density of the traps that have energy levels near the conduction and valence band edges. When the capacitor is biased near the flat-band conditions, the Fermi level is in the middle of the bandgap, and hence, most of the mid-gap traps are filled up. Therefore, it can be concluded that the higher depletion capacitance of the non-annealed wafer is due to the charge stored in the mid-gap traps at the silicon/oxide interface.

Another feature is the ringing at the high end of the inversion region. As the FGA temperature is increased, the spike width gets smaller, and then it goes up again. This signals that the trap density tends to increase with the increasing FGA temperature beyond 450 C. A qualitative description of the observed behavior is shown in Figure 4-11.

The I-V measurements also support these results. Figure 4-12 shows the I-V data for 3 different FGA conditions. The peaks in these plots have the same origins as the ringing in the QS C-V characteristics, and as it can be seen it is the highest for the non-annealed capacitor, it is small for the 400 C annealed capacitor, and is large again for the 550 C annealed capacitor. These results are in agreement with the FGA dependence of the interface trap density illustrated in Figure 4-11.

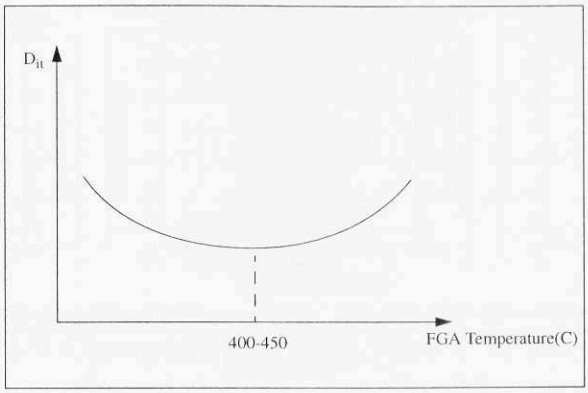


Figure 4-11: FGA Temperature Dependence of the Interface Trap Density

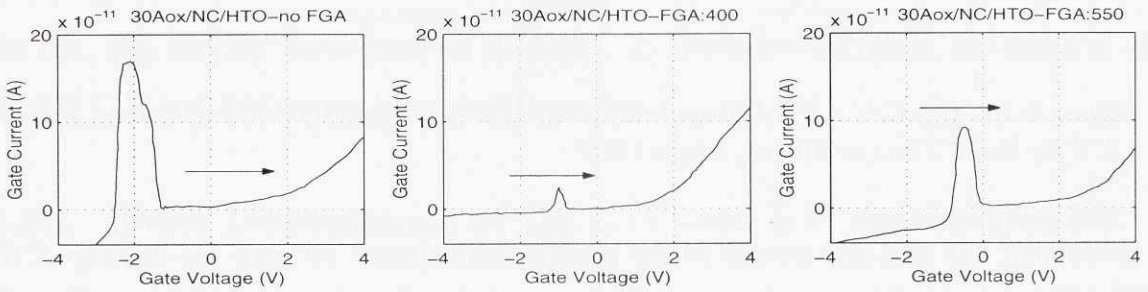


Figure 4-12: IV : FGA Dependence of I-V Characteristics: Hold Time = 10 sec; Light OFF

Based on the results of the QS C-V and I-V measurements, it can be concluded that the anomalies in the inversion region are due to the charging/discharging of the silicon-oxide band edge interface traps. The interface between the oxide and the silicon substrate is higher in quality in a conventional MOS device. The high interface trap density of the NC-MOS capacitors might be due to the nanocrystal deposition or to the UHV-CVD cluster tool that is used to grow the thermal oxide. Oxide growing techniques are optimized for special furnaces, and growing the oxide in the cluster tool is likely to produce poorer results.

4.3.2 Light Dependence of QS C-V and I-V Characteristics

The peculiar features in the inversion regime QS C-V and I-V characteristics can be explained by studying the effect of light. Figure 4-13 shows the light dependence of

the QS C-V characteristics for three different NC-MOS capacitors with different gate stack compositions. The voltage sweep scheme is -4V→4V, the hold time is 10 s, and the sweep rate is 0.1 V/s.

As it is discussed in Section 4.2.4, the major effect of the light is to increase the minority carrier concentration. The first column of plots in Figure 4-13 corresponds to the control capacitors. As it can be seen the single spike in the upper plot goes away if the light is kept on during the measurement. The holes generated by light flood into the interface traps at the silicon-oxide interface, and when the traps are flooded, trap charging/discharging that gives rise to the spike cannot take place. Another possible explanation for the spike can be the interface traps between the HTO and the thermal oxide, but to charge/discharge these traps, some tunneling must occur, and therefore, if the origin of the spike were the interface traps, it would not go away so easily by turning on the light.

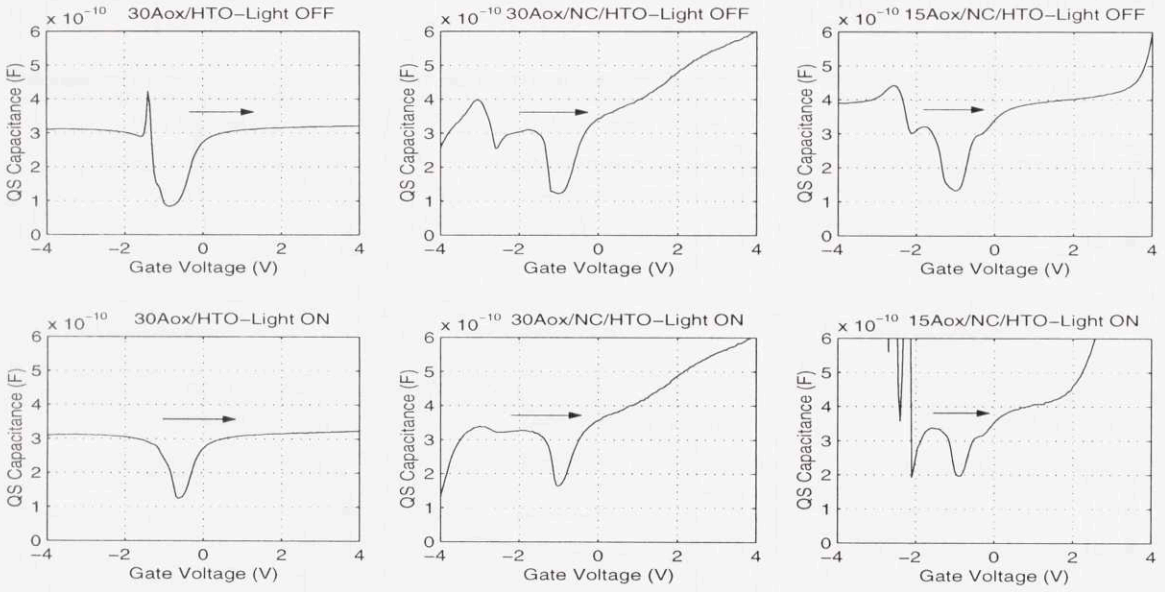


Figure 4-13: QS C-V : Light Dependence of QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec

The plots in the middle column in Figure 4-13 show the light dependence for a 30 A NC-MOS capacitor. As it can be seen the double-peaks go away when the light is on. With a similar argument, these peaks must be due to the charging/discharging of the silicon-oxide interface traps.

