INTEGRATION OF CARBON AND SILICON BASED NANOELECTRONICS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Udayan Ganguly
August 2006
INTEGRATION OF CARBON AND SILICON BASED NANOELECTRONICS

Udayan Ganguly, Ph. D.
Cornell University 2006

Nonvolatile memory technology has shown tremendous technological progress in the recent years. With the need for every higher memory density, the EEPROM structure has been subjected to aggressive scaling. Currently 65 nm devices are hitting production where the floating gate has been replaced by nitride traps for continued scaling. However the operational voltage has not scaled as aggressively. In this dissertation, various different structural variations of the EEPROM have been explored from experimental and theoretical perspectives. The electrostatics of nanocrystal memories with both metal and semiconductor nanocrystals have been analytically modeled to demonstrate enhanced field asymmetry in the path of least action in tunneling oxide. The larger polarization of the metal nanocrystal versus a semiconductor nanocrystal has been used to explain faster programming for metal nanocrystal memories compared to the semiconductor nanocrystal variety. The analytical model provides design intuition unlike numerical models for structural optimization. To provide a solution for the size-dependent variation of coulomb blockade self-assembled nanocrystals, molecules have been suggested as nano-floating gates in non-volatile memory. Carbon molecules like fullerenes (C60) have been integrated in the MOS gate stack. Charge injection into molecular orbital has been observed as repeatable steps in electron injection versus charging voltage data to demonstrate a successful molecular interface with CMOS. Finally, an ultra-narrow channel memory device fashioned out of self–assembled carbon nanotube (CNT)
channel with self aligned metal nanocrystals is proposed as an alternative to memory scaling. The device demonstrates large memory window and single-electron sensitivity. The single electron sensitivity at room temperature is confirmed by analysis of the transport in the CNT channel using non-equilibrium Green’s function (NEGF) formalism. The large number of charges stored in the nanocrystal at sub-5V gate bias for the 100 nm gate stack is demonstrated experimentally and explained from the electrostatic analysis of the CNT-nanocrystal memory. The ability to store large number of charges per nanocrystal and the ability of sense each of these charge indicates the possibility of a multi-level memory, a way of enhancing functional density while relaxing conventional memory scaling constraints.
BIOGRAPHICAL SKETCH

Udayan Ganguly opened his lungs to an early morning in Calcutta, India on April 25th, 1977. Born to a school teacher and a musician/singer, his world was soon propped up on piles of storybooks and long-play records. Learning to read, holding a tune and playing soccer in puddles of tepid water (inhabited by tadpoles) with the para (local) children of the pre-globalization “City of Joy” made a lasting impression on him. He bungled the spelling of ‘rubber’ twelve times in a row using all but the right combination of the vowels, before Father Prefect of St. Lawrence High School realized that the boy could be admitted to the Jesuit school in pursuit of scholastic/general excellence. Udayan made friends quickly with all. His thespian talents were thwarted early in life when he was cast in Gulliver’s Travels as the boat. He made the most out of the part, invented a storm and cast his classmates sitting on him to the mercy of the angry seas. Academically he was interested in everything - from language to science, math and art. However when the time came to choose a stream, he was intimidated by the tall tomes of artistic wisdom and settled for science. After all, life was too short for a long read and when one had the option of reading a formula or a pithy section, why belabor the enlightenment?

The next battle at hand was the Joint Entrance Exam to the Indian Institutes of Technology (IIT). Armed with a keen eye for physics, some mathematics and the ability to solve chemistry problems that employed the unitary method, Udayan took the plunge and without much fanfare, the hurdle was passed. The gates of IIT, beyond which high-technology and the wilderness of the national park coexisted in a strange harmony, gleamed in welcome. IIT Madras was a different beast. Udayan got a ‘C’ despite his best efforts and with an 82/100 score in chemistry as he was RG’ed (relative graded) out by hordes of newly-arrived despo’s (desperately ambitious persons). With academic pressure of this magnitude, different people react differently.
While some get stuffed further into their hostel rooms, Udayan’s life started blooming outside of it, delocalizing over the entire campus. He walked the tight-rope curricular (academics) and extra-curricular (soccer, music, arts and design and relentless ruminations) activities and secured the realizations that have shaped his outlook to life and beyond. It is also here that he discovered his wife Aditi during such an extra-curricular moment.

IIT provided the spring board to pursue graduate studies at Cornell University. After a brief two and a half years with the late Prof. Krusius, Udayan was ready to answer his ‘true’ calling- research in small devices where you could see electrons jump around. The caller came in the form of Prof. Edwin Kan. Three years of interesting course work, long discussions with friends (graduate students and professors), days of fabrication and measurements, a few ‘eurekas’ and a truckloads of ‘dammits’ has culminated in the portal out of graduate life. After being distinctly uncomfortable with the idea of joining the industry, Udayan was offered a post doctoral fellowship at NASA Ames Research Center which he plans to do for the next couple of years. He is smugly satisfied by the way his life has turned out so far…
To my parents, Indrani and Kamal, my sister, Debrupa, and my wife Aditi
ACKNOWLEDGMENTS

A dissertation is a product of perseverance and a fairly large chunk of a lifetime. It is unimaginable that such an effort can be made without the support of a host of people. I knew when I was making friends and being blessed with so many favors that there is no free lunch. It would be a grueling list of acknowledgements. First and foremost, I would like to extend my gratitude to my advisor, Prof. Edwin Kan, for being himself- generous, understanding, sharp and supportive and for being a steadfast friend to me. My collaborators come next. Dr. Yuegang Zhang, Intel, Prof. Jing Guo, University of Florida at Gainesville, Prof. Paul McIntyre, Stanford, Prof Paul McEuen, Cornell, are thanked for the encouragement and the opportunity to share knowledge and expertise. I would like to thank Prof. Thompson, Prof. Malliaras, Prof. Ast, Prof. Rana, Prof. Wiesner and Prof. Umbach for all the discussions, suggestions and criticism both scientific and career-wise. Prof. Krusius is thanked for making me strong, critical and calm and Prof. Pollock for making me feel secure after the shock of Prof. Krusius’ passing. Amongst fellow graduate students, I have a long list of those who have kept me challenged intellectually, lent me a hand under experimental exigencies and kept me generally entertained: Venkat Narayanan, Blake Jacquot, Chungho Lee, Tuo-Hung Hou, Jami Meteer, Jinsook Kim, Weiping Ni, Gen Pei, Yu-min Nick Shen, Myongseob Kim, Jaegoo Lee, Zengtao Liu, Anirudh Gorur-Seetharam, Pingshan Wang, Ava Wang, Nini Munoz, Felix Lee (Kan Group); Raghav Sreenivasan (McIntyre Group, Stanford); Herman Lopez, Shida Tan, Robert Chen, Michael Stewart (Intel Research); Swaroop Kommera, Kevin Dezfulian, Wayne Woods, Lin Hao (Krusius Group); Aravind Kumar, Chris Liu, Helena Silva, Ali Gokirmak, Lei Xue (Tiwari Group); Amit Pharkya (Eastman Group); Shankar Radhakrishnan, Rajesh Duggirala, Khurshad Araz, Serhan Ardanuc, Hui Li, Alper Bozkurt, Abhishek Ramkumar, Xi Chen (Lal Group); Faisal Ahmed, Jahan Dawlaty
(Rana Group) Vera Sazonova, Sami Rosenblatt, Markus Brink, Ken Boznick, Xinjian Zhou, Jun Zhu (McEuen Group); Kirill Bolotin, Ferdinand Kuemmeth, Abhay Pasupathy (Ralph Group); Phani Kumar, Huaqiang Liu, Ho-Young Cha, Goutam Koley, Lori Lepak (Spencer Group) Anand Pappu (Apsel Group);. I would also like to thank Mandy Esch, Daron Westly, Gregory Simelgor, Garry Bordonnaro, Jerry Drumheller, Phil Infante, Dan Woodie and Sam Wright, Kathy Springer and everyone else at Cornell Nanofabrication Facility and Jonathan Shu (Cornell Nanoscale Systems) for all the help.

A big cheer goes out to all my friends for the innumerable good-times: Anirban, Rohit, Kuntal, and Sovan – otherwise known as the Mamoos; Sharmistha, Bhargavi, Kusumananchi, Shankar, Janani, Phani, Sowmya, Duggi, Praveena, Robin, Sharvari, Saikat, and and Zia; and my room mates Swagato, Arijit, and Abhishek Nadamani. A lot of love is in order for the field hockey club for all the sweat and blood and the long drivers before and (whew!) after the games: Pritam, Frank, Jonathan, Julia, Amber, Molly among others. Abhishek Kammula, Atul Maharshi and the rest of the Maple Hill soccer group – thank you for all the competitive soccer in sveltering Ithaca summers.

Familial support has also been a big part of my sense of well being that has contributed in no small measure towards my perseverance in research. I would be remiss if I did not thank my semi-literate late house-keeper, Shibani Mallick (Mashi) for tirelessly waking up at 3 in the morning everyday to lovingly make coffee, to help me into my undergraduate alma mater. I salute my Didang, Dada, Maa, Baba, Debrupa and Rukmi. Kinku, Mama and Maima, Auta, Minju, Shejdada, Lilu, Chhotadada, Chhotdadubhai and family, Chhotomama and Mami, and all the Mashis, Meshos and cousins, Pishi, Pishai, Sanju, Laketown relatives and Buban for the unaccountable love and affection. Thanks are also due to my wife Aditi’s parents,
Amit, Anupama, Chachas, Chachis, Mamas, Mamis and all my new relatives for their warmth and good wishes. Of course I cannot forget a “thank you” to my schoolmates, Arun, PV, Anand, Ranbir, and Muku, for the brotherhood.

In the end, I would like to show off this dissertation to my wife, Aditi. I cannot believe my luck for having snared her into ensuring a lifelong’s worth of happiness.
# TABLE OF CONTENTS

1. INTRODUCTION ........................................................................................................... 1
2. THREE-DIMENSIONAL ANALYTICAL MODELING OF NANOCRYSTAL MEMORY ELECTROSTATICS ................................................................. 13
3. MOLECULAR REDOX OF FULLERENES IN SOLID-STATE IN AN EEPROM-TYPE DEVICE ...................................................................................... 32
4. CARBON NANOTUBE BASED NON-VOLATILE MEMORY WITH CHARGE STORAGE IN METAL NANOCRYSTALS ........................................... 42
5. HIGH SENSITIVITY AND NON-LINEARITY OF CARBON-NANOTUBE-BASED CHARGE SENSORS ........................................................................ 55
6. ASYMMETRIC ELECTRIC FIELD ENHANCEMENT IN NANOCRYSTAL BASED NANOTUBE/NANOWIRE MEMORIES ........................................... 71
7. PERFORMANCE EVALUATION OF NANOCRYSTAL BASED NANOTUBE MEMORY BY TUNNELING DIELECTRIC ENGINEERING ....................... 93
8. CONCLUSIONS AND FUTURE WORK ........................................................................ 113
LIST OF FIGURES

Figure 2.1 (a) Cross-sectional view of gate stack for nanocrystal memory analytical modeling with design parameters-- \( a \): nanocrystal radius; \( d \): control oxide thickness; \( t \): tunneling oxide thickness and \( pitch \): lattice spacing of square lattice; (b) Scanning electron micrograph of Au nanocrystal distribution on tunneling oxide (2-3 nm thick SiO\(_2\)) showing fluctuations in position and size for self–assembled nanocrystals. Distribution data as a function of assembly conditions has been published in Ref. 4.

Figure 2.2 Dielectric sphere of radius \( a \) centered at origin with dielectric constant \( \varepsilon_{NC} \) is shown. The electric fields from (a) a point charge \( q \) at \( z=d \) and (b) a spherical shell of radius \( a \) of charge density \( \sigma \) and total charge \( q \) with dielectric constant \( \varepsilon_{NC} \) at \( z=d \) polarize the dielectric sphere at the origin. Region \( D \) is the matrix with dielectric constant \( \varepsilon_D \) in which dielectric spheres and charge are embedded, region \( NC \) is inside the dielectric sphere on the left and region \( NC' \) is inside the dielectric sphere with the charge shell on the right.

Figure 2.3 Superposition of the sources of potential: \( E_0 \): the electric field in gate stack due to parallel plate capacitor without nanocrystals, \( P \): the polarization of the nanocrystal and \( \sigma \): uniform surface charge of nanocrystal. Both the uniform surface charge and polarization form images on the conducting planes.

Figure 2.4 Comparison of analytically calculated potential (dash-dot line) with 3D FEM Poisson solver (solid line) of a 3x3 uniform metal nanocrystal array and 1D solution (dashed line) for (a) write condition (b) erase condition (c) retention condition (d) write saturation condition. The solid and dash-dot line are almost overlaid.

Figure 2.5 Comparison of analytically calculated potential (dash-dot line) with 3D FEM Poisson solver (solid line) of a 3x3 uniform semiconductor nanocrystal array and 1D solution (dashed line) for (a) write condition (b) erase condition (c) retention
condition (d) write saturation condition. The solid and dash-dot line are almost overlaid.

Figure 3.1 (a) CV measurements for electron injection into fullerenes based MOS capacitors at T=10K showing non-uniform CV curve shift i) 4Å C₆₀ ii) 6Å C₆₀ iii) C₇₀ (b) Flat band voltage vs. charging voltage reflecting non-uniform charging i) 4Å C₆₀ ii) 6Å C₆₀ iii) C₇₀

Figure 3.2 (a) CV measurements for hole injection into fullerenes based MOS capacitors at T=10K showing non-uniform CV curve shift i) 4Å C₆₀ ii) 6Å C₆₀ iii) C₇₀ (b) Flat band voltage vs. charging voltage reflecting non-uniform charging i) 4Å C₆₀ ii) 6Å C₆₀ iii) C₇₀

Figure 3.3 Charge on fullerene vs. potential on fullerene with respect to CB edge of Si at Si/SiO₂ a) electron injection for samples i) 4Å C₆₀ ii) 6Å C₆₀ iii) C₇₀ b) hole injection for samples i) 4Å C₆₀ ii) 6Å C₆₀ iii) C₇₀

Figure 4.1 (a) A cross section illustration of SWNT-based memory device; (b) The AFM image of the boxed area in inset shows SWNT between Au source/drain leads; inset shows the layout micrograph; (c) AFM shows self-aligned nanocrystals formed over a pair of crossed SWNT and another on the left.  

Figure 4.2 Hysteresis measurements for the single sweep direction from $V_{CH}$ to 2 V. Control sample measurements are shown in (a) at T=300K with $\Delta V_T/\Delta V_{CH} = 1.2 \text{ V} / 3 \text{ V}=0.4$, and (c) at T=10K with $\Delta V_T/\Delta V_{CH} = 0.2 \text{ V} / 2.5 \text{ V}=0.08$. Nanocrystal CNTFET measurements are shown in (b) at T=300K with $\Delta V_T/\Delta V_{CH} = 2 \text{ V} / 3 \text{ V}=0.67$ and (d) at T=10K with $\Delta V_T/\Delta V_{CH} = 1.4 \text{ V} / 3.4 \text{ V} = 0.41$.  

Figure 4.3 After the metal nanocrystal CNTFET device was charged at $V_{CH} = -5.5\text{V}$ for 10s, current relaxation in time was measured with $V_G = -1\text{V}$ and $V_{SD} = 300\text{mV}$ at 10K and 300K: (a) log ($I_D$) vs. time; and (b) log ($I_D$) vs. log (time) for the same measurement. Steps in $I_D$ at T=10K indicated single hole discharge events from
charged nanocrystals. ..................................................48

Figure 4.4 (a) Advancing $V_{CH}$ in 50 mV steps shows aggregation of $I_D-V_G$ curves; (b) Extraction of $V_G$ at constant $I_D=0.95$ nA resulted in steps in $V_G$ due to the combined effect of Coulomb blockade in nanocrystals and single charge sensitivity of SWNT conductance. .................................................................49

Figure 5.1 A CNT charge sensor with a single charge dot at the middle of the channel. (a) A schematic plot of the cross section of the modeled device. (b) The simulated channel conductance vs. the gate voltage on log (left) and linear (right) scales for different amounts of charge on the dot, $Q_d=0$ to $3e$, at $e/step$. (c) The first conduction and valence bands versus the position along the channel for $Q_d=0$ to $3e$ at $V_G=0.2V$. The source and drain Fermi level is set at $E_F=0$. The diameter of the dot is 8nm.....60

Figure 5.2 (a) The simulated $V_T$ shift due to single electron addition to the charge dot versus the bottom oxide thickness when the charge dot is modeled as a metallic ball (the solid line) and a uniformly charged ball (the dashed line). (b) Room temperature measurement of the CNT memory device right after charging (the curve on the left) and fully discharged (the curve on the right). The sub-threshold slope (dashed line) should be the same in both curves. The $I_D-V_G$ curve for the charged device, however, shows a series of threshold voltage shift with magnitude of ~100mV-200mV whose direction suggests that these shifts are due to discharge events during measurements. .................................................................61

Figure 5.3 Position-dependent response of the sensor. (a) The channel conductance vs. the position of the charge dot when $Q_D=2e$ for the device as shown in Figure 1a. The dashed line shows the channel conductance with $Q_d=0$ for reference. The first subband profile (the solid line) and the channel transmission versus energy (the dashed line) for (b) no charge dot, (c) a charge dot at $x_{dot}=L_{ch}/10$, and (d) a charge dot at $x_{dot}=L_{ch}/2$. The gate voltage $V_G=0.2V$ and the drain voltage $V_D=0$. .................................................65
Figure 5.4 A CNT charge sensor with two charge dots. (a) The simulated device structure. (b) The channel conductance versus the position of dot 2, \( x_2 \). Dot 1 is fixed at \( x_1 = L_{ch}/3 \). The solid line is for \( Q_1 = Q_2 = 2e \), and the dashed line is for \( Q_1 = 2e \) and \( Q_2 = 0 \). (c) The first subband profile (the solid line) and the channel transmission (the dashed line) for \( x_1 = L_{ch}/3 \), \( Q_1 = 2e \), and \( Q_2 = 0 \). (d) The first subband profile and transmission for \( x_1 = L_{ch}/3 \), \( x_2 = 2/3 L_{ch} \), and \( Q_1 = Q_2 = 2e \). The gate voltage \( V_G = 0.2V \) and the drain voltage \( V_D = 0 \). The charge dots are modeled as metallic balls.

