Study of leakage and degradation in metal nanocrystal-embedded gate stacks

Lwin, Zin Zar

2012

Lwin, Z. Z. (2012). Study of leakage and degradation in metal nanocrystal-embedded gate stacks. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/53456>

<https://doi.org/10.32657/10356/53456>

**STUDY OF LEAKAGE AND DEGRADATION IN
METAL NANOCRYSTAL-EMBEDDED GATE STACKS**

ZIN ZAR LWIN

School of Electrical and Electronic Engineering

A thesis submitted to Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2012

To my beloved parents & teachers

စုစုပေါင်း ၃၀ ရက် ၂၀၁၂ တွင် ဘဝတမ်းသို့ ပြောင်းသွားခဲ့သော ကျေးဇူးရှင်ဖေဖေ၊

ကျေးဇူးရှင်မေမေ နှင့်

သင်ဆရာ မြင်ဆရာ ကြားဆရာ တို့အား

ဤကျမ်းဖြင့် ကန်တော့ပါ၏။

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my research supervisors, Prof. Pey Kin Leong and Assoc. Prof. Zhang Qing from Nanyang Technological University, Singapore for giving me the opportunity to pursue higher degree in the field of nanoscience and technology.

I am thankful to Prof. Pey Kin Leong for his valuable guidance and encouragement since day-one of my post-graduate studies. I consider myself very fortunate to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. His valuable advice and relentless technical guidance with his immense knowledge benefited me at many stages of this research project.

I am grateful to Prof. Zhang Qing who is always kind and helpful. His wide knowledge in device physics and technical discussions has given me fruitful insights during the course of this thesis work. Also, I am indebted to him for providing kind support and precious opportunity to work on nanoscale characterization using Kelvin Force Microscopy.

I would like to thank our collaborators, Prof. Souvik Mahapatra and his research team from the India Institute of Technology Bombay for providing the high quality metal nanocrystal wafers for our study. I am thankful to Dr. Navab Singh from the Institute of Microelectronics, Singapore and Dr. Lap Chan, Dr. Ng Chee Mang and Mr. Leong Kam Chew from the Globalfoundries Singapore for their sharing of knowledge and technical discussions. Also, I'd like to thank Dr. Michel Bosman from the Institute of Materials

Research and Engineering, Singapore and Dr. Liu Qing, Prof. Gan Chee Lip from School of Materials Science and Engineering, Nanyang Technological University, Singapore for their sincere guidance and support on nanoscale characterization techniques using TEM/EELS and FIB.

I am also grateful to the technicians in N²FC Clean Room I, Clean room II and Semiconductor Characterization lab II for their great help and useful supports. I also appreciate Ms. Yeo Hwee Bin from Nanyang Technological University and Ms. Geraldine Teo from Singapore University of Technology and Design for the efficient administrative support. I would like to acknowledge Singapore EDB-Globalfoundries and Singapore University of Technology and Design research grant for their financial support during my Ph.D study.

I feel pretty lucky that I always have nice friends around me. Special thanks go to Dr. Li Xiang, Mr. Chen Yining, Mr. Nagarajan Raghavan, Mr. Liu Wen Hu, Mr. Shubhakar, Mr. Qin Hai Lang, and Ms. Wu Xing for their support and sharing of knowledge with me.

I am especially grateful to my husband, Dr. Li Hong, for coming into my life and bringing the joy along, and supporting to pursue my ambitions. Most of all, I would like to express my deepest gratitude to my parents and family, especially to my late beloved father Saw Lwin who used to encourage me to be strong and persistent with enthusiasm whenever I feel down. Without their continuous love, guidance, sacrifice, support, blessings and encouragement, I would not be where I am today.

Zin Zar Lwin

TABLE OF CONTENTS

ACKNOWLEDGEMENT	i
SUMMARY	vii
LIST OF FIGURES	xi
LIST OF TABLES	xx
LIST OF ABBREVIATIONS	xxi
CHAPTER ONE Introduction.....	1
1.1 Background	1
1.2 Motivation	3
1.3 Objectives	5
1.4 Organization of Thesis	6
1.5 Contributions	7
CHAPTER TWO Literature Review	14
2.1 Introduction	14
2.2 Overview of NC-Based Gate Stacks Reliability.....	14
2.2.1 Size Scaling in NCs	16
2.2.2. NC Spacing Effect	21
2.3 Overview of Gate Dielectric Reliability	24
2.3.1 Gate Dielectric Breakdown	25
2.3.1.1 SiO ₂ Gate Dielectric	25
2.3.1.2 High-κ Gate Dielectric	29

2.4 Summary	35
CHAPTER THREE Experimental Procedure and Details	41
3.1 Introduction	41
3.2 Sample Information	42
3.3. Electrical Characterization	44
3.3.1 Dielectric Stressing Methodologies	46
3.3.2 Relaxation Current Measurement	50
3.3.3 Capacitance-Voltage Measurement	53
3.4 Nanoscale Characterization using Scanning Probe Microscopy	55
3.4.1 Atomic Force Microscopy	57
3.4.2 Electrostatic Force Microscopy	59
3.4.2 Kelvin Force Microscopy	64
3.5 Summary	67
CHAPTER FOUR Metal Nanocrystal in Degraded Dielectric Gate Stacks	72
4.1 Introduction	72
4.2 Experiments	73
4.3 Charging and Discharging Characteristics of MNCs in Degraded Dielectric Gate Stacks	74
4.3.1 Methodology and Approach	74
4.3.2 One-Layer Breakdown & I_g Blockade	78
4.3.3 Proposed Model	83
4.3.4 Discharging of NCs in Degraded Dielectric	87
4.4 Dielectric Breakdown Layer Detection	93

4.4.1 Methodology and Approach.....	93
4.4.2 Results and Discussions	95
4.5 Summary	99
CHAPTER FIVE Localized Charge Transport and Lateral Charge Diffusion in Single-Layer Metal Nanocrystal Embedded Gate Stacks	104
5.1 Introduction	104
5.2 Experiments.....	105
5.3 Kelvin Force Microscopy Characterization	107
5.3.1 Tip-to-Sample Distance Dependent KFM Resolution.....	111
5.4 Localized Charge Trapping and Charge Decay in Single-Layer MNCs	114
5.4.1 Charge Decay Fitting Laws and Characteristic Time Constant	119
5.4.2 Inter-Dot Tunneling Probability	122
5.4.3 Possible Charge Decay Paths in Single-Layer MNCs	123
5.5 Discharging Current at Elevated Temperatures	124
5.6 Charge Transport Mechanisms in Single-Layer MNCs.....	127
5.7 Summary	130
CHAPTER SIX Localized Charge Transport in Dual-Layer Metal Nanocrystal Embedded Gate Stacks	135
6.1 Introduction	135
6.2 Experiments.....	136
6.3 Vertical Charge Loss Mechanisms	139
6.3.1 Temperature Accelerated Retention	139
6.3.2 Gate Bias Accelerated Retention.....	141

6.3.3 Charge Distribution in Dual-Layer MNCs.....	144
6.3.4 Vertical Charge Loss Mechanisms in Dual-Layer MNCs	150
6.4 Charge Transport Mechanisms in Dual-Layer MNCs	153
6.4.1 KFM Analysis of Charge Decaying Characteristics.....	153
6.4.2 Charge Transport in Dual-Layer MNCs	160
6.5 Summary	164
 CHAPTER SEVEN Conclusion and Recommendations	 169
7.1 Conclusion.....	169
7.2 Recommendations.....	172
 PUBLICATIONS.....	 174

SUMMARY

Nanocrystal (NC) memories have attracted a lot of research attentions as a promising candidate to reduce defect-related charge loss and to overcome scaling limitation in conventional floating gate memories. Most work on metal nanocrystal (MNC) has focused primarily on the materials and fabrication issues, while the reliability aspects have been greatly ignored, such as inter-dot tunneling among NCs, the charge loss mechanisms and their performances in the degraded dielectric gate stacks which are of utmost importance for the device optimization and future scaling. This work focuses on the degradation and reliability of MNC embedded gate stacks, contributes to a deeper understanding of MNC performances in the degraded dielectric gate stacks and its charge transport behaviors, and reports extensively on the reliability characterization of both single-layer and dual-layer MNC embedded gate stacks.

To enable strong gate coupling and low operating voltage, NCs are conventionally embedded in the high- κ layer. Therefore, the overall dielectric gate stack in conventional NC memories typically consists of a tunneling oxide layer and a NC-embedded high- κ layer. It is thus important to understand the performance of MNCs in the case of the individual dielectric layer degradation. Firstly, we report the influence of dielectric degradation on charging and discharging characteristics of MNCs embedded in high- κ /SiO₂ gate stacks. It is found that the charging and discharging phenomena of the MNCs and leakage mechanism in the degraded gate stacks are strongly dependent on the lateral charge tunneling/hopping among the NCs. Our experimental results show that the localized breakdown not only affects charge holding capability of the affected NCs, but

also provides a leakage path for the charges stored in the surrounding NCs. It indicates the existence of lateral charge diffusion in the MNCs.

Next, the post-breakdown characteristics of the conduction path formed in the individual different dielectric layer (either SiO_2 or NC-embedded Al_2O_3) are identified. The first layer to breakdown is determined based on the physics underlying the Coulomb charging energy in relation to thermal energy gained by electrons at low voltage and in the very low temperature regime ranging from 11K to 70K. With this approach, the average trap (defect) size in NC-embedded high- κ and SiO_2 is further analyzed. It is noted that breakdown in SiO_2 leads to lateral charging/discharging among NCs while in Al_2O_3 , it leads to spontaneous breakdown of bi-layer gate stacks owing to high localized trap generation rate around the high- κ dielectric grain boundary and local electric field enhancement in the vicinity of MNCs.

To further confirm the lateral charge diffusion of MNCs observed in the degraded dielectric gate stacks, the localized charge transport and lateral charge diffusion phenomena in MNCs are investigated by Kelvin Force Microscopy (KFM) characterization. The results reveal that vertical charge loss and lateral charge diffusion are two competing mechanisms and they can be identified by discharging current measurements at elevated temperatures and KFM characterization. It is found that the MNC with higher work function has a lower inter-dot charge tunneling probability, which is favorable for improved retention in memory applications. However, the vertical charge loss during the initial decay period is a trade-off and hence it could be minimized by using dual-layer (DL) MNC structure wherein electron de-trapped from one-NC layer could be “re-trapped” in the adjacent NC layer.

Therefore, the DL-MNC devices with high work function MNCs are chosen to study their intrinsic charge loss mechanisms and inter-dot tunneling and inter-layer tunneling characteristics. A comparative study is performed on the DL devices with inter-layer dielectric (ILD) thickness variation. The temperature and gate-bias accelerated retention measurements and KFM approaches are used. It is found that charge loss in DL structures is mainly due to the internal electric field induced by trapped electrons which depends on the ILD thickness and MNC spacing. When the ILD thickness is comparable to MNCs spacing, the charge distribution in two-MNC layers are similar and the internal electric field induced by charges stored in the MNCs is higher. It leads to the fact that internal-electric-field assisted tunneling dominates at lower temperature. In contrast, when the ILD thickness is larger than the average MNC spacing, less charges are injected in the neighboring MNC layer, resulting in a reduction of both internal electric field and oxide trap alignment possibility with the traps inside the MNCs. Moreover, a thick ILD reduces the electron tunneling probability from one MNC-layer to another during retention. Our findings suggest that an optimized DL-MNCs embedded memory cell could be achieved by defining the ILD thickness larger than the average MNC-spacing for enhancement of retention ability in MNC embedded gate stacks. It implies the possibility of reducing MNC spacing in scaled memory devices by using DL-MNCs with controlled ILD thickness to achieve both increased memory window and prolonged retention.

In summary, the dielectric degradation and fundamental reliability issues in MNC embedded gate stacks are investigated in details and comprehensive understanding of

inter-dot tunneling and charge loss mechanisms in MNC embedded gate stack is achieved for optimum memory cell design with retention reliability perspective.

LIST OF FIGURES

Figure	Caption	Page
2.1	Figure 2.1. Band diagram illustration for MNC memory in (a) Erase and (b) Program processes [2]. In this illustration, erase operation is performed with a positive control gate voltage and write operation is carried out by applying a negative control gate voltage.	15
2.2	Density of states per volume and energy for a 3-D semiconductor (blue), 2-D quantum well (red), 1-D quantum wire (green) and 0-D quantum dot (black) [3].	17
2.3	(a) Single-electron charging energy of Au-NCs as a function of NC diameter and (b) flatband voltage shift in NCs with various diameters [5].	18
2.4	(a) Conduction band minimum up-shift of the semiconductor NCs and Fermi level up-shift of the metal NC and (b) retention time for various metal and semiconductor NC memories as a function of NC size [6].	20
2.5	Effect of NC spacing on (a) single-electron charging energy and (b) flatband voltage shift as a function of programming time in Au-NCs with diameters of 5 nm [5].	22
2.6	Program fail probability of various cell areas ($W \times L$) as a function of NC density [10].	22
2.7	Schematic of a percolation path formed in the gate dielectric due to the accumulation of defects or electronic traps [26].	27
2.8	The oxygen deficiency A-A' line profile in the percolation path of the dielectric breakdown samples with various breakdown hardness. The inserted diagram shows the A-A' cross section of a percolation in a breakdown oxide [30].	29
2.9	Comparison of relaxation current decay in various high- κ dielectric films and SiO_2 [38].	31
2.10	Illustration of the atomic re-arrangements as a particle in a double potential well model. (a) A particle in a double-well potential $V(x)$, (b) The particle “fall” downstream by classical mechanics in the presence of a strong electric field ξ eliminating the barrier. Also, the particle can still	33

	tunnel to adjacent well by (c) quantum-mechanical tunneling if the electric field is not strong enough to remove the potential barrier and (d) by thermally hopping over the barrier [41].	
2.11	Relaxation current of Ru-NC and ITO-NC embedded ZrHfO films and corresponding control sample as reported in [48].	34
3.1	Schematic process-flow of MNC-embedded gate stacks used in this study.	43
3.2	Schematic cross-sections (not-to-scale) of (a) control high- κ , and (b) single- and (c) dual-layer MNC-embedded gate stacks.	44
3.3	(a) Suss MicroTec PM8 200 mm wafer probe station and (b) Keithley 4200 semiconductor parameter analyzer.	45
3.4	Typical example of I_g - V_g ramping test with a voltage increment of 0.05V in determining V_{BD} of Au-NC-embedded gate stack.	47
3.5	The relationship of a voltage acceleration factor λ versus V_g reported in Ref [6].	47
3.6	Evolution of leakage current using a constant voltage stress ($V_{g-stress} = 5.5V$) under substrate injection in the Au-NC embedded device.	49
3.7	Evolution of leakage current using a successive CVS ($V_{g-stress}=5.5V$) with the compliance current ranging from 100 nA to 100 μA in Au-NC embedded device.	50
3.8	(a) Semi-log and (b) log-log plots of I_{relax} at zero gate voltage after the removal of a 5 sec, +3MV/cm stress in Au-MNC embedded sample and control sample.	52
3.9	A sample CV curve showing Program (P) and Erase (E) characteristics of DL-MNCs devices. The P/E transients were obtained from the shift in the flatband voltage (V_{FB}) of the CV curves.	54
3.10	(a) Setup of Cypher Scanning Probe Microscopy with a controller and computer. (b) Front-view of the Cypher Microscopy with the enclosure door opens.	56
3.11	Schematic illustration of basic AFM setup from Ref. [29].	57
3.12	A Pt-coated Si probe used in this work [31].	58

3.13	Illustration of EFM images at various polarities of detection voltage V_{EFM} after charge injection (negative charge injection in this case). (a) The first EFM image taken with $V_{\text{EFM}} = -1\text{V}$ at $t=136$ sec upon charge injection. The darker spot appeared at the injected charge cloud area due to repulsive force gradient. (b) The subsequent EFM image taken with $V_{\text{EFM}} = +1\text{V}$ at $t=227$ sec. The brighter spot appeared at the injected charge cloud area due to attractive force gradient. (c) The related profile of phase shift across the injected charge cloud.	62
3.14	Appearance of halo image at injected charge cloud area when $V_{\text{EFM}} = 0\text{V}$. The image is taken at $t=320$ sec after a negative charge injection.	63
3.15	(a) Typical plan view KFM image at the lift height of 50 nm upon negative tip injection on the NC embedded device. (b) The potential profile across the injected charge cloud (A-A') in (a).	67
4.1	(a) Schematic of the MNC-embedded bi-layer gate stack used in our study. (b) Plan view TEM micrograph of Au-NCs. (c) Cross-sectional view TEM micrograph showing existence of MNCs in the Al_2O_3 layer.	74
4.2	I_g - V_g ramping tests in the Au-NCs embedded samples to determine dielectric breakdown voltage (V_{BD}).	75
4.3	(a) Evolution of I_g using CVS, $V_{g\text{-stress}} = 5.5\text{V}$, in substrate injection mode with I_{g1} ranging from 100nA to 100mA. (b) I_g - V_g at different stages of progressive breakdown.	77
4.4	(a) Evolution of I_g using a CVS with $V_{g\text{-stress}} = 4.5\text{V}$ in substrate injection mode. (b) I_g - V_g before stress and after a 2-layer breakdown, respectively.	78
4.5	A new 3-stage I_g evolution under CVS in an Au-NC sample denoted as Sample 1. (a) I_g under 5V CVS (substrate injection) with I_{g1} ranging from 50 to 500 nA. Stage 1 - I_g fluctuations, Stage 2 - rapid jump in I_g owing to a one-layer breakdown, and Stage 3 - I_g blockade are seen. (b) I_g - V_g trends at different stages of progressive breakdown and after I_g blockade, respectively. The arrows delineate I_g evolution during progressive breakdown and after I_g blockade (<i>i.e.</i> , large decrement of I_g).	79

4.6	3-stage I_g evolution under CVS in Au NC-embedded samples. (a) Sample 2 with $I_{g-max} = 200$ nA under CVS of 4.8V. (b) Sample 3 with $I_{g-max} = 500$ nA under $V_{g-stress}=5.5$ V. The insets show the same data in a logarithmic ordinate scale and the horizontal dotted lines are guides to eyes.	82
4.7	3-stage I_g evolution in Ir-NC embedded gate stacks under CVS = 2.6V. The inset shows the same data in a logarithmic ordinate scale.	83
4.8	Schematics of (a) defect generated in dielectric film during CVS, (b) lateral charging through a percolation path, (c) leakage components before occurrence of breakdown, $I_g=I_1+I_2$ where I_1 is the direct/FN/trap assisted tunneling (TAT) current, and I_2 is the resonant tunneling current through NCs, and (d) after one-layer breakdown and lateral charging happened, I_1 decreases due to the increased electrostatic potential between the charged NCs. I_2 is eliminated by the Coulomb blockade effect. In this illustration, we assume that the breakdown path presents in the SiO ₂ layer.	84
4.9	Simulation of localized electric-field enhancement at the breakdown path and SiO ₂ /high- κ interface. For simplicity, we ignored the presence of NCs.	85
4.10	Schematic energy band diagrams of (a) a lateral charging process between a charged NC (NC1) and an uncharged neighboring NC (NC2) and (b) a discharging process through the breakdown path. The breakdown path is assumed to be in the SiO ₂ layer.	86
4.11	I_{relax} at different V_{on} (from 2 to 5V with step of 0.5V) of an unstressed fresh device.	87
4.12	I_{relax} before stress and after I_g blockade of two different breakdown hardness samples: (a) Sample-2 with breakdown hardness of 200 nA under $V_{g-stress}=4.8$ V and (b) Sample-3 with breakdown hardness of 500 nA under $V_{g-stress}=5.5$ V.	88
4.13	Estimation of discharged amount from I_{relax} before stress and after I_g blockade.	89
4.14	Schematic energy band diagram of 3 nm Au-NC in between Al ₂ O ₃ and SiO ₂ dielectric.	91

4.15	Radial distribution of a percolation path. The diameter of the breakdown path in SiO ₂ /SiON gate dielectrics is estimated to be 30 nm – 50 nm at the early stage of a breakdown [7]. All dimensions are shown to scale.	92
4.16	I_g evolution under CVS in substrate injection mode with $V_{g-stress}=5V$. (a) Spontaneous I_g evolution pattern denoted as Type-A and (b) a new 3-stage I_g evolution pattern with leakage drop denoted as Type-B are shown.	94
4.17	Arrhenius plot of the gate leakage current in the low voltage regime ($V_g=1V$) for fresh sample and Type-A, Type-B mode one layer breakdown samples. The gate leakage current from 11 to ~70K are linear in fresh and Type-B mode samples while temperature dependence behavior is more drastic in Type-B mode samples. The inset presents a plot of activation energy for transport regime above 150K in Type-A mode sample.	96
4.18	Weibull probability plot of the time to failure for Al ₂ O ₃ and SiO ₂ layer breakdown in Au-NCs embedded dual layer gate stack.	98
5.1	Schematic (not to-scale) of MNC-embedded gate stack used in (a) KFM characterization and (b) transient discharging current measurement.	105
5.2	Cross-sectional view TEM micrographs of (a) Au-NC sample and (b) Pt-NC sample. The average size (diameter) of the NCs is ~ 3nm. (c) Plan view TEM micrograph showing the distribution of MNCs.	106
5.3	Illustration of KFM methodology: (a) charge injection on the NC-embedded dielectric surface and (b) surface potential measurement in the lift mode.	108
5.4	Plan view and 3D view KFM images with associated potential profile just after (a) negative tip injection and (b) positive tip injection. The negative (positive) tip injection results in negative (positive) trapped charge which appeared as dark (bright) spot in plan view KFM images.	109
5.5	The time dependence of charge decay in Au-NC embedded samples acquired at different t_s , the time interval between the completion of charge injection and arriving first potential measurement at the center of the localized charge spot.	110

5.6	(a) The plan view KFM image at lift height of 5 nm after charging the sample. The single line scanning along A-A' was performed to investigate the impact of lift scans height on KFM surface potential. (b) Plan view and (c) 3D view KFM images at various lift height: increasing from 5 nm to 300 nm then decreasing to 5 nm as indicated with arrow.	113
5.7	(a) Plan-view and (b) 3D-view KFM images and (c) AFM topography images at t=1, 6 and 15 mins after a negative tip injection into an Au-NC embedded sample.	115
5.8	(a) Evolution of potential profile at t=10 mins and t=83 mins after charge injection in Pt- and Au- samples, normalized to peak at t=10mins. (b) Schematic energy band diagram of the two adjacent NCs during retention.	116
5.9	Time dependence of normalized potential to the peak at t=0 for the Au (the closed symbols) and Pt samples (the open symbols) after charge injection of -4V, -6V,-8V (blue, red and black symbols) for 10sec. The time dependence of charge decay in the control sample (<i>i.e.</i> , without NCs) is shown for comparison.	118
5.10	Charge decay fitting with exponential laws in Pt-MNCs-embedded gate stacks. The time constant τ for different fitting laws are shown. Only the stretch exponential function can extrapolate close to actual decay data point which is measured at t=30000 sec.	119
5.11	(a) Stretched exponential fitting to the charge decay of Pt and Au samples after charge injection of -4V,-6V,-8V (blue, red and black symbols) for 10sec. (b) Characteristic time constant τ and (c) normalized FWHM as a function of time in Au and Pt samples at various injection voltages. (d) Zoom-in results of the FWHM evolution during the initial stage of retention in the Au and Pt samples.	121
5.12	Illustration of the possible charge decay paths during retention in SL-MNC embedded gate stacks: P_x corresponds to inter-dot tunneling and P_z refers to vertical charge loss. The vertical charge loss towards surface state is ignored here since the decrease in both potential peak and diameter (<i>i.e.</i> , decrease in KFM resolution) during the initial stage of retention suggests the charges diffuse further away from the tip rather than towards surface states.	124
5.13	Transient discharging current (I_{relax}) measurement in (a) Pt and (b) Au samples at temperatures ranging from 300 to 450 K. The inset shows the	126

	change in magnitude of the slope n with temperature from the curve fitting: $I_{\text{relax}}=at^{-n}$.	
5.14	Schematic of dominant charge decay paths in SL-MNCs embedded sample. E_x and E_z correspond to the localized electric field in lateral and vertical directions, respectively. L_x and L_z represent inter-dot spacing and tunnel oxide thickness, respectively.	127
5.15	(a) The potential profile across the KFM image of injected negative charges. The slope of potential profile (dV/dx) indicates the localized electric field E_x between adjacent MNCs. Region-A represents the MNCs with maximum potential and E_x is minimum, while Region-B corresponds to the steepest slope of the potential profile where E_x is maximum. (b) and (c) show the energy band diagrams of adjacent MNCs at Region-A and -B, respectively.	128
6.1	Schematic (not to-scale) of DL-MNC embedded gate stack (a) for KFM characterization and (b) with Pt top electrode for microscopic electrical measurement.	137
6.2	(a) Plan view TEM micrograph of Pt-MNCs in DL device after forming the first MNC layer. Cross-sectional TEM micrographs of (b) DL2 and (c) DL3 samples with different ILD thickness.	138
6.3	Sample CV curve showing shift in the flatband voltage from “fresh” to “program” (P) state after applying +18V for 100 ms at room temperature in DL2 (open symbols) and DL3 (closed symbols) devices.	139
6.4	(a) Charge retention in DL2 and DL3 devices at temperature ranging from 298-398K. (b) Box plot of the retention time for 5% charge loss.	140
6.5	Schematic energy band diagram of DL device under an application of (a) positive bias and (b) negative bias to the top electrode during retention. An application of a positive bias accelerates the charge leakage through Al_2O_3 control oxide, while a negative bias enhances the carrier detrapping across SiO_2 tunnel oxide.	142
6.6	Retention charge loss from the Program state at various gate biases for 15 mins in the DL2 and DL3 devices.	143
6.7	(a) Potential profiles of the DL2 and DL3 samples after negative charge injection. (b) Evolution of FWHM and $1/e^2$ width with various potential peaks (<i>i.e.</i> , various injection voltages).	146

6.8	Preferred paths of charging in DL-MNCs structure under gate injection. P_1 represents charge diffusion to adjacent MNCs in the same MNC-layer and P_2 corresponds to the charging of bottom layer through top layer MNCs. P_3 shows the possibility of direct charge injection to bottom MNC layer depending on the vertical alignment of MNCs in DL gate stacks.	147
6.9	Schematic illustration of charge distribution in two MNC-layers of (a) DL3 and (b)DL2 depending on the ILD thickness under gate (AFM tip) injection.	149
6.10	The schematic energy band diagrams according to the charge distribution and internal electric field during retention in (a) DL2 and (b) DL3 devices. Labels 1-4 correspond to the possible charge loss processes through trap-assisted-tunneling, thermally assisted tunneling, thermionic emission and direct tunneling through rectangular barrier, respectively.	151
6.11	The change in (a) potential peak and (b) FWHM as a function of retention time in the DL3 samples with various injection voltages. The right column of (b) presents the zoom-in view of FWHM evolution with time. The arrow shows the trend of the initial amount of the injected charge (from smallest to largest).	155
6.12	The change in (a) potential peak and (b) FWHM as a function of retention time in the DL2 samples with various injection voltages. The arrow shows the trend of the initial amount of the injected charge (from smallest to largest).	156
6.13	(a) Transient discharging (relaxation) current measurement in the DL3 sample at temperature ranging from 300 to 450 K. (b) The change in the magnitude of slope n with temperature from the curve fitting.	158
6.14	Illustration of possible charge decay paths during retention in DL-MNC embedded gate stacks: P_x corresponds to inter-dot tunneling to MNCs in the same layer; while, P_{z-1} and P_{z-2} refer to inter-layer tunneling through ILD and vertical charge loss through SiO_2 , respectively.	159
6.15	Schematic illustration of unequal charge distribution in two MNC-layers upon gate injection. E_x and E_z are corresponding to the change in localized electric field in lateral and vertical directions. L_x and t_{ILD}	160

	represent inter-dot spacing and inter-layer dielectric thickness between two MNC-layers, respectively.	
6.16	(a) Stretched exponential fitting to the charge decay of DL2 and DL3 samples after various negative charge injection (-4 to -10V) for 10 sec. Solid symbols correspond to charge decay of DL3 and open symbols represent charge decay of DL2. (b) Box plot of the stretched exponent β in the DL3 and DL2 samples.	162

LIST OF TABLES

Table	Caption	Page
4.1	Stressing condition of Au-NC samples.	90
4.2	I_g evolution modes of Au-NCs samples under CVS.	96
5.1	Summary of charge transport in SL-MNCs embedded gate stacks.	130
6.1	Summary of charge transport in DL-MNCs embedded gate stacks.	163

LIST OF ABBREVIATIONS

Abbreviations	Full expressions
Al_2O_3	Aluminium oxide
AFM	Atomic force microscopy
Au	Gold
BTI	Bias temperature instability
CAFM	Conductive atomic force microscopy
CMOS	Complimentary metal-oxide-semiconductor
Co	Cobalt
CTF	Charge trap flash
C-V	Capacitance-voltage
CVS	Constant voltage stress
DL	Dual layer
EDS	Electron dispersive spectrometry
EELS	Electron energy loss spectrometry
EFM	Electrostatic force microscopy
FG	Floating gate
FN	Fowler-Nordheim (tunnelling)
FWHM	Full-width-at-half maximum
Ge	Germanium
HFCV	High frequency C-V
I_g	Gate (top electrode) leakage current
I_{gl}	Gate leakage compliance current limit
IL	Interfacial layer
ILD	Inter-layer dielectric

IPD	Inter-poly dielectric
ISSG	<i>In-situ</i> steam generation
KFM	Kelvin force microscopy
MNC	Metal nanocrystal
MOSFET	Metal-oxide-semiconductor field effect transistor
NC	Nanocrystal
Ni	Nickel
Pt	Platinum
PVD	Physical vapor deposition
RTP	Rapid thermal process
Si	Silicon
SILC	Stress induced leakage current
SiN	Silicon nitride
SiO ₂	Silicon dioxide
SL	Single layer
SONOS	Silicon-Oxide-Nitride-Oxide-Silicon
TANOS	Silicon-Oxide-Aluminium Oxide-Tantalum Nitride
SPM	Scanning probe microscopy
t _{BD}	Time-to-dielectric breakdown
TDDB	Time dependent dielectric breakdown
TEM	Transmission electron microscopy
V _{BD}	Dielectric breakdown voltage
V _{BDA}	Average breakdown voltage
V _{g-stress}	Stressing voltage

CHAPTER ONE

INTRODUCTION

1.1 Background

With increasing demand in the consumers' electronics, the capacity of flash memories continuously grows in the nonvolatile memory technology by scaling the dimensions of floating gate (FG) cell and using multi-level-cell storage [1-4]. A conventional FG flash memory stores charges in the inter-poly dielectric (IPD) in between two thick layers of dielectric materials, isolating it from the channel and control gate. Although the thin tunnel oxide is required to reduce the operating voltages, the thickness of the oxide cannot be too thin to prevent charge leakage during retention. Also the IPD in conventional FG memories must be scaled with the tunnel oxide to maintain adequate coupling to the tunnel oxide. Without compromising data retention requirement, the reduction in the tunnel oxide thickness is a major challenge due to stress-induced-leakage-current (SILC) related defect generation and pinhole defect formation in the tunnel oxide. The requirement of high- κ IPD to reduce the thickness of IPD in the planar FG results in more reliability concerns [5-9]. Therefore, the scaling of both tunnel oxide and IPD remains as challenges [10] to maintain high gate coupling ratio in FG flash memories. Hence the physical scaling of the FG cell is reaching a limit at the sub-30 nm technology node. According to ITRS 2011 [10], the charge trap flash (CTF)

(SONOS/TANOS) based memory is expected to replace FG flash at sub-30 nm technology node. In the CTF based memory, the charges are stored in the bulk traps of silicon nitride (SiN) and hence the defect related charge loss and coupling problem in FG cell will be reduced. However, the retention loss especially after cycling is the main challenge in CTF devices and it is related to shallow trap depth of the SiN storage layer which is an inherent material property and is difficult to control [11-12]. Therefore, the flash memory device using many discrete charge-storage nodes such as nanocrystals (NCs) has received considerable research attention. It is considered as one of the possible alternatives to replace the FG devices because of its superior memory performance, high scalability, similar memory process as CTF and reduction in mask count [13] over existing FG flash memory process.