The last column in Figure 4-13 is from a 15 A NC-MOS capacitor. It is more difficult to obtain quasi-static data with 15 A capacitors because of the large tunneling current density. The hump at about 0 V is due to the large tunneling current density of 15 A capacitors. Unlike the other features, the hump does not go away with light, and therefore it is not a trap-related feature. The origin of the hump is most likely the electron tunneling from the channel into the nanocrystals. The tunneling rate is much higher for 15 A capacitors, and therefore the hump is there only with a thin oxide.

The discussion so far is well-supported by the I-V data. Figure 4-14 shows the I-V data taken from the same set of wafers. As it can be seen, there is a single spike for the control capacitor, and it goes away with light. The situation is the same for the 30 A capacitor. On the other hand, there are two spikes for the 15 A capacitor, and only one of these spikes goes away with the light, and the other spike is not affected by turning on the light.

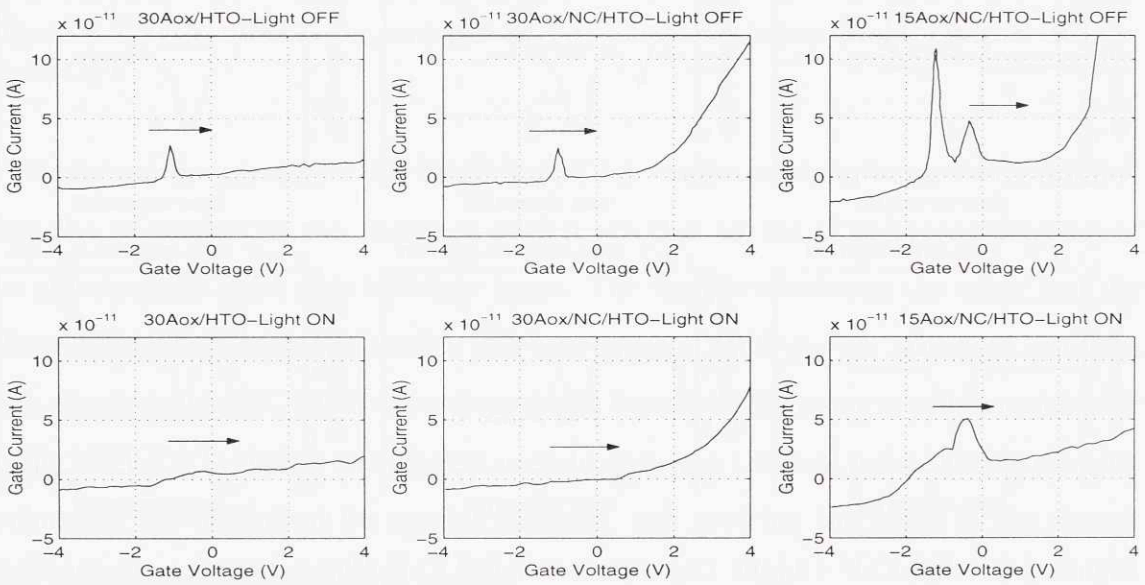


Figure 4-14: I-V : Light Dependence of I-V Characteristics: Hold Time = 10 sec

The features that change with light are due to the interface traps, and the features that don't change with light are due to the direct tunneling of electrons. Therefore, the spike that remains with light for the 15 A capacitor must be due to electron

tunneling, and it should have the same origin as the hump in the QS C-V data at about 0 V.

4.3.3 Sweep Rate Dependence of QS C-V Characteristics

The QS C-V sweep rate dependence measurements help to gain insight for the time constants related with the charging/discharging of the interface traps. Three different sweep rates are used: 0.05 V/s, 0.1 V/s and 0.5 V/s. Figure 4-15 shows the sweep rate dependence of the QS C-V characteristics for three different NC-MOS capacitors with different gate stack compositions. The voltage sweep scheme is -4V→4V, the hold time is 10 s, and the light is off during the measurements.

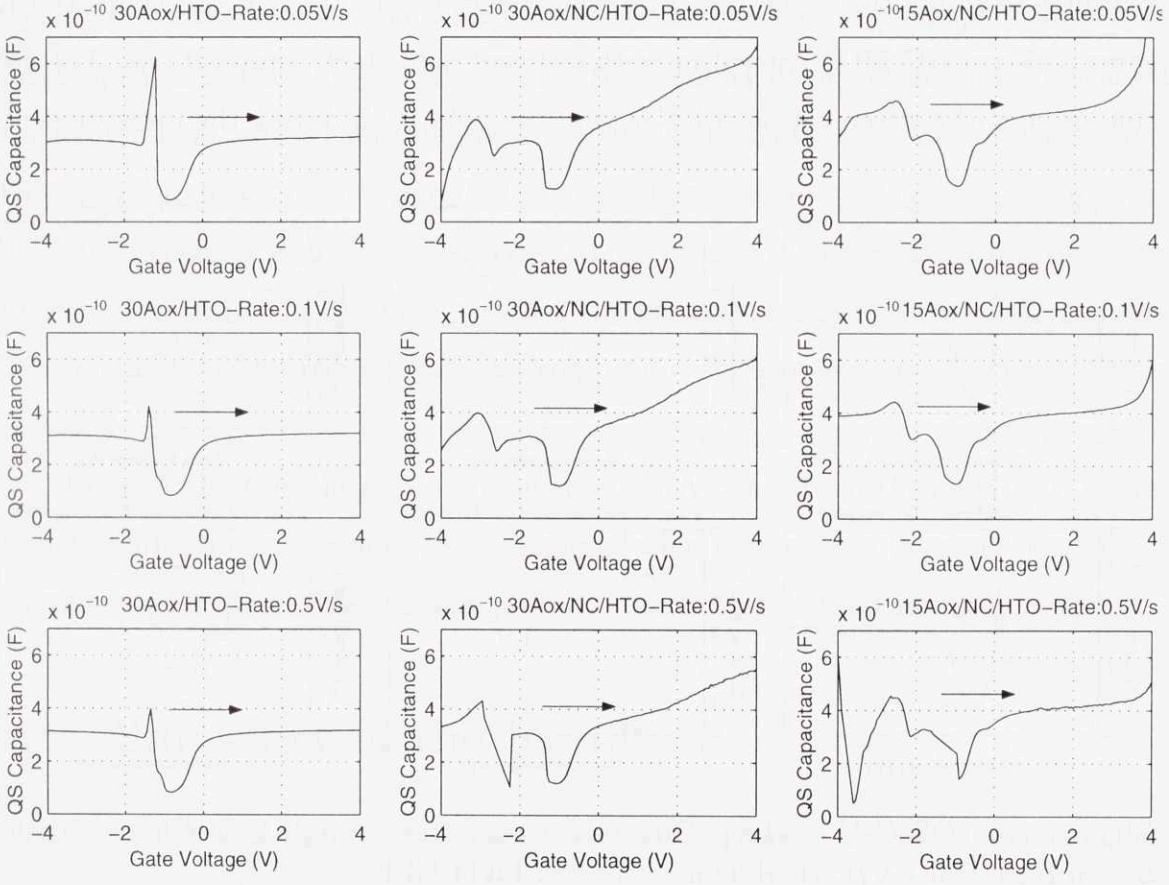


Figure 4-15: QS C-V : Sweep Rate Dependence of QS C-V Characteristics: Hold Time = 10 sec; Light OFF