Figure 6.1 Basic CNT-NC memory schematic showing (a) bottom-gate and (b) top-gate structures.

Figure 6.2 a) 3D structure for CNT-NC memory with the back-gated CNTFET with 3 NCs aligned to the CNT. b) The NC EEPROM structure in 3D showing a 3×3 array of NCs. The cross sections of the different memory structures c) a bottom-gate CNT-NC memory d) a top-gated CNT-NC memory e) a NC EEPROM are shown. The typical structural parameters used in the electrostatics simulations are as follows: NCs diameter: 6 nm; the CNT diameter 2 nm; pitch: 12 nm; tunneling dielectric thickness \( d_t \): 3 nm (Si EEPROM) and 4 nm (NC/CNT memory); distance between the center of the NC to the gate, \( d_c \): 30 nm.

Figure 6.3 Potential profile for capacitive coupling calculation (solid line), self capacitance calculation (dashed line) and capacitive coupling calculation without NC (dash-dotted line) for a CNT NC memory with various parameters \( d_c : 30 \) nm; \( d_t : 5 \) nm; NC diameter: 6 nm; CNT diameter: 2 nm for (a) top-gate and (b) bottom gate structures.

Figure 6.4 Retention condition potential profile along the extended path of least action. The solid (single NC case) and dashed (3 aligned NC case) lines show potential profiles when considering the CNT to be a dielectric. The dotted (single NC case) and dash-dotted (3 aligned NC case) lines are for assuming that the CNT as a conductor
which is similar to the self capacitance calculations……………………………….78

Figure 7.1 (a) AFM image of 4 nm of ALD Al\textsubscript{2}O\textsubscript{3} deposited on a nanotube. (b) Scan height along the CNT shows surface roughness approximately equal to film thickness due to inhomogeneous nucleation and growth of ALD Al\textsubscript{2}O\textsubscript{3} on the nanotube surface. ………………………………………………………………………………………….97

Figure 7.2 (a) Electrical characterization of a CNTFET covered with ALD Al\textsubscript{2}O\textsubscript{3}, before (blue curve) and after (red curve) the deposition of 1.2 nm Au for NC-formation, shows that the nanotube cannot be turned off after Au deposition. The on-current level remains roughly same. (b) Electrical characterization of a CNTFET with a bare CNT without ALD Al\textsubscript{2}O\textsubscript{3} cover, before (blue curve) and after (red curve) the deposition of 1.2 nm Au for NC-formation, shows that the nanotube cannot be turned off after Au deposition just as when it was covered with ALD Al\textsubscript{2}O\textsubscript{3}. (c) SEM of the CNT after 1.2 nm Au deposition for NC-formation. The presence of the nanotube is indicated by an approximate linear arrangement of the NCs in the random NC formation between by the arrows. Inset shows close-up of NCs arranged on either side of the NC avoiding the NC itself………………………………………………….98

Figure 7.3 Program rate measurements of NC sample and control sample with evaporated SiO\textsubscript{2} tunneling dielectric at 10K. There are two slopes evident. A strongly pulse voltage dependent slope for fast program times is separated from a pulse voltage independent slope for slow program times by a knee at approximately ~ms pulse widths. The NC sample produces larger threshold voltage shift for the same bias. For example, for $V_T$ shift of 2V, the NC sample requires pulse voltage of 11 V (indicated by blue arrow) compared to the 15 V (indicated by red arrow) for the control sample. The pulse independent slope is about 2× for the NC sample compared to the control sample. At pulse voltage $\leq$ 5 V, there is no $V_T$ shift for the control sample compared to a 1V $V_T$ shift for the NC-sample. ……………………………………………….103
Figure 7.4 Program threshold voltage shift measurement of control samples in (a) and (c) and NC sample in (b) and (d) with evaporated SiO\textsubscript{2} tunneling dielectric. (a) and (b) show measurements at room temperature and (c) and (d) show measurements at 10K for two different pulse widths: 100 µs (blue-line) and 1 s (red-line) as a function for pulse voltage. Both samples at room temperature show two regimes denoted by red roman numerals demarcated by solid red vertical lines for the pulse width of 1s: (i) FN tunneling or Poole-Frenkel tunneling at lower biases and (ii) saturation at high biases. However at pulse widths of 100 µs, two regimes denoted by blue roman numerals demarcated by solid blue vertical lines, can be seen: (i) No charge injection at low bias and (ii) FN tunneling or Poole-Frenkel tunneling at a higher bias. For the low temperature curve, Poole-Frenkel conduction is eliminated and Fowler Norheim tunneling dominates. Both samples at 10 K show three regimes for the pulse width of 1 s demarcated by solid red vertical lines denoted by red roman numerals: (iii) Direct tunneling at lower biases, (iv) FN tunneling at medium biases, and (v) saturation at high biases. At pulse widths of 100 µs, the curves are similar to that at room temperature for the same timescale. They show two regimes indicated by blue roman numerals demarcated by solid blue vertical lines: (i) No charge injection at low bias and (ii) FN tunneling at a higher bias. ..............................................................104

Figure 7.5 Program rate measurement of NC sample and control sample with composite tunneling dielectric of 4 nm ALD Al\textsubscript{2}O\textsubscript{3} and 3 nm evaporated SiO\textsubscript{2} tunneling dielectric at 10K. The effect of NC is not evident clearly as trap based charging dominates. ..............................................................106

Figure 7.6 Retention measurements for NC sample with composite tunneling dielectric of 4 nm ALD Al\textsubscript{2}O\textsubscript{3} and 3 nm evaporated SiO\textsubscript{2} tunneling dielectric at room temperature and at 85°C. (a) Room temperature I\textsubscript{D}V\textsubscript{G} measurements before (blue), after charging (green) and after 9000 s after charging (black) and at some intermediate
times. (b) The threshold voltage shift with time is plotted and extrapolated to find total retention time. The same process is repeated for figure (c) and (d) at 85°C. The retention values are approximately $\geq 10^5$ s (3 days) at room temperature while at 85°C retention is $\geq 6 \times 10^3$ s.
LIST OF TABLES

Table 6.1: Capacitive coupling and self capacitance calculations for various CNT-NC memory and NC planar Si memory structure for the program condition. .............82

Table 6.2: Potential drop in the tunneling dielectric in the path of least action for $d_c = 30$ nm for CNT-NC memory and Si planar memory for retention condition with 5 electrons in the NCs.................................................................83
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>1 dimensional</td>
</tr>
<tr>
<td>2D</td>
<td>2 dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>3 dimensional</td>
</tr>
<tr>
<td>AFM</td>
<td>Atomic force microscopy</td>
</tr>
<tr>
<td>ALD</td>
<td>Atomic layer deposition</td>
</tr>
<tr>
<td>BG</td>
<td>Bottom-gate</td>
</tr>
<tr>
<td>CB</td>
<td>Conduction band</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complimentary metal oxide semiconductor technology</td>
</tr>
<tr>
<td>CNT</td>
<td>Carbon nanotubes</td>
</tr>
<tr>
<td>CNTFET</td>
<td>Carbon nanotube field-effect-transistor</td>
</tr>
<tr>
<td>CVD</td>
<td>Chemical vapor deposition</td>
</tr>
<tr>
<td>CvMOS</td>
<td>Chemical neuron MOS</td>
</tr>
<tr>
<td>EEPROM</td>
<td>Electrically erasable programmable read-only memory</td>
</tr>
<tr>
<td>EFM</td>
<td>Electrostatic force microscopy</td>
</tr>
<tr>
<td>FEM</td>
<td>Finite element method</td>
</tr>
<tr>
<td>FET</td>
<td>Field effect transistors</td>
</tr>
<tr>
<td>FinFET</td>
<td>FET with a fin-like channel</td>
</tr>
<tr>
<td>GIDL</td>
<td>Gate induced drain leakage</td>
</tr>
<tr>
<td>HFCV</td>
<td>High frequency capacitance voltage</td>
</tr>
<tr>
<td>LDOS</td>
<td>Local density of states</td>
</tr>
<tr>
<td>LOCOS</td>
<td>Local Oxidation of Silicon</td>
</tr>
<tr>
<td>MFP</td>
<td>mean free path</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MOS</td>
<td>Metal oxide semiconductor technology</td>
</tr>
<tr>
<td>NC</td>
<td>Nanocrystal</td>
</tr>
<tr>
<td>NCM</td>
<td>Nanocrystal memories</td>
</tr>
<tr>
<td>NEGF</td>
<td>Non-equilibrium Green’s function</td>
</tr>
<tr>
<td>ONO</td>
<td>Oxide-nitride-oxide</td>
</tr>
<tr>
<td>P/E</td>
<td>Program/erase</td>
</tr>
<tr>
<td>QCM</td>
<td>Quartz crystal microbalance</td>
</tr>
<tr>
<td>RIE</td>
<td>Reactive ion etching</td>
</tr>
<tr>
<td>SEM</td>
<td>Scanning electron microscopy</td>
</tr>
<tr>
<td>SET</td>
<td>Single electron transistors</td>
</tr>
<tr>
<td>SOI</td>
<td>Silicon-on-insulator</td>
</tr>
<tr>
<td>SWNT</td>
<td>Single-wall carbon nanotubes</td>
</tr>
<tr>
<td>TG</td>
<td>Top-gate</td>
</tr>
<tr>
<td>VLSI</td>
<td>Very-large-scale-integration</td>
</tr>
<tr>
<td>WKB</td>
<td>Wentzel Kramers Brillouin method</td>
</tr>
</tbody>
</table>
## LIST OF SYMBOL

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0$</td>
<td>Fermi distribution function,</td>
</tr>
<tr>
<td>$V(\vec{r})$</td>
<td>the potential,</td>
</tr>
<tr>
<td>$\Sigma_{S,D}$</td>
<td>source/drain self-energy matrix for the metal-CNT contacts</td>
</tr>
<tr>
<td>$\Phi_{Bn}$</td>
<td>Schottky barrier height for electrons at CNT metal contact</td>
</tr>
<tr>
<td>$(r, \theta, \phi)$</td>
<td>coordinates</td>
</tr>
<tr>
<td>$a$</td>
<td>nanocrystal diameter</td>
</tr>
<tr>
<td>$C_{CH}$</td>
<td>capacitance of the capacitor during injection</td>
</tr>
<tr>
<td>$C_{OX}$</td>
<td>total oxide capacitance</td>
</tr>
<tr>
<td>$C_{tunnel-ox}$</td>
<td>capacitance due to a parallel-plate capacitor with the tunneling oxide</td>
</tr>
<tr>
<td>$D$</td>
<td>electric displacement vector</td>
</tr>
<tr>
<td>$d_c$</td>
<td>the control dielectric thickness</td>
</tr>
<tr>
<td>$d_{CNT}$</td>
<td>diameter of CNT</td>
</tr>
<tr>
<td>$d_t$</td>
<td>tunneling dielectric thickness</td>
</tr>
<tr>
<td>$E$</td>
<td>electric field vector</td>
</tr>
<tr>
<td>$E_0$</td>
<td>uniform field,</td>
</tr>
<tr>
<td>$E_{FS}, E_{FD}$</td>
<td>the source and drain Fermi levels</td>
</tr>
<tr>
<td>$E_g$</td>
<td>band gap</td>
</tr>
<tr>
<td>$F$</td>
<td>Frequency</td>
</tr>
<tr>
<td>$G_D$</td>
<td>channel conductance</td>
</tr>
<tr>
<td>$H$</td>
<td>Hamiltonian matrix</td>
</tr>
<tr>
<td>$I_D$</td>
<td>Drain Current</td>
</tr>
</tbody>
</table>
$L_{ch}$ channel length

$L_D$ Debye length

$p$ pitch

$q$ charge

$Q_F$ charge injection into fullerene ($)$

$Q_i$ charge in the $i^{th}$ quantum dot

$r$ position vector

$t$ tunneling oxide thickness

$T(E)$ the channel transmission

$t_{bot}$ bottom gate control oxide

$U$ Self-consist potential

$U_S, U_D, U_G$ electrostatic potential on the source, drain and bottom gate

$V_{CH}$ charging voltage ($)$

$V_F$ voltage on the fullerene ($)$

$V_{FB}$ shifted flat band voltage

$V_{FB0}$ ideal flat band voltage

$V_G$ Gate voltage

$V_T$ threshold voltage

$\Delta V_T/\Delta V_{CH}$ relative $V_T$ shift efficiency

$\Phi(r)$ potential

$\delta(r-r')$ delta function

$\delta_{dipole}$ the dipole as a pair of opposite point charges separated by a distance

$\varepsilon_D$ dielectric constant of the dielectric
CHAPTER 1

INTRODUCTION

1.1 Literature Review

Nanotechnology has emerged as a realm of promise and understanding for modern science and technology. Even though the classical description of the world is generally pertinent, the quantum mechanical nature of the world is becoming increasingly tangible as scientists routinely probe materials, devices and phenomena in the nanoscale. The flash memory [1], a commercial technology [2, 3], takes advantage of this quantum mechanical nature to inject/tunnel and store charge quanta in the potential well surrounded by a finite dielectric barrier. The presence/absence of the stored charge can be read out to distinguish between memory states. A literature review is generally in order to put the dissertation work in perspective of contemporary science and technology. In our case, a concise state-of-the-art has been described in the introduction of individual chapters. Additionally the doctoral dissertation of Chungho Lee [4] provides a reasonably current survey of the literature for nanocrystal memories. With the immediate surroundings established, this thesis attempts to implement further nanoscale engineering to integrate lower dimensional systems like nanoparticles and nanowires/nanotubes. While inclusion of nanocrystals as charge storage nodes strongly enhances the tunneling dielectric scaling prospects for low voltage operations due to purely physical reasons [4], the concomitant coulomb blockade and other quantum effects of nano-engineered systems can be harnessed to explore further performance enhancement. This thesis demonstrates the ability to control and detect the precise number of excess charge confined in the nanoparticle using the small length scale effects enables prospects of functional
densification where rather than coding information as merely the presence or absence of charge, the information stored as the number of charge leads to multi-level memory. In this regard, a brief overview of the state-of-the-art for nanoelectronics and related topics such fabrication/assembly technology has been provided to put this research into a larger perspective.

Single electron transistors (SET) [5] have been demonstrated where the resonant tunneling through a nanoparticle is modulated by a gate electrode and depends on the charge occupancy in the nanoparticle. Some progress has been made in terms of room temperature devices [6-8]. Some clever circuit implementations have also been proposed taking advantage of the discreteness of the charge storage in the nanoparticles [9, 10]. However, the low current readout of the devices provides challenge in terms of large scale integration, clock-speed, and integration. In contrast, the nanocrystal memory can be used to detect single charge and yet provide large output current drive for large scale integration in two possible configurations. The first one is to make the nanoparticles uniform in position and size. Simple uniform self assembled array of Au nanoparticles [11] as well as more complex 2 dimensional (2D) arrays using bimodal ensembles [12] has been demonstrated. Phase separated block-co-polymer based templates have also been used to produce such long range ordering [13]. Another important connection i.e. the registration to lithographic structures for the self-assembled structures has been demonstrated using block-co-polymer chemistry [14, 15]. An alternative idea to the uniform metal nanocrystal array is their replacement with molecules like fullerenes. Fullerenes chemistry and physics has been studied extensively [16]. Fullerenes have been studied as the conducting quantum dot in SET devices [17, 18]. Thin-films of C$_{60}$ have also been used as the channel for a field effect transistor (FET) [19]. We have attempted to integrate C$_{60}$ into the MOS gate stack. The second method to detect individual charge efficiently is to be able to
implement a 1-dimensional (1D) semiconducting channel [20, 21] which can be either nanowire or nanotubes. The challenge of integrating 1D element into devices has been addressed considerably. Nanowire fabrication has been demonstrated with both axial and radial modulation and control of material [22, 23]. Carbon nanotube growth and properties have been extensively research and compiled [24]. Uniformity of properties in as produced or filtered material is currently being explored. Conformal dielectric deposition on nanotube has been recently demonstrated using atomic layer deposition [25] which is a remarkable achievement since the nanotube provides a very inert surface for chemical deposit nucleation. To enable large scale integration, alignment of the 1D structures have been attempted both during growth using electric field, template-assistance and glass flow [26-28] and ex situ assembly using electric field, Langmuir-Blodgett technique, solvent flow and block co-polymer based phase separation [29-31]. Single vertical surround gate ZnO nanowire field effect transistors (FET) with prospects for high density integration has been demonstrated [32]. From a systems perspective, while the uniformity of self-assembled structures are promising for memory type uniform structures [33], interconnection for logic and therefore coding information on demand is still an open issue. Interestingly biological agents have been suggested for such directed assembly [34]. However there is scope for extended exploration before realization of self assembled logic-type devices for very large scale integration (VLSI).

1.2 Organization

The temporal progression of the work would provide a different perspective of this dissertation. Carbon molecules like nanotubes (CNT) and fullerenes (C_{60}, C_{70} etc) are a class of molecules with a wide range of electronic properties. These are self-assembled nanostructures that complement the CMOS top-down fabrication scheme.
Integration of such molecules on the CMOS platform provides an interface or a communication link between the nanoscale systems ($C_{60}$, CNT) and top-down electronics to exploit a new basket of properties. Exploration of this idea in the realm of non-volatile memory devices and charge sensors has been central to the development of this dissertation. There are two distinct sub-areas: a) nano-floating gate engineering and b) channel engineering.

