Abstract

Lazar, Heather Rebecca. Mobility Degradation of Advanced CMOS Devices. (Under the direction of Dr. Veena Misra).

As alternative materials are being pursued for CMOS gate dielectric scaling to below 1 nm, challenges arise that warrant investigation for greater understanding. Mobility values, in particular, of ultra thin SiO$_2$ and high k MOSFET devices have been reported to be low and the reason behind this degradation is still unclear. Furthermore, a study on the effects of implementation of candidate metal gate electrodes on mobility values has not yet been performed. Since many of the properties of these materials are still under scrutiny, accurate determination of their electrical and analytical characteristics is needed.

In this work, an investigation of the mobility of several candidate metal gate electrodes on high k dielectrics was performed and compared with that of SiO$_2$ and polysilicon. This work was separated into two components. The first study compared three metal gate electrodes on HfO$_2$ with that of SiO$_2$ in the 2 nm EOT range fabricated at North Carolina State University. Standard and advanced electrical characterization, including $D_{it}$ and bulk trapping characteristics, was performed including two and three level charge pumping and other pulsed methods. It was found that mobility values were very similar for the metal gate electrodes both on high k and SiO$_2$. The second study analyzed the mobility values for aggressively scaled EOT (< 2nm) devices with a metal and polysilicon gate electrode indicating a
severely degraded mobility. Both the interface and the bulk trapping behavior of these devices were also analyzed and used to correct the mobilities. However, corrections to mobility for both trapping and interface trap density did not recover the mobility of these devices. The mechanism responsible for mobility degradation beyond trapping in these aggressively scaled MOSFETs will be discussed. Finally, steps necessary to enhance mobility will be presented.
Mobility Degradation Of Advanced CMOS Gatestacks

by

Heather Rebecca Lazar

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

ELECTRICAL AND COMPUTER ENGINEERING

Raleigh

2005

APPROVED BY:

_______________________________
Dr. Jayant Baliga

_______________________________
Dr. Gerald Lucovsky

_______________________________
Dr. Griff Bilbro

_______________________________
Dr. Eric Vogel

Chair of Advisory Committee – Dr. Veena Misra
To my parents ...
Biography

Heather Rebecca Lazar was born a Gemini in Newport News, VA on June 6, 1975 at 2:40 AM to Douglas Robert and Virginia Mary Lazar. She has two older brothers, David and Gregory, and a younger brother, Erich and sister, Nicole. Moving along the East Coast, most of her life was spent in Columbia, SC before deciding to move back north and attend The Ohio State University in Columbus, OH where she became an avid Buckeye fan. After graduating summa cum laude with honors with a Bachelor of Science in Electrical Engineering in 1998, she decided to partake of the Wolfpack at North Carolina State University in Raleigh, NC. Finishing her Masters of Science degree in Electrical Engineering in 2000, she continued her studies under the direction of Dr. Veena Misra where she completed her Doctor of Philosophy in Electrical Engineering in 2005.
Acknowledgements

I wish to thank my advisor, Dr. Veena Misra, for her guidance and support through my graduate career. I appreciate her excitement and knowledge of the field as well as her perseverance when things became rough. An advisor and a friend, I would never have made it through without her. I would also like to thank my committee members, Dr. Eric Vogel, Dr. B. Jayant Baliga, Dr. Griff Bilbro, and Dr. Gerald Lucovský for their advice and patience.

In such a large school with such an outstanding facility there are so many who work together in this graduate process. None of my work would have been possible without the help of the lab staff as well as my fellow students. A special thanks to Kaye Bailey for dealing with silly paperwork. To those students both current and past, especially Chad Young, Kingsuk Maitre, Nivedita Biswas, Qian Zhao, Rashmi Jha, Yanxia Li, Jaehoon Li, Bei Chen, Mais Homsi, Greg Heuss, Huicai Zhong, Andy Oberhofer, and Edwin Barry. Your discussions and interests have made the last years memorable.

No words can express my sincere gratitude for my parents’ support and love. They are the lights of my life and have given me the strength to reach for any goal I choose. I only hope they are as proud of me as I am thankful for them. Also monumental in this process were my siblings. Their support when things were bad and encouragement when things were good never failed to sustain me. I am also blessed with an invaluable extended family and close, loving friends. Thank you all for your support and interest in my journey. You made it all the more rewarding.
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td></td>
<td>List of Tables</td>
<td>xiv</td>
</tr>
<tr>
<td></td>
<td>Chapter 1 Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>Scaling Technology</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Desired Properties</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>Current Issues of Advanced Gatestacks</td>
<td>3</td>
</tr>
<tr>
<td>1.4</td>
<td>Advanced Gatestack Mobility</td>
<td>5</td>
</tr>
<tr>
<td>1.5</td>
<td>Objective</td>
<td>6</td>
</tr>
<tr>
<td>1.6</td>
<td>Overview of Dissertation</td>
<td>7</td>
</tr>
<tr>
<td>1.7</td>
<td>Figures</td>
<td>9</td>
</tr>
<tr>
<td>1.8</td>
<td>References</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Chapter 2 Advanced Electrical Characterization</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Mobility Scattering Mechanisms</td>
<td>16</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Phonon Scattering</td>
<td>17</td>
</tr>
<tr>
<td>2.1.2</td>
<td>Coulomb Scattering</td>
<td>19</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Surface Roughness</td>
<td>20</td>
</tr>
<tr>
<td>2.1.4</td>
<td>Multiple Scattering Mechanisms</td>
<td>20</td>
</tr>
<tr>
<td>2.2</td>
<td>General Effective Mobility Measurements</td>
<td>21</td>
</tr>
<tr>
<td>2.3</td>
<td>Hall Mobility</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Series Resistance Correction</td>
<td>24</td>
</tr>
<tr>
<td>2.5</td>
<td>Other Resistances</td>
<td>26</td>
</tr>
<tr>
<td>2.6</td>
<td>Ultra Thin Oxide Mobility</td>
<td>26</td>
</tr>
<tr>
<td>2.7</td>
<td>Mobility Corrections for Alternative Materials</td>
<td>29</td>
</tr>
<tr>
<td>2.8</td>
<td>Sources of Error in Mobility Measurements</td>
<td>31</td>
</tr>
<tr>
<td>2.9</td>
<td>Pulsed Techniques</td>
<td>32</td>
</tr>
<tr>
<td>2.9.1</td>
<td>Two Level Charge Pumping</td>
<td>32</td>
</tr>
<tr>
<td>2.9.2</td>
<td>Three Level Charge Pumping</td>
<td>35</td>
</tr>
<tr>
<td>2.9.3</td>
<td>Pulsed Drain Current</td>
<td>36</td>
</tr>
<tr>
<td>2.10</td>
<td>Figures</td>
<td>38</td>
</tr>
<tr>
<td>2.11</td>
<td>References</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Chapter 3 Experimental Description</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Standard MOSFET Process Flow</td>
<td>49</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Hafnium Oxide</td>
<td>49</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Metal Gate Deposition Process</td>
<td>50</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.4</td>
<td>Mobility values</td>
<td>112</td>
</tr>
<tr>
<td>6.5</td>
<td>The Role of Interfacial and Bulk Traps</td>
<td>113</td>
</tr>
<tr>
<td>6.6</td>
<td>Other Scattering Mechanisms</td>
<td>116</td>
</tr>
<tr>
<td>6.7</td>
<td>Thin High K Mobility Degradation</td>
<td>119</td>
</tr>
<tr>
<td>6.8</td>
<td>Figures</td>
<td>125</td>
</tr>
<tr>
<td>6.9</td>
<td>References</td>
<td>139</td>
</tr>
<tr>
<td>7.1</td>
<td>Goals of the Study</td>
<td>141</td>
</tr>
<tr>
<td>7.2</td>
<td>Interfacial Layer Conclusions</td>
<td>141</td>
</tr>
<tr>
<td>7.3</td>
<td>Candidate Metal Gates Conclusions</td>
<td>142</td>
</tr>
<tr>
<td>7.4</td>
<td>Aggressively Scaled MOSFET Conclusions</td>
<td>143</td>
</tr>
<tr>
<td>7.5</td>
<td>Major Contributions of this Work</td>
<td>145</td>
</tr>
<tr>
<td>7.6</td>
<td>Future work</td>
<td>145</td>
</tr>
<tr>
<td>Appendix A</td>
<td>List of Publications</td>
<td>148</td>
</tr>
<tr>
<td>Appendix B</td>
<td>GEM Mask</td>
<td>151</td>
</tr>
<tr>
<td>Appendix C</td>
<td>High K Dielectrics on SiC</td>
<td>176</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.1. Equivalent oxide thickness scaling in 2004 International Technology Roadmap for Semiconductors ................................................................. 9

Figure 1.2. Bandgap as a function of dielectric constant for several alternative high k dielectrics .......................................................... 10

Figure 1.3. Energy band diagrams for threshold voltage requirements for NMOS and PMOS devices using (a) midgap and (b) dual metal gates ........................................... 10

Figure 1.4. Effective mobility as a function of interfacial layer thickness for various dielectrics and processing conditions ...................................................... 11

Figure 2.1. Soft optical phonon limited mobility for various dielectric constant materials .......................................................... 38

Figure 2.2. Plot of overall mobility with labeled regions limited by various scattering mechanisms ................................................. 38

Figure 2.3. Schematic of experimental setup to measure (a) gate to channel capacitance for Split CV, (b) gate to substrate capacitance, and (c) total gate capacitance ......................................................................................................... 39

Figure 2.4. Gate to channel capacitance as a function of gate bias for a device with 3 nm SiO$_2$ and TaN gate. The integral under this curve gives inversion charge, Q$_{inv}$ .......................................................... 40

Figure 2.5. Circuit model used by capacitance meter showing (a) parallel model and (b) corrected parallel model with series resistance, $R_s$ ........................................................................................................ 40

Figure 2.6. Resistances of MOSFET channel ........................................................................................................ 41

Figure 2.7. A decrease in mobility as a function of effective field is seen for ultra thin oxides after about 1.9 nm$^2$ .......................................................... 41

Figure 2.8. Effective mobility extracted using leakage correction involving the subtraction of two drain currents taken at different drain biases .................................................. 42

Figure 2.9. A decrease in mobility as a function of effective field is seen for ultra thin oxides even after $N_s$ correction .................................................. 42
Figure 2.10. Schematic of experimental setup for square wave two level charge pumping. .................................................................43
Figure 2.11. Schematic of square wave pulse used in two level charge pumping. ........................................................................43
Figure 2.12. Base level sweep charge pumping for 6 nm SiO$_2$. ..................................................44
Figure 2.13. Amplitude sweep charge pumping on 6 nm SiO$_2$. ..................................................44
Figure 2.14. Schematic of waveform used for three level charge pumping for an n-channel device. ...............................................................................45
Figure 2.15. Schematics of the Inversion-Charge Pumping (ICP) setup used to extract the full inversion charge in long channel MOSFETs using the device symmetry$^{24}$ ..........................................................45
Figure 2.16. Schematic of the inverter circuit used to characterize the drain current through a pulsed technique$^{28}$ ..............................................................46
Figure 2.17. Effect of lowered charge trapping on mobility extraction when using the Inversion Charge Pumping technique on a device with 2.1 nm SiO$_2$ $^{24}$ ..........................................................46
Figure 3.1 Schematic of UHV sputtering system ........................................................................58
Figure 3.2. Plot of atomic ratio of nitrogen to tantalum versus the flow ratio of N$_2$ to Ar$^4$ ........................................................................58
Figure 3.3. Atomic percent of Ta and Ru in RuTa alloys as a function of sputtering power$^6$. ..............................................................................59
Figure 3.4. Workfunction of RuTa alloys as a function of Ta sputtering power$^6$ ...................................................................................59
Figure 4.1. SIMS showing hydrogen and oxygen concentration of RPECVD silicon nitride. .................................................................69
Figure 4.2. Capacitance versus voltage for SiO$_2$ and Si$_3$N$_4$ NMOS and PMOS devices .................................................................69
Figure 4.3. Capacitance as a function of voltage for varying thickness of Si$_3$N$_4$. ........................................................................70
Figure 4.4. Gate leakage current as a function of field across the SiO$_2$ or Si$_3$N$_4$ for NMOS and PMOS devices ........................................70
Figure 4.5. Transconductance times EOT versus effective field across the dielectric for NMOS and PMOS Si$_3$N$_4$ and SiO$_2$ devices in proportionality to channel mobility.......................................................71

Figure 4.6. Schematic of Carrier Separation technique in NMOS device for isolating currents due to holes from electrons.............................................................71

Figure 4.7. Carrier separation for SiO$_2$ NMOS and PMOS devices showing typical behavior.................................................................................................72

Figure 4.8. Carrier separation for Si$_3$N$_4$ NMOS and PMOS devices...................72

Figure 4.10. Normalized defect density as a function of the over-coordination of the interface for PMOS devices...............................................................73

Figure 4.11. Transconductance values for 6 Å SiO$_2$ interfacial layer under Si$_3$N$_4$. ................................................................................................................74

Figure 5.1. CV curves for metal gate electrodes on SiO$_2$ and HfO$_2$. The solid symbols represent HfO$_2$ and the open symbols represent SiO$_2$ gate dielectrics..........................................................91

Figure 5.2. SIMS analysis of W/Ru/HfO$_2$ gatestacks from the topside at 900°C and compared with as deposited. .................................................................91

Figure 5.3. Drain currents of metal gate electrodes on (a) SiO$_2$ and (b) HfO$_2$. ....92

Figure 5.4. Leakage currents versus accumulation and inversion fields for metal gates on HfO$_2$.........................................................................................93

Figure 5.5. Charge pumping current versus base sweep voltage for TaN on HfO$_2$. ........................................................................................................93

Figure 5.6. Interface charge (N$_{it}$) as a function of amplitude bias for metal gate electrodes on HfO$_2$.................................................................94

Figure 5.7. D$_{it}$ measured from two level charge pumping at 100 kHz for metal gates on SiO$_2$ and HfO$_2$. .................................................................94

Figure 5.8. Pulsed drain current measurements for metal gates on SiO$_2$...........95

Figure 5.9. Pulsed drain current measurements for metal gates on HfO$_2$. .........95

Figure 5.10. D$_{it}$ as a function of bandgap for metal gate electrodes on HfO$_2$ extracted from three level charge pumping.................................96
Figure 5.11. Hole mobility for Ru and polysilicon MOSFETs with SiO$_2$ gate dielectric.................................................................96

Figure 5.12. Electron mobility for RuTa and polysilicon MOSFETs with SiO$_2$ gate dielectric.................................................................97

Figure 5.13. Interface trap density for Ru, RuTa and polysilicon MOSFETs as measured by the AC conductance method.................................97

Figure 5.14. Mobilities of metal gates on SiO$_2$ dielectrics extracted by Split CV analysis.................................................................98

Figure 5.15. Mobilities of metal gates on HfO$_2$ dielectrics extracted by Split CV analysis.................................................................98

Figure 5.16. Mobility of metal gates on SiO$_2$ and HfO$_2$ as a function of inversion charge .................................................................99

Figure 5.17. Mobility of metal gates on SiO$_2$ as a function at 20°C and 100°C. ......99

Figure 5.18. Mobilities for metal gate electrodes on HfO$_2$ as a function of increasing temperature. .................................................100

Figure 5.19. Peak mobility as a function of temperature for metal gates on SiO$_2$ and HfO$_2$ dielectrics.........................................................100

Figure 5.20. Coulombic scattered mobility component as a function of inversion charge for various doping and interface charge densities using the values in [16]. .................................................................101

Figure 5.21. TEM of metal gate electrodes on HfO$_2$ ..............................................102

Figure 6.1. Electron mobility for TaSiN gate electrode MOSFETs$^1$ .......................125

Figure 6.2. Capacitance of TiN and polysilicon gate electrodes on 30 Å HfO$_2$......125

Figure 6.3. Drain current as a function of drain voltage for metal gate and polysilicon on thin HfO$_2$ .................................................................126

Figure 6.4. Gate leakage current for metal gate and polysilicon on thin HfO$_2$ as a function of effective dielectric field.................................126

Figure 6.5. Interfacial traps as a function of bias measured by amplitude sweep charge pumping for metal and polysilicon gates on thin HfO$_2$ .................................................................127
Figure 6.6. Drain current versus gate voltage resulting from pulsed drain current measurements.................................................................127

Figure 6.7. Tunneling from pulsed drain current measurements for thin and thicker interfacial layers. .........................................................................................................................128

Figure 6.8. Split CV extracted electron mobilities as a function of effective field for TiN and Polysilicon gate electrodes on ultrathin HfO$_2$...........128

Figure 6.9. Effective electron mobility as a function of inversion charge for TiN and polysilicon gate electrodes on ultrathin EOT HfO$_2$...........129

Figure 6.10. Split CV and NCSU Hauser Mob2D extraction of electron mobility for TiN and polysilicon gate electrodes..................................................130

Figure 6.11. Effective electron mobility using D$_{II}$ corrections illustrated by Yale or by converting to a capacitance for TiN and polysilicon gate electrodes on thin HfO$_2$.................................................131

Figure 6.12. Mobility extracted from DC or pulsed drain current in conjunction with Mob2D.................................................................................................132

Figure 6.13. Effective electron mobility as a function of effective field for 20 °C, 100 °C, and 150 °C for both TiN and polysilicon gate electrons on thin HfO$_2$.................................................................................................133

Figure 6.14. Peak mobility versus temperature for aggressively scaled MOSFETs and compared with candidate metal gates on HfO$_2$ from Chapter 5...........................................................................133

Figure 6.15. Slope of peak mobility versus temperature plotted as a function of EOT for HfO$_2$ dielectrics.................................................................134

Figure 6.16. Mobility as a function of effective field calculated by the Hauser NCSU Mob2D model for different interface scattering parameters..................................................................................134

Figure 6.17. Modeled peak mobility versus temperature for different interface scattering and surface roughness parameter values using the Hauser NCSU Mob2D model.................................................................135

Figure 6.18. TEM of aggressively scaled MOSFETs.................................................................136

Figure 6.19. Effective mobility and calculated phonon limited mobility using Matthiessen’s rule for polysilicon and TiN gates on thin HfO$_2$...........137
Figure 6.20. Corrected effective mobility and calculated phonon limited mobility using Matthiessen’s rule for polysilicon and TiN gates on thin HfO$_2$. ...............................................................137

Figure 6.21. Mobility values versus EOT for both SiO$_2$ and high k with polysilicon and metal gates amalgamated from various published reports$^{12-15}$ .................................................................138
List of Tables

Table 2.1. Permittivities for various dielectrics\textsuperscript{4} ................................................................. 19
Table 3.1. Summary of standard MOSFET process flow. ............................................................. 56
Table 3.2. Gate electrode sputtering conditions of standard MOSFET process .............................................................. 57
Table 3.3. Brief summary of aggressively scaled MOSFET process flow. ......................... 57
Table 5.1. Extracted values for metal gate electrodes on SiO\textsubscript{2} and HfO\textsubscript{2} ................... 79
Table 5.2. Constants used to extract average D\textsubscript{it} from two level base sweep charge pumping .................................................................................. 81
Table 5.3. Parameters used in three level charge pumping to determine volume trap density of metal gate electrodes on HfO\textsubscript{2} .................................................... 84
Table 6.1. Extracted values of ultrathin high k devices with metal and polysilicon gate electrodes ........................................................................................................ 107
Table 6.2. Average D\textsubscript{it} values for TiN and polysilicon on thin HfO\textsubscript{2} using two level charge pumping ............................................................................. 109
Table 6.3. Charge density results from pulsed drain current measurements of TiN and polysilicon gate electrodes on thin HfO\textsubscript{2} ................................................. 111
Chapter 1 Introduction

1.1 Scaling Technology

In order to satisfy demand for increased circuit speed and packing density with decreased power dissipation, MOSFET device dimensions are being scaled according to the 2003 International Technology Roadmap for Semiconductors\(^1\). Decreasing device feature sizes shown in Figure 1.1 lead to short channel effects and reduced gate control which may be minimized with an equivalent gate oxide thicknesses (EOT) of less than 1 nm for a 70 nm channel length proposed for high performance devices in 2006. At this thickness of SiO\(_2\), devices are hampered by increased gate leakage, questionable reliability, large polysilicon depletion effects, and boron penetration into the channel. Due to the limitations of SiO\(_2\) as well as polysilicon, introducing alternative materials into the CMOS process may be the only means to continue scaling into the next decade.

Gate leakage current reduces the on-off current ratio of the device, creating heat and possibly breaking down the device. While increasing the thickness of the SiO\(_2\) will decrease this leakage, the capacitance and consequently the speed and drive of the device will decrease. One approach to decrease the leakage current while maintaining an appropriate capacitance is to replace SiO\(_2\) with a material possessing a higher dielectric constant denoted high k dielectrics. These high k dielectrics reduce leakage current by providing a physically thicker film while electrically behaving as a thinner dielectric. The most investigated dielectrics today are hafnium based dielectrics including HfO\(_2\), HfSiO\(_4\), and HfSiON which provide a
dielectric constant ranging from 12–30 with reasonable thermal stability on Si.\textsuperscript{2-4} However, work is also being performed on zirconium, aluminum, lanthanum, and yttrium based dielectrics\textsuperscript{5-13}. High k dielectrics have proven to decrease EOT while also decreasing leakage current. Both chemical and physical methods are being used to deposit these dielectrics.

Polysilicon depletion can also adversely decrease the capacitance of the device and only becomes a larger percentage of the thickness as the EOT is decreased. Increasing the doping does minimize the depletion capacitance but this only leads to greater boron penetration and threshold voltage shifts. Furthermore, the limit of polysilicon doping is \( \sim 10^{21} \text{ cm}^{-3} \) which may not be enough to recover the capacitance. It is suggested that replacing the gate electrode with metals, which have at least an order of magnitude higher carrier concentration and do not suffer from depletion, will tackle these issues. Much investigation is ongoing concerning alternative metal gate electrodes with work on titanium, tantalum, ruthenium, tungsten as well as their nitrides, silicides, and alloys\textsuperscript{14-18}. Metal gate electrodes have been successfully incorporated into the CMOS process and owing to their high carrier concentration show no signs of depletion.

### 1.2 Desired Properties

There are many properties of high k dielectrics and metal gate electrodes which are desired to meet the needs of the ITRS. For instance, the reduction in gate leakage current warrants a dielectric constant value 8-30 and a barrier height between the substrate and dielectric to be greater than 1.0 eV\textsuperscript{19}. Too high a
dielectric constant will cause large fields to be placed along the channel edges and too low a barrier height will increase the tunneling across the gate. There is a general trend that as the dielectric constant increases, the bandgap decreases as indicated by Figure 1.2. This leads to lower barrier heights and hence greater thermal emission and tunneling. Therefore, both of these fundamental properties should be examined before consideration as an alternative dielectric. Furthermore, the dielectric should have low oxygen diffusivity and good thermal stability with silicon so that the interfacial reaction with the underlying semiconductor is minimized.

Metal gate properties which are desired include having an appropriate workfunction for NMOS and PMOS. This means that the workfunction should be within 0.2 eV of the conduction and valence band edge of silicon as described by Figure 1.3. It should also have good thermal stability with the underlying dielectric, signifying that the choice of high k dielectric is vital from a metal processing standpoint. Metal gate candidates should also have a low diffusivity to oxygen and other dopants as well as possessing a high enough carrier concentration so that depletion effects are negligible.

