Electrical characterization of high-k gate dielectrics for advanced CMOS gate stacks

Yi Ming Ding
New Jersey Institute of Technology

Follow this and additional works at: https://digitalcommons.njit.edu/dissertations

Part of the Electrical and Electronics Commons

Recommended Citation
Ding, Yi Ming, "Electrical characterization of high-k gate dielectrics for advanced CMOS gate stacks" (2016). Dissertations. 83.
https://digitalcommons.njit.edu/dissertations/83

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at Digital Commons @ NJIT. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital Commons @ NJIT. For more information, please contact digitalcommons@njit.edu.
Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use” that user may be liable for copyright infringement.

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation.

Printing note: If you do not wish to print this page, then select “Pages from: first page # to: last page #” on the print dialog screen.
The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.
ABSTRACT

ELECTRICAL CHARACTERIZATION OF HIGH-K GATE DIELECTRICS FOR ADVANCED CMOS GATE STACKS

by

Yi Ming Ding

The oxide/substrate interface quality and the dielectric quality of metal oxide semiconductor (MOS) gate stack structures are critical to future CMOS technology. As SiO$_2$ was replaced by the high-k dielectric to further equivalent oxide thickness (EOT), high mobility substrates like Ge have attracted increasing in replacing Si substrate to further enhance devices performance. Precise control of the interface between high-k and the semiconductor substrate is the key of the high performance of future transistor. In this study, traditional electrical characterization methods are used on these novel MOS devices, prepared by advanced atomic layer deposition (ALD) process and with pre and post treatment by plasma generated by slot plane antenna (SPA).

MOS capacitors with a TiN metal gate/3 nm HfAlO/0.5 nm SiO$_2$/Si stacks were fabricated by different Al concentration, and different post deposition treatments. A simple approach is incorporated to correct the error, introduced by the series resistance ($R_s$) associated with the substrate and metal contact. The interface state density ($D_{it}$), calculated by conductance method, suggests that $D_{it}$ is dependent on the crystalline structure of hafnium aluminum oxide film. The amorphous structure has the lowest $D_{it}$ whereas crystallized HfO$_2$ has the highest $D_{it}$.

Subsequently, the dry and wet processed interface layers for three different p type Ge/ALD 1nm-Al$_2$O$_3$/ALD 3.5nm-ZrO$_2$/ALD TiN gate stacks are studied at low temperatures by capacitance-voltage (CV), conductance-voltage (GV) measurement and
deep level transient spectroscopy (DLTS). Prior to high-k deposition, the interface is treated by three different approaches (i) simple chemical oxidation (Chemox); (ii) chemical oxide removal (COR) followed by 1 nm oxide by slot-plane-antenna (SPA) plasma (COR&SPAO); and (iii) COR followed by vapor O$_3$ treatment (COR&O$_3$). Room temperature measurement indicates that superior results are observed for slot-plane-plasma-oxidation processed samples.

The reliability of TiN/ZrO$_2$/Al$_2$O$_3$/p-Ge gate stacks is studied by time dependent dielectric breakdown (TDDB). High-k dielectric is subjected to the different slot plane antenna oxidation (SPAO) processes, namely, (i) before high-k ALD (Atomic Layer Deposition), (ii) between high-k ALD, and (iii) after high-k ALD. High-k layer and interface states are improved due to the formation of GeO$_2$ by SPAO when SPAO is processed after high-k. GeO$_2$ at the interface can be degraded easily by substrate electron injection. When SPAO is processed between high-k layers, a better immunity of interface to degradation was observed under stress.

To further evaluate the high-k dielectrics and how EOT impacts on noise mechanism time zero 1/f noise is characterized on thick and thin oxide FinFET transistors, respectively. The extracted noise models suggest that as a function of temperatures and bias conditions the flicker noise mechanism tends to be carrier number fluctuation model (McWhorter model). Furthermore, the noise mechanism tends to be mobility fluctuation model (Hooge model) when EOT reduces.
ELECTRICAL CHARACTERIZATION OF HIGH-K GATE DIELECTRICS FOR ADVANCED CMOS GATE STACKS

by
Yi Ming Ding

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering

August 2016
ELECTRICAL CHARACTERIZATION OF HIGH-K GATE DIELECTRICS FOR ADVANCED CMOS GATE STACKS

Yi Ming Ding

Dr. Durgamadhab Misra, Dissertation Advisor
Professor of Electrical and Computer Engineering, NJIT

Dr. Leonid Tsybeskov, Committee Member
Professor of Electrical and Computer Engineering, NJIT

Dr. Marek Sosnowski, Committee Member
Professor of Electrical and Computer Engineering, NJIT

Dr. Alan E. Delahoy, Committee Member
Research Professor of Physics, NJIT

Dr. Purushothaman Srinivasan, Committee Member
Member of Technical Staff, Globalfoundries

Dr. Kandabara Tapily, Committee Member
Senior Engineering, Tokyo Electron Limited
BIOGRAPHICAL SKETCH

Author: Yi Ming Ding
Degree: Doctor of Philosophy
Date: August 2016

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering, New Jersey Institute of Technology, Newark, NJ, 2016
- Master of Science in Electrical Engineering, New Jersey Institute of Technology, Newark, NJ, 2014
- Bachelor of Engineering in Measurement Technology and Instrument, Nanjing University of Science and Technology, Nanjing, P. R. China, 2011

Major: Electrical Engineering

Presentations and Publications:


感谢我的父母和激励过我的人
To my parents and those encouraged me
ACKNOWLEDGMENT

I would like to express my sincere appreciation to my dissertation and academic advisor, Dr. Durgamadhab Misra. It’s his effort, which brings me to the door of semiconductor characterization. Under his guidance, I have not only gained the knowledge but also professional communication skills. Besides, I would like to thank my committee members who gave me valuable suggestion during the research, Dr. Leonid Tsybeskov, Dr. Marek Sosnowski, Dr. Alan E. Delahoy, Dr. Purushothaman Srinivasan and Dr. Kandabara Tapily.

I would like to also thank Dr. Kandabara Tapily, Dr. Robert Clark, Dr. Steve Consiglio, Cory Wajda, Dr. Gert Leusink and other researchers of TEL Technology Center, America, LLC, Albany, NY, for supplying devices used in this research and sponsorship for this research.

During my research study, I spent a great deal of time with my colleagues, and they also assisted me with my work when it was necessary. They are, but no limited to: Dr. Md Nasir Uddin Bhuyian, Dr. Zi Meng Cheng, Dr. Xue Hai Tan, Peter Paliwoda, Sromana Mukhopadhyay and Shilpa, Mitra.

Finally, I have to express my love to my parents, who encouraged me to achieve a Ph. D degree. Without their encouragement and support, I could not have achieved this.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation and Objectives</td>
<td>5</td>
</tr>
<tr>
<td>1.2 Proposal Organization</td>
<td>6</td>
</tr>
<tr>
<td>2 LITERATURE REVIEW: HIGH-K MATERIAL AND ITS APPLICATION ON Si/Ge SUBSTRATE</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Requirement of High-k and the Necessary Property of Implementation</td>
<td>8</td>
</tr>
<tr>
<td>2.1.1 Permittivity, Bandgap, and Band Alignment to Silicon</td>
<td>10</td>
</tr>
<tr>
<td>2.1.2 Thermodynamic Stability</td>
<td>11</td>
</tr>
<tr>
<td>2.1.3 Film Morphology and Deposition Method</td>
<td>12</td>
</tr>
<tr>
<td>2.1.4 Interface Quality and Bulk Defects</td>
<td>13</td>
</tr>
<tr>
<td>2.1.5 Gate Compatibility and Process Compatibility</td>
<td>15</td>
</tr>
<tr>
<td>2.2 Interfacial Layer Scavenging</td>
<td>17</td>
</tr>
<tr>
<td>2.3 High-k Application on Ge Channel</td>
<td>22</td>
</tr>
<tr>
<td>3 MOS CAPACITOR DEVICE MODELLING AND ELECTRICAL CHARACTERIZATION METHOD</td>
<td>28</td>
</tr>
<tr>
<td>3.1 MOS Capacitor Device Modeling</td>
<td>28</td>
</tr>
<tr>
<td>3.2 Electrical Characterization Method</td>
<td>36</td>
</tr>
<tr>
<td>3.2.1 Low Frequency CV</td>
<td>36</td>
</tr>
<tr>
<td>3.2.2 CV Hysteresis</td>
<td>38</td>
</tr>
<tr>
<td>3.2.3 Conductance Method</td>
<td>40</td>
</tr>
<tr>
<td>3.2.4 Deep Level Transient Spectroscopy (DLTS)</td>
<td>44</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS
(Continued)

<table>
<thead>
<tr>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.5 Flicker Noise</td>
</tr>
<tr>
<td>3.2.6 IV Temperature Dependence</td>
</tr>
<tr>
<td>3.2.7 Time Dependent Dielectric Breakdown (TDDB)</td>
</tr>
<tr>
<td>4 OXIDE STRUCTURE-DEPENDENT INTERFACIAL LAYER DEFECTS OF HfAlO/SiO₂/Si STACK ANALYZED BY CONDUCTANCE METHOD</td>
</tr>
<tr>
<td>4.1 Research Motivation for Al Doped HfO₂ as High-k Dielectrics</td>
</tr>
<tr>
<td>4.2 Device Information and Experimental Procedure</td>
</tr>
<tr>
<td>4.3 CV, GV, EOT, Dielectric Constant, Leakage Current Density, ( D_{it} ) and ( V_{FB} )</td>
</tr>
<tr>
<td>4.4 Discussion on the Role of Aluminum in HfO₂</td>
</tr>
<tr>
<td>5 ELECTRICAL CHARACTERIZATION OF DRY AND WET PROCESSED INTERFACE LAYER IN Ge/HIGH-K DEVICES</td>
</tr>
<tr>
<td>5.1 Research Motivation for Interface Treatment of High-k/Ge</td>
</tr>
<tr>
<td>5.2 Device Information and Experimental Procedure</td>
</tr>
<tr>
<td>5.3 EOT, ( V_{FB} ), ( N_A ), Low Temperature CV, Surface Potential and ( D_{it} ) Estimation</td>
</tr>
<tr>
<td>5.4 Results and Discussion</td>
</tr>
<tr>
<td>5.4.1 Impact of Interface Treatment</td>
</tr>
<tr>
<td>5.4.2 Deep Level Transient Spectroscopy (DLTS)</td>
</tr>
<tr>
<td>5.4.3 I-V Characteristics</td>
</tr>
<tr>
<td>5.4.4 Evidence of Edge Capacitance in Pseudo-Accumulation Region</td>
</tr>
<tr>
<td>5.5 Conclusion</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

(Continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 IMPACT OF SLOT PLANE ANTENNA ANNEALING ON CARRIER TRANSPORT MECHANISM AND RELIABILITY ON ZrO$_2$/Al$_2$O$_3$/Ge GATE</td>
<td>98</td>
</tr>
<tr>
<td>6.1 Research Motivation of SPAO on High-k Layer</td>
<td>98</td>
</tr>
<tr>
<td>6.2 Device Information and Experimental Procedure</td>
<td>99</td>
</tr>
<tr>
<td>6.3 Carrier Transport Mechanism</td>
<td>101</td>
</tr>
<tr>
<td>6.4 Time Dependent Dielectric Breakdown Performance</td>
<td>108</td>
</tr>
<tr>
<td>6.5 Conclusion</td>
<td>113</td>
</tr>
<tr>
<td>7 FLICKER NOISE MECHANISM ON THICK AND THIN OXIDE FINFET</td>
<td>114</td>
</tr>
<tr>
<td>7.1 Motivation of Noise Model Study on FinFET</td>
<td>114</td>
</tr>
<tr>
<td>7.2 Device Information and Measurement Method</td>
<td>116</td>
</tr>
<tr>
<td>7.3 Results and Discussion</td>
<td>117</td>
</tr>
<tr>
<td>7.3.1 Bias Dependence of Flicker Noise</td>
<td>117</td>
</tr>
<tr>
<td>7.3.2 Short and Long Channel Effect on Flicker Noise</td>
<td>118</td>
</tr>
<tr>
<td>7.3.3 Noise Model Extraction at Room Temperature</td>
<td>119</td>
</tr>
<tr>
<td>7.3.4 Temperature Effect on Flicker Noise Model</td>
<td>121</td>
</tr>
<tr>
<td>7.3.5 Planar Structure VS FinFET</td>
<td>126</td>
</tr>
<tr>
<td>7.4 Conclusion</td>
<td>128</td>
</tr>
<tr>
<td>8 SUMMARY AND FUTURE WORK</td>
<td>129</td>
</tr>
<tr>
<td>8.1 Summary</td>
<td>129</td>
</tr>
<tr>
<td>8.2 Future Work</td>
<td>132</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS
(Continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFERENCES</td>
<td>133</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of Relevant Properties for High-k Candidates</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Material Characteristics of Alternative Channel Materials for Ge and Si</td>
<td>24</td>
</tr>
<tr>
<td>4.1</td>
<td>List of Process Information of 13 HfAlO Samples</td>
<td>62</td>
</tr>
<tr>
<td>4.2</td>
<td>List of Results of 13 HfAlO Samples</td>
<td>68</td>
</tr>
<tr>
<td>5.1</td>
<td>Estimated EOT, $V_{FB}$, $D_{it}$, and $J_g$ Information for Three Samples</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>Arrhenius Plot Fitting Results</td>
<td>94</td>
</tr>
<tr>
<td>6.1</td>
<td>EOT, $V_{FB}$, $D_{it}$, and Interfacial Layer Type for Three Samples</td>
<td>99</td>
</tr>
<tr>
<td>6.2</td>
<td>Transport Mechanisms are Summarized for Both GEI and SEI Mode</td>
<td>107</td>
</tr>
<tr>
<td>7.1</td>
<td>Information of Devices Under Test</td>
<td>116</td>
</tr>
<tr>
<td>8.1</td>
<td>Comparison of Characterization Method for $D_{it}$</td>
<td>130</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparison of ITRS 2007 and 2008 Update for the trends of printed (resist) and physical gate lengths</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>2011 ITRS “Equivalent Scaling” process technologies timing, overall roadmap technology characteristics (ORTC) MPU/High-performance ASIC half pitch and gate length trends and timing, and industry “Nodes”</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Schematic of ALD process. (a) Substrate surface has natural functionalization or is treated to functionalize the surface. (b) Precursor A is pulsed and reacts with surface. (c) Excess precursor and reaction by-products are purged with inert carrier gas. (d) Precursor B is pulsed and reacts with surface. (e) Excess precursor and reaction by-products are purged with inert carrier gas. (f) Steps 2–5 are repeated until the desired material thickness is achieved</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Trend of EOT of the gate insulator for various versions of ITRS</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Static dielectric constant versus band gap for candidate gate oxides</td>
<td>11</td>
</tr>
<tr>
<td>2.3</td>
<td>TEM image of crystallization in HfO$_2$/SiO$_2$ dielectrics with (a) 40% HfO$_2$ and (b) 80%</td>
<td>12</td>
</tr>
<tr>
<td>2.4</td>
<td>(a) TEM cross section of a high-k/metal gate stack where the BIL (bottom interfacial layer) was removed by post anneal technique (b) High field electron mobility degradation as a function of EOT from literature data. The mobility values are taken at at $E_{\text{eff}} = 1$ MV/cm</td>
<td>14</td>
</tr>
<tr>
<td>2.5</td>
<td>Electron mobility in devices with high-k gate oxides</td>
<td>15</td>
</tr>
<tr>
<td>2.6</td>
<td>Energy diagrams of threshold voltages for nMOS and pMOS devices using (a) midgap metal gates and (b) dual metal gates</td>
<td>16</td>
</tr>
<tr>
<td>2.7</td>
<td>EWF change with anneal temperature. Four different compositions of Al based n-metals deposited on thin TiN layer followed by barrier layer and Al-fill</td>
<td>17</td>
</tr>
<tr>
<td>2.8</td>
<td>Schematic of metal high-k gate stack</td>
<td>18</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.9</td>
<td>Bright field micrograph of the gate stack from STEM. The interface to Si/HfO$_2$ shows a weak contrast. Al and Si were found by EELS above TiN. Inset: Schematic gate stack structure and Al scavenging process. During silicidation Al distributes uniformly into poly-Si. In between NiSi and TiN, SiO$_x$ and AlO$_x$ accumulation is found.</td>
<td>19</td>
</tr>
<tr>
<td>2.10</td>
<td>Schematics of direct and remote scavenging techniques in the literature.</td>
<td>21</td>
</tr>
<tr>
<td>2.11</td>
<td>Schottky barrier heights obtained experimentally for various metals with different vacuum work-functions. In case of Ge, the Fermi-level is strongly pinned near the valence band edge.</td>
<td>25</td>
</tr>
<tr>
<td>2.12</td>
<td>Schematic showing the effect of annealing on ZrO$_2$/Ge gate stack (a, b) and HfO$_2$/Ge gate stack (c, d), before (a, c) and after (b, d) annealing at 500°C. HfGe$_2$ in (d) is speculation and has not been observed.</td>
<td>26</td>
</tr>
<tr>
<td>3.1</td>
<td>(a) Cross section of a MOS capacitor. (b) Energy-band diagram of the MOS capacitor showing the energy barrier between metal and SiO$_2$ and between silicon and SiO$_2$. The metal is aluminum and the silicon is p-type.</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>(a) Electrical field configuration in the SiO$_2$ and silicon surface calculated from Equations (3.1-3.5) as function of x. (b) Potential distribution in the SiO$_2$ and the silicon surface calculated from Equations (3.1-3.5) corresponding to the electric field configuration in (a).</td>
<td>32</td>
</tr>
<tr>
<td>3.3</td>
<td>Cross section of a MOS capacitor showing a simple equivalent circuit. The capacitor labeled $C_s$ is shown variable to denote its bias dependence.</td>
<td>33</td>
</tr>
<tr>
<td>3.4</td>
<td>To illustrate the operation of a MOS capacitor, capacitance as a function of bias, and the corresponding energy-band diagrams are shown in accumulation, depletion, inversion, and at flatbands. Energy values $E_c$ and $E_v$ are the conduction and valence band edges, respectively.</td>
<td>35</td>
</tr>
<tr>
<td>3.5</td>
<td>Equivalent circuits for conductance measurements; (a) MOS-C with interface trap time constant $\tau_{it} = R_{it}C_{it}$, (b) simplified circuit of (a), (c) measured circuit, (d) including series $r_s$ resistance and tunnel conductance $G_t$, capacitance and resistance from border trap $C_{BT}$, $R_{it}$, capacitance from side of measured area $C_{L}$.</td>
<td>36</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>Low frequency method applied on HfAlO/Si devices. $D_{it}$ is calculated in the range where the pointer indicates (depletion region). The dispersion in the negative range (accumulation region) is due to border traps.</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td>CV hysteresis are taken by sweep voltage from accumulation to inversion, and sweep back immediately.</td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>$dI/dV$ is the derivative from DC measurement of IV characteristics. This conductance is due to tunneling current. Small conductance signal is due to both tunneling current and interface states. When frequency is low enough, it converges to $dI/dV$.</td>
<td></td>
</tr>
<tr>
<td>3.9</td>
<td>(a) MOS-C model with the interface trap time constant ($\tau_{it} = R_{it}C_{it}$), where $C_{ox}$ is the oxide capacitance, $C_s$ is the substrate capacitance, $R_{it}$ and $C_{it}$ are the interface defects induced resistance and capacitance, respectively. (b) Simplified circuit of (a), $C_p$ and $G_p$ are the equivalent capacitance and conductance in the substrate. (c) Circuit used for measurement after the correction, $G_c$ is the corrected conductance and $C_c$ is the corrected capacitance.</td>
<td></td>
</tr>
<tr>
<td>3.10</td>
<td>Screen shot of Labview program for $D_{it}$ measurement.</td>
<td></td>
</tr>
<tr>
<td>3.11</td>
<td>Two voltages, namely $V_a$ and $V_b$, are alternatively applied on the gate stack. Capacitance was sampled at two different time $t_1$ and $t_2$ after voltage was returned to $V_a$. Normally, an observation of increased capacitance will happen if there is majority carrier emission from the traps at the interface.</td>
<td></td>
</tr>
<tr>
<td>3.12</td>
<td>Band diagram at two different bias conditions for p type substrate are plotted for both MOS structure ($SiO_2/p$ type Si) and PN junction (n type CdS/p type CdTe): (a) MOS structure, bias voltage = 0 V, (b) MOS structure, bias voltage = -0.5 V (voltage difference between Si and SiO$_2$).</td>
<td></td>
</tr>
<tr>
<td>3.13</td>
<td>Implementation of the rate window concept with a double boxcar integrator. The output is the average difference of the capacitance amplitudes at sampling times $t_1$ and $t_2$.</td>
<td></td>
</tr>
<tr>
<td>3.14</td>
<td>Carriers in channel are trapped and detrapped from dielectric.</td>
<td></td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**  
*(Continued)*