The second chapter, however, describes the ponderings on the device physics of nanocrystal memories (NCM). In literature, the NCM has been modeled in 1D both in electrostatics (Poisson’s Equation) and the quantum transport (Schrödinger’s Equation) [35]. At most quasi-1D electrostatics has been considered by adding the 3D coulomb blockade term [36]. The quantum transport has been treated mostly 1D [36] or quasi-1D where the separation in Cartesian coordinates has been applied for analytical simplification [37]. However, the nanocrystal (NC) in a MOS gate stack introduces a large 3-dimensionality (3D) in the potential profile [38]. The enhanced electric field asymmetry due to the field focusing effect of the high dielectric constant (metal) nanocrystal explains the low-voltage operation compared the monolithic floating gate memory. The second chapter is devoted to the analytical electrostatic model developed that takes into account a finite number of interacting nanocrystals [39]. The solution of the time-dependent Schrödinger equation for the exact potential is the other part of the transport calculation. A numerical implementation of Bardeen’s Transfer Hamiltonian formalism for a transport calculation for exact 3D potential profile, which is a possible option for self-consistent calculations, is currently ongoing [40]. The basic reason to attempt a multi-dimensional approach is that it can easily accommodate NC-CNT quantum transport in the NC-CNT memory device discussed below which is severely 3D is structure.
The third chapter presents the first experimental investigations in the integration of carbon molecules in the CMOS gate stack [41, 42]. Nanocrystal memory (NCM) provides aggressive tunneling oxide scaling and high reliability over monolithic floating gates of flash memory. Nanocrystals (NCs) have a finite shape and size variation that leads to non-uniform coulomb blockade. Mono-disperse charge storage nodes can provide uniform charging energy and multi-level memory. Consequently, inherently mono-disperse C_{60} and C_{70} molecules have been proposed to replace the nanocrystals. The process to integrate the fullerenes in a CMOS gate stack have paved the way for the first demonstration of the control of molecular redox in the solid state which can be used for both fundamental and technological explorations.

The next four chapters (43-45) develop the different perspectives on the NC-CNT memory device. Narrow channel based memories have been studied as functional limits of the flash memory technology which enhances density, memory window, retention and reliability at the cost of considerable process complexity. In chapter 4, a nanocrystal based memory where the nanotube replaces the narrow silicon channel is introduced [43]. The fabrication and measurement of the device is described. The chapter 5 employs the non-equilibrium Green’s function (NEGF) based transport calculations in the nanotube perturbed by the potential of the charged nano-floating gates done in collaboration with Prof. Jing Guo, to explain the experimental results [44]. On the other hand, the NC-CNT transport has been explored from the electrostatics perspective in chapter 6 [45]. Chapter 7 is a survey of the room temperature implementation of the NC-CNT memory currently being pursued. Atomic layer deposition (ALD) of high-k dielectrics and chemical vapor deposition (CVD) of SiO_{2} are being explored for ultra-thin (3-7 nm) tunneling dielectric with acceptable defect density is required for room temperature operation.
The chapter 8 concludes the dissertation with a spate of future research topics which include scientific investigations of charge transport between lower dimensional systems e.g. nanocrystals, nanotubes as well as sensing, memory applications.
REFERENCES


2.1 Abstract

The integration of a 2-dimensional (2D) distribution of discrete nanoscale floating gates in the nonvolatile memory gate stack produces significant 3-dimensional (3D) electrostatic effects in contrast to the conventional flash memory modeling where a 1-dimensional (1D) treatment is often sufficient. We have developed an analytical model for 3D electrostatics, which can not only enhance design intuition for device optimization, but also provide convenient integration with a Schrödinger solver for self-consistent transport calculations since it is independent of discretization requirements. The model is validated by comparing with a finite-element Maxwell equation solver. The 3D analytical model has a much lower root-mean-square error than the 1D formulation for electrostatic potentials and fields in the tunneling path.

2.2 Introduction

Nanoscale discrete charge storage for nonvolatile memory devices has been explored from both experimental and modeling perspectives in the last decade. From the materials point of view, charge storage has been explored in traps of oxide-nitride-oxide (ONO) gate stack [1] and in self-assembled nanocrystals of semiconductors (e.g., Si, Ge, etc) [2, 3] and metals (e.g., Au, Ag, Pt, Ni, etc) [4-6]. Various dielectrics have also been evaluated such as SiO$_2$ and high-$\kappa$ dielectrics [7, 8]. The electrostatic models currently used for tunneling calculations are essentially 1-dimensional (1D), considering only the least-action tunneling path directly under the nanocrystal [9, 10].
Figure 2.1 (a) Cross-sectional view of gate stack for nanocrystal memory analytical modeling with design parameters-- $a$: nanocrystal radius; $d$: control oxide thickness; $t$: tunneling oxide thickness and $pitch$: lattice spacing of square lattice; (b) Scanning electron micrograph of Au nanocrystal distribution on tunneling oxide (2-3 nm thick SiO$_2$) showing fluctuations in position and size for self–assembled nanocrystals. Distribution data as a function of assembly conditions has been published in Ref. 4.
However, experimental observation of low program/erase (P/E) voltages and large memory window of nanocrystal memories [11] in comparison with monolithic floating gate and trap-based ONO gate stacks [12] indicates differences that can only be explained by considering the discreteness and 3-dimensional (3D) nature of the array of nanocrystals. The electrostatic correction due to the 3D potential perturbation from the nanocrystals is actually significant in comparison with the 1D model [13]. In fact, neither the electric field nor the charging capacitance is accurately represented in the 1D model, both key to accurate formulations for direct [14] and Fowler-Nordheim tunneling [15]. Although 3D numerical electrostatic solutions can be employed, the discretization requirements for convergence and accuracy can be computationally expensive and often need special symmetry assumptions for the reduction of computational costs that are unable to accommodate a finite system of nanocrystals with non-uniform size and positional distributions. Additionally, design intuition can not be easily obtained from such opaque numerical calculations. In the present work, a simple but accurate analytical 3D electrostatics model for the nanocrystal memory gate stack has been developed, where the electrostatic contributions to the potential due to different sources of potential, e.g. charge, potential on electrodes, polarization, etc, are separated and evaluated analytically, and finally integrated using superposition.

2.3 Formulation

The nanocrystal memory gate stack is represented by a distribution of spherical nanocrystals formed by self-assembly on top of a tunneling dielectric grown on a silicon substrate as shown in Fig. 2.1(a). The nanocrystals have a positional and size distribution [4] evident from the scanning electron micrograph (SEM) in Fig. 1(b). The stack is further covered with control dielectric with a metal gate on top. In
general, the potential relaxation or band-bending in the silicon substrate needs to be solved self-consistently. We approximate the substrate as a ground plane. This is quite accurate in the inversion and accumulation regimes (i.e. during P/E operations). When the substrate is in the depletion regime, the ground-plane approximation offers useful upper-bounds for electric field profiles without having to implement self-consistent solutions for the potential. The 3D electrostatics of a single metal/dielectric sphere between conducting plates have accurate analytical formulation by the solution of an infinite set of equations [16, 17]. However electrostatic modeling of a nanocrystal memory device requires the consideration of not a single nanocrystal but a distribution of electrostatically interacting nanocrystals on a tunneling oxide as shown in Fig. 1(b). While a finite element Poisson solver can easily solve for a regular 2-dimensional (2D) array of nanocrystals using symmetry and periodicity considerations, an irregular array would require the treatment of the entire space. In such a case of a finite distribution of nanocrystals especially for devices beyond the 65 nm node or to study the effects of nanocrystals at the boundaries, an analytical solution can be a useful guide. Our formulation for a finite nanocrystal array involves several electrostatic solution techniques (subsections A-G) discussed below.

A. Method of images: This method is used to solve for electrostatic potential for a conducting plane at V=0 V at z=0 with a charge \( q \) at distance \( z=d \). The conducting surface can be replaced by an image charge \(-q\), placed at \( z=-d \). By anti-symmetry the potential at the surface is zero\(^{18}\). Similarly, a charge distribution \( \rho(r) \) on a grounded conducting plane can be similarly treated by using an image charge distribution. The potential due to a charge between two such conducting planes can be equivalently treated with an infinite series of image charges. A polarized dielectric between two conducting planes can be regarded as a bound surface and volume charge [18] which can also be accommodated in the method of images as a charge distribution.
Figure 2.2 Dielectric sphere of radius $a$ centered at origin with dielectric constant $\varepsilon_{NC}$ is shown. The electric fields from (a) a point charge $q$ at $z=d$ and (b) a spherical shell of radius $a$ of charge density $\sigma$ and total charge $q$ with dielectric constant $\varepsilon_{NC}$ at $z=d$ polarize the dielectric sphere at the origin. Region $D$ is the matrix with dielectric constant $\varepsilon_D$ in which dielectric spheres and charge are embedded, region $NC$ is inside the dielectric sphere on the left and region $NC'$ is inside the dielectric sphere with the charge shell on the right.
Figure 2.3 Superposition of the sources of potential: $E_0$: the electric field in gate stack due to parallel plate capacitor without nanocrystals, $P$: the polarization of the nanocrystal and $\sigma$: uniform surface charge of nanocrystal. Both the uniform surface charge and polarization form images on the conducting planes.
B. Modeling for metallic and semiconducting nanocrystals in a dielectric of permittivity $\varepsilon_D$: A dielectric sphere (nanocrystal) embedded in another dielectric with dielectric constant $\varepsilon_D$ can manifest its metallic or semiconducting nature of the nanocrystal by its dielectric constant $\varepsilon_{NC}$. In case of a metal, we can set $\varepsilon_{NC} \rightarrow \infty$. For semiconducting nanocrystal, if its extrinsic Debye length $L_D >> 2a$, where $a$ is the radius of the nanocrystal, then the dielectric constant of the semiconductor nanocrystal, $\varepsilon_{NC}$, contains all information necessary for polarization calculations.

C. Polarization of a dielectric sphere due to the charge shell: In spherical polar coordinates $(r, \theta, \phi)$ with a dielectric sphere centered at the origin as shown in Fig. 2.(a), the potential $\Phi(r)$ produced by the polarization of a dielectric sphere due to point charge at $r'$ can be solved from Poisson’s equation. We need to solve two equations with appropriate boundary conditions as shown in Fig. 2.2(a). In region $D$ inside the dielectric, we need to solve the Poisson’s Eq. (1), where the $\delta(r-r')$ is a delta function showing that the charge $q$ is present only at $r=r'$ and in region $NC$ inside the nanocrystal we need to solve the Laplace Eq. (2)

$$\nabla^2 \Phi_D = \left( \frac{q}{\varepsilon} \right) \delta(r-r') \quad (2.1)$$

$$\nabla^2 \Phi_{NC} = 0 \quad (2.2)$$

In cylindrical polar coordinates $(\rho, \phi, z)$ with $q$ is on the z-axis, i.e. $r'=d\hat{z}$ as shown in Fig. 2.2(a), the analytical solution in spherical polar coordinates is of the form

$$\Phi_D = \frac{q}{4\pi\varepsilon_D|r-r'|} + \sum_{l=0}^{\infty} B_l r^{-(l+1)} P_l(\cos \theta) \quad (2.3)$$

$$\Phi_{NC} = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) = \frac{q}{4\pi\varepsilon_D|r-r'|} + \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) \quad (2.4)$$
where $P_l$ are Legendre polynomials and $A_l$, $B_l$ and $A'_l$ are constants. The effect of polarization in Eq. (3) and (4) can be separated from the potential due to the point charge $q$, as the summation terms containing $A_l'$ and $B_l$. The boundary conditions at the nanocrystal/dielectric interface are based on the continuity of normal component of the electric displacement $D$ and of the tangential component of the electric field $E$ to the interface:

$$
-\varepsilon_D \left( \frac{\partial \Phi}{\partial r} \right)_{r=a^+} = -\varepsilon_{NC} \left( \frac{\partial \Phi_{NC}}{\partial r} \right)_{r=a^-} \tag{2.5}
$$

$$
-\frac{1}{a} \left( \frac{\partial \Phi}{\partial \theta} \right)_{r=a^+} = -\frac{1}{a} \left( \frac{\partial \Phi_{NC}}{\partial \theta} \right)_{r=a^-} \tag{2.6}
$$

Using these boundary conditions, the solution for $A_l$, $A_l'$ and $B_l$ is given by

$$
A_l = -\frac{q}{4\pi\varepsilon_D} \frac{1}{d^{l+1}} \frac{2l+1}{(\varepsilon_{NC} / \varepsilon_D + 1)^l + 1} \tag{2.7}
$$

$$
B_l = -\frac{q}{4\pi\varepsilon_D} \frac{a^{2l+1}}{d^{l+1}} \frac{(\varepsilon_{NC} / \varepsilon_D - 1)^l}{(\varepsilon_{NC} / \varepsilon_D + 1)^l + 1} \tag{2.8}
$$

$$
A'_l = -\frac{q}{4\pi\varepsilon_D} \frac{1}{d^{l+1}} \frac{(\varepsilon_{NC} / \varepsilon_D - 1)^l}{(\varepsilon_{NC} / \varepsilon_D + 1)^l + 1} \tag{2.9}
$$

The above calculation can be easily modified for a spherical shell at $r'$ with charge $q$ uniformly distributed as opposed to a point charge $q$ as shown in Fig. 2.2(b). The potential $\Phi_D$ for Eq. (1) contains an additional constant term as shown in Eq. (10) in square brackets.
\[ \Phi_D = \frac{q}{4\pi \varepsilon_D} \left| r - r' \right| - \frac{q}{4\pi a} \left( \frac{1}{\varepsilon_D} - \frac{1}{\varepsilon_{NC}} \right) + \sum_{l=0}^{\infty} B_l r^{-l+1} P_l(\cos \theta) \]  

(2.10)

Due to the constraint of the boundary conditions in Eq. (5) and (6), \( \Phi_{NC} \) also contains the same additional constant term.

\[ \Phi_{NC} = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) - \frac{q}{4\pi a} \left( \frac{1}{\varepsilon_D} - \frac{1}{\varepsilon_{NC}} \right) \]  

(2.11)

The values of \( A_l \) and \( B_l \) are unaffected by the constant because they are determined by taking derivatives.

**D. Approximating the polarization due to a point charge:** The potential outside the dielectric sphere \((r \geq a)\), due to the polarization due to the point charge \( q \), can be approximated by

\[ \Phi_D = \frac{q}{4\pi \varepsilon_1} \left( \frac{\varepsilon_{NC}}{\varepsilon_D} - 1 \right) \frac{a^3}{r^2} P_1(\cos \theta) + \frac{a^3}{d^2} \left( \frac{a^5}{d^3} \right) \]  

In the first-order approximation with \( d >> a \), the polarization is equivalently caused by a uniform field \( E_{eff} \), which is exactly the electric field at center of the dielectric sphere due to the charge \( q \) at \( r' = d\hat{z} \).

\[ E_{eff} = -\frac{q}{4\pi \varepsilon_D d^2} \hat{z} \]  

(2.13)

Hence, for \( d >> a \), the polarization of the dielectric sphere at the origin is independent of the detailed nature of the electric field and can be approximated by a uniform electric field. This argument can be extended for an arbitrary charge distribution instead of a point charge as long as the \( d >> a \) assumption is satisfied.
E. Polarization due to a uniform field: In this case the polarization by the uniform field treatment is exact [18].

F. Polarization of a dielectric sphere due to a uniformly polarized sphere: The polarization effect of a uniformly polarized sphere on a dielectric sphere has been treated using uniform field approximation in our implementation as described in subsection D. Alternatively, it is possible to treat the uniformly polarized sphere as a dipole, and then describe the dipole as a pair of opposite point charges separated by a distance \( \delta_{dipole} \ll a \). We can then use the polarization due to point charge formulation as shown in subsection C and superposition principle for the two point charges to get higher order corrections.

G. Superposition: The final results are obtained by superposition [18] of the effect of different sources of potential shown in Fig. 2.3.

   a) Charge stored in nanocrystals between conductive planes shown as \( \sigma \).
   b) Potential difference between the conductive parallel planes creating to the uniform field, \( E_0 \).
   c) Polarization of the nanocrystals between the conductive planes because of electric field due to
      i. charges in nanocrystals and the image charges from the parallel conductive planes
      ii. potential difference between the parallel conductive planes
      iii. polarization of other nanocrystals and images of the polarized nanocrystals from the parallel conducting planes

For polarization of a nanocrystal, the significance of the terms described in (c)(i-iii) above, is proportional to the magnitude of the electric fields. In this regard, the influence of (iii) is the smallest because electric field due to the polarized source decays much faster with distance than that for (i) and (ii). Hence the correction due
to (iii) can be treated to first order using uniform electric approximation as discussed previously. However it is necessary to treat the polarization due to (i) and (ii) exactly since their electric fields are significant.

2.4. Results, Validation and Discussion

To validate our model, we compare the analytical solution for a regular array of 3×3 nanocrystals, in a number of different memory operations for both metal and silicon nanocrystals with calculations from an FEM-based Poisson solver [19]. The typical parameters used for the model as indicated in Fig. 2.1(a) were nanocrystal radius $a$ of 3 nm, pitch $p$ of 12 nm, control oxide thickness $d_c$ of 33 nm, and tunneling oxide thickness $t$ of 3 nm. The memory operation conditions are:

1. Start of Programming: $q=0e$ and $V_G=5V$
2. Start of Erasing: $q=5e$ and $V_G=-5V$
3. Retention with a built-in voltage of 1V: $q=5e$ and $V_G=1V$
4. Saturation of programming: $q=5e$ and $V_G=5V$

We have assumed that the silicon substrate as a ground plane because during inversion and accumulation it has a high carrier density. The presence of nanocrystals can cause small potential variations in Si/SiO$_2$ interface which creates a percolation [20] based current flow. However, the intricate pathway of least resistive path is created by the exponential dependence of carrier concentration on the potential. Hence, while the carrier concentration can vary significantly the potential fluctuations are quite small at the Si/SiO$_2$ interface. Hence while in memory operation conditions 1 and 2, the assumption of the silicon substrate as ground plane is mostly valid, the conditions 3 and 4, where inversion/accumulation cannot be assumed, the model only serves as an upper-bound estimation of the electric field. This is because in such cases the silicon substrate is under depletion, the electric field can penetrate into the substrate causing...
Figure 2.4 Comparison of analytically calculated potential (dash-dot line) with 3D FEM Poisson solver (solid line) of a 3x3 uniform metal nanocrystal array and 1D solution (dashed line) for (a) write condition (b) erase condition (c) retention condition (d) write saturation condition. The solid and dash-dot line are almost overlaid.
Figure 2.5 Comparison of analytically calculated potential (dash-dot line) with 3D FEM Poisson solver (solid line) of a 3x3 uniform semiconductor nanocrystal array and 1D solution (dashed line) for (a) write condition (b) erase condition (c) retention condition (d) write saturation condition. The solid and dash-dot line are almost overlaid.
the potential to relax gradually over a Debye length. Therefore, setting the substrate as a ground plane will give higher electric fields than the case of the depleted substrate.