The plots in the first column correspond to a control capacitor. As it can be seen,

the spike gets smaller with the increasing sweep rate. The traps have less time to charge/discharge when the sweep rate is higher, hence the spike is smaller for higher sweep rates. The results of the sweep rate measurements support the observation that the strange features in the inversion QS C-V characteristics are due to the interface traps.

4.3.4 Sweep Direction Dependence of QS C-V Characteristics

The discussion so far has been on forward sweeps. In this section, QS C-V sweeps in the reverse direction are compared with the QS C-V sweeps in the forward direction. Figure 4-16 shows the sweep direction dependence of the QS C-V characteristics for three different NC-MOS capacitors with different gate stack compositions. The hold time is 10 s, the sweep rate is 0.1 V/s, and the light is off during the measurements.

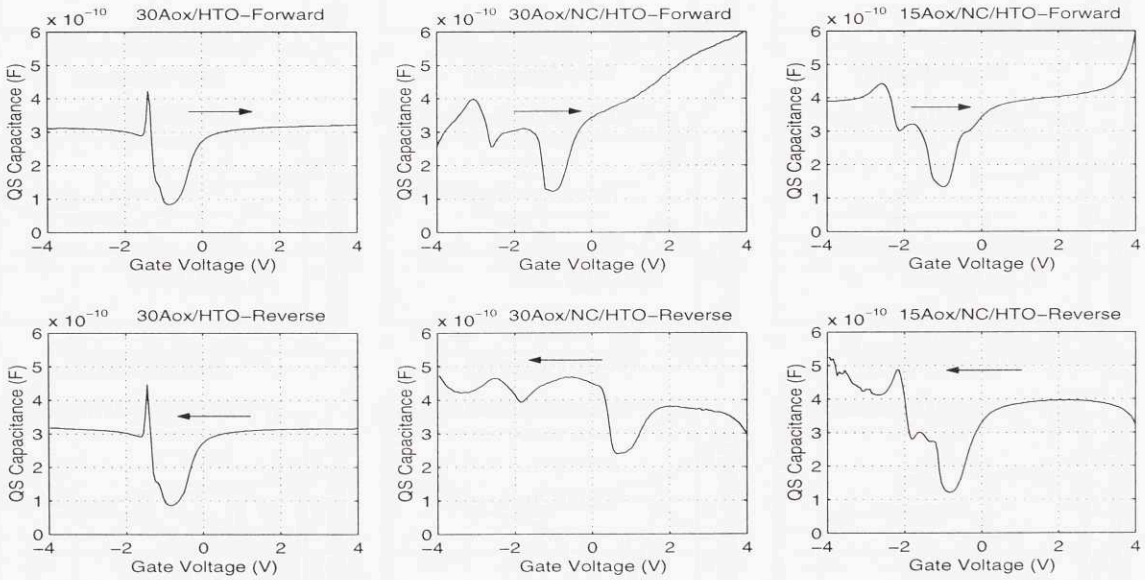


Figure 4-16: QS C-V : Sweep Direction Dependence of QS C-V Characteristics: Sweep Rate = 0.1 V/s; Hold Time = 10 sec; Light OFF

The most significant difference is the change in the accumulation region slope for the QS C-V curve obtained from 30 A NC-MOS capacitors. In the forward sweep, the accumulation capacitance has a positive slope. In the reverse sweep, it is almost flat.

The origin of the positive slope is the electron tunneling from the nanocrystals into the channel. When electrons tunnel into the nanocrystals, and stay there, the amount of the gate charge that is being compensated in the nanocrystals will be larger, and this will increase the capacitance between the two device terminals since the distance from the gate to the nanocrystals is smaller than the distance from the gate to the channel, and the dielectric constants of the HTO and the thermal oxide are identical.

The reverse sweep does not exhibit this positive slope. Prior to this sweep, the device is held at 4 V for 10 s, a time long enough to fill up the nanocrystals such that most of the gate charge is compensated by the electrons in the nanocrystals. As the sweep starts, the electrons in the nanocrystals tunnel back into the channel which would normally decrease the QS capacitance. However, as it is seen in the figure, the capacitance stays constant. This means that the discharging of the nanocrystals does not happen as fast as the charging. Using the memory device terminology discussed in Chapter 1, it can be said that the erase time is much longer than the write time. This is an important observation that might put some constraints on the usage of the NC-MOS capacitor as a memory device. Namely, since the erase and write times are not compatible, it is very difficult to use the nanocrystal memory as main memory since main memories require fast and compatible write/erase times as discussed in Section 1.1.3.

Although the tunneling rate is higher in 15 A capacitors, their QS C-V curves do not exhibit a positive slope in the accumulation region since the retention time is very short, and the charge tunneling into the nanocrystals leaks out very quickly.

4.4 Summary of the Results

This section gives a summary of the results that have been discussed so far. It has been shown that the hysteresis is due to the charge stored either in the nanocrystal quantum dot states or in the interface traps that surround the nanocrystals. The tunneling in the accumulation bias regime is electron tunneling from the channel into the nanocrystals. The tunneling in the inversion bias regime is hole tunneling from

the channel into the nanocrystals. There is no significant tunneling in other bias regimes. The tunneling rate is higher with a thinner oxide, but the retention time is also lower since the charge can leak out more easily. A higher voltage and a longer hold time causes a larger tunneling current density. Charging of the nanocrystals is faster than the discharging. It is also demonstrated that the interface trap density is reduced by the FGA process. There is a U-shaped relation between the interface trap density and the temperature which is shown in Figure 4-11.

Chapter 5

Discussion

The goal of this thesis is to understand the characteristics of the memory effect which is evident in NC-MOS capacitors. Several factors such as the oxide thickness, the FGA condition and the measurement conditions can influence the flat-band voltage shift, which is the major evidence of the memory effect, and these are discussed in the previous chapters. In this chapter, a more quantitative approach is taken to analyze the memory effect. The flat-band voltage shift and the corresponding nanocrystal charge are calculated for a particular HF C-V measurement with certain conditions which are given below. The data shown in the last chapter is used to extract the electrical thickness, the flat-band voltage shift and the oxide charge amount in the NC-MOS capacitors.