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.15</td>
<td>To illustrate source of noise, McWhorter model is based carrier number fluctuation, and Hooge model is based on mobility fluctuation. Unified model correlated them by surface scattering process.</td>
</tr>
<tr>
<td>4.1</td>
<td>Schematic of the HfO$_2$–Al$_2$O$_3$ phase diagram, the symbols represent the crystalline phases identified by electron diffraction in the films with different amounts of Al$_2$O$_3$ after annealing at 1400 °C.</td>
</tr>
<tr>
<td>4.2</td>
<td>Typical MOS structure in this study. (a) cap structure, (b) sandwich structure.</td>
</tr>
<tr>
<td>4.3</td>
<td>Steps calculating $C_c$, $G_c$ and $D_{it}$.</td>
</tr>
<tr>
<td>4.4</td>
<td>Capacitance data of the sample fabricated by depositing 10 cycles of HfAlO as a cap layer on the 30 cycles of HfO (T3-A2.4%), where the corrected capacitance, $C_c$, show less frequency dispersion as compared to the measured capacitance, $C_m$. Inset figure is the corresponding band diagram when gate voltage equals to -1 V.</td>
</tr>
<tr>
<td>4.5</td>
<td>Conductance data of the sample fabricated by depositing 10 cycles of HfAlO as a cap layer on the 30 cycles of HfO (T3-A2.4%), where the corrected conductance, $G_c$ in the accumulation region is much smaller than in the depletion region. However, the measured conductance, $G_m$ show frequency dispersion in the accumulation region. Inset figure is the corresponding band diagram when gate voltage equals to -0.5 V.</td>
</tr>
<tr>
<td>4.6</td>
<td>$\frac{G_p}{\omega}$ is plotted by sweeping both frequency and gate voltage. The inset shows band diagram. $E_c$, $E_i$, $E_F$ and $E_v$ are the conduction band, intrinsic level, Fermi-level and valence band, respectively.</td>
</tr>
<tr>
<td>4.7</td>
<td>The corrected conductance ($G_c$) values of 13 different samples are plotted as a function of gate voltage.</td>
</tr>
<tr>
<td>4.8</td>
<td>Leakage current density measured at gate voltage ($V_{FB}$-1 V). The results of 13 samples which were annealed at the temperatures 680 °C and 800 °C, are plotted as a function of dielectric constant.</td>
</tr>
<tr>
<td>4.9</td>
<td>The interface state density ($D_{it}$) values are plotted as a function of the dielectric constant of 13 samples.</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4.10</td>
<td>The interface state density ($D_{it}$) as a function of the Al/(Al+Hf) % of 13 samples are plotted. Data points within the dashed line are the 3 samples where Al was introduced by HfAlO as the cap layer on HfO$_2$. Data points in solid line are the 2 samples where Al was introduced by sandwiching HfAlO inside the HfO$_2$ stack.imento.</td>
</tr>
<tr>
<td>4.11</td>
<td>The leakage current density ($J_g$) measured at the gate voltage ($V_{FB}$-1 V) as a function of the Al/(Al+Hf) % of 13 samples is shown as a function of Al concentration. Dashed dot line and dashed line follow the trend of $J_g$ varies with different Al doping level at 680 °C PDA and 800 °C PDA, respectively.</td>
</tr>
<tr>
<td>5.1</td>
<td>Structure of MOS capacitor of high-k/Ge.</td>
</tr>
<tr>
<td>5.2</td>
<td>(a) Measured capacitance ($C_m$) of Chemox sample observed at different frequencies is plotted as a function of gate bias. (b) Corrected capacitance ($C_c$) of Chemox sample measured at different frequencies is plotted as a function of gate bias.</td>
</tr>
<tr>
<td>5.3</td>
<td>Corrected capacitance ($C_c$) at different frequencies and different temperatures is plotted as a function of gate bias for sample Chemox. (a) at five different temperature, (b) at 100 K and 300 K. Both frequency dispersion and temperature dispersion can be observed.</td>
</tr>
<tr>
<td>5.4</td>
<td>Corrected capacitance ($C_c$) of Chemox sample measured by different frequencies at 100 K is plotted as a function of gate bias. Frequency dispersion can be observed.</td>
</tr>
<tr>
<td>5.5</td>
<td>$C_{it}$ ($C_{it} = C_{acc}-C_{FB}$) as function of temperature (1/kT) at fixed frequency (100 KHz).</td>
</tr>
<tr>
<td>5.6</td>
<td>Corrected 1 MHz capacitances ($C_c$) of three samples are plotted as a function of gate voltage at room temperature. Device area is 40 μm × 40 μm.</td>
</tr>
<tr>
<td>5.7</td>
<td>Corrected capacitance ($C_c$) is plotted as function of gate voltage at 100 K, (a) COR&amp;SPA$	ext{O}_x$, (b) COR&amp;$	ext{O}_3$.</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

(Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8</td>
<td>(a) Interface state density ( D_{it} ) is plotted as function of bandgap for three samples at two temperatures (100 K and 300 K), other ( D_{it} ) as a function of temperature are within this range; (b) Calculated ( D_{it} ) of three samples as a function of gate voltage by capacitance spectroscopy method and conductance method.</td>
</tr>
<tr>
<td>5.9</td>
<td>Deep level transient spectrum for three samples, ( C_{t1} ) and ( C_{t2} ) are two capacitance values sampled at two different time ( (t_1 &lt; t_2) ), ( T ) is temperature, ( \ln ) is nature logarithm, ( \tau ) is time constant calculated as ( \frac{t_2 - t_1}{\ln(t_2/t_1)} ), (a) simple chemical oxidation (Chemox), (b) chemical oxide removal (COR) followed by 1 nm oxide by slot-plane-antenna (SPA) plasma (COR&amp;SPAO(_x)), (c) COR followed by vapor O(_3) treatment (COR&amp;O(_3)). (d) is Arrhenius plot for all three samples.</td>
</tr>
<tr>
<td>5.10</td>
<td>Comparison of gate leakage current density ( J_g ), (a) ( J_g ) as a function of gate voltage, (b) ( J_g ) as a function of EOT for different splits.</td>
</tr>
<tr>
<td>5.11</td>
<td>Corrected capacitance ( C_c ) measured at -1.5 V (pseudo-accumulation) were measured at two different frequencies, 1 MHz and 1 KHz, respectively. (a) At 1 MHz, the capacitance is contributed from edge and it has linear dependence of device perimeter (b) At 1 MHz, the capacitance is contributed from interface states in area and it has linear dependence of device area.</td>
</tr>
<tr>
<td>6.1</td>
<td>Application of SPAO at different stages of gate stack deposition.</td>
</tr>
<tr>
<td>6.2</td>
<td>Current density ( J_g ) is plotted as a function of gate voltage ( V_g ) at four different temperatures (25 °C, 50 °C, 80 °C, 110 °C) for three different samples, (a) Ge/SPAO/( \text{Al}_2\text{O}_3/\text{ZrO}_2 ), (b) Ge/( \text{Al}_2\text{O}_3/\text{SPAO/ZrO}_2 ), (c) Ge/( \text{Al}_2\text{O}<em>3/\text{ZrO}<em>2/\text{SPAO} ). (d) ( J_g ) as a function of ( V_g - V</em>{FB} ) is plotted for the above three samples at 25 °C where ( V</em>{FB} ) is the flat band voltage.</td>
</tr>
<tr>
<td>6.3</td>
<td>( \ln(J_gE^2) ) is plotted as function of ( 1/E ) as FN tunneling for sample Ge/( \text{Al}_2\text{O}_3/\text{ZrO}_2/\text{SPAO}, J_g ) is current density and ( E ) is electric field. (a) gate electron injection mode (GEI), (b) substrate electron injection mode (SEI).</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>6.4</td>
<td>(a) Current density $\ln(J_g)$ is plotted as function of $1/kT$ for sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO, in SEI mode. (b) Slope value ($S = qaE - E_a$), which is obtained from (a) is then plotted as function electric field $E$ (SEI mode).</td>
</tr>
<tr>
<td>6.5</td>
<td>Hopping Conduction and Poole-Frenkel emission (PF) were used to fit the I-V characteristics for sample Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ (solid symbol) and sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$ (open symbol) at both GEI mode (black symbol, squares) and SEI mode (red symbol, circles). (a) Slope value ($S = qaE - E_a$) is plotted as function electric field $E$. (b) Slope value ($S = \beta E^{1/2} - q\phi$) is plotted as function $E^{1/2}$, where $\beta = q^3/\pi\varepsilon_r\varepsilon_0$. The slope value $S$ is the fitting value from $\ln(J_g)$ versus $1/kT$.</td>
</tr>
<tr>
<td>6.6</td>
<td>Carrier transport mechanisms are explained in band diagram for both GEI mode and SEI mode. (a) GEI mode, band diagram is simulated under $V_g = -1.5$ V. (b) SEI mode, band diagram is simulated under $V_g = 1$ V.</td>
</tr>
<tr>
<td>6.7</td>
<td>$\beta$ values are around 0.7 and similar for all samples and stress conditions, which were obtained by Weibull plot.</td>
</tr>
<tr>
<td>6.8</td>
<td>Charge to breakdown value ($Q_{BD}$) is plotted as function of gate voltage ($V_g$) for both GEI mode (solid symbol) and SEI mode (open symbol).</td>
</tr>
<tr>
<td>6.9</td>
<td>Gate leakage current ($I_g$) is plotted as function sampling time for all samples.</td>
</tr>
<tr>
<td>7.1</td>
<td>Noise spectra (average value of eight dies) are plotted for both (a) nFinFET and (b) pFinFET. Bias conditions are at low overdrive voltage in linear region.</td>
</tr>
<tr>
<td>7.2</td>
<td>Normalized $S_{id} (S_{id}/I_d^2)$ sampled at 25 Hz is plotted against (a) drain current ($I_d$), and (b) overdrive voltage ($V_g - V_{th}$). Device under test is thin nFinFET.</td>
</tr>
<tr>
<td>7.3</td>
<td>The normalized $S_{id} (S_{id}/I_d^2*WL)$, is plotted for all four type FinFET at both linear and saturation regions. At relative large gate bias, $\sim V + 0.2$ V, thin oxide FinFET shows larger discrepancy than thick oxide due to both short channel and LDD effect.</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>7.4</td>
<td>Noise mechanisms were extracted at low gate bias in linear region by fitting normalized $S_{id}$ versus $V_g V_{th}$ . (a) n/p thick FinFET, (b) n/p thin FinFET.</td>
</tr>
<tr>
<td>7.5</td>
<td>Left axis shows oxide trap density ($N_t$) against overdrive voltage in linear region for thick nFinFET, right axis shows Hooge empirical parameter ($\alpha_H$) against overdrive voltage ($V_g - V_{th}$) in linear region for thick pFinFET and thin n/p FinFET.</td>
</tr>
<tr>
<td>7.6</td>
<td>Normalized $S_{id}$ ($S_{id}/I_d^2$) is plotted as function of overdrive voltage ($V_g - V_{th}$) at linear bias condition under three different temperatures (-40 °C, 25 °C, 125 °C) for four types of devices (a) thick nFinFET, (b) thick pFinFET, (c) thin nFinFET, (d) thin pFinFET. The data is an average of eight devices.</td>
</tr>
<tr>
<td>7.7</td>
<td>Normalized $S_{id}$ ($S_{id}/I_d^2$) is plotted as function of overdrive voltage ($V_g - V_{th}$) at both linear and saturation bias conditions for thick pFinFET.</td>
</tr>
<tr>
<td>7.8</td>
<td>Slope of flicker noise spectra, $\gamma$ ($S_{id} = \frac{1}{f^\gamma}$) plotted for four devices at different temperatures (-40 °C, 25 °C, 125 °C) and bias conditions (linear and saturation). Slope values are average value obtained from eight devices at both the bias conditions.</td>
</tr>
<tr>
<td>7.9</td>
<td>(a) Normalized $S_{id}$ ($S_{id}/I_d^{2<em>WL}$) is plotted against different structure at different current bias condition (linear region) for nFET; (b) Normalized $S_{id}$ ($S_{id}/I_d^{2</em>WL}$) is plotted against drain current ($I_d$) at linear region for nFET.</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

The gate length of Complementary Metal Oxide Semiconductor (CMOS) technology has been scaled to 14 nm (channel length) and is commercially available in 2014 from Intel, which is a three years delay from the predicted year from ITRS 2007 report (Figure 1.1) [1]. Since then, the aggressive scaling has slowed down due to transistor reaching its physical limits, one such issue is high leakage current between source and drain. To address the S/D off-leakage current, industries have been finding solutions. As Figure 1.2 shows, three approaches are being followed to continue scaling of chips [2]: (i) gate stack material, (ii) channel material, and (iii) device architecture. The oxide layer of gate stack is the leading candidate as devices below 10 nm [3]. Metal gate high-k (MGHK) has been implemented to boost chip performance while physical thickness is kept thick enough (~5 nm) to prevent large direct tunneling current. Different materials other than Si, have been considered as a substrate over decades at research level, which have higher carrier mobility [4]. Nevertheless, only strained silicon has been implemented so far, since Ge or GaAs are still in their research stage. 3D structures like FinFET have been successfully implemented in the industry to improve drain-induced barrier lowering (DIBL) (short channel effect) [5]. This dissertation is focused on electrical characterization methods on these new novel nano devices and their electrical property under different process conditions. Critical issues need to be addressed before the implementation of these devices in the semiconductor industry. Additionally, compared to traditional CMOS devices, other physical phenomena are occurring since the
device is in nano range.

![Comparison of ITRS 2007 and 2008 Update for the trends of printed (resist) and physical gate lengths.](image)

**Figure 1.1** Comparison of ITRS 2007 and 2008 Update for the trends of printed (resist) and physical gate lengths.

Source: [1].

While CMOS technology is scaling down, the process for high-k deposition and annealing have also improved. Atomic layer deposition (ALD) method has significant advantage over alternative deposition methods, such as chemical vapor deposition (CVD) and various physical vapor deposition (PVD) techniques, due to its conformity and control over materials thickness and composition. These desirable characteristics originate from self-saturating nature of ALD processes [6]. For deposition of dielectric (high-k material), ALD has been used in this research. Figure 1.3 shows the schematic of ALD process.
During or after the ALD deposition, the quality of dielectric in terms of EOT and interface state density can be improved by exposing them to a slot-plane-antenna (SPA) plasma with various gases such as O₂ or inert gases [7]. The SPA plasma provides a high density plasma at low electron temperature, where the radicals diffuse from the plasma generation region to the wafer surface. SPA plasma is also a very low damage plasma process compared to conventional inductively coupled plasma (ICP) or electron cyclotron resonance (ECR) plasma [7]. It was found that the SPA plasma helps better film densification as well as improved interfacial layer growth. The dielectric is prevented from crystallization at low annealing temperatures. Consequently, oxygen cannot diffuse through boundary of crystalline and interfacial layer thickness is kept thin.
Figure 1.3 Schematic of ALD process. (a) Substrate surface has natural functionalization or is treated to functionalize the surface. (b) Precursor A is pulsed and reacts with surface. (c) Excess precursor and reaction by-products are purged with inert carrier gas. (d) Precursor B is pulsed and reacts with surface. (e) Excess precursor and reaction by-products are purged with inert carrier gas. (f) Steps 2–5 are repeated until the desired material thickness is achieved.

Source: [6].

Several electrical characterization methods were used to evaluate the interface quality or oxide quality in terms of defects. These methods can be categorized into two major groups, i.e., evaluation of capacitance of the gate stack and leakage current through the dielectric. Because of existing defects, experimental results can deviate from theoretical calculations. On the other hand, these deviations can be utilized to evaluate defects in the device. Characterization methods such as conductance method [8], capacitance-voltage (CV) at various low temperatures, flicker noise, capacitance transient spectroscopy, deep level transient spectroscopy (DLTS) [9], CV hysteresis, and time dependent dielectric breakdown (TDDB) [10] are used in this work.
1.1 Motivation and Objectives

The devices under test in this research involve high-k/Si (next generation), high-k/Ge, and MGHK-FinFETs. One of the critical issues in MGHK, is the high interface state density ($D_{it} \sim 10^{12} \text{ cm}^{-2}\text{eV}^{-1}$) [11] compared to traditional SiO$_2$/Si system ($\sim 10^{10} \text{ cm}^{-2}\text{eV}^{-1}$) [12]. Conventional SiO$_2$/Si system prevailed over decades due to its perfect interface quality due to thermally grown SiO$_2$ on Si substrate [12]. Introducing high-k with metal gate (HKMG) brings additional reliability issue like threshold voltage degradation ($\Delta V_{\text{th}}$) after bias temperature instability (BTI) stress in both nMOS and pMOS transistors [13]. This is due to the degradation of both interface and high-k gate dielectrics. If silicon substrate is replaced by other materials, it is necessary to address the interface defects density before expected mobility can be achieved [14]. Therefore, it is imperative to characterize the high-k dielectric layer and the interface quality.

The objective of this research is to use various electrical characterization techniques to study interface quality and high-k dielectrics deposited by various process condition. This provides the comprehensive information on defects, such as density, energy level, time constant and how they interact with other parameters (like flat band voltage $V_{\text{FB}}$, flicker noise, and dielectric life time). Both theoretical model and experimental work are involved in the research work. Different evaluation methods can provide a good analytical approach to study the dielectrics in the gate stacks. The correlation of experimental data from different methods can enhance the understanding the defects behavior. Furthermore, this study will discuss the advantages and disadvantages of various techniques, since each method has its own limitation like sensitivity, range, different extracted parameters and difficulty of implementation.
1.2 Dissertation Organization

Chapter 2 reviews state of art MOS-capacitor. Introduce why high-k layer is a requirement for future transistor scaling, and explain the basic requirement to scale EOT below sub-nm range. New interface control technique was introduced. high-k/Ge was discussed as well as for its high mobility and its subsequent interface challenge before it can be fully considered for commercial use.