In 1995, S. Tiwari *et al.*[14] proposed NC memories using Si-NC embedded in SiO₂ tunnel and control oxide films. With the discrete storage node structure, it is believe that the charge leakage during retention is reduced and hence the thicknesses of tunnel and control oxide layers can be scaled down. This can lead to a reduction in the operating voltage required for flash memory operation. In the early stages of NC memories research, various type of semiconductor NCs (such as Si, Ge, and SiGe [14-17]) have been extensively reported for memory applications. Though semiconductor NCs are compatible with CMOS process, their small work function leads to small memory windows and quantum confinement results in poor retention. In 2009, the Si-NC based memories have been commercially released as embedded flash products [18]. However the inability to control the Si-NC size has limited their application in NOR memories, whose size scaling is not as aggressive as NAND. To overcome the limitations of the

semiconductor NCs, devices based on metal nanocrystal (MNC) are proposed [19-20] and the performances of various metal NCs (such as Au, Pt, Ni, Co, *etc.* [21-24]) have been studied. Due to large density of states around the Fermi level of metals, quantum confinement effect is reduced and hence barrier height for the stored charges increases [25], providing larger memory window and longer retention than semiconductor NCs. Moreover, availability of various work functions in metals and a strong coupling between the channel and MNCs provide significant advantages over semiconductor NCs.

1.2 Motivation

Though MNCs have shown better performance than the semiconductor NCs, the reported memory window is still smaller than CTF memory devices. This is because the NC density is relatively small (20-25% area coverage and NC density of less than $1 \times 10^{12} \text{ cm}^{-2}$ in the previous reported devices) and non-uniform size distribution [26]. Therefore, the large number of NC density and uniform size control are required to provide a significant threshold voltage shift. This is more important in the scaled devices because the random placement of NCs could result in the number fluctuations which affect device performance. Moreover, the NC density must be much higher (at least one order of magnitude) than the density of background charges (*i.e.*, interfacial states charges) so that the memory operations can be dominated by the embedded-NCs [27]. Therefore, the requirement of large number density is inevitable in the NC memory devices.

On the other hand, a large number of NC densities could cause inter-dot tunneling among adjacent NCs leading to charge loss through lateral paths to the contact regions and/or through the defect chain in the oxide. Since the reduction of defects-related charge loss is one of the fundamental motivations of NC memory research, it is critical to understand the lateral charge diffusion phenomenon and charge transport mechanisms in NC embedded gate stack.

The NCs are conventionally embedded in the high- κ dielectric, such as HfO_2 [28-29], HfAlO [21, 30], and Al_2O_3 [31-34], to increase the coupling between NC and control gate and to reduce the operating voltages. For high quality tunnel oxide, thermally grown SiO_2 is widely used because of its large bandgap and low interface trap density. Therefore, the overall gate dielectric stack in NC memory includes SiO_2 and NC-embedded high- κ layer. The gate dielectric reliability of SiO_2 and high- κ materials has been an intensive research topic aiming to reduce leakage current in CMOS generations. The time-dependent dielectric breakdown (TDDB), threshold voltage instability, bias-temperature instability (BTI), *etc.*, have been studied to understand the breakdown and degradation mechanisms. However, there are a few researches on the dielectric degradation and failure of NC embedded high- κ/SiO_2 gate stacks. The systematic understanding on the degradation and breakdown issues of the NC embedded high- κ/SiO_2 dual layer gate stack is critical for the future design and development of the next generation non-volatile memory devices.

A lot of research has been done to improve performance of MNC devices from the material and fabrication aspects and most attention has been given to incorporating metals in the gate stack. However, there is still lack of extensive study on reliability

aspects, such as inter-dot tunneling among NCs, the charge loss and charge transport mechanisms and their performances in the degraded dielectric which are of utmost importance for the device optimization and future scaling.

1.3 Objectives

The main objective of this study is to focus on the degradation and reliability of MNC embedded gate stacks, to contribute to a deeper understanding of MNC performance in the degraded dielectric gate stacks and its charge transport behaviors, and to report extensively on the reliability characterization of both single-layer (SL) and dual-layer (DL) MNC embedded gate stacks. The ultimate goals of this study are (1) to investigate the failure mechanisms and impact of individual dielectric layer degradation in the NC-embedded SiO₂/high- κ gate stacks, (2) to understand localized charge transport and lateral charge diffusion of SL and DL MNCs embedded gate stacks, and (3) to study their charge loss mechanisms. With all the critical results, we aim to envision an optimized MNC embedded memory cell design for enhancement of retention ability in NC memory devices.

1.4 Organization of Thesis

This thesis consists of seven chapters. The structure of the thesis is as follows:

Chapter *One* introduces the background, the motivation and objectives of this work, as well as the major contributions of the research.

Chapter *Two* is a literature review on the related researches that has been done. It provides a literature survey on the study of breakdown in SiO₂ and high- κ gate dielectric and gives an overview of reliability challenges in nanocrystal embedded oxides.

Chapter *Three* presents the experimental procedures and details. It describes the details of the devices used in our study and explains the electrical characterization techniques applied in the research, including nanoscale characterization using scanning probe microscopy.

Chapter *Four* provides the detailed study of MNCs in the degraded dielectric gate stacks. It describes a comprehensive study of charging and discharging characteristics of MNCs in the degraded dielectric gate stacks. This is followed by the individual dielectric breakdown layer detection in MNC embedded high- κ /SiO₂ gate stacks.

Chapter *Five* focuses on the study of localized charge transport and lateral charge diffusion in single layer MNC embedded gate stacks using electrical and nanoscale characterization approaches.

Chapter *Six* discusses the charge transport in dual layer MNC-embedded gate stacks. The study of charge distribution in two MNC-layers and the charge loss mechanisms are presented.

Chapter *Seven* concludes the main findings in this dissertation and it gives the recommendations for future study in the reliability of NC embedded gate stacks.

1.5 Contributions

Our studies provide unique insights into the leakage and degradation mechanisms of MNC embedded high- κ /SiO₂ gate stack through a series of reliability investigations and contribute to a deeper understanding of characteristics of MNCs in the degraded dielectric gate stacks and their charge transport behaviors.

The influences of dielectric degradation on charging and discharging characteristics MNCs and the individual dielectric layer degradation behaviors in NC-embedded bi-layer dielectric gate stacks are studied. The novelty of this work is the investigation of how many storage nodes would be involved when the dielectric is degraded in NC based memory gate stack, which could be a limiting factor for a future scaling. We believe this research will contribute to further improvement of promising NC-embedded gate stacks from a reliability perspective since the previous understanding predicted that the breakdown in the tunneling dielectric or blocking oxide is a localized phenomenon which affects limited NCs. Also, it provides the physical evidence of lateral

charge diffusion phenomenon in MNCs embedded gate stack with MNC area coverage of ~30%.

A new methodology to detect the sequence of dielectric layer breakdown in MNC embedded high- κ /SiO₂ gate stack is proposed based on the Coulomb charging energy in relation to thermal energy gained by electrons at low voltage and in the very low temperature regime ranging from 11K to 70K. With this approach, the average trap size in NC-embedded Al₂O₃ and SiO₂ is further analyzed. The presented results may pave way for further investigation in optimizing the gate stacks so as to improve robustness of these NC memory stacks due to dielectric degradation.

The lateral charge diffusion of MNCs observed in the degraded dielectric stacks is further confirmed by the nanoscale characterization using Kelvin Force Microscopy (KFM). With this technique, the real time evolution of localized charge transport in MNC embedded devices is studied. The transient discharging current measurements are carried out at various temperatures to decouple lateral charge diffusion and vertical charge loss of charge decay observed in the KFM. From a comparative study of MNCs with different work functions, we find that the MNCs with higher work function have a lower inter-dot charge tunneling probability, which is favorable for improving retention in memory applications. However, the vertical charge loss during the initial decay period is a trade-off and it could be minimized by using a dual-layer MNC structure. With the continual scaling down of memory devices, characterization of the trapped charge distribution and charge decay mechanisms in submicron resolution gate stacks is crucial and our study shows the localized charge transport behaviors of MNC embedded devices for implementation of reliable NC-based flash memory technology.

Since the performance of SL-MNCs clearly indicates the need for DL-MNCs for future NC based memory, the DL-MNCs with high work function MNCs are chosen to study their intrinsic charge loss mechanisms, inter-dot tunneling and inter-layer tunneling characteristics. It is found that charge loss in DL structures is mainly due to the internal electric field induced by trapped electrons which depends on the inter-layer dielectric (ILD) thickness and MNC spacing. When the ILD thickness is comparable to MNCs spacing, the charge distribution in two-MNC layers are similar and the internal electric field induced by charges stored in the MNCs is higher. It leads to the fact that internal-electric-field assisted tunneling dominates and significant charge loss occurs, especially at lower temperatures. In contrast, when the ILD thickness is larger than the average MNC spacing, less charge is injected into the neighboring MNC layer, resulting in a reduction of both internal electric field and oxide trap alignment possibility with the traps inside the MNCs. Moreover, a thick ILD reduces the electron tunneling probability from one MNC-layer to another during retention. Our findings suggest that an optimized DL-MNCs embedded memory cell could be achieved by defining the ILD thickness larger than the average MNC-spacing for enhancement of retention ability in MNC embedded gate stacks. It implies the possibility of reducing MNC spacing in scaled memory devices by using DL-MNCs with controlled ILD thickness to achieve both increased memory window and prolonged retention.

The above work has been published in a world-renowned conference, *i.e.*, *IEEE International Reliability Physics Symposium*, as well as prestigious journals: *Electron Device Letters* and *Applied Physics Letters*.

References

- [1] K. Shimizu, K. Narita, H. Watanabe, E. Kamiya, Y. Takeuchi, T. Yaegashi, S. Aritome, and T. Watanabe, "A novel high-density $5F^2$ NAND STI cell technology suitable for 256 Mbit and 1 Gbit flash memories," *IEEE Tech. Dig. IEDM 1997*, pp. 271-274.
- [2] S. Lee, Y. T. Lee, W. K. Han, D. H. Kim, M. S. Kim, S. H. Moon, H. C. Cho, J. W. Lee, D. S. Byeon, Y. H. Lim, H. S. Kim, S. H. Hur, and K. D. Suh, "A 3.3V 4Gb four-level NAND flash memory with 90nm CMOS technology," *IEEE Tech. Dig. ISSCC 2004*, pp. 52-53.
- [3] J. H. Park, S. H. Hur, J. H. Lee, J. T. Park, J. S. Sel, J. W. Kim, S. B. Song, J. Y. Lee, J. H. Lee, S. J. Son, Y. S. Kim, M. C. Park, S. J. Chai, J. D. Choi, U. I. Chung, J. T. Moon, K. T. Kim, K. Kim, and B. I. Ryu, "8Gb MLC (multi-level cell) NAND flash memory using 63nm process technology," *IEEE Tech. Dig. IEDM 2004*, pp. 873-876.
- [4] M. Noguchi, T. Yaegashi, H. Koyama, M. Morikado, Y. Ishibashi, S. Ishibashi, K. Ino, K. Sawamura, T. Aoi, T. Maruyama, A. Kajita, E. Ito, M. Kishida, K. Kanda, K. Hosono, S. Miyamoto, F. Ito, G. Hemink, M. Higashitani, A. Mak, J. Chan, M. Koyanagi, S. Ohshima, H. Shibata, H. Tsunoda, and S. Tanaka, "A high-performance multi-level NAND flash memory with 43 nm-node floating-gate technology," *IEEE Tech. Dig. IEDM 2007*, pp. 445-448.
- [5] N. Mielke, "Reliability of floating-gate flash memories," *IEEE Intl. Integrated Reliability Workshop 2007*, p. 151.
- [6] A. Ghetti, C. M. Compagnoni, F. Biancardi, A. L. Lacaita, S. Beltrami, L. Chiavarone, A. S. Spinelli, and A. Visconti, "Scaling trends for random telegraph noise in deca-nanometer flash memories," *IEEE Tech. Dig. IEDM 2008*, p. 4.
- [7] C. Yi-Chou and R. Liu, "Future trend of flash memories," *IEEE Memory Technology, Design and Testing MTDT 2007*, pp. 17-18.
- [8] J. De Vos, D. Wellekens, I. Debusschere, J. Van Houdt, S. R. A. Van Aerde, P. R. Fischer, and P. M. Zagwijn, "Scaling of floating gate electrode for sub-40nm flash technologies," *Proceedings of the 38th European Solid-State Device Research Conference*, 2008, pp. 123-125.

- [9] C. Y. Lu, K. Y. Hsieh, and R. Liu, "Future challenges of flash memory technologies," *Microelectronic Engineering*, vol. 86, no. 3, pp. 283-286, 2009.
- [10] ITRS 2011, "International Technology Roadmap for Semiconductors," available at <http://www.itrs.net>
- [11] C. Sandhya, U. Ganguly, K. K. Singh, P. K. Singh, C. Olsen, S. M. Seutter, R. Hung, G. Conti, K. Ahmed, N. Krishna, J. Vasi, and S. Mahapatra, "Nitride engineering and the effect of interfaces on Charge Trap Flash performance and reliability," *Proceedings of the Intl. Reliab. Phys. Symp.*, 2008, pp. 406-411.
- [12] C. Sandhya, U. Ganguly, N. Chattar, C. Olsen, S. M. Seutter, L. Date, R. Hung, J. M. Vasi, and S. Mahapatra, "Effect of SiN on performance and reliability of charge trap flash (CTF) under Fowler-Nordheim tunneling program/erase operation," *IEEE Electron Device Lett.*, vol. 30, no. 2, pp. 171-173, 2009.
- [13] R. Muralidhar et al., "A 6 V embedded 90 nm silicon nanocrystal non-volatile memory," *IEEE Tech. Dig. IEDM 2003*, pp. 601-604.
- [14] S. Tiwari, F. Rana, K. Chan, H. Hanafi, C. Wei, and D. Buchanan, "Volatile and non-volatile memories in silicon with nano-crystal storage," *IEEE Tech. Dig. IEDM 1995*, pp. 521-524.
- [15] H. I. Hanafi and S. Tiwari, "A controlled threshold low power nano-crystal memory," *Proceedings of the 25th European Solid State Device Research Conference*, 1995, pp. 209-212.
- [16] H. C. Jing, Q. W. Ying, J. Y. Won, Y. Yee-Chia, G. Samudra, D. S. Chan, Y. D. An, and K. Dim-Lee, "Nonvolatile flash memory device using Ge nanocrystals embedded in HfAlO high- κ tunneling and control oxides: Device fabrication and electrical performance," *IEEE Trans. Electron Devices*, vol. 51, no. 11, pp. 1840-1848, Nov.2004.
- [17] R. Gupta, W. J. Yoo, Y. Wang, Z. Tan, G. Samudra, S. Lee, D. S. H. Chan, K. P. Loh, L. K. Bera, N. Balasubramanian, and D. L. Kwong, "Formation of SiGe nanocrystals in HfO₂ using in situ chemical vapor deposition for memory applications," *Appl. Phys. Lett.*, vol. 84, no. 21, pp. 4331-4333, 2004.
- [18] J. Yater, M. Suhail, S. T. Kang, J. Shen, C. Hong, T. Merchant, R. Rao, H. Gasquet, K. Loiko, B. Winstead, S. Williams, M. Rossow, W. Malloch, R.

- Syzdek, and G. Chindalore, "16Mb Split Gate Flash Memory with Improved Process Window," *IEEE Intl. Memory Workshop 2009*, pp. 1-2.
- [19] Z. Liu, M. Kim, V. Narayanan, and E. C. Kan, "Process and device characteristics of self-assembled metal nano-crystal EEPROM," *Superlattices and Microstructures*, vol. 28, pp. 393-399, 2000.
- [20] Z. Liu, C. Lee, V. Narayanan, G. Pei, and E. C. Kan, "Metal nanocrystal memories. I. Device design and fabrication," *IEEE Trans. Electron Devices*, vol. 49, no. 9, pp. 1606-1613, Sept.2002.
- [21] K. C. Chan, P. F. Lee, and J. Y. Dai, "Single electron tunneling and Coulomb blockade effect in HfAlO/Au nanocrystals/HfAlO trilayer nonvolatile memory structure," *Appl. Phys. Lett.*, vol. 92, no. 14, pp. 143117-143120, 2008.
- [22] T. Jiun-Yi, C. Cheng-Wei, W. Sheng-Yu, W. Tai-Bor, H. Kuang-Yeu, and L. Rich, "Memory characteristics of Pt nanocrystals self-assembled from reduction of an embedded PtOx ultrathin film in metal-oxide-semiconductor structures," *Appl. Phys. Lett.*, vol. 85, no. 13, pp. 2595-2597, 2004.
- [23] J. L. Jong, Y. Harada, W. P. Jung, and K. Dim-Lee, "Nickel nanocrystal formation on HfO₂ dielectric for nonvolatile memory device applications," *Appl. Phys. Lett.*, vol. 86, no. 10, pp. 103505-103507, 2005.
- [24] L. Jianlin, Z. Dengtao, and Z. Yan, "Charge storage in a metal-oxide-semiconductor capacitor containing cobalt nanocrystals," *Solid-State Electronics*, vol. 50, no. 2, pp. 268-271, 2006.
- [25] W. Guan, S. Long, M. Liu, Q. Liu, Y. Hu, Z. Li, and R. Jia, "Modeling of retention characteristics for metal and semiconductor nanocrystal memories," *Solid-State Electronics*, vol. 51, pp. 806-811, 2007.
- [26] P. K. Singh, "Metal Nanocrystal Memory for NAND Flash Applications", Ph.D dissertation, Indian Institute of Technology, Bombay, 2009.
- [27] S. Oda, and S. Huang *Silicon Nanocrystal Flash Memory*, in *Silicon Nanocrystals: Fundamentals, Synthesis and Applications* edited by L. Pavesi and R. Turan (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010), pp. 395-438.

- [28] J. L. Jong, W. Xuguang, B. Weiping, L. Nan, and K. Dim-Lee, "Theoretical and experimental investigation of Si nanocrystal memory device with HfO_2 high- κ tunneling dielectric," *IEEE Trans. Electron Devices*, vol. 50, no. 10, pp. 2067-2072, 2003.
- [29] T. H. Ng, V. Ho, L. W. Teo, M. S. Tay, B. H. Koh, W. K. Chim, W. K. Choi, A. Y. Du, and C. H. Tung, "Fabrication and characterization of a trilayer germanium nanocrystal memory device with hafnium dioxide as the tunnel dielectric," *Thin Solid Films*, vol. 462-463, pp. 46-50, 2004.
- [30] J. H. Chen, Y. Q. Wang, W. J. Yoo, Y. C. Yeo, G. Samudra, D. S. H. Chan, A. Y. Du, and D. L. Kwong, "Nonvolatile flash memory device using Ge nanocrystals embedded in HfAlO High- κ tunneling and control oxides: Device fabrication and electrical performance," *IEEE Trans. Electron Devices*, vol. 51, no. 11, pp. 1840-1848, 2004.
- [31] Q. Wang, Z. T. Song, W. L. Liu, C. L. Lin, and T. H. Wang, "Synthesis and electron storage characteristics of isolated silver nanodots on/embedded in Al_2O_3 gate dielectric," *Applied Surface Science*, vol. 230, no. 1-4, pp. 8-11, 2004.
- [32] Z. Min, C. Wei, D. Shi-Jin, L. Zhi-Ying, H. Yue, L. Zhong-Wei, and D. W. Zhang, "Physical and electrical characterization of atomic-layer-deposited Ru nanocrystals embedded into Al_2O_3 for memory applications," *J. Phys. D: Appl. Phys.*, vol. 41, no. 3, p. 032007, 2008.
- [33] S. Maikap, S. Z. Rahaman, and T. C. Tien, "Nanoscale (EOT = 5.6 nm) nonvolatile memory characteristics using n-Si/ SiO_2 / HfAlO nanocrystal/ Al_2O_3 /Pt capacitors," *Nanotechnology*, vol. 19, no. 43, p. 435202, 2008.
- [34] P. K. Singh, G. Bisht, K. K. Singh, R. Hofmann, N. Krishna, and S. Mahapatra, "Metal Nanocrystal Memory with Pt Single- and Dual-Layer NC with Low-Leakage Al_2O_3 Blocking Dielectric", *IEEE Electron Device Lett.*, vol. 29, pp. 1389-1391 Dec. 2008.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter provides a review of the recent progress in the area of semiconductor and metal nanocrystal (NC) based floating gate flash memories and the associated reliability challenges reported in the literature. Also, it presents a brief overview of the dielectric breakdown phenomenon in conventional silicon dioxide and high- κ gate stacks.

2.2 Overview of NC-Based Gate Stacks Reliability

The memory cell structures employing discrete charge storage nodes were firstly introduced in 1995 [1] to replace the continuous floating gate layer in the conventional flash memories. The NC-based memory devices can be classified into two major categories based on the NC material: semiconductor and metal NC. The F/N tunneling serves as the main mechanism for Program/Erase (P/E) processes in NC memory. However, in Si-NC memory, the strong electric field cannot be confined in one oxide layer when there is a relatively thick tunnel oxide. In MNC memory, a wide range of

available work functions allow flexible selection of the work function of both MNC and the control gate such that the barrier height, and hence, the turn-on electric field for FN tunneling from the NC and the control gate can efficiently be engineered [2].

Figure 2.1. Band diagram illustration for MNC memory in (a) Erase and (b) Program processes [2]. In this illustration, erase operation is performed with a positive control gate voltage and write operation is carried out by applying a negative control gate voltage.

As illustrated in the band diagram of MNC memory in P/E processes (Fig. 2.1), it can be seen that F-N tunneling probability is very strongly dependent on the work functions of MNC and control gate since they modulate both the height and width of the barrier. For instance, if the work function of the gate is larger than the electron affinity of Si, tunneling will be limited within tunnel oxide and erasing can be performed by extracting electrons from the MNCs. Similarly, when the work function of the gate is smaller than the electron affinity of Si, only a small control gate bias is required to initiate tunneling only in the control oxide and extra electrons can be added to the MNCs.

After selection of the work function of control gate, the threshold voltage can be varied by choosing an appropriate MNC work function.

Replacing continuous floating gate layer by NCs and the choice of NC material attract a number of new physical research topics, such as quantum effects (charge confinement and Coulomb blockade), band engineering, NC density and size control. This section reviews the work done on development of NC memories and the associated challenges.

2.2.1 Size Scaling in NCs

When the size of the nanocrystals scales down to a few nanometers, the Coulomb blockade effect and quantum confinement effect become significant. The energy levels in bulk material are very close so that they are described as continuous. When the size approaches the material's Exciton Bohr Radius, the energy levels become discrete (*i.e.*, finite separation between energy levels) due to the confinement of electron wavefunction. Fig. 2.2 shows the density of states per volume and energy for 3-D semiconductor, 2-D quantum well, 1-D quantum wire and 0-D quantum dot reported by B. V. Zeghbroeck [3].

Figure 2.2. Density of states per volume and energy for a 3-D semiconductor (blue), 2-D quantum well (red), 1-D quantum wire (green) and 0-D quantum dot (black) [3].

When the NC is charged with one elementary charge by the tunneling electron, the single electron charging energy E_c arises which is given by [4]

$$E_c = \frac{e^2}{C} \quad (2.1)$$

where e is the elementary charge of 1.6×10^{-19} Coulomb and

$$C = 4\pi\epsilon_0\epsilon_{ox}R \quad (2.2)$$

is the self capacitance of the NC with a radius R . The self charging energy of the NC raises the electrostatic potential U with

$$U = \frac{e}{C} \quad (2.3)$$

which induces an electric field. It prevents tunneling of a second electron, unless the applied voltage is large enough to overcome it and this effect is named as Coulomb blockade.

Fig. 2.3(a) shows the single-electron charging energy of MNCs as a function of NC diameter [5]. It can be found that charging energy decreases as the NC diameter increases, which is solely attributed to the larger NC capacitance with increasing NC size. As a result, more charges could be stored in the NC with larger diameter providing a larger memory window (Fig. 2.3(b)). Therefore, if the NC number density remains the same and the inter-dot tunneling is negligible, an increase in NC diameter possesses the advantages both of programming and retention characteristics.

(a)

(b)

Figure 2.3. (a) Single-electron charging energy of Au-NCs as a function of NC diameter and (b) flatband voltage shift in NCs with various diameters D [5].

The Coulomb blockade effect raises the electrostatic potential of the NC, while the quantum confinement effect shifts the NC energy band edge upward [6], resulting in a reduction of the energy band offset between the NC and the surrounding dielectric, which leads to retention issues in a long term. According to Kubo theory [7], the average energy level spacing δ in a metal NC can be calculated as:

$$\delta = \frac{4E_F}{3M} \quad (2.4)$$

where E_F is the Fermi potential of the bulk metal, and M is the total number of valence electrons in the NC. Using equation (2.4), the energy level spacing δ for a given metal NC with various diameters can be calculated. In semiconductor NC, the quantum confinement effect results in widening of the band gap and the upward shift of the conduction band minimum. The analytical law for the upward shift of conduction band minimum in Si-NC is derived [6] as:

$$E_C(d_{Si-NC}) = E_C(\infty) + \frac{1.39}{d_{Si}^2 + 1.79d_{Si-NC} + 0.668} \quad (2.5)$$

where $E_C(\infty)$ is the conduction band minimum for bulk Si and d_{Si-NC} is the diameter of Si-NC in nm. Fig. 2.4(a) shows the conduction band minimum up-shift of the semiconductor NCs and the Fermi level up-shift of the metal NC as a function of NC size. It can be seen that the quantum confinement effect in metal-NC is greatly reduced which is attributed to a large density of states around the Fermi level of metals. The Fermi level up-shift in metal NC is only obvious for a NC size less than 2 nm in diameter. For the state of the art MNC memories, the diameter is always larger than 2 nm [2].

Figure 2.4. (a) Conduction band minimum up-shift of the semiconductor NCs and Fermi level up-shift of the metal NC and (b) retention time for various metal and semiconductor NC memories as a function of NC size [6].

The effect of size scaling in retention ability of metal and semiconductor NC memory is reported as shown in Fig. 2.4(b) [6]. The retention time is defined as the time when the 50% of charge loss from the NCs. It shows that the retention time in metal NC insignificantly changes with NC size, while in semiconductor NCs, the retention time decreases notably with reducing NC size. The difference in retention ability of metal and semiconductor NCs is ascribed to weaker quantum confinement effect in MNC and higher energy barrier for electron from the MNC to the substrate.

2.2.2. NC Spacing Effect

From the threshold voltage shift in Si-NC memory due to electron storage [8]

$$\Delta V_{th} = \frac{n_1 p q}{\epsilon_{ox}} \left(t_{control} + \frac{\epsilon_{ox} d}{2 \epsilon_{Si}} \right) \quad (2.6)$$

where n_1 is the NC number density, p is the average number of electron stored in the NC, $t_{control}$ is the control oxide thickness and d is the average diameter of NC, it is clear that greater n_1 , p and d are desirable for the NC memory with a larger memory window. Since the density of NCs is inversely proportional to the NC spacing, the requirement of large number density results in a closer dielectric spacing between adjacent NCs. As shown in Fig. 2.5(a), the single-electron charging energy in metal NCs increases with decreasing NC spacing [5]. It indicates that the Coulomb blockade effect in nanoparticles becomes larger with increasing NC density. This can be explained by the smaller NC capacitance due to the smaller NC unit cell area, increasing the single-electron charging energy. Fig. 2.5(b) shows the effect of NC density on the memory window as a function of programming pulse time [5]. The device with a denser NC array provides a larger memory window assuring less fluctuation of the memory window across devices when scaling to smaller cell areas [9].