As it is discussed in Section 2.2.1, a quantum mechanical method is needed to study the electrostatics of the NC-MOS capacitors. Self-consistently solving the electrostatics problem by iterating the Schrodinger's equation with the Poisson's equation will be the focus of future research in this field. As it is discussed in Section 2.2.1, in this thesis work an approximation method is used to calculate the flat-band shift and the amount of charge stored in the nanocrystals. This method is very similar to the well-known oxide charge approximation used with MOS capacitors. Namely, the charge in the nanocrystals is approximated as a continuous sheet of charge that is located at the mid-point of the nanocrystal, and then the HF forward and reverse sweep branches of the NC-MOS capacitors are compared with the flat-band voltage

of an ideal capacitor to approximate the flat-band shift.

The derivations are made for all three gate stack compositions that are annealed at 400 C. To minimize the off-set effects of the DC current, the interpolations are based on the HF C-V data. Signal-to-noise ratio is maximized by using the smallest size dots, which are 0.4064 mm in diameter. The hold time is 10 s, the step delay time is 0.1 s, and the light is off during the measurements.

5.1 Electrical Thickness

The electrical thickness is defined as the equivalent oxide thickness of the gate stack, and it is derived from the HF C-V capacitance in accumulation. The goal of calculating the electrical thickness is to verify the predicted gate layer thickness. The gate stack layers all have sub-10 nm range thicknesses, and since the fabrication in this length scale is prone to errors, the estimated gate stack thicknesses should be verified. As it was discussed in Chapter 2, an NC-MOS capacitor has two different types of regions between the gate and the back contact. This is shown in Figure 2-2. The two regions have different thickness because of the discrete nanocrystal array. The electrical thickness will be compared to the predicted thickness which is calculated by using the continuous sheet charge approximation. In this approximation, it will be assumed that the nanocrystals form a continuous film of 6 nm thickness on the tunnel oxide.

The HF capacitance in accumulation can be written as the capacitance of an equivalent oxide:

$$C_{ox} = \frac{\varepsilon_{ox} A}{t_{ox}} \tag{5.1}$$

where A is the area of the capacitor, ε_{ox} is the oxide dielectric constant and t_{ox} is the oxide thickness that is being investigated. Using this equation t_{ox} can be written as:

$$t_{ox} = \frac{\varepsilon_{ox} A}{C_{ox}} \tag{5.2}$$

The area of the capacitor is 1.3×10^{-3} cm^2, and the dielectric constant of SiO_2 is $3.9 \times 8.854 \times 10^{-14}$ F·cm^{-1}. The oxide capacitance can be extracted from the C-V plots.

The approximated value of the electrical thickness is compared to the predicted processing thickness. There are two approaches to calculate the predicted oxide thickness. The first one is based on the uniform layer approximation. To proceed, the thickness of the nanocrystal layer is scaled by the ratio of the silicon and oxide dielectric constants, and an equivalent oxide thickness is obtained:

$$\begin{aligned} t_{NC}^{eq} &= \frac{\varepsilon_{ox}}{\varepsilon_{si}} t_{NC} \\ &= \frac{3.9 \times 8.854 \times 10^{-14}}{11.9 \times 8.854 \times 10^{-14}} t_{NC} \\ &\cong \frac{1}{3} t_{NC} \end{aligned} \tag{5.3}$$

The equivalent oxide thickness of the nanocrystals is then added to the predicted tunnel oxide and control oxide thicknesses to find the total predicted oxide thickness. As it was discussed before, the predicted tunnel oxide thickness is either 15 A or 30 A, and the predicted control oxide thickness is 90 A.

A problem with this approach is that when the nanocrystal sheet density is 5×10^{11} cm^{-2}, the nanocrystals occupy 18% of the total capacitor area. Therefore, more than 80% of the capacitor area is free of nanocrystals, and it has the gate stack composition of the tunnel oxide and the control oxide. Therefore, it is more reasonable to expect that the predicted thickness will look more like the thickness of the nanocrystal-free region. However, the measured values of the electrical thickness are different for nanocrystal and control devices. This can also be due to the errors in the fabrication process, but to a first order approximation, this thesis will present the predicted oxide thicknesses based on the uniform layer approximation. The electrical and predicted thicknesses for three different NC-MOS capacitors are shown in Table 5.1.

The measured electrical thickness is usually within 80% agreement with the predicted values. The error is due to the use of UHV-CVD cluster tool. The oxide growth

Table 5.1: Electrical and Predicted Oxide Thickness

	Oxide Capacitance(pF)	Electrical t_{ox}(A)	Predicted t_{ox}(A)
30 A ox / 90 A HTO	310	145	120
30 A ox / 60 A NC / 90 A HTO	360	125	140
15 A ox / 60 A NC / 90 A HTO	345	130	125

and deposition processes are optimized for furnace tools, and the use of UHV-CVD cluster tool might result in a higher error rate.

5.2 Flat-band Voltage Shift

The flat-band voltage shift is the major evidence for the memory effect. It is possible to extract the flat-band voltage from the data by using the continuous sheet charge approximation. The HF capacitance under the flat-band condition can be written as:

$$C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\varepsilon_{si} A}} \tag{5.4}$$

where L_D is the extrinsic Debye length. The extrinsic Debye length for silicon is a function of the donor density in the substrate and the temperature:

$$L_D = \left(\frac{kT\varepsilon_{si}}{q^2 N_D} \right)^{1/2} \tag{5.5}$$

where N_D is the donor density which is 10^{16} cm^{-3} in this case. Using this value, the Debye length at room temperature can be found as:

$$\begin{aligned} L_D &= \left(\frac{0.0259\text{eV} \cdot 11.9 \times 8.854 \times 10^{-14}\text{Fcm}^{-1}}{(1.6 \times 10^{-19}\text{C})^2 10^{16}\text{cm}^{-3}} \right)^{1/2} \\ &= 4.14 \times 10^{-6} \text{ cm} \end{aligned} \tag{5.6}$$

The flat-band capacitance for three different gate stacks is shown in Table 5.2.