1.3 Current Issues of Advanced Gatestacks

With the introduction of novel materials as the gate dielectric and the gate electrode, serious problems arise either during or as a result of the processing. The first challenge is that the deposition of high k dielectrics on Si tends to form interfacial layers which are SiO$_2$ rich if not stoichiometric SiO$_2$. This hinders the ultra
thin EOTs needed in future devices. Titanium oxide (TiO$_2$), for example, showed interfacial layers even at low temperatures$^{20,21}$ and thus was ruled out as a possible alternative for SiO$_2$. Nitrogen based material or nitrogen annealing can be used to decrease or avoid completely the interfacial layer$^{12,22}$.

Since the dopants of the source and drain need to be activated, high temperatures are required and these can lead to unwanted reactions in the (i) dielectric, (ii) between the dielectric and Si, (iii) between the metal gate and the dielectric, and (iv) possibly even metal diffusion into the channel. For example, HfSiO$_4$ is known to phase separate into layers of SiO$_2$ and HfO$_2$ after annealing due to the high oxygen diffusivity of the material$^3$. This is an undesirable effect which increases EOT. However, there have been reports that HfSiO$_4$ and ZrSiO$_4$ have been stable up to 950°C$^{23,24}$. Possible inclusion of N or even Al have been proposed to deter the phase separation and may also impede dopant diffusion$^{25,26}$.

High temperatures also change the film morphology of the dielectric and can result in crystalline or polycrystalline materials. The large grain boundaries associated with a polycrystalline film allow carriers and dopants to move easily through the material which will increase the leakage or create unwanted threshold voltage shifts. Currently, the impact of crystallization of the dielectric is not fully understood$^{27}$. Reactions between the dielectric and polysilicon may occur under high temperature processing which can also create flatband and threshold voltage shifts$^{28-30}$. This may also be the cause of higher EOTs with polysilicon gates than with metal gates.
Finally, reactions between the polysilicon or metal gate electrode with the dielectric may result in Fermi level pinning and further the threshold voltage shift problem\textsuperscript{28, 29}. Fermi level pinning occurs when states at the interface are charged and cause a dipole to form driving the band alignment to change so that zero dipole will exist\textsuperscript{31}. This tends to shift the Fermi level towards the charge neutrality level and hence pinning of the Fermi level occurs at the interface. Reports of polysilicon Fermi level pinning with hafnium based dielectrics have been published \textsuperscript{32-35} but currently candidate metal gates do not indicate this behavior\textsuperscript{36, 37}.

### 1.4 Advanced Gatestack Mobility

One of the greatest issues with high k dielectrics is that both polysilicon and metal gate electrodes is that the devices show very poor mobilities\textsuperscript{11, 18, 38}. Furthermore, it has been shown that as the thickness of the dielectric decreases, the mobility values decrease as well\textsuperscript{39}. This can be seen in Figure 1.4 where the mobility values of various dielectrics with different processing conditions are plotted against the interfacial layer thickness. Different mechanisms can ‘scatter’ the electrons as they drift along the channel and degrade the effective surface mobility. Since the lattice of the substrate is comprised of Si atoms which thermally vibrate in phonon energy quanta, phonon scattering can reduce the momentum of the electrons and hence, reduce its mobility. In addition, coulombic scattering due to ionized impurities in the channel, as well as interface trapped charge or charge within the oxide can occur. It can be screened by the inversion layer in the channel especially under high fields when there are a large number of inversion carriers.
However, at lower transverse fields where this mechanism dominates, coulomb scattering is proposed as one of the main reasons for the low mobilities of high k dielectrics\textsuperscript{40-43}. Finally, a surface roughness scattering mechanism dominant at high transverse fields, where electrons are pulled tightly to the dielectric–Si interface, causes mobility to decrease. All these scattering mechanisms contribute to decreasing the mobility and the mechanism with the lowest mobility dominates the total mobility according to Mathiessen’s rule.

\[
\frac{1}{\mu_{tot}} = \frac{1}{\mu_{phonon}} + \frac{1}{\mu_{Coulomb}} + \frac{1}{\mu_{rough}} + \ldots \tag{1.1}
\]

The above mechanisms are very well established for thick EOT SiO\textsubscript{2} devices with polysilicon electrodes. As the dielectric becomes thinner, however, many of these mechanisms will have to be reevaluated. Also, with the implementation of high k dielectrics and metal gate electrodes, these mechanisms may not be sufficient to explain the observed results. Other mechanisms such as soft optical phonons or charge trapping from traps located within the bulk of the high k dielectric are being investigated as sources of mobility degradation\textsuperscript{44, 45}.

1.5 Objective

The purpose of this work is to determine the effects of emerging candidate metal gate electrodes and alternative dielectrics on device performance, especially mobility. Device parameters will be extracted from fully characterized MOSFET devices and qualitative descriptions of the materials will be proposed.
1.6 Overview of Dissertation

This dissertation will correlate electrical characterization results with the material properties to obtain an understanding of the mobility behavior of advanced gate stacks.

Chapter 2 provides a detailed introduction to electrical characterization of advanced MOSFETs with a discussion on factors affecting MOSFET mobility.

Chapter 3 is a description of the experiments used to fabricate the devices used in this work.

Chapter 4 describes the impact of an interfacial layer on the mobility of Si$_3$N$_4$ MOSFETs.

Chapter 5 discusses the differences in mobility between metal gates on SiO$_2$ and alternative dielectrics.

Chapter 6 is a study on the mobility values of aggressively scaled EOT MOSFET devices.

Chapter 7 concludes the work and proposes future work in this area.

Appendix A lists the publications and presentations of this author.

Appendix B explains the GEM mask created by the author for use with charge pumping analysis and is used in the process flow of the standard MOSFETs results obtained for this dissertation.
Appendix C investigates high k dielectrics on wide bandgap semiconductor, SiC which was the initial work of the author.
1.7 Figures

Figure 1.1. Equivalent oxide thickness scaling in 2004 International Technology Roadmap for Semiconductors.
Figure 1.2. Bandgap as a function of dielectric constant for several alternative high k dielectrics.

Figure 1.3. Energy band diagrams for threshold voltage requirements for NMOS and PMOS devices using (a) midgap and (b) dual metal gates.
Figure 1.4. Effective mobility as a function of interfacial layer thickness for various dielectrics and processing conditions.\textsuperscript{39}
1.8 References

33. C. S. Kang et al., "Improved thermal stability and device performance of ultra-thin (EOT less than or equal 10a) gate dielectric MOSFETs by using hafnium oxynitride (HfO_2N_y)," in International Symposium on VLSI Technology Technical Digest, Honolulu, HI, United States, 2002.
44. M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, "Effective electron mobility in Si inversion layers in metal--oxide--semiconductor systems with a high-


Chapter 2 Advanced Electrical Characterization

Electrical characterization measurements performed on MOSFET devices are necessary to determine the parameters which describe the quality of the device. With the advent of high k dielectrics and other alternative materials, electrical characterization techniques may need altering to account for different dominant mechanisms occurring during the measurement. Since capacitance – voltage and DC current – voltage measurements are standard in this field, they will not be discussed. However, this chapter will provide information on electrical characterization of advanced MOSFET devices including mobility, interfacial trap density and bulk trapping measurements.

2.1 Mobility Scattering Mechanisms

One of the most important characteristics of a MOSFET device is the channel mobility. A measure of the ease with which the inversion carriers travel along the channel, the mobility is sensitive to the quality of the channel, the dielectric, and even the gate electrode. Properties such as charges, lattice vibrations, and localized surface potentials can scatter the moving carriers and decrease their mobility, reducing the effectiveness of the device. Because low reported mobility values for devices with ultrathin SiO₂ or high k dielectrics have proven to be one of the greatest challenges for implementation, many of the effects of these different scattering mechanisms are still being investigated to determine why such degradation is occurring\(^1\)\(^\text{2}\). In order to understand more clearly the behavior of
effective mobility, it is important to break down the mechanisms which can affect the inversion layer mobility. Among the scattering mechanisms, coulomb scattering, phonon scattering, and surface roughness are the most prevalent in CMOS devices.

2.1.1 Phonon Scattering

Phonon scattering of the inversion carriers originates from the acoustic and optical mode vibrations of the lattice. As the temperature increases, lattice vibrations also increase and thus the phonon scattering rate also increases. At temperatures below room temperature, phonon scattering can be ignored. However, at room temperature, phonon scattering plays an important role on the inversion layer mobility. It has been previously shown experimentally by Takagi that mobility limited by phonon scattering can be determined from

\[ \mu_{ph} = A E_{eff}^{-0.3} T^{-1.75} \]  

where \( A \) is a constant calculated to be 2 x 10^5 and 6.1 x 10^4 for electron and hole mobility on (100) Si, respectively, \( E_{eff} \) is the effective field, and \( T \) is the temperature in Kelvin.

Recently, Fischetti predicted that the remote phonon scattering originating from the gate dielectric can not be neglected in MOSFET devices with high k materials. Since the bonds in most high k dielectrics are metal-oxygen bonds, they are highly polarizable and result in a large static permittivity, \( \varepsilon_{ox}^0 \), desirable for advanced CMOS scaling. However, inherent in these soft bonds are low energy lattice oscillations which are optical in nature and derive from their ionic characteristics. These low energy phonons trigger frequent emission and absorption
processes by thermal electrons in the inversion layer, resulting in a remote phonon scattering mechanism which, due to the dipole field of the insulator, decays away from the bulk of the insulator into the inversion layer. Conversely, SiO$_2$ yields hard Si-O bonds and a reduced ionic polarization with a high energy of optical phonons. It is difficult for thermal electrons to emit excitations at these energies especially at room temperature where the number of absorbed thermally excited phonons is low.

The scattering strength of the optical phonons can be found using

$$S = \hbar \omega_{SO} \left[ \frac{1}{\varepsilon_{Si}^\infty + \varepsilon_{ox}^\infty} - \frac{1}{\varepsilon_{Si}^\infty + \varepsilon_{ox}^0} \right]$$  \hspace{1cm} (2.2)$$

where $\omega_{SO}$ is the frequency of the surface optical phonon and $\varepsilon_{Si}^\infty$ is the optical permittivity of the semiconductor, silicon. Since high k dielectrics have such a large difference between their static permittivity, $\varepsilon_{ox}^0$, and their optical permittivity, $\varepsilon_{ox}^\infty$, the resulting scattering strength is large. For SiO$_2$, this difference is small given the hard bond nature of Si-O, so the scattering strength of SiO$_2$ optical phonons is small. Typical numbers for various dielectrics are given in Table 2.1. Notice that the difference in permittivity is least for SiO$_2$ and greatest for ZrO$_2$ and HfO$_2$. Also note that ZrSiO$_4$ has a lower difference than ZrO$_2$ indicating that the scattering strength should be less for silicates.
Table 2.1. Permittivities for various dielectrics$^4$.

<table>
<thead>
<tr>
<th>Quantity $\varepsilon_{ox}^o$</th>
<th>SiO$_2$</th>
<th>Al$_2$O$_3$</th>
<th>ZrO$_2$</th>
<th>HfO$_2$</th>
<th>ZrSiO$_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{ox}^\infty$</td>
<td>2.5</td>
<td>3.2</td>
<td>4.0</td>
<td>5.0</td>
<td>4.20</td>
</tr>
<tr>
<td>$\varepsilon_{ox}^o - \varepsilon_{ox}^\infty$</td>
<td>1.4</td>
<td>9.3</td>
<td>20.0</td>
<td>17.0</td>
<td>7.55</td>
</tr>
</tbody>
</table>

Since the scattering rate and the scattering strength of optical phonons in high $k$ dielectrics is larger than that of SiO$_2$, remote phonon scattering is expected to cause a greater degradation in the effective mobility of inversion carriers in MOSFETs with high $k$ insulators. This is illustrated in Figure 2.1 where the theoretical soft optical limited mobility is plotted as a function of dielectric constant for a constant inversion density of $1.54 \times 10^{12}$ cm$^{-2}$. With the exception of AlN, as the dielectric constant is increased, the soft optical limited mobility decreases, perhaps indicating an unavoidable trend of high $k$ dielectric MOSFETs. One of the objectives of this work is to confirm the affinity of alternative dielectrics in MOSFET devices and lower mobility values.

2.1.2 Coulomb Scattering

Scattering from charged centers such as fixed charges, interface traps in the gate dielectric, and ionized impurities in the depletion layer of the substrate is designated as coulomb scattering. As the field increases and more carriers enter the inversion layer, screening can occur and mobility limited by coulomb scattering increases$^3$. This means that coulomb scattering limited mobility is dominant at lower fields where the surface carrier density is small. However, increased doping
concentration and fixed charge density decrease the mobility arising from coulomb scattering. The temperature dependence of coulomb scattered mobility increases as temperature increases since the velocity of the inversion carriers increase and interact less effectively with stationary impurities.

2.1.3 Surface Roughness

Since the silicon-insulator interface is not perfect, the surface potential can fluctuate on the molecular level and the small deviations of the interface from the ideal plane can cause perturbations in the inversion carrier mobility. On a larger scale, the surface roughness of the interface decreases the mobility of the inversion layer. As the field is increase, the inversion carriers are pulled more tightly to the interface and surface roughness limited mobility becomes increasingly smaller. Therefore, the mobility limited by surface roughness has an inverse relationship with effective field with a power, \( \gamma \), close to 2 as seen in eqn. (2.3)\(^3,5\).

\[
\mu_{sr} \propto E^{-\gamma} \tag{2.3}
\]

2.1.4 Multiple Scattering Mechanisms

When multiple scattering mechanisms are present, the overall mobility is generally described by Matthiessen’s rule:

\[
\frac{1}{\mu} = \sum_j \frac{1}{\mu_j} = \frac{1}{\mu_{ph}} + \frac{1}{\mu_{coul}} + \frac{1}{\mu_{sr}} + \ldots \tag{2.4}
\]

Since there is a reciprocal relationship of this sum, the effective mobility is dominated by the process with the lowest value of mobility. Coulomb scattered
limited mobility is dominant at lower fields since there is no screening of the inversion layer and surface roughness mobility is dominant at higher fields since the carriers are pulled closely to the imperfect interface. This leaves phonon scattering dominant at fields which have intermediate values. These regions are labeled in Figure 2.2, a plot of the overall mobility as a function of the effective field. As temperature increases, this curve shifts downward. However, there may be an increase in the lower field region since coulomb scattering decreases as temperature increases.

2.2 General Effective Mobility Measurements

In order to speculate on the different scattering mechanisms and their effects on channel mobility in any device, accurate measurements of high k mobility must be performed. In a long channel device, for a sufficiently low drain voltage, \( V_d \), such that the transistor is operating in the linear region, the drain current can be described by

\[
I_d = \frac{W}{L} \mu_{eff} Q_{inv} (V_g) V_d
\]  

where \( W \) and \( L \) are the precise channel width and length of the device, respectively, \( \mu_{eff} \) is the effective mobility of the carriers in the inversion layer, and \( Q_{inv} \) is the inversion charge density, which is a function of the gate bias. Rearranging eqn. (2.5), \( \mu_{eff} \) can be determined from

\[
\mu_{eff} = \frac{L}{W} \frac{I_d (V_g)}{Q_{inv} (V_g)} = \frac{L}{W} \frac{g_d}{Q_{inv}}
\]  

where the drain conductance \( g_d \) is defined as
\[ g_d = \frac{I_d}{V_d} \]  

(2.7)

or more specifically,

\[ g_d = \frac{\partial I_d}{\partial V_{ds}} \bigg|_{V_{gs}=const \tan t} \]  

(2.8)

By far, the most important parameter in eqn. (2.6) is \( Q_{inv} \). There are several ways to measure inversion charge. At one time, \( Q_{inv} \) was approximated by

\[ Q_{inv} = C_{ox} (V_{gs} - V_t) \]  

(2.9)

where \( C_{ox} \) is the oxide capacitance and \( V_t \) is the threshold voltage of the device. Since the threshold is not exactly known and some of the charges may be trapped at or near the interface, this can lead to inadequate estimations of the true inversion charge, especially near \( V_t \). A better approximation is to use the Split CV method to integrate as in

\[ Q_{inv} = \int_{-\infty}^{V_{gs}} C_{gc} \, dV_{GS} \]  

(2.10)

where \( C_{gc} \) is the gate to channel capacitance. \( C_{gc} \) is measured by connecting the capacitance meter between the gate and the source-drain while grounding the substrate as seen in the schematic of Figure 2.3 (a). At voltages below threshold, the capacitance is equal to the total overlap capacitance of the gate to the source and drain, \( 2C_{ov} \). As the channel starts to invert and inversion charge appears, the capacitance increases. As the voltage is increased further, the capacitance saturates to some inversion value which is a combination of the overlap capacitance and the channel capacitance, \( 2C_{ov} + C_{ch} \). A typical \( C_{gc} - V_{gs} \) of a 3.0 nm device with \( \text{SiO}_2 \) dielectric can be seen in Figure 2.4. Subtracting \( 2C_{ov} \) and integrating this curve gives \( Q_{inv} \) as a function of \( V_{gs} \) also seen in Figure 2.4. The equipment setup for
measuring gate to substrate and total gate capacitance can be seen in Figure 2.3 (b) and (c), respectively. Integrating gate to substrate capacitance will result in depletion charge, $Q_d$.

The drain conductance, $g_{d}$, can be determined from eqn. (2.7) using a single low drain bias, usually 10–50 mV, and measuring the DC drain current characteristics. It is important to note that using a drain bias introduces an error in $\mu_{eff}$, especially near $V_t$, since the gate to channel capacitance is measured with no drain bias and inversion charge decreases as drain bias increases. However, a drain bias is obviously needed to measure the drain current and calculate $g_d$ so it is suggested to use as low a value as possible.

### 2.3 Hall Mobility

The Hall effect measurement of an MOS structure provides values of mobile inversion carrier density, $Q_{inv}$, and Hall mobility, $\mu_H$. In a semiconductor, a magnetic field, $B_z$, is applied perpendicular to the direction of motion which deflects the carriers by way of the Lorentz force. The deflection causes a Hall voltage across the sample given by

$$V_H = \frac{R_H IB_z}{d}$$  \hspace{1cm} (2.11)

where $I$ is the current through the sample, $d$ is the sample thickness, and $R_H$ is the Hall coefficient. The Hall coefficient is equal to $R_H = r_H/(qp) = - r_H/(qn)$ where $r_H$ is the Hall factor and $n$ or $p$ are the carrier concentrations. For an MOS Bridge-type Hall structure which has two inversion layer contact tabs along the channel $\mu_H$ is given by
\[ \mu_H = \frac{V_H}{B_z, l R \frac{W}{L}} \]  

(2.12)

where \( W \) and \( L \) are the transversal and longitudinal separation of the tabs and \( R \) is the longitudinal resistance between the two tabs.

The relationship between Hall mobility and conductive mobility is given by

\[ \mu_H = \gamma \mu_c \]  

(2.13)

where \( \gamma \) is the Hall scattering factor and generally lies between 1 and 2\(^6\). For most mobilities determined by the Hall effect, \( \gamma = 1 \). In SiC substrates, the Hall mobility was used to delineate the mechanisms responsible for the difference in extracted effective mobility over that of the calculated Hall mobility. Low reported effective mobility values were attributed to charge trapping from a large number of interfacial states at the SiC – SiO\(_2\) interface\(^9, 10\). Similar results have been published for Si with high k dielectrics where correcting for the trapped carriers in interfacial states resulted in similar effective mobility as those obtained from Hall mobility measurements\(^11\). Furthermore, the Hall mobilities extracted were smaller for high k dielectrics compared with those of SiO\(_2\) still leaving the cause of mobility degradation unclear. Given these observations, no Hall structures were fabricated for Hall mobility measurements in this work.

### 2.4 Series Resistance Correction

Series resistance (\( R_s \)) can cause serious errors in capacitance – voltage measurements\(^6, 12\). The source of \( R_s \) arises from the probing contacts, the backside contact, the resistance of the quasi-neutral bulk silicon between the backside contact
and the depletion layer edge at the silicon surface underneath the gate, or any an extreme nonuniform doping distribution in the silicon underneath the gate. Minimization of $R_s$ can be accomplished by improved processing or using low frequencies so that the effect of $R_s$ is negligible. However, the high frequency measurement can be corrected for series resistance if the value is known.

In order to measure $R_s$, the device is biased into accumulation where the oxide capacitance and series resistance dominate all other admittance elements in the device. High frequencies are used to measure $R_s$ so that interface trap response is minimal. The real part of the measured admittance of the structure at high frequency and at the accumulation bias is $R_s$ and can be calculated from

$$R_s = \frac{G_{ma}}{G_{ma}^2 + \omega^2 C_{ma}^2}$$  \hspace{1cm} (2.14)

where $G_{ma}$ and $C_{ma}$ are the measured conductance and capacitance in accumulation, respectively.

$R_s$ can cause errors especially if the capacitance meter assumes the device is represented by the parallel equivalent circuit shown in Figure 2.5 (a). When the series resistance is included in this model, both the measured capacitance and conductance are affected as shown in Figure 2.5 (b). In this case, the corrected capacitance is

$$C_c = \frac{C_m}{\left(1 - G_m R_s\right)^2 + \omega^2 C_m^2 R_s^2}$$  \hspace{1cm} (2.15)

and the corrected conductance is
where \( \omega = 2\pi f \) and \( f \) is the frequency. These relationships indicate that the series resistance modifies the measured capacitance and conductance and this modification is worse for high frequencies and high conductance values.

### 2.5 Other Resistances

The channel of a MOSFET device is also plagued with resistances. In Figure 2.6, these resistances are labeled as the channel resistance, \( R_{ch} \), the source and drain resistance, \( R_{sd} \), and the contact resistance, \( R_c \). The effect of these resistances is widely known and can be determined through the appropriate test structures mentioned in Appendix B. The contact resistance, especially, becomes more important as channel lengths become shorter and they become comparable to the channel resistance. In this case, measuring the contact resistance from test structures and determining the voltage across the channel, \( V_{d} \) = \( V_d - V_c \) where \( V_c \) is the voltage across the two contacts, \( V_c = I_d 2R_c \). Similar measures can be taken if \( R_{sd} \) is large.

### 2.6 Ultra Thin Oxide Mobility

As the oxide thickness decreases to around or below 30 Å, the mobility also decreases as seen in Figure 2.7, particularly in the low field region. It is unclear what is the primary cause for this degradation but the most widely held theory is that there is remote coulombic scattering from the polysilicon gate electrode\(^2, 13-16\). The gate electrode and the interface of the gate electrode with the gate dielectric play a
more prominent role in the channel characteristics as the gate dielectric thickness decreases. This is because the effect of charge on the channel is inversely proportional to the distance from the gate. The polysilicon depletion present when the device is in inversion can cause remote charge scattering and there has been evidence of polysilicon gate electrode roughness at the gate-oxide interface which can also scatter carriers in the channel\textsuperscript{17}. It is expected that metal gate electrodes will not suffer from remote coulomb scattering since their carrier concentrations are large enough that the Thomas-Fermi screening length is very small. Thus scattering from ionized impurities in the metal gate is not present.