The reduction in the NC density and cell area for the scaling of NC memory also results in an increase in program fail probability; in which the average threshold voltage V_{th} is lower than the program verify level, typically 75% of ΔV_{th} (Fig. 2.6) [10]. The significant increase in the program fail probability with reducing NC density and cell area is attributed to the lower average number of NCs controlling cell state.

Figure 2.5. Effect of NC spacing on (a) single-electron charging energy and (b) flatband voltage shift as a function of programming time in Au-NCs with diameters of 5 nm [5].

Figure 2.6. Program fail probability of various cell areas ($W \times L$) as a function of NC density [10].

Therefore, it can be seen that larger NC spacing and density could improve the memory window and retention characteristics. However, these cannot be implemented simultaneously because a larger NC density corresponds to a smaller NC spacing and stronger lateral charge transfer among NCs. Also, larger NC diameter and greater NC spacing limit the scaling of cell areas. Motorola reported implementation of Si-NC based 4 Mbit prototype using a conventional 90 nm process technology in 2003 [11]. However, it is found to be challenging to commercialize. This is because the scaling in NC-based floating gate flash transistors gives concern of number of NCs embedded in each NC-based floating gate memory. If the number of NCs is reduced in the scaled cell, statistical variation between transistors could arise. The decrease in the NC size and increase in the NC density might solve the problem [12]. However, NC size scaling is only possible for the metal NC device, since the quantum confinement effect in Si-NC is much severe than MNC. It is very clear that the requirement of large number density is inevitable in the NC memory devices. On the other hand, increasing the nanocrystal density will increase the inter-dot tunneling among adjacent NCs. Since the main motivation of introducing NC technology in memory application is to reduce SILC induced charge leakage, allowing a significant reduction of the tunnel oxide thickness. The strong inter-dot tunneling among NCs could give rise to charge transfer towards the SILC site bringing the cell to retention failure.

Some researchers used AFM [13, 14] / EFM [15-18] or Kelvin Probe technique [19] to investigate the possible carrier transport mechanisms of localized charge storage of NCs in SiO₂ matrix. The increase in the area of localized charge spot is mainly attributed to the charge diffusion from charged NC to neighbor uncharged NCs [16, 17].

Moreover, the polarity of neighboring charges has strong influence on the charge diffusion rate of NCs [16]. The same polarity of neighboring charges leads to a longer charge decay, while the opposite polarity promotes the charge dissipation, leading to a shorter charge decay time. Q. Yu. et al. [20] reported the lateral charge diffusion of Si-NC in oxide using a device simulation approach. It is found that the flat band voltage shift in charged NC-embedded device could easily affect the flat band voltage of neighbor device with a NC-embedded dielectric spacing of 25 nm in between them [20]. It suggested that the lateral charge diffusion can cause the interference among the neighboring devices. Since the lateral charge diffusion could have impact on the charge retention capability of memory device, as well as interaction with the neighboring devices, a good understanding on this is necessary for the NC-embedded device application.

2.3 Overview of Gate Dielectric Reliability

In this section, the gate dielectric breakdown phenomena in SiO₂ and high- κ dielectric, which are conventionally used as a tunnel oxide and a control oxide, respectively, in the NC-based memory devices, are reviewed. Also, the dielectric relaxation current, which is widely used to access the gate dielectric integrity, is briefly discussed.

2.3.1 Gate Dielectric Breakdown

Gate dielectric breakdown is generally defined as the loss of the insulating properties of gate dielectric. In other words, the occurrence of a dielectric breakdown event corresponds to a substantial decrease in the resistances of some of the local areas in gate dielectric. As reported in all the related studies, the occurrence of a dielectric breakdown event is often associated with a sudden, abrupt increase in gate leakage current I_g . The overall process of stressing the gate oxide under a constant voltage or current and continuously creating electronics traps within the gate oxide until the occurrence of a dielectric breakdown event is known as time-dependent dielectric breakdown (TDDB) [21].

The effect of dielectric breakdown on device performance is generally described by two major types of breakdown modes, namely soft and hard breakdowns [22, 23]. The situation where the device is partially failed but it still can operate with a degraded performance is named as a soft breakdown. In contrast, the case where an abrupt jump of leakage current leads to a totally failed device is addressed as a hard breakdown. Though a universally accepted criterion to define soft and hard breakdown events still remains ambiguous, one of the common methods to differentiate soft and hard breakdown is based on the magnitude of post-breakdown conduction.

2.3.1.1 SiO₂ Gate Dielectric

Silicon dioxide (SiO₂) has been intensively used as a tunnel oxide in floating gate flash memories to insulate the channel from the floating gate. With a large energy

bandgap of 8.9 eV, SiO_2 is considered as an excellent insulator and tunnel barrier for flash memories. Furthermore, the interface between Si and SiO_2 is very stable resulting in low electronic states at the Si- SiO_2 interface and good mobility of electrons and holes.

Since SiO_2 is widely used not only in the floating gate flash memories, but also as a gate oxide in the MOSFETs, the dielectric breakdown phenomenon and defects in SiO_2 have been analyzed with several physical mechanisms. The thermo-chemical, anode hole injection and hydrogen release models are some of the well-known gate dielectric intrinsic breakdown models to address the defect generation in SiO_2 . When the defect density in the oxide reaches a critical threshold, a percolation path will be formed within the gate oxide which is described as the percolation model [24, 25]. Fig. 2.7 illustrates the general schematic of a percolation path formed in the gate oxide due to the accumulation of generated defects/traps. The cathode and anode are electrically connected allowing huge amount of leakage current to surge through the percolation path. According to J. H. Stathis [25], the percolation model with a non-uniform distribution of defective sites may be a more appropriate model for gate dielectric breakdown, because defects are preferentially generated near the substrate and the gate electrode.

Figure 2.7. Schematic of a percolation path formed in the gate dielectric due to the accumulation of defects or electronic traps [26].

The statistics of gate dielectric breakdown are usually described using the Weibull distribution [27], which is given by

$$F(x) = 1 - \exp \left[- \left(\frac{x}{\eta} \right)^\beta \right] \quad (2.7)$$

where β is known as the slope parameter. The Weibull distribution is an extreme-value distribution in $\ln(x)$ and is appropriate for problems involving a weakest link. From elementary statistics, if the probability for any one unit fails is P , then the probability for any one of X independent units fails is

$$F = 1 - (1 - P)^X \quad (2.8)$$

which can be reorganized as

$$W = \ln[-\ln(1 - F)] = \ln(X) + \ln(-\ln(1 - P)) \quad (2.9)$$

W is known as the Weibit. Note that β is an important parameter in the evaluation of the gate dielectric reliability. One of the advantages of the percolation model is to establish a direct correlation between β and the critical number of defects N required in triggering the occurrence of a dielectric breakdown event, following the relationship [28]

$$\beta = \alpha \cdot N \quad (2.10)$$

where α is the power law exponent of SILC bulk trap generation and $N = t_{ox}/a_0$, in which t_{ox} is the oxide thickness and a_0 is the size of the defect (trap).

In 2008, our research group has unveiled the chemical nature of the percolation path in gate dielectric SiO_2 [29, 30]. The site-specific chemical nature of dielectric breakdown path was studied by STEM and EELS analysis and it was found that the physical defects in the percolation path are related to oxygen deficiencies, and they are radially distributed from the center of the breakdown path to surrounding areas. As shown in Fig. 2.8, the radial distribution of defects in the percolation path is of diameter 30-55 nm depending on the dielectric breakdown hardness and the Si–O composition changes from $\text{SiO}_{1.76}$ to $\text{SiO}_{x \rightarrow 0}$ as moving to the center of the percolation path. Therefore, it is believed that the oxygen deficiency is the key signature for the structural change at molecular level in the breakdown path. Chemical bond breakage and the local Joule heating due to large current surging through the percolation path are believed to be the main driving forces leading to the oxygen dissociation and washed-out.

Figure 2.8. The oxygen deficiency A-A' line profile in the percolation path of the dielectric breakdown samples with various breakdown hardness. The inserted diagram shows the A-A' cross section of a percolation in a breakdown oxide [30].

2.3.1.2 High- κ Gate Dielectric

With the motivations of stronger control gate coupling and reduction in operating voltage, the high- κ materials are introduced as a control dielectric in the flash memories. The breakdown characteristics of the NC-embedded high- κ are not fully understood, but are known to be inferior to the thermally grown SiO_2 [31].

The reliability study of high- κ gate dielectric receives more research attentions in the application of gate oxide in the sub-65 technology node of MOSFETs. McPherson *et al.*[32] reported that the ultimate breakdown strength (E_{bd}) of a dielectric material decreases with increasing dielectric constant κ and has a relationship of E_{bd} proportional

to $k^{-1/2}$. It is suggested that a very high local electric field in the high- κ dielectrics tends to distort/weaken the polar molecular bonds, making them more susceptible to bond breakage by standard Boltzmann processes and/or by hole-capture, resulting in lower E_{bd} [32]. It has been reported [33] that for a similar physical thickness of SiO₂ and high- κ dielectric, Weibull slope β in high- κ (HfO₂) is smaller than that in SiO₂ indicating reliability of the gate dielectric (the larger the β value, the better reliability). Moreover, the breakdown in high- κ is intrinsic and can be explained by the percolation model as well [34].

2.3.1.3 Dielectric Relaxation Current

Dielectric relaxation current (I_{relax}) measurement is used as one of the approaches to access gate dielectric integrity, especially in the high- κ dielectric [35-38]. It is a transient displacement current after the sudden removal of an applied constant bias on the gate and follows the direction of dV/dt . Due to the low conductivity of dielectric film, dielectric relaxation is a slow process in which I_{relax} decays with time following the Curie-von Schweidler law [39]

$$I_{relax} = at^{-n} \quad (2.11)$$

where n is the slope close to 1 in the log-log scale.

Fig. 2.9 compares the decay of the relaxation current with time of various high- κ dielectric films with respect to SiO₂ [38]. The relaxation current of high- κ films is found to be one to two orders of magnitudes larger than that of SiO₂. It is consistent with the

atomic configuration of high- κ dielectrics, in which the oscillating dipoles are easily created in high- κ dielectrics [37] due to lack of symmetry in metal-oxygen bonds. As a result, I_{relax} is easily detectable in the high- κ dielectric comparing with SiO_2 .

Figure 2.9. Comparison of relaxation current decay in various high- κ dielectric films and SiO_2 [38].

Reisinger *et al.* and Jameson *et al.* proposed that the relaxation current is due to the dielectric material polarization/relaxation [38, 40] induced by carrier hopping in double potential well [41]. Wolters *et al.* [42], Dumin *et al.* [43], and Bachhofer *et al.* [44] attributed the dielectric relaxation observed in SiO_2 and Si_3N_4 to charge trapping and detrapping in the gate dielectric. Z. Xu and coworkers [45] attributed the relaxation current of $\text{SiO}_2/\text{Al}_2\text{O}_3$ and $\text{SiO}_2/\text{HfO}_2$ gate stacks to the electron trapping and detrapping in the high- κ dielectric stacks. Overall, the physical nature of dielectric relaxation current

can be explained by the polarization/relaxation and/or charge trapping/detrapping in the gate dielectric. According to dielectric polarization/relaxation model [37, 46], an applied external bias across the dielectric film separates the bound charges, resulting in polarization and a compensating internal field. Once the external bias is removed, the internal bound charges are neutralized by the hopping of mobile charges while the polarization and internal field still remain in the dielectric film, leading to I_{relax} decaying with time. In the charge trapping/detrapping model [45], an application of an external bias causes the electrons to tunnel into the traps at different spatial locations or energy levels from the cathode. These traps can be located in the high- κ film and/or near the interface of SiO_2 and high- κ . When the bias is switched off, electrons emit out of the dielectric stacks generating a power law relaxation current. The mechanism of dielectric polarization/relaxation and charge trapping/detrapping is difficult to distinguish [41] because they might not be physically distinctive. As described by Anderson [47], the trapped charges in the dielectric are trapped in pairs presumably in the negative-U centers due to the lack to spin resonance. Then, the motions of electron pairs cause the atomic re-arrangement with an appropriately chosen local dipole. Jameson [41] describes the atomic re-arrangements as a particle in a double well, with the two wells representing the configurations of the atoms before and after the rearrangement (Fig. 2.10).

Figure 2.10. Illustration of the atomic re-arrangements as a particle in a double potential well model. (a) A particle in a double-well potential $V(x)$, (b) The particle “fall” downstream by classical mechanics in the presence of a strong electric field ζ eliminating the barrier. Also, the particle can still tunnel to adjacent well by (c) quantum-mechanical tunneling if the electric field is not strong enough to remove the potential barrier and (d) by thermally hopping over the barrier [41].

In the double well model, when electric field ζ distorts the double well in Fig. 2.10(a) by a term $-e\zeta x$, the “upstream” well is shifted-up by energy $\mu\epsilon/2$ and the “downstream” well is reduced by energy $\mu\epsilon/2$. The barrier disappears when the field is sufficiently large (Fig. 2.10(b)) causing the particle to “fall” by classical mechanics. When the field is not strong enough to remove the barrier, the particle still can tunnel into downstream by quantum-mechanical tunneling (Fig. 2.10(c)) and by thermally hopping over the barrier (Fig. 2.10(d)). Upon the removal of bias, the double wells revert to their original shapes and the particles that tunneled downstream now tunnel back to their original wells, yielding the current in opposite direction. According to the double well model, the current due to relaxation a single dielectric layer can be described as [41]

$$J_{relax} = 2\sigma_0 \frac{V}{t_{ox}} \left(3 + \ln \frac{t}{t_0} \right) \frac{t_0}{t}, t > t_0 \quad (2.12)$$

where t_{ox} is the dielectric thickness, V is an applied external bias while t_0 and σ_0 are material constants, respectively.

Figure 2.11. Relaxation current of Ru-NC and ITO-NC embedded ZrHfO films and corresponding control sample as reported in Ref. [48].

The dielectric relaxation current approach is mostly used to detect dielectric integrity of pure high- κ film. Some of the researchers, such as Yang *et al.* [48] and Chen and coworkers [49], extended the application of dielectric relaxation current method to investigate the failure process of NC embedded high- κ dielectric. It is reported that initial I_{relax} in the NC embedded high- κ film is larger than the pure high- κ film (Fig. 2.11). Also, the decay rate of I_{relax} strongly depends on the embedded NC material suggesting the charges can be trapped in the NCs deeply or loosely depending on the material properties. Moreover, it takes longer for the NC embedded film's I_{relax} to reach the final value than the control sample does. It is believed that the extra I_{relax} must be attributed to the embedded NCs. However, the high I_{relax} of the NC-embedded film is not directly related

to the amount of charges stored in the NCs because their I_{relax} decay rates are not related to the initial I_{relax} 's [48].

2.4 Summary

This chapter reviewed the current understanding and challenges associated with the NC-based memory and gate dielectric reliability. Choosing metal as a NC material, rather than semiconductor counter-part, provides a superior device performance due to little or no quantum confinement, better size scaling, higher work function and improved coupling to channel. The requirement of large NC density is inevitable in the NC-based memory. It points out the importance of study on lateral charge diffusion among NCs and optimized device structures for future scaling. Furthermore, we have reviewed the reliability of gate dielectric material mostly used in the NC-based memories. However, it is obvious that there is still lack of studies on the dielectric degradation mechanisms of MNC based dual layer (high- κ /SiO₂) gate stacks.

References

- [1] S. Tiwari, F. Rana, K. Chan, H. Hanafi, C. Wei, and D. Buchanan, "Volatile and non-volatile memories in silicon with nano-crystal storage," in *IEEE Tech. Dig. IEDM 1995*, pp. 521-524.
- [2] Z. Liu, C. Lee, V. Narayanan, G. Pei, and E. C. Kan, "Metal nanocrystal memories. I. Device design and fabrication," *IEEE Trans. Electron Devices*, vol. 49, pp. 1606-1613, 2002.
- [3] B. V. Zeghbroeck, *Principles of Semiconductor Devices*, Colorado University, Boulder, USA, 2004.
- [4] M.-P. Lu and M.-J. Chen, "Oxide-trap-enhanced Coulomb energy in a metal-oxide-semiconductor system," *Physical Review B*, vol. 72, p. 235417, 2005.
- [5] T. H. Hou, C. Lee, V. Narayanan, U. Ganguly, and E. C. Kan, "Design Optimization of Metal Nanocrystal Memory - Part I: Nanocrystal Array Engineering," *IEEE Trans. Electron Devices*, vol. 53, pp. 3095-3102, 2006.
- [6] W. Guan, S. Long, M. Liu, Q. Liu, Y. Hu, Z. Li, and R. Jia, "Modeling of retention characteristics for metal and semiconductor nanocrystal memories," *Solid-State Electronics*, vol. 51, pp. 806-811, 2007.
- [7] R. Kubo, "Electronic Properties of Metallic Fine Particles. I.," *Journal of the Physical Society of Japan*, vol. 17, pp. 975-986, 1962.
- [8] R. A. Rao, M. A. Sadd, R. F. Steimle, C. T. Swift, H. Gasquet, and M. Stoker, "Silicon Nanocrystal Nonvolatile Memory," in *Nanotechnology for Electronics Materials and Devices*, edited by A. Korkin, J. Labanowski, E. Gusev, and S. Luryi, (Springer, New York, USA, 2007), pp. 171-195.
- [9] L. Perniola, B. D. Salvo, G. Ghibaudo, A. F. Para, G. Pananakakis, V. Vidal, T. Baron, and S. A. Lombardo, "Modeling of the programming window distribution in multnanocrystals memories," *IEEE Trans. Nanotechnology*, vol. 2, pp. 277-284, 2003.
- [10] R. Gusmeroli, C. M. Compagnoni, and A. S. Spinelli, "Statistical constraints in nanocrystal memory scaling," *Microelectronic Engineering*, vol. 84, pp. 2869-2874, 2007.

- [11] R. Muralidhar, R. F. Steimle, M. Sadd, R. Rao, C. T. Swift, E. J. Prinz, J. Yater, L. Grieve, K. Harber, B. Hradsky, S. Straub, B. Acred, W. Paulson, W. Chen, L. Parker, S. G. H. Anderson, M. Rossow, T. Merchant, M. Paransky, T. Huynh, D. Hadad, C. Ko-Min, and B. E. White, Jr., "A 6 V embedded 90 nm silicon nanocrystal nonvolatile memory," in *IEEE Tech. Dig. IEDM 2003*, pp. 26.2.1-26.2.4.
- [12] R. J. Walters, "Silicon Nanocrystals for Silicon Photonics", Ph.D. dissertation, California Institute of Technology, 2007.
- [13] E. A. Boer, L. D. Bell, M. L. Brongersma, H. A. Atwater, M. L. Ostraat, and R. C. Flagan, "Charging of single Si nanocrystals by atomic force microscopy", *Applied Physics Letters*, vol. 78, p. 3133, 2001.
- [14] E. A. Boer, M. L. Brongersma, H. A. Atwater, R. C. Flagan, and L. D. Bell, "Localized charge injection in SiO₂ films containing silicon nanocrystals", *Applied Physics Letters*, vol. 79, p. 791, 2001.
- [15] D. M. Schaadt, E. T. Yu, S. Sankar, and A. E. Berkowitz, "Charge storage in Co nanoclusters embedded in SiO₂ by scanning force microscopy," *Applied Physics Letters*, vol. 74, pp. 472-474, 1999.
- [16] C. Y. Ng, T. P. Chen, H. W. Lau, Y. Liu, M. S. Tse, O. K. Tan, and V. S. W. Lim, "Visualizing charge transport in silicon nanocrystals embedded in SiO₂ films with electrostatic force microscopy", *Applied Physics Letters*, vol. 85, p. 2941, 2004.
- [17] C. Guillemot, P. Budau, J. Chevrier, F. Marchi, F. Comin, C. Alandi, F. Bertin, N. Buffet, C. Wyon, and P. Mur, "Imaging of stored charges in Si quantum dots by tapping and electrostatic force microscopy", *Europhys. Letters*, vol. 59 (4), pp. 566–571, 2002.
- [18] C. Y. Ng, T. P. Chen, M. S. Tse, V. S. W. Lim, S. Fung, and Ampere A. Tseng, "Influence of silicon-nanocrystal distribution in SiO₂ matrix on charge injection and charge decay", *Letters*, vol. 86, p. 152110, 2005.
- [19] N. Shimizu, M. Ikeda, E. Yoshida, H. Murakami, S. Miyazaki and M. Hirose, "Charging States of Si Quantum Dots as Detected by AFM/Kelvin Probe Technique", *Japanese Journal of Applied Physics*, vol.39, p. 2318, 2000.

- [20] Q. Yu, Y. Liu, T. P. Chen, Y. F. Yu, J. I. Wong, and M. Yang, "Modeling of lateral charge transfer in Si nanocrystals in SiO₂ thin film", *Journal of Applied Physics*, vol. 111, p. 073707, 2012.
- [21] D. L. Crook, "Method of Determining Reliability Screens for Time Dependent Dielectric Breakdown," *Proceedings of the Intl. Reliab. Phys. Symp.*, 1979, pp. 1-7.
- [22] J. Sune, G. Mura, and E. Miranda, "Are soft breakdown and hard breakdown of ultrathin gate oxides actually different failure mechanisms?," *Electron Device Letters*, vol. 21, pp. 167-169, 2000.
- [23] T. Pompl, C. Engel, H. Wurzer, and M. Kerber, "Soft breakdown and hard breakdown in ultra-thin oxides," *Microelectronics Reliability*, vol. 41, pp. 543-551, 2001.
- [24] R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, "A consistent model for the thickness dependence of intrinsic breakdown in ultrathin oxides ", in *IEEE Tech. Dig. IEDM 1995*, pp. 863-866.
- [25] J. H. Stathis, "Percolation models for gate oxide breakdown," *Journal of Applied Physics*, vol. 86, pp. 5757-5766, 1999.
- [26] V. L. Lo, K. L. Pey, C. H. Tung, and X. Li, "Multiple Digital Breakdowns and Its Consequence on Ultrathin Gate Dielectrics Reliability Prediction," in *IEEE Tech. Dig. IEDM 2007*, pp. 497-500.
- [27] E. Y. Wu, E. J. Nowak, R. P. Vollertsen, and L. K. Han, "Weibull breakdown characteristics and oxide thickness uniformity," *IEEE Trans. Electron Devices*, vol. 47, pp. 2301-2309, 2000.
- [28] A. T. Krishnan and P. E. Nicollian, "Analytic Extension of the Cell-Based Oxide Breakdown Model to Full Percolation and its Implications," *Proceedings of the Intl. Reliab. Phys. Symp.*, 2007, pp. 232-239.
- [29] X. Li, C. H. Tung, and K. L. Pey, "The radial distribution of defects in a percolation path," *Applied Physics Letters*, vol. 93, p. 262902, 2008.
- [30] X. Li, C. H. Tung, K. L. Pey, and V. L. Lo, "The physical origin of random telegraph noise after dielectric breakdown," *Applied Physics Letters*, vol. 94, p. 132904, 2009.

- [31] C. Sire, S. Blonkowski, M. J. Gordon, and T. Baron, "Statistics of electrical breakdown field in HfO₂ and SiO₂ films from millimeter to nanometer length scales," *Applied Physics Letters*, vol. 91, p. 242905, 2007.
- [32] J. McPherson, J. Kim, A. Shanware, H. Mogul, and J. Rodriguez, "Proposed universal relationship between dielectric breakdown and dielectric constant," in *IEEE Tech. Dig. IEDM 2002*, pp. 633-636.
- [33] Y. H. Kim, K. Onishi, C. S. Kang, H. J. Choi, R. Neih, S. Gopalan, R. Choi, J. Han, S. Krishnan, and J. C. Lee, "Area dependence of TDDDB characteristics for HfO₂ gate dielectrics," *Electron Device Letters*, vol. 23, pp. 594-596, 2002.
- [34] Y. H. Kim, K. Onishi, C. S. Kang, H. J. Cho, R. Choi, S. Krishnan, M. S. Akbar, and J. C. Lee, "Thickness dependence of Weibull slopes of HfO₂ gate dielectrics," *Electron Device Letters*, vol. 24, pp. 40-42, 2003.
- [35] L. Wen, Y. Tao, K. Yue, L. Jiang, Y. Jiong, and K. Way, "Charge trapping and dielectric relaxation in connection with breakdown of high- κ gate dielectric stacks," *Applied Physics Letters*, vol. 88, p. 202904, 2006.
- [36] R. Wan, J. Yan, Y. Kuo, and J. Lu, "Dielectric Breakdown and Charge Trapping of Ultrathin ZrHfO/SiON High- κ Gate Stacks," *Japanese Journal of Applied Physics*, vol. 47, pp. 1639-1641, 2008.
- [37] L. Wen, K. Yue, and K. Way, "Dielectric relaxation and breakdown detection of doped tantalum oxide high-k thin films," *Device and Materials Reliability, IEEE Transactions on*, vol. 4, pp. 488-494, 2004.
- [38] H. Reisinger, G. Steinlesberger, S. Jakschik, M. Gutsche, T. Hecht, M. Leonhard, U. Schroder, H. Seidl, and D. Schumann, "A comparative study of dielectric relaxation losses in alternative dielectrics," in *IEEE Tech. Dig. IEDM 2001*, pp. 12.2.1-12.2.4.
- [39] A. K. Jonscher, "Dielectric relaxation in solids," *Journal of Physics D: Applied Physics*, vol. 32, p. R57, 1999.
- [40] J. R. Jameson, P. B. Griffin, A. Agah, J. D. Plummer, H. S. Kim, D. V. Taylor, P. C. McIntyre, and W. A. Harrison, "Problems with metal-oxide high- κ dielectrics due to $1/t$ dielectric relaxation current in amorphous materials," in *IEEE Tech. Dig. IEDM 2003*, pp. 4.3.1-4.3.4.

- [41] J. R. Jameson, W. Harrison, P. B. Griffin, and J. D. Plummer, "Double-well model of dielectric relaxation current," *Applied Physics Letters*, vol. 84, pp. 3489-3491, 2004.
- [42] D. R. Wolters and J. J. v. d. Schoot, "Kinetics of charge trapping in dielectrics," *Journal of Applied Physics*, vol. 58, pp. 831-837, 1985.
- [43] D. J. Dumin and J. R. Maddux, "Correlation of stress-induced leakage current in thin oxides with trap generation inside the oxides," *IEEE Trans. Electron Devices*, vol. 40, pp. 986-993, 1993.
- [44] H. Bachhofer, H. Reisinger, E. Bertagnolli, and H. v. Philipsborn, "Transient conduction in multielectric silicon--oxide--nitride--oxide semiconductor structures," *Journal of Applied Physics*, vol. 89, pp. 2791-2800, 2001.
- [45] Z. Xu, L. Pantisano, A. Kerber, R. Degraeve, E. Cartier, S. D. Gendt, M. Heyns, and G. Groeseneken, "A study of relaxation current in high- κ dielectric stacks," *IEEE Trans. Electron Devices*, vol. 51, pp. 402-408, 2004.
- [46] M. Schumacher and R. Waser, "Curie - Von Schweidler behaviour observed in ferroelectric thin films and comparison to superparaelectric thin film materials," *Integrated Ferroelectrics*, vol. 22, pp. 109-121, 1998.
- [47] P. W. Anderson, "Model for the Electronic Structure of Amorphous Semiconductors," *Physical Review Letters*, vol. 34, pp. 953-955, 1975.
- [48] C. H. Yang, Y. Kuo, R. Wan, C. H. Lin, and W. Kuo, "Failure analysis of nanocrystals embedded high- κ dielectrics for nonvolatile memories," *Proceedings of the Intl. Reliab. Phys. Symp.*, 2008, pp. 46-49.
- [49] Y. N. Chen, K. L. Pey, K. E. J. Goh, Z. Z. Lwin, P. Singh, and S. Mahapatra, "Study of the Charge Leakage of Dual Layer Pt Metal Nanocrystal-based high- κ /SiO₂ Flash Memory Cell - A Relaxation Current Point of View " in *18th International Symposium on the Physical and Failure Analysis on integrated circuits (IPFA)*, 2011, pp. 38-41.

CHAPTER THREE

EXPERIMENTAL PROCEDURES AND DETAILS

3.1 Introduction

This chapter discusses the details of electrical and nanoscale characterization approaches and equipments used in the study of degradation and reliability in MNC-embedded high- κ /SiO₂ gate stacks. The various stressing methodologies to induce the dielectric breakdown in NC-based gate stacks are thoroughly described. These methodologies allow us to understand the characteristics of MNC in the different stages of progressive dielectric breakdown. Other reliability study tools such as dielectric relaxation approach and capacitance-voltage technique are described. The nanoscale characterization using Kelvin Force Microscopy as compared to Electrostatic Force Microscopy technique is presented as a key platform to examine the real time evolution of localized charges trapped in the MNCs.