The 3D Poisson solution for a 3×3 nanocrystal case between conducting planes using Maxwell 3D® finite element solver [19] was implemented and solutions have been compared with the analytical results. The root mean square (RMS) error between the analytical solution and the numerical results ranged from 11.3 - 25mV for the tunnel oxide and 11 - 28mV for entire gate stack as shown in Fig. 2.4 for the metal nanocrystal case. This is around the thermal voltage of 25 mV at room temperature. For the silicon nanocrystal case shown in Fig. 2.5, the RMS error is even lower in the tunneling oxide, ranging from 2.5 - 8.6mV, which is most critical for tunneling calculations. The total RMS error ranges from 10.3 - 21.9mV. These results have been compared to electrostatics calculations assuming 1-dimensionality as is popularly used for memory operations modeling [9, 10] and equal stored charge density on floating gate. The RMS error is as high as 0.36V in the case of Si nanocrystals and 0.5V for metal nanocrystals. This is a direct consequence of the inability of the 1D calculation to account for the non-negligible 3D potential variation in the parallel plate capacitor due to the insertion of the 3D nanocrystal. Charging capacitances of nanocrystals are also severely erroneous for 1D calculation.

In terms of computational cost, the non-optimized MATLAB implementation of the analytical formulation took about a tenth of the time of the optimized C-based implementation of a commercial Maxwell solver for the chosen mesh density required for good accuracy in the uniform 3×3 nanocrystal case. Furthermore, the MATLAB program was able to compute the potential and electric field information at any point without interpolation on the grid points as that required in the FEM solver. Finally it maybe computationally prohibitive to calculate realistic nanocrystal distributions in the FEM solver if there are many nanocrystals under the gate stack as the calculation
cost is superlinearly proportional to the number of grid points which scales directly with the geometrical complexity (i.e. volume of gate stack, number, distribution of nanocrystals etc). In comparison, the analytical formulation can be readily extended for the non-uniform nanocrystal array because the number of terms is only linearly proportional to the number of nanocrystals.

The analytical equations provide not only better understanding of the observation of lower write/erase voltage in metal nanocrystal memories [11] due to field enhancement under the nanocrystal but also useful insights into memory design trade-offs based on the electrostatics. In terms of the number density, increased nanocrystal density provides larger storage site density. But there exits a trade-off from the electrostatics point of view. While a single nanocrystal causes field enhancement, other nanocrystals in the vicinity, polarized by $E_0$, produce an electric field which counteracts the polarization by opposing $E_0$ inside the nanocrystal and hence, reducing the field enhancement effect. The nanocrystal number density, size and shape should have a considerable impact on the electrostatics. While size optimization is possible within the model, introduction of arbitrary shapes should require finite element based analysis. Such concerns will be addressed elsewhere. Finally, the model provides a clear separation of various contributions to the electrostatics that can be independently used to study trends and optimize design. To calculate the transport explicitly for P/E operations, the self-consistent solution of the electrostatics with Schrödinger’s equation is required. Given that the electrostatics can be described in a functional form now instead of numerical values at discretization points, there exists a distinct possibility of solving the transport analytically albeit under further simplifying assumptions, which will help design optimization and compact model formation. In any case, integration with a numerical implementation of a Schrödinger solver will be conveniently facilitated, especially for a 1D solver in the
path of least action where the limited electrostatic information required can be more efficiently provided by the analytical formulation, without the calculations involving the entire 3D volume.

2.5. Conclusions

We demonstrate a simple but accurate analytical description of the electrostatics of nano-crystal floating-gate memories. The model has been validated using 3D numerical Poisson solver. Some important scaling issues like nanocrystal size, shape and number density can be derived using the proposed description. Being non-iterative and discretization independent, the analytical solution can provide an accurate description of the electrostatics at minimal computational expense and can be coupled with a Schrödinger solver for tunneling rate calculations. Separation of the contribution of the various sources of potential provides better understanding of trends for design optimization.


CHAPTER 3

MOLECULAR REDOX OF FULLERENES IN SOLID-STATE IN AN EEPROM-TYPE DEVICE

3.1 Abstract

Molecular interface with CMOS is an area indispensable to the enhancement of our understanding of the nano-scale world. We report the integration of fullerenes in CMOS gate stack and demonstrate a functional molecular interface by effecting molecular redox operations through non-volatile charge injection in an EEPROM-type device. The gate stack of the MOS capacitor consists of a tunneling thermal oxide. A sub-monolayer of fullerenes is deposited. Then the control oxide is deposited and finally the gate metal is patterned. Charge injection occurs at a specific potential of the fullerene molecules with respect to the conduction band of Si at the Si/SiO$_2$ interface, independent of the concentration of the fullerene sub-monolayer. This strongly indicates molecular redox in solid state that is electrostatically controllable. Such molecular interfaces can be used to enhance the spatial sensitivity of chemical sensors like the CvMOS to be able to interface with macromolecular systems.

3.2 Introduction

A challenge in nanotechnology is to be able to build interfaces from the macro-world to communicate with the nano-scale world. Molecular interfaces with CMOS, is an area that proposes to use the silicon platform to electronically address molecules. Molecules of carbon, like fullerenes and nanotubes have aroused a great deal of scientific curiosity. Carbon molecules are known for their chemical stability and can be found in myriad forms from spheres to nanotubes. They exhibit a wide range of
electronic properties which can be further modified by using the rich chemistry of carbon. Carbon nanotubes have been used in a wide range of electrical sensing application as well as a logic and memory elements [1]. In the present paper, we report our ability to control the charge state of a carbon molecule embedded in the dielectric of a CMOS gate stack. From a technological point of view, such interfaces can have a tremendous impact. Carbon nanotubes, chemically similar to fullerenes, used as a floating gate as in a C\textsubscript{ν}MOS [2], whose charge is controlled by a CMOS interface, can act as a electrical probe with nanometric spatial resolution. Charge injection of this precise nature constrained by chemistry can also pave way for multi-level charge storage and high density memory.

### 3.3 Experimental Set-up

A MOS capacitor is the ideal set-up to measure charge injection as a function of injection voltage. The structure of the device has been shown elsewhere [3]. Briefly, the device consists of a silicon doped p-type to $2 \times 10^{17}$ cm\textsuperscript{-3} density. After making a standard LOCOS isolation as described elsewhere [3], a 2.3 - 2.7 nm dry thermal oxide is grown on the active area. The fullerenes C\textsubscript{60} and C\textsubscript{70} are evaporated thermally to a few angstroms, typically 4Å and 6Å, as measured by the quartz crystal microbalance (QCM). The thicknesses are calculated from the QCM output using the parameters for graphite. Experimentally we observe that these depositions result in sub-monolayer fullerene thin-films. Evaporated oxide is then deposited without venting the system to form the control oxide of about 20 nm thickness. Finally chromium is patterned to form the gate. For the C\textsubscript{70} experiments, the capacitor structure was changed to ensure simpler process and higher throughput. No isolations structures like LOCOS were used. The entire gate stack was created on an unpatterned doped Si wafer. Finally the gate metal consisting of 40 nm Cr and 10 nm Au was
patterned by lift-off. The Au on Cr helps make good electrical contacts under low temperature conditions when the probes become very brittle. The devices were measured using the Keithley 590 high frequency capacitance voltage (HFCV) measurement system.

### 3.4 Room Temperature Measurements

Room temperature CV measurements reported previously [3], show low interface states and negligible silicon contamination from the integration of fullerenes. However the evaporated oxide show charge injection from the metal gate. This process can be understood from the fact the evaporated oxide has high trap density and the traps can conduct electrons in a Poole Frenkel mechanism. The electron conduction in oxide is hence temperature activated. Therefore, a low temperature would decrease the current conduction and charge injection into the control oxide.

### 3.5 Low Temperature Measurements (T=10K)

Low temperature measurements were performed on a Desert Cryogenics™ probe station. The measurements sample without any fullerenes show no gate injection into the control oxide [3].

**C\textsubscript{60} devices:** The low temperature HFCV measurements were performed to study charge injection into fullerenes (Q\textsubscript{F}) as a function of the charging voltage (V\textsubscript{CH}). The V\textsubscript{CH} was applied for 10s and HFCV measurement was performed. The V\textsubscript{CH} is stepped in steps of 0.5V or finer steps corresponding CV curves were obtained to measure the flat-band voltage (V\textsubscript{FB}) shift. The V\textsubscript{FB} shift is used as a measure the sign and magnitude of the injected charge density as a function of V\textsubscript{CH}. Two different depositions of C\textsubscript{60} were measured. Sample A contains 4Å thick C\textsubscript{60} sub-monolayer and
Figure 3.1 (a) CV measurements for electron injection into fullerenes based MOS capacitors at T=10K showing non-uniform CV curve shift i) 4Å C$_{60}$ ii) 6Å C$_{60}$ iii) C$_{70}$ (b) Flat band voltage vs. charging voltage reflecting non-uniform charging i) 4Å C$_{60}$ ii) 6Å C$_{60}$ iii) C$_{70}$
Figure 3.2 (a) CV measurements for hole injection into fullerenes based MOS capacitors at T=10K showing non-uniform CV curve shift i) 4Å C$_{60}$ ii) 6Å C$_{60}$ iii) C$_{70}$ (b) Flat band voltage vs. charging voltage reflecting non-uniform charging i) 4Å C$_{60}$ ii) 6Å C$_{60}$ iii) C$_{70}$
sample B contains 6Å of C₆₀. Electron injection experimental results are shown in Fig.
3.1 (a) (i) and (ii). Both devices show clear aggregation of HFCV curves for certain
\( V_{CH} \), which means charge injections are forbidden for this \( V_{CH} \) range. This non-
uniform charging is reflected in the \( V_{FB} \) vs. \( V_{CH} \) curves shown in Fig. 3.1 (b) (i) and
(ii) extracted from the HFCV measurements. For hole injection, the data is not that
clear for sample A as shown in Fig. 3.2 (a) (i) but the sample B as shown in Fig. 3.2
(a) (ii) shows clear aggregation of the CV curves for certain \( V_{CH} \) and this reflects in
the \( V_{FB} \) vs \( V_{CH} \) plot as shown in Fig. 3.2 (b) (ii).

**C₇₀ devices:** The devices with C₇₀ show very similar aggregation at a different \( V_{CH} \)
range both for electrons as shown in Fig. 3.1 (a) (iii) and holes as shown is Fig. 3.2 (a)
(iii). The \( V_{FB} \) vs. \( V_{CH} \) curves, in Fig. 3.3 (a), reflect similar non-uniform CV curve
shift as a function of \( V_{CH} \).

### 3.6 Discussion

There Fig. 3.1 (b) and 3.2 (b) show from the \( V_{FB} \) shift that charge is being
injected into the oxide at different \( V_{CH} \). However, to be sure that the charge is stored
in the molecule, we have to understand the injection process. At thermal equilibrium,
provided by shining UV radiation, the molecule is at electronic equilibrium with
silicon. Conduction band (CB) edge is populated by electrons at low temperature
under inversion. Any injection of electrons into the fullerene from the Si should occur
when the molecular orbital energy levels align with the CB edge of Si at the Si/SiO₂
interface. Hence, when we scan \( V_{CH} \), we are essentially scanning the energy levels of
the fullerene across the CB edge of Si at the Si/SiO₂ interface, which is where the
injected charges originate. So if we can calculate the voltage on the fullerene (\( V_F \))
during injection with respect to the CB of Si at the Si/SiO₂ interface as a function of
\( V_{CH} \) and plot the charge injection into fullerene (\( Q_F \)) vs. \( V_F \), we should obtain a step
function. To extract $V_F$ and $Q_F$, we consider the band diagram shown elsewhere [3].

The total charge in the oxide $Q_F$ can be computed from equation (3.1). The total charge on the gate during injection $Q_T$ and eventually the $V_F$ can be computed from the equations (3.2) and (3.3), where $V_{FB0}$ is the ideal flat band voltage, $C_{CH}$ is the capacitance of the capacitor during injection, $C_{OX}$ is the total oxide capacitance and $C_{tunnel-ox}$ is the capacitance due to a parallel-plate capacitor with the tunneling oxide.

\[ Q_T = C_{CH} V_{CH} \]  
\[ Q_F = (V_{FB} - V_{FB0}) C_{OX} \]  
\[ V_F = \frac{Q_T - Q_F}{C_{tunnel-ox}} \]

The $Q_F$ vs. $V_F$ plots extracted from the various measurements are shown in Fig. 3.3. At the onset it is important to note a few things. The electron injection for different concentrations of $C_60$ shows clear steps. The potential for the steps are independent of concentration at 0.8V [5]. For the hole injection, the sample B shows some blockade. The Fig. 3 (a) (ii) is probably dominated by traps. The $C_{70}$ devices show clear steps both for electron and hole injection. The observation of step charging clearly indicates a very special configuration. Firstly, the blockade part of the step shows that traps are not active in the charge injection process. Secondly, concentration independent $V_F = 0.8V$ specific injection implies that the electrons are being injected simultaneously into a specific energy state at a surface density of $1.5 \times 10^{12}$ cm$^{-2}$ across the 250 µm x 250 µm area of a capacitor. Thirdly the existing traps are not participating at low temperatures as they were at room temperature [3] probably because injection into the small cross-sectional area is much more inefficient [4]. Hence the charge injection occurs into the molecule. Electrons and holes are, thus, tunneled into the fullerenes and the observed phenomenon is clearly molecular redox in solid state. Previously, we
Figure 3.3 Charge on fullerene vs. potential on fullerene with respect to CB edge of Si at Si/SiO\textsubscript{2}  

a) electron injection for samples i) 4Å \textit{C}_{60} ii) 6Å \textit{C}_{60} iii) \textit{C}_{70}  
b) hole injection for samples i) 4Å \textit{C}_{60} ii) 6Å \textit{C}_{60} iii) \textit{C}_{70}
have shown that the energy values obtained for the C$_{60}$ based device steps is reasonable [3]. Here it is important to note that the lower symmetry of the C$_{70}$ should be responsible for lower sharpness of the steps. Orientation of molecule may have effect on injection.

3.6 Conclusion

We report the first successful integration of two different fullerenes in the CMOS gate stack. The integrity of fullerenes is indicated by the observation of clear steps during charge injection. Finally we have experimentally observed molecular redox. Such controlled charge injection into molecules in the solid state demonstrates a functional interface between macroscopic devices and the molecular world.
REFERENCE


Note: There is a numerical discrepancy in the $V_F$ extraction with [3]. In our previous work [3], tunneling oxide thickness was measured in as 2-3 nm using 3 wavelength Rudolph Ellipsometer. The calculations were done using 2 nm tunnel oxide. Currently, better measurement using a spectroscopic ellipsometer shows that the oxide is 2.7 nm for the C$_{60}$ experiments and 2.4 nm for the C$_{70}$ experiments. This does not disturb our arguments that the injections occur into molecular levels of the fullerenes since the different tunneling oxide thickness just scales $V_F$ to a different value. This $V_F$ is same for all devices with the same fullerene when the same tunnel oxide thickness is used for $V_F$ extraction.
4.1 Abstract

Ultra thin and narrow semiconductor body on insulator allows aggressive scaling of nonvolatile memories for low power, low read/write voltage, high retention and high density in comparison with bulk devices. We have fabricated memory cells with single-wall carbon nanotubes as channels and gold nanocrystals as charge storage nodes. The devices have large memory windows with low voltage operations and single-electron-controlled drain currents. Coulomb blockade in nanocrystals combined with single charge sensitivity of the nanotube field-effect-transistor can potentially enable multi-level operations. Measured retention time is longer than 6200 s at 10 K, but is only about 800 s at room temperature due to the high leakage in evaporated tunnel oxide used in this study. Better dielectric on nanotubes is expected to greatly improve the room temperature performance for the nanotube memory device.

4.1 Introduction

The growth potential for nonvolatile memory has driven the scaling of flash memories towards ever higher bit density operations [1]. Since the invention of discrete charge storage in the metal-oxide-semiconductor (MOS) field-effect-transistor (FET) structure [2], material integration, such as metal nanocrystals [3] and high-k dielectrics [4], has been used to enhance tunneling oxide reliability and charge retention. To optimize the device scaling from the transistor structure point of view, channel modifications such as ultra-thin body silicon-on-insulator (SOI) [5], double
gate [6] and FinFET [7], have been used to improve the electrostatic coupling and to minimize the short channel effects. However, body thickness variation from etch non-ideality has raised concerns on process control and parameter yield. Single-wall carbon nanotubes (SWNTs) with mono-dispersed nanoscale cross-sections have been proposed to resolve the body thickness variation with trap based charge storage in nitride [8,9], but the memory window is relatively small with high operational voltage due to the limited density and cross section of the nitride traps. In this letter, we report a new nonvolatile memory structure based on a SWNT sensing channel and metal nanocrystals as charge storage media. The gate electrode regulates the charging and discharging of the metal nanocrystals from the channel. Each charged nanocrystal imposes an extra local potential on the SWNT channel, and hence alters its electrical conduction. We observed that the SWNT channel could sense the charge on the nanocrystals with precision down to the single-electron level.