Also with thinner gate dielectrics the gate dielectric leakage increases and tends to distort the $I_d - V_{gs}$ characteristics. The drain current can be corrected for gate leakage by using a weighing factor\textsuperscript{18} which partitions the leakage current to either the source or the drain and assumes that the substrate current is negligible. To determine the weighing factor, the gate current supplied from the source, $I_{gs}$, and from the drain, $I_{gd}$ is measured at zero drain bias. The weighing factors are calculated as

$$\alpha_d = \frac{I_{gd}(V_d = 0, V_g)}{I_g(V_d = 0, V_g)} \quad \text{and} \quad \alpha_s = \frac{I_{gs}(V_d = 0, V_g)}{I_g(V_d = 0, V_g)} \quad (2.17)$$

which means that $\alpha_d + \alpha_s = 1$ and the drain current is computed assuming that the weighing factor is weakly drain bias dependent in the linear region.
It is common to assume $\alpha_s = \alpha_d = 0.5$ if the device is symmetrically processed\(^\text{19}\). This correlates to 50% partitioning of the gate leakage current to the drain. This leakage current correction will be used in this work unless otherwise stated.

Another way to correct for gate leakage is to use the true equation for the drain conductance specified in eqn. (2.8) where two drain currents are measured at two different drain biases, subtracted, and divided by the difference of the drain biases\(^\text{11, 20}\).

A more accurate measure of $\mu_{\text{eff}}$ can be seen in Figure 2.8 when using these approaches.

Other features of thin oxide devices which can cause mobility degradation are surface roughness\(^\text{17}\), interface plasmon scattering\(^\text{4}\), wave function penetration scattering\(^\text{21}\), and impurity scattering due to impurity penetration from the gate. Still another theory is that the inversion charge is inaccurately measured, especially since direct tunneling current is so large at these thicknesses\(^\text{22}\). Since the leakage current affects both the drain conductance, $g_D$, and the inversion charge density, $Q_{\text{inv}}$ which are calculated from the $I_d - V_{gs}$ and $C_{gc}$ curves, respectively, an appropriate channel length is determined to minimize the effect. In reference \(^\text{22}\), it was determined that a 10 $\mu$m x 10 $\mu$m device could accurately give the inversion charge if the channel current was an average of the drain and source current and the
surface inversion charge was an average of the charge at the drain and source as seen in eqn. (2.20) and eqn. (2.21).

\[ I_{ch} = \frac{I_d + I_s}{2} \quad (2.20) \]

\[
N_s = \frac{N_{Source} + N_{Drain}}{2} = \frac{1}{2q} \left( \int_{-\infty}^{V_g} C_{gc} dV + \int_{-\infty}^{V_d-V_g} C_{gc} dV \right) \quad (2.21)
\]

Even with the accurate calculation of inversion charge and the minimization of the effect of leakage current on the calculations, the mobility values were still seen to decrease as a function of oxide thickness as seen in Figure 2.9. This suggests other scattering mechanisms inherent in ultra thin oxide devices are the reason for this mobility degradation when compared with thick oxide universal mobility.

### 2.7 Mobility Corrections for Alternative Materials

Beyond the general and high leakage current corrections for aggressively scaled MOSFET devices with SiO₂ dielectrics, there may be other corrections necessary for devices with alternative high-K and metal gate materials in order to accurately determine the inversion charge and the reasons for any degradation. One such correction is for the high level of interface traps (D_{it}) that are usually present with high k dielectrics. \(^{11, 23}\) D_{it} can respond to the AC signal in the split capacitance measurements, adding to the measured value and leading to an overestimation of \(Q_{inv}\). This effect can be minimized by using a high frequency capacitance measurement. Note that using a higher frequency is limited by the capacitance meter and its accuracy at this frequency. Still, even when using a high
frequency, $D_{it}$ can respond to the slow DC ramp and traps near the band edges can respond to the AC bias, especially if they are highly doped substrates.

There are two ways to correct for $D_{it}$. One technique developed at Yale utilizes a calculated gate to channel capacitance from $^{11}$. 

$$C_{gc} = \left( \frac{1}{C_{ox}^{-1}} + \frac{1}{C_{inv}^{-1}} + \frac{C_{d}}{C_{ox} C_{inv}} \right)^{-1}$$  \hspace{1cm} (2.22)$$

where $C_{ox}$ is the oxide capacitance, $C_{d}$ is the depletion capacitance and the inversion capacitance is defined as

$$C_{inv} = \frac{dQ_{inv}}{d\psi_s}$$  \hspace{1cm} (2.23)$$

and assuming that the traps in the inversion layer can not follow the high frequency ac signal or that the interface trap capacitance, $C_{it}$, can be neglected. Thus, $D_{it}$ values do not have to be known. Integration of $C_{gc}$ in eqn. (2.22) results in the true inversion charge which can be used to calculate the mobility. This is similar to the Hauser Mob2D model developed here at NCSU$^{20}$.

However, knowledge of $D_{it}$ can be very meaningful to understanding the degradation. For example, if the $D_{it}$ as a function of surface band bending is known, the correction for interfacial trapping above can be taken a step further. The resulting capacitance due to the interface traps is

$$C_{it} = q \frac{\partial D_{it}}{\partial \psi_s}$$  \hspace{1cm} (2.24)$$

so that the adjusted gate to channel capacitance can be found from
The integration of this equation provides the true inversion charge to be used in calculating the mobility. In this work, both the Yale method and the Dit-Cit method are used with (i) average Dit and (ii) Dit as a function of the surface band bending. It is important to note that the above methods still rely on a DC $I_d - V_{gs}$ measurement which may be prone to bulk trapping. To this end, pulsed measurements were also performed and are discussed in Section 2.9.3.

2.8 Sources of Error in Mobility Measurements

While corrections can be made for resistances and leakage current there may still be some errors which persist in making mobility measurements utilizing corrections. One source of error may be in the extracted parameters. Dielectric thickness plays an important role in the corrections and the need to know oxide and depletion capacitance in order to determine inversion charge accurately from the gate to channel capacitance becomes highly important. Underestimating EOT, especially, will cause the inversion charge to be overestimated and hence a lower value of mobility to be reported. Errors in extracted doping values will have a similar effect on calculating accurate inversion charge and depletion charge needed for effective field measurements. Other sources of error include the extracted Dit values which may be overestimated due to the bulk type trapping within the dielectric which can be detected with charge pumping measurements discussed below.
2.9 Pulsed Techniques

Pulsed characterization techniques are also being studied to gauge the impact of traps located within many alternate dielectrics being pursued. These techniques have the advantage of varying the response time of various traps within the device depending on the frequency of the pulse. By using pulsed measurements with time ranges on the order of $\mu$s, the effect of charge trapping on carrier mobility can be minimized$^{24}$.

2.9.1 Two Level Charge Pumping

In order to fully understand how to characterize mobility through pulsed techniques, an explanation on the response nature of interfacial traps is necessary. The most popular method in determining interface trap density ($D_{it}$) values in MOSFETs is the charge pumping technique$^{25}$. In two level charge pumping a periodic square wave is applied to the gate while the substrate current is monitored. The amplitude of the square wave is such that the device is driven from inversion to accumulation. The source and drain can be tied together and grounded or a small reverse bias can be applied. A schematic of the experimental setup is shown in Figure 2.10 for an n-channel device. It is important to note that the square wave is, in fact, a trapezoidal waveform with a finite rise and fall time.

Two level charge pumping relies on the Shockley Reed Hall statistics of electron and hole recombination. When the device is in strong inversion and the Fermi level is very close to the conduction band edge, minority carriers populate the inversion layer and fully occupy the interfacial states. As the device falls from
inversion to accumulation during the finite fall time of the pulse, the minority carriers in the inversion layer drift to the source and drain. Additionally, interface traps near the conduction band edge have enough time to emit their minority carriers and they begin to drift toward the source and drain. However, deeper into the bandgap, the minority carriers do not have sufficient time to escape their traps and remain captured. As the device moves into the accumulation regime, majority carriers now flow into the channel. Some of these carriers are captured by those interface traps which are still filled with minority carriers, recombining until all the states are now filled with majority carriers. A similar process occurs again as the device is pulsed back into inversion and the majority carriers unable to escape the traps deeper into the band gap recombine with minority carriers.

Over one entire period of the applied square wave, the net influx of majority carriers is measured as substrate current or charge pumping current, \( I_{cp} \). From eqn. (2.26), \( I_{cp} \) can be related to the average density of interface traps at the semiconductor – dielectric interface.

\[
I_{cp} = 2qD_{\phi}fA_{G}kT \left[ \ln(V_{th}n_i \sqrt{\sigma_n \sigma_p}) + \ln \left( \frac{|V_{FB} - V_T|}{\Delta V_g} \right) \right] \tag{2.26}
\]

where \( q \) is the electron charge, \( f \) is the frequency, \( A_G \) is the gate area, \( V_{th} \) is the thermal velocity of the carriers, \( n_i \) is the intrinsic carrier density, \( \sigma_n \) and \( \sigma_p \) are the capture cross section coefficients for electrons and holes, \( \Delta V_g \) is the peak to peak amplitude of the pulse, and \( t_r \) and \( t_f \) are the rise and fall times seen in Figure 2.11.

Two level charge pumping derives its name from the dual level nature of the square wave pulse. The two levels are denoted the base level, \( V_{gbl} \), and the peak
level, $V_{gbh}$ and can be seen in Figure 2.11. There are two different types of square wave two level charge pumping methods. The first method varies the base level, $V_{gb}$ while keeping the amplitude constant. In this case, denoted base sweep charge pumping, only specific levels of $V_{gb}$ will span the entire band gap and result in a saturation of $I_{cp}$. However as $V_{gb}$ is increased further, the bandgap span becomes less and $I_{cp}$ decreases again. The resulting plot of $I_{cp}$ versus $V_{gb}$ for a 6 nm SiO$_2$ can be seen in Figure 2.12. The other method varies the amplitude of the pulse while keeping $V_{gb}$ constant and is called amplitude sweep charge pumping. In this case, as soon as the amplitude allows the Fermi level to span from inversion to accumulation, any further increase in amplitude does not change $I_{cp}$. The resulting $I_{cp}$ versus $V_{gb}$ plot for the same 6 nm SiO$_2$ can be seen in Figure 2.13 illustrating the plateau in $I_{cp}$. From the maximum $I_{cp}$ of these plots, average $D_{it}$ can be determined.

There exists another recombination process which can affect the measured $I_{cp}$. If the fall time is too short, the inversion layer charge will not have enough time to drift back to the source and drain when moving into accumulation. Instead, some fraction of the minority carriers will recombine with the incoming majority carriers resulting in an increased substrate current and an overestimation of the average $D_{it}$. This additional geometrical current is dependent on the geometry of the device and can be minimized by the use of small channel lengths and appropriately long rise and fall times. For all measurements made in this thesis, constraints were taken to avoid any unnecessary geometrical effects.
Three Level Charge Pumping

It is more interesting to understand asymmetric or increasing $D_{it}$ as a function of the bandgap ($D_{it}(E)$) instead of the average $D_{it}$ across the bandgap. Three level charge pumping is a method which can be used to determine $D_{it}(E)$\(^{26}\). In this charge pumping technique, the device is pulsed from inversion to an intermediary voltage, $V_{step}$, near midgap and then to accumulation. The waveform applied to the gate of an n-channel device can be seen in Figure 2.14. An inverse of this waveform must be used with a p-channel device.

Operation is similar to two level charge pumping until the device is biased at $V_{step}$. Again, the interface traps are filled with minority carriers while the device is in inversion. When the device is pulsed to $V_{step}$, those traps which are above the Fermi level emit their minority carriers as long as $t_{step}$ is appropriately long. Now only those traps which are below the Fermi level are available to recombine with the majority carriers as they enter when the device is pulsed into accumulation. In this way, $I_{cp}$ is a measure of how many traps are filled with minority carriers when the device is biased at $V_{step}$. If $V_{step}$ is swept from midgap to the majority band edge then $I_{cp}$ can be related to bandgap energy through the surface potential as seen in eqn. (2.27).

\[
D_{it}(q\psi_{step}) = \frac{1}{qfA_{eff}} \frac{dI_{cp}}{d(q\psi_{step})}
\] (2.27)

where $f$ is the frequency of the pulse, $A_{eff}$ is the effective area of the gate and $\psi_{step}$ is surface band bending when the gate is at $V_{step}$. The frequency is the inverse of the total period, $T$, which is the sum of all the times at the various levels and the rise and fall times to pulse to those levels.
This method has the advantage that capture cross section coefficients do not have to be known to determine $D_{it}(E)$. However, geometric effects are still valid in three level charge pumping and the effects increase as $V_{step}$ is reduced. Care must be taken to minimize the geometric current arising from short $t_{step}$ or long channels. All three level charge pumping measurements performed in this thesis used devices and parameters which avoided geometrical effects as much as possible. Three level charge pumping has not yet been performed on metal gates with high k dielectrics until this work.

2.9.3 Pulsed Drain Current

Using pulsed techniques, the inversion charge can be determined and the effects of trap response can be minimized. The inversion charge on long channel devices is measured with the amplitude sweep charge pumping technique having fast rise and fall times on the order of 10 ns. This exploits the geometrical component usually seen as an undesirable effect of charge pumping. Under these settings, the long channel is rapidly pinched off and forces a large fraction of the inversion charge to recombine in the substrate. By understanding the symmetry of the channel, the full inversion charge can be calculated from the schematic seen in Figure 2.15 as

$$N_{inv} = 2N_{ICP-a} - N_{ICP-b}$$ \hspace{1cm} (2.28)

To determine the drain current, a resistor, $R_L$, is placed on the drain and biased with $V_{DD}$ as in an inverter circuit seen in Figure 2.16. The drain and gate
voltages are extracted using a digital oscilloscope as the gate is biased with a pulse. The current as a function of gate voltage can be extracted using

\[ I_D = \frac{V_{DD}}{V_d} \left( \frac{V_{DD} - V_D}{R_L} \right) \]  

(2.29)

Figure 2.17 shows the combination of eqn. (2.28) with eqn. (2.29) into the equation for effective mobility and the result of the lowered impact of charge trapping on a device with SiO₂ as the gate dielectric.

In this work, 10 ns rise and fall times could not be put into practice so another method was used to understand the effects of bulk trapping on mobility. That is to combine the pulsed drain current calculated from eqn. (2.29) with the Hauser NCSU Mob2D model²⁷. In this way, the extracted mobility is free from bulk traps and interfacial traps.

All the methods discussed in this chapter will be utilized to analyze the differences in dielectrics and metal gates on advanced MOSFET devices. Next, the processing methods of the fabricated devices will be explained.
2.10 Figures

Figure 2.1. Soft optical phonon limited mobility for various dielectric constant materials.

Figure 2.2. Plot of overall mobility with labeled regions limited by various scattering mechanisms.
Figure 2.3. Schematic of experimental setup to measure (a) gate to channel capacitance for Split CV, (b) gate to substrate capacitance, and (c) total gate capacitance.
Figure 2.4. Gate to channel capacitance as a function of gate bias for a device with 3 nm SiO$_2$ and TaN gate. The integral under this curve gives inversion charge, $Q_{\text{inv}}$.

Figure 2.5. Circuit model used by capacitance meter showing (a) parallel model and (b) corrected parallel model with series resistance, $R_s$. 
Figure 2.6. Resistances of MOSFET channel.

Figure 2.7. A decrease in mobility as a function of effective field is seen for ultra thin oxides after about 1.9 nm$^2$. 

$R_{c}$ $R_{sd}$ $R_{ch}$ $R_{sd}$
Figure 2.8. Effective mobility extracted using leakage correction involving the subtraction of two drain currents taken at different drain biases\textsuperscript{11}.

Figure 2.9. A decrease in mobility as a function of effective field is seen for ultra thin oxides even after $N_s$ correction\textsuperscript{22}. 
Figure 2.10. Schematic of experimental setup for square wave two level charge pumping.

Figure 2.11. Schematic of square wave pulse used in two level charge pumping.
Figure 2.12. Base level sweep charge pumping for 6 nm SiO₂.

Figure 2.13. Amplitude sweep charge pumping on 6 nm SiO₂.
Figure 2.14. Schematic of waveform used for three level charge pumping for an n-channel device.

Figure 2.15. Schematics of the Inversion-Charge Pumping (ICP) setup used to extract the full inversion charge in long channel MOSFETs using the device symmetry\textsuperscript{24}.
Figure 2.16. Schematic of the inverter circuit used to characterize the drain current through a pulsed technique.  

Figure 2.17. Effect of lowered charge trapping on mobility extraction when using the Inversion Charge Pumping technique on a device with 2.1 nm SiO$_2$. 
2.11 References


7. Soldini *Solid State Electronics* vol 25, no 9, p833, 1982


27. as discussed at IEEE Semiconductor Interface Specialists Conference (SISC) 2004.

Chapter 3 Experimental Description

This chapter will provide details on the different samples used in this study. Description of the process flow and deposition processes are included for the candidate metal gate electrodes on high k and SiO$_2$ standard MOSFETs as well as the aggressively scaled MOSFETs with metal gate and polysilicon gate electrodes.

3.1 Standard MOSFET Process Flow

The standard MOSFET process flow can be seen in Table 3.1. Since charge pumping analysis was to be performed, a new mask had to be designed with appropriate channel lengths and structures to determine other important parameters such as channel length modulation and contact resistances. This GEM mask set, designed in part by the author, and the devices used for these samples can be seen in Appendix B. The highest temperature seen by these devices was 950 °C for 10 seconds in order to activate the dopants. The metal gate electrodes were patterned by liftoff process and a forming gas anneal (FGA) at 400 °C for 30 minutes was performed as the last step.

3.1.1 Hafnium Oxide

Hafnium oxide (HfO$_2$) was chosen as the alternative dielectric due to its useful characteristics. It has a medium value dielectric constant ~25 and since it has a rather high heat of formation ($\Delta H_f = 271$ Kcal/mol) it forms a stable oxide on silicon$^1$. The bandgap is about 5.68 eV which provides a great enough barrier for tunneling
probability and it is resistive to impurity diffusion and intermixing at the interface due to its high density (9.68 g/cm$^3$). Finally, Hf can reduce the native SiO$_2$ and has been shown to be compatible with an n+ polysilicon gate electrode$^2$.

HfO$_2$ was deposited using a DC magnetron sputtering system located at the University of Texas at Austin, TX$^3$. A thin layer of Hf is deposited and annealed in vacuum which acts as an oxidation barrier for the subsequent HfO$_2$ deposition. Then the HfO$_2$ is reactively sputtered in Ar and O$_2$ ambient where the oxygen flow is regulated to control the interface quality and the growth of the interfacial layer. The target thickness of HfO$_2$ was 45 Å.

### 3.1.2 Metal Gate Deposition Process

Metal gate electrodes were deposited using an eight inch UHV RF magnetron sputtering system located at North Carolina State University. The basic schematic can be seen in Figure 3.1. A single wafer is loaded and the load lock chamber is subsequently pumped until a base pressure of $\sim$1x$10^{-6}$ Torr is reached. The wafer is then transferred into the deposition chamber and placed on a rotating sample holder. Then the main chamber is pumped until a base pressure of $10^{-8}$ Torr is reached for sputtering. The metal targets are typically pre-sputtered prior to deposition in order to remove the surface layer that may have oxidized/poisoned during prior deposition or system standby condition.

During processing, the Ar flow is 40 sccm and the main chamber valve is adjusted to stabilize the operating pressure at 5.5 mTorr. The RF power used for sputtering ranges from 50 to 100 Watts depending on the deposited material, its
alloy composition, and the desired deposition rate. No intentional heating is applied to the substrate during deposition. Typically, a tungsten cap is sputtered directly after the metal gate electrode deposition in order to decrease the probing resistance and provide an additional barrier layer for implantation damage to the underlying dielectric. The target thickness for the gate electrode is 400 Å while the target thickness for the W cap is 500 Å. A table of sputtering conditions for the metal gates used in the standard MOSFET process can be found in Table 3.2.

### 3.1.3 Tantalum Nitride

Tantalum nitride (TaN) is an attractive material for use as a gate electrode. Ta has a workfunction near the conduction band of silicon, \( \Phi_m \approx 4.1 \) eV, so it is appropriate for n-channel devices. Incorporating nitrogen in the film provides an excellent barrier to oxygen diffusion which can help deter any interfacial SiO\(_2\) growth. However, the inclusion of nitrogen does increase the workfunction\(^4\). The quantity of nitrogen incorporated into the film is related to the partial pressure of nitrogen within the sputtering ambient during the deposition. This is seen in Figure 3.2 where the nitrogen concentration in the film determined by Auger Electron Spectroscopy (AES) is plotted as a function of the N\(_2\) flow rate. The concentration of nitrogen increases as the flow rate increases until about 10% partial pressure of N\(_2\). At this point, the surface ends up engrossed in nitrogen and the concentration saturates.

In the standard MOSFET process flow, the TaN gate electrodes were deposited using reactive sputtering from a Ta target in 5% N\(_2\) in Ar which results in
40% nitrogen in the film. Following the sputtering of TaN at 100 Watts for 12 minutes in 2 sccm N\textsubscript{2}, a W cap is sputtered at 100 Watts for 12 minutes to (i) increase thickness of the metal gate electrode, (ii) provide a good probing contact to the underlying metal, and (iii) reduce the diffusion of oxygen into the underlying layers. TaN deposited using 5% N\textsubscript{2} flow rate has been shown to have a workfunction of 4.5 eV on SiO\textsubscript{2} devices extracted by both CV and barrier height analysis\textsuperscript{4}. It is also important to note that nitrogen can diffuse to the dielectric and possibly the dielectric – silicon interface at higher temperatures where it will result in threshold shifts as well as contribute to interface traps.

3.1.4 Ruthenium Tantalum Alloys

Ruthenium tantalum (RuTa) alloys are also appealing as an n-channel gate electrode. Ru has a workfunction near the valence band which is appropriate for p-channel devices. However, alloying the metal with Ta which, as seen in section 3.1.3, has a workfunction near the conduction band, can lead to a workfunction appropriate for n-channel devices. The alloy allows a suitable gate electrode without the instability of low workfunction metals. RuTa alloys can be deposited by co-sputtering both Ru and Ta targets simultaneously\textsuperscript{5, 6}. Varying the sputtering power of each gun can result in different concentrations of the alloy as seen in Figure 3.3.

Varying the composition in the alloy also varies the workfunction. This can be seen when the workfunction is plotted as a function of Ta sputtering power in Figure 3.4. The higher the sputtering power, the higher the Ta concentration and the lower the workfunction. The workfunctions range from 5.1 eV for a pure Ru metal gate to
4.2 eV for a pure Ta metal gate. In the standard MOSFET structures, both Ru$_{50}$Ta$_{50}$ and Ru$_{90}$Ta$_{10}$ were used as metal gate electrodes which have been found to have workfunctions very close to their pure Ru and pure Ta counterparts. Ru$_{50}$Ta$_{50}$ on SiO$_2$ has a workfunction, $\Phi_m$=5.1 eV suitable for p-channel devices and Ru$_{90}$Ta$_{10}$ on SiO$_2$ has a workfunction, $\Phi_m$=4.3 eV suitable for n-channel devices.