3.2 Sample Information

The metal (Au/Pt) NCs-embedded in the MOS capacitors with $\text{Al}_2\text{O}_3/\text{SiO}_2$ dual layer dielectric are used in this study. In this work, the Al_2O_3 dielectric is used though its dielectric constant ($\kappa \sim 9$) is relatively smaller than other high- κ materials (HfO_2 : 25, TiO_2 : 80, *etc.*). This is because Al_2O_3 provides the largest bandgap and the best thermal stability [1], which are important factors to consider for the application of high- κ as a control dielectric in NC-based memory devices. The device fabrication process started with surface cleaning of Si wafer. Then, the tunneling oxide was grown thermally using *in-situ* steam generation (ISSG) process in an Applied Materials Rapid Thermal Process (RTP) chamber. The MNC formation was realized by depositing a thin film of metal using DC sputtering in an Applied Materials' Endura Sputtering system. The surface energy minimization leads to balling up of deposited metal thin film into NCs. Then the sample was annealed in an N_2 ambient at high temperatures in the RTP chamber and discrete NCs were formed on the tunneling oxide layer. Following the MNC formation, an optimized Al_2O_3 film was deposited using pulse-DC sputtering. Then post-deposition-anneal was performed to improve the high- κ quality. The detailed fabrication procedures can be found in Ref. [2]. For electrical characterization, an array of top electrodes with a diameter of 150 μm was formed by 100-nm-thick Pt or Au/Ti deposited using electron beam evaporator. The overall process flow of MNC embedded gate stack is illustrated in Fig. 3.1.

Figure 3.1. Schematic process flow of MNC-embedded gate stacks used in this study.

In dual-layer MNC samples, Al₂O₃ was deposited as an inter-layer dielectric (ILD) after formation of the first MNC-layer by metal thin film deposition and annealing. Then, another thin metal film was deposited and annealed to form the second MNC layer. Following the MNC formation steps, Al₂O₃ was deposited as a control dielectric using reactive sputter deposition technique. The fabrication of high- κ control sample was realized by depositing a pure high- κ layer of 10 nm Al₂O₃ with a 1 nm interfacial layer

(IL) SiO_2 on the Si substrate. The schematic cross-sections of high- κ control sample, single- and dual-layer MNC embedded gate stacks used in this work are illustrated in Fig. 3.2.

Figure 3.2. Schematic cross-sections (not-to-scale) of (a) control high- κ , and (b) single- and (c) dual-layer MNC-embedded gate stacks.

3.3. Electrical Characterization

As shown in Fig. 3.3, a 200 mm Suss MicroTec PM8 wafer probe station and a Keithley 4200-SCS semiconductor characterization system were used for micrometer-scale reliability characterization. The Suss MircoTec probe station equipped with a thermo-chuck allows the gate dielectric stressing and all the related electrical measurements of CMOS and MOSFET devices. The four source-monitor units (SMUs) with a current resolution of 1 fA from Keithley 4200-SCS were connected to four probes of the probe station, respectively, via Kelvin tri-axial cables with low leakage current.

These probe heads of probe station steadily held the tungsten probes which could separately contact to the device terminals.

For the low temperature electrical characterization in the NC-embedded devices, a cryogenic chamber with a CTI-Cryogenics 8200 compressor was used which enabled varying the temperature from the room temperature down to $\sim 10\text{K}$.

(a)

(b)

Figure 3.3. (a) Suss MicroTec PM8 200 mm wafer probe station and (b) Keithley 4200 semiconductor parameter analyzer.

3.3.1 Dielectric Stressing Methodologies

The dielectric stressing methodologies play a vital role to control the evolution of progressive breakdown and to investigate the correlated dielectric failure mechanisms [3]. Among various stressing methodologies in the study of dielectric reliability, constant voltage stress (CVS) and successive CVS approaches were chosen to study the dielectric degradation in MNC-embedded devices.

3.3.1.1 Stressing Voltage Determination

In order to characterize dielectric failure in the reasonable time frame and to study the behaviors of leakage current (I_g) at different stages of the progressive breakdown, the determination of stressing gate voltage is crucial. A constant voltage to stress the gate dielectric must be lower than the dielectric breakdown voltage (V_{BD}) such that the dielectric breakdown event does not occur instantly. Therefore, before the gate oxide stressing was carried out in our experiments, V_{BD} was determined from I_g - V_g ramp test sweeping. V_g was increased from 0 V to a higher value with a voltage increment of 0.05-0.1V while substrate is grounded in MOS capacitor structure.

A typical example of I_g - V_g ramping test in NC-embedded gate stack is shown in Fig. 3.4. It can be seen that I_g increases with increasing V_g due to the quantum-mechanical tunneling [4] and FN tunneling [5] of the charge carriers through the gate oxide. The abrupt increase in I_g indicated the occurrence of a dielectric breakdown event, and the corresponding V_g was treated as V_{BD} . This procedure was repeated for about 10-20 samples to define average breakdown voltage V_{BDA} .

Figure 3.4. Typical example of I_g - V_g ramping test with a voltage increment of 0.05V in determining V_{BD} of Au-NC-embedded gate stack.

Figure 3.5. The relationship of a voltage acceleration factor λ versus V_g reported in Ref [6].

Knowing V_{BDA} from the sampling of V_{BD} , a constant voltage to stress the gate dielectric ($V_{g-stress}$) is determined using [6]

$$\ln(t_{BD}) = \ln(t_{BD}') + \lambda(V_{BDA} - V_{g-stress}) \quad (3.1)$$

where t_{BD} and t_{BD}' are the time-to-dielectric breakdown of gate oxide at $V_{g-stress}$ and V_{BDA} , respectively. λ is a voltage acceleration factor which can be estimated from the λ versus V_g relationship shown in Fig. 3.5, as reported in Ref. [6].

3.3.1.2 Constant Voltage Stress and Successive Constant Voltage Stress

Constant-voltage stress (CVS) is a widely used stressing methodology in the study of gate dielectric degradation and reliability [7-9]. In a standard CVS, a constant voltage is applied to the gate electrode of a device (*i.e.*, $V_g = V_{g-stress}$) until specific dielectric breakdown criteria are met, while the substrate terminal is grounded in the MOS capacitor structure. The typical evolution of I_g in the single layer Au-NCs embedded gate stacks using a standard CVS is shown in Fig. 3.6. Under CVS, I_g increases with time since the electronic traps (defects) are built-up in the gate dielectric by electron tunneling current, resulting in the generation of stress-induced-leakage-current (SILC) [10]. The time-dependent dielectric breakdown (TDDB) is the overall process of defect generation in the gate oxide and the occurrence of dielectric breakdown event [11].

Figure 3.6. Evolution of leakage current using a constant voltage stress ($V_{g\text{-stress}}=5.5\text{V}$) under substrate injection in the Au-NC embedded device.

Based on the standard CVS approach, a successive CVS, which is a current-limited stressing methodology, was developed [7]. Firstly, a suitable $V_{g\text{-stress}}$ with low gate leakage compliance current limit (I_{gl}) which is typically two-times of magnitude larger than the pre-stressed leakage current level, is applied until the leakage current reaches I_{gl} . Then, the stress is continued by increasing I_{gl} to higher values without changing the stressing gate voltage. This method is useful to study the evolution of different stages of breakdown, the associated microstructural damages, and the effect of the post-breakdown gate leakage current on the device performance. The evolution of I_g in a single layer Au-NCs embedded gate stacks using a successive CVS in accumulation mode is shown in Fig. 3.7.

Figure 3.7. Evolution of leakage current using a successive CVS ($V_{g-stress}=5.5V$) with the compliance current ranging from 100 nA to 100 μA in Au-NC embedded device.

For the current limited CVS, the current overshoot effect is expected to be minimum ($<0.1\%$ typically [12]). However, for a standard CVS without compliance current, it may have the current overshoot effect because the data point can only be collected at a minimum delay of 0.28 sec. The current overshoot effect in CVS process can be controlled by using a much lower stressing voltage than the breakdown voltage.

3.3.2 Relaxation Current Measurement

The leakage current in this work is the current leaking through the capacitor when a gate bias is applied, while the relaxation current (I_{relax}) is the current leaking through the

capacitor upon the sudden removal of an applied gate bias (V_{on}). In the I_{relax} measurement, V_{on} ranging from $\pm 3\text{V}$ to $\pm 5\text{V}$ (depending on the thickness of dielectric) was applied for a short period of 5 sec to achieve the detection of sufficient current and to avoid the defect generation in the oxide.

Fig. 3.8 shows semi-log and log-log scale plots of I_{relax} in the control sample and MNC-embedded samples. The first data point of I_{relax} can only be recorded at ~ 5 sec upon removal of V_{on} due to the feedback delay in the probe station. As discussed in Section 2.3 (*i.e.*, the Literature Review Section on I_{relax}), I_{relax} in SiO_2 is negligible compared to that in high- κ film since dipoles are easily created and charge trapping/detrapping can easily occur in the high- κ dielectric. Therefore, I_{relax} in the control sample (IL $\text{SiO}_2/\text{Al}_2\text{O}_3$ 10nm) is dominated by the Al_2O_3 high- κ dielectric. As shown in Fig. 3.8(a), the magnitude of I_{relax} in the MNC embedded sample is about 1 order of magnitude larger than that in the control sample. This suggests that the total relaxation current in MNC embedded gate stacks may comprises two components: $I_{\text{relax-HK}}$ and $I_{\text{relax-MNC}}$.

Form the log-log scale plot of I_{relax} in the control sample and samples with embedded MNCs (Fig. 3.8(b)), I_{relax} is in the direction of dV_g/dt and follows the Curie-von Schweidler law [13]:

$$I_{\text{relax}} = at^{-n} \quad (3.2)$$

where n is the slope in the log-log scale plot. In the control sample, the slope n is close to unity which is consistent with the previous reports of the relaxation current in the high- κ dielectric. However, it is noted that the slope n in the MNC samples deviates

considerably from unity. It is therefore reasonable and justified to postulate that current decay in the MNC-samples is “not” entirely due to relaxation effects “alone”.

(a)

(b)

Figure 3.8. (a) Semi-log and (b) log-log plots of I_{relax} at zero gate voltage after the removal of a 5 sec, +3MV/cm stress in Au-MNC embedded sample and control sample.

According to the potential well model [14], the current due to relaxation from a single dielectric layer is

$$J_{\text{relax}} = 2\sigma_0 \frac{V}{t_{ox}} \left(3 + \ln \frac{t}{t_0} \right) \frac{t_0}{t}, t > t_0 \quad (3.3)$$

where t_{ox} is the thickness of the dielectric, V is the applied external bias, while t_0 and σ_0 are material constants, respectively. When the MNCs are embedded in the $\text{SiO}_2/\text{Al}_2\text{O}_3$ gate stack, the embedded MNC could modify the localized field across the gate stack [15], and hence the overall relaxation current is believed to be the combined effects of material dependent relaxation and localized electric field from the embedded MNCs. Using I_{relax} as a reliability study tool in the MNC embedded gate stacks will be discussed details in Chapter Four and Five.

3.3.3 Capacitance-Voltage Measurement

The Capacitance-Voltage (C-V) technique is usually applied to measure the device performance of memory gate stacks, such as memory window, endurance and retention characteristics. A two terminal C-V analyzer was connected to the top electrode and substrate, via the chuck of the prober and drove the high-frequency (100kHz-1MHz) AC signals to the NC-embedded MOS capacitor. The program (P) and erase (E) states of NC devices can be extracted from the shift in the flatband voltage (V_{FB}) of CV curve. Fig. 3.9 shows a typical high frequency CV (HFCV) curve of a DL-Pt MNCs embedded device.

Figure 3.9. A sample CV curve showing Program (P) and Erase (E) characteristics of DL-MNCs devices. The P/E transients were obtained from the shift in the flatband voltage (V_{FB}) of the CV curves.

Since the basic memory performance (such as memory window, endurance and retention) of single- and dual-layer MNC-embedded devices used in this study have been investigated by P. K. Singh *et al.* [16-20], the C-V technique was applied here only to define the charge loss mechanisms in-conjunction with other reliability study tools which will be discussed in Chapter Six.

3.4 Nanoscale Characterization using Scanning Probe Microscopy

As shown in Fig. 3.10, a Cypher Scanning Probe Microscopy (SPM) from Asylum Research was used for nanoscale reliability characterization. The system is equipped with a Cypher microscope, a controller and a computer. The microscope is where the actual imaging takes place and it comprises five basic components: enclosure, chassis, camera, scanner and backpack. The controller includes necessary electronics for controlling the scan motion and for acquiring image data from microscope. The computer is the main interface for controlling microscope and for communication to controller.

The Cypher (SPM) system is capable of Atomic Force Microscopy (AFM), Electrostatic Force Microscopy (EFM) and Kelvin Force Microscopy (KFM) characterizations. The EFM and KFM techniques which are derived from AFM could provide nano-scale electrostatic measurements. Although EFM and KFM are based on the electrostatic forces measurement, their implementations, output signals, advantages and drawbacks are different [21]. The EFM and KFM approaches have been used in the previous reports of exploring the spreading of charges in the dielectric layers with or without nanocrystals [22-28]. This section will provide a comprehensive description of both techniques used in the MNC based gate stacks as a guide to understand the literature and recent development related to the measurement of local charge transport properties of devices based on NCs.

(a)

(b)

Figure 3.10. (a) Setup of Cypher Scanning Probe Microscopy with a controller and computer. (b) Front-view of the Cypher Microscopy with the enclosure door opens.

3.4.1 Atomic Force Microscopy

In 1986, the first Atomic Force Microscopy was invented by Binnig, Quate and Gerber to capture local information by scanning the sample surface using a mechanical probe. There are two scanning modes of AFM: (1) contact mode and (2) AC mode imaging. In the contact mode, the conductive tip which is attached to the cantilever-end lands on the sample lightly. This mode is typically used for charge injection into the sample. In the AC mode imaging, the tip is in the intermittent contact with the sample by oscillating the cantilever at or near the resonant frequency. As the tip approaches to the sample surface, it interacts with the sample through short range forces such as van der Waals forces. It provides additional interactions which change the resonance frequency of the tip, thereby oscillation amplitude and phase lag are varied. The cantilever deflection is measured from the reflection of an incident laser on the cantilever, which is collected by the photodiodes as shown in Fig. 3.11.

Figure 3.11. Schematic illustration of basic AFM setup from Ref. [29]

Since feedback loop is used in AFM to keep oscillation amplitude by varying the tip-to-sample distance using a piezoelectric crystal, the change in the piezoelectric signal with position generates the surface topography. Depending on the attractive and repulsive regions of van der Waals forces, AC mode imaging can be further classified into non-contact mode and intermittent-contact mode, respectively. The AC mode imaging in AFM is typically used for the scanning of topographic images.

For conducting electric measurements in nano-meter scale, the EFM and KFM techniques with a metallic probe are widely used. In this work, a commercial Pt-coated Si probe (Olympus AC240TM) was used to measure electrostatic force gradients (in EFM) or surface potentials (in KFM) (Fig. 3.12). It provides high electrical conductivity (~ 1.5 orders of magnitude higher than heavily doped Si probes) with low probe resistance ($\sim 350 \Omega$) for electric measurements and prevents oxidation on the probe surface [30]. The cantilever is $240 \mu\text{m}$ long and the tip attached to the exact end of the cantilever has a radius of $28 \pm 10 \text{ nm}$ [31] which could reveal the sample surface precisely both electrically and topographically.

Figure 3.12. A Pt-coated Si probe used in this work [31].

3.4.2 Electrostatic Force Microscopy

The AFM is used to scan the surface topography, while, the EFM technique allows imaging the electrical properties of surface, by oscillating the tip at distance beyond the short range van der Waals interactions so that only electrostatic forces remains. The EFM image is obtained using a two-pass technique. In the first pass, the surface topography is obtained in the intermittent contact mode using AFM setup. In the second pass, the tip is lifted up in a fixed distance z from the sample and scans along the surface topography obtained during first-pass. In the EFM, the tip is biased at a detection voltage (V_{EFM}) with respect to the substrate. The measurement records the shift in the cantilever resonance frequency Δf with respect to its nominal value f_0 and produces the phase shift of the cantilever oscillation ($\Delta\phi$) which represents the long-range electrostatic force gradient detected by the tip. $\Delta\phi$ can be related to Δf at the resonance frequency using $\Delta\phi/2Q_{\text{lever}} = \Delta f/f_0$ under the condition of $\Delta f < f_0/2Q_{\text{lever}}$ [21], where Q_{lever} is the cantilever quality factor. The basic relation between cantilever resonance frequency shift Δf and the electrostatic force gradient $\frac{\partial F}{\partial z}(z_0)$ detected by the tip can be expressed as [21]

$$\frac{\Delta f}{f_0} = -\frac{1}{2} \frac{\frac{\partial F}{\partial z}(z_0)}{k} \quad (3.4)$$

where k is the cantilever spring constant. Considering the two plate capacitor forms in between the tip and the sample with the potential energy $U = \frac{1}{2} CV^2$, the force gradient on the tip can be described as [21]

$$\frac{\partial F}{\partial z}(z_0) = \frac{1}{2} \frac{\partial^2 C}{\partial z^2} V_{ts}^2 \quad (3.5)$$

where V_{ts} is the difference in potential of tip and sample. If the tip-substrate work function difference leads to the surface potential V_s and the tip is biased at V_{EFM} during scanning, the force gradient is [21]

$$\frac{\partial F}{\partial z}(z_0) = \frac{1}{2} \frac{\partial^2 C}{\partial z^2} (V_{EFM} - V_s)^2 \quad (3.6)$$

When there is charge Q presented in the sample, the total force gradient on the tip is [21]

$$\frac{\partial F}{\partial z}(z_0) = \frac{1}{2} \frac{\partial^2 C}{\partial z^2} [(V_{EFM} - V_s)^2 - 2(V_{EFM} - V_s)V_Q + V_Q^2] \quad (3.7)$$

where the first term $(V_{EFM} - V_s)^2$ represents the capacitive force resulted from the tip voltage which always corresponds attractive force gradient because of the square term. The second term $(V_{EFM} - V_s)V_Q$ relates to the Coulombic force between charge Q presented in the sample and the charge at EFM tip which can be either attractive or repulsive, while, the third term V_Q^2 is the Coulombic force between the charge Q and its image charge which always results in attractive force gradient because of the square term. Depending on the net attractive or repulsive force gradients, the EFM image appears as either bright (attractive) or dark (repulsive) spot at the area where the charge Q presented in the sample.

Using EFM technique, the charges are injected into the Au-NC embedded sample with a conductive Pt tip to investigate evolution of trapped charges in the NCs. The charge decay measurements were carried out by varying the polarity of V_{EFM} at a fixed

distance ($z=50$ nm), as shown in Fig. 3.13. Firstly, negative charges were injected into the sample by a tip voltage of -6 V for 15 sec. The first EFM image was obtained with V_{EFM} of -1 V at 136 sec right after charge injection. The relative negative phase shift at the charge cloud area can be seen due to the repulsive force gradient. Then, a second EFM image was obtained with V_{EFM} of $+1$ V at 227 sec and the relative positive phase shift was found due to the attractive force gradient. It is interesting to note that the magnitude of the phase shift with $t=227$ sec is larger than that with $t=137$ sec. It shows that the detection of the charge cloud height (magnitude of phase shift) depends on the polarity of V_{EFM} during the EFM scanning. Since the phase shift is supposed to remain or decrease with time (due to the charge holding or decaying nature), the unusual increase in the magnitude of the phase shift with time under the positive polarity of V_{EFM} suggests the attractive force gradient tends to dominate in the EFM signal, which is in a good agreement with capacitive force equation (3.7) as discussed above.

Figure 3.13. Illustration of EFM images at various polarities of detection voltage V_{EFM} after charge injection (negative charge injection in this case). (a) The first EFM image taken with $V_{\text{EFM}} = -1\text{V}$ at $t = 136\text{ sec}$ upon charge injection. The darker spot appeared at the injected charge cloud area due to repulsive force gradient. (b) The subsequent EFM image taken with $V_{\text{EFM}} = +1\text{V}$ at $t = 227\text{ sec}$. The brighter spot appeared at the injected charge cloud area due to attractive force gradient. (c) The related profile of phase shift across the injected charge cloud.

Subsequently at $t=320$ sec, the EFM image was obtained at zero tip bias ($V_{\text{EFM}}=0\text{V}$) as shown in Fig. 3.14. It can be seen that a charge ring (halo image) appears at the injected charge cloud area. The experiment was repeated at various locations of the sample and it is noted that the charge ring always appeared at a later stage of charge decay (>300 sec) when the detection voltage is zero. When $V_{\text{EFM}}=0\text{V}$ and V_Q is negative, the second term in the capacitive force equation contributes to the repulsive force gradient which will appear as a darker spot in the EFM image. However, the brighter spot is observed at the center of the charge cloud (halo image in Fig. 3.14). At the center of the charge cloud, the charge density is the highest. As a result, $V_Q^2 > 2(V_{\text{EFM}}-V_s)V_Q$ and the attractive force could dominate the EFM signal resulting in the charge ring (halo image) in the EFM imaging. This is an important artifact to understand in the use of EFM technique in charge decay analysis. The appearance of halo image is not due to the change in the type of charge present in the sample or discharging; rather it owes to the detection approach used in the EFM system.

Figure 3.14. Appearance of halo image at injected charge cloud area when $V_{\text{EFM}} = 0\text{V}$. The image is taken at $t=320$ sec after a negative charge injection.

It is reported that EFM technique could provide single electron resolution [32, 33] under high vacuum and low temperature environments, which is the main advantage over KFM technique where the lateral resolution is known to be lower than EFM due to the side capacitance effect [21]. Our experimental results indicate that (1) the polarity of V_{EFM} plays an important role in the detection of charges presented in the sample and (2) the magnitude of V_{EFM} cannot be too small for the detection of charge decay. However, applying large bias during EFM scanning could result in interaction between the tip and the sample which may complicate the charge transport analysis in the NC based devices. Therefore, KFM technique is explored which will be discussed in the following section.

3.4.2 Kelvin Force Microscopy

The KFM technique is based on the electrostatic excitation of cantilever using a potential feedback loop to nullify the tip oscillation amplitude, while EFM technique is based on the mechanical excitation of cantilever using a detection voltage V_{EFM} . The basic operating principle of KFM can be understood as follows. Firstly, the probe is driven electrically with an AC bias. Any potential difference between the tip and the sample causes the probe to oscillate and these oscillations are then cancelled by a potential feedback loop using a built-in voltage generator. The voltage required to nullify the probe oscillation appears as surface potential map in KFM image.

The AC and DC components of potential applied on the tip result in an electrostatic force $F(z)$ acting on the cantilever which can be expressed as [34]:

$$F(z) = \frac{1}{2} \frac{\partial C}{\partial z} [V_{dc} + V_{ac} \sin(\omega t) - V_{sp}]^2 \quad (3.8)$$

where z is the tip-to-substrate distance, C is the tip-substrate capacitance and V_{sp} is the surface potential. The above equation can be re-arranged as

$$F(z) = \frac{1}{2} \frac{\partial C}{\partial z} \left[(V_{dc} - V_{sp})^2 + \frac{V_{ac}^2}{2} + 2(V_{dc} - V_{sp}) \times V_{ac} \sin \omega t - \left(\frac{V_{ac}^2}{2} \cos 2\omega t \right) \right] \quad (3.9)$$

From the equation (3.9), it can be seen that there are three force terms contributing to the net electrostatic force. The first force has no frequency dependent, thus static. The second and third forces oscillate at angular frequencies 1ω and 2ω , respectively. The three force terms can be expressed by:

$$\begin{aligned} F_{dc} &= \frac{1}{2} \frac{\partial C}{\partial z} \left[(V_{dc} - V_{sp})^2 + \frac{V_{ac}^2}{2} \right] \\ F_{\omega} &= \frac{\partial C}{\partial z} (V_{dc} - V_{sp}) \times V_{ac} \sin \omega t \\ F_{2\omega} &= -\frac{1}{4} \frac{\partial C}{\partial z} V_{ac}^2 \cos 2\omega t \end{aligned} \quad (3.10)$$

Since the spectral component of the force at the angular frequency ω is proportional to the amplitude of cantilever oscillation at ω , F_{dc} and $F_{2\omega}$ are out of cantilever resonance. The potential feedback loop in KFM system provides V_{dc} to make F_{ω} component zero and the feedback potential V_{dc} to nullify the cantilever oscillation amplitude corresponds to sample voltage ($V_{sp}=V_s+V_Q$). This generates the real time measurements of the local surface potentials on the sample. The data collected using KFM technique includes tip-substrate work function difference, trapped charge in the

sample and any permanent or applied voltage between the tip and the sample [34]. Therefore, the KFM technique is generally considered as a pseudo-quantitative technique. Though it directly gives a mapping of local surface potential, the number likely consists of superposition of more than one physical quantity. Similar to EFM approaches; KFM technique uses two-pass procedure: the first pass captures the surface topography and the second pass retraces the topography at a constant distance from the sample surface to record the surface potential. Due to the side capacitance of the conductive tip, the resolution of KFM depends on the distance between the tip and the sample. The tip-sample distance dependent KFM resolution will be discussed in Chapter Five.

Fig. 3.15 shows a KFM image and its related potential profile in the NC embedded device after a negative tip injection. Unlike in the EFM imaging, there is no external bias required to scan the surface potential minimizing charge interaction between the tip and the sample. More importantly, no halo artifact is observed in the KFM imaging. Furthermore, the potential profile of local charges can be directly measured using KFM approach. It efficiently simulates the charging and local charge transport in NC embedded devices, providing qualitative information on the charge holding capability of NC-embedded devices. Therefore, the KFM technique is applied in this work to investigate real time evolution of localized charge transport in single- and dual-layer MNC embedded devices which will be discussed details in Chapter Five and Six.

Figure 3.15. (a) Typical plan view KFM image at the lift height of 50 nm upon negative tip injection on the NC embedded device. (b) The potential profile across the injected charge cloud (A-A') in (a).

3.5 Summary

In summary, the details of the devices, the electrical characterization approaches and nanoscale characterization techniques are described in this chapter. Moreover, the equipment used for reliability characterization is introduced. The various electrical characterization techniques and dielectric stressing methodologies are discussed. The EFM and KFM techniques used for the study of localized charge transport in MNC-based gate stacks are discussed in details and their implementation, advantages and drawbacks are briefly described.

References

- [1] P. K. Singh, K. K. Singh, R. Hofmann, K. Armstrong, S. Mahapatra, and N. Krishna, "Aluminum Oxide thin films using Sputtering technique as blocking oxide for Flash memory applications," in *MRS Spring Symposium 2008* (San Francisco).
- [2] P. K. Singh, "Metal Nanocrystal Memory for NAND Flash Applications," Ph.D dissertation, Indian Institute of Technology, Bombay, 2009.
- [3] L. J. Tang, K. L. Pey, C. H. Tung, R. Ranjan, and W. H. Lin, "Study of breakdown in ultrathin gate dielectrics using constant voltage stress and successive constant voltage stress," *Microelectronic Engineering*, vol. 80, pp. 170-173, 2005.
- [4] S. H. Lo, D. A. Buchanan, Y. Taur, and W. Wang, "Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFET's," *IEEE Electron Device Letters*, vol. 18, pp. 209-211, 1997.
- [5] M. Lenzlinger and E. H. Snow, "Fowler-Nordheim Tunneling into Thermally Grown SiO₂," *Journal of Applied Physics*, vol. 40, pp. 278-283, 1969.
- [6] A. Berman, "Time-Zero Dielectric Reliability Test by a Ramp Method," *Proceedings of the Intl. Reliab. Phys. Symp.*, 1981, pp. 204-209.
- [7] B. P. Linder, J. H. Stathis, R. A. Wachnik, E. Wu, S. A. Cohen, A. Ray, and A. Vayshenker, "Gate oxide breakdown under Current Limited Constant Voltage Stress," *IEEE VLSI Technology Symposium*, 2000, pp. 214-215.
- [8] B. P. Linder, D. J. Frank, J. H. Stathis, and S. A. Cohen, "Transistor-limited constant voltage stress of gate dielectrics," *IEEE VLSI Technology Symposium*, 2001, pp. 93-94.
- [9] K. L. Pey, C. H. Tung, M. K. Radhakrisnan, L. J. Tang, Y. Sun, X. D. Wang, and W. H. Lin, "Correlation of failure mechanism of constant-current-stress and constant-voltage-stress breakdowns in ultrathin gate oxides of n-MOSFETs by TEM," *Microelectronics Reliability*, vol. 43, pp. 1471-1475, 2003.

- [10] D. J. DiMaria and E. Cartier, "Mechanism for stress-induced leakage currents in thin silicon dioxide films," *Journal of Applied Physics*, vol. 78, pp. 3883-3894, 1995.
- [11] D. L. Crook, "Method of Determining Reliability Screens for Time Dependent Dielectric Breakdown," *Proceedings of the Intl. Reliab. Phys. Symp.*, 1979, pp. 1-7.
- [12] Keithley Instruments, Inc. 4200-SCS Semiconductor Characterization System - QuickStart Manual, 2003. Document Number: 4200-903-01, Rev. A.
- [13] A. K. Jonscher, "Dielectric relaxation in solids," *Journal of Physics D: Applied Physics*, vol. 32, p. R57, 1999.
- [14] J. R. Jameson, W. Harrison, P. B. Griffin, and J. D. Plummer, "Double-well model of dielectric relaxation current," *Applied Physics Letters*, vol. 84, pp. 3489-3491, 2004.
- [15] C. Lee, U. Ganguly, V. Narayanan, T. Hou, J. Kim, and E. C. Kan, "Asymmetric electric field enhancement in nanocrystal memories," *IEEE Electron Device Letters*, vol. 26, pp. 879-881, 2005.
- [16] P. K. Singh, G. Bisht, K. Auluck, M. Sivatheja, R. Hofmann, K. K. Singh, and S. Mahapatra, "Performance and Reliability Study of Single-Layer and Dual-Layer Platinum Nanocrystal Flash Memory Devices Under NAND Operation," *IEEE Trans. Electron Devices*, vol. 57, pp. 1829-1837, 2010.
- [17] P. K. Singh, G. Bisht, R. Hofmann, K. Singh, N. Krishna, and S. Mahapatra, "Metal Nanocrystal Memory With Pt Single- and Dual-Layer NC With Low-Leakage Al_2O_3 Blocking Dielectric," *IEEE Electron Device Letters*, vol. 29, pp. 1389-1391, 2008.
- [18] P. K. Singh, G. Bisht, S. Mahapatra, R. Hofmann, and K. Singh, "Dual Layer Pt Metal Nanocrystal Flash for Multi-Level-Cell NAND Application," in *IEEE Intl. Memory Workshop*, 2009, pp. 1-4.
- [19] P. K. Singh, R. Hofmann, K. K. Singh, N. Krishna, and S. Mahapatra, "Performance and Reliability of Au and Pt Single-Layer Metal Nanocrystal Flash Memory Under NAND (FN/FN) Operation," *IEEE Trans. Electron Devices*, vol. 56, pp. 2065-2072, 2009.