Once the flat-band capacitance is known, the flat-band voltage can be extracted from the data by considering that the flat-band voltage is the voltage where the HF capacitance is equal to the flat-band capacitance. There are two flat-band voltages

Table 5.2: Calculated Flat-Band Capacitance

	Oxide Capacitance(pF)	Flat-band Capacitance(pF)
30 A ox / 90 A HTO	310	160
30 A ox / 60 A NC / 90 A HTO	360	172
15 A ox / 60 A NC / 90 A HTO	345	169

for each -4V→4V→-4V sweep, one for the forward sweep and another for the reverse sweep. The flat-band voltage is used to determine the flat-band voltage shift from the flat-band voltage of an ideal MOS capacitor. The flat-band voltage for an ideal MOS capacitor is the metal-semiconductor work-function difference, Φ_{MS}:

$$\Phi_{MS} = \Phi_M - (q\chi - \frac{E_g}{2} - q\,|\phi_n|) \tag{5.7}$$

Here, Φ_M is the metal work-function (4.1 eV), and the quantity in parenthesis is the semiconductor work-function where χ is the electron affinity (4.15 V), E_g is the silicon bandgap (1.12 eV), and ϕ_n is the bulk potential which is given as:

$$|\phi_n| = \frac{|E_f - E_i|}{q} = \frac{kT}{q} \ln \frac{N_D}{n_i} \tag{5.8}$$

where n_i is the intrinsic electron density, and N_D is the dopant density. Bulk potential corresponding to a dopant density of 10^{16} cm^{-3} is 0.36 V. Using these values yields:

$$\Phi_{MS} = -0.25 \text{ eV} \tag{5.9}$$

Table 5.3 shows the flat-band voltage and the corresponding shift for both branches of the three gate stacks.

Table 5.3: The Flat-Band Voltage Shifts

	V_{FB} Forward	ΔV_{FB} Forward	V_{FB} Reverse	ΔV_{FB} Reverse
30 A ox / 90 A HTO	-0.4 V	-0.15 V	-0.4 V	-0.15 V
30 A ox / 60 A NC / 90 A HTO	-0.57 V	-0.32 V	1.12 V	1.37 V
15 A ox / 60 A NC / 90 A HTO	-0.58 V	-0.33 V	-0.45 V	-0.2 V

As it can be seen, the flat-band voltage difference between the left and right branches of the 30 A capacitor is more than 2 V. This is an important observation

because a voltage shift of 2 V is large enough to build a floating-gate memory device. The memory circuits that are specialized for commercial continuous floating gate devices such as Flash memories can easily sense a flat-band voltage difference of 2 V. Therefore, the hysteresis effect in the 30 A capacitor makes it very promising as a real memory device. The lower hysteresis of the 15 A capacitors makes them less desirable for use in a memory device. The reason for the lower hysteresis is the lower retention rate which is discussed at the end of the next section.

5.3 The Nanocrystal Charge

In this section, the nanocrystal charge is estimated by using the flat-band voltage shift calculated in the previous section. The relation between the flat-band voltage shift and the charge stored in the nanocrystals can be approximated as [4]

$$\Delta V_{FB} = \frac{q\bar{v}n_{nc}}{\varepsilon_{ox}}\left(t_{cnt} + \frac{1}{2}\frac{\varepsilon_{ox}}{\varepsilon_{si}}t_{NC}\right) \tag{5.10}$$

where q is the unit charge, $\bar{v}$ is the average number of electrons (holes) per nanocrystal[1], n_{nc} is the nanocrystal sheet density, t_{cnt} is the control oxide thickness, and t_{NC} is the thickness of the nanocrystal layer. This equation is derived by assuming that the charge in the nanocrystals is an average distance of $t_{cnt} + \frac{1}{2}t_{NC}$ away from the gate. The $\frac{1}{2}$ factor is an approximation that is used instead of integrating the charge density over the entire nanocrystal. The basic assumption is that the charge centroid is at the nanocrystal center. Using this equation and a nanocrystal sheet density of 5×10^{11} cm^{-2}, each additional charge into the nanocrystals will change the flat-band voltage by 0.225 V. The nanocrystal charge density, Q'_{nc}, can be expressed as

$$Q'_{nc} = q\bar{v}n_{nc} \tag{5.11}$$

[1]The sign of $\bar{v}$ is positive for holes and negative for electrons.

Table 5.4 shows the calculated values of the nanocrystal charge density, Q'_{nc}, and the average charge stored per nanocrystal, $\bar{v}$. The calculations are based on a nanocrystal sheet density of 5×10^{11} cm^{-2}.

Table 5.4: The Nanocrystal Charge

	Q'_{nc}(nC/cm^2):Forward	$\bar{v}$:Forward	Q'_{NC}(nC/cm^2):Reverse	$\bar{v}$:Reverse
30 A ox / 60 A NC / 90 A HTO	110	1.5	473	6
15 A ox / 60 A NC / 90 A HTO	114	1.5	69	1

The results verify that the 30 A capacitors have much higher retention rates than the 15 A capacitors. The charge stored in the nanocrystals tends to leak when the voltage is about the flat-band voltage since the charge in the nanocrystals can always tunnel back into the channel. Therefore, it can be concluded that during the HF sweep, as the voltage gets closer to the flat-band voltage of the device, the charge stored in a 15 A capacitor can much more easily tunnel back into the channel, and this means that the retention time that is discussed in Chapter 1 is much smaller for 15 A capacitors. This is also in agreement with the theoretical results which predict that the tunnel rate is exponentially dependent on the oxide thickness.

Chapter 6

Conclusion

This thesis studies the memory effect in nanocrystal embedded MOS capacitors. The nanocrystal embedded MOS capacitor (NC-MOS) is the core structure of a new type of memory which is called the nanocrystal memory. The nanocrystal memory is proposed as a discrete floating gate memory device which operates using the quantum mechanical tunneling current between a discrete layer of nanocrystals and the MOS channel. A nanocrystal is a quantum dot of silicon, and it is surrounded by interface traps. Nanocrystals are embedded in the oxide of a MOS capacitor by using a chemical vapor deposition technique.

The NC-MOS capacitor exhibits memory effect which is evident in the hysteresis that is observed in the HF and QS C-V data. The hysteresis is due to the change in the flat-band voltage. The flat-band voltage changes when the charge level inside the nanocrystals changes by the effect of the direct tunneling process. The charge in the nanocrystals can be stored either at the discrete states of the quantum dot or at the interface trap states.

Two different device design parameters are used in this thesis: the tunnel oxide thickness and the forming gas anneal condition. The tunnel oxide thickness is the distance between the nanocrystal and the MOS channel, and it is a determinant factor in the operating speed and the retention rate. There is a trade-off between the operating speed and the retention rate. The tunneling rate is much higher with a thinner oxide, and a higher tunneling rate is necessary for fast operation. However,

when a thinner oxide is used, the retention rate is much lower since the charge in the nanocrystals can leak out much faster when the barrier thickness is lowered. The forming gas anneal process is used to reduce the interface trap density. The hysteresis slightly decreases with FGA which is evidence for charge storage in the interface traps.

In this thesis, three types of measurements are conducted: HF and QS capacitance-voltage and current-voltage measurements. The HF C-V measurements are used to study the memory effect. It is observed that the memory effect is dependent on the measurement conditions such as bias scheme, sweep rate, light and hold time. The results of the HF C-V measurements reveal that the nanocrystal charging is mostly due to electron tunneling in accumulation regime, and the nanocrystal discharging is due to hole tunneling in the inversion regime.