It has been shown that Ru based alloys impart more damage to the devices which may increase the trap density of the underlying dielectric$^6,7$. This is due to the energy transfer from the plasma of a co-sputtered deposition system. Ru$_{50}$Ta$_{50}$ and Ru$_{90}$Ta$_{10}$ have also shown to preserve their workfunction difference on HfO$_2$. In the standard MOSFET process described in this dissertation, Ru$_{50}$Ta$_{50}$ was formed by co-sputtering Ru and Ta at 50 Watts each for 12 minutes followed by sputtering a W cap at 100 Watts for 22 minutes and 20 seconds. The Ru$_{90}$Ta$_{10}$ gate electrode was deposited by co-sputtering Ru at 90 Watts and Ta at 10 Watts for 12 minutes followed by the same capping layer as the Ru$_{50}$Ta$_{50}$ alloy.

### 3.2 Aggressively scaled MOSFETs

MOSFETs with aggressive EOT (<20 Å) were fabricated at Sematech using ALD HfO$_2$ and either MOCVD TiN or standard polysilicon deposition as the gate electrode. A brief summary of the process flow can be seen in Table 3.3. Prior to HfO$_2$ deposition, the substrates received an HF (0.2 % HCl) sacrificial oxide etch followed by an ozonated clean in an overflow of ozonated water for 10 minutes at 23°C with a concentration of 20 ppm and an HCl concentration of 0.2 %. A chemical oxide was formed during the O$_3$ pre-gate cleaning which is about 0.8 nm.
A target thickness of 30 Å of HfO$_2$ was deposited using ALD using alternate cycles of HfCl$_4$ and H$_2$O and followed by a post deposition anneal in N$_2$ at 600 °C for 60 seconds. The effect of the electrode on mobility was examined through the parallel processing of gate stacks having both polysilicon and a thin TiN under a polysilicon gate electrodes. Polysilicon gate electrodes were formed by depositing 200 nm of undoped amorphous Si (a – Si) using a silane – based CVD furnace at 550ºC. The TiN and polysilicon gate electrodes were formed in a NH$_3$/TDEAT – based CVD process at using a 340 ºC for 20 seconds to deposit ~100 Å of TiN followed by a silane – based RTP CVD process at 620ºC for 120 seconds to deposit 1800 Å of undoped a – Si. After electrode formation, the polysilicon was doped by ion implantation and activated with an anneal at 1000ºC for 10 seconds in which the Si is converted to polysilicon. The EOT ranges from 10-15 Å$^9,10$.

### 3.3 Role of Interfacial Layer

In both sets of MOSFET structures, there is undoubtedly an interfacial layer which may be intentional or unintentional underneath the dielectric and resides directly in contact with the silicon substrate. The interfacial layer formation in the standard MOSFETs seem to be the result of the reactive sputtering process and appears to be a silicate$^3,11,12$. Work has gone into trying to minimize this interfacial layer so as to minimize EOT$^{13}$. More work has been done to nitridize the interface or the interfacial layer to suppress oxide growth and improve performance$^{14-16}$. With the transport from The University of Texas at Austin, TX, more interfacial growth or small changes in the dielectric could be possible.
In the aggressively scaled MOSFETs, the interfacial layer is a result of the \( \text{O}_3 \) cleaning process and whether intentional or unintentional, the final composition of the resulting interface is unclear. The 0.8 nm thickness seems questionable when the extracted EOT thickness is only 1 nm suggesting that the interfacial layer is not \( \text{SiO}_2 \). If the high processing temperatures lead to Hf diffusion and some silicate formation, the extracted EOT may make more sense.

It is important to note that all results are coming from stacked structures. The interfacial layer involvement in the electrical characteristics has to be included concurrently with the high k dielectric. More information on the role of the interfacial layer on MOSFETs and the impact of a \( \text{Si}_3\text{N}_4 \) interfacial layer on mobility will be presented in the next chapter. Furthermore, the requirement of a \( \text{SiO}_2 \) rich interfacial layer will be justified. Conclusions will be drawn on the requirements of this interfacial layer in high k devices.
### Table 3.1. Summary of standard MOSFET process flow.

<table>
<thead>
<tr>
<th>Step</th>
<th>Process Description</th>
<th>Other Process Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wafer scribing</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>RCA clean</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Field oxidation</td>
<td>1000 °C, 45 min, ~3200 Å</td>
</tr>
<tr>
<td>4</td>
<td>JTB clean</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Pre-coat bake</td>
<td>115 °C, 5 min</td>
</tr>
<tr>
<td>6</td>
<td>Photolithography – GEM Active Mask</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Descum</td>
<td>30 sec</td>
</tr>
<tr>
<td>8</td>
<td>Field oxide etch</td>
<td>10 sec over etch</td>
</tr>
<tr>
<td>9</td>
<td>Strip Resist</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>RCA clean</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Sacrificial oxide growth</td>
<td>100 Å</td>
</tr>
<tr>
<td>12</td>
<td>Sacrificial oxide etch</td>
<td>1% HF</td>
</tr>
<tr>
<td>13</td>
<td>Gate dielectric formation</td>
<td>Thermal oxide controls ~25 Å</td>
</tr>
<tr>
<td>14</td>
<td>Photolithography – GEM Gate Mask</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Descum</td>
<td>30 sec</td>
</tr>
<tr>
<td>16</td>
<td>Metal gate deposition</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Metal gate lift off</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Low temperature oxide deposition</td>
<td>~150 Å</td>
</tr>
<tr>
<td>19</td>
<td>Source and drain implantation</td>
<td>As, 5 x 10^{15} cm^{-2}, 20 keV</td>
</tr>
<tr>
<td>20</td>
<td>Rapid thermal anneal for S/D activation</td>
<td>950 °C, 10 sec</td>
</tr>
<tr>
<td>21</td>
<td>Low temperature oxide deposition</td>
<td>~3200 Å</td>
</tr>
<tr>
<td>22</td>
<td>Pre-coat bake</td>
<td>115 °C, 5 min</td>
</tr>
<tr>
<td>23</td>
<td>Photolithography – GEM Contact Mask</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Oxide contact hole etch</td>
<td>BOE, 5 min 30 sec</td>
</tr>
<tr>
<td>25</td>
<td>Strip resist</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Pre-coat bake</td>
<td>115 °C, 5 min</td>
</tr>
<tr>
<td>27</td>
<td>Photolithography – GEM Metal Mask</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Contact metal evaporation</td>
<td>Ti (50 nm)/Al (200 nm)</td>
</tr>
<tr>
<td>29</td>
<td>Contact metal lift off</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Backside oxide etch</td>
<td>BOE swab</td>
</tr>
<tr>
<td>31</td>
<td>Forming gas anneal</td>
<td>400 °C, 30 min</td>
</tr>
</tbody>
</table>
Table 3.2. Gate electrode sputtering conditions of standard MOSFET process

<table>
<thead>
<tr>
<th>Metal Gate</th>
<th>Metal Gate Conditions</th>
<th>Tungsten Cap Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaN</td>
<td>100 W, 2 sccm N₂, 12 min</td>
<td>100 W, 12 min</td>
</tr>
<tr>
<td>Ru₅₀Ta₅₀</td>
<td>50 W Ru, 50 W Ta, 12 min</td>
<td>100 W, 22 min, 20 sec</td>
</tr>
<tr>
<td>Ru₉₀Ta₁₀</td>
<td>90 W Ru, 10 W Ta, 12 min</td>
<td>100 W, 22 min, 20 sec</td>
</tr>
</tbody>
</table>

Table 3.3. Brief summary of aggressively scaled MOSFET process flow.

<table>
<thead>
<tr>
<th>Step</th>
<th>Process Description</th>
<th>Process Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HF last treatment</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>O₃ chemical oxide formation</td>
<td>~0.8 nm</td>
</tr>
<tr>
<td>3</td>
<td>ALD HfO₂ formation</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Post deposition anneal</td>
<td>N₂, 600 °C, 60 sec</td>
</tr>
<tr>
<td>5</td>
<td>Gate electrode formation</td>
<td>~100 Å MOCVD TiN and/or ~1800 Å polySi</td>
</tr>
<tr>
<td>6</td>
<td>S/D activation</td>
<td>N₂, 1000 °C, 10 sec</td>
</tr>
</tbody>
</table>
3.5 Figures

Figure 3.1 Schematic of UHV sputtering system

Figure 3.2. Plot of atomic ratio of nitrogen to tantalum versus the flow ratio of $N_2$ to $Ar$.

Figure 3.2. Plot of atomic ratio of nitrogen to tantalum versus the flow ratio of $N_2$ to $Ar$. 

Figure 3.3. Atomic percent of Ta and Ru in RuTa alloys as a function of sputtering power⁶.

Figure 3.4. Workfunction of RuTa alloys as a function of Ta sputtering power⁶.
3.6 References


Chapter 4 The Role of the Interfacial Layer

4.1 Introduction

While scaling requires the need for alternative materials, the implementation of new materials brings about increasing processing obstacles. Scaling to ultra thin effective oxide thicknesses (EOT) is hampered by the formation of an interfacial layer which results when oxygen diffuses to the silicon interface\(^1\). This interfacial layer is currently a crucial point in the implementation of high k dielectrics in that it can be exploited for mobility values. However, minimizing the thickness of the interfacial layer is needed. One solution to reducing the interfacial layer is to provide nitrogen at the interface either by annealing prior to deposition or depositing a silicon nitride (Si\(_3\)N\(_4\)) layer at that interface\(^2,3\). While Si\(_3\)N\(_4\) has only a dielectric constant of \(\sim 8\), it does have a wide bandgap which can also aid in suppressing the amount of tunneling across the interface. It is also beneficial in blocking the penetration of boron from the polysilicon gate electrode through to the channel. With the realization of Si\(_3\)N\(_4\) as an interfacial layer for high k dielectrics, it is important to understand how it impacts MOSFET mobility.

4.2 Silicon Nitride Experimentation

Many previous experiments showed that Si\(_3\)N\(_4\) was a viable option as a dielectric\(^4-6\). In particular, the RPECVD nitride films have proven to be of good quality in both NMOS and PMOS devices\(^7,8\). In this work, silicon nitride films were formed in a cluster tool so that the entire gatestack was formed \textit{in situ} and before
exposure to atmosphere. First, a 200 nm field oxide was grown on 100 mm substrates and patterned into active areas. A 10 nm sacrificial oxide was grown and removed with 2% HF solution immediately before insertion into the cluster tool. The Si$_3$N$_4$ layers were formed by RPECVD at 400 Watts, 300 mTorr, and 300 °C using silane (SiH$_4$) and either ammonia (NH$_3$) or nitrogen (N$_2$) in Ar plasma$^{6-8}$. Single layer nitrides with varying thicknesses of 1.5 to 3.0 nm were formed on both n-type and p-type silicon substrates. Stacked layers were also formed with an SiO$_2$ interface below the Si$_3$N$_4$ so that the properties of ultra thin nitride layers could be studied using conventional characterization techniques. To reduce the hydrogen content and promote formation of additional Si-N bonds, the films were annealed at 900 °C in vacuum after deposition$^6, 9$. Control oxides of varying thicknesses were also prepared by thermal oxidation for comparison. The devices were then transferred to another chamber under vacuum where polysilicon was deposited at 700 °C using rapid thermal processing. Surface channel NMOS and PMOS devices were formed by conventional techniques and a forming gas anneal was performed at 400 °C after metallization.

4.3 Material Analysis

Because hydrogen is unwanted in the film due to the hysteresis effects it causes, material analysis techniques were performed to determine the amount of elemental concentrations in the films. Secondary Ion Mass Spectroscopy (SIMS) of the RPECVD nitride films is shown in Figure 4.1 indicating a hydrogen content less than 15% throughout the film. The SIMS also shows that the oxygen concentration
is less than 3% with only a slight increase at the interface of the silicon substrate. This contrasts with other nitrides deposited via LPCVD techniques such as Jet Vapor Deposition (JVD) where the oxygen concentration is ~20% creating an SiO$_2$ rich nitride having a significant interfacial layer$^5$. Since there is some exposure to the cleanroom environment the surface may become contaminated before being loaded into the RPECVD and this could be the cause of the slight increase in oxygen concentration near the interface. From the physical thickness determined by Transmission Electron Spectroscopy (TEM) and the electrical thickness determined from the capacitance, the dielectric constant is around 7.3.

4.4 Electrical Measurements

Capacitance measurements are shown in Figure 4.2. The data was fit to the Hauser NCSU CVC least squares model extracting an effective oxide thickness (EOT) of 2.1 nm and 2.3 nm for the NMOS and PMOS devices, respectively$^{10}$. Not only is there a flatband voltage shift between the SiO$_2$ and Si$_3$N$_4$ gate dielectrics in the NMOS, which has been attributed to fixed charge$^{11-13}$ but the PMOS transistor shows even greater flatband shift and erroneous inversion characteristics. The PMOS device behaves as if no inversion layer has formed. Additionally, the PMOS capacitance versus voltage is plotted for varying thicknesses of Si$_3$N$_4$ interface covered by 5 nm of SiO$_2$ to gauge the impact of decreasing Si$_3$N$_4$ interfacial layer. As shown in Figure 4.3, even a nitride thickness of > 5 Å can drastically influence the inversion characteristics for PMOS devices. From the gate leakage current data shown in Figure 4.4 as a function the voltage across the dielectric that even despite
a physically thicker dielectric, the leakage current in the negative voltage regime is greater for Si$_3$N$_4$ for both NMOS and PMOS devices.

Furthermore, plotting the gate transconductance times the EOT against the gate voltage, a value proportional to channel mobility, the Si$_3$N$_4$ NMOS device shows a significantly lower peak mobility but enhanced high field mobility over that of SiO$_2$. This is similar to what has been published on oxynitrides$^{11, 12}$. However, the Si$_3$N$_4$ PMOSFET shows negligible mobility values. The extremely low PMOS mobility, lack of inversion capacitance, and increased gate leakage at negative voltages for Si$_3$N$_4$ indicate that the hole behavior is considerably degraded.

In order to understand these characteristics, a technique called carrier separation$^{14}$ was applied to isolate currents due to electrons and holes. A transistor is biased in inversion and the source and drain are tied together. The gate bias creates a channel and as carriers tunnel out of the channel into the gate dielectric they are replaced by carriers from the source/drain regions. Measurement of gate, channel and substrate current provides a measure of the electron and hole current components. A schematic of the measurement setup is given in Figure 4.6. In a normal transistor, the channel current should equal the gate current until large gate voltages when the substrate current becomes equal to the gate current as seen for the SiO$_2$ dielectrics in Figure 4.7.

The Si$_3$N$_4$ NMOS devices show similar behavior to that of SiO$_2$ as seen in the right side of Figure 4.8. However, for the Si$_3$N$_4$ PMOS devices seen on the other side, the channel current starts out very low and unequal to the gate current, indicating that the channel holes are not participating in carrier conduction in this
voltage regime. Also, the substrate current is very large and equal to the gate current which means the electron current is the dominating process. After the gate voltage reaches ~-1.5 V the channel current starts to increase to gate current values.

4.5 Si$_3$N$_4$ Donor Traps

To explain these results, it is assumed the that Si$_3$N$_4$ possesses a large density of donor type traps located in the bandgap near the valence band of silicon. An illustration of this can be found in Figure 4.9. The traps are neutral and occupied when the Fermi level is above them as in the case of the NMOS devices. When the Fermi level starts to move toward the valence band, as in the case of PMOS inversion, the traps empty and become positively charged. The high density of charges can screen the gate charge until all the traps become unpopulated. This occurs at about -1.5V which is the threshold voltage extracted from the drain current versus gate voltage characteristics. At this voltage, the surface potential is such that the onset of inversion occurs in the channel.

The screening of the gate charge from the high density of traps below midgap prevents a voltage from dropping in the semiconductor. This screening also explains the elevated leakage current observed in the Si$_3$N$_4$ devices at negative voltages. A voltage across SiO$_2$ dielectric, $V_{ox}$, is

$$V_{ox} = V_{app} - V_{sub}$$

(4.1)

where $V_{app}$ and $V_{sub}$ are the applied and substrate voltages, respectively. But the voltage across the Si$_3$N$_4$ dielectric, $V_{nitr}$, is
\[ V_{nitr} \approx V_{app} \]  

(4.2)

since the substrate potential is not allowed to form because of the donor traps screening. \( V_{nitr} \) is greater than \( V_{ox} \) and thus a larger field is placed across the nitride compared to that of SiO\(_2\) with a similar EOT. Therefore, larger current is observed in the negative voltage regime for Si\(_3\)N\(_4\).

A possible physio-chemical explanation for the Si – Si\(_3\)N\(_4\) interface defect formation may lie in the bonding constraints at the Si – Si\(_3\)N\(_4\) interface. It has been reported that bonding constraint theory calculations can relate defect formation to bonding coordination at the interface of two materials\(^{12}\). This is shown in Figure 4.10 where the defect density was calculated from the over-coordination of the interface for PMOS devices. Bonding constraints arise due to differences between the average number of bonds per atom on each side of the interface\(^{13}\). The calculations for various Si based dielectrics has been performed and reported. It was found that when the average coordination was < 3.0, good quality interfaces were obtained as indicated by the Si – SiO\(_2\) interface, which has an average coordination of 2.8. However, when the average coordination was > 3, defective interfaces are obtained as indicated by the Si – Si\(_3\)N\(_4\) interface, which has an average coordination of 3.5. It is assumed that the over coordination of the Si\(_3\)N\(_4\) – Si system compared with that of the SiO\(_2\) – Si system results in a trap formation which negatively impacts the electrical properties of the MOSFET devices.
4.6 The Role of the Interfacial Layer

A thin layer of SiO$_2$ was interposed between the Si$_3$N$_4$ and the Si substrate as an interfacial layer. The thickness of the monolayer was $\sim 6$ Å as measured *in situ* by Auger Electron Spectroscopy (AES). The average coordination per atom for this system is calculated to be 3, indicating a low defect density forming interface. The transconductance values for both the NMOS and PMOS devices are plotted in Figure 4.11. With the thin interfacial layer both the NMOS and PMOS devices show complete recovery of the mobility values. This indicates that 6 Å of SiO$_2$ is enough to reduce the interface state density near the conduction and valence band edges. Stated another way, the presence of oxygen assists in reducing the defect density by reducing the average bonding coordination as systematic experiments have shown$^{15}$. In the world of higher k dielectrics, this interfacial layer will provide a crucial role in reducing interface state density and increasing mobility and will be discussed in subsequent chapters.
4.7 Figures

Figure 4.1. SIMS showing hydrogen and oxygen concentration of RPECVD silicon nitride.

Figure 4.2. Capacitance versus voltage for SiO$_2$ and Si$_3$N$_4$ NMOS and PMOS devices.
Figure 4.3. Capacitance as a function of voltage for varying thickness of Si$_3$N$_4$.

Figure 4.4. Gate leakage current as a function of field across the SiO$_2$ or Si$_3$N$_4$ for NMOS and PMOS devices.
Figure 4.5. Transconductance times EOT versus effective field across the dielectric for NMOS and PMOS Si₃N₄ and SiO₂ devices in proportionality to channel mobility.

Figure 4.6. Schematic of Carrier Separation technique in NMOS device for isolating currents due to holes from electrons.
Figure 4.7. Carrier separation for SiO₂ NMOS and PMOS devices showing typical behavior.

Figure 4.8. Carrier separation for Si₃N₄ NMOS and PMOS devices.
Figure 4.9. Illustration of donor type traps near the valence band of Si$_3$N$_4$ on Si.

Figure 4.10. Normalized defect density as a function of the over-coordination of the interface for PMOS devices.\(^{16}\)
Figure 4.11. Transconductance values for 6 Å SiO$_2$ interfacial layer under Si$_3$N$_4$. 
4.8 References


Chapter 5 Mobility of Candidate Metal Gates on SiO$_2$ and High k Dielectrics

5.1 Introduction

Concurrent with the study of various alternative dielectrics for aggressively scaled devices, metal gate electrodes also need to be highly investigated in order to meet the needs of the International Technology Roadmap for Semiconductors$^1$. As discussed in Chapter 1, when scaling progresses, the polysilicon gate depletion is becoming increasingly problematic, not only adversely affecting the EOT but also possibly degrading mobility through remote Coulomb scattering$^{2-5}$. In contrast, metal gate electrodes have high carrier concentrations and show no signs of depletion. It is also suggested that metal gate electrodes can screen soft optical phonons present in high k dielectrics, effectively increasing the mobility of these devices over that of polysilicon$^3$. As the search for the perfect high k dielectric is pursued, so is that of the perfect metal gate. It is important in this pursuit to examine the differences, if they exist, between different metal gates in order to justify their use and qualify their performance. This chapter provides a study of the mobility of emerging metal gate electrodes on both SiO$_2$ and HfO$_2$ with emphasis on any differences between a metal on either dielectric or differences between the metal gate electrodes, themselves.
5.2 Basic Electrical Results

Extraction of MOSFET characteristics using capacitance – voltage (CV) and current – voltage (IV) analysis were performed on n-channel devices with either SiO$_2$ or HfO$_2$ gate dielectrics and metal gate electrodes. The devices were fabricated per conditions outlined in Chapter 3. An HP 4284a LCR Meter was used for capacitance and conductance measurements at a constant frequency or voltage and an HP 4155b Semiconductor Parameter Analyzer was used to measure the DC current as a function of voltage. The Hauser NCSU CVC program was used to determine parameters such as flatband voltage ($V_{fb}$), effective oxide thickness (EOT), and threshold voltage ($V_t$).^6

From the CV plots of all three metal gates on SiO$_2$ and HfO$_2$ shown in Figure 5.1, it can be seen that the devices show good device characteristics. The slight variation in minimum capacitance is due to the different device area measured for Ru$_{50}$Ta$_{50}$ on SiO$_2$. From the maximum capacitance in inversion, $C_{inv}$, and accumulation, $C_{acc}$, for each metal, it can be seen that these devices show no sign of gate depletion. Also, there is an observable thickness difference for each sample. Device areas of 50 $\mu$m x 50 $\mu$m (or 50 $\mu$m x 10 $\mu$m in the Ru$_{50}$Ta$_{50}$ on SiO$_2$ case) were chosen for mobility extraction to avoid the effects of overlap and series resistance. Frequencies of 100 kHz were chosen so that the AC $D_{ii}$ response was minimized and the LCR meter was not limited by inaccuracies. Extracted values for EOT, $V_{fb}$, and $V_t$ are shown in Table 5.1. The devices show similar EOT and $V_t$ values on their high k counterparts as on SiO$_2$. As was previously reported, the shifts in flatband between the HfO$_2$ and SiO$_2$ are similar for each metal gate.
However, the observed differences are less than those observed on MOS capacitors which were also subjected to similar processing temperatures but without a tungsten capping layer\(^7\). This could be due to some diffusion from the capping layer since there is evidence of W diffusion into the underlying metal gate and dielectric at 900\(^\circ\)C in Figure 5.2 where Secondary Ion Mass Spectrometry (SIMS) was performed from the topside of the gatestack. This intermixing of W along with some Hf diffusion into the metal gate has also been shown in SIMS performed from the backside. Since W has a higher workfunction, the intermixing will increase the workfunction of the underlying metal gate electrode toward midgap causing the extracted flatband voltages to be shifted from their expected values.