4.2 Experimental

The proof-of-concept device consisted of a back-gated carbon nanotube field-effect-transistor (CNTFET) [10] with metal nanocrystals embedded in the dielectric layer near the SWNT to enable the nonvolatile memory function (Fig. 4.1a). The CNTFET was fabricated using a process similar to that described previously [10]. Dry thermal oxide was grown at 1100ºC to a thickness of 100 nm on a heavily doped p-type Si wafer with surface resistivity of 0.005 Ω-cm. SWNTs of diameters 1.1 - 3.7 nm were grown at 950ºC in methane by chemical vapor deposition (CVD) from patterned catalyst pads. Source/drain lead structures were photo-lithographically patterned using 5 nm Cr and 50 nm Au layers shown in Fig. 4.1 (b). For the memory superstructure, a 5 nm SiO₂ layer was deposited on the CNTFET using electron-beam evaporation. The evaporated oxide thickness could be accurately controlled to a few.
Figure 4.1 (a) A cross section illustration of SWNT-based memory device; (b) The AFM image of the boxed area in inset shows SWNT between Au source/drain leads; inset shows the layout micrograph; (c) AFM shows self-aligned nanocrystals formed over a pair of crossed SWNT and another on the left.
angstroms with variation less than 20% across the wafer measured by ellipsometry and root-mean-square (RMS) roughness of 2.6 Å measured by atomic force microscopy (AFM). The Au nanocrystal layer was formed by self-assembly from an electron-beam evaporated Au thin-film of 1.2 nm nominal thickness [11], and covered by a capping passivation SiO$_2$ layer of 30 nm thickness using plasma enhanced CVD. As a final step, the source/drain and gate pads were opened for measurements by reactive ion etching (RIE).

4.4 Physical Characterization

The self-assembled nanocrystals were characterized using scanning electron microscopy (SEM) and AFM. SEM for nanocrystals on thermal oxide showed an 8 nm average diameter. The AFM image, as shown in Fig 4.1 (c), indicated a certain degree of self-alignment of nanocrystals on top of SWNT. This self-alignment phenomenon could be explained by the preferred nanocrystal formation at a convex SiO$_2$ surface from SWNT underneath due to surface energy minimization [3]. This effect should relieve the nanocrystal number fluctuation problem when the channel length is scaled down to tens of nanometers.

4.5 Electrical Characterization

Room-temperature measurements for SiO$_2$ encapsulated CNTFETs without Au nanocrystals showed good p-type FET characteristics with on/off current ratio larger than $10^4$. Writing of nonvolatile charges was performed with various charging voltages ($V_{CH}$) for 5 s. Drain current ($I_D$) – gate voltage ($V_G$) sweeps were performed to determine the threshold voltage $V_T$. Fig. 4.2 (a) and (b) show the measurement results at T=300K of a control CNTFET device without any metal nanocrystals and a device
with Au nanocrystals, respectively. The relative $V_T$ shift efficiency, defined by $\Delta V_T/\Delta V_{CH}$, shows a larger value of 0.67 for CNTFET with nanocrystals in comparison to 0.4 for the control device. The memory effect in the control device was attributed to the traps in the evaporated tunneling oxide used in this study as a result of facility limitation due to carbon and catalyst metal contamination restrictions. High quality dielectric deposition on nanotubes had been demonstrated in other types of top-gate structures [12,13], which should greatly reduce the oxide trap density and trap-assisted leakage current if used in the present memory device structure [11]. Nevertheless, the trap assisted leakage current follows the Poole-Frenkel conduction model which is thermally activated. Hence, in low-temperature measurements we could eliminate the effect of oxide traps and obtain the memory characteristics intrinsic to the tunneling between the SWNT and metal nanocrystals.

Low-temperature $I_D-V_G$ measurements for various charging voltages were performed at 10 K. Negligible $V_T$ shift was observed for control devices without nanocrystals (Fig. 4.2 (c)), indicating that the oxide traps were inactive at 10 K. On the other hand, large hysteresis with $\Delta V_T/\Delta V_{CH}$ about 0.5 due to charge injection into nanocrystals was consistently observed in memory devices with nanocrystals (Fig. 4.2 (d)). This charging efficiency was much higher than that in other trap-based CNT memory devices [9], where $\Delta V_T/\Delta V_{CH}$ was about 0.02, comparable to the efficiency we observed in the control device. The fact that tunneling into nanocrystals is much more preferred to traps is mainly because the nanocrystal has a much larger capture cross-section. In the absence of inelastic phonon scattering at low temperatures, cross-sectional effects dominate in tunneling efficiency [14].

An important difference between 1D and 2D channels with discrete charge storage is that the 1D channel current is mainly controlled by the maximum barrier in the channel and hence can be modulated by a single charged nanocrystal near the 1D
Figure 4.2 Hysteresis measurements for the single sweep direction from $V_{CH}$ to 2 V. Control sample measurements are shown in (a) at $T=300K$ with $\Delta V_T/\Delta V_{CH} = 1.2$ V / 3 V=0.4, and (c) at $T=10K$ with $\Delta V_T/\Delta V_{CH} = 0.2$ V / 2.5 V=0.08. Nanocrystal CNTFET measurements are shown in (b) at $T=300K$ with $\Delta V_T/\Delta V_{CH} = 2$ V / 3 V=0.67 and (d) at $T=10K$ with $\Delta V_T/\Delta V_{CH} = 1.4$ V / 3.4 V =0.41.
Figure 4.3 After the metal nanocrystal CNTFET device was charged at $V_{CH}=-5.5\text{V}$ for 10s, current relaxation in time was measured with $V_G=-1\text{V}$ and $V_{SD}=300\text{mV}$ at 10K and 300K: (a) log ($I_D$) vs. time; and (b) log ($I_D$) vs. log (time) for the same measurement. Steps in $I_D$ at T=10K indicated single hole discharge events from charged nanocrystals.
Figure 4.4 (a) Advancing $V_{CH}$ in 50 mV steps shows aggregation of $I_D-V_G$ curves; (b) Extraction of $V_G$ at constant $I_D=0.95$ nA resulted in steps in $V_G$ due to the combined effect of Coulomb blockade in nanocrystals and single charge sensitivity of SWNT conductance.
channel, known as the “bottleneck” effect [5]. In the 2D channel case, the least resistive path controls the current in the percolation process [15]. In our device structure, the “bottleneck” effect is even stronger due to the geometrical difference between the nanocrystals and CNT, which further contributes to the large memory window and the single-electron sensibility of the device.

The signature of the single-electron effect was observed in the charge retention characteristics. We charged the nanocrystals with holes by applying $-5.5 \text{ V}$ to $V_G$ for 10 s, and then set $V_G$ to -1 V and the drain voltage $V_D$ at 0.3 V while monitoring $I_D$ as a function of time. In Fig. 4.3, $I_D$ did not relax exponentially in time but clear steps were observed which indicated discrete discharge from the nanocrystals. The high noise in $I_D$ was postulated as 1/f noise reported in SWNT [16], which could be strong at 1 Hz. Nevertheless, single charge in nanocrystals clearly affected the SWNT conductance. To investigate charge injection as a function of $V_{CH}$, $V_{CH}$ was varied from -2 V to -5.4 V in 50 mV steps. For each $V_{CH}$, $I_D-V_G$ curves were measured with $V_G$ sweeping from $V_{CH}$ to 2 V in 10 mV resolution. In Fig. 4.4 (a), kinks were observed in the $I_D-V_G$ curves where $I_D$ switched tracks rightward to follow the previous $I_D-V_G$ curves with less initial $|V_{CH}|$. These kinks represented single-charge loss in the “bottleneck” nanocrystal during $I_D-V_G$ measurements. Similar single charge conductance modulation was also observed in narrow Si-channel nanocrystal memory [17]. Coulomb blockade in the nanocrystals was noticed by the clear aggregation of $I_D-V_G$ curves shown in Fig. 4.4 (a). The readout voltages at an arbitrarily fixed current level of 0.95 nA was plotted in Fig. 4.4 (b), where step-like $V_G$ shift for different $V_{CH}$ was an alternative way to show the discrete charging/discharging effect. The eye openings in Fig. 4.4 (a) and the steps in Fig. 4.4 (b) of about 0.13 V suggest that the current is controlled by the nanocrystal with the most effective potential barrier due to channel location, number of charge and proximity. A discharge of the most critical
“bottleneck” nanocrystal would hand the control of $I_D$ to the second “bottleneck”, giving a $V_T$ shift equal to the difference of their potential barriers on the CNTFET. Simulation based on the non-equilibrium Green’s function [18] reveals that a single charge could cause a $V_T$ shift of about 90 mV for the experimental geometry [19], which further strengthens our postulation.

As shown in Figs. 4.3 (a) and (b), the low-temperature retention was better than 6200 s, however, the present poor retention at room temperature (about 800 s for 1 V memory window) was due to the trap-laden evaporated oxide, which provided charge leakage paths from nanocrystals to the SWNT channel. This should not be viewed as a fundamental limitation as discussed previously. On the contrary, after process optimization and work function engineering [20], nanocrystal CNTFET memories could have excellent retention characteristics in view of the bottleneck effect [5].

4.6 Conclusions

In conclusion, we have demonstrated a metal nanocrystal CNTFET memory with high $V_T$ shift efficiency and large memory window. The device shows clear single-electron sensitivity and Coulomb blockade charging, which could potentially enable multi-level data storage. Room temperature devices require the integration of tunneling dielectrics with better quality. Further improvement of the present device structure, for example, by using a more efficient top gate stack, could eventually lead to high density, low voltage and low power nonvolatile memory applications.
REFERENCE


[18] Unpublished communication with Jing Guo, University of Florida.

CHAPTER 5

HIGH SENSITIVITY AND NON-LINEARITY OF CARBON-NANOTUBE-BASED CHARGE SENSORS

5.1 Abstract

A self-consistent atomistic simulation method is established to investigate carbon nanotubes (CNTs) as the charge sensing channels in nonvolatile memory and sensor applications. The theoretical simulation agrees with the experimental results, demonstrating that the sensor can achieve high sensitivity down to a single-electron charge at room temperature. The electrical response of the CNT charge sensor strongly depends on the channel position of a small amount of discrete sensing charges, due to the near ballistic transport and quantum interference in a submicron meter CNT channel. This result differs from those for conventional CMOS charge-based sensors, in which carrier transport is diffusive and electrical response is expected to be independent of the position of the sensing charge. The high sensitivity and nonlinearity of CNT charge-based sensors are important features that need to be considered in the design of CNT memories and sensors.

5.2 Introduction

The excellent electrical, optical, and mechanical properties of single-wall carbon nanotubes (CNTs) have made them attractive for various device applications [1]. CNTs have been recently explored for applications of biomolecular sensors [2] and non-volatile memory cells [3, 4]. In both applications, the CNT channel acts as a charge-based sensor, whose conductance varies with the amount of charge in the nanocrystals or the biomolecules. A CNT charge sensor has several distinct
advantages. First, electrostatic coupling between the CNT and the environment is very strong because all carbon atoms of the CNT channel are located on the surface. Second, the potential barrier created by sensing charge has a bottleneck effect on the one-dimensional carrier transport, which results in a large modulation on the channel conductance. Despite rapid experimental progress on CNT biomolecular sensors and nanocrystal memories [2-4], little is known about linearity of the sensors. Device simulators of CNT sensors have not been available, which prevents theoretical investigation of device physics and exploration of design optimization.

In this paper, a self-consistent atomistic simulation approach is established for detailed simulation of CNT charge-based sensors. We demonstrate single electron sensitivity at room temperature through both theoretical calculations and experimental measurements. Due to near ballistic transport and quantum interference of electron waves in CNTs, the simulated sensor response strongly depends on the channel position of the sensing charge. As a result, the sensor behaves nonlinearly to a small amount of discrete charges randomly distributed along the channel. This work establishes the theoretical foundation for designing CNT charge-based sensors in memory and sensor applications.

5.2 Approach

Fig. 1a shows the modeled device structure of the CNT charge sensor. The device consists of a back-gated CNT field-effect transistor (FET) and a charge dot buried in the top oxide. It is similar to the experimental memory cell [3], where the CNT is used as the sensing channel and metal nanocrystals as the charge storage elements. The intrinsic (22, 0) zigzag CNT channel has a channel length of $L_{ch}=200\text{nm}$, and a diameter of $d_{CNT}=1.7\text{nm}$, which results in a band gap of $E_g=0.5\text{eV}$. The metal/CNT Schottky barrier height for electrons is $\Phi_{bn}=0.12\text{eV}$. The charge dot is buried in $\text{SiO}_2$
with a dielectric constant of $\kappa = 3.9$. The bottom SiO$_2$ thickness is $t_{\text{bot}}=20\text{nm}$. The simulations were performed at the room temperature, $T=300\text{K}$. The above values are nominal, and are varied parametrically to explore different issues.

In a CNT-based memory cell or a biomolecular sensor, the sensing charge distributes on metal nanocrystals or biological molecules. We modeled the following two limits for the dielectric response of the sensing charge. (i) The sensing charge is distributed on a metallic ball to maintain an equal potential, which is relevant to nanocrystal memory applications [3]. (ii) The sensing charge is uniformly distributed on a dielectric ball, which is relevant for biomolecular sensor applications [2].

The mean free path (mfp) in CNTs at low source-drain bias is long ($\sim 1\mu\text{m}$) [6]. Ballistic transport was assumed for the simulated submicron CNT channel of $L_{ch}=200\text{nm}$. An atomistic description of the CNT with a p$_Z$-orbital basis set was used. Atomistic simulation in real space is computationally expensive, but significant computational savings can be achieved using a mode space approach [7]. The mode space approach is valid when the potential variation around the tube is smaller than the spacing between the neighboring subbands ($\sim 0.25\text{eV}$). A phenomenological treatment of the metal/CNT contacts was used [7]. The model well describes the experimentally measured data for a short-channel CNTFET [8]. The retarded Green’s function of the CNT channel is computed as,

$$G(E) = [EI - H - U - \Sigma_S - \Sigma_D]^{-1}, \quad (5.1)$$

where $H$ is the Hamiltonian matrix of the CNT channel, $U$ is the self-consist potential, and $\Sigma_{S,D}$ is the source/drain self-energy matrix for the metal-CNT contacts. After the retarded Green’s function is obtained, the local density of states (LDOS) are computed
and filled according to the Fermi levels of the two contacts. The charge density of the CNT channel is calculated by integrating the filled LDOS over energy [9].

The electrostatic potential in Eqn. 5.1 was computed by self-consistently solving a 3D Poisson equation with the following boundary condition. The electrostatic potential on the source, drain and bottom gate are fixed at $U_S$, $U_D$, and $U_G$, respectively, as determined by the bias condition. The charge on the charge dot and that on the CNT channel (as computed by the quantum-transport equation), are input parameters to the Poisson equation. The boundary element method [10] was used to efficiently solve the 3D Poisson equation. Discretization was performed on the surface of the source, drain, and gate electrodes, as well as the CNT channel and nanocrystals. The boundary element method relates the potential, $V(\vec{r})$, to the charge density, $\rho(\vec{r}')$, through electrostatic kernel, $K(\vec{r} - \vec{r}')$ [11],

$$V(\vec{r}) = \int d\vec{r}' \rho(\vec{r}') K(\vec{r} - \vec{r}'), \quad (5.2)$$

from which the potential of the CNT channel was solved.

The iteration between the NEGF transport equation and the 3D Poisson equation continues until self-consistency is achieved. The source-drain current is then computed as [9],

$$I_D = \frac{4e}{h} \int dE \cdot T(E) \cdot \left[ f_0(E - E_{FS}) - f_0(E - E_{FD}) \right], \quad (5.3)$$

where $T(E)$ is the channel transmission computed by the NEGF formalism, $f_0$ is the Fermi distribution function, $E_{FS}$ and $E_{FD}$ are the source and drain Fermi levels, respectively. The channel conductance at zero source-drain bias is computed as,

$$G_D = \frac{\partial I_D / \partial V_D}{|_{V_D=0}} = \frac{4e^2}{h} \int dE T(E) \left[ - \frac{\partial f_0(E - E_F)}{\partial E} \right], \quad (5.4)$$

where $E_F = E_{FS} = E_{FD}$ is the Fermi level.
5.3 Results and Discussion

We first explore the sensitivity of the charge sensor as shown in Fig. 5.1 (a). Fig. 5.1 (b) plots the simulated channel conductance vs. the bottom gate voltage on the log scale (left) and linear scale (right) for different number of electrons on the charge dot. The log scale plot indicates that addition of a single electron charge increases the threshold voltage, $V_T$, by about 45meV, which is about 2 times larger than the thermal voltage at the room temperature. (The threshold voltage is here defined as the gate voltage at which the channel conductance reaches $G_D=0.1\mu\text{S}$. Notice that the log($I_D$) vs. $V_G$ curves in Fig. 5.1 (b) parallel with each other at low $V_G$, the threshold voltage shift is insensitive with the specific value of $G_D$ used to define $V_T$.) The linear scale plot shows that addition of a single electron charge results in considerable change of channel conductance when the device is biased near the threshold voltage. Because varying the amount of sensing charge by a single electron results in a $V_T$ shift larger than the thermal voltage, the simulation results indicate that the sensor has the potential to achieve single-electron sensitivity at room temperature.