- [20] P. K. Singh, K. K. Singh, R. Hofmann, K. Armstrong, N. Krishna, and S. Mahapatra, "Au nanocrystal flash memory reliability and failure analysis," *Proceedings of the International Symposium on the Physical and Failure Analysis, IPFA 2008*, pp. 1-5.
- [21] T. Melin, M. Zdrojek, and D. Brunel, in *Scanning Probe Microscopy in Nanoscience and Nanotechnology*, edited by B. Bhushan (Springer-Verlag, Heidelberg, 2010), pp. 89-125.
- [22] F. Marchi, R. Dianoux, H. J. H. Smilde, P. Mur, F. Comin, and J. Chevrier, "Characterisation of trapped electric charge carriers behaviour at nanometer scale by electrostatic force microscopy," *Journal of Electrostatics*, vol. 66, pp. 538-547, 2008.
- [23] D. M. Schaadt, E. T. Yu, S. Sankar, and A. E. Berkowitz, "Charge storage in Co nanoclusters embedded in SiO₂ by scanning force microscopy," *Applied Physics Letters*, vol. 74, pp. 472-474, 1999.
- [24] S. D. Tzeng and S. Gwo, "Charge trapping properties at silicon nitride/silicon oxide interface studied by variable-temperature electrostatic force microscopy," *Journal of Applied Physics*, vol. 100, p. 023711, 2006.
- [25] T. Mélin, H. Diesinger, D. Deresmes, and D. Stiévenard, "Electric force microscopy of individually charged nanoparticles on conductors: An analytical model for quantitative charge imaging," *Physical Review B*, vol. 69, p. 035321, 2004.
- [26] C. Guillemot, P. Budau, J. Chevrier, F. Marchi, F. Comin, C. Alandi, F. Bertin, N. Buffet, W. Ch, and P. Mur, "Imaging of stored charges in Si quantum dots by tapping and electrostatic force microscopy," *Europhysics Letters*, vol. 59, p. 566, 2002.
- [27] C. Dumas, L. Ressier, J. Grisolia, A. Arbouet, V. Paillard, G. BenAssayag, S. Schamm, and P. Normand, " KFM detection of charges injected by AFM into a thin SiO₂ layer containing Si nanocrystals " *Microelectronic Engineering*, vol. 85, pp. 2358-2361, 2008.

- [28] G. H. Buh, H. J. Chung, and Y. Kuk, "Real-time evolution of trapped charge in a SiO₂ layer: An electrostatic force microscopy study," *Applied Physics Letters*, vol. 79, pp. 2010-2012, 2001.
- [29] <http://www.chem.ucla.edu/dept/Faculty/gimzewski/nnano2.html>.
- [30] http://probe.olympus-global.com/en/product/omcl_ac240tm_w2/.
- [31] <http://www.asylumresearch.com/Probe/AC240TM,Olympus>.
- [32] T. D. Krauss and L. E. Brus, "Charge, Polarizability, and Photoionization of Single Semiconductor Nanocrystals," *Physical Review Letters*, vol. 83, pp. 4840-4843, 1999.
- [33] R. Stomp, Y. Miyahara, S. Schaer, Q. Sun, H. Guo, P. Grutter, S. Studenikin, P. Poole, and A. Sachrajda, "Detection of Single-Electron Charging in an Individual InAs Quantum Dot by Noncontact Atomic-Force Microscopy," *Physical Review Letters*, vol. 94, p. 056802, 2005.
- [34] K. Jones, "Asylum Research Kelvin Probe Microscopy User Guide, Rev. 662 " 2010.

CHAPTER FOUR

METAL NANOCRYSTALS IN DEGRADED DIELECTRIC GATE STACKS

4.1 Introduction

In the conventional nanocrystal (NC) memories, the dielectric gate stack usually contains a tunneling oxide and a NC-embedded high- κ layer in order to improve coupling between the storage nodes and the control gate. Also, the use of high- κ control dielectric in NC-memories can efficiently reduce the operating voltages and P/E speeds. Since the cumulative effect of the charge trapping/detrapping process in the dielectric during the program/erase cycling operation reduces the endurance capability [1], it is very crucial to understand the charging and discharging of NCs in the degraded dielectric gate stacks for future scaling. Many investigations have been carried out on the dielectric breakdown and retention characteristics of metal and semiconductor NCs [1-3]. However, there is a lack of systematic study focusing on understanding the behaviors of individual dielectric layer degradation in NC-embedded bi-layer gate stacks that is important for device optimization of NC-embedded memories.

In this chapter, the dielectric breakdown behaviors in MOS capacitor structures

with metal nanocrystals (MNCs) embedded in bi-layer gate stacks ($\text{SiO}_2/\text{Al}_2\text{O}_3$) are investigated. Using a unique stressing methodology that induces a breakdown path in one dielectric layer, the charging and discharging phenomenon of the MNCs and leakage mechanism in the degraded gate stacks are studied. The methodology to detect the dielectric breakdown layer in a MNC-embedded bi-layer gate stack is explored. The post-breakdown conduction mechanisms and non-localized nature of discharging from MNCs via the breakdown path are reported.

4.2 Experiments

The samples used in this work were Au-MNCs embedded in a bi-layer gate stack consists of SiO_2 (50Å) and Al_2O_3 (60Å) on an n-Si substrate with an As doping of $3 \times 10^{19} \text{ cm}^{-3}$. Fig. 4.1(a) shows a schematic of MNCs-embedded bi-layer gate stacks. From the plan view TEM micrograph shown in Fig. 4.1(b), the extracted density of NC is about $3 \times 10^{12} \text{ cm}^{-2}$. The average size of the NCs is around 3 nm, as displayed in Fig. 4.1(c).

Figure 4.1. (a) Schematic of the MNC-embedded bi-layer gate stack used in our study. (b) Plan view TEM micrograph of Au-NCs. (c) Cross-sectional view TEM micrograph showing existence of MNCs in the Al_2O_3 layer.

4.3 Charging and Discharging Characteristics of MNCs in Degraded Dielectric Gate Stacks

4.3.1 Methodology and Approach

First of all, the breakdown voltage (V_{BD}) was determined from the I_g - V_g ramping test with a voltage increment of 0.05 V, as shown in Fig. 4.2. The breakdown voltage was found to be 6-6.5V for Au-NC embedded $\text{SiO}_2/\text{Al}_2\text{O}_3$ gate stacks.

Figure 4.2. I_g - V_g ramping tests in the Au-NCs embedded samples to determine dielectric breakdown voltage (V_{BD}).

In order to achieve sequential breakdown of the individual dielectric layers in the NC-embedded SiO_2 /high- κ gate stacks, current limited CVS was applied [4]. The stressing was automatically halted when the gate (*i.e.*, top electrode) current, I_g , reached a current compliance, I_{gl} , and then, I_g - V_g measurements were carried out. This cycle was repeated by setting a higher I_{gl} without changing $V_{g\text{-stress}}$ to investigate the different stages of progressive breakdown with decreasing percolation resistance (*i.e.*, increased breakdown hardness). The use of a low I_{gl} followed by higher values of I_{gl} helps to “arrest” the progression in breakdown of one-layer and its influence on the degradation in the next layer.

Fig. 4.3(a) shows the evolution of I_g with a current limited CVS at $V_{g\text{-stress}}=5.5$ V in a substrate injection mode (at room temperature). I_{gl} was initially fixed at 100 nA. After I_g reached 100 nA (solid line), I_{gl} was increased to 500 nA (open circles) and the stressing was continued, while $V_{g\text{-stress}}$ remained constant at 5.5 V. The similar procedure was carried out for $I_{gl} = 1 \mu\text{A}$, $10 \mu\text{A}$ and 0.1 A , respectively. During the electrical stress test, I_g - V_g measurement was performed on the fresh sample, after one-layer breakdown and after two-layer breakdown, respectively, as illustrated in Fig. 4.3(b).

The one- and two-layer breakdowns are differentiated by comparing the increase in the leakage current with respect to the fresh I_g . Comparing with the fresh I_g , I_g after an one-layer breakdown increases by about 2-3 orders of magnitude, while I_g after a two-layer breakdown increases by more than 7 orders of magnitude. Thus, a low I_{gl} (typically ~2 orders of I_g that is induced by the pre- breakdown stress) was used to induce one-layer breakdown first and the post- breakdown I_g - V_g measurements were performed to differentiate one-layer breakdown from two-layer breakdown.

(a)

Figure 4.3. (a) Evolution of I_g using CVS, $V_{g\text{-stress}} = 5.5\text{V}$, in substrate injection mode with I_{gl} ranging from 100nA to 100mA. (b) I_g - V_g at different stages of progressive breakdown (BD).

4.3.2 One-Layer Breakdown & I_g Blockade

As shown in Figs. 4.4 and 4.5, two different types of I_g evolution were observed under CVS.

(a)

(b)

Figure 4.4. (a) Evolution of I_g using a CVS with $V_{g\text{-stress}} = 4.5\text{V}$ in substrate injection mode. (b) I_g - V_g before stress and after a 2-layer breakdown (BD), respectively.

(a)

(b)

Figure 4.5. A new 3-stage I_g evolution under CVS in an Au-NC sample denoted as Sample 1. (a) I_g under 5V CVS (substrate injection) with I_{gl} ranging from 50 to 500 nA. Stage 1 - I_g fluctuations, Stage 2 - rapid jump in I_g owing to a one-layer breakdown, and Stage 3 - I_g blockade are seen. (b) I_g - V_g trends at different stages of progressive breakdown and after I_g blockade, respectively. The arrows delineate I_g evolution during progressive breakdown and after I_g blockade (*i.e.*, large decrement of I_g).

Fig. 4.4(a) shows the leakage current evolution under CVS with $V_{g\text{-stress}}=4.5\text{V}$ in Au-NC embedded device. The current fluctuation can be seen during initial stage of stressing at $t < 200$ sec, where I_g increases to $I_{g\text{-max}}$ of 10 nA (*i.e.*, about 1.8 times of pre-stress I_g) and recovered to a level comparable to pre-stress state. This auto-recovery from high to low I_g in NC-embedded gate stacks is believed to be governed by Joule heating induced oxygen vacancies passivation [5, 6]. This is because the amount of oxygen vacancies in the percolation path increases with dielectric breakdown hardness [10] and the expelled oxygen travels in the form of ions which can be trapped in the metal/oxide interface: metal electrode/oxide and MNC/oxide [9]. When the local current density increases in the percolation path, a significant Joule heating could occur and some of the trapped oxygen ions passivate the oxygen vacancy defects by thermal diffusion. Since the auto-recovery behavior can be seen only at low $I_{g\text{-max}}$, it is believed that the oxygen ions can only passivate limited defects in the oxide which are near metal/oxide interface (*i.e.*, within the diffusion length of oxygen ions). At $t\sim 800$ sec, after an instantaneous, huge jump of I_g to about 2.5 times of pre-stress state (*i.e.*, the occurrence of dielectric breakdown event), I_g enters a high conduction state. The post- breakdown I_g - V_g shown in Fig. 4.4(b) suggests that a two-layer dielectric breakdown occurred.

Fig. 4.5 illustrates a new 3-stage I_g evolution, which is different from the usual I_g evolution during a progressive breakdown, as depicted in Fig. 4.4. In Stage 1, I_g fluctuation was seen (black solid circles) when I_{g1} was set to 50 nA. After I_{g1} was increased to 200 nA, I_g experienced a sudden, abrupt increase (blue solid triangles), denoted as Stage 2. Then, the experiment was halted and the post- breakdown I_g - V_g measurement was performed, as shown in Fig. 4.5(b). I_g increased by 2 orders of

magnitude in comparison with the pre-stress I_g , suggesting that the rapid jump in I_g in Stage 2 is owing to a one-layer breakdown. Then the stressing was continued by increasing I_{gl} (300 nA and 500 nA) and I_g increased with I_{gl} . Nevertheless, when the stressing was continued at $I_{gl} > 500$ nA, a subsequent abrupt decrease of I_g by ~ 2 orders of magnitude occurred and after which, I_g stayed at a very low level (green open circles) in Stage 3 for a prolonged time. This phenomenon in Stage 3 is named as “ I_g blockade”. Similar behaviors were observed in other Au (Fig. 4.6) and Ir (Fig. 4.7) NC-embedded gate stacks. As shown in Fig. 4.6, it is noted that the I_g blockade Samples 2 and 3 have different breakdown hardness (*i.e.*, I_{g-max} just before I_g blockade occurred) with 200 nA and 500 nA, respectively. It is noted that the “ I_g blockade” behavior is different from dielectric breakdown recovery since I_g at Stage 3 could be much lower than the pre-stress leakage current level (see Fig. 4.6(b) and Fig. 4.7).

(a)

(b)

Figure 4.6. 3-stage I_g evolution under CVS in Au NC-embedded samples. (a) Sample 2 with $I_{g-max} = 200$ nA under CVS of 4.8V. (b) Sample 3 with $I_{g-max} = 500$ nA under $V_{g-stress} = 5.5$ V. The insets show the same data in a logarithmic ordinate scale and the horizontal dotted lines are guides to eyes.

Figure 4.7. 3-stage I_g evolution in Ir-NC embedded gate stacks under CVS = 2.6V. The inset shows the same data in a logarithmic ordinate scale.

4.3.3 Proposed Model

As shown in Fig. 4.8, we propose a lateral charging model via the percolation (*i.e.*, breakdown) path formed in one layer of the $\text{SiO}_2/\text{Al}_2\text{O}_3$ stack to interpret the I_g blockade phenomenon. The I_g fluctuation observed in Stage 1 in Figs. 4.5-4.7 is closely related to charging and discharging of traps during the CVS stressing. When one layer is broken down in Stage 2, a localized electric field enhancement occurs at the breakdown path/NC interface, as shown in Fig. 4.9. This results in a lateral charging to the NCs nearby via the breakdown path so that all energy states in the affected NCs were occupied by the injected electrons from the substrate injection stress until no more electron injection into the NCs is possible. After the NCs are charged through the breakdown path, I_g decreases owing to the

increased electrostatic repulsive potential among the charged NCs and the Coulomb Blockade effect in these charged NCs.

Figure 4.8. Schematics of (a) defects generated in dielectric film during CVS, (b) lateral charging through a percolation path, (c) leakage components before occurrence of breakdown, $I_g = I_1 + I_2$ where I_1 is the direct/FN/trap assisted tunneling (TAT) current, and I_2 is the resonant tunneling current through NCs, and (d) after one-layer breakdown and lateral charging happened, I_1 decreases due to the increased electrostatic potential between the charged NCs. I_2 is eliminated by the Coulomb blockade effect. In this illustration, we assume that the breakdown path presents in the SiO₂ layer.

Figure 4.9. Simulation of localized electric-field enhancement at the breakdown path and SiO₂/high- κ interface. For simplicity, we ignored the presence of NCs.

Fig. 4.10 shows the schematic energy band diagrams of the proposed lateral charging model when SiO₂ is assumed to have been broken down. For simplicity, the energy band diagram of a charged NC (NC1) with an empty neighboring NC (NC2) is considered. When a positive bias is applied to the top electrode, electrons in the substrate can tunnel through the tunneling oxide and charge NC1. After a one-layer breakdown, electrons can flow through the breakdown path and then charge NC1 easily. The charging process will continue when electron gains energy that is larger than the charging energy [8],

$$E_c = e^2/C \quad (4.1)$$

where $C=4\pi\epsilon_0\epsilon_r R$ is the self-capacitance of the charged NC with radius R . After NC1 has been charged, the electrostatic potential of NC1 increases, and then the lateral charging to

the nearby NC2 could occur. However, after NCs are completely charged under the applied constant voltage, further charging or tunneling through the tunneling oxide is forbidden by the energy barrier E_c , leading to an I_g blockade. Upon the removal of bias, the charges stored in NCs could discharge through the breakdown path gradually, as illustrated in Fig. 4.10(b).

Figure 4.10. Schematic energy band diagrams of (a) a lateral charging process between a charged NC (NC1) and an uncharged neighboring NC (NC2) and (b) a discharging process through the breakdown path. The breakdown path is assumed to be in the SiO₂ layer.

4.3.4 Discharging of NCs in Degraded Dielectric

The I_g blockade phenomenon enables us to study charging and discharging characteristics of NCs and their lateral interaction through the breakdown path. The relaxation current, I_{relax} , due to charge trapping/detrapping and dielectric polarization (relaxation) [9-11], was measured after the removal of a stress (5 sec, $V_{\text{on}} = +3\text{V}$) on the sample of interest (*e.g.* after the occurrence of the I_g blockade). It is noted that I_{relax} is governed by the Curie-von Schweidler law [12] and the slope of I_{relax} is independent of the V_{on} magnitude, as illustrated in Fig. 4.11.

Figure 4.11. I_{relax} at different V_{on} (from 2 to 5V with step of 0.5V) of an unstressed fresh device.

In the NC-embedded bi-layer gate stacks, I_{relax} is determined mainly by the NCs and the high- κ layer since I_{relax} of SiO_2 layer is negligible [13]. Fig. 4.12 shows I_{relax} for Sample-2 (Fig. 4.6(a)) and Sample-3 (Fig. 4.6(b)) before stressing and after the occurrence

of an I_g blockade. We propose that in the fresh sample before CVS, I_{relax} is mainly contributed by the discharging of the as-deposited shallow traps and polarization/relaxation in the high- κ layer. But after the appearance of I_g blockade, a larger and flatter (*i.e.*, longer decay time) I_{relax} was observed, implying that the discharging of the electrons from NCs in the vicinity of the breakdown path has occurred. We believe that I_{relax} discharge takes place through the breakdown path as well.

Figure 4.12. I_{relax} before stress and after I_g blockade of two different breakdown hardness samples: (a) Sample-2 with breakdown hardness of 200 nA under $V_{g\text{-stress}}=4.8$ V and (b) Sample-3 with breakdown hardness of 500 nA under $V_{g\text{-stress}}=5.5$ V.

As shown in Fig. 4.12, the slope of $\log(I_{\text{relax}})-\log(t)$ curves after the occurrence of the I_g blockade in Samples 2 and 3 are 0.60 and 0.83, respectively. It is noted that smaller slopes indicate longer discharging time (t_d). The transient discharging time constant (τ) is known as the time needed for I_{relax} to decay to 37% (or $1/e$) of its initial value. As shown in Fig. 4.13, it will take $t_d=5\tau$ (for constant slope) to reach a completely discharged state in NCs.

Figure 4.13. Estimation of discharged amount from I_{relax} before stress and after I_g blockade.

In the fresh samples before CVS, t_d is ~ 60 sec for all three samples. On the other hand, the estimated t_d for Samples 2 and 3 at Stage 3 are 120 and 100 sec, respectively. Since the gate is biased at 0 V during the I_{relax} measurement, the charge density and electric field will decrease as more electrons tunneling out of the NCs [14]. Hence the decay slope is supposed to decrease with time. Therefore, t_d extrapolated here is the minimum discharging time. The difference in t_d after the I_g blockade among the 2 different samples is ascribed to the different breakdown hardness of the percolation path. Comparing Samples 2 and 3 under the same stressing conditions (see Table. 4.1), the I_g blockade in Sample 2 occurred after I_g reached $I_{g-\text{max}}=200$ nA, while in Sample 3, the I_g blockade occurred after $I_{g-\text{max}}=500$ nA was reached, suggesting that Sample 3 has experienced a harder breakdown than that of Sample 2. As a consequence, I_g of Stage 3 in

Sample 3 is one order of magnitude lower than that of the fresh sample, while that in Sample 2 is similar to its pre-TDDB stress induced I_g (refer to the insets in Fig. 4.6). Thus, it is believed that sample with higher breakdown hardness has more lateral distribution of defects in the dielectric [15]; and thus more NCs can be charged and discharged more easily because of a more conductive path. The number of discharged electrons during the I_g blockade period can be estimated from Figs. 4.12(a) and (b) and is found to be $4\sim 8 \times 10^7$ in Samples 2 and 3.

TABLE. 4.1. STRESSING CONDITION OF AU-NC SAMPLES

<i>Sample^a</i>	<i>Stressing condition</i>	<i>I_{g-max} before I_g blockade</i>
1 (Fig. 4.5)	Current limited CVS	500 nA
2 (Fig. 4.6(a))	CVS	200 nA
3 (Fig. 4.6(b))	CVS	500 nA

^a. Au-NC in SiO₂/Al₂O₃ dielectric stack

Since the electronic energy levels are discrete in nanoparticles because of the electron wave function confinement in miniaturized volumes of the nanoparticles, the average electronic energy level spacing between successive quantum levels, δ , for an Au NC with diameter of 3 nm can be estimated from Kubo gap equation [16],

$$\delta = 4E_F/3m \quad (4.2)$$

where E_F is the Fermi energy of bulk Au and m is the total number of valence electrons in the Au nanocrystal. Thus, for an individual Au nanoparticle with 3 nm in diameter, the total number of valence electrons is close to ~ 900 and the estimated value of δ would be ~ 8 meV.

Figure 4.14. Schematic energy band diagram of 3 nm Au-NC in between Al_2O_3 and SiO_2 dielectric.

As shown in Fig. 4.14, the potential well depth E_p of 3nm Au-MNC sandwiched between SiO_2 (8.9eV) and Al_2O_3 (9eV) can be estimated from

$$E_p = \varphi_{Au} - \chi_{ox} - E_c - E_1 \quad (4.3)$$

where φ_{Au} is the workfunction of Au-MNC, χ_{ox} is the electron affinity of oxide, E_1 is the first energy level in quantum dot. With $\delta = 8$ meV and a calculated potential well depth, $E_p = 4.05$ eV, we estimate that each Au NC can store a “maximum” of one thousand electrons assuming there is no electron detrapping and each energy level is occupied by

two (spin-up / spin-down) electrons. This corresponds to “at least” $\sim 40 \times 10^3$ NCs affected by the breakdown path in Sample 2 and 3. However, as depicted in Fig. 4.15, the number of NCs just above (or below) the breakdown spot is only 11~50 NCs as the localized breakdown occurs in one-layer dielectric could have radial distribution of 30~50 nm [7]. It indicates that the discharge radius is much larger (up to ~ 650 nm for a given NC density of $3 \times 10^{12} \text{ cm}^{-2}$) than the diameter of the breakdown spot. Therefore, a very localized breakdown path will not just affect the charge-holding capability of the NCs located just above (or below) the breakdown spot. Rather, additional charges stored in the surrounding NCs will also tend to leak through the breakdown path gradually, indicating that the discharging of NCs via the breakdown path is not a localized phenomenon.

Figure 4.15. Radial distribution of a percolation path. The diameter of the breakdown path in SiO_2/SiON gate dielectrics is estimated to be 30 nm – 50 nm at the early stage of a breakdown [7]. All dimensions are shown to scale.

4.4 Dielectric Breakdown Layer Detection

4.4.1 Methodology and Approach

As discussed in Section 4.3, two different types of I_g evolution trends were observed under CVS in NC-embedded $\text{SiO}_2/\text{Al}_2\text{O}_3$ gate stacks. In short, I_g blockade behavior (denoted as Type-B in Fig. 4.16(b)) and the spontaneous I_g evolution without I_g blockade behavior (denoted as Type-A in Fig. 4.16(a)) were observed. We have discussed that I_g blockade phenomenon is ascribed to the lateral charging of the NC through the percolation path formed in one layer of the $\text{SiO}_2/\text{Al}_2\text{O}_3$ stack. Comparing Type-A and Type-B mode I_g evolution trends, one can clearly see that their first layer of breakdown (*i.e.*, either in SiO_2 tunnel oxide or NC-embedded Al_2O_3), is different.

In the NC-embedded gate stacks, the electrostatic energy required to add one electron to the quantum dot with a self-capacitance $C=4\pi\epsilon_0\epsilon_r R$, where R is the NC radius, is related to a single-particle Coulomb blockade voltage, $V_0=q/C$, below which tunneling is suppressed [17]. An average diameter of 3 nm Au-NCs embedded in an Al_2O_3 dielectric layer with permittivity of $\epsilon_r \sim 9$ has a self-capacitance $C \sim 1.5$ aF and a single-particle Coulomb blockade voltage $V_0 \sim 53$ mV with $q=e/2$. It is found that the charging energy $E_c = e^2/C$ is much larger (~ 17 -112 times) than the thermal energy ($k_B T$) gained by electrons at low temperature regime (11-70K). The tunneling is suppressed at the temperatures where thermal energy is less than the charging energy (typically $E_c > 10 k_B T$) [18]. Therefore, in 3-5 nm Au-NCs embedded gate stacks, electron transport would be essentially temperature independent, especially at low temperatures. As the NCs are embedded in the Al_2O_3 layer, temperature dependency is used as an indicator of the layer

which breaks down first. In other words, temperature sensitive tunneling behavior is expected if the current transport is through SiO_2 layer [19]. In contrast, if high- κ is the surviving layer, electron transport could be independent of temperature, especially at very low temperatures as Au-NCs are embedded in the high- κ layer. With this approach, we would be able to differentiate SiO_2 induced breakdown from Al_2O_3 induced breakdown.

Figure 4.16. I_g evolution under CVS in substrate injection mode with $V_{g\text{-stress}}=5\text{V}$. (a) Spontaneous I_g evolution pattern denoted as Type-A and (b) a new 3-stage I_g evolution pattern with leakage drop denoted as Type-B are shown.

4.4.2 Results and Discussions

Fig. 4.17 shows the Arrhenius plot of fresh sample, Type-A and Type-B mode one-layer breakdown samples in a low voltage regime. It can be clearly seen that I_g increases with temperature in Type-A samples, while the fresh sample and Type-B samples show temperature independent electron transport behavior especially in the very low temperature regime (11-70 K). Therefore, the surviving layer in Type-A sample is SiO_2 and that in Type-B sample is Al_2O_3 . In other words, the I_g blockade behaviors in Type-B samples are owing to SiO_2 induced breakdown and I_g evolutions of Type-A samples are caused by Al_2O_3 induced breakdown, as summarized in Table. 4.2.

Figure 4.17. Arrhenius plot of the gate leakage current in the low voltage regime ($V_g=1V$) for fresh sample and Type-A, Type-B mode one layer breakdown samples. The gate leakage current from 11 to ~70K are linear in fresh and Type-B mode samples while temperature dependence behavior is more drastic in Type-B mode samples. The inset presents a plot of activation energy for transport regime above 150K in Type-A mode sample.

TABLE 4.2. IG EVOLUTION MODES OF AU-NCs SAMPLES UNDER CVS

<i>Leakage conduction mode</i>	<i>Behavior</i>	<i>1st breakdown layer</i>
Type-A	Spontaneous breakdown	NC-embedded Al ₂ O ₃
Type-B	I_g blockade	SiO ₂

The activation energy (E_A) was extracted for Type-A sample to further confirm our findings. In the Arrhenius plot of I_g in Type-A sample, strong ($T > 150\text{K}$) and weak ($T < 150\text{K}$) temperature dependent regimes can be observed. The activation energy in the high temperature regime varies with the applied gate voltage, as depicted in the inset of Fig. 4.17. It is noted that the maximum E_A (at $V_g=0\text{V}$) is 42 meV, which agrees well with the literature reported results for SiO_2 transport [20]. Among 32 devices tested and analyzed, ~56% of them indicated that the NC-embedded Al_2O_3 was the first layer to break down (Type-A) during substrate injection stress, while ~20% of them suggested a SiO_2 breakdown (Type-B). The rest of the samples showed no clear evidence of Type-A or Type-B mode for a stress time of 1000 sec.

Figure 4.18 shows the Weibull plot of individual Al_2O_3 and SiO_2 breakdown detected using the post-breakdown I_g signatures and temperature dependency discussed above. The Weibull slope β is used to indicate the number of traps (N) needed for percolation breakdown and is expressed as $\beta = \alpha \cdot N$, where α is the power law exponent for SILC bulk trap generation and $N = t_{\text{ox}} / a_0$, in which t_{ox} is the thickness of oxide and a_0 is the size of the defect (trap) [21]. It is noted that the Weibull slope of NC-embedded Al_2O_3 failure is smaller than that of SiO_2 . Given that the SILC bulk trap generation rate α is 0.35 [22] in both Al_2O_3 and SiO_2 films, the trap size can be estimated to be 0.8 nm and 1.71 nm for the SiO_2 and Al_2O_3 traps, respectively. However, as the bulk trap generation rate α in the high- κ layer could be higher than 0.35 because of the presence of grain boundaries that is a region of excess inter-grain open volume [23-25], and the embedded NCs (large in size) that can serve as traps, the average trap size in NC embedded Al_2O_3 is supposed to be larger than 1.71 nm.

Figure 4.18. Weibull probability plot of the time to failure for Al₂O₃ and SiO₂ layer breakdown in Au-NCs embedded dual layer gate stack.