**Table 5.1. Extracted values for metal gate electrodes on SiO\(_2\) and HfO\(_2\).**

<table>
<thead>
<tr>
<th></th>
<th>SiO(_2)</th>
<th>HfO(_2)</th>
<th>SiO(_2)</th>
<th>HfO(_2)</th>
<th>SiO(_2)</th>
<th>HfO(_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOT [Å]</td>
<td>26.3</td>
<td>22.2</td>
<td>23.8</td>
<td>28.4</td>
<td>29.4</td>
<td>26.7</td>
</tr>
<tr>
<td>(V_{fb}) [V]</td>
<td>-0.51</td>
<td>-0.37</td>
<td>-0.67</td>
<td>-0.46</td>
<td>-0.48</td>
<td>-0.31</td>
</tr>
<tr>
<td>(V_t) [V]</td>
<td>1.07</td>
<td>1.07</td>
<td>1.11</td>
<td>1.10</td>
<td>1.24</td>
<td>1.22</td>
</tr>
</tbody>
</table>

The drain currents of these devices can be seen in Figure 5.3 and illustrate excellent working devices. The inset of Figure 5.3 (a) shows Ru\(_{50}\)Ta\(_{50}\) on SiO\(_2\) on a separate scale due to the difference in area. Finally, gate leakage currents as a function of effective dielectric field can be seen in Figure 5.4 for metal gates on HfO\(_2\) along with a direct tunneling model of leakage current in accumulation. The leakage currents are similar for all metal gates. The orders of magnitude difference in accumulation leakage current from an equal SiO\(_2\) EOT direct tunneling model
indicates that the dielectric is physically thicker and therefore has a higher dielectric constant.

5.3 Advanced Electrical Results

Interfacial trap densities and bulk trapping characteristics can be determined utilizing the characterization techniques outlined in Chapter 2. This section will investigate the similarities and differences in trapping behavior between the different metal gate electrodes on both SiO$_2$ and HfO$_2$.

5.3.1 Interface States

It is assumed that interfacial states are present in these devices and are caused either by (i) the inherent nature of the HfO$_2$ – Si system, (ii) the sputtering damage from the deposition of the dielectric, (iii) the sputtering damage from deposition of the metal gate, or (iv) diffusion of the metal gate to the interface. Therefore, both SiO$_2$ and HfO$_2$ devices may be affected by interfacial trapping. To ascertain the $D_{it}$ values, the two level charge pumping technique described in Chapter 2 was performed on MOSFETs with areas of 25 $\mu$m x 5 $\mu$m. An HP 8812a pulse generator was used to create the square wave and a Keithley 4200 was used to measure the substrate current. Typical base sweep charge pumping current ($I_{cp}$) versus base level bias ($V_{gb}$) are shown in Figure 5.5 for TaN on HfO$_2$ noting that as frequency increases the $I_{cp}$ also increases as expected. Amplitude sweep charge pumping was also performed and the extracted interface charge ($N_{it}$) versus amplitude bias can be seen in Figure 5.6 for all the metal gate electrodes on HfO$_2$. 
The large deviation in the Ru$_{50}$Ta$_{50}$ sample below an amplitude of 1.5 due to the larger gate leakage found in this sample. This value was subtracted from the entire curve which shifts the Ru$_{50}$Ta$_{50}$ sample to similar values as the other metal gates.

Important constants used to extract the average $D_{it}$ are listed below. Capture cross section coefficients ($\sqrt{\sigma_n\sigma_p}$) were recently published for similar HfO$_2$ devices by performing charge pumping measurements as a function of frequency and determining $\sigma_N$ and $\sigma_p$ individually by varying the rise and fall times$^8$. The reported average capture cross sectional coefficients for HfO$_2$ were more than an order of magnitude higher than SiO$_2$ and $\sigma_n$ was found to be $2.0 \times 10^{-13}$ which means that asymmetric trapping exists and that electrons are more likely to be captured above the middle of the bandgap.

Table 5.2. Constants used to extract average $D_{it}$ from two level base sweep charge pumping.

\[
\begin{align*}
\text{SiO}_2 \sqrt{\sigma_n\sigma_p} &= 5 \times 10^{-16} \text{ cm}^2 \\
\text{HfO}_2 \sqrt{\sigma_n\sigma_p}^9 &= 9.4 \times 10^{-15} \text{ cm}^2 \\
V_a &= 2.0 \text{ V} \\
V_{th} &= 1 \times 10^{7} \text{ V} \\
n_i &= 1.5 \times 10^{10} \text{ cm}^{-3} \\
t_r, t_f &= 20 \text{ ns}
\end{align*}
\]

The $D_{it}$ extracted from two level charge pumping can be seen in Figure 5.7. These values were similar for all the frequencies measured. All $D_{it}$ values are greater than $10^{11}$ eV$^{-1}$ cm$^{-2}$ however, it was also observed that the $D_{it}$ values were
higher for SiO$_2$ than for HfO$_2$. This is attributed to the fact that SiO$_2$ is physically thinner than HfO$_2$ and the damage caused by sputtering the gate electrode can penetrate further into the dielectric causing more damage to the interface. It is also important to note that the $D_{it}$ values are higher for the RuTa alloys than for TaN, especially on SiO$_2$ where $D_{it}$ values are greater by almost a magnitude. As previously stated, RuTa alloys are co-sputtered and will cause greater damage than that of the reactively sputtered TaN. This is most likely the reason for the larger $D_{it}$ values.

5.3.2 Pulsed Drain Current

Chapter 2 introduced a technique to determine if any bulk trapping exists in high k devices which may degrade mobility\textsuperscript{10}. By biasing the gate with a square pulse having a fast rise time where bulk traps can not respond and measuring the drain current of the full pulse, via a resistor, a shift in $V_t$ results if bulk trapping is present. The pulsed drain current measurement was applied to these samples to determine if any threshold voltage instability exists due to bulk trapping. Pulse widths of 100 $\mu$s and rise and fall times of 10 $\mu$s were used with a resistor value of 300 $\Omega$ and $V_{dd}$ of 100 mV. The short rise time was chosen so that negligible tunneling occurred through the unintentional interfacial layer. These setup parameters allowed for about a 50 mV drop across the resistor.

Results from the extracted drain current can be seen in Figure 5.8 for SiO$_2$ and Figure 5.9 for HfO$_2$. While bulk trapping has been reported in high k dielectrics, it is not expected in SiO$_2$\textsuperscript{10}. However, both HfO$_2$ and SiO$_2$ dielectrics exhibit bulk
trapping in a $V_t$ shift. This can be explained by the physical nature of the sputtered metal gate deposition process. Since the SiO$_2$ samples are physically thinner than the high k samples, more damage from the metal gate deposition is imparted to the dielectric and thus bulk traps are created. In the HfO$_2$ samples, all metal gates show very similar bulk trapping characteristics. The HfO$_2$ is the most probable origin of bulk trapping in these devices, however, sputtering damage associated with the deposition of the high k and the metal gate electrode could also be playing a role in promoting this behavior.

### 5.4 Three Level Charge Pumping

Since mobility is known to decrease with increasing interface state density, knowing the $D_{it}$ can aid in determining their impact. However, average numbers do not relay enough information about the interface and can overestimate $D_{it}$. Also, since the effect of interface traps is greatest near threshold where the amount of trapped charge is a large fraction of the total integrated charge, it is helpful to know the $D_{it}$ as a function of the bandgap$^{11}$. For the first time three level charge pumping is performed on three different candidate metal gate electrodes with high k dielectrics and their results are compared.

The parameters used for the three level charge pumping measurement can be seen in Table 5.3. An HP 8816A function generator was used to create the three level waveform and a Keithley 6517 electrometer was used to measure the substrate current. The device areas were 25 μm x 5 μm and $D_{it}$ values were obtained using the three level charge pumping technique description in Chapter 2. Results from
extraction of $D_{it}$ can be seen in Figure 5.10 for metal gates on HfO$_2$. All the metal gates show similar values and standard increase in $D_{it}$ towards the bandgap edge. The TaN does show a plateau in $D_{it}$ value near the conduction band edge whereas the RuTa alloys increase to higher values. Also, there is indication of trap localization for all the metal gates on high k which is consistent with previously reported results$^{8,12}$. Despite these slight differences, these metal gates on HfO$_2$ show the same interface trapping characteristics.

**Table 5.3. Parameters used in three level charge pumping to determine volume trap density of metal gate electrodes on HfO$_2$**

\[
\begin{align*}
  t_{step} & = 740 \, \mu s \\
  t_n, t_f & = 2 \, \mu s \\
  t_{th}, t_l & = 0.1 \, \mu s \\
  f & = 1343 \, Hz 
\end{align*}
\]

5.5 Mobility

It has been reported that metal gates on SiO$_2$ have not shown large differences in mobility when compared with polysilicon$^{13}$. Figure 5.11 and Figure 5.12 show the mobility values for Ru, RuTa, and polysilicon gate electrodes on 30 – 32 Å EOT SiO$_2$ for both NMOS and PMOS devices extracted by the Hauser NCSU Mob2D program$^{14}$. Hole mobilities for the Ru PMOS devices demonstrated values very close to polysilicon gates and the universal mobility. RuTa NMOS devices did show some low field mobility degradation when compared with polysilicon which was attributed to the large $D_{it}$ found in these devices.
Figure 5.13 compares the $D_{it}$ values of SiO$_2$ with both Ru and RuTa to that of polysilicon extracted by conductance method (G-ω) showing the large values found in the RuTa samples. It was assumed that the damage from co-sputtering both Ru and Ta, which resulted in a higher sputtering power density, was the reason for the increased trap density and hence the lowered mobility values. While these EOTs are thicker than the MOSFETs studied in this chapter the mobility values indicated that metal gates on SiO$_2$ showed lower mobility than their polysilicon counterparts. The rest of this chapter will investigate further the effects of metal gates on both SiO$_2$ and HfO$_2$ dielectrics.

In order to investigate further any degradation in mobility due to other metal gates on high k dielectrics and to determine the role of $D_{it}$, mobility values for the two different RuTa alloys and the TaN gate electrode MOSFETs were extracted using Split CV analysis described in Chapter 2. A correction for leakage was performed assuming that 50% of the gate leakage moves to the drain which is a valid assumption given the symmetric processing conditions of these samples. Figure 5.14 shows the extracted mobilities for TaN, Ru$_{50}$Ta$_{50}$, and Ru$_{90}$Ta$_{10}$ on SiO$_2$. All metal gates are well behaved and suffer from only a slight degradation of mobility compared to the universal mobility curve. Since the degradation is similar for all three metal gate electrodes, it is safe to assume that it is sputtering related damage associated with the metal gate deposition. Nevertheless, these results prove that metals are viable as gate electrodes in MOSFET devices and that mobility values do not suffer with SiO$_2$ as the gate dielectric.
Next, the mobility of these metal gates on HfO$_2$ was extracted. The results of Split CV mobility extraction on HfO$_2$ can be seen in Figure 5.15 as a function of effective electric field. The effective field plot shows the comparison of these mobilities with the universal mobility curve. These mobilities were found to be very similar to each other and only slightly degraded from the universal mobility curve. The TaN shows the lowest mobility values of the three metal gate electrodes which may be attributed to the fact that TaN possesses the thinnest EOT. A plot of all metal gates on SiO$_2$ and HfO$_2$ is shown in Figure 5.16 as a function of inversion charge so the electric field shifts do not detract from the mobility values. From this figure, the mobility values of all metal gates on SiO$_2$ and HfO$_2$ behave similarly.

5.6 Mobility at Higher Temperatures

As discussed in Chapter 2, there are several mechanisms which can scatter the channel mobility. The temperature dependence varies according to the individual mechanism. Coulombic scattered mobility increases with increasing temperature while phonon scattering is inversely proportional to temperature and surface scattered mobility is independent of temperature. It is from these effects that phonon scattered mobility dominates at higher temperatures even at lower fields and the effective channel mobility decreases when temperature is increased. In order to understand the response of the degradation mechanisms, Split CV analysis was performed at higher temperatures. Temperature dependent mobility can also help gauge how these devices will perform under extreme operating conditions.
All measurements necessary for Split CV mobility extraction were performed at room temperature (20 °C), 100 °C and 200°C. The leakage current for the SiO\(_2\) devices at 200°C became too large for analysis of the mobility and thus only the mobilities at 20°C and 100°C are included. The results can be seen in Figure 5.17 for SiO\(_2\) and in Figure 5.18 for HfO\(_2\). All metal gates show mobility degradation as temperature is increased as expected from phonon scattering dependent mobility. However, the decrease in mobility with temperature is greater for the SiO\(_2\) case. If the peak mobility is plotted as a function of temperature as seen in Figure 5.19, the slopes of metal gates on SiO\(_2\) is larger than that of HfO\(_2\). It is important to note, however, that because there are only two points for the SiO\(_2\) case, there is more error in determining the slope. The higher slope indicates that the high k mobility, while showing similar trends as that of SiO\(_2\), does so at a slower rate. This is similar to results seen on HfO\(_2\) gate dielectrics with Al gates\(^{15}\). It is ascertained that the decreased dependence on temperature of HfO\(_2\) mobility compared with that of SiO\(_2\) is attributed to the low phonon energy of optical phonons in the bulk of the HfO\(_2\). Since this energy is easily excited at these temperatures, the scattering rate from the optical phonons has a weak dependence on temperature and therefore the mobility component limited by these soft optical phonons also has a weak dependence on temperature. The weak dependence could also be attributed to a more dominant mechanism such as scattering due to interface traps or surface roughness.
5.7 Scattering Mechanisms

Results in this chapter have shown that the mobilities of metal gates both on SiO$_2$ and on HfO$_2$ show comparable values to that of the universal curve. The slight degradation can be attributed to the damaged encountered from the physical deposition of the metal gate and is noticeable in the obvious values of $D_{it}$. Even with the resulting $D_{it}$ values, however, the mobilities do not show the decreased values expected. Therefore, corrections for $D_{it}$ and bulk trapping would not indicate significant improvement, especially since the effect of substrate dopants and the coulombic scattering from their ionization plays a dominant role.

It has been proposed that the coulombic scattered mobility follows an inverse relationship to that of a linear combination of doping density and interface charge as seen in eqn. (5.1)$_{16}$.

$$\mu_{coul} = \frac{\mu_o}{\alpha N_{it} + L_{th} N_A}$$  \hspace{1cm} (5.1)

where $N_A$ and $N_{it}$ are the doping density and interface density, respectively, $L_{th}$ is related to the screening length, $\alpha$ is a fitting parameter, and $\mu_o$ is the unscreened mobility. Using the parameter values determined in [16] and plotting the coulombic mobility component as a function of inversion charge for various doping densities and interface charge, it can be illustrated that as the doping density approaches $10^{18}$ cm$^{-3}$, the effect of $10^{11}$ cm$^{-2}$ interface charge does very little to reduce mobility$^{16, 17}$. This is shown in Figure 5.20. It is possible that because the doping level is high in the MOSFETs for this study that the mobility values are limited solely by coulombic scattering.
Taking that into account, however, the mobility values are still higher than what has been previously reported for high k devices. There are several reasons which for this unexpected mobility value. The first is that an unintentional interfacial layer grew during high temperature processing. This is evident from the Transmission Electron Spectroscopy (TEM) of the metal gate electrodes on HfO$_2$ in Figure 5.21. Notice that the interfacial layer for the TaN gate is thinner than the RuTa alloys, however, the interfacial layers for all the metal gates in this study may be thick enough to provide mobility recovery as seen previously with Si$_3$N$_4$ in Chapter 4. Not all of the interfacial layer is expected to be SiO$_2$ since extracted thicknesses from the TEM range from 25 to 30 Å which are greater than the EOT values. This suggests that the interfacial layer is a medium k dielectric such as a silicate (HfSiO$_4$). This could account for the discernable $D_{it}$ values and the mobility values typically found with these dielectrics.

Another possibility is that the metal gate electrode is screening the soft optical phonons associated with the high k dielectric as proposed in [3]. Previous reports have theorized that metal gates can provide higher mobility$^{18}$ due to the screening of surface phonon scattering from coupling with the inversion channel. If the interfacial layer is a silicate or some dielectric with lower k value, the effect of optical phonons in the high k bulk may be less since they are removed from the interface. It has also suggested that the phonon component of mobility is less in HfSiO$_4$ than that of HfO$_2$.$^{3}$

Whether or not metal screening is occurring can only be ascertained by comparing the mobility of a metal gated MOSFET with that of a polysilicon gate on
ultra thin EOTs since the polysilicon electrode should not be able to screen any optical phonon coupling with inversion charge due to it low carrier concentration and hence large screening length. Given the circumstances in this chapter, metal gate electrodes do not impact mobility values either on SiO$_2$ or on HfO$_2$ and are a viable alternative to polysilicon. It is unclear from this study, however, if phonon screening is occurring or if these mobility values will persist on thinner dielectrics. This will be discussed in the following chapter on thinner dielectrics.
Figure 5.1. CV curves for metal gate electrodes on SiO$_2$ and HfO$_2$. The solid symbols represent HfO$_2$ and the open symbols represent SiO$_2$ gate dielectrics.

Figure 5.2. SIMS analysis of W/Ru/HfO$_2$ gatestacks from the topside at 900°C and compared with as deposited.
Figure 5.3. Drain currents of metal gate electrodes on (a) SiO$_2$ and (b) HfO$_2$. 
Figure 5.4. Leakage currents versus accumulation and inversion fields for metal gates on HfO$_2$.

Figure 5.5. Charge pumping current versus base sweep voltage for TaN on HfO$_2$. 
Figure 5.6. Interface charge ($N_{it}$) as a function of amplitude bias for metal gate electrodes on HfO$_2$.

Figure 5.7. $D_{it}$ measured from two level charge pumping at 100 kHz for metal gates on SiO$_2$ and HfO$_2$. 
Figure 5.8. Pulsed drain current measurements for metal gates on SiO$_2$.

Figure 5.9. Pulsed drain current measurements for metal gates on HfO$_2$. 
Figure 5.10. $D_I$ as a function of bandgap for metal gate electrodes on HfO$_2$ extracted from three level charge pumping.

Figure 5.11. Hole mobility for Ru and polysilicon MOSFETs with SiO$_2$ gate dielectric.
Figure 5.12. Electron mobility for RuTa and polysilicon MOSFETs with SiO$_2$ gate dielectric.

Figure 5.13. Interface trap density for Ru, RuTa and polysilicon MOSFETs as measured by the AC conductance method.
Figure 5.14. Mobilities of metal gates on SiO$_2$ dielectrics extracted by Split CV analysis.

Figure 5.15. Mobilities of metal gates on HfO$_2$ dielectrics extracted by Split CV analysis.
Figure 5.16. Mobility of metal gates on SiO$_2$ and HfO$_2$ as a function of inversion charge.

Figure 5.17. Mobility of metal gates on SiO$_2$ as a function at 20°C and 100°C.
Figure 5.18. Mobilities for metal gate electrodes on HfO$_2$ as a function of increasing temperature.

Figure 5.19. Peak mobility as a function of temperature for metal gates on SiO$_2$ and HfO$_2$ dielectrics.
Figure 5.20. Coulombic scattered mobility component as a function of inversion charge for various doping and interface charge densities using the values in [16].
Figure 5.21. TEM of metal gate electrodes on HfO$_2$. 
5.9 References


Chapter 6  Mobility of Aggressively Scaled High k MOSFETs

6.1  Introduction

In Chapter 5, the mobility values of various metal gates were only slightly degraded from the universal mobility and did not vary with metal type. The high values reported suggest that the resulting EOT thickness and thick interfacial layer of those samples dominate the mobility behavior. Now the focus is turned to the differences between metal gate and polysilicon gate electrodes on aggressively scaled (<20 Å EOT) MOSFETs with high k dielectrics. It has been reported that mobility decreases as SiO$_2$ thickness decreases for polysilicon gated MOSFET devices as discussed in Chapter 1. The cause for this decrease is still under discussion but the most widely held theories are that coulombic scattering or remote phonon scattering from the gate electrode causes mobility degradation with EOT < 19 Å.

Ultrathin high k dielectrics with metal gate electrodes have been studied previously and resulted in low mobilities. MOSFETs were fabricated with a TaSiN/HfSiON/HfO$_2$ stack resulting in 16 Å EOT$^1$ and mobilities extracted using the Hauser NCSU mob2D model$^2$ are shown in Figure 6.1. These devices show severely degraded mobility which was attributed to interfacial charge and surface roughness from the physical nature of the deposition. While these samples had a different dielectric and gate electrode and were deposited using well optimized sputtering conditions, the final mobility results confirm other reports of thin high k dielectrics. Therefore, it is important to understand how processing conditions,
dielectric properties and metal gate electrodes impact mobility by making a direct comparison between thin high k with metal and polysilicon gate electrodes since a complete study of why degradation is occurring for metal gate electrodes on high k dielectrics is lacking. This chapter will investigate any differences between metal gate and polysilicon gate electrodes on thin HfO$_2$ using electrical and analytical results.

### 6.2 Electrical Characteristics

The aggressively scaled MOSFETs described in Chapter 3 may assist in deciding the true nature of high k dielectric mobility degradation and whether metal gate electrodes can alleviate that degradation. CV curves comparing the TiN and polysilicon gate electrodes on ~30 Å of physical HfO$_2$ are shown in Figure 6.2. As with the samples discussed in the previous chapter, large 20 µm x 20 µm devices were measured at 100 kHz to minimize short channel effects and equipment inaccuracies. From these plots, it can be seen that the TiN shows a lower EOT and slightly less negative flatband voltage ($V_{fb}$). The TiN does not demonstrate any of the gate depletion seen in the polysilicon sample. However, the TiN sample shows signs of leakage illustrated by the increase in capacitance at higher inversion biases due to the thinner EOT.
Table 6.1. Extracted values of ultrathin high k devices with metal and polysilicon gate electrodes.

<table>
<thead>
<tr>
<th></th>
<th>TiN/Poly</th>
<th>Poly</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOT [Å]</td>
<td>11.8</td>
<td>15.6</td>
</tr>
<tr>
<td>$V_{fb}$ [V]</td>
<td>-0.77</td>
<td>-0.83</td>
</tr>
<tr>
<td>$V_t$ [V]</td>
<td>0.53</td>
<td>0.55</td>
</tr>
</tbody>
</table>

The Hauser NCSU CVC extracted values for EOT, $V_{fb}$ and $V_t$ can be seen in Table 6.1 and clearly show lower EOT values and less flatband shift for the metal gate. Speculation of why lower EOTs are seen in TiN with the same growth conditions as the polysilicon probably derives from the fact that nitrogen is a very good barrier to oxygen diffusion and may inhibit any interfacial growth during high temperature processing. There has also been evidence of a scavenging effect exhibited by HfO$_2$ layers on SiO$_2$\textsuperscript{3-5}. If the high k dielectric is annealed at low O$_2$ partial pressures and has oxygen vacancies, then the SiO$_2$ can decompose and fill those vacancies, reducing the thickness of the SiO$_2$. Still further evidence suggests that reactive metals, such as Ti and Ta, can dissolved oxygen ions which diffuse across the high k dielectric\textsuperscript{6} reducing or removing the interfacial layer upon anneal. Furthermore, there may be less reaction of TiN and HfO$_2$ than with a silicon containing gate electrode which may account for the thinner EOT and lower $V_{fb}$.

Finally, polysilicon gates have been shown to exhibit pinning of the Fermi level to midgap workfunctions\textsuperscript{7,8}. The difference in threshold voltage of these two samples is a negligible 20 meV. However, we would expect that given the values of band edge polysilicon and midgap TiN workfunctions this difference should be about
500 meV. Therefore, Fermi level pinning is believed to be occurring in these structures.