The log scale plot in Fig. 5.1 (b) shows that variation of $Q_d$ from 0 to $1e$ approximately results in the same $V_T$ shift as the further single-electron addition events. The linear scale plot, however, shows that the change of channel conductance is smaller. Fig. 5.1 (c), which plots the first conduction and valence band profiles, explains the reason. The channel conductance is limited by either metal/CNT Schottky barriers or the sensing barrier at the middle of the CNT. When the CNTFET is biased above the threshold ($V_G>V_T$) and the middle barrier is small (for $Q_d=0$ to $1e$), the resistance of the larger source/drain Schottky barriers dominates. Varying the amount of the sensing charge only changes the channel conductance slightly. When more electrons are in presence on the charge dot, the sensing barrier plays a more important role, which leads to a larger reduction of the channel conductance.
Figure 5.1 A CNT charge sensor with a single charge dot at the middle of the channel. (a) A schematic plot of the cross section of the modeled device. (b) The simulated channel conductance vs. the gate voltage on log (left) and linear (right) scales for different amounts of charge on the dot, $Q_d=0$ to $3e$, at $e/\text{step}$. (c) The first conduction and valence bands versus the position along the channel for $Q_d=0$ to $3e$ at $V_G=0.2\text{V}$. The source and drain Fermi level is set at $E_F=0$. The diameter of the dot is 8nm.
Figure 5.2 (a) The simulated $V_T$ shift due to single electron addition to the charge dot versus the bottom oxide thickness when the charge dot is modeled as a metallic ball (the solid line) and a uniformly charged ball (the dashed line). (b) Room temperature measurement of the CNT memory device right after charging (the curve on the left) and fully discharged (the curve on the right). The sub-threshold slope (dashed line) should be the same in both curves. The $I_D-V_G$ curve for the charged device, however, shows a series of threshold voltage shift with magnitude of ~100mV-200mV whose direction suggests that these shifts are due to discharge events during measurements.
In order to confirm that single-electron sensitivity can be achieved, we compared the simulated threshold voltage shift, $\Delta V_T$, to the experimentally measured value for a CNT-based metal nanocrystal memory cell [3]. Fig. 5.2 (a), which plots the threshold voltage shift vs. the bottom oxide thickness $t_{ox}$, shows that when the bottom oxide thickness increases, $\Delta V_T$ increases. The reason is that thicker bottom oxide results in weaker electrostatic coupling between the control gate and the channel, and requires a larger $\Delta V_T$ to compensate the same addition of a single electron charge on the dot. For the experimental bottom oxide thickness of $t_{ox} \approx 100\text{nm}$, the simulated $\Delta V_T \approx 65\text{meV}$ at room temperature when the metal nanocystal is modeled as a metallic ball. The simulations at a low temperature result in much steeper $G_D-V_G$ curves, but a similar $\Delta V_T$. The simulated results are in reasonable agreement with the experimentally measured $\Delta V_T \approx 100\text{meV}$ at low temperature [3].

While direct extraction of $V_T$ shift from the measured $I_D-V_G$ curves due to single electron discharging from the nanocrystal to the CNT is possible at low temperature (T=10K) [3], this technique can not be used to establish single electron sensitivity at room temperature. The reason is that the Poole-Frenkel charging and discharging events [12] from the oxide traps are much more significant at room temperature, and the $I_D-V_G$ curves are much more noisy. A new technique was developed in this study to experimentally establish single electron sensitivity at room temperature. We compare the $I_D-V_G$ measurements for the CNT nonvolatile memory cell [3] at the charged state (right after the charging operation) and the fully discharged state (devices relaxed to equilibrium after a long storage in clean inert ambient), as shown in Fig. 5.2 (b). The subthreshold slope should be similar for both cases. While the discharged device shows the subthreshold slope clearly, the charged device shows the same subthreshold
slope in segments that depict parallel shifts in the \( I_D-V_G \) curve. The parallel shifts with a magnitude of \(-100-200\text{mV}\) can be due to the single-electron discharge events from either a metal nanocrystal or a closer oxide trap to the CNT, which results in a larger \( V_T \) shift. In the literature, room temperature single-electron sensitivity has also been demonstrated for quantum-dot and nanowire devices using complex fabrication techniques [13].

The high sensitivity makes it possible for a CNT channel to sense a small amount of discrete charges possessed by one or a few storage nodes near the channel. We next explore how the sensor response depends on the position of single or a few sensing electrons. The solid line in Fig. 5.3 (a) shows the channel conductance vs. the channel position of a charge dot modeled as a metallic ball with \( Q_D=2e \) for the device as shown in Fig. 5.1 (a). The dashed line shows the channel conductance without the charge dot \((Q_D=0)\) for reference. The electrical response depends strongly on the position of the charge dot, especially when the dot is located near the center along the channel. Fig. 5.3 (b), which plots the conduction band profile and channel transmission without the charge dot, shows that resonant tunneling through double barriers at the metal/CNT contacts results in closely spaced peaks in the transmission vs. energy curve. Fig. 5.3 (c) plots the band profile and transmission when the charge dot is located at \( x=L_{ch}/10 \). A potential barrier is created by the charge dot and the channel is divided into two quantum wells. The transmission depends on the alignment of the energy levels in two quantum wells. Because the width and depth of two quantum wells are quite different, the transmission at low energies significantly reduces. The channel conductance is determined by the transmission near the source/drain Fermi level, \( E_F=0 \), and the reduced transmission at low energies result in a drop of the channel conductance. In contrast, if the charge dot is located at the center of the channel, as shown in Fig. 5.3 (d), the two quantum wells are symmetric and the
energy levels in the wells exactly aligns, which results in a significantly larger channel conductance. The dependence of the electrical response on the sensing charge position stems from near ballistic transport and quantum interference in a submicron CNT channel. Changing the charge position modifies the potential profile and electron wave function in the channel, and therefore, changes the channel conductance.

In order to see whether the position-dependent response remains when multiple charge dots are in presence, we simulated a charge sensor with two charge dots (with \( Q_1=Q_2=2e \)), as shown in Fig. 4a. The channel position of dot 1 is fixed at \( x_1=L_{ch}/3 \), and the position of dot 2, \( x_2 \), is varied. The solid line in Fig. 5.4 (b) plots the channel conductance vs. \( x_2 \). The dashed line is the channel conductance when only dot 1 exists (\( Q_2=0 \)) for reference. The electrical response still strongly depends on the channel position of dot 2. Surprisingly, the channel conductance increases if an additional charge dot is added at the right position, \( x_2=2/3L_{ch} \). This phenomenon can be again explained in terms of resonant tunneling through multiple barriers, as shown by Figs. 5.4 (c) and 5.4 (d), which plot the band profiles and transmissions for a single dot and that for two dots, respectively. The width and the depth of two quantum wells are quite different when only one barrier is created by dot 1 at \( x_1=L_{ch}/3 \) (as shown in Fig. 5.4 (c)), which diminishes the alignment of energy levels in the two wells and reduces the transmission. After an additional barrier is imposed by dot 2 at \( x_2=2/3L_{ch} \), the widths of three quantum wells in the CNT channel are the same, which facilitates the alignment of eigenenergy levels in the quantum wells. The transmission and channel conductance increase.

For a conventional CMOS charge sensor, carrier transport in the channel is diffusive, and we expect that the electrical response of the sensor is independent of the position of the sensing charge. In contrast, carrier transport in a submicron meter CNT channel at low biases is near ballistic. Quantum interference occurs in the channel
Figure 5.3 Position-dependent response of the sensor. (a) The channel conductance vs. the position of the charge dot when $Q_D=2e$ for the device as shown in Fig. 1a. The dashed line shows the channel conductance with $Q_D=0$ for reference. The first subband profile (the solid line) and the channel transmission versus energy (the dashed line) for (b) no charge dot, (c) a charge dot at $x_{dot}=L_{ch}/10$, and (d) a charge dot at $x_{dot}=L_{ch}/2$. The gate voltage $V_G=0.2V$ and the drain voltage $V_D=0$. 
Figure 5.4 A CNT charge sensor with two charge dots. (a) The simulated device structure. (b) The channel conductance versus the position of dot 2, $x_2$. Dot 1 is fixed at $x_1=L_{ch}/3$. The solid line is for $Q_1=Q_2=2e$, and the dashed line is for $Q_1=2e$ and $Q_2=0$. (c) The first subband profile (the solid line) and the channel transmission (the dashed line) for $x_1=L_{ch}/3$, $Q_1=2e$, and $Q_2=0$. (d) The first subband profile and transmission for $x_1=L_{ch}/3$, $x_2=2/3 L_{ch}$, and $Q_1=Q_2=2e$. The gate voltage $V_G=0.2V$ and the drain voltage $V_D=0$. The charge dots are modeled as metallic balls.
because the phase of electron wave is preserved. As a result, the electrical response of the CNT sensor strongly depends on the position of the sensing charge. For an ideally linear sensor, electrical response should vary linearly with the total amount of sensing charge, regardless with the position of the sensing charge. In this sense, a CNT charge sensor is nonlinear because of its charge-position-dependent response. This nonlinearity is more significant for single or a few electron sensing, where the locations of a few discrete charge are difficult to be controlled.

A simple estimation shows that the single electron charging energies of the CNT segments separated by the sensing barriers are in the same order of magnitude as the room-temperature thermal energy for the modeled devices. The single electron effect, therefore, can be important. The NEGF formalism, which treats quantum effects and 3D electrostatic effects, however, does not treat the single electron effect. Single electron charging results in decrease of the channel conductance due to Coulomb blockade. As the charge dot moves from the center of the channel to the end of the channel in Fig. 5.3, the length of the shorter CNT segment decreases. The conductance decreases more due to smaller charging capacitance and more severe Coulomb blockade. The single electron effect would rather increase the position-dependent response of the CNT sensing channel.

5.4 Conclusions

Both theoretical and experimental results show that ultra-high sensitivity down to a single electron charge can be achieved by a CNT sensing channel at room temperature. The electrical response of the sensor strongly depends on the position of the sensing charge because of near ballistic transport and quantum interference in a submicron CNT channel. This phenomenon is different from the conventional CMOS sensor, where carrier transport through the channel is diffusive, and is especially
important for single or a few electron sensing. High sensitivity and nonlinearity of CNT-based charge sensors are important features that need to be considered in future design of CNT-based nanocrystal memories and biological sensors.
REFERENCE


CHAPTER 6

ASYMMETRIC ELECTRIC FIELD ENHANCEMENT IN NANOCRYSTAL BASED NANOTUBE/NANOWIRE MEMORIES

6.1 Abstract

The metal nanocrystal (NC)/carbon nanotube (CNT) based non-volatile memory has been proposed recently in comparison to the micro-fabricated Si channel and Si NCs in ultra-narrow channel structure. The electrostatics of metal NC/CNT devices during memory operations differ significantly from the metal NC memory with planar silicon channel. In this paper we present the theoretical analysis on the three-dimensional (3D) electrostatics of a NC/CNT device during memory operations, to illustrate the experimentally-observed large number of charge storage at low gate bias (5V) despite a 100 nm thick bottom-gate control dielectric. NCs are electrostatically more strongly coupled to the two-dimensional (2D) gate electrode than to the one-dimensional (1D) channel, even when the NCs are in much closer proximity to the 1D channel for efficient tunneling and low-voltage program operation. Under the retention condition, the NC/CNT devices have lower electric field across tunneling oxide than that in the case of a 2D channel. This enhanced electric field asymmetry with respect to program versus retention operations indicates better ratio between retention/program times. Together with the large number of electrons stored per NC, this asymmetry can be utilized either to reduce the operating voltage or to reduce statistical fluctuation of the information storage.
6.2 Introduction

In comparison with DRAM (dynamic random-access memory), EEPROM (electrically programmable erasable read-only memory) has significant advantages on geometrical scaling due to its 1-transistor cell and current readout mechanism. For EEPROM, low voltage operation, power scaling, and enhanced retention/program time ratio are the main challenges in technology scaling. Discrete nanoscale floating gates by semiconductor and metal nanocrystals (NCs) have been proposed to replace the thin-film floating gate for reliable tunneling oxide scaling in low voltage operations [1-5]. A similar device with charge stored in Si$_3$N$_4$ traps rather than NCs has also been explored extensively [6]. Metal NCs have advantages over semiconducting NCs from both electrostatic (larger electric field asymmetry) and transport perspectives (larger density of states) [2-4]. High-k dielectrics have been used to replace SiO$_2$ in the tunneling and control dielectrics [4, 5] to further optimize the tunneling barrier and electrostatics. Another important structural variation explored has been the ultra-narrow sensing channel, where sub-10 nm cross-sectional channels have been fabricated in silicon-on-insulator (SOI) structures [7, 8] to obtain high charge sensitivity and reliability. Memory cells with carbon nanotube (CNT) as the sensing channel and Si$_3$N$_4$ traps as charge-storage have also been reported, utilizing the self-assembled CNT channel of nano-scale width [9]. Recently memory cells based on metal NC and CNT have been demonstrated with further improved characteristics [10] despite a 100 nm bottom-gate control dielectric. The channel transport in CNT-NC memories has been modeled by the non-equilibrium Green’s functions (NEGF) in the presence of charges stored in the NC [11]. In this paper, we theoretically explore the CNT-NC memory from the three-dimensional (3D) electrostatics perspective to demonstrate the electric field asymmetry with respect to the program versus retention operations, in an attempt to illustrate the low voltage
Figure 6.1 Basic CNT-NC memory schematic showing (a) bottom-gate and (b) top-gate structures
Figure 6.2 a) 3D structure for CNT-NC memory with the back-gated CNTFET with 3 NCs aligned to the CNT. b) The NC EEPROM structure in 3D showing a 3×3 array of NCs. The cross sections of the different memory structures c) a bottom-gate CNT-NC memory d) a top-gated CNT-NC memory e) a NC EEPROM are shown. The typical structural parameters used in the electrostatics simulations are as follows: NCs diameter: 6 nm; the CNT diameter 2 nm; pitch: 12 nm; tunneling dielectric thickness $d_t$: 3 nm (Si EEPROM) and 4 nm (NC/CNT memory); distance between the center of the NC to the gate, $d_c$: 30 nm.
operation experimentally demonstrated in ref. [10]. This analysis further illustrates the significant advantages of this structure over planar Si channels.

### 6.3 Electrostatic Model

The basic bottom-gate (BG) and top-gate (TG) memory structures are shown in Fig. 6.1. The bottom-gate device [10] consists of a CNT on top of thermal SiO$_2$ grown on the degenerately doped Si substrate that acts as a gate electrode. The CNT is then covered by the tunneling dielectric of 4 nm thickness, and followed by metal NCs self-assembly. A 30 nm thick SiO$_2$ passivates the entire device. The top-gate memory structure has a similar process except that passivation dielectric is replaced by the control dielectric and the top gate is patterned subsequently. The key electrostatic features can be illustrated by the program/erase and retention memory operations, where the quantum transport between the CNT channel and the NC floating gates is controlled by the gate bias. The tunneling model requires 3D self-consistent solution of the Schrödinger (transport) and Poisson (electrostatics) equations. Tunneling formalisms such as the transfer Hamiltonian method have been used under simplified quasi-one-dimensional assumptions for calculating the quantum transport [12]. The NEGF approach [13] is another possible formalism for the same. In the planar silicon channel and floating gate case, the one-dimensional (1D) WKB method in the least-action path has been reasonably accurate with the underlying notion that the least-action path is the dominant tunneling path, and the electric field and potential drop in this path should dictate the tunneling kinetics [14, 15]. For clear comparison with the planar devices, we restrict our analysis to the 3D electrostatics of the CNT-NC memory in the least-action path to demonstrate the enhanced electric-field asymmetry inherent in this nanoscale structure. For simplicity, we assume the gradual channel approximation [16] similar to that in MOSFETs to ignore the contribution of the
source-drain potential. We simulate CNT-NC memory structure and the Si channel based NC memory structure as shown in Fig. 6.2 (a) and (b) respectively. The various parameter values used in the electrostatics simulations for the different structures are excerpted from typical values in the available experiments for the purpose of verification. The NC diameter of 6 nm, the CNT diameter of 2 nm, and the NC pitch of 12 nm are chosen. The tunneling dielectric thickness $d_t$ is 3 nm for Si EEPROM and 4 nm for the NC/CNT memory, which derives from the non-ideal experimental reality of the difference in the dielectric quality on Si and on CNT. The control dielectric thickness $d_c$ (defined for the ease of calculation as the distance between the center of the NC and the gate) is typically 30 nm for both the Si EEPROM and NC/CNT memory simulations. These values are used unless otherwise stated. A range of $d_c$ values of 27 nm, 50 nm, and 100 nm, are also explored to study the effect of control dielectric thickness to the electrostatics, which is very strong for the two-dimensional (2D) channel case, but is surprisingly weak for the 1D channel case. We treat the CNT electrostatics at its two asymptotic cases [17]: (a) as a conductor in the ‘ON’ state during programming where its Fermi potential is set by the source/drain contact, and (b) as a dielectric in the ‘OFF’ state during retention.

### 6.4 Program/Erase (P/E) Operations

During the ‘program’ operation, the CNT under accumulation can be viewed as a conductor. The electrostatics in the memory structure can be conveniently separated into two different contributions – a) the self-capacitance of the NC in the specific geometry and b) the capacitive coupling of the NC to the gate electrode and the channel. This separation of different sources of potentials, i.e., the stored charge and the applied biases, is a corollary of the superposition principle in electrostatics [18]. The two contributions counterbalance each other in the sense that the gate
actuates the potential of the NC through the capacitive coupling term to overcome the charging energy due to the the NC self-capacitance. We analyze the NC/CNT memory in two 3D configurations of the bottom gate (Fig. 6.2 (a) and (c)) and the top gate (Fig. 6.2 (d)) structures. The results are compared with electrostatics in the planar Si channel case (Fig. 6.2 (b) and (e)), where the electrostatics is computed by assuming that the Si substrate is a conductor. For program conditions, this is generally valid. The potential profiles along the vertical dashed lines in figs. 2 (c) and (d) coincide with the least-action path for tunneling, and have also been plotted in figs. 3 and 4 for write and retention conditions.

A. SELF-CAPACITANCE:

The self-capacitance of the NC can be defined as the capacitance due to addition of a charge into the NC in the given structure. The common approximation is based on the NC in an infinite medium, which relates to the potential energy to bring a charge from infinity to the NC. This energy, however, is affected by the specific surroundings. In the absence of an analytical model, it can be calculated numerically using a 3D Poisson solver [19] as the potential the NC settles to due to the addition of charge when the channel and the gate are both grounded. The situation is very similar to the retention condition except here the channel is assumed to be conducting. A comparison of this energy should reveal difference in charging energy penalty in various structures. In the general case of comparable NC coupling to channel and gate, the total number of steady-state charges in the NC is reached when the emission current from NC to gate equals to the injection current from channel to NC [14]. However, the much larger thickness and lower electric field of the control dielectric, especially in the NC/CNT memory, cannot sustain any traceable amount of either direct or Fowler-Nordheim tunneling compared to the tunneling dielectric in the P/E
Figure 6.3 Potential profile for capacitive coupling calculation (solid line), self capacitance calculation (dashed line) and capacitive coupling calculation without NC (dash-dotted line) for a CNT NC memory with various parameters $d_c: 30$ nm; $d_t: 5$ nm; NC diameter: 6 nm; CNT diameter: 2 nm for (a) top-gate and (b) bottom gate structures.
Figure 6.4 Retention condition potential profile along the extended path of least action. The solid (single NC case) and dashed (3 aligned NC case) lines show potential profiles when considering the CNT to be a dielectric. The dotted (single NC case) and dash-dotted (3 aligned NC case) lines are for assuming that the CNT as a conductor which is similar to the self capacitance calculations.
voltages considered here. In this case, the charging energy essentially controls the number of electrons that can be accommodated in a potential well, which can be filled until the NC Fermi level exceeds that of the source of injection i.e. the CNT channel under bias condition. Hence, we can use the charging energy calculated in 3D electrostatics to derive our device comparisons without invoking the tunneling current calculations.