The results of the measurements are in agreement with the theoretical predictions of Chapter 2. The energy band diagrams suggested in Chapter 2 show the band bending at the silicon surface, and the direct tunneling is explained based on the surface potential condition. It is shown that the potential barrier to electron tunneling is 3.2 eV, and the potential barrier to hole tunneling is 4.7 eV. A direct effect of this is that the electron tunneling current density is much larger which is also evident in the larger flat-band voltage shift of the accumulation-charged branch of the HF C-V curve compared to the inversion-charged HF C-V curve.

In Chapter 5, a more quantitative approach is used to calculate the flat-band voltage shift and the amount of the charge stored in the nanocrystals. The results show that the flat-band voltage can be changed by more than 2 V for 30 A capacitors. As it is discussed in that chapter, this voltage shift is large enough to use this device as a floating gate memory device. On the other hand, the flat-band voltage shift for the 15 A capacitors is much smaller, and this demonstrates the trade-off between write/erase speed and the retention rate.

The future direction in the research of NC-MOS capacitors includes the improvement of the device fabrication to produce nanocrystals with more persistent sizes and trap densities. TEM measurements are necessary to characterize the density, size and the shape of the nanocrystals. Once the fabrication technique is improved, the

NC-MOS electrostatics can be investigated more rigorously by using more reliable devices that are designed with different gate stack compositions and FGA conditions. Future theoretical work includes a better characterization of these mesoscopic devices by the help of quantum, statistical and semiconductor physics.

Nanocrystal memories are one of the first examples of mesoscopic electronic devices that are built with conventional semiconductor technology. As the device dimensions are getting smaller, it is predicted that the mesoscopic limits will eventually be reached, and nanocrystal memories are certainly one of the pioneering applications of electronic devices that operate in this exciting new size regime.

Appendix A

NC-MOS Capacitor Run-Sheet for the Fabrication Facility

NCCAP-AL2

Run authorized for sign-in by DUKOVIC on 09/15/99
RUN SHEET NAME: NCCAP-AL2
SUBMISSION DATE: 09/09/99 – 14:31
CREATION DATE: 09/15/99 RUN
TYPE: CMOS
RUN PURPOSE: Nanocrystal Al Dot Capacitors
INITIATOR: Eralp Atmaca, Ken Rim, Jeff Welser
PHONE: 2661, 2946, 3250
LOCATION: 38-129
USERID OF OWNER: EATMACA
NUMBER OF OPTICAL MASK LEVELS: 0
NUMBER OF E-BEAM MASK LEVELS: 0
LITH MASK SET:

RUN WAFER SPECIFICATIONS

NUMBER OF RUN WAFERS: 15
WAFER SOURCE: FACILITY
WAFER HISTORY:
WAFER VENDOR:
LOT NUMBER:
RESISTIVITY: 0.6 - 1.5
WAFER TYPE: n
DIAMETER: 125

WAFER COMMENTS

The resistivity isn't crucial – something close to 0.1 is fine (n-type).

ROUTING SHEET

1. HOT WAFER IDENTIFICATION - RECORD WAFER LOT ID NO.'S ON RUNSHEET
2. HOT CLEAN - HUANG CLEAN - Co-ordinate for loading into UHV-CVD tool
3. LPCVD Non-standard process - THICKNESS: 3 nm
4. LPCVD Non-standard process - THICKNESS: 6 nm
5. LPCVD Non-standard process - THICKNESS: 9 nm
6. HOT CLEAN - HUANG CLEAN - Co-ordinate for loading into UHV-CVD tool
7. LPCVD Non-standard process - THICKNESS: 1 nm
8. LPCVD Non-standard process - THICKNESS: 6 nm
9. LPCVD Non-standard process - THICKNESS: 9 nm
10. HOT CLEAN - HUANG CLEAN - Co-ordinate for loading into UHV-CVD tool
11. LPCVD Non-standard process - THICKNESS: 3 nm
12. LPCVD Non-standard process - THICKNESS: 9 nm
13. HOT ANNEAL - TEMPERATURE 400 C ; TIME 30 min - FGA 350-550
14. HOT ANNEAL - TEMPERATURE 450 C ; TIME 30 min - FGA 350-550
15. HOT ANNEAL - TEMPERATURE 550 C ; TIME 30 min - FGA 550-900
16. HOT ANNEAL - TEMPERATURE 600 C ; TIME 30 min - FGA 550-900
17. METAL EVAP - EVAP AL DOTS - Thickness: 300 nm
18. METAL WET ETCH - ETCH BACKSIDE - Backside films: 9nm ox / 6nm Si-NC / 3nm ox
19. METAL EVAP - EVAP BACKSIDE AL - Thickness: 300nm

PROCESSES REQUIRING SPECIAL SIGN-OFF:

3. LPCVD Non-standard process - THICKNESS: 3 nm
4. LPCVD Non-standard process - THICKNESS: 6 nm
5. LPCVD Non-standard process - THICKNESS: 9 nm
7. LPCVD Non-standard process - THICKNESS: 1 nm
8. LPCVD Non-standard process - THICKNESS: 6 nm
9. LPCVD Non-standard process - THICKNESS: 9 nm
11. LPCVD Non-standard process - THICKNESS: 3 nm
12. LPCVD Non-standard process - THICKNESS: 9 nm

WAFER SPLITS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Hot Wafer Ident.															
2	Hot Clean						X	X	X	X	X	X	X	X	X	X
3	LPCVD						X	X	X	X	X	X	X	X	X	X
4	LPCVD						X	X	X	X	X	X	X	X	X	X
5	LPCVD						X	X	X	X	X	X	X	X	X	X
6	Hot Clean	X	X	X	X	X						X	X	X	X	X
7	LPCVD	X	X	X	X	X						X	X	X	X	X
8	LPCVD	X	X	X	X	X						X	X	X	X	X
9	LPCVD	X	X	X	X	X						X	X	X	X	X
10	Hot Clean	X	X	X	X	X	X	X	X	X	X					
11	LPCVD	X	X	X	X	X	X	X	X	X	X					
12	LPCVD	X	X	X	X	X	X	X	X	X	X					
13	Hot Anneal	X		X	X	X	X		X	X	X	X		X	X	X
14	Hot Anneal	X	X		X	X	X	X		X	X	X	X		X	X
15	Hot Anneal	X	X	X		X	X	X	X		X	X	X	X		X
16	Hot Anneal	X	X	X	X		X	X	X	X		X	X	X	X	
17	Metal Evap															
18	Metal Wet Etch															
19	Metal Evap															

---

1. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

HOT WAFER IDENTIFICATION - RECORD WAFER LOT ID NO.'S ON RUNSHEET
TOOL: BAR CODE READER recif
SIGNOFF AND DATE: TRACKED BY USERID MEYER ON SEPTEMBER 20, 1999

---

2. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
					X	X	X	X	X	X	X	X	X	X

HOT CLEAN - HUANG CLEAN
TOOL: HUANG CLEAN HOOD
COMMENTS: Co-ordinate with Kevin for loading into UHV-CVD cluster tool.
MEASUREMENTS SPECIFICATIONS: The run contains 15 product wafers.