Drain currents for the two samples are shown in Figure 6.3 as a function of drain voltage for several gate voltages. The drain current for TiN can be seen to be slightly higher than that of the polysilicon gate possibly due to the thinner EOT. The gate leakage current measured on these devices are shown in Figure 6.4 as a function of effective dielectric field. The open symbols are the SiO₂ direct tunneling model for the same EOT for these devices. True high k behavior can be seen from the lower leakage current as compared to the model. The polysilicon gate also shows lower leakage due to its larger EOT either from thicker interfacial layer or another layer resulting from a reaction of the polysilicon gate electrode with the HfO₂.

6.3 Advanced Electrical Results

Interfacial trap densities and bulk trapping characteristics can be determined utilizing the characterization techniques outlined in Chapter 2. This section will investigate the similarities and differences in trapping behavior between the TiN and polysilicon gate electrodes.

6.3.1 Interface State Density

Because these samples have a dielectric which was deposited by ALD and a metal gate deposited by CVD, the amount of interface states created by the deposition process should be minimal. In order to determine interface state density
(\(D_{it}\)), charge pumping was performed on devices which were \(10 \, \mu m \times 0.5 \, \mu m\) in area. A similar setup as that in Chapter 5 was used for these measurements. Amplitude sweep measurements were performed and the results are shown in Figure 6.5. TiN shows high charge pumping current in the low amplitude bias region which can be attributed to gate leakage from the thin EOT. The value at the low bias was subtracted and then interfacial traps were determined as seen from the shifted TiN values. The results, shown in Table 6.2, indicate that TiN has a slightly larger \(D_{it}\) value than the polysilicon gate but both samples have \(D_{it}\) which is greater than \(10^{11}\) \(eV^{-1} \, cm^{-2}\) using similar parameters discussed in Chapter 5. Three level charge pumping measurements could not be performed on these devices because of high gate leakage current.

Table 6.2. Average \(D_{it}\) values for TiN and polysilicon on thin HfO\(_2\) using two level charge pumping.

<table>
<thead>
<tr>
<th>Gate</th>
<th>(D_{it} [eV^{-1} cm^{-2}])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly</td>
<td>(2 \times 10^{11})</td>
</tr>
<tr>
<td>TiN</td>
<td>(5 \times 10^{11})</td>
</tr>
</tbody>
</table>

In the amplitude sweep charge pumping technique, as the amplitude bias increases, traps further away from the interface can respond. These traps, referred to as bulk traps or slow traps are located within the bulk of the dielectric. This has been reported as a main source of mobility degradation for high \(k\) dielectrics especially\(^9\). Both samples have evidence of bulk trapping indicated by the increase
in charge pumping at higher biases and so the pulsed drain current technique was used to measure the bulk trapping characteristics of the thin HfO$_2$ samples.

6.3.2 Pulsed Drain Current

In order to analyze the bulk trapping in these dielectrics, pulsed drain current ($I_d$) measurements were performed. As discussed in Chapter 2, applying a pulse to the gate of an inverted transistor and measuring the voltage across a resistor at the drain node using an oscilloscope can be used to determine the drain current and the amount of charge trapping the device exhibits. Results from the pulsed $I_d$ versus gate voltage can be seen in Figure 6.6 using a 300 Ohm resistor, $V_{DD}$ of 100 mV, and 10 $\mu$m x 1 $\mu$m transistors. The rise and fall times were 5 $\mu$s and the pulsed width was 100 $\mu$s. The fast rise and fall time are selected to avoid any charge trapping induced by tunneling through the thin interfacial layer and was used previously in devices with EOTs < 20 Å$^{10}$. Similar charge trapping results are obtained from the metal and polysilicon gate electrodes indicating minimal charge trapping behavior.

Despite the exaggerated bulk trapping properties in the charge pumping measurements, the pulsed $I_d$ shows minimal bulk trapping from the high k dielectric. This may be due to the fact that interfacial layer is extremely thin and a large degree of tunneling is occurring during the rise times. As seen in Figure 6.7, the tunneling across the thinner interfacial layer, which is assumed for the TiN stack, may allow trapping to occur on the rise of the gate voltage, skewing the pulsed drain current results. The end result would not be a current free of bulk traps. On the other hand,
given that the polysilicon has a thicker interfacial layer, it would be safe to assume that the rate of tunneling may be lower and therefore it has a more accurate pulsed $I_d$ result. Since both $I_d$ curves indicate some bulk trapping an ultra fast rise time in the nanosecond range may be required for a true measurement of the bulk trapping characteristics. Additionally, the interfacial layer composition of the two gate electrodes may not be SiO$_2$ and could account for the bulk trapping behavior.

Using the voltage difference between the rising and falling pulsed $I_d$ curves at 100 $\mu$A, an estimate of charge density, $Q_{bt}$, can be determined from

$$Q_{bt} = C\Delta V$$  \hspace{1cm} (6.1)

where $C$ is the capacitance per unit area and $\Delta V$ is the voltage difference between the curves. The results are seen in Table 6.3 showing similar values for both TiN and polysilicon gates. The effects of this bulk trapping on mobility will be discussed later in this chapter but it is questionable that a charge density on the order of $10^{12}$ cm$^{-2}$ will have much impact on the mobility values of these devices because these traps are located away from the Si interface.

<table>
<thead>
<tr>
<th>Gate</th>
<th>$\Delta V$ [mV]</th>
<th>$Q_{bt}$ [cm$^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly</td>
<td>168</td>
<td>$1.7 \times 10^{12}$</td>
</tr>
<tr>
<td>TiN</td>
<td>148</td>
<td>$2.1 \times 10^{12}$</td>
</tr>
</tbody>
</table>
6.4 Mobility values

Mobility values of both polysilicon and TiN on HfO$_2$ were extracted by split CV analysis and using a leakage current correction that assumes 50% of the gate leakage comes from the drain as seen in eqn. 2.8. This is a valid assumption given the symmetric processing of these samples. The results are plotted as a function of effective field in Figure 6.8 and compared to the universal mobility. There is a shift in effective field from the slight difference in flatband voltage and EOT. Figure 6.9 shows the same mobility as a function of inversion charge showing that both samples have similar mobility values. However, both TiN and polysilicon devices have severely degraded mobility from the universal mobility values and compared to the thicker EOT devices studied in Chapter 5. The TiN shows a slightly lower mobility than the polysilicon gate possibly due to the thinner EOT. However, the bulk of this chapter will closely investigate the differences and similarities between the two gates on the thin high k dielectric to determine how much of an effect a thin EOT and different trapping characteristics has on mobility.

As mentioned in Chapter 2, the Hauser NCSU Mob2D program is similar to split CV analysis in that is uses a DC drain current versus gate voltage measurement and so will have some response to bulk trapping and $D_{it}^2$. The split CV technique relies on a capacitance curve to determine inversion charge, which is also subject to both $D_{it}$ and bulk traps. Alternatively, the Hauser NCSU Mob2D model uses an ideal CV curve to determine inversion charge and does not include $D_{it}$ or bulk trapping in that calculation. The comparison of these two analysis methods should prove how much $D_{it}$ and bulk trapping impact the mobility.
In Figure 6.10, the extracted mobility is shown for both the Hauser NCSU Mob2D model and split CV analysis. The Hauser NCSU Mob2D model extracts a surface roughness coefficient of \( \sim 26 \, \AA^2 \) and an interface scattering factor of \( \sim 1.2 \times 10^{11} \) cm\(^{-2} \) for both polysilicon and TiN gate electrodes. The TiN gatestack shows a larger difference between the Mob2D and split CV analysis but the polysilicon does not show much difference. Since polysilicon is assumed to have a thicker interfacial layer, the \( D_{it} \) and bulk trapping characteristics of this stack should not impact the mobility as much. The TiN gate electrode, on the other hand, has a thinner interfacial layer so the impact of interfacial traps and bulk traps will be more. In the case of TiN, the split CV analysis is overestimating the inversion charge which is scattered due to \( D_{it} \) and bulk trapping and can be corrected using the Mob2D model and a calculated capacitance. The overestimation causes a reduction in the mobility values of the TiN on thin HfO\(_2\) samples.

**6.5 The Role of Interfacial and Bulk Traps**

Since average interfacial trap density was extracted for both TiN and polysilicon, its particular effect on mobility can be determined. Chapter 2 illustrated two ways to correct for \( D_{it} \). One technique, developed at Yale, uses a calculated ideal capacitance to determine inversion charge which does not involve any extracted \( D_{it} \) values. The second technique converts the average \( D_{it} \) into a capacitance and then calculates the inversion charge using the measured capacitance curve. These two techniques can be compared with the thin HfO\(_2\) samples to see how much \( D_{it} \) impacts the mobility. The results of both correction
methods are seen in Figure 6.11 for both gate electrodes. In the case of TiN, the correction for average $D_{it} \sim 5 \times 10^{11}$ eV$^{-1}$ cm$^{-2}$ shows some increase in mobility. The Yale method shows even more indication of increased mobility which is explainable by the understanding of the method. In the Yale method, a theoretical split capacitance curve is calculated from the true inversion charge. For this case, the theoretical curve will exclude all effects from the $D_{it}$. Therefore, the Yale method is more accurate than just including a capacitance due to an average $D_{it}$.

In Figure 6.11, the polysilicon does not show any change using either correction for average $D_{it} \sim 2 \times 10^{11}$ eV$^{-1}$ cm$^{-2}$ or the theoretical capacitance calculation method of Yale. Again, this could be due to the interfacial layer thickness or gatestack composition difference between the TiN and polysilicon gate electrodes. The $D_{it}$ extracted from charge pumping in the TiN devices may be including more traps away from the thinner interfacial layer since it has a higher tunneling current. Thus the $D_{it}$ value may be overestimated and hence mobility would increase if these interface traps were taken into account. The polysilicon devices have a thicker interfacial layer and thus its $D_{it}$ value may be more accurate. The overestimation of $D_{it}$ for the TiN samples could cause the correction to have greater mobility.

According to published empirical relationships, with substrate doping levels of $10^{18}$ cm$^{-3}$ $D_{it}$ values of $10^{12}$ eV$^{-1}$ cm$^{-2}$ and lower should not impact the mobility as discussed in Chapter 5. This is due to the dominance of the coulombic scattering from ionized dopants at high doping levels. Since the $D_{it}$ values obtained for both polysilicon and TiN are on the order of $10^{11}$ eV$^{-1}$ cm$^{-2}$, their mobility values will not
change when these interface traps are taken into account. However, both the Yale method and the Hauser NCSU Mob2D method are also considering an ideal capacitance curve to calculate the inversion charge. Therefore even the bulk trapping from inversion charge is removed from these results and hence shows a slight increase in mobility, especially for the TiN case where the interfacial layer is thin. Using an average $D_{it}$ value correction does also show some increase in mobility for TiN but as stated above, this value may be overestimated, especially in the low inversion carrier density region where the effects of $D_{it}$ have greater impact.

The mobility analysis and corrective methods investigated so far still use a DC $I_d - V_g$ curve which gives traps located within the bulk of the dielectric enough time to respond. Using the pulsed drain current and applying the Hauser NCSU Mob2D model, the mobility can be extracted with a mobility that is free of bulk traps and $D_{it}$. Refer to Figure 6.6 which shows the results from Pulsed $I_d$ measurements discussed in Section 6.3.2. Bulk trapping densities were $\sim 10^{12}$ cm$^{-2}$ which were not thought to influence the mobility since it is removed from the Si interface. To prove this, the mobility extracted from Hauser NCSU Mob2D using a DC $I_d$-$V_g$ and the pulsed $I_d$ measurement was compared. Only the upward trace of the Pulsed $I_d$ curve was used since this trace is not influenced by bulk trapping assuming the rise time of 5 $\mu$s was appropriate for the interfacial layer thicknesses of these devices.

The results of the Hauser NCSU Mob2D comparison using DC $I_d$-$V_g$ and Pulsed $I_d$ measurements are shown in Figure 6.12. The TiN sample shows a slight increase in mobility while the polysilicon sample indicates no change in mobility value. Given that the TiN and polysilicon gate electrodes show the same bulk
trapping characteristics, this is further evidence that the polysilicon gate interfacial layer is thicker than the TiN. The impact of the bulk trapping is not seen in the mobility value of the polysilicon gate because the trapping is occurring away from the interface. On the other hand, the TiN gate has a thinner interfacial layer and so the bulk trapping will scatter the inversion carriers due to their closer proximity to the channel and therefore correcting for it should improve the mobility.

After all mobility values accounting for bulk trapping and interface trap density were calculated, the mobility behavior of both the metal gate and the polysilicon gate are similar but still degraded from the universal mobility. However, in Chapter 2 it was discussed that according to EOT values the mobility for TiN should be lower than the polysilicon since its EOT is thinner. This will be discussed in the final section.

6.6 Other Scattering Mechanisms

As mentioned, other factors could be dominating the mobility which does not allow the detrimental effects from interface trapped charge or bulk traps to be seen. One mechanism is the coulombic scattering of the ionized dopants in the channel given that the doping of these samples is very high. The mobility may be starting out at a degraded value because of the ionized impurities in the channel so that any more scattering from interfacial traps goes unnoticed. This could be the reason that the mobility values do not change drastically even when the high D_{it} values are taken into account.
The degraded mobility could also be attributed to a more dominant mechanism such as scattering due to interface traps or surface roughness. Surface roughness may be also playing a role in the mobility degradation. Since the mobility due to surface roughness is inversely proportional to the square of the effective field, the impact of surface roughness should only be a factor at higher fields. While surface roughness was not directly measured, the Hauser NCSU Mob2D model does extract a roughness coefficient which is sensitive to the step height times the step length and is given in Angstroms squared \([\text{Å}^2]\). A value greater than \(~25 \text{ Å}^2\) means that the roughness coefficient is detrimental to the mobility at higher fields. Both the TiN and polysilicon gated HfO\(_2\) samples have a roughness coefficient \(~26 \text{ Å}^2\) which should not influence the mobility. However, it is important to note that the interfacial layer of these devices may not be a stoichiometric SiO\(_2\) or that the deposition of the high k and its subsequent annealing sequences change the average roughness factor even at the internal interface, something which has not been investigated as a source of mobility degradation at this point.

To investigate the role of scattering due to lattice vibrations and to verify that the mechanism with a high k is similar to that of SiO\(_2\), measurements were performed at higher temperatures. Both the split CV and DC current measurements were performed at 20 °C, 100 °C, and 150 °C and mobility was extracted using split CV analysis. No corrections were made because at higher temperatures, the mobility limited by coulombic scattering increases. Therefore, the uncorrected split CV mobility will have less effect of high k bulk trapping and \(D_t\) at higher temperatures. It can be seen in Figure 6.13 that as a function of increasing
temperature, the uncorrected mobility for both TiN and polysilicon gate electrodes on thin HfO$_2$ decrease. This indicates that the role of phonon scattering from the silicon lattice is similar to that of SiO$_2$. As with the results in Chapter 5, peak mobility versus temperature plots were constructed and shown in Figure 6.14. The slopes were compared with that of the standard MOSFETs fabricated with candidate metal gate electrodes on HfO$_2$ showing similar if not slightly lower slope values. The slope of the peak mobility of all the HfO$_2$ dielectrics studied in this dissertation versus temperature is then plotted as a function of extracted EOT and shown in Figure 6.15. A general trend is realized that as the EOT is decreased the temperature dependence of mobility is less.

Again, it has been suggested that a lower temperature dependence of mobility with HfO$_2$ is due to the scattering rate of the soft optical phonons present in the high k dielectric which has a weaker dependence on temperature than that of SiO$_2$. However, to investigate whether other scattering factors could also cause a weaker temperature dependence, various mobilities were calculated using the Hauser NCSU Mob2D model for SiO$_2$ at higher temperatures and increased fitting parameter values. In Figure 6.16, the modeled mobility at three different temperatures for a SiO$_2$ device with surface roughness coefficient equal to 25 Å$^2$ and interface scattering factor equal to $1.5 \times 10^{10}$ cm$^{-2}$ is shown by the open symbols. These values constitute ideal parameters for a SiO$_2$ device with good mobility characteristics having the same EOT and doping as the devices studied in this chapter. If the surface roughness coefficient is increased by an order of magnitude, the temperature dependence of the mobility decreases drastically. The peak
mobility is plotted as a function of temperature for the ideal interface scattering factor and surface roughness coefficient in Figure 6.17. If the surface roughness coefficient is increased to 30 Å$^2$ the slope of the line goes down. If the interface scattering factor is increased to $1.2 \times 10^{11}$ cm$^{-2}$, the slope of the line becomes negligibly small. This indicates that the high interface scattering factor and not phonon limited mobility is playing the more dominant role in decreasing the temperature dependence of the high k dielectric.

6.7 Thin High K Mobility Degradation

Many possible reasons for the degradation of the mobility have been investigated on these aggressively scaled MOSFETs. Even with corrections made for $D_{it}$ and high k bulk trapping, the mobility still appears lower than that of the thicker EOT samples studied previously and seriously degraded from the universal mobility curve used in industry. With the thicker EOT samples, a thick interfacial layer resulted after the high temperature processing of the devices and recovered the mobility. While an interfacial layer is still possible with the thinner samples and evidence that even 6 Å of SiO$_2$ is enough to fully recover the low Si$_3$N$_4$ mobility values discussed in Chapter 4, these aggressively scaled high k devices do not show the recovery expected. Again, it is possible that the interfacial layer is (i) not thick enough to recover the mobility, (ii) that it is of substandard quality, or (iii) that it is a high k film which promotes further degradation. The entire stack may, in fact, be at the mercy of the interfacial layer which is dominating the mobility characteristics.
Transmission Electron Spectroscopy (TEM) was used to ascertain the physical structure of these gatestacks. The results are shown in Figure 6.18 for the polysilicon and TiN gate electrodes on thin HfO$_2$. Two main observations can be made here: (i) the TiN gated sample has a slightly thinner interfacial layer than the poly-Si gated sample and (ii) the polysilicon gate electrode has a reaction layer between the polysilicon and the bulk high k. The interfacial layer is about 1 nm for the polysilicon gate electrode but appears to be slightly thinner for the TiN gated sample. This indicates that the interfacial layer is a medium k dielectric given that the EOT is only somewhat thicker than 1 nm. Furthermore, TEM images do not give any information on composition of the layer so differences could persist between the two gate electrodes. In addition, the reaction layer observed between the polysilicon and the gate dielectric would tend to decrease the mobility since defects and traps are likely present at that interface. However, it is observed that the interfacial layers of TiN and polysilicon are similar.

Since the interfacial layer is greater than 6 Å and assuming that is enough to recover the mobility, the dominating feature then would be the high k dielectric. As described in Chapter 2, the Hall mobility and hence Hall inversion carrier density determination can not account for the low mobility values seen in these devices. To understand the lower mobility values of these devices and because their interfacial layer thickness are similar, the different scattering mechanisms of mobility were separated and the phonon mobility was extracted using Matthiessen’s rule such that
where $\mu_{ph}$ is the mobility due to both bulk substrate and soft optical phonons in the high k dielectric, $\mu_{coul}$ is the mobility due to coulombic scattering, $\mu_{sr}$ is the mobility due to surface roughness, and $\mu_{eff}$ is the effective mobility measured in these samples.

The coulombic component of mobility is comprised of scattering from ionized impurities in the channel as well as from interface states located at the interface and fixed charge located within the bulk of the dielectric. The impact of ionized impurities and interface state density was discussed in Chapter 5 and should be very similar for both the polysilicon and TiN gatestacks in this study. The mobility due to surface roughness should also be similar between the two different gate electrodes given that the surface roughness coefficient extracted from the Hauser NCSU Mob2D model is the same and also observing the slope for the mobilities are similar for both samples. The surface roughness component of mobility is also well published and depends on the inverse square of the effective field. Therefore, it will only affect the high field mobility and does not degrade the mobility in the region studied.

Subtracting, then, the coulombic component from the uncorrected calculated mobility leaves the mobility due to phonon scattering for TiN and polysilicon gate electrodes and can be seen in Figure 6.19. It is obvious that the phonon limited mobility is moderately higher for the polysilicon gate electrode than for the TiN gate electrode. However, the above extractions involved using DC $I_d$-$V_g$ curves which will have errors due to trapping. To this end, pulsed drain current corrected mobility was used and the coulombic component was subtracted to determine the phonon limited mobility due to phonon scattering.
mobility which is shown in Figure 6.20. As can be seen, the extracted phonon limited mobilities are now very similar for both TiN and polysilicon gated MOSFETs.

As was discussed in Chapter 1, one proposed model for the lower mobility observed with high-K is the low energy soft-optical phonon modes that couple with the inversion electrons and lower the mobility. The energies of these modes for high k dielectrics such as HfO$_2$ and ZrO$_2$ are about ten times lower than SiO$_2$ owing to the soft nature of the metal – oxygen bonds. For example, the phonon modes for HfO$_2$ range from $12.4 - 48.4$ meV, whereas those for SiO$_2$ vary from $144 - 165$ meV. In addition, in the temperature range used for measurements, the thermal energies of the electrons do not exceed $\sim 40$ meV and therefore the modes for SiO$_2$ are not reached. However, since the phonon modes for HfO$_2$ and ZrO$_2$ are near the thermal energy of the carriers at these temperatures (kT), these modes are easily captured.

It has also been theorized that the gate electrode itself plays a role in the scattering ability of these soft optical phonons. If the carrier concentration in the gate is high, then the carriers in the gate can assist in screening the soft optical phonon modes thereby minimizing the scattering events experienced by the channel electrons. This is especially true for thinner dielectrics where the screening occurs more effectively. The high carrier concentration in the gate leads to higher plasmon frequencies which then lead to effective coupling of these low energy soft optical phonons. This effect is more pronounced when the total thickness of the stack becomes thinner. However, if the gate carrier concentration is low, such as in the case of polysilicon gate electrodes, this screening does not occur and the channel electrons experience a larger degree of scattering. It should be noted that some
screening also occurs by the electrons in the substrate themselves, however this occurs at higher inversion charge densities\textsuperscript{11}. Therefore the use of metal gates will have a larger plasmon frequency which will then assist in effective screening of the soft-optical phonon modes.

Finally, the thickness of the interfacial layer will also play a role in determining the degree of soft optical phonon limited mobility. A thicker interfacial layer (or a more SiO\textsubscript{2} rich layer) will assist in exponentially decaying the phonon scattering before it reaches the channel electrons. To this end, a thicker interfacial layer will assist in recovering the mobility of the channel electrons. However, the composition of the interfacial layer will also play a role in this screening process. In this Chapter, the TiN gated device exhibited similar mobility as the polysilicon gated device even with a thinner EOT. This suggests two possible mechanisms: (i) metal gate induced screening of optical phonons which improves the phonon limited mobility or (ii) some other scattering mechanism, such as compensated charge or fixed charge at the top interface reduces the polysilicon gated device mobility. It should be reemphasized that the compensated or fixed charge induced scattering cannot be corrected for experimentally. Given that the interfacial layer of the TiN is similar to that of the polysilicon with a possible composition difference, the similar phonon mobilities extracted above suggest that metal gate screening is not occurring in the TiN structures. This states that yet again, the high k dielectric and the device processing are the limiting factor in mobility.