Chapter 3 reviews different electrical characterization methods that used in this research work.

Chapter 4 discusses the interface properties in a metal oxide semiconductor capacitor (MOS-C) device with a 3 nm HfAlO/0.5 nm SiO₂/Si stack prepared by various processing conditions. Different Al doping, different post deposition annealing temperatures and different deposition steps were discussed in this chapter. EOT and $V_{FB}$ were obtained from capacitance-voltage measurements. The correcting of the error introduced by the series resistance ($R_s$) associated with the substrate and contact is also discussed here.

Chapter 5 discusses the dry and wet processed interface layer properties for three different p type Ge/ALD 1nm-$\text{Al}_2\text{O}_3$/ALD 3.5nm-$\text{ZrO}_2$/ALD TiN gate. Several parameters like EOT, flat band voltage, bulk doping, and surface potential as a function of gate voltage are reported. It is also discussed that the high frequency capacitance of TiN/$\text{ZrO}_2$/Al₂O₃/p-Ge gate stacks measured in the accumulation region depends on the device area after substrate resistance correction.

Chapter 6 deals with the TiN/$\text{ZrO}_2$/Al₂O₃/p-Ge gate stacks subjected to the different SPAO (Slot Plane Antenna Oxidation) annealing conditions, namely, (i) before
high-k ALD (Atomic Layer Deposition), (ii) between high-k ALD, and (iii) after high-k ALD. After XPS (X-ray Photoelectron Spectroscopy) and EOT (estimated by capacitance voltage) measurement. The carrier transport mechanisms on these samples were extracted at high field range to reveal how SPAO can effective remove traps in high-k layer. The reliability of interfacial layer (GeOₓ/GeO₂) is evaluated by TDDB performance under substrate electron injection condition.

Chapter 7 characterizes time zero 1/f noise on different oxide thickness of thick and thin oxide FinFET transistors to study how EOT impacts the noise mechanism. At room temperature, the flicker noise mechanisms were determined at linear region to study the bias dependence of FinFET. The noise models were subsequently extracted. The impact of temperature on the noise level and noise mechanism was discussed as well.

Chapter 8 summarized the overall work of this research and discusses future work.
CHAPTER 2
LITERATURE REVIEW: HIGH-K MATERIAL AND ITS APPLICATION ON Si/Ge SUBSTRATE

In this chapter, historical requirement of implementation of high-k dielectric will be reviewed. The basic requirement is to improve the high-k and interface property if high-k can be successful implemented on silicon substrate. The current status of high-k dielectrics will be discussed in relation to different methods to further scale EOT below 0.7 nm. Finally, the challenge of enhancing the properties of high-k dielectrics and its application on germanium channel is discussed.

2.1 Requirement of High-k and the Necessary Property of Implementation

The first point-contact transistor was invented by Bardeen, Brattain, and Shockley invented in 1947. In acknowledgement of this accomplishment, Shockley, Bardeen, and Brattain were jointly awarded the 1956 Nobel Prize in Physics “for their researches on semiconductors and their discovery of the transistor effect.” [15]. Interestingly, the first transistor was germanium. However, researchers discovered that material property of Si, which is same as in group IV as Ge, is better than Ge, when considering the formation of gate oxide: SiO$_2$ can be grown on its surface, which is a very good insulator. On the other hand, Ge does not form this oxide layer on its surface so easily and GeO$_2$ is hydroscopic and not thermally stable [16]. Additionally, Si is abundant, for example, sand is a source of Si.

In 2004, EOT (oxide thickness calculated by using dielectric constant of SiO$_2$) was scaled to ~1.2 nm [1] as Figure 2.1 indicates, as oxide layer was the first layer went
to sub 1 nm range. However, ultrathin SiO$_2$ suffers from direct-tunneling current which increases exponentially as thickness decreases [17]. The large gate leakage current exhausts battery rapidly. To overcome gate leakage problems, addition of N into SiO$_2$ has been used either by post deposition annealing in nitrogen ambient or forming a nitride/oxide stack structure. As incorporating nitrogen into SiO$_2$, it not only increases the dielectric constant but also acts as a better barrier preventing boron penetration from poly silicon gate. SiON served as a transition stage between high-k and SiO$_2$, which has maxim dielectric constant less than 8 [18].

Figure 2.1 Trend of EOT of the gate insulator for various versions of ITRS.

Source: [1].

High-k materials were first studied on memory device. Table 2.1 listed high-k
candidates and their relevant properties. Before it can be implemented in industry, the following issues has to be considered first: (a) permittivity, bandgap, and band alignment to silicon, (b) thermodynamical stability, (c) film morphology and deposition method, (d) interface quality and bulk defects, (e) gate compatibility and process compatibility [3].

### Table 2.1 Comparison of Relevant Properties for High-k Candidates

<table>
<thead>
<tr>
<th>Material</th>
<th>Dielectric constant</th>
<th>Band gap $E_G$ (eV)</th>
<th>$\Delta E_c$ (eV) to Si</th>
<th>Crystal Structure(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO$_2$</td>
<td>3.9</td>
<td>8.9</td>
<td>3.2</td>
<td>Amorphous</td>
</tr>
<tr>
<td>Si$_3$N$_4$</td>
<td>7</td>
<td>5.1</td>
<td>2</td>
<td>Amorphous</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>9</td>
<td>8.7</td>
<td>2.8$^a$</td>
<td>Amorphous</td>
</tr>
<tr>
<td>Y$_2$O$_3$</td>
<td>15</td>
<td>5.6</td>
<td>2.3$^a$</td>
<td>Cubic</td>
</tr>
<tr>
<td>La$_2$O$_3$</td>
<td>30</td>
<td>4.3</td>
<td>2.3$^a$</td>
<td>Hexagonal, cubic</td>
</tr>
<tr>
<td>Ta$_2$O$_5$</td>
<td>26</td>
<td>4.5</td>
<td>1-1.5</td>
<td>Orthorhombic</td>
</tr>
<tr>
<td>TiO$_2$</td>
<td>80</td>
<td>3.5</td>
<td>1.2</td>
<td>Tetr.($^c$) (rutile, anatase)</td>
</tr>
<tr>
<td>HfO$_2$</td>
<td>25</td>
<td>5.7</td>
<td>1.5$^a$</td>
<td>Mono.$^b$, tetr.$^c$, cubic</td>
</tr>
<tr>
<td>ZrO$_2$</td>
<td>25</td>
<td>7.8</td>
<td>1.4$^a$</td>
<td>Mono.$^b$, tetr.$^c$, cubic</td>
</tr>
</tbody>
</table>

$^a$Calculated by Robertson. [19]  
$^b$Mono.=$^b$monoclinic.  
$^c$Tetrag.=$^c$tetragonal.

Source: [3].

### 2.1.1 Permittivity, Bandgap, and Band Alignment to Silicon

In order to have a good insulating property, it is suggested that conduction band offset (CBO) between high-k and substrate should be larger than 1 eV to inhibit Schottky emission, and it is same for valence band offset as well. Considering different work function between substrate and the high-k dielectric, specifically a bandgap of 4 eV is necessary to avoid serious leakage current and breakdown. Figure 2.2 shows a relationship between dielectric constant and bandgap. Unfortunately, large K material like TiO$_2$ suffers from a low bandgap problem, and its large dielectric constant will give extra
fringe capacitance between gate and source/drain, which is not desirable. In Figure 2.2, it suggests that ZrO$_2$ and HfO$_2$ are good candidates since they have bandgap larger than 4 eV and its dielectric constant is still large enough for further EOT scaling.

![Figure 2.2 Static dielectric constant versus band gap for candidate gate oxides.](image)

Source: [19].

### 2.1.2 Thermodynamic Stability

In gate first CMOS processes, the gate stacks must undergo rapid thermal annealing (RTA) at 1000 °C for 5s. This requires that the gate oxides must be thermally and chemically stable with the contacting materials [20]. From this point of view, HfO$_2$ has better thermal stability than ZrO$_2$ [21]. Additionally, as ZrO$_2$ and HfO$_2$ thin films were grown by atomic layer deposition, the structural and electrical behavior of the films were somewhat precursor-dependent, revealing better insulating properties in the films grown from oxygen-containing precursors, therefore the HfO$_2$ films showed lower leakage.
compared to ZrO$_2$ [22].

### 2.1.3 Film Morphology and Deposition Method

It is desirable to have an amorphous high-k layer after necessary processing treatments due to several benefits of amorphous structure. Polycrystalline gate dielectrics are not favored as gate oxide layer since grain boundaries serve as high-leakage paths. In addition, grain size and orientation can cause significant variation in $K$ value. However, most high-k materials except Al$_2$O$_3$ will form polycrystalline film either during deposition or after thermal treatment if temperature is high enough. HfO$_2$ or ZrO$_2$, crystallize at much lower temperatures at $\sim$400 °C and $\sim$300 °C, shown in Figure 2.3 [23].

![TEM image of crystallization in HfO$_2$/SiO$_2$ dielectrics with (a) 40% HfO$_2$ and (b) 80%. Source: [23].](image)

**Figure 2.3** TEM image of crystallization in HfO$_2$/SiO$_2$ dielectrics with (a) 40% HfO$_2$ and (b) 80%.

Source: [23].