When there is a percolation in one of the dielectric stack layers, a localized electric field enhancement occurs at the SiO₂/Al₂O₃ interface if the stress is maintained. Therefore, in the case of SiO₂ breakdown, the enhancement in electric field may be attributed to a lateral charging of the near-by NCs through the breakdown path so that all the energy states in the affected NCs are occupied by the injected electrons from substrate injection stress until no more electron injection into NCs is possible. As a consequence, the I_g blockade behavior is observed in the SiO₂ breakdown case. Hence, if there is no further increase in voltage (*i.e.*, stressing voltage remain unchanged), Al₂O₃ is less likely to breakdown as the tunneling current is suppressed by the Coulomb blockade effect. In contrast, when Al₂O₃ is the first layer to breakdown, a localized electric field enhancement at the vicinity of interface could result in a spontaneous breakdown of the complete bi-layer gate stack. The failure of NC-embedded high- κ layer could correspond to localized electric field enhancement in the vicinity of NCs [26] and high localized trap

generation rate around grain boundaries [23-25].

4.5 Summary

In summary, we studied the dielectric breakdown behaviors and its impact on device characteristics in MOS capacitor structures with MNCs embedded dual-layer ($\text{SiO}_2/\text{high-}\kappa$) gate stacks. Our study shows the possible effects on the charging and discharging of multiple MNCs due to localized breakdown either in the tunnel oxide layer or the NC-embedded high- κ layer. We believe this finding will contribute to further improvement of promising NC-embedded gate stacks from a reliability perspective since the previous understanding predicted that breakdown in the tunneling dielectric or blocking oxide is a localized phenomenon which affects limited number of NCs.

The important results are summarized as follows:

- (1) The charging and discharging of MNCs embedded in $\text{SiO}_2/\text{high-}\kappa$ gate stacks can take place easily via a percolation path formed in the dielectric. It is found that the localized breakdown not only affects charge holding capability of the affected NCs, but also provides a leakage path for the charges stored in the surrounding NCs. Therefore, the discharging of NCs through the breakdown path is a non-localized phenomenon.
- (2) From the post-breakdown characteristic of the conduction path formed in the different dielectric layers of NC-embedded $\text{SiO}_2/\text{high-}\kappa$ gate stack, SiO_2 and NC-embedded Al_2O_3 dielectric breakdowns are successfully detected based on the physics underlying the Coulomb charging energy related to thermal energy gained by electrons at low voltage

and in the very low temperature regime (from 11 to 70K). With this approach, the Weibull slope β which is a key indicator of the number of traps needed for percolation breakdown, for SiO₂ and NC-embedded high- κ failures were further analyzed. The average trap size in NC embedded Al₂O₃ is found to be much larger (>1.71 nm) than that in SiO₂ (0.8 nm), which could be due to the presence of grain boundaries in Al₂O₃ film and the embedded NCs. Our experimental results show that SiO₂ breakdown leads to a lateral charging/discharging among NCs while the Al₂O₃ percolation leads to spontaneous breakdown of the complete bi-layer gate stack that is observed in majority of the tested samples.

References

- [1] P. K. Singh, G. Bisht, M. Sivatheja, C. Sandhya, G. Mukhopadhyay, S. Mahapatra, et al., "Reliability of single and dual Layer Pt nanocrystal devices for NAND flash applications: A 2-region model for endurance defect generation," in Reliability Physics Symposium, 2009 IEEE International, 2009, pp. 301-306.
- [2] P. K. Singh, G. Bisht, R. Hofmann, K. Singh, N. Krishna, and S. Mahapatra, "Metal Nanocrystal Memory With Pt Single- and Dual-Layer NC With Low-Leakage Al_2O_3 Blocking Dielectric," *Electron Device Letters*, IEEE, vol. 29, pp. 1389-1391, 2008.
- [3] W. Guan, S. Long, M. Liu, Q. Liu, Y. Hu, et al., "Modeling of retention characteristics for metal and semiconductor nanocrystal memories," *Solid-State Electronics*, vol. 51, pp. 806-811, 2007.
- [4] B. P. Linder, J. H. Stathis, R. A. Wachnik, E. Y. Wu, S. A. Cohen, A. Ray and A. Vayshenker, "Gate oxide breakdown under current limited constant voltage stress", *IEEE VLSI Technology Symposium*, 2000, p. 214.
- [5] N. Raghavan, K. L. Pey, W. H. Liu, X. Wu, and X. Li, "Unipolar recovery of dielectric breakdown in fully silicided high- κ gate stack devices and its reliability implications," *Applied Physics Letters*, vol. 96, p. 142901.
- [6] Y. N. Chen, K. L. Pey, K. E. J. Goh, Z. Z. Lwin, P. Singh, and S. Mahapatra, "Study of automatic recovery on the metal nanocrystal-based $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stack," *Applied Physics Letters*, vol. 98, pp. 083504, 2011.
- [7] X. Li, C. H. Tung, and K. L. Pey, "The radial distribution of defects in a percolation path," *Applied Physics Letters*, vol. 93, p. 262902, 2008.
- [8] P. Beecher, A. J. Quinn, E. V. Shevchenko, H. Weller, and G. Redmond, "Insulator-to-Metal Transition in Nanocrystal Assemblies Driven by in Situ Mild Thermal Annealing," *Nano Letters*, vol. 4, pp. 1289-1293, 2004.
- [9] Y. Chia-Han, K. Yue, W. Rui, L. Chen-Han, and K. Way, "Failure analysis of nanocrystals embedded high- κ dielectrics for nonvolatile memories," in

- Reliability Physics Symposium, 2008. IRPS 2008. IEEE International, 2008, pp. 46-49.*
- [10] R. Wan, J. Yan, Y. Kuo, and J. Lu, "Dielectric Breakdown and Charge Trapping of Ultrathin ZrHfO/SiON High- k Gate Stacks," *Jpn. J. Appl. Phys.*, vol. 47, pp. 1639-1641, 2008.
 - [11] X. Zhen, L. Pantisano, A. Kerber, R. Degraeve, E. Cartier, S. De Gendt, *et al.*, "A study of relaxation current in high- κ dielectric stacks," *IEEE Trans. Electron Devices*, vol. 51, pp. 402-408, 2004.
 - [12] K. J. Andrew, , "Dielectric relaxation in solids," *Journal of Physics D: Applied Physics*, vol. 32, p. R57, 1999.
 - [13] L. Wen, K. Yue, and K. Way, "Dielectric relaxation and breakdown detection of doped tantalum oxide high- κ thin films," *Device and Materials Reliability, IEEE Transactions on*, vol. 4, pp. 488-494, 2004.
 - [14] W. Guan, S. Long, M. Liu, Q. Liu, Y. Hu, *et al.*, "Modeling of retention characteristics for metal and semiconductor nanocrystal memories," *Solid-State Electronics*, vol. 51, pp. 806-811, 2007.
 - [15] X. Li, C. H. Tung, K. L. Pey, and V. L. Lo, "The physical origin of random telegraph noise after dielectric breakdown," *Applied Physics Letters*, vol. 94, p. 132904, 2009.
 - [16] C. N. R. Rao, G. U. Kulkarni, P. J. Thomas, and P. E. Peter, "Size-Dependent Chemistry: Properties of Nanocrystals," *Chemistry - A European Journal*, vol. 8, pp. 28-35, 2002.
 - [17] R. Parthasarathy, X.-M. Lin, and H. M. Jaeger, "Electronic Transport in Metal Nanocrystal Arrays: The Effect of Structural Disorder on Scaling Behavior," *Physical Review Lett.*, vol. 87, p. 186807, 2001.
 - [18] G. Schmid (Ed.), "Nanoparticles: From Theory to Applications," Wiley-VCH, Weinheim, 2010.
 - [19] P. Lundgren and M. O. Andersson, "Temperature dependence confirmation of tunneling through 2-6 nm silicon dioxide," *Solid-State Electronics*, vol. 39, pp. 1143-1147, 1996.
 - [20] R. G. Southwick, J. Reed, C. Buu, H. Bui, R. Butler, G. Bersuker, and W. B.

- Knowlton, "Temperature (5.6-300K) Dependence Comparison of Carrier Transport Mechanisms in HfO₂ and SiO₂ MOS Gate Stacks", *IEEE IRW*, 2008, pp. 48-54.
- [21] A. T. Krishnan and P. E. Nicollian, "Analytic extension of the cell-based oxide breakdown model to full percolation and its implications," in *Proc. Int. Reliab. Phys. Symp.*, 2007, pp. 232–239.
- [22] C. Petit and D. Zander, "Low voltage stress-induced leakage current and traps in ultrathin oxide (1.2-2.5 nm) after constant voltage stresses," *Semiconductor Science and Technology* 22 (10), 1165 (2007).
- [23] R. Ranjan, K. L. Pey, C. H. Tung, L. J. Tang, G. Groeseneken, L. K. Bera, and S. De Gendt, "A comprehensive model for breakdown mechanism in HfO₂ high- κ gate stacks," in *IEDM Tech. Dig.*, 2004, pp. 725-728.
- [24] Y. C. Ong, D. S. Ang, K. L. Pey, S. J. O. Shea, K. E. J. Goh, C. Troadec, C. H. Tung, T. Kawanago, K. Kakushima, and H. Iwai, "Bilayer gate dielectric study by scanning tunneling microscopy," *Appl. Phys. Lett.*, vol. 91, p. 102905, 2007.
- [25] N. Raghavan, K. L. Pey, K. Shubhakar and M. Bosman, "Modified Percolation Model for Polycrystalline High- κ Gate Stack With Grain Boundary Defects," *IEEE Electron Device Lett.*, vol. 32, pp. 78-80, 2011.
- [26] C. Lee, U. Ganguly, V. Narayanan, T. H. Hou, J. Kim, and E. C. Kan, "Asymmetric electric field enhancement in nanocrystal memories," *IEEE Electron Device Lett.*, vol. 26, pp. 879-881, 2005.

CHAPTER FIVE

LOCALIZED CHARGE TRANSPORT AND LATERAL CHARGE DIFFUSION IN SINGLE-LAYER METAL NANOCRYSTAL EMBEDDED GATE STACKS

5.1 Introduction

As discussed in Chapter 4, we have found that the strong inter-dot tunneling among NCs could lead to substantial charge leakage, not only from the NCs just above the degraded dielectric, but also from the surrounding NCs. Though NC memories have shown the potential for enhanced device scalability and reduced defect related charge loss [1], the lateral charge diffusion among NCs could provide limitation for future scaling. Moreover, the vertical charge loss directly influences the retention capability of the memory gate stacks. Therefore, a systematic study of lateral charge diffusion and vertical charge loss, which are critical degradation mechanisms in NC-based flash memory, is crucial.

In this chapter, we present a study on the localized charge transport in MNC-embedded gate stacks in real time by Kelvin Force Microscopy (KFM). A two-step analysis approach is demonstrated that enables an in-depth understanding of charge trapping and charge decay mechanisms in MNC-embedded high- κ /SiO₂ gate stacks. The two competing mechanisms of vertical charge loss and lateral charge diffusion are explored by discharging current measurement at elevated temperatures and the KFM characterization.

5.2 Experiments

Figure 5.1. Schematic (not to-scale) of MNC-embedded gate stack used in (a) KFM characterization and (b) transient discharging current measurement.

To understand the role of MNCs on the charge retention properties, we carried out a comparative study on two types of MNC stacks with different work functions - Pt (~5.3 eV) and Au (~5 eV) [2, 3]. These NCs were embedded in a bi-layer stack comprising a

SiO_2 (~ 50 Å) layer as tunnel oxide and an Al_2O_3 (~ 60 Å) layer as the blocking oxide on an n-Si substrate with As-doping of $3 \times 10^{19} \text{ cm}^{-3}$. The Au/Ti (100 nm/ 30 nm) pads with $\varnothing 150 \mu\text{m}$ were deposited as top electrodes for the transient discharging current measurement.

Figure 5.2. Cross-sectional view TEM micrographs of (a) Au-NC sample and (b) Pt-NC sample. The average size (diameter) of the NCs is $\sim 3 \text{ nm}$. (c) Plan view TEM micrograph showing the distribution of MNCs.

Fig. 5.1 illustrates the schematic of MNC-embedded gate stack used in our study. Figs. 5.2(a) and (b) show the cross-sectional view TEM micrographs of Au and Pt

samples. It is seen that the average diameter of the NCs is around 3 nm. The TEM sample thickness is ~100 nm and the MNCs seem to be overlapped due to the shadowing effect. The plan view TEM micrograph in Fig. 5.2(c) shows that the MNCs are isolated and the number density is about $\sim 3 \times 10^{12} \text{ cm}^{-2}$. The two-step analysis approach is demonstrated by KFM characterization on the blanket sample and transient discharging current, I_{relax} measurement on the macroscopic capacitors with a diameter 150 μm Au electrode.

5.3 Kelvin Force Microscopy Characterization

To ‘visualize’ the real time evolution of the charges trapped in the MNCs, KFM measurement was performed at room temperature. The charge injection was carried out with a contact mode using a Pt coated conductive tip (Olympus AC240TM) with a nominal spring constant 2 N/m and a nominal radius $28 \pm 10 \text{ nm}$. For the charge injection process, the tip was brought in contact with the sample surface by reducing the amplitude set-point to near zero (Fig. 5.3(a)). Then a voltage of -10 to 10V was applied for 10s to the tip while the substrate was grounded. During charge injection, the maximum contact area was estimated to be $2.46 \times 10^{-3} \mu\text{m}^2$ assuming the maximum contact radius r_c is equivalent to tip radius R_{tip} . If the field emission is only from the part of the tip which is in contact with the sample, the effective contact area A_{eff} is expected to be smaller since the contact radius is supposed to be smaller than the tip radius which can be expressed as [4]:

$$r_c = \left(\frac{3}{4} (k_1 + k_2) F_c R_{\text{tip}} \right)^{\frac{1}{3}} \quad (5.1)$$

and therefore

$$A_{\text{eff}} = \pi r_c^2 = \pi \left(\frac{3}{4} (k_1 + k_2) F_c R_{\text{tip}} \right)^{\frac{2}{3}} \quad (5.2)$$

where F_c is contact force and $k_i = (1 - \nu_i^2)/\epsilon_i$, where ϵ_i and ν_i are Young's moduli and Poisson's ratios of the tip and sample surface, respectively.

After charging the NCs, KFM images were acquired in two-pass procedures [5] (Fig. 5.3(b)). In the first pass, the tapping mode scan was carried out along a single line to provide the surface topography (1 in Fig. 5.3(b)). In the second pass, the tip was raised above the surface at a fixed distance (~ 50 nm) (2 in Fig. 5.3(b)) and the localized variation of surface potential was measured along the same topographic profile (3 in Fig. 5.3(b)).

Figure 5.3. Illustration of KFM methodology: (a) charge injection on the NC-embedded dielectric surface and (b) surface potential measurement in the lift mode.

Figure 5.4. Plan view and 3D view KFM images with associated potential profile just after (a) negative tip injection and (b) positive tip injection. The negative (positive) tip injection results in negative (positive) trapped charges which appeared as dark (bright) spot in plan view KFM images.

Due to the time required for setting up the KFM measurement, it is impossible to acquire potential profile of injected charge cloud at the time of the completion of charge injection. It takes at least 60 sec to reach the center of localized charge spot after the charge injection. Fig. 5.4(a) and (b) show the plan view and 3D view KFM maps of induced surface potential after negative charge injection and positive charge injection, respectively. The potential profile was acquired along a horizontal cross section of a KFM map. Then, the background

potential was subtracted from the potential profile, resulting in the relative potential profile which is related to the amount of charge stored in the NCs. The result analysis was based on the change in surface potential peak (ΔV_{max}) measured upon charge injection and its evolution with time. The charge decaying process was monitored and the full-width-at-half maximum (FWHM) of the potential profile was analyzed. It is found that the positive (negative) bias applied to the AFM tip resulting in a positively (negatively) trapped charge in the sample. This indicates that the charge transfer occurs between the tip and the dielectric layer rather than from the Si-substrate which is in good agreement with other literature reports [6-8].

Figure 5.5. The time dependence of charge decay in Au-NC embedded samples acquired at different t_s , the time interval between the completion of charge injection (-6V, 10sec) and arriving first potential measurement at the center of the localized charge spot.

Fig. 5.5 shows the time dependence of charge decay acquired at different t_s , the time interval between completion of charge injection and reaching first potential measurement at the center of the localized charge spot, in the Au-NCs embedded sample. The experiment was carried out at different locations so that no initial charges were present before the charge injection. The higher potential was observed for smaller t_s . Assuming the MNC distribution are rather uniform for a localized charge spot with a diameter of ~500 nm, it is found that the influence of varying t_s has negligible effect on the charge decay. It suggests there is no significant charge transfer between the tip and the sample at a fixed distance of 50 nm.

5.3.1 Tip-to-Sample Distance Dependent KFM Resolution

In the MNC embedded gate stacks, the influence of tip-to-sample distance on the KFM measurement is illustrated, as shown in Fig. 5.6. The plan view KFM image was acquired at a lift height of 5 nm after charging the sample (Fig. 5.6(a)). Then the single line scanning was carried out at various tip-to-sample distances (5-300 nm) across the center of the charge spot as shown in Figs. 5.6(b) and (c). It is found that the lateral resolution increases with decreasing tip-to-sample distance.

The dependence of KFM resolution on the tip-to-sample distance could be understood as follows. In the KFM system, the electrostatic force, F_{ES} , is used to provide potential feedback. In the case of parallel plate electrodes, F_{ES} is given by [9]

$$F_{ES} = \frac{\epsilon_0 A}{2} \times \frac{V_{DC}^2}{d^2} \quad (5.3)$$

where V_{DC} is the applied DC bias between two electrodes, A is the electrode area and d is the distance between two electrodes. Therefore, it can be seen that the electrostatic force which is proportional to the square of the electric field depends on the applied bias and the separation between two electrodes.

The three-dimensional shape of KFM tip results in non-uniform electrostatic force on the surface of the tip and the sample. However, the overall F_{ES} between the tip and the sample still can be assumed as [9]

$$F_{ES} \propto \frac{[V_{DC} - (\phi_t - \phi_s)]^\rho}{d_{ts}^\gamma} \quad (5.4)$$

where ϕ_t and ϕ_s are work functions of tip and sample, d_{ts} is the distance between the tip apex and the sample. ρ and γ are related to the dependence of electrostatic force on the sample bias and the tip-to-sample distance, respectively. In the parallel plates, the electrostatic force dependence of tip-sample force shows a quadratic inverse d_{ts} dependence (*i.e.*, $F_{ES} \propto d_{ts}^{-\gamma}$ with $\gamma=2$). Whereas for a KFM system, it is reported [9,10] that the tip-sample force decreases much more slowly than in the parallel plate case (*i.e.*, $\gamma < 2$). It suggests that the effective tip-sample distance is strongly correlated to the geometry of the tip which is ascribed to the tip apex and the cone contribution [9, 10]. It is obvious that the lateral resolution increases with the decreasing tip-sample distance. However, the reduction in the tip-sample distance could cause charge transfer between the tip and sample during probing, which complicates the charge decay analysis in NC embedded samples.

Figure 5.6. (a) The plan view KFM image at lift height of 5 nm after charging the sample. The single line scanning along A-A' was performed to investigate the impact of lift scans height on KFM surface potential. (b) Plan view and (c) 3D view KFM images at various lift height: increasing from 5 nm to 300 nm then decreasing to 5 nm as indicated with arrow.

Therefore, the reproducible KFM measurement was acquired by

- (1) recording the potential profile at a constant lift height of 50 nm throughout the experiments to compromise the tip-to-sample distance dependent lateral resolution and charge transfer between the tip and the sample,

- (2) using the same tip throughout the particular experiment to compare and contrast the results, as tip geometries have a strong influence on the charge injection trends and the KFM measurement [11], and
- (3) performing under the same environmental conditions, such as ambient pressure and humidity.

5.4 Localized Charge Trapping and Charge Decay in Single-Layer MNCs

Fig. 5.7 represents a series of 2D and 3D view KFM potential images and topographic AFM images on a NC-embedded sample after charge injection (V_{inj}) with -8V for 10 sec. The increase in the apparent height at injected charge cloud area can be seen in the first a few topographic images ($t < 10$ mins) as shown in Fig. 5.7(c)). During the tapping mode topography scanning, the presence of injected charges provides an additional electrostatic force [12]. At the localized charge spot, the AFM tip retracts from the surface to maintain the imposed oscillation amplitude during the trapping mode scanning. It results in an increase in the height in the topographic images. The 2D and 3D view KFM images clearly show an increase in the diameter and a decrease in the peak of the potential distribution with time.

Figure 5.7. (a) Plan-view and (b) 3D-view KFM images and (c) AFM topography images at $t=1$, 6 and 15 mins after a negative tip injection into an Au-NC embedded sample.

(a)

(b)

Figure 5.8. (a) Evolution of potential profile at $t=10$ mins and $t=83$ mins after charge injection in Pt- and Au- samples, normalized to peak at $t=10$ mins. (b) Schematic energy band diagram of the two adjacent NCs during retention.

Fig. 5.8 (a) illustrates the potential profiles acquired along a horizontal cross section of KFM images of Pt- and Au-NC samples at 10mins and 83mins after charge

injection (*i.e.*, retention of trapped charges), respectively. For Pt, the broadening of the potential profile is not so significant but peak height obviously decreases. In Au-embedded samples, broadening of the potential profile with the reduction in the peak height was observed during the decay. The decrease in the potential peak intensity and the increase in the size of the potential distribution provide a clear evidence of the lateral charge diffusion with time [13, 14]. The schematic energy band diagram of two adjacent NCs during retention is illustrated in Fig. 5.8(b). NC₁ corresponds to the NC with higher potential energy and NC₂ is the neighboring NC with lower potential energy. Electrons are added to the NC when their potential energy is high enough to overcome addition energy (energy level spacing ΔE and the charging energy e^2/C) and when occupied electronic states in one NC align with available empty states in the adjacent NC [15]. This allows electrons to tunnel to adjacent NCs, which is detected as a decrease in the potential peak and an increase in the FWHM of the potential distribution.

Fig. 5.9 shows the decrease in the potential peak with time in MNC-embedded samples and the control sample for comparison. The control sample (*i.e.*, without NCs) consists of 10 nm Al₂O₃ and 1 nm interfacial layer SiO₂ on the Si substrate. In the control sample, the charge decay is very fast and not detectable after retention time of 15 mins. In contrast, the MNC-embedded samples show significantly longer charge decay, suggesting that most of the injected charges are trapped in the MNCs rather than the high- κ oxide layer. The charge decay in KFM measurement accounts for both lateral and vertical charge loss, which is different from the macroscopic capacitor based electrical measurement that only sense the vertical charge loss. To see the effect of injection voltage on the charge decay, the KFM experiments were repeated with various injection

voltages (-4 to -8V). It is noted that $\Delta V_{\text{-max}}$ is directly proportional to the absolute injection voltage. However, the charge decay trend is independent of the amount of injected charges; rather it depends only on the material property of the embedded MNCs (Fig 5.8). Under the same experimental conditions such as AFM tip, tip-sample distance and ambient environment, we have obtained consistent results at various locations on the sample, suggesting that the size and distribution of MNCs are rather uniform.

Figure 5.9. Time dependence of normalized potential to the peak at $t=0$ for the Au (the closed symbols) and Pt samples (the open symbols) after charge injection of -4V, -6V, -8V (blue, red and black symbols) for 10sec. The time dependence of charge decay in the control sample (*i.e.*, without NCs) is shown for comparison.

5.4.1 Charge Decay Fitting Laws and Characteristic Time Constant

In order to get a better insight into the charge decay mechanisms in the MNC-embedded gate stacks, the change in potential peak with time (1-5000 sec) was fitted with

(1) Single exponential function: $V(t) = V_0 \cdot \exp\left(-\frac{t}{\tau}\right)$,

(2) Double exponential function: $V(t) = V_1 \cdot \exp\left(-\frac{t}{\tau_1}\right) + V_2 \cdot \exp\left(-\frac{t}{\tau_2}\right)$, and

(3) Stretched exponential function: $V(t) = V_0 \cdot \exp\left[-\left(\frac{t}{\tau}\right)^\beta\right]$

as shown in Fig. 5.10.

Figure 5.10. Charge decay fitting with exponential laws in Pt-MNCs-embedded gate stacks. The time constant τ for different fitting laws are shown. Only the stretch exponential function can extrapolate close to actual decay data point which is measured at $t=30000$ sec.

For a limited fitting range ($\sim 5 \times 10^3$ sec), the experimental data points are fitted well with double exponential and stretch exponential functions. However, only the stretch exponential function can extrapolate close to actual decay data point which is measured at $t=30000$ sec. Therefore, the characteristic decay time constant τ in Pt- and Au-MNCs embedded samples is computed by fitting the measured data to the stretched exponential function [16-18]: $V(t)=V_0.\exp(-t/\tau)^\beta$ (Fig. 5.11(a)).

The stretched exponential function was first proposed by Kohlrausch (1863, 1876) and later introduced by Williams and Watts (1970), and it is known as Kohlrausch-Williams-Watts (KWW) function [16-18]. This function describes the slow relaxation in a number of distinct physical systems. The parameter β varies in the range of 0-1 and $\beta=1$ corresponds to a simple exponential decay. It is found that the charge decay in both Pt and Au samples is best fitted with a stretch factor of $\beta \sim 0.52$ while $\beta \sim 1$ in the control sample. Kita and Koizumi [19] reported that the width of the loss peak increases rapidly as β decreases and the response becomes effectively “flat” for $\beta \leq 0.3$. That is, the decrease in the β values is related to the broadening of the loss peaks; which indicates the coupling strength and/or strong interaction of the rigidity of the system [20]. Therefore, $\beta \sim 0.52$ observed in Pt and Au sample indicates a high coupling strength and strong interaction of MNC embedded system.

As can be seen in Fig. 5.11(b), τ for the Pt MNCs is ~ 4 times larger than that of the Au MNCs sample. The measured FWHM as a function of time is plotted in Figs. 5.11 (c) and 5.11(d), where the slope of the curve indicates the rate of lateral charge diffusion. The Au sample exhibits a much higher rate of diffusion compared to the Pt sample.

Figure 5.11. (a) Stretched exponential fitting to the charge decay of Pt and Au samples after charge injection of -4V, -6V, -8V (blue, red and black symbols) for 10sec. (b) Characteristic time constant τ and (c) normalized FWHM as a function of time in Au and Pt samples at various injection voltages. (d) Zoom-in results of the FWHM evolution during the initial stage of retention in the Au and Pt samples.

5.4.2 Inter-Dot Tunneling Probability

The tunneling coefficient T for the Pt and Au-NC samples was evaluated using the conventional WKB approximation. For a small T , the transmission through a rectangular barrier is expressed as [21]

$$T = 16 \frac{E}{U_0} \left(1 - \frac{E}{U_0}\right) e^{-2KL} \quad (5.5)$$

where U_0 is the potential well height of the NC, E is the energy of electron in the NC, L is the spacing between adjacent NCs and K is the wave vector in the oxide:

$$K = \frac{2\pi}{h} \sqrt{2m_e(U_0 - E)} .$$

Given the spacing between NCs of ~ 3 nm and effective mass of electron m_e in the Al_2O_3 oxide to be $0.35m_0$ [21], it is found that the tunneling probability of escaping electrons from one NC to an adjacent NC through the Al_2O_3 spacing dielectric in the Au sample is ~ 3.3 times larger than that in the Pt sample owing to deeper potential well (*i.e.*, higher work function) of Pt-MNCs. Interestingly, the calculated ratio between the inter-dot tunneling probabilities of these two samples (~ 3.3) is in good agreement with the experimental results of the time constant presented earlier: τ of Pt sample is about 4 times larger than that in Au sample. This suggests that electron tunneling between the MNCs is the physical reason why the potential distribution flattens with the time after the charge injection.

5.4.3 Possible Charge Decay Paths in Single-Layer MNCs

As discussed above, the decrease in the potential peak and increase in FWHM with time are clear evidences of lateral charge diffusion. However, as can be seen in Figs. 5.11(a) and (d), during the initial decay period (<300sec), it is interesting to note that both FWHM and potential peak decreased in the Pt sample, but remained unchanged in the Au samples. If the amount of total charge is the main concern, the decrease in both potential peak and diameter suggests a vertical charge loss behavior. Since KFM resolution is tip-sample distance dependent (see Section 5.3.1), the vertical charge loss behavior (*i.e.*, decrease in KFM resolution) during the initial stage of charge decay is most likely caused by tunneling back of charge to the Si-substrate and/or Si-SiO₂ interface rather than towards surface states.

Therefore, two possible mechanisms of charge decay in single layer (SL) - MNC-embedded gate stacks can be defined as illustrated in Fig. 5.12: (1) inter-dot tunneling which leads to lateral charge diffusion (P_x) and (2) vertical charge loss through tunneling back of charge to the Si substrate (P_z). During the initial stage of retention (<300sec), the decrease in both potential peak and diameter is more significant in the Pt samples, suggesting that the vertical charge loss is the dominant mechanism ($P_z, \gg P_x$) during the initial stage of retention. While for the Au sample, the decrease in potential peak with an almost constant FWHM indicates that P_z and P_x are comparable during the initial decay period.

Figure 5.12. Illustration of the possible charge decay paths during retention in SL-MNC embedded gate stacks: P_x corresponds to inter-dot tunneling and P_z refers to vertical charge loss. The vertical charge loss towards surface state is ignored here since the decrease in both potential peak and diameter (*i.e.*, decrease in KFM resolution) during the initial stage of retention suggests the charges diffuse further away from the tip rather than towards surface states.