This is clearly shown when the peak mobility values are compared for several different dielectrics and metal gate electrodes from published reports in the last
several years shown in Figure 6.21\textsuperscript{12-15}. For SiO\textsubscript{2}, the trend is that as the EOT decreases beyond 19 Å, the mobility values decrease, indicating either some remote scattering from the gate or something inherent in thin SiO\textsubscript{2} which presents lower values\textsuperscript{12}. A possible trend is drawn for the SiO\textsubscript{2} case. Reports of various high k devices with either metal or polysilicon gates are also shown with drastically lower mobility values. However, because these values may be dependent on an interfacial layer, as stated above, the trend of their mobility values may be similar to what is being observed with ultra thin SiO\textsubscript{2} and, in fact, do show a similar trend. It is important to note that both metal gates and polysilicon gates show low values of mobility on high k dielectrics.

Although the above results show that polysilicon and TiN gated MOSFETs have similar mobility values, the results are nevertheless surprising since the TiN gated MOSFETs have significantly lower EOT values than the polysilicon gated MOSFETs (11.8Å vs. 15.6Å) and were therefore expected to have lower mobilities due the thinner EOT. Therefore, it is viable that metal gate electrodes prove to be the same, if not better, than polysilicon gate electrodes on high k devices with the same or even more aggressively scaled EOTs.
6.8 Figures

Figure 6.1. Electron mobility for TaSiN gate electrode MOSFETs.

Figure 6.2. Capacitance of TiN and polysilicon gate electrodes on 30 Å HfO₂.
Figure 6.3. Drain current as a function of drain voltage for metal gate and polysilicon on thin HfO$_2$.

Figure 6.4. Gate leakage current for metal gate and polysilicon on thin HfO$_2$ as a function of effective dielectric field.
Figure 6.5. Interfacial traps as a function of bias measured by amplitude sweep charge pumping for metal and polysilicon gates on thin HfO$_2$.

Figure 6.6. Drain current versus gate voltage resulting from pulsed drain current measurements.
Figure 6.7. Tunneling from pulsed drain current measurements for thin and thicker interfacial layers.

<table>
<thead>
<tr>
<th>SiO₂</th>
<th>HfO₂</th>
<th>TiN</th>
</tr>
</thead>
</table>

Thinner interfacial layer means more trapping.

<table>
<thead>
<tr>
<th>SiO₂</th>
<th>HfO₂</th>
<th>Poly</th>
</tr>
</thead>
</table>

Thicker interfacial layer means less trapping.

Figure 6.8. Split CV extracted electron mobilities as a function of effective field for TiN and Polysilicon gate electrodes on ultrathin HfO₂.

- HfO₂/TiN/Poly
- HfO₂/Poly
- Univ

Split CV
20 μm x 20 μm

Effective Electron Mobility [cm²/V·s] vs. Effective Field [MV/cm]
Figure 6.9. Effective electron mobility as a function of inversion charge for TiN and polysilicon gate electrodes on ultrathin EOT HfO$_2$. 
Figure 6.10. Split CV and NCSU Hauser Mob2D extraction of electron mobility for TiN and polysilicon gate electrodes.
Figure 6.11. Effective electron mobility using $D_{it}$ corrections illustrated by Yale or by converting to a capacitance for TiN and polysilicon gate electrodes on thin HfO$_2$. 

131
Figure 6.12. Mobility extracted from DC or pulsed drain current in conjunction with Mob2D.
Figure 6.13. Effective electron mobility as a function of effective field for 20 °C, 100 °C, and 150 °C for both TiN and polysilicon gate electrons on thin HfO$_2$.

Figure 6.14. Peak mobility versus temperature for aggressively scaled MOSFETs and compared with candidate metal gates on HfO$_2$ from Chapter 5.
Figure 6.15. Slope of peak mobility versus temperature plotted as a function of EOT for HfO\(_2\) dielectrics.

Figure 6.16. Mobility as a function of effective field calculated by the Hauser NCSU Mob2D model for different interface scattering parameters.
Figure 6.17. Modeled peak mobility versus temperature for different interface scattering and surface roughness parameter values using the Hauser NCSU Mob2D model.
Figure 6.18. TEM of aggressively scaled MOSFETs.
Figure 6.19. Effective mobility and calculated phonon limited mobility using Matthiessen's rule for polysilicon and TiN gates on thin HfO$_2$.

Figure 6.20. Corrected effective mobility and calculated phonon limited mobility using Matthiessen's rule for polysilicon and TiN gates on thin HfO$_2$. 
Figure 6.21. Mobility values versus EOT for both SiO$_2$ and high k with polysilicon and metal gates amalgamated from various published reports$^{12-15}$. 

- Takagi-LSI (SiO$_2$/Poly)
- Tsai-Intel (HfO$_2$/Poly & Metal)
- Chau-Intel (HfO$_2$/Metal)
- Lazar-NCSU (HfO$_2$/Poly & Metal)
- Kim-Sematech (HfO$_2$/Poly)
6.9 References


Chapter 7 Summary and Future Work

In this work, candidate metal gate electrodes on high k devices were investigated to see if they meet the needs of the ITRS. Robust electrical and analytical techniques were used to understand the characteristics of these devices focusing on the charge trapping and channel mobility. This chapter summarizes the key results as well as proposing future work in this area.

7.1 Goals of the Study

The main goal of this study was to compare and contrast the mobility characteristics of candidate metal gate electrodes and high k dielectrics. At the time of this work, HfO\textsubscript{2} was chosen as the high k dielectric because of its propensity in literature as well as availability. The metal gates studied included TaN, Ru\textsubscript{50}Ta\textsubscript{50}, Ru\textsubscript{90}Ta\textsubscript{10}, as well as comparing TiN to polysilicon gate electrodes on aggressively scaled devices. These metal gates show real promise as alternatives to polysilicon as a gate electrode and have been studied extensively as such. Therefore, their properties, while still not fully understood, have familiarity with experience. Finally, advanced characterization techniques were examined and put into practice for use in future device qualification.

7.2 Interfacial Layer Conclusions

MOSFETs with Si\textsubscript{3}N\textsubscript{4} gate dielectrics were studied to understand how nitrided interfaces impact the channel mobility and the implementation of scaled EOTs.
RPECVD Si$_3$N$_4$ gatestacks with polysilicon gate electrodes were formed *in situ* so that no part was exposed to atmosphere until after the gate was deposited. Mobility values for NMOS transistors showed reduced peak values and enhanced high field values. PMOS devices showed nonexistent inversion characteristics and mobility values. This was explained by a high density of donor type states near the valence band which shifts the flatband values and reduces the involvement of hole carriers in the creating the inversion layer. Finally, if a 6 Å SiO$_2$ layer was inserted between the silicon and the Si$_3$N$_4$ layer, full recover of the mobility was observed for both NMOS and PMOS devices.

### 7.3 Candidate Metal Gates Conclusions

Next, the basic electrical properties of three candidate metal gate electrodes on high k dielectrics were put forth and compared with those of the same metal gates on SiO$_2$ gate dielectrics of similar EOTs. TaN, Ru$_{50}$Ta$_{50}$, and Ru$_{90}$Ta$_{10}$ on HfO$_2$ and SiO$_2$ dielectrics were fabricated with EOTs ranging from 22 to 29 Å with the thinnest EOT having the TaN gate electrode. The metal gates showed similar shifts in flatband voltage between HfO$_2$ and SiO$_2$ and but did not show similar results to previously studied MOS capacitors. The HfO$_2$ dielectrics indicated high k behavior having larger physical thickness and reduced leakage current.

$D_{it}$ values were obtained using two level amplitude sweep charge pumping indicating expected higher values for the RuTa alloys and especially for the SiO$_2$ dielectrics as expected from increased sputtering damage of the co-sputtered RuTa alloys and physically thinner SiO$_2$ dielectrics even after high temperature annealing.
For the first time, three level charge pumping was performed for the different metal gate electrodes on high k dielectrics. Results indicated increased $D_{it}$ toward the conduction band edge with signs of a localized trap concentration near midgap matching recent published results.

The mobility values reported for the different metal gates on SiO$_2$ were similar to the universal mobility values indicating that metal gates are a viable alternative to polysilicon. The mobility of the gate electrodes on high k showed very good characteristics and only slightly degraded from universal values. This was attributed to the damage from the physical deposition process. While $D_{it}$ values were high, mobility values did not suffer, indicating that the coulombic scattering of mobility was playing a dominant role. Mobilities at higher temperatures showed that the high k values decreased at a faster rate than those of SiO$_2$ with increasing temperature. Similar effective mobility values between high k and SiO$_2$ were attributed to an interfacial layer formation or some amount of remote phonon screening from the gate electrode.

7.4 Aggressively Scaled MOSFET Conclusions

MOSFETs with EOT < 20 Å were fabricated at Sematech with TiN and polysilicon gate electrodes. The polysilicon gate electrode resulted in 16 Å EOT while the TiN gate electrode resulted in 11 Å EOT. However, similar flatband voltages indicated that Fermi level pinning was occurring in these structures. Again, high k behavior was observed in both devices in a reduction of the leakage current when compared to similar SiO$_2$ EOT values. The interface state density values for
both TiN and polysilicon gate electrodes were \( \sim 10^{11} \text{ eV}^{-1} \text{ cm}^{-2} \) and they both exhibited similar minimal pulsed drain properties.

The mobility extracted by Split CV analysis with no corrections indicates severely degraded values. Even with \( D_{it} \) and bulk trapping considerations, these values did not increase to universal mobility values. However, the role of the interfacial layer may play a crucial role in these values. The temperature dependence of the mobility of high k dielectrics showed a decreasing trend as EOT was decreased indicating that the phonon limited mobility and interface scattering were dominating the behavior. Phonon limited mobility was extracted and shown to be similar for both polysilicon and metal gated MOSFETs.

Since the interfacial layer and bulk high k layer thicknesses were similar, as shown in TEM images, and both the TiN and polysilicon gate electrode devices possessed the same phonon limited mobility values, it was determined that metal gate screening is not occurring for the TiN structures. After comparing the mobility values of these devices with other published reports of metal and polysilicon gates on high k dielectrics as well as ultra thin SiO\(_2\), the degraded mobility observed is dominated by the interfacial layer thickness and composition. Still, mobility values on thinner EOT devices with metal gate electrodes showed similar values to that of the polysilicon gate electrode. Given this, it is conclusive that metal gate electrodes can be utilized on high k dielectrics of the aggressively scaled devices planned for future technologies.
7.5 Major Contributions of this Work

- Investigated thin IL with medium k dielectrics MOSFETs resulting in a recovery of mobility values with 6 Å SiO$_2$.
- Performed three level charge pumping on various metal gate electrodes on high k dielectrics.
- Compared mobility values of different emerging metal gate electrodes on high k dielectrics resulting in similar values near universal.
- Compared mobility values of metal and polysilicon gate electrode on ultra thin EOT devices resulting in degraded mobility attributed to the thin interfacial layer and bulk high k properties and no evidence of metal gate screening.
- Analyzed the temperature mobility dependence of mobility for metal gate electrodes on high k devices resulting in a decreasing dependence on temperature with decreasing EOT attributed to interfacial scattering.
- Compared and contrasted the various mobility correction techniques for advanced gatestacks resulting in a trap free technique which showed that the mobility is highly dependent on the interfacial layer thickness and composition.

7.6 Future work

From this work, the application of metal gate electrodes on high k dielectrics even for devices with aggressively scaled EOTs is realized. However, more work can be done to understand the degradation of mobility. First, the investigation of
metal penetrating through to the channel of the device needs to be performed. Next, measurements of mobility at lower temperatures where bulk phonon mobility is negligible can rule out this portion of the phonon mobility from calculated values. Also, faster rise times in the pulsed drain current measurements will accurately determine bulk trapping effects on mobility. In addition, the effects of compensated charge should be analyzed and applied to these devices despite the fact that it is difficult to measure. Any sources of strain or bandgap narrowing in the channel should also be investigated as a source of mobility degradation. Finally, a better understanding of the surface roughness from analytical techniques on both the silicon interface as well as internal interfaces will help in determine the full impact of the surface roughening component of mobility.

In conjunction with the above proposed studies, PMOS devices need to be fabricated to discern their mobility values and mechanisms. Finally, reliability of these structures should be investigated to serve as a convincing argument for implementation. With thorough answers to the experiments listed in this section and the results of the study detailed in this dissertation, the mechanisms behind the resulting mobility of advanced CMOS devices will be established.
Appendices
Appendix A List of Publications

A.1 Published Works


A.2 Conference Presentations and Posters


Appendix B  GEM Mask

B.1 Overview

This mask set is intended for metal gate electrode and/or high k MOS devices. The complete process flow uses five masks however, a two, three, four, or five mask flow can be utilized to obtain the desired devices or flow complexity. A light field copy of the gate mask is included for polysilicon electrodes. All other masks are dark field and all masks assume positive resist for area definition. A substrate contact is included for alternate semiconductors or substrates which necessitate epitaxial contact. There are 99 12 x 12 mm chips and one blank chip to be used for TEM or other material analysis for a total of 100 chips. Two alignment marks are located on every chip and there are two global marks for the entire mask. Each mask layer can be aligned to any previous layer. The chip contains both overlap and nonoverlap capacitors, transistors, diodes, gated diodes, antenna structures, four point probe structures, TLM, and contact chains. This manual is split into three sections. The Mask section explains the mask nomenclature and alignment marks. The Process Flow section will list a general process for using the various mask sets. Finally, the Devices section will detail each device on the chip.

B.2 Masks

Each mask is 5.0 x 5.0 x 0.90 inches with soda lime substrate and chrome coating. They have been mirrored so that the correct reading is when the chrome
side is facing down. All masks are dark field except for the polysilicon mask, which is light field, and a direct copy of the gate mask. The tolerance of each mask is ± 0.5 μm. In order of process flow, the masks are: Active, Gate (or Poly), Substrate, Contact, and Metal. The alignment marks are also explained in this section. Feature sizes stated below exclude alignment marks and substrate contact.

B.2.1 Active

A dark field mask for definition of active areas isolated by field oxide. The minimum feature size is 10 μm and the minimum pitch is 20 μm. The maximum feature size is 4750 μm (metal only capacitors area).

B.2.2 Gate

A dark field mask for definition of gate electrodes assuming patterned liftoff. The minimum feature size is 3 μm (FET) or 12 μm (capacitor) and the minimum pitch is 40 μm. The maximum feature size is 500 μm.

B.2.3 Poly

A light field mask for definition of gate electrodes assuming patterned etch. The minimum feature size is 3 μm (FET) or 12 μm (capacitor) and the minimum pitch is 40 μm. The maximum feature size is 500 μm.
B.2.4 Substrate

A dark field mask for definition of substrate contacts or substrate based structures (e.g. TLM). The minimum feature size is 5 \( \mu \text{m} \) and the minimum pitch is 100 \( \mu \text{m} \). The maximum feature size is 9075 \( \mu \text{m} \) (substrate contact).

B.2.5 Contact

A dark field mask for definition of contact vias isolated by LTO. The minimum feature size is 10 \( \mu \text{m} \) and the minimum pitch is 50 \( \mu \text{m} \). The maximum feature size is 490 \( \mu \text{m} \).

B.2.6 Metal

A dark field mask for definition of metal contacts to be used as probing pads. The minimum feature size is 12 \( \mu \text{m} \) and the minimum pitch is 10 \( \mu \text{m} \) except for 5 \( \mu \text{m} \) in TLM. The maximum feature size is 500 \( \mu \text{m} \).

B.2.7 Alignment Marks

Each mask has an alignment mark directly left and right of the horizontal center and is labeled GLOBAL on each side of each mark. This alignment mark is a duplicate of the marks on each chip. Each layer can be aligned to any previous layer. Great care was taken to make the alignment marks easy to find by encompassing the mark in a large dark border. Also, time was invested to make the
marks easy to align. Therefore, the alignment mark on the mask is a dark cross and is aligned to an open cross on a previous layer. Subsequently, however, the alignment marks for aligned layers will be invisible after processing.

B.3 Process Flow

Process flows with varying mask sets are given below from most to least complex flow. Details of which devices are available or not available with the various mask flows are given in Devices section.

B.3.1 Five Mask

A general process for a complete five mask flow is listed below for both metal gate and polysilicon gate. It is assumed that metal gate definition will use patterned liftoff and polysilicon gate will use patterned etch. This mask flow will result in transistors with a substrate contact.
### Metal Gate

1. Substrate clean
2. Field oxide growth (~ 3000 Å)
3. Active Mask photolithography
4. Field oxide etch
5. Gate dielectric deposition
6. Gate Mask photo
7. Metal gate deposition
8. Metal gate liftoff
9. Ion implantation
10. Substrate Mask photo
11. Field oxide etch
12. LTO deposition
13. Contact Mask photo
14. LTO etch
15. Metal Mask photo
16. Metal contact deposition
17. Metal contact liftoff

### Polysilicon Gate

1. Substrate clean
2. Field oxide growth (~ 3000 Å)
3. Active Mask photolithography
4. Field oxide etch
5. Gate dielectric deposition
6. Polysilicon gate deposition
7. Poly Mask photo
8. Polysilicon etch
9. Ion implantation
10. Substrate Mask photo
11. Field oxide etch
12. LTO deposition
13. Contact Mask photo
14. LTO etch
15. Metal Mask photo
16. Metal contact deposition
17. Metal contact liftoff

### B.3.2 Four Mask

A general process for a four mask flow with the substrate step omitted is listed below for both metal gate and polysilicon gate. It is assumed that metal gate definition will use patterned liftoff and polysilicon gate will use patterned etch. Use this mask flow is transistors but no substrate contacts are needed.
<table>
<thead>
<tr>
<th>Metal Gate</th>
<th>Polysilicon Gate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Substrate clean</td>
<td>1. Substrate clean</td>
</tr>
<tr>
<td>2. Field oxide growth (~ 3000 Å)</td>
<td>2. Field oxide growth (~ 3000 Å)</td>
</tr>
<tr>
<td>3. Active Mask photolithography</td>
<td>3. Active Mask photolithography</td>
</tr>
<tr>
<td>4. Field oxide etch</td>
<td>4. Field oxide etch</td>
</tr>
<tr>
<td>5. Gate dielectric deposition</td>
<td>5. Gate dielectric deposition</td>
</tr>
<tr>
<td>6. Gate Mask photo</td>
<td>6. Polysilicon gate deposition</td>
</tr>
<tr>
<td>7. Metal gate deposition</td>
<td>7. Poly Mask photo</td>
</tr>
<tr>
<td>8. Metal gate liftoff</td>
<td>8. Polysilicon etch</td>
</tr>
<tr>
<td>9. Ion implantation</td>
<td>9. Ion implantation</td>
</tr>
<tr>
<td>10. LTO deposition</td>
<td>10. LTO deposition</td>
</tr>
<tr>
<td>11. Contact Mask photo</td>
<td>11. Contact Mask photo</td>
</tr>
<tr>
<td>12. LTO etch</td>
<td>12. LTO etch</td>
</tr>
<tr>
<td>14. Metal contact deposition</td>
<td>14. Metal contact deposition</td>
</tr>
<tr>
<td>15. Metal contact liftoff</td>
<td>15. Metal contact liftoff</td>
</tr>
</tbody>
</table>

**B.3.3 Three Mask**

A general process for a three mask flow with just the active, gate/poly, and substrate masks is listed below for both metal gate and polysilicon gate. It is assumed that metal gate definition will use patterned liftoff and polysilicon gate will use patterned etch. This mask flow should be used for capacitors only with a substrate contact. The substrate mask must be repeated for metal contact to the substrate.
<table>
<thead>
<tr>
<th>Metal Gate</th>
<th>Polysilicon Gate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Substrate clean</td>
<td>1. Substrate clean</td>
</tr>
<tr>
<td>2. Field oxide growth (~ 3000 Å)</td>
<td>2. Field oxide growth (~ 3000 Å)</td>
</tr>
<tr>
<td>3. Active Mask photolithography</td>
<td>3. Active Mask photolithography</td>
</tr>
<tr>
<td>4. Field oxide etch</td>
<td>4. Field oxide etch</td>
</tr>
<tr>
<td>5. Gate dielectric deposition</td>
<td>5. Gate dielectric deposition</td>
</tr>
<tr>
<td>6. Gate Mask photo</td>
<td>6. Polysilicon gate deposition</td>
</tr>
<tr>
<td>7. Metal gate deposition</td>
<td>7. Poly Mask photo</td>
</tr>
<tr>
<td>8. Metal gate liftoff</td>
<td>8. Polysilicon etch</td>
</tr>
<tr>
<td>10. Field oxide etch</td>
<td>10. Field oxide etch</td>
</tr>
<tr>
<td>11. Substrate Mask photo</td>
<td>11. Substrate Mask photo</td>
</tr>
<tr>
<td>12. Metal deposition</td>
<td>12. Metal deposition</td>
</tr>
</tbody>
</table>

**B.3.4 Two Mask**

A general process for a two mask flow with just the active and gate/poly masks is listed below for both metal gate and polysilicon gate. It is assumed that metal gate definition will use patterned liftoff and polysilicon gate will use patterned etch. Use this mask flow if only capacitors and no substrate contacts are needed.
Metal Gate
1. Substrate clean
2. Field oxide growth (~ 3000 Å)
3. Active Mask photolithography
4. Field oxide etch
5. Gate dielectric deposition
6. Gate Mask photolithography
7. Metal gate deposition
8. Metal gate liftoff

Polysilicon Gate
1. Substrate clean
2. Field oxide growth (~ 3000 Å)
3. Active Mask photolithography
4. Field oxide etch
5. Gate dielectric deposition
6. Polysilicon gate deposition
7. Poly Mask photolithography
8. Polysilicon etch

B.4 Devices

Each chip is 11 x 11 mm with a 1 mm distance separating every chip. Size labels are read as *active-gate* so that a capacitor labeled 50-100 has a 50 x 50 μm active area and a 100 x 100 μm gate area. MOSFET size labels are read *width/length* so a MOSFET labeled 25/5 is 25 μm in width and 5 μm in length. In general, labels are on the top left side of the structures and size labels on the top right. The exception to this are labels and size labels which are rotated by 90° and directly left of the structures. A large crossbar substrate contact spans the chip allowing the contact to be as large as possible and accessing many of the structures. There is a metal and gate border around every chip to visually separate the chips. These devices as well as the entire chip and alignment marks can be seen in the Figures Section B.6. A table listing all the device structures, dimensions, quantities, and labels is shown in the Table B.3. A table listing the availability of each structure for the various mask flows can also be seen in Table B.2 as some devices will not be available with the process flows detailed in the Process Flow section.
B.4.1 Overlap Capacitors

Typical capacitor structures isolated by field oxide which utilize either a 100 x 100 µm gate pad in the case of 10 x 10, 30 x 30, and 50 x 50 areas or a 5 µm gate overlap in the case of 100 x 100 or 200 x 200 areas. Overlap capacitors are available in any mask process flow.

B.4.2 Varying Overlap Capacitors

Due to the overlap of the gate, there is an overlap capacitance, $C_{ov}$, in parallel with the oxide capacitance. To ascertain this overlap capacitance, capacitor structures with varying gate overlaps are included. $C_{ov}$ can be extrapolated by plotting the measured capacitance versus the overlap and determining the quadratic intercept at zero overlap. Field oxide thickness can also be determined electrically from $C_{ov}$. The overlap dimensions provided are 1, 3, 5, 7, and 10 µm on each side for 10 x 10, 30 x 30, and 50 x 50 µm areas. A table giving the overlap areas for the above dimensions can be seen in Table B.1. Varying overlap capacitors are available in any mask process flow.