The crystalline temperature of dielectrics can be increased by incorporating other impurities, which is first studied by van Dover [24]. TiO$_x$ dielectric with addition of Nd,
Tb and Dy dopants shows lower leakage current and comparable permittivity [24]. In chapter 4, HfO$_2$ was incorporated by aluminum to achieve an amorphous dielectric layer. Deposition methods for high-k film can be categorized into two major approaches based on the reaction mechanism during preparation, namely CVD (chemical vapor deposition) and PVD (Physical Vapor Deposition) processes. CVD-based approaches include metal-organic chemical vapor deposition (MOCVD), plasma-enhanced chemical vapor deposition (PECVD), atomic-layer chemical vapor deposition (ALCVD or ALD) [6] and so on. Among them, ALD is the most popular method for growing high-k layer as it is the only method that can control thickness accurately in nanometer range.

2.1.4 Interface Quality and Bulk Defects

Interface between high-k and substrate must have highest electrical property, flatness, and low interface state density $D_{it}$. Fixed charges present at the interface can cause flat band voltage shift. Large $D_{it}$ degrades mobility by surface scattering mechanism. Most of the high-k materials reported $D_{it}$ range from $10^{11}$ to $10^{12}$ cm$^{-2}$eV$^{-1}$ [25, 26], which is much higher than conventional thermal grown SiO$_2$ [27]. Interface treatment, therefore is necessary before depositing high-k layer to obtain a low $D_{it}$ interface, of which thickness is usually about two atomic layer of SiO$_x$. On the contrary, overall EOT value is strongly depend on thickness of interfacial layer. The thickness of interfacial layer will impact overall mobility of carriers in the channel as shown in Figure 2.4.
Figure 2.4 (a) TEM cross section of a high-k/metal gate stack where the BIL (bottom interfacial layer) was removed by post anneal technique (b) High field electron mobility degradation as a function of EOT from literature data. The mobility values are taken at at $E_{\text{eff}} = 1 \text{ MV/cm}$. 

Source: [28].

Similar to interface defects, bulk defects formed in high-k oxides during deposition also causes degraded transistor performance and it is reported that high-k materials is intrinsically defective since their bonding cannot relax easily [29]. The bonding in high-k oxides is ionic. The nature of intrinsic defects in ionic oxides differs from those in SiO$_2$. The oxygen vacancies, oxygen interstitials, or oxygen deficiency defects due to possible multiple valence of the transition metal [29]. Oxides can be of significant problem. Moreover, high-k oxides achieve their high-k value because of the low-lying soft polar modes. These modes could be a limit on scattering, which doesn’t happen in SiO$_2$ [30]. Figure 2.5 compares mobility in MOSFET with different dielectrics.
2.1.5 Gate Compatibility and Process Compatibility

Conventional polysilicon as gate material doesn’t work properly when high-k is incorporated. This is due to Fermi level pinning problem which did not allow the p-type poly Si gate to work effectively [31]. The solution came with the introduction of the metal gate such as TiN. A metal layer was inserted to mitigate above mentioned issues by inhibiting the formation of direct interface between the high-k and poly-silicon. This structure was later named as “metal inserted poly silicon” (MIPS) and became the foundation of the high-k/metal gate stack research. MIPS allowed the industry to maintain the legacy self-aligned nature of the transistor fabrication process flow (Ion implantations, silicide) stemmed from the etching of the gate poly Si [28].
Metal gates such as TiN and Pt have been used as gate metal to prevent reaction at the gate interface (gate/high-k). TiN is called as midgap metal, as its work function places at the mid of Silicon bandgap as Figure 2.6(a) indicates. This approach will have a symmetric threshold voltage $V_{th}$ for both n and p MOSFET at a cost of large $V_{th} \sim 0.5$ V, which is too large for low power and new sub 100 nm technology. On the other hand, dual metal approach is not feasible. Al is not a feasible electrode metal for nMOS because it will reduce nearly any oxide gate dielectric to form an Al$_2$O$_3$-containing interface layer. For pMOS, Pt is not a practical choice for the gate metal, since it is not easily processed, and does not adhere well to most dielectrics [3].

Moreover, the band bending in the silicon channel was not fully controlled by the metal gate’s intrinsic work function. For this reason, the “effective work function” (EWF) concept replaced the “vacuum work function”. To discover proper EWF for nMOS and pMOS, respectively, it was discovered that EWF of TiN can be modified by Al diffusion.
into TiN as shown in Figure 2.7 [28].

![Figure 2.7 EWF change with anneal temperature. Four different compositions of Al based n-metals deposited on thin TiN layer followed by barrier layer and Al-fill. Source: [28].](image)

**Figure 2.7** EWF change with anneal temperature. Four different compositions of Al based n-metals deposited on thin TiN layer followed by barrier layer and Al-fill.

Source: [28].

### 2.2 Interfacial Layer Scavenging

Hf-based high-k material is the most popular material in industrial now. HfO$_2$ has a dielectric constant around 20. The EOT for the first generation HKMG device is approximately 1.0 nm. However, sub-nm EOT is required at those advanced technology nodes beyond 7 nm. A typical HKMG stack structure contains a silicon oxide based interfacial layer (IL), a high-k dielectric, followed by a metal gate electrode as shown in Figure 2.8. The high-k layer and interfacial layer are equivalent to two capacitors in serial combination. Thus, the total EOT of the HKMG stack can be expressed as Equation (2.2).
\[ EOT = EOT_{high-k} + EOT_{IL} \]  \hspace{1cm} (2.2)

![Diagram of metal high-k gate stack.](image)

**Figure 2.8** Schematic of metal high-k gate stack.

By understanding of Equation (2.2), it is expected that reducing thickness of high-

k layer can simply give a lower EOT. However, it has been reported that a critical

thickness of Hf-based oxide, and mobility degradation becomes a serve problem when

the thickness is below 3.5 nm [32]. The above discussion leaves researcher with other

three approaches to reduce EOT value: (i) search other high-k material, with larger value;

(ii) increase dielectric constant of interfacial layer; (iii) reduce the thickness of interfacial

layer;

Since a serial combination of two capacitors, the overall capacitance will be

limited by smaller one. It is apparent that approaches dealing with interfacial layer will be

practical to reduce EOT. In this section, the technique named interfacial layer scavenge

will be discussed, the aim of which is to boost dielectric constant of interfacial layer. The

technique can be used widely, but HfO$_2$/SiO$_2$/Si gate stack will be used as an example.

Most process grows a thin quality SiO$_2$ on the Si substrate to achieve a better interface
before depositing high-k HfO$_2$. When EOT comes below 1 nm, the thickness of IL became the main challenge, since SiO$_2$ has low dielectric constant compared to HfO$_2$. The idea of scavenge is to replace the IL by subsequent annealing process with a different IL composition (higher K value material). Equation (2.2) is a specific scavenging reaction [33], Al is the chosen scavenging element.

$$2Al + 1.5SiO_2 \rightarrow Al_2O_3 + 1.5Si$$  \hspace{1cm} (2.2)

**Figure 2.9** Bright field micrograph of the gate stack from STEM. The interface to Si/HfO$_2$ shows a weak contrast. Al and Si were found by EELS above TiN. Inset: Schematic gate stack structure and Al scavenging process. During silicidation Al distributes uniformly into poly-Si. In between NiSi and TiN, SiO$_x$ and AlO$_x$ accumulation is found.

Source: [33].
As Figure 2.9 shows, during the annealing process, scavenging element Al can diffuse throughout metal gate and high-k layer and finally reach with IL layer (SiO$_2$). Al$_2$O$_3$ was formed after reaction, therefore overall EOT is increased since Al$_2$O$_3$ has a larger dielectric constant than SiO$_2$. The inset figure of Figure 2.9 shows Al layer was capped on the top of the gate stack. This kind of approach is called remote scavenging technique. On the other hand, for direct scavenge, scavenging metals are incorporated within the high-k layers either by direct capping like structure presented in Chapter 4. Both techniques are visualized in Figure 2.10. The advantage of direct scavenge is that it does not require such a strong driving force of the reaction [34]. The drawback of this approach is as following [34]: (i) Effective work function change either by inherently low vacuum work function of scavenging metals or by formation of fixed charges and/or interface dipoles; (ii) excessive carrier mobility degradation and leakage current increase; and (iii) increase in $D_{it}$. 
It is possible to cause IL scavenging reaction without out-diffusion of the scavenging elements into the high-k layer. Such a process enables EOT scaling without extrinsic degradation in carrier mobility and leakage current and with no change in EWF. If Al layer was capped on the top of the metal layer like Figure 2.9 and among the possible metal options for EOT scavenging, Al has the ability to reduce TiO$_2$ and SiO$_2$ leaving HfO$_2$ unreacted [33]. As will be presented in Chapter 4, Aluminum incorporated in HfO$_2$ can reduce EOT at the cost of larger interface state density, which is the consequence of increased dielectric constant of IL. However, the results were attributed to dielectric constant of HfO$_2$ after subtracting IL thickness (assume IL is SiO$_2$). This is because it is not possible to measure dielectric constant of IL. Reader should refer to both explanations (either EOT or dielectric constant of HfO$_2$) at his/her favor with a scavenging technique concept [35].
But what is the minimum value of EOT that can be achieved in the future? The state of art of EOT of HF-base dielectrics is reduced to as low as 0.42 nm [36]. However, it also brings the following question to the table: what if there is no interfacial layer or what is the relationship between thickness of IL and carrier mobility in channel. As Figure 2.1 indicates, EOT or thickness of interfacial layer are somewhat correlated. Tatsumura studied how EOT impacts mobility [37]. The scavenging technique should optimized EOT value without server degradation on carrier mobility. Extreme IL scaling (zero-IL) ends up with loss of EWF control and penalty in reliability issues like poor TDDB and BTI performance.