As a measure of the self-capacitance, the NC potential with 5 stored electrons is calculated with the gate and the CNT grounded. The simplest case of charging a sphere embedded in infinite SiO$_2$ with 5 electrons causes a $-0.61$ V shift in potential energy when ground is at infinity. In the NC planar memory case, the potential is constrained by the grounded parallel plates rather than by assuming zero potential at infinity. In this case, as shown in Fig. 6.2 (b), the potential on the single NC with 5 electrons is $-0.46$ V. This can be understood in terms of the fact that the two grounded plates can be replaced by an infinite set of image charges [3]. The proximity of image charges of the nearest grounded plate (positive charges in this case) produces the main contribution to the potential increase. To include NC interactions, we can calculate the electrostatics for an NC array. For simplicity, a 3×3 NC array is used to include the nearest-neighbor effects for the central NC. The potential for the central NC in this case is then $-0.67$ V. The nearest-neighbor NCs between grounded plates also can be resolved as an infinite set of image charges. The decrease in potential is because the relative proximity of similarly charged NCs than their corresponding image charges. The detailed electrostatic analysis of this system is presented elsewhere [20]. In comparison, the top-gate and bottom-gate NC/CNT structures are simulated. To include the effect of the 1D channel and the NC polarization, the potential profiles for the single NC case is shown in Fig. 6.3, where the NC potential is $-0.51$ V for the top-gate case and $-0.52$ V for the bottom-gate. This value for the 1D channel and 2D gate
structure (1D/2D constraint) is somewhere between that of a parallel plate capacitor (2D/2D constraint) and of an infinite dielectric (no constraints), which can be rationalized as the difference in the ability to constrain the potential by the different dimensions. To include the nearest-neighbor NC effects, 3 NCs aligned to the CNT channel as shown in Fig. 6.2 (a) are calculated. The potential on the central NC is $-0.68 \text{ V}$ for the bottom-gate case and $-0.65 \text{ V}$ for the top-gate case. These values are more negative than the single NC case also due to inter-NC electrostatic interactions discussed for the 2D channel case. From the data in Table 6.1, although we can observe some variation in the self-capacitance, in comparison with the capacitive coupling, as we shall see in the next section, the potential due to self-capacitance is not as sensitive to various geometries considered, i.e., 2D/1D channels, top-gate/bottom-gate configurations, and variation in control oxide thickness ($d_c$ values of 27 nm, 30 nm, 50 nm, and 100 nm). Irrespective of the structures, all potentials based on self-capacitance fall in the small range of $0.58 \text{ V} \pm 0.1 \text{ V}$, which seems logical for the metal NC spheres because the potential dies down quickly as inverse of distance.

**B. Capacitive Coupling:**

In the ideal case of for NC between two infinitely large parallel plates, the capacitive coupling for NC can be reasonably estimated by the 1D model where the capacitance is inversely proportional to the distance from the respective electrodes. However, if we replace the point with a metal sphere, then the polarization of the sphere makes the capacitive coupling depart significantly from the 1D electrostatics and 3D electrostatics is essential for planar NC memory analysis [3, 20]. The replacement of the planar 2D channel with a 1D channel adds further requirements for full 3D electrostatic treatment. The capacitive coupling of the NC to the 1D channel, which is very different in nature from the coupling to the 2D control gate, is
Table 6.1: Capacitive coupling and self capacitance calculations for various CNT-NC memory and NC planar Si memory structure for the program condition.

<table>
<thead>
<tr>
<th>Structural Parameters</th>
<th>Potential on NC (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d_c$ (nm)</td>
</tr>
<tr>
<td>1NC-CNT BG</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>1NC-CNT TG</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>3NC-CNT BG</td>
<td>30</td>
</tr>
<tr>
<td>3NC-CNT TG</td>
<td>30</td>
</tr>
<tr>
<td>1 NC-Si</td>
<td>30</td>
</tr>
<tr>
<td>3×3 NC -Si</td>
<td>30</td>
</tr>
<tr>
<td>Parameters from ref. 10: dt=4 nm NC dia = 8 nm</td>
<td></td>
</tr>
<tr>
<td>3NC-CNT BG</td>
<td>100</td>
</tr>
<tr>
<td>3×3 NC -Si</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 6.2: Potential drop in the tunneling dielectric in the path of least action for \( d_c = 30 \text{ nm} \) for CNT-NC memory and Si planar memory for retention condition with 5 electrons in the NCs.

<table>
<thead>
<tr>
<th>Structural Parameters</th>
<th>Channel Condition</th>
<th>Potential drop (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1NC-CNT TG</td>
<td>Conducting</td>
<td>−0.509</td>
</tr>
<tr>
<td>3NC-CNT TG</td>
<td>Conducting</td>
<td>−0.645</td>
</tr>
<tr>
<td>1NC-CNT TG</td>
<td>Dielectric</td>
<td>−0.35</td>
</tr>
<tr>
<td>3NC-CNT TG</td>
<td>Dielectric</td>
<td>−0.39</td>
</tr>
<tr>
<td>1NC-Si</td>
<td>Conducting</td>
<td>−0.46</td>
</tr>
<tr>
<td>3×3 NC -Si</td>
<td>Conducting</td>
<td>−0.67</td>
</tr>
<tr>
<td>3×3 NC –Si</td>
<td>Doping ( 10^{18} \text{ cm}^{-3} )</td>
<td>−0.58</td>
</tr>
</tbody>
</table>
imperative to understand the low voltage operations of the NC/CNT memory. As in the case of the self-capacitance, in the absence of an analytical expression, the capacitive coupling can be calculated numerically [19] to obtain the potential on the NC due to the potential difference between the control gate and the channel. We simulate the cases of 5 V control-gate bias, grounded channel and no stored charge in the NC, which corresponds to the program operation at its onset.

In the case of a simple parallel plate capacitor but without NC as shown in Fig. 6.2 (b), the potential at the interface of tunneling oxide and control oxide is 0.42 V. The inclusion of a single metal NC, i.e. the inclusion of the polarization of the metal NC, results in a potential drop of 0.81 V in the path of least action due to repulsion of equipotential contours out of the metal NC. To include nearest-neighbor NC interaction, we consider the similar 3×3 NC array as shown in Fig. 6.2 (b). The potential on the central nanocrystal is reduced to 0.74 V. The electric field due to the polarization of the nearest neighbor NCs counteracts the electric field due to the capacitor in the central NC to cause a decrease in NC polarization and hence lowers the NC potential. These effects have been analytically discussed in details elsewhere [20]. By further adding the 3D effects from the CNT channel, the potential for a single NC 4 nm away from a grounded CNT is a remarkable 2.35 V for the 5 V gate potential for the structure shown in Fig. 6.1a. The NC is strongly coupled to the control gate in spite of its proximity to the CNT unlike the case of the planar channels. This effect is a direct result from the 1D nature of the CNT compared to the 2D control gate. This can be rationalized as follows. In the absence of the metal sphere, near the 1D channel the electric field varies as inverse of distance while near the 2D gate the electric field is about constant. Hence going from the 1D channel to the 2D gate the potential changes monotonically and faster near the 1D channel as shown in Fig. 6.3a. Hence the potential on any point will be more strongly coupled to the gate than the channel in
contrast to the case of the 2D channel, where the potential is a simple function of the distance. The polarization effect of the metal NC further causes the equipotential contours to be repelled out of the metal, compressed into a larger potential drop into the tunneling dielectric. The NC potential for the program condition due to the capacitive coupling for the various geometries considered is shown in Table 6.1. The capacitive coupling is similar for top-gate and bottom-gate structures. This is reasonable since near the 1D channel any radial direction is more or less electrostatically equivalent and hence as long as the NC is close to the channel, the top-gate or bottom-gate configurations should perform similarly. For structures with \( d_c \) of 30 nm as shown in Table 6.1, the capacitive coupling of the 1D channel is 3 times more effective than the 2D channel case. To compare the performance of structures equivalent to the experimental structures [10], we simulate both 1D and 2D channel cases with parameters of 8 nm NC diameter, 4 nm tunneling dielectric and 100nm \( d_c \). In the case of the 1D channel the NC is at 1.86V while in the case of the 2D channel the NC is at 0.39V. In this case, 1D channel is 4 times more effective than the 2D channel case. From the control dielectric scaling perspective, the \( d_c \) of 100 nm in the 1D channel case is equivalent to a scaled \( d_c \) of 14 nm for the 2D channel case. We clearly see that unlike the 2D planar memory devices, the NC-gate capacitive coupling is a weak function of the control dielectric thickness as seen from comparing different \( d_c \) for the various structures as shown in Table 6.1. This weak dependence on control dielectric thickness is consonant with the idea advanced earlier that the potential imposed by the 1D channel weakens radially very quickly which lets the gate couple much stronger to the NC. The polarization of the NC additionally enhances this effect. Hence, the control dielectric thickness weakly affects the capacitive coupling compared to the 2D channel case. These results demonstrate that capacitive coupling in case of the 1D channel is much more effective than the case of the 2D channel.
To validate against experimental results [10], we use the results from the bottom gate CNT-NC structure simulations with parameters $d_c$ of 100 nm, $d_t$ of 4 nm and NC diameter of 8 nm same as the structure in ref. 10. At 5V gate bias, the NC potential is at 1.86 V as noted previously. To account for the band offsets we do the following. The valence band energy of the CNT is approximately aligned with the Cr (the source/drain metal) Fermi level at 4.5 eV since the CNT turns on at $V_G \sim 0$V when no charge has been injected [10]. The Au (NC) work function is 5.1V. The relative well depth for hole injection is $-0.6$ eV at the beginning of programming. The total hole well depth is therefore $1.86V - 0.6V = 1.26V$ at 5V gate bias. The potential due to self-capacitance is calculated as $-104$ mV per electron. Hence at 5V gate bias a maximum of $\sim 12$ electrons can be present in the NC. Experimentally, threshold voltage ($V_T$) shift of about 1.24V has been observed for a charging voltage for 5.0 V, where $V_T$ shift of 130 mV/electron has been measured [10]. This indicates the detection of 9.5 stored electrons per NC, which approaches but does not exceed the maximum number of holes that can be injected from our electrostatic model. The small decrease in the expected hole injection is mainly due to read disturbance in the experimental setup. The comparison to the experiments corroborates the validity of our model, including the weakly coupled assumption.

The erase operation is very similar to the program operation. Unlike the silicon channel case, the CNT channel is usually ambipolar. This behavior can be understood from the conduction mechanism in CNT based FET where the band bending near the contacts, similar to the case of gate induced drain leakage (GIDL) [21], causes the channel to be on under erase conditions but populated with carriers of the opposite charge, which could provide fast erase. In any case, the erase can be a block erase [22] and as such has lower speed requirements from the memory architecture perspective.
6.5 Retention Operation

Under the retention condition, the gate can be considered grounded. The nanotube in the ‘OFF’ state has no free charge and can be modeled as a dielectric. The dielectric constant of a semiconducting CNT has been explored by some researchers theoretically [19, 23]. We approximate the system with the CNT in the ‘OFF’ state essentially as a system where the nanotube is absent or equivalently, has the dielectric constant close to the medium (3.9 for SiO$_2$ compared for calculated CNT dielectric constant of 1.0–2.5 [23, 24]). This effective alteration of the structure produces a potential drop of $-0.35$ V (single NC case) and $-0.39$ V (with nearest neighbor NC interaction, i.e. the 3 NC case) between the CNT and the NC as shown in Fig. 6.4 and Table 6.2. This is a 40% reduction in the potential difference in the least-action path compared to the condition when the CNT is considered a conductor. The potential drop in the least-action path compares favorably to the planar Si memory where the potential drop is $-0.46$V and $-0.67$V for 1 NC and 3×3 NC cases respectively as shown in Table 6.2. In this case, however, the assumption of the Si channel as a conductor cannot be valid any more. However, this calculation serves as an lower bound for retention time estimation since part of the potential drop will occur in the depleted Si channel, and hence decrease the electric field in the tunneling oxide. To obtain a better estimate, a doping of $10^{18}$ cm$^{-3}$ was used to self-consistently calculate the depletion charge in Si for the retention condition. Despite electric field penetration into Si, a maximum potential drop of $-0.58$ V along the least-action path between the NC and the Si/SiO$_2$ interface was calculated for the 3×3 NC cases. This is still much larger than the NC/CNT potential difference under the same retention condition.
6.5 Discussion

The electrostatics elucidates certain salient features of the NC/CNT memory devices. First, the NC/CNT structure provides an important design advantage for P/E operations. In the planar-channel devices, there is an inherent trade-off in the control dielectric scaling, where the decreasing control dielectric thickness enhances gate coupling to the NC and the gate leakage at the same time. In the NC/CNT structure, the control dielectric scaling has much more design space in light of the strong electrostatic gate-NC coupling, where the NC/CNT memory produces 3 times larger potential drops during programming than the corresponding planar case for the technologically reasonable $d_c$ of 30 nm. Second, for retention condition, the CNT becomes electrically floated, and the field crowding effect to the CNT small perimeter is much reduced. Hence the surface electric field is significantly smaller. The potential drop between the CNT and NC is lower than that between the NC and the planar channel under similar retention conditions. These two observations indicate the very desirable enhancement of electric field asymmetry in the NC/CNT memory. Third, a large gate-NC coupling provides the possibility of larger number of electrons for storage than the comparable planar NC memory structure, as long as the quantum confinement effect is not dominant. This indicates the possibility of enhanced reliability of information storage as the number of electrons stored per NC increases when low-voltage P/E operations are desirable.

6.5 Conclusion

The CNT-NC memory electrostatics has been analyzed from the memory operation perspective. The structural asymmetry of the CNT-NC memory translates to highly enhanced electric field asymmetry for better retention/program time ratio.
Larger number of electrons per nanocrystal observed experimentally at low programming gate bias for a 100 nm thick gate stack, have been explained through the above analysis. Inherent advantages of the structure have been discussed in view of electrostatics, which differentiates it from planar device structures.
REFERENCES


CHAPTER 7

PERFORMANCE EVALUATION OF NANOCRYSTAL BASED NANOTUBE MEMORY BY TUNNELING DIELECTRIC ENGINEERING

7.1 Abstract

The nanocrystal (NC) carbon nanotube (CNT) memory provides an interesting device structure for single-electron based memory operations. In this study, the speed for program/erase (P/E) operations is evaluated by tunneling rate measurements between the CNT and the NC which is disturbed by the presence of traps in the tunneling dielectric. Measurements of charge transport between a 1-dimensional (1-D) CNT channel and a 3-dimensional (3-D) NC or a 0-dimensional (0-D) trap is of fundamental physical interest. The tunneling rate as a function of pulse magnitude and duration indicates two different regimes: direct tunneling and Fowler Nordheim (FN) tunneling that can be utilized for different memory operational modes. The effect of NC and traps, distinguished clearly from the measurements, provides the evaluation of the performance of trap vs. NC for charge storage. Fast ~µs program was observed for -9 V bias pulses. Retention enhancement over the previous experiments is also explored by engineering the tunneling dielectric. While single high-k dielectric films are shown to be unreliable, thin composite dielectric film using atomic layer deposited Al₂O₃ using a buffer layer of evaporated SiO₂ was used to increase the room temperature retention by 2 orders of magnitude to about 3 days.
7.2 Introduction

The one-transistor (1T) flash memory cell has demonstrated significant density enhancement for non-volatile memory applications [1]. As flash memory scaling faces physical and fundamental limitations, various solutions are being proposed to advance memory technology. Engineering the floating gate has been a major focus. There are attempts to replace the thin-film floating gate by charge storage in silicon nitride traps [2]. Self-assembled semiconducting nanocrystal floating gates have been alternately proposed to reduce the operating voltage of the memory cell by enhanced tunneling dielectric scaling [3, 4]. Further, metal nanocrystals, which provide better electrostatic, high density of states for faster tunneling and choice of different work-functions for better retention, have been implemented [5-7]. A parallel direction of innovation has been engineering the channel. Nanowire based channels in conjunction with semiconductor nanocrystal based storage has been explored [8, 9]. Significant enhancements in density, retention and memory window enhancement have been indicated as the channel approached a one dimensional (1-D) nanowire [8]. The fabrication of 1-D channel by top-down approach involves expensive sub-lithographic fabrication but can be easily realized using bottom-up self assembly. Single wall carbon nanotubes (SWNT) based channels have been used to study the prototype memory device. Trap-based storage has been demonstrated using a SWNT channel [10]. Recently we have demonstrated a SWNT memory with charge storage in metal nanocrystals self-aligned to the SWNT [11]. The unique electrostatics of this device due to the interplay of the nanoscale 1-D channel and the larger metal nanocrystal that leads to sub-5V operations despite a thick 100 nm control dielectric and 5 nm tunneling dielectric [12]. Room temperature single electron sensing has been explored by engineering the electrostatics and SWNT transport [13]. The SWNT device retention was about 900 s due to the poor quality of the evaporated SiO$_2$ used as the
tunneling dielectric [11]. In this report, we present the efforts to improve the tunneling dielectric. P/E operations have been explored as a function of the various tunneling dielectric in different tunneling regimes. The relative effects of traps versus NCs are analyzed. Implications for device design are then summarized.