B.4.3 Antenna Overlap Capacitors

To assess damage imparted to the device by the overlapped areas, antenna structures are provided with large gate pads. The active areas are 50 x 50 µm and the gate pads are 200 x 200 µm or 500 x 500 µm corresponding to a 75 or 225 µm
overlap, respectively. Antenna capacitors are only available in a two or three mask process flow.

### B.4.4 Nonoverlap Capacitors

To avoid an overlap capacitance, $C_{ov}$, nonoverlap capacitors are included on this chip. The field oxide etched area is $200 \times 200 \, \mu m$ and $50 \times 50$ or $100 \times 100 \, \mu m$ gate areas are provided. Nonoverlap capacitors are only available in a two or three mask process flow.

### B.4.5 Metal Only Nonoverlap Capacitors

Capacitors which are not isolated by field oxide are also provided by etching a large active area. They are available in $30 \times 30$, $50 \times 50$, and $100 \times 100 \, \mu m$ areas. Metal only nonoverlap capacitors are only available in a two or three mask process flow.

### B.4.6 Field Oxide Capacitors

Capacitors on field oxide are provided to electrically determine the field oxide thickness as well as process flow corruptions. $400 \times 400$ and $500 \times 500 \, \mu m$ areas were designed for an assumed field oxide thickness $\sim 3000 \, \AA$. Field oxide capacitors are available in any mask flow.
B.4.7 Diodes

Gate metal – semiconductor diodes are provided in 50 x 50 and 100 x 100 \( \mu m \) areas. Diodes are only available in a four or five mask flow.

B.4.8 Gated Diodes

Gated diodes are provided with two different doped areas of 50 and 75 \( \mu m \) overlap. The gate area is 50 x 50 \( \mu m \). Gated diodes are only available in a four or five mask flow.

B.4.9 MOSFETs

MOSFETs are provided with various sizes of widths and lengths but similar proportions. They all have 40 x 60 \( \mu m \) S/D contact vias except for one of the 10 x 10 \( \mu m \) and 25 x 25 \( \mu m \) transistors rows which have 30 x 30 \( \mu m \) S/D contact and are marked by a hash (> and < or \( \approx \)). Every transistor has 30 x 90 \( \mu m \) gate contact vias and 100 x 100 \( \mu m \) metal pads for probing ease. The sizes include 10 x 3, 10 x 5, 10 x 10, 25 x 5, 25 x 10, 25 x 25, 50 x 5, 50 x 10, and 50 x 25 \( \mu m \) and are denoted by W/L. Transistors are only available in a four or five mask flow.

B.4.10 Four Point Probe

Greek cross structures are provided for four point probe sheet resistance measurements. There are four point probe structures for both substrate and
gate/poly resistivity determination in sizes of \( W \times L = 50 \times 50 \) and \( 100 \times 100 \) \( \mu \text{m} \). Substrate four point probe structures are only available in a four or five mask process flow. Metal four point probe structures are available in any mask process flow.

**B.4.11 Transfer Length Method**

Transfer length method (TLM) structures are provided for both doped and undoped substrate resistance measurements. By plotting the resistance between two contacts as a function of the spacing between the contacts sheet resistance, contact resistance, and specific contact resistance can be extrapolated. Specific parameter values are \( L = 75 \ \mu \text{m}, \ Z = 150 \ \mu \text{m}, \) and \( W = 190 \ \mu \text{m} \). The distances range from \( 5 \) to \( 50 \ \mu \text{m} \) in steps of \( 5 \ \mu \text{m} \). See *Semiconductor Material and Device Characterization* 2nd Edition by D. K. Schroeder, John Wiley and Sons, 1998, p. 157 for more information. TLM structures are only available in a four or five mask process flow.

**B.4.12 Contact Chains**

Two contact chains for both metal (undoped) and substrate (doped) process control are included. A simple measurement of this structure is an immediate pass or fail of the semiconductor or metal processing. However, knowing the sheet resistance of the semiconductor, the contact resistance can be determined by summing the total resistance between any two contacts. There are a total of 12
contacts (7 islands), \( Z = 50 \, \mu\text{m}, W = 100 \, \mu\text{m}, \) and \( d = 150 \, \mu\text{m}. \) See *Semiconductor Material and Device Characterization* 2nd Edition by D. K. Schroeder, John Wiley and Sons, 1998, p. 146 for more information. Metal contact chains are only available in a five mask process flow but substrate contact chains are available in a four or five mask process flow.

**B.4.13 Blank Chip**

There is a blank chip in the lower left quadrant of the mask which when completed will result in only gate metal directly on gate dielectric directly on an undoped substrate. All other layers will be etched away and is available in any mask flow. This chip can be used for material analysis purposes such as TEM.
### B.5 Tables

Table B.1. Overlap areas for varying overlap capacitors.

<table>
<thead>
<tr>
<th>Capacitor Area</th>
<th>Overlap Area [µm²]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 µm</td>
</tr>
<tr>
<td>10</td>
<td>44</td>
</tr>
<tr>
<td>30</td>
<td>124</td>
</tr>
<tr>
<td>50</td>
<td>204</td>
</tr>
</tbody>
</table>
Table B.2. Structures which are available (A) or not available (NA) according to the two, four, or five mask process flows described in Process Flow section.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Five Mask</th>
<th>Four Mask</th>
<th>Three Mask</th>
<th>Two Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap Capacitors</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Varying Overlap Capacitors</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Antenna Overlap Capacitors</td>
<td>NA</td>
<td>NA</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Nonoverlap Capacitors</td>
<td>NA</td>
<td>NA</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Metal Only Nonoverlap Capacitors</td>
<td>NA</td>
<td>NA</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Field Oxide Capacitors</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Diodes</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Gated Diodes</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Transistors</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Metal/Poly 4 Point Probes</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Substrate 4 Point Probes</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Doped Transfer Length Methods</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Undoped Transfer Length Methods</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Metal/Poly Contact Chains</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Substrate Contact Chains</td>
<td>A</td>
<td>A</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Substrate Contact</td>
<td>A</td>
<td>NA</td>
<td>A</td>
<td>NA</td>
</tr>
<tr>
<td>Blank Chip</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>
Table B.3. GEM chip device structures, sizes, quantities, and labels

<table>
<thead>
<tr>
<th>Structure</th>
<th>Size (µm)</th>
<th>Quantity</th>
<th>Label</th>
<th>Size Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap Capacitors 100 x 100 µm metal/gate pad</td>
<td></td>
<td></td>
<td>CAP</td>
<td>10-100</td>
</tr>
<tr>
<td>10 x 10</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 x 30</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 x 50</td>
<td>200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overlap Capacitors 5 µm overlap</td>
<td></td>
<td></td>
<td>CAP</td>
<td>200-210</td>
</tr>
<tr>
<td>10 x 10</td>
<td>100</td>
<td></td>
<td>OVERLAP</td>
<td>1,3,5,7,10</td>
</tr>
<tr>
<td>30 x 30</td>
<td>190</td>
<td></td>
<td>OVERLAP</td>
<td>1,3,5,7,10</td>
</tr>
<tr>
<td>50 x 50</td>
<td>48</td>
<td></td>
<td>OVERLAP</td>
<td>1,3,5,7,10</td>
</tr>
<tr>
<td>Varying Overlap Capacitors 1, 3, 5, 7, 10 µm overlap</td>
<td></td>
<td></td>
<td>OVERLAP</td>
<td></td>
</tr>
<tr>
<td>100 x 100</td>
<td>200 x 200</td>
<td>20</td>
<td>ANT</td>
<td>50-200</td>
</tr>
<tr>
<td>Antenna Overlap Capacitors 50 x 50 µm active area</td>
<td></td>
<td></td>
<td>ANT</td>
<td>50-500</td>
</tr>
<tr>
<td>50 x 50</td>
<td>100</td>
<td></td>
<td>NONOVERLAP</td>
<td>200-50</td>
</tr>
<tr>
<td>100 x 100</td>
<td>100</td>
<td></td>
<td>NONOVERLAP</td>
<td>200-100</td>
</tr>
<tr>
<td>Metal Only Nonoverlap Capacitors 30 x 30</td>
<td>30</td>
<td>50</td>
<td>METAL ONLY CAP</td>
<td>30</td>
</tr>
<tr>
<td>No field oxide isolation</td>
<td>50</td>
<td>50</td>
<td>METAL ONLY CAP</td>
<td>50</td>
</tr>
<tr>
<td>100 x 100</td>
<td>100</td>
<td>50</td>
<td>METAL ONLY CAP</td>
<td>100</td>
</tr>
<tr>
<td>Field Oxide Capacitors 400 x 400</td>
<td>400</td>
<td>10</td>
<td>FIELD OXIDE</td>
<td>400</td>
</tr>
<tr>
<td>500 x 500</td>
<td>500</td>
<td>10</td>
<td>FIELD OXIDE</td>
<td>500</td>
</tr>
<tr>
<td>Diodes</td>
<td>50</td>
<td>100</td>
<td>DIODE</td>
<td>50</td>
</tr>
<tr>
<td>100 x 100</td>
<td>100</td>
<td></td>
<td>DIODE</td>
<td>100</td>
</tr>
<tr>
<td>Gated Diodes 50 x 50 µm active area</td>
<td>50</td>
<td>16</td>
<td>G DIODE</td>
<td>50-50</td>
</tr>
<tr>
<td>Transistors</td>
<td>10 x 3</td>
<td>5</td>
<td>FET W/L</td>
<td>10/3</td>
</tr>
<tr>
<td>10 x 10</td>
<td>30</td>
<td></td>
<td>FET W/L</td>
<td>10/5</td>
</tr>
<tr>
<td>25 x 5</td>
<td>30</td>
<td></td>
<td>FET W/L</td>
<td>25/5</td>
</tr>
<tr>
<td>25 x 10</td>
<td>30</td>
<td></td>
<td>FET W/L</td>
<td>25/10</td>
</tr>
<tr>
<td>50 x 5</td>
<td>30</td>
<td></td>
<td>FET W/L</td>
<td>50/5</td>
</tr>
<tr>
<td>50 x 10</td>
<td>30</td>
<td></td>
<td>FET W/L</td>
<td>50/10</td>
</tr>
<tr>
<td>Metal/Poly 4 Point Probe Greek Cross</td>
<td>50</td>
<td>4</td>
<td>4 POINT METAL</td>
<td>50</td>
</tr>
<tr>
<td>Substrate 4 Point Probe Greek Cross</td>
<td>50</td>
<td>4</td>
<td>4 POINT SUBSTRATE</td>
<td>50</td>
</tr>
<tr>
<td>Metal/Poly Undoped Contact Chain</td>
<td>-</td>
<td>2</td>
<td>DOPED TLM</td>
<td>-</td>
</tr>
<tr>
<td>Undoped Transfer Length Method</td>
<td>-</td>
<td>2</td>
<td>UNDOPED TLM</td>
<td>-</td>
</tr>
<tr>
<td>Metal/Poly Undoped Contact Chain</td>
<td>-</td>
<td>4</td>
<td>METAL CHAIN</td>
<td>-</td>
</tr>
<tr>
<td>Substrate Doped Contact Chain</td>
<td>-</td>
<td>4</td>
<td>SUBSTRATE CHAIN</td>
<td>-</td>
</tr>
</tbody>
</table>
B.6 Figures

Figure B.1. GEM Mask chip.
Figure B.2. Layout for alignment marks on GEM chip.
Figure B.3. Layout of 100 x 100 μm overlap capacitors with 110 x 100 μm gate pads.

Figure B.4. Layout of 10 x 10 μm overlap capacitor with 100 x 100 μm gate pads and varying overlaps of 1 to 10 μm.
Figure B.5. of antenna overlap capacitors with 500 x 500 and 200 x 200 μm gate pads.
Figure B.6. Layout of nonoverlap capacitors of dimensions 50 x 50 and 100 x 100 μm with 200 x 200 μm active areas.

Figure B.7. Layout of metal only nonoverlap capacitors of dimensions 30 x 30, 50 x 50, and 100 x 100 μm.
Figure B.8. Layout of field oxide capacitors of dimensions 400 x 400 and 500 x 500 μm.

Figure B.9. Layout of diodes of dimensions 50 x 50 and 100 x 100 μm.
Figure B.10. Layout of gated diodes of dimension 50 x 50 µm.
Figure B.11. Layout of individual 50 x 50 \( \mu \)m transistors with 40 x 60 \( \mu \)m S/D contact via (A) and 30 x 30 \( \mu \)m S/D contact via (B). Each transistor has 30 x 90 \( \mu \)m gate contact via and 100 x 100 \( \mu \)m metal pads.

Figure B.12. Layout of four point probe greek cross structure of dimensions 50 x 50 and 100 x 100 \( \square \)m for both substrate and metal gate resistivity calculation.

Figure B.13. Layout of transfer length method structures of for both doped and undoped substrate resistance measurements.
Figure B.14. of contact chain structure of for both substrate and metal gate.
Appendix C  High K Dielectrics on SiC

This paper was submitted to *Applied Physics Letters* and accepted for publication in June 2001.

C.1 Abstract

**Characteristics of Metal Organic Remote Plasma Chemical Vapor Deposited Al$_2$O$_3$ Gate Stacks on SiC Metal Oxide Semiconductor Devices**

H. R. Lazar, V. Misra

Department of Computer and Electrical Engineering

R. S. Johnson, G. Lucovsky

Department of Physics

North Carolina State University, Raleigh, NC 27695

Metal organic remote plasma chemical vapor deposited (RPCVD) SiO$_2$/Al$_2$O$_3$ stacks were deposited on 6H p-type silicon SiC to fabricate a high $k$ gate stack SiC MOS capacitors. Capacitance-voltage (C-V) and current-voltage (I-V) measurements were performed. C-V characteristics showed excellent properties at room and higher temperatures. Samples exhibited a slight negative flatband shift.
from which net oxide charge ($Q_{ox}$) was calculated. Low leakage currents were observed even at high temperatures. I-V characteristics of $\text{Al}_2\text{O}_3$ were superior to those observed on AlN and $\text{SiO}_2$ dielectrics on SiC.

C.2 $\text{Al}_2\text{O}_3$ on SiC

$\text{SiC}$ has long been studied for high power, high frequency, and high temperature applications.\textsuperscript{1-3} It’s high breakdown field, high saturation velocity and relative high electron mobility make it a candidate for fabricating high power devices with reduced power loss. High power, however, leads to high fields placed on the semiconductor. Since the field across the dielectric is proportional to the SiC electric field, a large field is dropped across the dielectric resulting in a low breakdown field. Using high K gate dielectric materials can reduce these fields and increase the lifetime of the device.\textsuperscript{4}

Issues surrounding the wide bandgap of SiC and interface defect density will define the materials that can be deposited or grown on SiC. Most dielectrics follow the trend of decreasing bandgap with increasing dielectric constant.\textsuperscript{5} $\text{Al}_2\text{O}_3$ and $\text{Si}_3\text{N}_4$ are exceptions to this rule exhibiting large dielectric constants and large bandgaps. There have been few successful reports of high $k$ dielectrics on SiC including $\text{Si}_3\text{N}_4$, oxynitrides, and AlN. However, their dielectric constants are lower than $\text{Al}_2\text{O}_3$ ($k \sim 10$).\textsuperscript{5-8} Lipkin et al deposited AlN on SiC and then oxidized the film to create AlN:O.\textsuperscript{4} It should be noted that after the oxygen anneal there would be a significant amount of nitrogen left in this film and, therefore,
can not be directly compared to Al$_2$O$_3$. In this work, Al$_2$O$_3$ gate stacks are deposited on SiC and electrical and interface characteristics are reported.

The 6H p-type SiC wafers used were aluminum doped p$^+$ epitaxial layers with a nominal doping of 5.5 X $10^{17}$ grown on p$^+$ 6H SiC substrates. These wafers were diced and a standard RCA clean was performed at 75 °C. A low temperature oxide (LTO) was deposited at 200 °C, 25 Watts, and at a pressure of 1 Torr with a target thickness of 50 Å. This SiO$_2$ interface layer was chosen since to date, the SiO$_2$-SiC interface has been reported to have the best properties. Furthermore, Al$_2$O$_3$ deposition directly on Si surfaces has shown degraded interface properties requiring the use of a thin SiO$_2$ interface layer. After the interface layer formation, Al$_2$O$_3$ was deposited using metal organic remote plasma chemical vapor deposition (RPCVD) at 300 °C, 30 W, and at a pressure of 300 mTorr with a target thickness of 200 Å to create a SiO$_2$/Al$_2$O$_3$ gate stack. Triethylaluminum tri-sec-butoxide (TEDA-TSB) in helium was used as the precursor for deposition. The Al$_2$O$_3$ films were then annealed at 800 °C in argon for 30 seconds in order to relax the chemical and structural nature of the film. Aluminum metal, 100 nm in thickness, was evaporated and patterned as a gate electrode to fabricate MOS capacitors. Measurements were taken before and after a forming gas anneal (FGA) at 400 °C for 30 minutes. Interface and electrical properties were determined by high frequency capacitance-voltage (C-V) using an HP 4284A Precision LCR Meter and current-voltage (I-V) measurements using an HP 4155B Semiconductor Parameter Analyzer. Room and high temperature measurements were taken using a
Tempronics temperature controller. All measurements were obtained under dark conditions.

Auger Electron Spectroscopy (AES) was performed at intermittent steps of a previous Al₂O₃ deposition on Si and on 0.6 nm of RPCVD SiO₂ on Si. Evolving chemical composition and bonds of the Si interface and the deposited Al₂O₃ thin film can be determined from this measurement. As seen in Figure C.1, the Al-O feature at ~ 55 eV does not change with increasing deposition time and its energy is consistent with aluminum bonded to oxygen. From this, it was concluded that Al atoms are not bonded to Si atoms at the interface and the Al is fully oxidized throughout the deposited film.

The capacitors fabricated with SiO₂/Al₂O₃ gate stacks show excellent electrical characteristics. The room temperature C-V curves of p-type 6H SiC before and after FGA are shown in Figure C.2. The effective oxide thickness (EOT) of the Al₂O₃ samples obtained from the HF accumulation capacitance was 135 Å. Ideal flatband voltage, assuming no oxide charge, for aluminum on 6H SiC can be calculated using the doping density stated above and the following values: $E_g = 3.0$ eV, $q\chi = 3.85$ eV,¹¹ and $n_i = 1.6\times10^{-6}$ cm⁻³.³ For p-type 6H SiC, ideal flatband is about −2.65 V. The extracted flatband voltage was −4.1 V before FGA and −3.3 V after FGA, corresponding to a net positive charge in the range of $10^{12}$ cm⁻². This agrees with values reported for both SiO₂ ¹²,¹³ and alternate dielectrics on SiC.⁴,⁷,⁸,¹⁴ A parallel shift of the experimental capacitance curve towards ideal flatband after FGA was attributed to a reduction in fixed charge. The stretch out of the C-V curves is attributed to the interface trap charge (Dit) contribution to the voltage in the p-type
samples, which agrees with values reported in literature. However, as shown in the inset of Figure C.2 minimal frequency dependence is observed due to the non-responsive nature of the interface traps. Recent data of RPCVD SiO$_2$/Al$_2$O$_3$ gate stacks on Si has indicated the presence of negative charge at the SiO$_2$-Al$_2$O$_3$ internal interface.\textsuperscript{9} In this work, the high $D_{it}$ values present at the SiO$_2$-SiC interface dominate the negative charge contribution coming from the SiO$_2$-Al$_2$O$_3$ internal interface and thereby result in a net negative flatband shift.

Cooper reported that states which lie 0.7 eV above the band edge (for p-type material) can not respond to the changes in DC bias and are not detected for measurements at room temperature.\textsuperscript{15} Therefore, in order to sample more of the interface states, higher temperatures must be utilized. C-V measurements taken at several temperatures for two different frequencies are shown in Figure C.3. The negative flatband shift observed as temperature increases is larger in value and opposite in direction than compared with the temperature dependence of the work function difference between Al and 6H SiC. Since the curves are shifting towards negative values, the flatband can be attributed to the generation of positive charge arising from either fixed charge or slow interface traps.\textsuperscript{16}

Gate leakage current was measured as a function of voltage for several temperatures. Current densities derived from these I-V characteristics for dielectric fields of 0.8 MV/cm and 1.7 MV/cm are plotted as a function of inverse temperature (1/K) and are shown in Figure C.4. The current density at the maximum temperature for a 1.7 MV/cm dielectric field is only $10^{-6}$ A/cm$^2$, which is lower than reported for AlN on SiC.\textsuperscript{6} The observed superior leakage currents of Al$_2$O$_3$ are a
product of two advantages of Al$_2$O$_3$ over AlN. The first advantage is the dielectric constant of Al$_2$O$_3$ is higher than AlN, resulting in a physically thicker film. Accounting for the 50 Å of LTO, the dielectric constant of Al$_2$O$_3$ was found to be ~ 9 whereas published reports of AlN range from 7.8-8.5.\textsuperscript{6-8} The second advantage is that the barrier height at the metal-Al$_2$O$_3$ interface is larger than that of AlN thereby decreasing tunneling probability of gate injected electrons in accumulation. Published metal work functions of gate electrodes on AlN were close to 4.1 eV, which are similar to the aluminum work function of this work. A higher dielectric constant and larger barrier height result in lower leakage currents and indicate that Al$_2$O$_3$ films are a strong candidate for SiC MOSFET devices.

In summary, metal organic RPCVD SiO$_2$/Al$_2$O$_3$ gate stack MOS capacitors were successfully fabricated on 6H p-type SiC. Well-behaved capacitance-voltage and current-voltage characteristics were demonstrated at 25 °C and above. Reduction in fixed charge was found after FGA and large $D_{it}$ screened the negative charge present at the SiO$_2$-Al$_2$O$_3$ interface, resulting in a net negative flatband shift. High temperature C-V showed increasing negative flatband shift with increasing temperature. High temperature I-V showed an increase in current with temperature but current densities are lower than reported values at equal fields. These results are encouraging for metal organic RPCVD Al$_2$O$_3$ gate stacks on other wide bandgap semiconductors such as GaN and GaAs.

Acknowledgement: The authors would like to thank the National Science Foundation (#9906255) for their support of this work.
Figure C.1. Auger electron spectrum of $\text{Al}_2\text{O}_3$ deposition on 0.6 nm of $\text{SiO}_2$ on HF last Si.
Figure C.2. Capacitance-Voltage plot of 6H p-type SiC MOS capacitor before (circles) and after (squares) forming gas anneal (400 °C). The stack consists of 50 Å SiO$_2$/200 Å Al$_2$O$_3$ and the areas are 100 μm x 200 μm. A parallel shift in flatband voltage shows a reduction in fixed charge.
Figure C.3. Capacitance-Voltage plot of 6H p-type SiC MOS capacitor at 25 °C (circles), 100 °C (squares), and 150 °C (diamonds) for (a) 10 kHz and (b) 1 MHz. Area is 100 μm x 200 μm and the parallel shift in flatband is attributed to the increase in the number of states involved at higher temperatures.

Figure C.4. Current Density as a function of inverse temperature [1/K] for 6H p-type SiC at fields of 0.8 MV/cm and 1.7 MV/cm. Values are less than reported values at equivalent fields.
C.4 References