### 2.3 High-k Application on Ge Channel

As mentioned earlier, the hydroscopic property of GeO$_2$ hindered development of Ge transistor. The development of high-k material reopened the door to Ge transistor, since a GeO$_2$ is not necessary to be presented as the gate dielectrics. Research had begun on Ge channel in 2000’s again because its hole mobility is four times as high as that of silicon and its electron mobility is twice as high as that of silicon. Table 2.1 shows material characteristics of alternative channel materials for Ge and Si. Additionally, the lattice constant of Ge is close to that of GaAs is expected to facilitate integration of III-V n-MOSFETs (GaAs) and optical devices on Ge substrates in the future [38]. However, Ge nMOSFET has not been implemented successfully as predicted electron mobility due to large density of interface states [14]. Some researchers believe this mobility degradation is due to the degraded Ge interface and is inherent to Ge [39], and Ge could be used only in pMOSFETs. Figure 2.11 shows that metal/p-Ge has fermi level pinning problem,
which is not only a problem for high-k/p-Ge, but also it is problematic for source/drain formation as well.

Although HfO$_2$ is widely used in Si system, it is not a good selection for Ge system. HfO$_2$ is unsuitable on a Ge substrate, since gate leakage current density is larger than ZrO$_2$ [40]. Ge can diffuse into the HfO$_2$ layer that results in the increase in gate leakage current if no appropriate interfacial layer presents. Nevertheless, ZrO$_2$, which was screened out because of low thermal stability with Si, is a good candidate for replacement metal gate integration since the thermal budget is greatly reduced compared to gate first integration. ZrO$_2$/Ge gate stacks can sustain and improve its electrical characteristics after annealing [40]. Figure 2.12 illustrates effect of annealing on ZrO$_2$/Ge and HfO$_2$/Ge gate stack.
Table 2.2 Material Characteristics of Alternative Channel Materials for Ge and Si

<table>
<thead>
<tr>
<th>Material characteristics of alternative channel materials</th>
<th>Ge</th>
<th>Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandgap, $E_g$(eV)</td>
<td>0.66</td>
<td>1.12</td>
</tr>
<tr>
<td>Electron affinity, $\chi$(eV)</td>
<td>4.05</td>
<td>4</td>
</tr>
<tr>
<td>Hole mobility, $\mu_h$(cm$^2$V$^{-1}$s$^{-1}$)</td>
<td>1900</td>
<td>450</td>
</tr>
<tr>
<td>Electron mobility, $\mu_e$(cm$^2$V$^{-1}$s$^{-1}$)</td>
<td>3900</td>
<td>1500</td>
</tr>
<tr>
<td>Effective density of states in valence band, $N_V$(cm$^{-3}$)</td>
<td>$6.0 \times 10^{18}$</td>
<td>$1.04 \times 10^{19}$</td>
</tr>
<tr>
<td>Effective density of states in conduction band, $N_C$(cm$^{-3}$)</td>
<td>$1.04 \times 10^{19}$</td>
<td>$2.8 \times 10^{19}$</td>
</tr>
<tr>
<td>Lattice constant, $a$(nm)</td>
<td>0.565</td>
<td>0.543</td>
</tr>
<tr>
<td>Dielectric constant, K</td>
<td>16</td>
<td>11.9</td>
</tr>
<tr>
<td>Melting point, $T_m$(°C)</td>
<td>937</td>
<td>1412</td>
</tr>
</tbody>
</table>

Source: [38].
Figure 2.11 Schottky barrier heights obtained experimentally for various metals with different vacuum work-functions. In case of Ge, the Fermi-level is strongly pinned near the valence band edge.

Source: [41].

To replaced Si by Ge for in future CMOS technology, researchers must find best passivation method for Ge to reduce density of interface states. Available surface passivation methods includes: epi-Si passivation, surface oxidation and/or nitridation, and S-passivation. Among these, plasma-based processed surface passivation followed by plasma-enhanced ALD for high-k layer showed highest gate stack quality [42]. To achieve ultimate scaling, another solution is to deposit dielectrics directly on Ge without incorporation of an IL, which typically has a much lower capacitance value than
expected. This is because, generally, an IL either intentionally or unintentionally formed during the high-k dielectrics deposition process or during post-deposition annealing process [41]. Nevertheless, Ming Lin et al. reported a 0.39 nm EOT with ultrathin GeON formed by remote plasma treatment, with a $D_{it}$ of $4 \times 10^{12}$ cm$^{-2}$eV$^{-1}$ [43].

![Image](image.png)

**Figure 2.12** Schematic showing the effect of annealing on ZrO$_2$/Ge gate stack (a, b) and HfO$_2$/Ge gate stack (c, d), before (a, c) and after (b, d) annealing at 500°C. HfGe$_2$ in (d) is speculation and has not been observed.

Source: [38].

Recently, GeO$_2$ passivation layer has been reconsidered as promising passivation layer due to its low $D_{it}$ ($6 \times 10^{10}$ cm$^{-2}$eV$^{-1}$) [44]. As mentioned earlier, GeO$_2$ is undesirable
because it is hygroscopic and water soluble. GeO$_2$ is thermally unstable and converts to volatile GeO$_x$ at approximately 430 °C. However, it is found that it is not necessary to form GeO$_2$, instead, GeO$_x$ can give a promising $D_{it}$ value and it is more stable [45, 46] as it can be controlled by post annealing process.

Since GeO is volatile [16], it is necessary to have a layer which can effectively prevent GeO volatilization and GeO growth via retarding the inter diffusion of Ge and O atoms. If GeO$_x$ is not passivated by other element like nitrogen, sulfur, silicon. Al$_2$O$_3$ can be considered as first oxide layer if EOT is not aggressively scaled below 0.7 nm as it has lower intrinsic oxygen permeability [45]. Besides, Al$_2$O$_3$ has a larger bandgap and conduction band offset, which helps effectively block electron injection from Ge substrate [47]. Moreover, Houssa et al. used first principle to calculate interface property, and found that Al-O-Ge bond tends to give a surface states free bandgap [48]. In this dissertation, a bilayer structure 1 nm Al$_2$O$_3$/3.5 nm ZrO$_2$ is used for Ge gate stack.

In summary, although the literature reported a low $D_{it}$ as $10^{11}$ cm$^{-2}$eV$^{-1}$, Ge nMOSFET still suffers low electron mobility and it can achieve approximately achieve a maximum 1.5 times of that of silicon. The underlying mechanism of low electron mobility is studied in progress and needs to be addressed before it can be deployed in future CMOS technology.
CHAPTER 3

MOS CAPACITOR DEVICE MODELLING AND ELECTRICAL CHARACTERIZATION METHOD

3.1 MOS Capacitor Device Modeling

In this section, the electrical model of MOS capacitor will be reviewed and discussed in order to explain how the electrical performance interacts with physics of MOS devices and its defects. Therefore, the electrical properties of devices reflects the parameters of defects and dielectrics. The information provided in this chapter lays the fundamental technique, which will be utilized in the following chapters. MOS capacitor (MOS-C) can be mathematically developed by Poisson equation in one dimension. Figure 3.1(a) shows the physical cross section of a MOS-C, while Figure 3.1(b) is band diagram of a traditional SiO₂/Si system.
**Figure 3.1** (a) Cross section of a MOS capacitor. (b) Energy-band diagram of the MOS capacitor showing the energy barrier between metal and SiO$_2$ and between silicon and SiO$_2$. The metal is aluminum and the silicon is p-type.

Source: [49].

When gate voltage is applied on gate related to substrate (Silicon), a charge density forms on the two side of oxide which is strongly dependent on the voltage on the substrate ($\phi_s$), which can be accumulation, depletion, or inversion. These nominations are corresponding to the carrier density close to the oxide/substrate interface. Accumulation is a condition when majority carrier density at interface is larger than that of bulk. On the contrary, inversion is a condition when minority carrier density is larger at the interface than that of majority carrier in the bulk. Depletion by its name means majority carriers are depleted at the surface, and only ionized charges are left. The carrier concentration at the interface can be given by:
\[ n = n_i \exp\left(\frac{E_F - E_i}{kT}\right) \]  
(3.1a)

\[ p = n_i \exp\left(\frac{E_i - E_F}{kT}\right) \]  
(3.1b)

\[ n = N_D, p = N_A \]  
(3.2)

where \( n \) is electron density, \( p \) is hole density, \( n_i \) is carrier density if silicon is intrinsic, \( E_F \) is energy level of Femi level, \( E_i \) is the intrinsic Fermi level, \( k \) is Boltzmann constant, \( T \) is temperature. When silicon is doped by either donor or acceptor, the fermi level will move either towards to conduction band or valence band to accommodate extra electrons or holes in the system. The original Fermi level is dependent on the doping concentration and dopants type, as long as silicon is not degenerate type, Equation (3.1) and Equation (3.2) hold.

When the system has a voltage change, the Fermi level related to band diagram is changed as shown in Equation (3.3a), where \( x \) is measured from bulk where the electric field is screen out by charge (no electric field). Voltage in the silicon substrate and charge density has to follow the Poisson equation as shown in Equation (3.3b), where \( \phi(x) \) is voltage at specific location \( x \), \( \rho(x) \) is charge density which is a combination of electrons, holes, ionized dopants \( (N_D^+ \text{ or } -N_A^-) \) as Equation (3.3c) shows.

\[ \phi(x) = E_F(x) - E_F(0) \]  
(3.3a)

\[ \frac{d^2\phi(x)}{dx^2} = -\frac{\rho(x)}{\varepsilon_s} \]  
(3.3b)

\[ \rho(x) = (-qn) + (qN_D^+) + (qp) + (-N_A^-) \]  
(3.3c)
Since the electric field cannot penetrate metal too much as substantial number of electrons is available in metal, and the voltage consumed in metal can be neglect for convenience. Electric field in oxide is constant and it has to follow Equation (3.4),

\[ C = \frac{Q}{V} \]  \hspace{1cm} (3.4)

\[ V_g = V_{ox} + \phi_s \]  \hspace{1cm} (3.5)

where \( C \) is the capacitance of dielectric and \( Q \) is the total charge in the silicon. Note this amount of charge is exactly same on the metal as well. Therefore, when there is a gate voltage applied on the SiO\(_2\)/Si system, the final potential profile and electric profile are determined by Equations (3.1-3.5), which is shown in Figure 3.2.
Figure 3.2 (a) Electrical field configuration in the SiO$_2$ and silicon surface calculated from Equations (3.1-3.5) as function of $x$. (b) Potential distribution in the SiO$_2$ and the silicon surface calculated from Equations (3.1-3.5) corresponding to the electric field configuration in (a).