MEASUREMENT	TARGET	MIN	MAX	HOLD
PARTICLES BEFORE	50	0	100	N
PARTICLES AFTER	0	0	100	Y

SIGNOFF AND DATE: TRACKED BY USERID MEYER ON SEPTEMBER 21, 1999

---

3. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
					X	X	X	X	X	X	X	X	X	X

LPCVD Non-standard process - Non-stnd process
THICKNESS: 3 nm
PROCESS: Thermal Oxide
TOOL: Furnace FLP-1 (Poly)
COMMENTS: Gate tunnel oxide growth in UHV-CVD cluster tool.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 23, 1999

---

4. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
					X	X	X	X	X	X	X	X	X	X

LPCVD Non-standard process - Non-stnd process
THICKNESS: 6 nm
PROCESS: Nanocrystals
TOOL: Furnace FLP-1 (Poly)
COMMENTS: Please deposit your best nanocrystals, and record conditions used here.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 23, 1999

---

5. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
					X	X	X	X	X	X	X	X	X	X

LPCVD Non-standard process - Non-stnd process
THICKNESS: 9 nm
PROCESS: HTO
TOOL: Furnace FLP-1 (Poly)
COMMENTS: Please deposit 9nm of high-quality, high temperature oxide (HTO).
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 23, 1999

---

6. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X						X	X	X	X	X

HOT CLEAN - HUANG CLEAN
TOOL: HUANG CLEAN HOOD
COMMENTS: Co-ordinate with Kevin for loading into UHV-CVD cluster tool.

MEASUREMENT	TARGET	MIN	MAX	HOLD
PARTICLES BEFORE	50	0	100	N
PARTICLES AFTER	0	0	100	Y

SIGNOFF AND DATE: TRACKED BY USERID MEYER ON SEPTEMBER 23, 1999

---

7. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X						X	X	X	X	X

LPCVD Non-standard process - Non-stnd process
THICKNESS: 1 nm
PROCESS: Thermal Oxide
TOOL: Furnace FLP-1 (Poly)
COMMENTS: Gate tunnel oxide growth in UHV-CVD cluster tool. Please grow 1.5 nm.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 24, 1999

---

8. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X						X	X	X	X	X

LPCVD Non-standard process - Non-stnd process
THICKNESS: 6 nm
PROCESS: Nanocrystals
TOOL: Furnace FLP-1 (Poly)
COMMENTS: Please deposit your best nanocrystals, and record conditions used here.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 24, 1999

---

9. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X						X	X	X	X	X

LPCVD Non-standard process - Non-stnd process
THICKNESS: 9 nm
PROCESS: HTO
TOOL: Furnace FLP-2 (Nitr)
COMMENTS: Please deposit 9nm of high-quality, high temperature oxide (HTO).
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 24, 1999

10. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X	X	X	X	X	X					

HOT CLEAN - HUANG CLEAN
TOOL: HUANG CLEAN HOOD
COMMENTS: Co-ordinate with Kevin for loading into UHV -CVD cluster tool.

MEASUREMENT	TARGET	MIN	MAX	HOLD
PARTICLES BEFORE	50	0	100	N
PARTICLES AFTER	0	0	100	Y

SIGNOFF AND DATE: TRACKED BY USERID DEMIC ON SEPTEMBER 24, 1999

---

11. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X	X	X	X	X	X					

LPCVD Non-standard process - Non-stnd process
THICKNESS: 3 nm
PROCESS: Thermal oxide
TOOL: Furnace FLP-1 (Poly)
COMMENTS: Gate tunnel oxide growth in UHV-CVD cluster tool.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 27, 1999

---

12. LPCVD SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X	X	X	X	X	X					

LPCVD Non-standard process - Non-stnd process
THICKNESS: 9 nm
PROCESS: HTO
TOOL: Furnace FLP-2 (Nitr)
COMMENTS: Please deposit 9nm of high-quality, high temperature oxide (HTO).
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 28, 1999

---

13. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X		X	X	X	X		X	X	X	X		X	X	X

HOT ANNEAL - FORMING GAS ANNEAL 350-500C
TEMPERATURE: 400 C
TIME: 30 min
TOOL: FURNACE B-3 SILCD ANN
COMMENTS: Pre-metal forming gas anneal.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 28, 1999

---

14. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X		X	X	X	X		X	X	X	X		X	X

HOT ANNEAL - FORMING GAS ANNEAL 350-500C
TEMPERATURE: 450 C
TIME: 30 min
TOOL: FURNACE A-3 ANNEAL
COMMENTS: Pre-metal forming gas anneal.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 29, 1999

---

15. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X		X	X	X	X		X	X	X	X		X

HOT ANNEAL - FORMING GAS ANNEAL 550-900C
TEMPERATURE: 550 C
TIME: 30 min
TOOL: FURNACE A-3 ANNEAL
COMMENTS: Pre-metal forming gas anneal.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 29, 1999

---

16. HOT SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X		X	X	X	X		X	X	X	X	

HOT ANNEAL - FORMING GAS ANNEAL 550-900C
TEMPERATURE: 600 C
TIME: 30 min
TOOL: FURNACE A-3 ANNEAL
COMMENTS: Pre-metal forming gas anneal.
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 29, 1999

---

17. F-METAL SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

METAL EVAP - EVAP AL DOTS
THICKNESS: 300 NM
TOOL: T-Evaporator (5)
COMMENTS: Thickness: 300 nm
SIGNOFF AND DATE: TRACKED BY USERID JWSJR ON SEPTEMBER 29, 1999

---

18. F-METAL SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

METAL WET ETCH - ETCH BACKSIDE
TOOL: Ultrafab Metal Hood
COMMENTS: Backside films: 9nm ox / 6nm Si-NC / 3nm ox
SIGNOFF AND DATE: TRACKED BY USERID JWSJR ON SEPTEMBER 29, 1999

---

19. F-METAL SECTOR

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

METAL EVAP - EVAP BACKSIDE AL
THICKNESS: 300 NM
TOOL: T-Evaporator (5)
COMMENTS: Thickness: 300nm
SIGNOFF AND DATE: TRACKED BY USERID KCHAN ON SEPTEMBER 30, 1999

Bibliography

[1] Paolo Pavan, Roberto Bez, Piero Olivo, and Enrico Zanoni. Flash memory cells–an overview. In *Proceedings of the IEEE*, number 8, pages 1248–1271, August 1997.

[2] Ashok K. Sharma. *Semiconductor Memories.* IEEE Press, New Jersey, 1997.

[3] C. T. Sah. Evolution of the mos transistor–from conception to vlsi. In *Proceedings IEEE*, number 76, page 1280, October 1988.