5.5 Discharging Current at Elevated Temperatures

To verify the vertical charge loss behaviors during the initial stage of retention, the transient discharging current measurement was performed using dielectric relaxation current (I_{relax}) approach as mentioned in Section 3.3.2 of Chapter 3. For electrical measurement, the samples with diameter 150 μm Au gate electrodes (*i.e.*, top electrodes) were used. The I_{relax} in the Pt and Au MNCs embedded samples were measured at elevated temperatures ranging from 300K to 450K by a Keithley 4200 semiconductor characterization system (Fig. 5.13). The I_{relax} was measured at zero gate voltage right after a removal of initial stressing bias. The electron flux direction is from the substrate towards the gate electrode under a positive gate bias. Upon a sudden removal of positive

gate bias, electrons tunnel back to the substrate and the tunneling rate decay with time. The I_{relax} behavior of the MNC-embedded high- κ /SiO₂ is similar to the I_{relax} of pure dielectric film, which follows the Curie-von Schweidler law [25]: $I_{\text{relax}}=at^{-n}$, with a slope n . In the NC-embedded dual layer gate stacks, I_{relax} is mainly contributed by the NCs and the high- κ layer since I_{relax} of SiO₂ is negligible [23, 26].

It is interesting to note that the slope n for the fast decay ($t < 60$ sec) becomes steeper as the temperature increases in the Pt samples while the opposite trend was observed in the Au samples (see insets of Fig. 5.13). The dependence of the slope n on the temperature should stem from the embedded NCs since the slope n in pure high- κ (*i.e.* without NCs) is independent of temperature [23, 27, 28]. At high temperatures, the electrons at the top energy level in NCs will gain enough thermal energy to diffuse to adjacent NCs (P_x) and/or tunneling back to Si substrate (P_z). Since the I_{relax} measurement senses the charge loss in vertical direction, the steeper I_{relax} slope in the Pt samples should be attributed to the tunneling of stored charges in NCs back to the substrate [27]. While in the Au samples, the shallower I_{relax} slope at higher temperatures suggests the charges de-trapped to the substrate is negligible compared with Pt samples, which could be due to faster inter-dot tunneling behavior observed in KFM analysis as discussed above.

Figure 5.13. Transient discharging current (I_{relax}) measurement in (a) Pt and (b) Au samples at temperatures ranging from 300 to 450 K. The inset shows the change in magnitude of the slope n with temperature from the curve fitting: $I_{\text{relax}} = at^{-n}$.

5.6 Charge Transport Mechanisms in Single-Layer MNCs

The KFM measurement shows two main possible charge decay paths in NC-embedded samples, *i.e.*, charge diffusion in lateral direction (P_x), charge de-trapped towards Si-substrate as discussed in Section 5.4. The transient discharging current measurement confirms the vertical charge loss behavior observed in the initial stage of retention during KFM measurement. As illustrated in Fig. 5.14, a physical model is proposed to explain the two competing mechanisms of charge decay observed in SL-MNC embedded gate stacks. The probability of charge tunneling to adjacent MNCs and/or Si-substrate depends on the difference in the localized electric field in lateral E_x and vertical E_z directions which relates to localized potential difference (ΔV) and the inter-dielectric spacing (L), as shown in Fig. 5.14.

Figure 5.14. Schematic of dominant charge decay paths in SL-MNCs embedded sample. E_x and E_z correspond to the localized electric field in lateral and vertical directions, respectively. L_x and L_z represent inter-dot spacing and tunnel oxide thickness, respectively.

Figure 5.15. (a) The potential profile across the KFM image of injected negative charges. The slope of potential profile (dV/dx) indicates the localized electric field E_x between adjacent MNCs. Region A represents the MNCs with maximum potential and E_x is minimum, while Region B corresponds to the steepest slope of the potential profile where E_x is maximum. (b) and (c) show the energy band diagrams of adjacent MNCs at Region-A and -B, respectively.

As shown in Fig. 5.15(a), the localized electric field between adjacent MNCs (E_x) can be estimated from the KFM potential profile. For a diameter of the potential distribution with ~ 500 nm and a given NC density of $3 \times 10^{12} \text{ cm}^{-2}$, there are at least ~ 6000 MNCs involved in charge decaying process. The two distinct regions are presented in Fig. 5.15 (a): A and B; where Region 'A' is MNCs with maximum potential and E_x is minimum, while Region 'B' corresponds to the steepest slope of potential profile where E_x is maximum. Figs. 5.15(b) and (c) show the energy band diagrams of adjacent MNCs at Region-A and -B, respectively.

At $t=0$ (*i.e.*, immediately after charge injection), the localized electric field in the vertical direction could be much larger than that in the lateral direction at Region 'A' where the electric field of adjacent MNCs, E_x is minimum (*i.e.*, $E_z \gg E_x$) and the vertical charge loss behavior could be a dominant process during the initial stage of the retention. As a result, the diameter of the potential profile is not expanding laterally though the potential peak decreases with time during the initial charge decay in KFM measurement of SL-MNCs embedded samples. This phenomenon is more easily observed in the sample with a lower inter-dot tunneling probability like Pt sample. Therefore, the Region A (*i.e.*, MNCs with minimum E_x) of the potential profile is the dominant mechanism to be responsible for the vertical charge loss under the conditions such that the localized vertical electric field is high enough to detrapp the electrons from MNCs towards the Si substrate through the tunnel oxide layer with thickness of ~ 5 nm, while Region B (*i.e.*, MNCs with maximum E_x) is responsible for the inter-dot tunneling.

In the later stage of the charge decay ($t > 300$ sec in this case), localized potential decreases with time as more electrons tunnel out of MNCs and asymmetric electric field

presents along lateral and vertical direction ($E_z < E_x$) due to a thicker tunneling oxide than average MNC spacing (*i.e.*, $L_z > L_x$). As a result, the inter-dot tunneling becomes dominant in the later stage of charge decay and the diameter of the potential profile increases with time. Therefore, the two competing mechanisms of lateral and vertical charge diffusion in SL-MNCs embedded gate stack could be defined as shown in Table. 5.1.

TABLE 5.1. SUMMARY OF CHARGE TRANSPORT IN SL-MNCs EMBEDDED GATE STACKS

<i>Inter-dot spacing (L_x) and tunnel oxide thickness (L_z)</i>	<i>Localized potential difference in lateral (ΔV_x) and vertical direction (ΔV_z)</i>	<i>Dominant conduction path during retention</i>	
		<i>Lateral</i>	<i>Vertical</i>
$L_x < L_z$	$\Delta V_x = \Delta V_z$	√	
	$\Delta V_x > \Delta V_z$	√	
	$\Delta V_x < \Delta V_z$	√	√

5.7 Summary

This chapter describes KFM and transient discharging current measurement as a viable approach to detect and monitor the lateral charge diffusion kinetics and vertical charge loss of localized trapped-charges in MNC based flash memory devices. The real time evolution of charge decay is monitored and their charge transport mechanisms are investigated. It is found that the inter-dot tunneling and vertical charge loss through backward tunneling to the substrate are two competing mechanisms in MNC embedded

gate stacks. MNCs with higher work function show a slower lateral charge diffusion, which is favorable for application in scaled memory devices with prolonged retention, even in the case of MNCs in defective dielectric. However, the vertical charge loss during the initial decay period is found to be a trade-off phenomenon. Therefore, next chapter focus on the dual-layer devices with Pt-MNCs as storage nodes since electron de-trapped from one-NC layer could be “re-trapped” in the adjacent NC layer, preventing charge loss and ensuring robust non-volatility.

References

- [1] S. Tiwari, F. Rana, K. Chan, H. Hanafi, C. Wei, and D. Buchanan, "Volatile and non-volatile memories in silicon with nanocrystal storage," in *IEEE Tech. Dig. IEDM 1995*, pp. 521-524.
- [2] B. deBoer, A. Hadipour, M. M. Mandoc, T. vanWoudenberg, and P. W. M. Blom, "Tuning of Metal Work Functions with Self-Assembled Monolayers," *Advanced Materials*, vol. 17, pp. 621-625, 2005.
- [3] J. L. Freeouf and J. M. Woodall, "Schottky barriers: An effective work function model," *Applied Physics Letters*, vol. 39, pp. 727-729, 1981.
- [4] T.G. Ruskell, R. Workman, D. Chen, D. Sarid, S. Dahl, S. Gilbert, "High resolution Fowler–Nordheim field emission maps of thin silicon oxide layers", *Applied Physics Letters*, vol. 68, pp. 93–95, 1996.
- [5] K. Jones, *Asylum Research Kelvin Probe Microscopy User Guide Rev. 662*, 2010.
- [6] J. T. Jones, P. M. Bridger, O. J. Marsh, and T. C. McGill, "Charge storage in $\text{CeO}_2/\text{Si}/\text{CeO}_2/\text{Si}(111)$ structures by electrostatic force microscopy," *Applied Physics Letters*, vol. 75, pp. 1326-1328, 1999.
- [7] D. M. Schaadt, E. T. Yu, S. Sankar, and A. E. Berkowitz, "Charge storage in Co nanoclusters embedded in SiO_2 by scanning force microscopy," *Applied Physics Letters*, vol. 74, pp. 472-474, 1999.
- [8] C. J. Kang, G. H. Buh, S. Lee, C. K. Kim, K. M. Mang, C. Im, and Y. Kuk, "Charge trap dynamics in a SiO_2 layer on Si by scanning capacitance microscopy," *Applied Physics Letters*, vol. 74, pp. 1815-1817, 1999.
- [9] T. Takahashi and S. Ono, "Tip-to-sample distance dependence of an electrostatic force in KFM measurement," *Ultramicroscopy*, vol. 100, pp. 287-292, 2004.
- [10] A. Liscio, V. Palermo, K. Mullen, and P. Samori, "Tip-Sample Interactions in Kelvin Probe Force Microscopy: Quantitative Measurement of the Local Surface Potential," *The Journal of Physical Chemistry C*, vol. 112, pp. 17368-17377, 2008.
- [11] A. Belarni, M. Lamhamdi, P. Pons, L. Boudou, J. Guastavino, Y. Segui, G. Papaioannou, and R. Plana, "Kelvin probe microscopy for reliability investigation

- of RF-MEMS capacitive switches," *Microelectronics Reliability*, vol. 48, pp. 1232-1236, 2008
- [12] C. Guillemot, P. Budau, J. Chevrier, F. Marchi, F. Comin, C. Alandi, F. Bertin, N. Buffet, W. Ch, and P. Mur, "Imaging of stored charges in Si quantum dots by tapping and electrostatic force microscopy," *Europhysics Letters*, vol. 59, p. 566, 2002.
- [13] R. Krishnan, Q. Xie, J. Kulik, X. D. Wang, S. Lu, M. Molinari, Y. Gao, T. D. Krauss, and P. M. Fauchet, "Effect of oxidation on charge localization and transport in a single layer of silicon nanocrystals," *Journal of Applied Physics*, vol. 96, pp. 654-660, 2004.
- [14] C. Y. Ng, T. P. Chen, V. S. W. Lim, and M. S. Tse, "Real time evolution of charge decay characteristics in silicon nanocrystals," in *IEEE Intl. Conference on Semiconductor Electronics* Kuala Lumpur, Malaysia., 2004.
- [15] L.P. Kouwenhoven, C.M. Marcus, P. L. Mceuen, S. Tarucha, R.M. Westervelt, and N.S. Wingreen, 1997, in *Mesoscopic Electron Transport*, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schön, Series E: Applied Sciences (Kluwer Academic, Dordrecht), Vol. 345, pp. 105–214.
- [16] R. Kohlrausch, *Pogg. Ann. Phys. Chem.* , vol. 91, p. 179, 1854.
- [17] M. Cardona, R. V. Chamberlin, and W. Marx, "The history of the stretched exponential function," *Annalen der Physik*, vol. 16, pp. 842-845, 2007.
- [18] G. Williams and D. C. Watts, "Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function," *Transactions of the Faraday Society*, vol. 66, pp. 80-85, 1970.
- [19] Y. Kita and N. Koizumi, "Evaluation of the relaxation intensity from dielectric loss data," *Advances in Molecular Relaxation and Interaction Processes*, vol. 15, pp. 261-272, 1979.
- [20] A. K. Jonscher, *Universal Relaxation Law*. London: Chelsea Dielectrics Press, 1996.
- [21] R. Dianoux, H. J. H. Smilde, F. Marchi, N. Buffet, P. Mur, F. Comin, and J. Chevrier, "Charging dynamics and strong localization of a two-dimensional electron cloud," *Nanotechnology*, vol. 18, p. 325403, 2007.

- [22] Y.-N. Xu and W. Y. Ching, "Self-consistent band structures, charge distributions, and optical-absorption spectra in MgO, α -Al₂O₃, and MgAl₂O₄," *Physical Review B*, vol. 43, pp. 4461-4472, 1991.
- [23] Z. Xu, L. Pantisano, A. Kerber, R. Degraeve, E. Cartier, S. De Gendt, M. Heyns, and G. Groeseneken, "A study of relaxation current in high- κ dielectric stacks," *IEEE Trans. Electron Devices*, vol. 51, pp. 402-408, 2004.
- [24] R. J. John, H. Walter, P. B. Griffin, and J. D. Plummer, "Double-well model of dielectric relaxation current," *Applied Physics Letters*, vol. 84, pp. 3489-3491, 2004.
- [25] A. K. Jonscher, "Dielectric relaxation in solids," *Journal of Physics D: Applied Physics*, vol. 32, p. R57, 1999.
- [26] J. R. Jameson, P. B. Griffin, A. Agah, J. D. Plummer, H. S. Kim, D. V. Taylor, P. C. McIntyre, and W. A. Harrison, "Problems with metal-oxide high- κ dielectrics due to $1/t$ dielectric relaxation current in amorphous materials," in *IEEE Tech. Dig. IEDM 2003*, pp. 4.3.1-4.3.4.
- [27] Y.N. Chen, K. E. J. Goh, X. Wu, Z. Z. Lwin, P.K. Singh, S. Mahapatra, and K. L. Pey, "Temperature-dependent Relaxation Current on Single and Dual layer Pt Metal Nanocrystal-based Al₂O₃/SiO₂ Gate Stack," *Journal of Applied Physics*, 2011 (under reviewed).
- [28] H. Reisinger, G. Steinlesberger, S. Jakschik, M. Gutsche, T. Hecht, M. Leonhard, U. Schroder, H. Seidl, and D. Schumann, "A comparative study of dielectric relaxation losses in alternative dielectrics," in *IEEE Tech. Dig. IEDM 2001*, pp. 12.2.1-12.2.4.

CHAPTER SIX

LOCALIZED CHARGE TRANSPORT IN DUAL-LAYER METAL NANOCRYSTALS EMBEDDED GATE STACKS

6.1 Introduction

As described in Chapter 5, single-layer (SL) MNCs with high work function show small lateral charge diffusion, which is favorable for improved retention. However, the vertical charge loss during initial decay period is a trade-off in memory application. Therefore, in this chapter, dual-layer (DL) MNC structure is proposed and the charge decay mechanisms and retention reliability are explored. It has been reported that the two metal NC-layers provide improved charge storage capacity and retention over the single-layer devices [1-3]. Though NCs are not exactly aligned on top of each other in the DL structure due to the random placement of NCs by self-assembly process, it still leads to an increase in NC density and effective area coverage, resulting in the reported increase in memory window. Moreover, the retention improvement in DL-NC devices is expected as the possibility of de-trapped charges from top NC layer could be re-trapped in the

lower NC layer. Since the DL-MNC structure has shown a high potential for memory applications, a full understanding of its retention performance and charge loss mechanisms is crucial. Moreover, in order to design memory cell with specific retention characteristic, understanding the charge distribution and its impact on the charge loss behavior are crucial for device optimization, reliability predictions and future scaling [4].

In this chapter, the localized charge transport and vertical charge loss mechanisms in DL-MNC memory devices are discussed. The impact of the inter-layer-dielectric (ILD) thickness between the two MNC-layers is investigated in details. The first section describes the study of vertical charge loss mechanisms by retention measurements at elevated temperatures. The charge leakage paths in the DL devices are investigated using gate-bias accelerated retention measurements. Since the electron detrapping should be dependent on the trapped electron distribution in the devices, the charge distributions in two MNC-layers are explored using Kelvin Force Microscopy (KFM) technique. Then, the localized charge transport in DL-MNCs is discussed in details.

6.2 Experiments

The DL-MNC devices were fabricated on a p-type Si substrate with a doping of $\sim 1 \times 10^{15} \text{ cm}^{-3}$. After the surface cleaning by HF, a 50 Å SiO_2 layer as tunnel oxide was thermally grown using an *in-situ* steam generation process in an Applied Materials Centura RTP tool. The MNC formation was realized by blanket Pt thin film (5Å) deposition using physical vapor deposition (PVD) technique and subsequent annealing in a RTP chamber, leading to a random distribution of NC position and size as a result of

dewetting (agglomeration). An optimized PVD deposited Al_2O_3 was used as inter-layer dielectric (ILD). A thin layer of Pt ($5\text{\AA}$) was then deposited again and annealed to form another MNC layer. A $120\text{\AA}$ optimized Al_2O_3 was then deposited and annealed to form the control dielectric. Then, $1000\text{\AA}$ Pt metal was deposited as control gate using shadow mask for macroscopic electrical measurements. In the experiment, the two DL devices were fabricated with a variation in the ILD thickness. More details of the fabrication procedures can be found in Ref. [5]. Fig. 6.1 shows the schematic of a DL-MNC embedded gate stack used in this study.

Figure 6.1. Schematic (not to-scale) of DL-MNC embedded gate stack (a) for KFM characterization and (b) with Pt top electrode for microscopic electrical measurement.

The plan view TEM micrograph in Fig. 6.2 (a) shows well isolated Pt-MNCs in the DL devices. From the cross-sectional view TEM micrographs shown in Figs. 6.2(b) and (c), the average diameter of MNCs is around 3 nm and the two MNCs-layers are clearly separated by a 2 nm and 3 nm thick Al_2O_3 ILD in DL2 and DL3, respectively.

The area coverage of MNCs was observed to be $\sim 26\%$ with a density of $\sim 3.4 \times 10^{12} \text{ cm}^{-2}$ in each MNC layer.

(a)

(b)

(c)

Figure 6.2. (a) Plan view TEM micrograph of Pt-MNCs in DL device after forming the first MNC layer. Cross-sectional TEM micrographs of (b) DL2 and (c) DL3 samples with different ILD thickness.

6.3 Vertical Charge Loss Mechanisms

The microscopic capacitor structure with Pt electrodes was used to measure the vertical charge loss in the DL devices. The temperature and gate-bias accelerated retention measurement approaches were utilized to investigate intrinsic charge loss mechanisms in DL structures.

6.3.1 Temperature Accelerated Retention

To study intrinsic charge loss mechanisms in DL-MNCs structure, the retention measurements were performed on “fresh” devices (*i.e.*, without endurance cycling) after programming at +18V for 100 ms. Fig. 6.3 shows “fresh” to “program” transient of the HFCV curves in DL devices. The charge loss was obtained by measuring flat band voltage shift (ΔV_{FB}) with time at 100 kHz.

Figure 6.3. Sample CV curve showing shift in the flatband voltage from “fresh” to “program” (P) state after applying +18V for 100 ms at room temperature in DL2 (open symbols) and DL3 (closed symbols) devices.

Figure 6.4. (a) Charge retention in DL2 and DL3 devices at temperature ranging from 298-398K. (b) Box plot of the retention time for 5% charge loss.

In Fig. 6.4(a), the charge retention behavior of the DL devices at elevated temperatures ranging from 298 to 398K is presented. Fig.6.4 (b) shows the box plot of the

retention time for a 5% charge loss in the DL2 and DL3 devices at various temperatures, respectively. The results clearly reveal that the charge retention in DL3 is better than that of DL2 at all measured temperatures. It is interesting to note that the charge loss in DL2 was independent of temperatures below 100°C.

6.3.2 Gate Bias Accelerated Retention

To separate the contributions of the tunnel oxide and control oxide in the retention deterioration characteristics, bias-accelerated charge leakage measurements were carried out by applying an external bias to the top electrode during retention. As illustrated in Fig. 6.5, an applied positive gate bias ($+V_g$) accelerates the electron detrapping across the Al_2O_3 control oxide layer, while an applied negative bias ($-V_g$) enhances the electron detrapping towards the SiO_2 tunnel oxide.

Fig. 6.6 shows the retention loss from the program state of DL2 and DL3 devices at various gate biases for 15 mins. Each data point was collected from the individual devices after an application of writing pulse. It is found that the charge loss increased with increasing V_g positively. For $V_g = 0 \sim -4$ V, insignificant charge loss was observed indicating the good quality of SiO_2 in the DL devices which results in a low leakage at small electric fields. For $V_g = -4 \sim -8$ V the slight increase in the retention loss was found due to an increase in the electric field across the tunnel oxide layer. The electron trapping during negative bias retention was found to be negligible by applying a gate voltage between -2 V and -8 V in the fresh devices.

(a)

(b)

Figure 6.5. Schematic energy band diagram of DL device under an application of (a) positive bias and (b) negative bias to the top electrode during retention. An application of a positive bias accelerates the charge leakage through Al_2O_3 control oxide, while a negative bias enhances the carrier detrapping across SiO_2 tunnel oxide.

Figure 6.6. Retention charge loss from the Program state at various gate biases for 15 mins in the DL2 and DL3 devices.

Therefore, the gate bias accelerated charge loss measurements suggest that (1) the DL3 devices have better charge retention than DL2, which is consistent with the temperature accelerated measurements and (2) the primary charge leakage path in the DL devices is through the Al_2O_3 control oxide. Moreover, a similar trend but different amount of charge loss in DL2 and DL3 was observed under a positive bias during retention, *i.e.*, in order to have the same amount of charge loss as DL2, a larger positive bias is required to be applied to DL3 during retention. Therefore, the experimental results from the gate bias accelerated retention measurements clearly convey the presence of higher internal electric field in DL2 as compared to DL3.

Considering the charge leakage path is through the control oxide in the DL devices, the localized electric field F across the Al_2O_3 control oxide can be defined as: $F = F_1 + F_2$, where F_1 is the electric field due to localized trapped charge in the top MNC layer and F_2 is the Coulomb-repulsion field due to the neighboring (bottom) MNC layer.

Assuming a capacitor formed between each MNC and the control gate with the sandwiched Al_2O_3 control oxide as insulator, F_1 can be expressed as [6]

$$F_1 = \frac{Q_{\text{NC}}}{C \cdot t_{\text{co}}} \quad (6.1)$$

where Q_{NC} is the amount of charge stored in the MNC, $C = (\frac{\epsilon_{\text{co}}}{t_{\text{co}}} \cdot A)$ is the capacitance across the control oxide t_{co} and A is the plane-collapsed area πR^2 for a MNC with radius R . Similarly, the Coulomb-repulsion field due to the neighboring NC layer, F_2 is given by the Coulomb's Law:

$$F_2 = \frac{Q_{\text{NC}}'}{4 \cdot \pi \cdot \epsilon_0 \cdot t_{\text{ILD}}^2} \quad (6.2)$$

where Q_{NC}' is the amount of charged stored in the “neighboring” MNC-layer and t_{ILD} is the ILD thickness. Thus, the internal electric field in the DL devices can be defined as

$$F = \frac{Q_{\text{NC}}}{C \cdot t_{\text{co}}} + \frac{Q_{\text{NC}}'}{4 \cdot \pi \cdot \epsilon_0 \cdot t_{\text{ILD}}^2} \quad (6.3)$$

which depends on the amount of charge stored in the top-MNC layer Q_{NC} , its neighboring bottom MNC layer Q_{NC}' , and the ILD thickness between two MNC layers t_{ILD} .

6.3.3 Charge Distribution in Dual-Layer MNCs

To further confirm higher internal electric field observed in DL2 from the gate bias accelerated retention measurement, the charge distribution in two MNC-layers were analyzed by KFM characterization since the distribution of trapped electrons in MNC-

layers contributes to the internal electric field across oxide. The blanket wafer (*i.e.*, the samples without top gate electrode) was used in this study. The charges were injected first over the surface of the dielectric film using a Pt coated conductive tip which was in contact with the surface during the charge injection process. After charging the MNCs, KFM images were acquired in a two-pass procedure [7]. From the potential images, we confirmed that the positive (negative) voltage stress applied to the AFM tip always results in a positively (negatively) trapped charge distributions. This indicates that the charge transfer occurs between the tip and the oxide layer rather than from the Si-substrate which is in good agreement with other literature reports [8-10] and the KFM characterization of SL-MNC devices (see Chapter 5). Therefore, the charge injection in the KFM characterization is always gate (AFM tip) injection, which is opposite from the device programming using microscopic C-V technique where the charge transfer is shown to be substrate injection.

Fig. 6.7(a) shows the potential profiles of the DL2 and DL3 samples upon negative charge injection. The full-width-at-half maximum (FWHM) (*i.e.*, 50% of the peak) and $1/e^2$ width (*i.e.*, 13.5% of the peak) of the DL2 and DL3 samples were extracted from the potential profiles acquired at ~60 sec right after charge injection. Fig. 6.7(b) presents the change in FWHM and $1/e^2$ width with various potential peaks (*i.e.*, various injection voltages). As the injection voltage increases, the diameter of charge cloud at FWHM and $1/e^2$ increases in both DL2 and DL3. It is interesting to note that the rate of increase in the $1/e^2$ width with injection voltages in DL3 is much more significant than that in DL2 (see Fig. 6.7(b)).

(a)

(b)

Figure 6.7. (a) Potential profiles of the DL2 and DL3 samples after negative charge injection. (b) Evolution of FWHM and $1/e^2$ width with various potential peaks (*i.e.*, various injection voltages).

Figure 6.8. Preferred paths of charging in DL-MNCs structure under gate injection. P_1 represents charge diffusion to adjacent MNCs in the same MNC-layer and P_2 corresponds to the charging of bottom layer through top layer MNCs. P_3 shows the possibility of direct charge injection to bottom MNC layer depending on the vertical alignment of MNCs in DL gate stacks.

The physical basis for increasing $1/e^2$ width more rapidly than the FWHM evolution can be attributed to the distance dependence KFM resolution. When the charges are closer to the KFM tip, it provides better lateral resolution of potential profile in KFM. It is worth noting that the injected charges could present in both MNC layers in DL devices. Depending on the charge distribution of two MNC layers, the detected potential profile could change accordingly. Therefore, as illustrated in Fig. 6.8, we propose the preferential path of charge injection model in dual-layer MNCs to explain the

potential profiles observed in the DL2 and DL3 samples using KFM characterization. Under the AFM tip (gate) injection, most of the carriers will be injected to the top MNC layer since the direct tunnel distance to the lower MNC layer is larger. However, the charges injected into the top NC layer can be de-trapped and re-trapped either in the neighbor MNC (P_1) or in the bottom NC layer (P_2) due to the presence of enhanced electric field between NCs during charge injection. Also, it is possible to charge the bottom layer directly from the tip depending upon the position of the bottom layer MNC (P_3). In either case, it is reasonable to assume that the top MNC layer will receive more charges than the bottom MNC layer under gate injection. Similarly, it is expected that more charges will be injected into the bottom MNC layer under the substrate injection mode.

The significant increase in the rate of $1/e^2$ width of the potential profile with the tip injection voltages in DL3 suggests that most of the carriers injected to the top MNC-layer were de-trapped and re-trapped in the adjacent MNCs in the same layer, rather than tunneling to the neighbor MNC-layer ($P_1 \gg P_2$) in DL3 (Fig. 6.9(a)). It depicts that the average MNC spacing is smaller than the ILD thickness of DL3 (<3 nm) and the difference in the charge distribution of two MNC-layers in DL3 is large. On the other hand, a similar trend (slope) of increase in both FWHM and $1/e^2$ width with the injection voltages in DL2 (see Fig. 6.7(b)) suggests the charge carries were most likely to be injected in both top and bottom MNC layers owing to thinner ILD (Fig. 6.9(b)). In this case, the average MNC spacing would be comparable or larger than the ILD thickness of DL2 (≥ 2 nm) and the difference in the injected charge distribution of two MNC-layers in DL2 is less compared to that in DL3.

Figure 6.9. Schematic illustration of charge distribution in two MNC-layers of (a) DL3 and (b) DL2 depending on the ILD thickness under gate (AFM tip) injection.

Assuming the average MNC spacing in the DL devices is to be the same because MNCs were deposited using the same process parameters, the average MNC spacing in the DL devices was estimated to be between 2 and 3 nm, based on the diameter ~ 500 nm potential profiles observed in the KFM characterization. Also, the similar charge distribution of the two MNC-layers in DL2 implies that there are more injected charges in the neighboring MNC layer compared with DL3, which could result in higher internal electric field in DL2 during retention due to a stronger repulsive field F_2 as discussed earlier in equation 6.3.

The KFM technique provides the physical evidence of the difference in the charge distribution of two MNC-layers depending on the ILD thickness. It is found that when the ILD thickness is greater than the MNC spacing (*i.e.*, DL3), fewer charges will be injected

in the neighbor MNC layer due to longer tunneling distance. As a result, the difference in the charge distribution of two MNC layers is larger.

On the other hand, when the ILD thickness is comparable to the MNC spacing (*i.e.*, DL2), both MNC layers could be charged and as a result, the charge distribution of two MNC layers is similar. It provides higher internal electric field across the oxide during retention due to the field induced by the trapped charges in the top-MNC and high repulsive field induced by the thin ILD.

6.3.4 Vertical Charge Loss Mechanisms in Dual-Layer MNCs

As described in Section 6.3.2, the gate-bias accelerated retention measurement indicates that the primary charge leakage path in the DL devices is through Al_2O_3 control oxide. To explain the charge loss through traps presented in the Al_2O_3 control oxide, the multi-phonon-assisted tunneling model [11] was adopted which considers electron-phonon coupling in the oxide. In the multiphonon assisted tunneling, a series of virtual states for the coupled electron-phonon system allows for the transition paths with an “increased” tunnel probability [12]. In the oxides with deep trap centers [11, 13, 14], the pure trap tunneling (horizontal path without phonon assistance) is rather unlikely [13]. Since the deep-level traps present in the Al_2O_3 [15-18], the capture and emission should be essentially a multiphonon process [11, 13, 14], *i.e.*, phonon-induced and accompanied by the emission of some phonons [13]. Therefore, the charge loss through Al_2O_3 control oxide occurs by electron tunneling to different trapping levels presented in the high- κ and the capture and emission rates of these traps depend on the phonon transition probability.