7.3 Device design and fabrication

The SWNT device described in [11], consists of a nanotube channel with Au/Cr source and drain, placed on a 100 nm dry thermal control oxide grown on a degenerately doped Si bottom gate. A 4-7 nm tunneling dielectric, deposited on the nanotube channel, separates self-assembled Au NCs formed by depositing a 1.2 nm thin layer of Au on top of the tunneling dielectric [11]. A PECVD SiO$_2$ blanket film is used to passivate the device. In the present study, the tunneling dielectric is engineered for better room temperature retention. The various options explored are:

i) 5 nm electron-beam evaporated SiO$_2$ for consistent comparison with previous batch,

ii) single dielectric film of Al$_2$O$_3$ of thickness 4-5 nm using the atomic layer deposition (ALD) technique,

iii) composite dielectric film of Al$_2$O$_3$ of 4 nm deposited on the CNT covered with buffer layer of 3 nm electron-beam evaporated SiO$_2$ for conformal ALD based dielectric deposition.

The single dielectric film deposited on CNT using ALD technique is investigated by AFM analysis and electrical characterization. The large roughness approximately equal to the thickness of the deposited dielectric can be observed from the AFM image of the nanotube covered with ALD Al$_2$O$_3$ shown in Fig. 7.1. The NC-formation step involved the deposition of a 1.2 nm of Au followed by an optional 400°C anneal in N$_2$ atmosphere for 10 s duration. Electrical characterization of CNT FET before and after
the NC-formation step indicates the adverse effects on device performance as shown in Fig. 7.2 (a). Even the CNT turn-on current remains roughly the same, it becomes very difficult to turn off the device. The degradation has been attributed to the Au contamination of the CNT due to the inadequate ALD film quality on the CNT. This hypothesis has been experimentally validated by electrically characterizing CNTFET devices with the deposition of Au directly on the CNT without the protective tunneling dielectric layer as shown in Fig. 7.2 (b) to demonstrate equivalent effect to NC formation in the presence of the high-k tunneling dielectric. To resolve the suspicion that the Au forms nanowires using the CNT as a template, SEM image of the device in Fig. 7.2 (c) shows that the Au nanocrystals form on either side of the CNT avoiding the CNT itself. Such a result is consistent with the weak interaction between nanotubes and Au indicated by the inability to form metal nanowires by evaporation of Au onto suspended CNT as a template [14]. The identical on-current before and after NC formation strongly indicates that the CNT is still responsible for the transport rather than percolation paths between the 2 µm channel through the NC at low bias. The on-current is not strongly bias dependent as should be the case for percolation (Poole-Frenkel) based current. It can be conjectured that the Au acts as a contaminant to cause Fermi level pinning and consequent device degradation in the gate control. Conformal dielectric ALD on nanotubes has been a challenge until recently and the above results to the effort of dielectric deposition on the CNT are consistent with efforts to deposit dielectric directly without any CNT functionalization or chemical treatment [15].

The inclusion of a buffer SiO₂ layer has been frequently used in high-k gate dielectric research. A 3 nm SiO₂ layer was used to cover the CNT before a 4 nm high-k Al₂O₃ deposition. ALD Al₂O₃ deposition was done by trimethyl aluminum (TMA) and H₂O at 300°C [16]. NC-formation step on the composite dielectric encapsulated
Figure 7.1 (a) AFM image of 4 nm of ALD Al$_2$O$_3$ deposited on a nanotube. (b) Scan height along the CNT shows surface roughness approximately equal to film thickness due to inhomogeneous nucleation and growth of ALD Al$_2$O$_3$ on the nanotube surface.
Figure 7.2 (a) Electrical characterization of a CNTFET covered with ALD Al₂O₃, before (blue curve) and after (red curve) the deposition of 1.2 nm Au for NC-formation, shows that the nanotube cannot be turned off after Au deposition. The on-current level remains roughly same. (b) Electrical characterization of a CNTFET with
a bare CNT without ALD Al₂O₃ cover, before (blue curve) and after (red curve) the deposition of 1.2 nm Au for NC-formation, shows that the nanotube cannot be turned off after Au deposition just as when it was covered with ALD Al₂O₃. (c) SEM of the CNT after 1.2 nm Au deposition for NC-formation. The presence of the nanotube is indicated by an approximate linear arrangement of the NCs in the random NC formation between by the arrows. Inset shows close-up of NCs arranged on either side of the NC avoiding the NC itself.
CNTFET produced minimal adversity on the device performance. These devices were subjected to memory operation characterization and the results have been presented here.

### 7.4 Programming Rate

The programming rate is a function of the programming pulse which has two parameters—pulse magnitude ($V_{PULSE}$) and duration ($t_{PULSE}$). The variation of the threshold voltage as a function of pulse duration for different pulse voltages has been shown in Fig. 7.3 for the devices with evaporated SiO$_2$ tunneling dielectric at 10K. The measurement data indicate the tunneling processes in the memory device. Two distinct slopes can be observed in the tunneling rate curves, which is especially clear at high pulse biases $V_{PULSE} \geq 5$V, separated by a knee at pulse width $t_{PULSE} \sim$ ms. The distinctive slopes indicate different tunneling mechanisms. We advance two observations from the measurements to clarify the tunneling mechanisms involved. First, in the low-bias condition, $V_{PULSE} \leq 5$V, the NC sample shows slow yet definite charge injection in the $t_{PULSE} >$ ms regime indicated by the threshold voltage shift of 1 V. In contrast, the threshold voltage shift is negligible ~0.2 V in the samples without NC, as is consistent with our pervious experiments [11]. The hypothesis that at low bias, the mechanism of charge injection is direct tunneling is further fortified by the argument that the metal NCs primarily provides a much larger density of states for higher tunneling rate compared to traps and hence produces the larger tunneling current at low bias. Second, the knee at $t_{PULSE} \sim$ ms separates portion of the curve with a slope that varies strongly with $V_{PULSE}$ for $t_{PULSE} \leq$ ms and that with a slope that is approximately independent of $V_{PULSE}$. The slope is rather constant for the sample device type (control or NC) even until low $V_{PULSE}$. Furthermore, the slope for $t_{PULSE} \geq$ ms is clearly 2× larger for NC sample compared to the control sample. We
hypothesize that for $t_{\text{PULSE}} \leq \text{ms}$, Fowler Nordheim (FN) tunneling dominates while for $t_{\text{PULSE}} \geq \text{ms}$, direct tunneling dominates [17, 18]. FN tunneling occurs at high bias and is strongly bias dependent, where the charge is injected into the SiO$_2$ and hence only the SiO$_2$ density of states is of relevance. The charge injections by FN tunneling are not completely equivalent for either control or NC samples because the NC sample produces a 35% higher maximum local electric field [12], to which the FN tunneling mechanism is very sensitive. This manifests in the observation that for the same threshold shift or equivalently the same total charge injection, the control sample requires 35% larger $V_{\text{PULSE}}$ than the NC sample. For example, for $V_T$ shift of 2 V, the NC sample requires $V_{\text{PULSE}} = 11$ V compared to the $V_{\text{PULSE}} = 15$ V that the control sample requires. Direct tunneling, on the other hand, has less bias dependence and will dominate at low biases. This can be observed from Fig. 7.3 that the slope for $t_{\text{PULSE}} \geq \text{ms}$ is clearly 2× larger for NC sample compared to the control sample. The enhanced memory window in the NC sample can thus be conjectured from the difference in the density of states between the metal NC and the trap.

A different point of view of the programming is presented in Fig. 7.4 at two pulse widths, $t_{\text{PULSE}} = 100$ µs and $t_{\text{PULSE}} = 1$ s for increasing $V_{\text{PULSE}}$, to demonstrate the difference between room temperature and 10K measurements. The room temperature measurements show two regimes as shown in Fig. 7.4 (a) and (b). FN tunneling along with Pool-Frenkel injection [20] result in the large slope with the pulse bias. At higher bias the traps are saturated with charge, relaxing the electric field in the tunneling dielectric and enhancing the electric field in the control dielectric to start charge loss. A decrease in the slope of the charging is observed as a result [21]. The 10K measurements show 3 regimes as shown in Fig. 7.4 (c) and (d). The low bias regime where the traps are inactive for the control sample and bias independent direct tunneling occurs in the NC samples. Then FN tunneling occurs in the $V_{\text{PULSE}}$ range of -
5 V to -8V approximately for the NC sample and -6 V to – 9 V approximately of the control sample. At higher biases, a charge loss to the gate is large and consequently, saturation is established as that in the room temperature measurements [21]. For the 100 µs measurement, the process is extremely non-equilibrium. Though there is a NC charging current for the pulse voltage, it is not large enough to effectively cause $V_T$ shift at the 100 µs pulse width. Only beyond pulse bias of approximately -7V is there $V_T$ shift for both control and NC samples at both temperatures.

Programming rate measurements for samples with composite tunneling dielectric in Fig. 7.5 show that even at 10K, the measurement cannot clearly resolve NC and control samples. This is attributed to the fact that though the electrostatic effective oxide thickness (EOT) of in the composite tunneling dielectric samples is comparable to the 4 nm evaporated SiO₂ samples, the physical thickness difference of 3 nm of the tunneling dielectric makes the NC direct tunneling current very weak. Hence, low bias NC direct tunneling signature is not observed. Electrostatic enhancement is also not observed clearly in the FN tunneling regime due to the physical thickness of the tunneling dielectric. The magnitude of threshold voltage shift is about 2 V for both control and NC samples, which is the same as the control of the evaporated SiO₂ tunneling dielectric sample. Hence, we conclude that charge injection into the NC in the composite dielectric sample was not observed.

7.5 Retention

Retention measurements were made at room temperature for composite tunneling dielectric samples. A pulse of $t_{\text{PULSE}} = 1s$ and $V_{\text{PULSE}} = 10V$ was used to charge the devices. $I_D V_G$ measurements of uncharged devices, devices immediately after charging and after some time after charging, at room temperature and at the elevated temperature of 85°C are presented in Fig. 7.6 (a) and (c). $I_D$-$t$ measurements
Figure 7.3 Program rate measurements of NC sample and control sample with evaporated SiO$_2$ tunneling dielectric at 10K. There are two slopes evident. A strongly pulse voltage dependent slope for fast program times is separated from a pulse voltage independent slope for slow program times by a knee at approximately ~ms pulse widths. The NC sample produces larger threshold voltage shift for the same bias. For example, for $V_T$ shift of 2V, the NC sample requires pulse voltage of 11 V (indicated by blue arrow) compared to the 15 V (indicated by red arrow) for the control sample. The pulse independent slope is about 2× for the NC sample compared to the control sample. At pulse voltage $\leq$ 5 V, there is no $V_T$ shift for the control sample compared to a 1V $V_T$ shift for the NC-sample.
Figure 7.4 Program threshold voltage shift measurement of control samples in (a) and (c) and NC sample in (b) and (d) with evaporated SiO$_2$ tunneling dielectric. (a) and (b) show measurements at room temperature and (c) and (d) show measurements at 10K for two different pulse widths: 100 $\mu$s (blue-line) and 1 s (red-line) as a function for pulse voltage. Both samples at room temperature show two regimes denoted by red roman numerals demarcated by solid red vertical lines for the pulse width of 1s: (i) FN
tunneling or Poole-Frenkel tunneling at lower biases and (ii) saturation at high biases. However at pulse widths of 100 µs, two regimes denoted by blue roman numerals demarcated by solid blue vertical lines, can be seen: (i) No charge injection at low bias and (ii) FN tunneling or Poole-Frenkel tunneling at a higher bias. For the low temperature curve, Poole-Frenkel conduction is eliminated and Fowler Norheim tunneling dominates. Both samples at 10 K show three regimes for the pulse width of 1 s demarcated by solid red vertical lines denoted by red roman numerals: (iii) Direct tunneling at lower biases, (iv) FN tunneling at medium biases, and (v) saturation at high biases. At pulse widths of 100 µs, the curves are similar to that at room temperature for the same timescale. They show two regimes indicated by blue roman numerals demarcated by solid blue vertical lines: (i) No charge injection at low bias and (ii) FN tunneling at a higher bias.
Figure 7.5 Program rate measurement of NC sample and control sample with composite tunneling dielectric of 4 nm ALD Al₂O₃ and 3 nm evaporated SiO₂ tunneling dielectric at 10K. The effect of NC is not evident clearly as trap based charging dominates.
Figure 7.6. Retention measurement for NC sample with composite tunneling dielectric of 4 nm ALD Al₂O₃ and 3 nm evaporated SiO₂ tunneling dielectric at room temperature and at 85°C. (a) Room temperature IᵥG measurements before (blue), after charging (green) and after 9000 s after charging (black) and at some intermediate times. (b) The threshold voltage shift with time is plotted and extrapolated to find total retention time. The same process is repeated for figure (c) and (d) at 85°C. The retention values are approximately ≥10⁵ s (3 days) at room temperature while at 85°C retention is ≥ 6×10³ s.
with the source/drain bias at 0.2V and the gate at 0V and extrapolation to determine the total retention time has been shown in Fig. 7.6 (b) and (d) for both temperatures. Typical room temperature retention of $\geq 10^5$ s is shown to be an improvement over 900 s retention using evaporated SiO$_2$ [11]. At 85°C, the retention is reduced to $\geq 6 \times 10^3$ s. This suggests an activation barrier or 0.74 eV for the discharge process and possibly related to the difference between the conduction band of the dielectric and traps. While the retention is enhanced, it is still far away from practical Flash memory applications. Reliable, defect-free dielectric deposition on carbon nanotubes remains a research challenge. Recently, voltage shift with time is plotted and extrapolated to find total retention time. The same process is repeated for figure (c) and (d) at 85°C. The retention values are approximately $\geq 8 \times 10^4$ s at room temperature while at 85°C retention is $\geq 6 \times 10^3$ s.
techniques to deposit high-k dielectric conformally on CNT using ALD has been reported [15]. Such techniques should help evaluate device performance of the NC-CNT memory. Theoretically retention enhancements [12] along with experimentally observed low temperature retention [11] should be pursued at room temperature by engineering reliable tunneling dielectric on nanotubes.

7.6 Conclusions

The program rate measurements are a necessary evaluation of the operational characteristics of the NC-CNT memory for the prototype memory structure. It demonstrates the tunneling characteristic of the memory device and shows the possibility of μs program times for large biases of $V_{PULSE} \geq 9$ V and for ~10ms program times for $V_{PULSE} \leq 5$ V. Furthermore this experiment demonstrates a method of measuring tunneling between a 1-D channel and 3-D metal NC or 0-D trap. These results should be interesting benchmarks for the 3-D tunneling theory and simulations that have hitherto not been developed. The retention measurements show initial results of 3 days at room temperature for essentially trap based storage. The injection into NC is not confirmed at room temperature since tunneling into traps at close proximity to the CNT dominates for the composite tunneling dielectric stack. Ultra-thin tunneling dielectric engineering to scale the dielectric should enhance direct tunneling rate for low bias operations and enhance retention.
REFERENCE:


[16] ALD of Al₂O₃ was performed at Cambridge Nanotech Inc., Boston MA. (http://www.cambridgetnanotech.com/)


CHAPTER 8

CONCLUSION AND FUTURE WORK

The dissertation focuses on advances in nanoscale memory technology for high density, low power and low voltage applications. The technological issues have been addressed from experimental and theoretical perspectives. The basic foundation of electrostatics of the non-volatile memory devices has been explored analytically for NCM and from numerical simulations for NC-CNT memory. The fabrication and measurements of the NC-CNT memory device, the theoretical analysis of the transport both within the 1D CNT channel and the transport between the CNT and NC are presented. The efforts to develop room temperature devices have been presented. In a sense, the dissertation makes a case for the NC-CNT memory from the device performance perspective. The ability to store information in a redundant way (by the multiplicity of nanocrystals) at low voltages, single electron resolution for both injection and sensing coupled, and with large output current drive at room temperature, is a benchmark for non-volatile memory performance. Being inherently nanoscale the fundamental limit to the large packing density of these memory devices is definitely sub-lithographic. All these advantages are however demonstrated from a single device performance perspective.

Very-large-scale-integration (VLSI) of these devices is the imminent engineering challenge and the future work should be done broadly towards addressing these issues. Significant advances are required for the realization of this technology at a product level. The first issue is the production of nanowires/nanotubes of tightly uniform properties. The second issue is of alignment and registration which is extremely important since the superstructure is fabricated in a top down fashion in
high volume manufacturing environment. The final pitch/density of the memory is currently lithography limited and hence the nanoscale only serves as a performance limit of the VLSI assembly. The third issue which is more pertinent to the memory technology is engineering the tunneling dielectric. Nanotubes are notoriously averse to dielectric deposition using any chemical means as the nanotube surface does not provide a chemical site (like an –OH bond sticking out) for the deposition of dielectrics. Very recently, dielectric deposition on CNT has been demonstrated. However dielectric stack optimization is definitely an important device level issue. The fourth issue is the nanocrystal size and positional control. Self assembled ordering or physical confinement based alignment are possible solutions. The technological realization of this memory is contingent upon the resolution of these glaring issues. However the basic end-goal for the pursuit i.e. high-density, low-power, low-voltage memories, has been clearly established in this exploratory study and the performance projections make a strong case for further investment in this technology.

For fullerenes in the MOS gate-stack, the results have been astonishingly clear from the basic science perspective. The demonstration of molecular redox in the stable solid state environment with single electron precision is a method to harness chemistry that has been a step in the direction of the CMOS-molecular interface. A possible next step is to integrate nanotubes to replace the extended floating gate of a CvMOS for a sensor with nanoscale spatial resolution. Such resolution is very critical for specific, sub-molecular chemical sensing. In an array form, such an assembly can resolve the potential landscape of a macromolecule that is being scanned over it in a fluid or immobilized on a surface.

In so far as metal nanocrystals are concerned, size, shape and density optimization can provide interesting results in enhancing electric field asymmetry. Another approach that goes hand in hand is the theoretical study of metal dielectric
semiconductor tunneling. However such an attempt coupled with experimental verifications in all honesty is only for the bravest of souls.

In conclusion, the work described in this dissertation provides quite a broad view of quantum tunneling based devices in nanoelectronics which focuses mainly on non-volatile memory applications. Novel structure have been introduced and explored. The prospects and issues have been discussed to provide a basis for further cultivation for researchers interested in this area.
REFERENCES