[4] S. Tiwari, F. Rana, K. Chan, L. Shi, and H. Hanafi. Single charge and confinement effects in nano-crystal memories. *Appl. Phys. Lett.*, 69(9):1232–1234, August 1996.

[5] S. Tiwari, F. Rana, H. Hanafi, and K. Chan A. Hartstein, E. F. Crabb. A silicon nanocrystal based memory. *Appl. Phys. Lett.*, 68(10):1377–1379, March 1996.

[6] S. Tiwari, F. Rana, K. Chan, H. Hanafi, W. Chan, and D. Buchanan. Volatile and non-volatile memories in silicon with nano-crystal storage. *IEDM*, 95:521–523, 1995.

[7] Kazuo Yano, Tomoyuki Ishii, Takashi Hashimoto, Takashi Kobayashi, Fumio Murai, and Koichi Seki. Room-temperature single-electron memory. *IEEE Transactions on Electron Devices*, 41:1628–1638, 1994.

[8] E. H. Nicollian and J. R. Brews. *MOS (Metal Oxide Semiconductor) Physics and Technology.* John Wiley & Sons, New York, 1982.

[9] J. A. Wahl, H. Silva, A. Gokirmak, A. Kumar, J. J. Welser, and S. Tiwari. Write, erase and storage times in nanocrystal memories and the role of interface states. 1999.

[10] Y. Shi, K. Saito, H. Ishikuro, and T. Hiramoto. Effects of traps on charge storage characteristics in metal-oxide-semiconductor memory structures based on silicon nanocrystals. *American Institute of Physics*, 84(4):2358–2360, August 1998.

[11] Sandip Tiwari, J. J. Welser, Arvind Kumar, and F. Rana. Statistical effects in memories using silicon quantum dots: Transport and interface state issues. *Silicon Nanoelectronics Workshop, Kyoto, Japan*, June 1999.

[12] A. Kohno, H. Murakami, M. Ikeda, H. Nishiyami, S. Miyazaki, and M. Hirose. Electron charging to a si quantum dot floating gate in mos structures. *Workshop on Silicon Nanofabrication and Nanodevices, Tokyo, Japan*, September 1999.

[13] Elizabeth Boer and Harry Atwater. Atomic force microscopy of silicon nanocrystals: Characterization and charging. *Research Report, Thomas J. Watson Laboratory of Applied Physics, California Institute of Technology, Pasadena, Ca.*

[14] F. Rana, S. Tiwari, and D. A. Buchanan. Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides. *Appl. Phys. Lett.*, 69(8):1104–1106, August 1996.

[15] Y. Taur, D. S. Zicherman, D. R. Lombardi, P. J. Rstle, C. H. Hsu, H. I. Hanafi, M. R. Wordeman, B. Davari, and G. G. Shadidi. A new shift and ratio method for mosfet channel-length extraction. *IEEE Electron Device Letters*, 13(5):267–269, May 1992.

[16] D. J. DiMaria and D. W. Dong. High current injection into sio_2 from si rich sio_2 films and experimental applications. *J. Appl. Phys.*, 51(5):2722–2734, May 1980.

[17] D. J. DiMaria, K. M. DeMeyer, C. M. Serrano, and D. W. Dong. Electrically-alterable read-only-memory using si-rich sio_2 injectors and a floating polycrystalline silicon storage layer. *J. Appl. Phys.*, 52(7):4825–4839, July 1981.

[18] D. J. DiMaria, E. Cartier, and D. Arnold. Impact ionization, trap creation, degradation, and breakdown in silicon dioxidefilms on silicon. *J. Appl. Phys.*, 73(7):3367–3383, April 1993.

[19] D. J. DiMaria and J. W. Stasiak. Trap creation in silicon dioxide produced by hot electrons. *J. Appl. Phys.*, 65(6):2342–2355, March 1989.

[20] I. Kim, S. Han, H. Kim, J. Lee, B. Choi, S. Hwang, D. Ahn, and H. Shin. Room temperature single electron effects in si quantum dot memory with oxide-nitride tunneling dielectrics. *IEEE*, pages 111–114, 1998.

[21] H. Sunamura, T. Sakamoto, Y. Nakamura, H. Kawaura, J. S. Tsai, and T. Baba. Single electron memory with carrier trapping sin layer. *Appl. Phys. Lett.*, (71):2038, 1997.

[22] F. Sabri and D. G. Hasko. Charge storage in gaas metal insulator semiconductor field effect transistor metal nanodot memory structures. *Applied Physics Letters*, 74(20):2996–2998, May 1999.

[23] A. Dutta, S. P. Lee, Y. Hayafune, S. Hatatani, S. Oda, and M. Willander. Observation of quantum effects and coulomb blockade in nanodevices with plasma derived nanocrystals. *Silicon Nanoelectronics Workshop, Kyoto, Japan*, June 1999.

[24] Y. Fu, M. Willander, F. Yun, and S. Oda. Electronic band structures in si nanocrystals with applications to dot-based mosfets. *Silicon Nanoelectronics Workshop, Kyoto, Japan*, June 1999.

[25] Yasuhiko Ishikawa, Masaaki Kosugi, Minoru Kumezawa, and Michiharu Tabe. Capacitance-voltage study of single-crystalline si layer and si dots on ultrathin buried sio_2. *Silicon Nanoelectronics Workshop, Kyoto, Japan*, June 1999.

[26] Kwangseok Han, Ilgweon Kim, and Hyungcheol Shin. Pmos-based si nano-crystal memory. *Silicon Nanoelectronics Workshop, Kyoto, Japan*, June 1999.

[27] K. Uchida, J. Koga, R. Ohba, S. Takagi, and A. Toriumi. Planar non-volatile memory with single-electron channel fabricated on a hyper-thin soi film. *Silicon Nanoelectronics Workshop, Kyoto, Japan*, June 1999.

[28] Robert F. Pierret. *Semiconductor Device Fundamentals.* Addison-Wesley, Reading, Massachusetts, March 1996.

[29] Richard S. Muller and Theodore I. Kamins. *Device Electronics for Integrated Circuits.* John Wiley & Sons, New York, 1977.

[30] Sheng S. Li. *Semiconductor Physical Electronics.* Microdevices–Physics and Fabrication Technologies. Plenum Press, New York, 1993.

[31] Franklin F. Y. Wang. *Introduction to Solid State Electronics.* North-Holland, New York, 1988.

[32] Ben G. Streetman. *Solid State Electronic Devices.* Prentice-Hall, New Jersey, 1980.

[33] D. J. DiMaria. Defects and impurities in thermal sio$_2$. In *The Physics of SiO_2 and its interfaces*, number 4, pages 160–176, March 1978.

[34] Marshall Wilson, Jack Lagowski, Alexandre Savtchouk, Lubek Jastrzebski, and John D'Amico. Cocos (corona oxide characterization of semiconductor) metrology: Physical pricipals and applications. In *ASTM Conference on Gate Dielectric Integrity*, January 1999.

2749-28