From the charge distribution analysis using the KFM technique, the difference in the retention behaviors of DL2 and DL3 observed under the temperature accelerated retention measurements could be explained by the energy band diagram shown in Fig. 6.10. It illustrates the possible charge loss processes through high- κ control oxide in the DL2 and DL3 devices based on the internal-electric field induced by charge distribution in the two MNC-layers and the distance between two MNC-layers. In this illustration, the charge distribution is considered under the substrate injection in the temperature accelerated retention measurements – where more charges are injected in the bottom MNC-layer and fewer charges are injected in the top MNC-layer especially in the sample with thicker ILD (*i.e.*, DL3).

Figure 6.10. The schematic energy band diagrams according to the charge distribution and internal electric field during retention in (a) DL2 and (b) DL3 devices. Labels 1-4 correspond to the possible charge loss processes through trap-assisted-tunneling, thermally assisted tunneling, thermionic emission and direct tunneling through rectangular barrier, respectively.

In DL2, the high internal electric field across oxide and more charges stored in the top MNC-layer comparing with DL3 could provide an alignment of oxide traps and the trap inside the NCs (see Fig. 6.10(a)). As a result, the discharging might easily occur through a trap-assisted-tunneling. Simultaneously, the thermally detrapped electrons could reach the oxide conduction band through thermally assisted tunneling and/or thermionic emission and are collected by the control gate which becomes dominant at higher temperature ($T > 100^\circ\text{C}$ in this case).

In DL3, the reduced internal electric field across Al_2O_3 and fewer electrons trapped in the top layer under substrate injection could result in a misalignment between the trapping states in the MNCs and the traps presented in the high- κ . Hence, “only” the thermally excited electrons have the high possibility to tunnel into the traps in the high- κ . As a consequence, the resulting charge loss is reduced and it strongly depends on temperature. Moreover, the tunneling probability P from one NC layer to another through rectangular barrier is exponentially proportional to the ILD thickness: $P \propto \exp(-2k_{ox}t_{ILD})$ where k_{ox} is the wave vector in the oxide. Hence, the larger ILD thickness in DL3 of 1 nm significantly reduces the electron tunneling probability from the bottom NC layer to the top NC layer minimizing the charge leakage through control oxide during retention.

The systematic investigation of vertical charge loss and the charge distribution in the DL-MNCs devices indicates that the charge loss in the DL structures is mainly due to the internal electric field induced by trapped electrons which depends on the ILD thickness and MNC spacing. When ILD thickness is comparable to MNCs spacing (*i.e.*, DL2), the charge distribution in two-MNC layers is similar and the internal electric field

induced by charges stored in the MNCs is higher. It leads to the fact that the internal-electric-field assisted tunneling dominates at lower temperatures (below 100°C). In contrast, when the ILD thickness is larger than the average MNC spacing (*i.e.*, DL3), less charge injected in the neighboring MNC layer, resulting in a reduction of both internal electric field and possibility of oxide trap alignment with the traps inside the MNCs. Moreover, the thick ILD reduces the electron tunneling probability from one MNC-layer to another during retention. Hence, it provides a significant improvement of retention especially at room temperature.

6.4 Charge Transport Mechanisms in Dual-Layer MNCs

The charge retention analysis on microscopic capacitor structures by electrical measurement (*such as* Capacitance-Voltage technique, Current-Voltage technique, *etc.*,) accounts for the charge leakage only in the vertical direction. In order to analyze the localized charge transport in both lateral and vertical directions during retention, the KFM approach is used.

6.4.1 KFM Analysis of Charge Decaying Characteristics

Fig. 6.11 presents the change in the localized potential peak and FWHM with time in the DL3 samples after charge injection with various injection voltages (-4 to -10V). The arrow indicates the trend of the amount of the initial injected charge (from smallest to largest). When lateral charge diffusion dominates during retention, the

diameter (FWHM) of the charge cloud is supposed to become larger with time. However, it can be seen that the change in FWHM is negligible during the initial stages of retention (<1000sec) (see Fig. 6.11(b)). We named this behavior of charge decay as “*vertical charge loss*” since the FWHM was unchanged though the potential peak was decreasing with time. This behavior was found to be more prominent in samples with higher injection voltages.

In contrast, the decrease in the potential peak and increase in the FWHM with retention time in the DL2 samples were observed, as shown in Fig. 6.12. It is noted that the “*vertical charge loss*” cannot be observed in the DL2 samples in all the injection voltages (-4 to -10V).

Figure 6.11. The change in (a) potential peak and (b) FWHM as a function of retention time in the DL3 samples with various injection voltages. The right column of (b) presents the zoom-in view of FWHM evolution with time. The arrow shows the trend of the initial amount of the injected charge (from smallest to largest).

(a)

(b)

Figure 6.12. The change in (a) potential peak and (b) FWHM as a function of retention time in the DL2 samples with various injection voltages. The arrow shows the trend of the initial amount of the injected charge (from smallest to largest).

We observed the similar behavior of the charge decay (unchanged FWHM with decrease in the potential peak during the initial stage of retention) as DL3 in the SL-Pt-MNCs devices, owing to the tunneling of the stored charges in the NCs back to the substrate (see Chapter 5). Since there is another MNC-layer existed in the DL structure,

there are two possibilities of observing unchanged FWHM with decreased potential peak in the DL3 devices: (1) diffusion of stored charges in the top-MNC layer to the neighboring (lower) MNC layer and/or (2) the charge tunneling backwards to substrate. In the SL-Pt-MNCs, the vertical charge loss towards Si-substrate is confirmed by the transient discharging current (relaxation current) measurement at elevated temperatures. Therefore, a similar experiment was repeated for the DL3 devices at temperatures ranging from 300 to 450K (Fig. 6.13(a)) and the change in the slope n from the fitting of $I_{\text{relax}}=at^{-n}$ [19] was monitored. Fig. 6.13(b) shows the change in the slope n at various temperatures. Unlike in the SL-Pt-MNCs (Chapter 5), it is found that the slope n tends to decrease with increasing temperature. As the temperature increases, the electrons at the topmost energy level in NCs gains enough energy to diffuse to the adjacent MNCs in the same layer and/or in the neighboring layer, and/or tunnel back to Si substrate. Therefore, the smaller I_{relax} slope at higher temperature indicates that the detrapping of the stored charges in NCs to the substrate is negligible in the DL devices.

Moreover, another evidence to exclude charge leakage path through the Si substrate in the DL devices can be found from the vertical charge loss study, as discussed in Section 6.3. It clearly indicates DL3 has better retention characteristic than DL2. However, no evidence of “*vertical charge loss*” is found in the charge decay of the DL2 devices (see Fig. 6.12). Hence, the negligible change in the FWHM with decreased potential peak found in the KFM characterization of the DL3 samples is not due to the charge tunneling backward to the Si-substrate, rather the inter-layer tunneling of the stored charges in the top MNC-layer to the neighboring layer is a possible reason.

Therefore, the vertical charge loss through tunneling backward to substrate is ignored in the KFM charge decay analysis of the DL structures.

(a)

(b)

Figure 6.13. (a) Transient discharging (relaxation) current measurement in the DL3 sample at temperature ranging from 300 to 450 K. (b) The change in the magnitude of slope n with temperature from the curve fitting.

Since the localized charge decay observed in the KFM analysis could account for both lateral and vertical charge loss, the possible charge decay paths in the DL-MNC structures can be identified as follows (Fig. 6.14):

- (1) Inter-dot tunneling (towards adjacent MNCs in the same layer) (P_x),
- (2) Inter-layer tunneling (towards MNCs in the neighboring layer) (P_{z-1}), and
- (3) Vertical charge loss through tunneling backward to Si-substrate (P_{z-2})

Based on the KFM analysis, P_{z-2} could be ignored in the DL structure. Thus, the KFM charge decay observed in the DL devices is dominated by inter-dot tunneling (P_x) and inter-layer tunneling (P_{z-1}).

Figure 6.14. Illustration of possible charge decay paths during retention in DL-MNC embedded gate stacks: P_x corresponds to inter-dot tunneling to MNCs in the same layer; while, P_{z-1} and P_{z-2} refer to inter-layer tunneling through ILD and vertical charge loss through SiO₂, respectively.

6.4.2 Charge Transport in Dual-Layer MNCs

As illustrated in Fig. 6.15, a physical model is proposed to explain the two dominant charge transport mechanisms of charge decay observed in the DL-MNCs embedded devices. Similar to the charge transport mechanism observed in SL-MNCs (see Section 5.6), the probability of the inter-dot tunneling and inter-layer tunneling in DL-MNCs depends on the difference in the localized electric field in lateral E_x and vertical E_z directions. The localized electric field in between any two adjacent MNCs relates to their localized potential difference (ΔV) and the inter-dielectric spacing (L_x and t_{ILD}), as shown in Fig. 6.15.

Figure 6.15. Schematic illustration of unequal charge distribution in two MNC-layers upon gate injection. E_x and E_z are corresponding to the change in localized electric field in lateral and vertical directions. L_x and t_{ILD} represent inter-dot spacing and inter-layer dielectric thickness between two MNC-layers, respectively.

From the study of the charge distribution in two MNC-layers using KFM, the difference in the charge distribution of two MNC-layers in the DL3 samples is more prominent than that in the DL2 samples. That is, more charges injected in the top MNC-layer and significant less charge injected in the lower MNC-layer in the DL3 samples under AFM tip injection. As a result, E_z could be much larger than E_x ($E_z \gg E_x$) during the initial stages of retention which leads to inter-layer tunneling towards adjacent MNC layer. Hence, we observed that FWHM remained almost unchanged though the potential peak was decreasing with time during the initial stages of retention ($t < 1000$ sec) in the DL3 samples. Since it is more obvious in the sample with increasing injected amount of charge as shown in Fig. 6.11(b), it suggests that the larger difference in the charge distribution of the two MNC-layers leads to the more inter-layer tunneling. Also, it is evident that the charge de-trapped from the charged MNCs could be re-trapped in the adjacent MNC layer in the DL-MNCs.

In the DL2 samples, FWHM increases with time throughout the charge decay measurement (60-5000sec) (Fig. 6.9(b)). Since the charge distributions in the two MNC layers in DL2 are comparable as discussed in Section 6.3, it is expected that $E_x \geq E_z$ and the lateral charge diffusion dominates the charge transport mechanism. It is found that the dominant charge transport mechanisms of DL2 and DL3 differ though the only variation is in the ILD thickness. In DL2, the inter-dot tunneling (lateral charge diffusion) is found to be a dominant mechanism. While, the inter-layer tunneling dominates during the initial stages of retention in DL3 which results in the inter-dot tunneling to slow down in the later stages of retention.

(a)

(b)

Figure 6.16 (a) Stretched exponential fitting to the charge decay of DL2 and DL3 samples after various negative charge injection (-4 to -10V) for 10 sec. Solid symbols correspond to charge decay of DL3 and open symbols represent charge decay of DL2. (b) Box plot of the stretched exponent β in the DL3 and DL2 samples.

Fig. 6.16 compares the time dependence of normalized potential peak in the DL2 and DL3 devices under various injection voltage (-4 to -10V). The solid lines present the stretched exponential fitting [20-22]: $V(t)=V_0.\exp(-t/\tau)^\beta$ to the charge decay and Fig.

6.16(b) indicates the range of fitting parameter β in DL3 and DL2 samples. It can be seen that the localized charge decay using KFM technique also indicates the better retention characteristic of DL3 compared to DL2.

The real time evolution of the localized charge transport in DL-MNCs indicates the evidence of inter-dot tunneling and inter-layer tunneling during retention. It strongly depends on the charge distribution of two MNC layers which relates to the MNC spacing, the ILD thickness and the localized electric field which can be summarized as shown in Table 6.1.

TABLE 6.1. SUMMARY OF CHARGE TRANSPORT IN DL-MNCs EMBEDDED GATE STACKS

<i>Inter-dot spacing (L_x) and Inter-layer dielectric (ILD)</i>	<i>Charge Distribution in two MNC-layers (KFM measurement)</i>	<i>Dominant conduction path during charge injection (KFM measurement)</i>		<i>Dominant conduction path during retention (KFM measurement)</i>		<i>Self-repulsive field induced vertical charge loss (Macroscopic C-V measurement)</i>
		<i>Inter-layer tunneling</i>	<i>Inter-dot tunneling</i>	<i>Inter-layer tunneling</i>	<i>Inter-dot tunneling</i>	
$L_x > \text{ILD}$ (e.g. DL2)	Comparable	√	√		√	Relatively Strong
$L_x < \text{ILD}$ (e.g. DL3)	Unequal		√	√		Relatively Weak

6.5 Summary

The localized charge transport behavior in DL-MNCs embedded gate stacks was directly imaged by KFM and the charge transport mechanisms in complex two-dimensional MNC-networks were identified. The nano-scale characterization of charge decay and macroscopic capacitor based retention measurements in DL-MNCs suggest the importance of spacing between adjacent MNCs and the ILD thickness between the two layers of MNCs on the lateral charge diffusion and retention characteristics of NC devices. It is found that the internal electric field induced by trapped electrons in the two MNC-layers is responsible for the vertical charge loss of the DL devices during retention. Our experimental results demonstrate that

- (1) when the ILD thickness is comparable to the average MNC spacing, the charge distribution in the two MNC layers is similar and the internal electric field induced by charges stored in the MNCs is higher, leading to internal electric field assisted tunneling dominating at lower temperatures (below 100°C in this case).
- (2) when the ILD thickness is larger than the average MNC spacing, less charges are injected in the neighboring MNC layer, resulting in a reduction of both internal electric field and possibility of oxide traps alignment with the traps inside the MNCs. Moreover, a thick ILD reduces the electron tunneling probability from one MNC-layer to another during retention.

Therefore, our findings suggest that an optimized DL-MNCs embedded memory cell could be achieved by defining the ILD thickness larger than the average MNC-spacing for enhancement of the retention ability in the MNC embedded gate stacks. It implies the possibility of reducing the MNC spacing in the DL structure of scaled memory devices by controlling the thickness of ILD. Moreover, the real time evolution of the localized charge transport in DL devices shows inter-layer tunneling behavior in DL-MNCs, providing an evidence of detrapped electron from one MNC-layer could be re-trapped in the adjacent MNC layer. This could be a physical reason of longer retention observed in the DL-MNC devices comparing with the SL-MNC structures. However, it could not be true for all the DL-NC devices, since our experimental results show that the retention and charge transport characteristics of DL devices strongly depend on the device structures (such as ILD thickness, MNC spacing) as well.

References

- [1] S. K. Samanta, P. K. Singh, Y. Won Jong, G. Samudra, Y. Yee-Chia, L. K. Bera, and N. Balasubramanian, "Enhancement of memory window in short channel non-volatile memory devices using double layer tungsten nanocrystals," in *IEEE Tech. Dig. IEDM 2005*, pp. 170-173.
- [2] P. K. Singh, G. Bisht, M. Sivatheja, C. Sandhya, G. Mukhopadhyay, S. Mahapatra, R. Hofmann, K. Singh, and N. Krishna, "Reliability of single and dual Layer Pt nanocrystal devices for NAND flash applications: A 2-region model for endurance defect generation," in *Proceedings of the Intl. Reliab. Phys. Symp.*, 2009, pp. 301-306.
- [3] R. Ohba, Y. Mitani, N. Sugiyama, and S. Fujita, "Impact of stoichiometry control in double junction memory on future scaling," in *IEEE Tech. Dig. IEDM 2004*, pp. 897-900.
- [4] A. Padovani, L. Larcher, P. Pavan, L. Avital, I. Bloom, and B. Eitan, "Id-Vgs-Based Tools to Profile Charge Distributions on NROM Memory Devices," *IEEE Trans. on Device and Materials Reliability*, vol. 7, pp. 97-104, 2007.
- [5] P. Singh, C. Sandhya, K. Auluck, G. Bisht, M. Sivatheja, G. Mukhopadhyay, S. Mahapatra, and R. Hofmann, "Applicability of dual layer metal nanocrystal flash memory for NAND 2 or 3-bit/cell operation: Understanding the anomalous breakdown and optimization of P/E conditions," in *Proceedings of the Intl. Reliab. Phys. Symp.*, 2010, pp. 981-987.
- [6] C. Buseret, A. Souifi, T. Baron, G. Guillot, F. Martin, M. N. Semeria, and J. Gautier, "Discharge mechanisms modeling in LPCVD silicon nanocrystals using C-V and capacitance transient techniques" *Superlattices and Microstructures*, vol. 28, pp. 493-500, 2000.
- [7] K. Jones, *Asylum Research Kelvin Probe Microscopy User Guide*, Rev. 662, 2010.
- [8] J. T. Jones, P. M. Bridger, O. J. Marsh, and T. C. McGill, "Charge storage in $\text{CeO}_2/\text{Si}/\text{CeO}_2/\text{Si}(111)$ structures by electrostatic force microscopy," *Applied Physics Letters*, vol. 75, pp. 1326-1328, 1999.

- [9] C. J. Kang, G. H. Buh, S. Lee, C. K. Kim, K. M. Mang, C. Im, and Y. Kuk, "Charge trap dynamics in a SiO₂ layer on Si by scanning capacitance microscopy," *Applied Physics Letters*, vol. 74, pp. 1815-1817, 1999.
- [10] D. M. Schaadt, E. T. Yu, S. Sankar, and A. E. Berkowitz, "Charge storage in Co nanoclusters embedded in SiO₂ by scanning force microscopy," *Applied Physics Letters*, vol. 74, pp. 472-474, 1999.
- [11] F. Schuler, R. Degraeve, P. Hendrickx, and D. Wellekens, "Physical description of anomalous charge loss in floating gate based NVM's and identification of its dominant parameter," in *Proceedings of the Intl. Reliab. Phys. Symp.*, 2002, pp. 26-33.
- [12] A. Schenk, in *Advanced Physical Models for Silicon Device Simulation* edited by S. Selberherr (Springer-Verlag Wien New York 1998), pp. 294-315.
- [13] A. Schenk, "A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon," *Solid-State Electronics*, vol. 35, pp. 1585-1596, 1992.
- [14] P. C. Sercel, "Multiphonon-assisted tunneling through deep levels: A rapid energy-relaxation mechanism in nonideal quantum-dot heterostructures," *Physical Review B*, vol. 51, pp. 14532-14541, 1995.
- [15] T. Sugizaki, M. Kobayashi, M. Ishidao, H. Minakata, M. Yamaguchi, Y. Tamura, Y. Sugiyama, T. Nakanishi, and H. Tanaka, "Novel multi-bit SONOS type flash memory using a high-k charge trapping layer," in *IEEE VLSI Tech. Dig.*, 2003, pp. 27-28.
- [16] S. Maikap, P. J. Tzeng, T. Y. Wang, H. Y. Lee, C. H. Lin, C. C. Wang, L. S. Lee, J. R. Yang, and M. J. Tsai, "HfO₂/HfAlO/HfO₂ Nanolaminate Charge Trapping Layers for High-Performance Nonvolatile Memory Device Applications," *Japanese Journal of Applied Physics*, vol. 46, pp. 1803-1807, 2007.
- [17] B. Govoreanu, D. Wellekens, L. Haspeslagh, J. De Vos, and J. Van Houdt, "Investigation of the low-field leakage through high-k interpoly dielectric stacks and its impact on nonvolatile memory data retention," in *IEEE Tech. Dig. IEDM 2006*, pp. 1-4.
- [18] Y. N. Novikov, A. V. Vishnyakov, V. A. Gritsenko, K. A. Nasyrov, and H. Wong, "Modeling the charge transport mechanism in amorphous Al₂O₃ with

- multiphonon trap ionization effect," *Microelectronics Reliability*, vol. 50, pp. 207-210.
- [19] A. K. Jonscher, "Dielectric relaxation in solids," *Journal of Physics D: Applied Physics*, vol. 32, p. R57, 1999.
- [20] R. Kohlrausch, *Pogg. Ann. Phys. Chem.* , vol. 91, p. 179, 1854.
- [21] G. Williams and D. C. Watts, "Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function," *Transactions of the Faraday Society*, vol. 66, pp. 80-85, 1970.
- [22] M. Cardona, R. V. Chamberlin, and W. Marx, "The history of the stretched exponential function," *Annalen der Physik*, vol. 16, pp. 842-845, 2007.

CHAPTER SEVEN

CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion

This thesis reports extensively on the degradation and reliability of metal nanocrystal (MNC)-embedded $\text{Al}_2\text{O}_3/\text{SiO}_2$ gate stacks, which are promising candidates for future flash memory applications. The major milestones achieved in this project can be summarized as follows:

- i. By inducing a breakdown path in one dielectric layer, the charging and discharging phenomenon of the MNCs and leakage mechanism in the degraded dielectric gate stacks are found to be strongly dependent on the lateral charge tunneling/hopping among the MNCs. We have found that the localized breakdown not only affects the charge holding capability of the affected MNCs, but also provides a leakage path for the charges stored in the surrounding MNCs. Thus, from the dielectric degradation study of the MNC with a number density of $3 \times 10^{12} \text{ cm}^{-2}$, we conclude that the discharging of MNCs via the breakdown path is not a localized phenomenon.

- ii. A methodology to detect the individual dielectric layer breakdown in any bi-layered MNC system was established. The individual dielectric layer breakdown is differentiated from the post-breakdown electrical signatures of the conduction path formed in the MNC-embedded dielectric layer and tunneling oxide layer. With this approach, the Weibull slope β , which is a key indicator of the number of traps needed for percolation breakdown in SiO_2 and MNC-embedded high- κ was explored. The average trap size in MNC embedded Al_2O_3 was found to be much larger (>1.71 nm) than that in SiO_2 (0.8 nm), which could be due to presence of the grain boundaries in Al_2O_3 film and embedded MNCs.
- iii. Real time monitoring of lateral charge diffusion phenomenon in MNC-embedded gate stacks was carried out using Kelvin Force Microscopy. A comparative study was performed on the MNCs with different work functions to investigate the localized charge transport and the lateral charge diffusion phenomenon. With the transient discharging current measurement at elevated temperatures and KFM characterization, we found that the inter-dot tunneling and vertical charge loss through backward tunneling to the substrate are the two competing mechanisms in the MNC-embedded gate stacks. MNCs with a higher work function show slower lateral charge diffusion, which is favorable for application in scaled memory devices with prolonged retention. However, the vertical charge loss during the initial decay period was observed in these samples.
- iv. The charge transport mechanisms in dual layer (DL)-MNC embedded gate stacks were identified. The real time evolution of localized charge transport in DL devices shows the inter-layer tunneling behavior in DL-MNCs, providing a

physical evidence of detrapped electrons from one MNC-layer could be re-trapped in the other MNC layer during retention. It is found that the internal electric field induced by trapped electrons in two MNC-layers is responsible for the vertical charge loss of DL devices during retention. The charge transport and the charge loss mechanisms in DL-MNC embedded devices are found to be strongly dependent on the spacing between adjacent MNCs and interlayer dielectric (ILD) thickness between the two layers of MNCs. Our results demonstrate that ILD thickness should be larger than the average MNC spacing in the optimized DL-MNC structures so that it could

- (1) provide re-distribution of charges (inter-layer tunneling) to the other MNC layer during retention (hence it slows down the lateral charge diffusion rate which is favorable to minimize defect related charge loss),
- (2) reduce self-repulsive field induced vertical charge loss which depends on both the ILD thickness and charge distribution of two-MNC layers (hence it provides a significant improvement of retention time), and
- (3) provide feasibility of reducing the MNC spacing in DL structure by controlling the ILD thickness, which is a bottleneck of scaling down the single-layer NC-based memory where MNC density and spacing could not be compromised to achieve both large memory window and negligible lateral charge diffusion.

7.2 Recommendations

We believe that this project could be further extended to realize an optimized MNC-based memory. Some of the possible directions for future investigations and studies are given as follows:

- Though the non-localized nature of MNC discharging through a dielectric breakdown path in degraded dielectric gate stack has been uncovered in this work, it can be further extended/proven by a powerful nano-characterization technique, via the Kelvin Force Microscopy (KFM) in an ultra-high vacuum environment. The conventional Conductive Atomic Force Microscopy (CAFM) technique could not be realized here unless tunnel oxide thickness is reduced below 2 nm, since the CAFM technique is based on the current feedback mechanism. With the KFM attached with high voltage kit, the localized dielectric breakdown spot could be induced and the charge diffusion kinetic via the dielectric breakdown path could be monitored in real time. Also, the activation energy of localized charge transfer in NC embedded gate stacks can be investigated using the KFM equipped with environmental chamber and temperature control.
- Using KFM/EFM techniques, the lateral interaction between adjacent MNCs could be further analyzed using the various parameters such as number density, dielectric spacing, diameters, *etc.*
- Applying the approaches presented in this work such as inducing dielectric degradation, breakdown layer detection, nanoscale characterization using KFM,

the influence of MNC distribution could be taken into account to study the current transport/charge diffusion/leakage and device degradation.

- In this work, we studied the influences of MNC spacing and ILD thickness on the performance of DL-MNC devices. It is found that ILD thickness should be relatively larger than the MNC spacing for a prolonged retention. It suggests the possibility of reducing the MNC spacing in the DL devices without affecting the retention ability. Therefore, the minimum MNC spacing allowed in the DL structures should be further explored which will be of the utmost importance for the application of MNC-based memory for future scaling.
- There is a wide range of MNC applications in memory devices. For instance, it serves as a discrete storage node in floating gate flash memories. Moreover, it can also be embedded in RRAM to reduce variation of switching voltage (SET/RESET). Therefore, the physical mechanisms governing dielectric breakdown and SET/RESET formation of conduction path in MNC-embedded RRAM would be of interest. For instance, the real time evolution of dielectric breakdown path formation in NC-embedded gate stack could be studied using an *in-situ* TEM and a high resolution TEM, as well as chemical analysis using EELS/EDS.

PUBLICATIONS

- [1] **Z. Z. Lwin**, K. L. Pey, Q. Zhang, M. Bosman, Q. Liu, C. L. Gan, P. K. Singh, and S. Mahapatra, "Study of Charge Distribution and Charge Loss in Dual-layer Metal-Nanocrystal-Embedded High- κ /SiO₂ Gate Stack", *Applied Physics Letters*, vol. 100, p. 193109 (2012).
- [2] **Z. Z. Lwin**, K. L. Pey, Q. Zhang, M. Bosman, Q. Liu, C. L. Gan, P. K. Singh, and S. Mahapatra, "Charge Transport in Metal-Nanocrystal-Embedded Gate Stacks", in *Institute of Physics Singapore (IPS Meeting)*, Singapore (2012).
- [3] **Z. Z. Lwin**, K. L. Pey, C. Liu, Q. Liu, Q. Zhang, Y. N. Chen, P. K. Singh, and S. Mahapatra, "Localized Charge Trapping and Lateral Charge Diffusion in Metal Nanocrystal-embedded High- κ /SiO₂ Gate Stack", *Applied Physics Letters*, vol. 99, p. 222102 (2011).
- [4] **Z. Z. Lwin**, K. L. Pey, N. Raghavan, Y. N. Chen, S. Mahapatra, "New Leakage Behavior and Unique Dielectric Breakdown Layer Detection in Metal Nanocrystals Embedded Dual Layer Gate Stack", *IEEE Electron Device Letter*, vol. 32, pp. 800-802 (2011).
- [5] Y. N. Chen, K. L. Pey, K. E. J. Goh, **Z. Z. Lwin**, P. Singh, and S. Mahapatra, "Study of Automatic Recovery on the Metal Nanocrystal-based Al₂O₃/SiO₂ Gate Stack," *Applied Physics Letters*, vol. 98, pp. 083504 (2011).

- [6] Y. N. Chen, K. L. Pey, K. E. J. Goh, **Z. Z. Lwin**, P. Singh, and S. Mahapatra, " Study of the Charge Leakage of Dual Layer Pt Metal Nanocrystal-based high- κ /SiO₂ Flash Memory Cell - A Relaxation Current Point of View", in *18th International Symposium on the Physical and Failure Analysis on integrated circuits (IPFA)*, Incheon, Korea, pp. 38-41 (2011).
- [7] K. L. Pey, N. Raghavan, X. Wu, W. Liu, X. Li, M. Bosman, K. Shubhakar, **Z. Z. Lwin**, Y.N. Chen, H. Qin, and T. Kauerauf, "Physical Analysis of Breakdown in High- κ /Metal gate Stacks using TEM/EELS and STM for Reliability Enhancement (invited)," in *7th Conference on Insulating Films on Semiconductors (INFOS)*, Grenoble – France, pp. 1365-1372 (2011).
- [8] **Z. Z. Lwin**, K. L. Pey, Y. N. Chen, P. K. Singh, and S. Mahapatra, "Charging and Discharging Characteristics of Metal nanocrystals in Degraded Dielectric Stacks," in *IEEE Proc. Int. Reliab. Phys. Symp. (IRPS)*, California, pp. 89-93 (2010).
- [9] Y. N. Chen, K. L. Pey, K. E. J. Goh, **Z. Z. Lwin**, P. K. Singh, and S. Mahapatra, "Tri-level Resistive Switching in Metal Nanocrystal-based Al₂O₃/SiO₂ Gate Stack," *IEEE Trans. Electron Devices*, vol. 99, pp. 1-5 (2010).
- [10] Y. N. Chen, K. L. Pey, K. E. J. Goh, **Z. Z. Lwin**, P. K. Singh, S. Mahapatra, Q. X. Wang and J. Zhu, "Post-Breakdown Recoverable Metal Nanocrystal-based Al₂O₃/SiO₂ Gate Stack for Non-volatile Memory," in *International Conference on Solid State Devices and Materials(SSDM)*, Japan (2010).