Source: [50]

The beauty of the MOS structure is that carrier concentration of the surface of semiconductor is a function of surface potential, and therefore a function of gate voltage
(\(V_g\)), which enables MOS works like a switch. Unlike metal, semiconductor has both electron and hole carrier and the density of them is subjected to bias voltage. When MOS-C is measured for its capacitance, small signal is more powerful since the voltage-charge characteristic of semiconductor is no longer linear like dielectric, as Equation (3.1-3.3) predicts. It strongly depends on surface potential, \(\phi(x)\). Therefore, it is more convenient to measure small signal capacitance of MOS structure rather than large signal. Capacitance of oxide \(C_{ox}\) is not voltage dependent, on the other hand, the capacity of substrate can be defined as \(C_s\), which is voltage dependent and cannot be measured directly. The equivalent circuit is shown in Figure 3.3.

![Figure 3.3 Cross section of a MOS capacitor showing a simple equivalent circuit. The capacitor labeled \(C_s\) is shown variable to denote its bias dependence.](source)[51].
Figure 3.4(a) shows a theoretical small signal capacitance voltage measurement of p type MOS-C. It has three parts, accumulation, depletion and inversion. When surface potential is large enough to bend the bands of Si to supply carriers, it can supply carriers like metal. The overall capacitance of MOS-C converges to $C_{ox}$. When band diagram of Si is biased to deplete holes, $C_s$ is provided by ionized charge in the substrate before substrate can be inverted. The above discussion has an assumption that device under test is quasi-static, and it is only valid when testing frequency is low enough. For majority carrier, carrier response time is the relaxation time of the semiconductor, and is generally in the range of pico seconds. Specifically, for a typical range of measured frequency, $C_s$ is not frequency dependent in accumulation region. However, the minority carrier response time in silicon depends upon the generation rate, and it can be lower than 1 Hz in undoped silicon substrate [52].

The interface between substrate and dielectric is not perfect as described in the ideal model because of dangling bonds as interface defects. These defects can response to small signal and therefore contribute to additional capacitance, which is parallel to $C_s$. It is also voltage and frequency dependent, which is named as $C_{it}$. $C_{it}$ at interface is easier to be measured in depletion region since capacitance in depletion region is very small as Figure 3.4(a) shows, and this additional $C_{it}$ is distinct enough to be detected. Since defect has a response time (time constant $\tau$) like minority carrier, it is mathematically modeled as a combination of resistor and capacitor in series to consider the energy loss at high frequency. Recently, there is observation of frequency dispersion at accumulation as well for thin dielectrics (EOT ~3 nm) in addition to gate tunneling current modeled as $G_t$, which can be attribute to the border trap [53]. Figure 3.5 shows a schematic modification
of equivalent circuits for conductance measurement. Figure 3.5(d) includes the contribution from border trap and edge capacitance, which are $C_{BT}$ and $C_L$, respectively.

**Figure 3.4** To illustrate the operation of a MOS capacitor, capacitance as a function of bias, and the corresponding energy-band diagrams are shown in accumulation, depletion, inversion, and at flatbands. Energy values $E_c$ and $E_v$ are the conduction and valence band edges, respectively.

Source: [54].
Figure 3.5 Equivalent circuits for conductance measurements; (a) MOS-C with interface trap time constant $\tau_{it} = R_{it}C_{it}$, (b) simplified circuit of (a), (c) measured circuit, (d) including series $r_s$ resistance and tunnel conductance $G_t$, capacitance and resistance from border trap $C_{BT}, R_{it}$, capacitance from side of measured area $C_L$.

3.2 Electrical Characterization Method

In this section, all the measurement method used in the work is discussed. For capacitance based technique, low frequency CV, CV hysteresis, conductance method (capacitance data is necessary), DLTS is reviewed. IV measurements, flick noise, IV as function of temperature and TDDB are also addressed.

3.2.1 Low Frequency CV

It is also called as quasi-static CV method or high/low frequency method, which is
developed by Berglund in 1966 [55]. This method is popular for its simple of implementation. Capacitance data at high and low frequency is used to calculate interface state density. This model is based on the following assumption: interface states cannot follow high frequency signal (normally 1 MHz), and they will follow low frequency signal (as low as possible, but 100 Hz is enough). Therefore, following Equation (3.6) is valid, and can be used to extract \( D_{it} \). Notice that if a sufficiently low frequency is used, there is no energy loses in the MOS system and resistor component can be neglected as shown in Figure 3.5. For high frequency, defects will manifest themselves as conductance rather than capacitance as Equation (3.6b). Combining Equation (3.6a) and Equation (3.6b) gives Equation (3.6c) where \( D_{it} \) is proportional to \( C_{it} \) as Equation (3.6d) indicates (every defects is capturing one electron). Some literatures use \( q^2 \) instead of \( q \) in Equation (3.6(c-d)). The difference is the understanding of the units, which is explained by Schrodor [56]. \( D_{it} \) is usually extracted in depletion region since \( C_s \) would be too large, and measured capacitance would be converged to \( C_{ox} \). The resolution of low frequency method is about \( 10^9 \) cm\(^{-2}\)eV\(^{-1}\).

\[
\frac{1}{C_{lf}} = \frac{1}{C_{ox}} + \frac{1}{C_s + C_{it}} \quad (3.6a)
\]

\[
\frac{1}{C_{hf}} = \frac{1}{C_{ox}} + \frac{1}{C_s} \quad (3.6b)
\]

\[
D_{it} = \frac{C_{ox}}{q} \left( \frac{C_{lf} / C_{ox}}{1 - C_{lf} / C_{ox}} - \frac{C_{hf} / C_{ox}}{1 - C_{hf} / C_{ox}} \right) \quad (3.6c)
\]

\[
D_{it} = qC_{it} \quad (3.6d)
\]
Figure 3.6 Low frequency method applied on HfAlO/Si devices. $D_{it}$ is calculated in the range where the pointer indicates (depletion region). The dispersion in the negative range (accumulation region) is due to border traps.

Figure 3.6 shows a typical low frequency measured of MOS device, the $D_{it}$ information can be calculated in the depletion region. How low the frequency can be is dependent on leakage current level of dielectric since leakage current can affect the reading of capacitance [57].

3.2.2 CV Hysteresis

CV hysteresis shown as Figure 3.7 is also a very simple technique to have quick evaluation of the dielectrics quality. Compare to low frequency CV, it can only measure
slow interface states and traps in oxide layer. It cannot extract fast interface stats information since they can follow to the DC sweep. Additionally, hysteresis measurement doesn’t consist small signal voltage information, therefore, it has different unit to present information of defects density. It has a symbol of $N_{it}$ (cm$^{-2}$) rather than $D_{it}$ (cm$^{-2}$eV$^{-1}$) to distinct them. The measurement is necessary to be taken at high frequency to remove the impact of capacitance values (vertically) from slow states.

![CV hysteresis](image)

**Figure 3.7** CV hysteresis are taken by sweep voltage from accumulation to inversion, and sweep back immediately.

The hysteresis value can be taken at their largest value or the difference of one reference value like $V_{FB}$. $N_{it}$ is proportional to this $\Delta V_{FB}$ as shown in Equation (3.7).
Moreover, \( N_{it} \) is representing overall defects density in the bandgap. And hysteresis measurement is strongly depending on voltage ramping speed and this is the disadvantage of this method.

\[
N_{it} = \frac{\Delta V_{FB} \times C_{ox}}{q}
\]  

(3.7)

3.2.3 Conductance Method

The \( D_{it} \) values can also be calculated by the conductance method as well [8]. It is known that interfacial defects do not immediately respond to the small signal change, but can be represented by a phase lag behind the signal. This can be modeled as a combination of resistor and capacitor in series, and their product is equal to interface time constant \( (\tau_{it}) \). \( \tau_{it} \) has the relationship with the physical model as shown in Equation (3.8),

\[
\tau_{it} = R_{it}C_{it} = f_o(c_p n_s)^{-1}
\]  

(3.8a)

\[
C_{it} = \frac{q^2}{KT} n_T f_o(1 - f_o)
\]  

(3.8b)

\[
\frac{1}{R_{it}} = G_{it} = \frac{q^2}{KT} n_T (1 - f_o)c_p n_s
\]  

(3.8c)

\[
f_o = \frac{1}{1 + g \times \exp[(E_T - E_F)/KT]}
\]  

(3.8d)

\[
C_{it} = q^2 D_{it}, q^2 = 1 \times 1.6 \times 10^{-19}(Coul^2)
\]  

(3.8f)

where \( g \) is degeneracy factor, \( E_T - E_F \) is the energy level difference between the trap energy level and the Fermi level, \( K \) is Boltzmann constant, \( T \) is temperature in kelvin. Capture
probability ($c_p$) is the product of defect cross-section and thermal velocity. $n_s$ is the majority carrier concentration in the depletion region. $n_T$ is the interface defect density. Equation (3.8(a-c)) shows capacitance contribution due to interface states, $C_{it}$, is almost constant and is proportional to $n_T$. In Equation (3.8d) it is expected that $f_o$ is almost constant, if it is assumed only defects near the Fermi level respond. For a shorter time constant $\tau_{it}$, $n_s$ is larger by shifting the corresponding Fermi level closer to the valence band edge. The interface conductance, $G_{it}$, on the other hand, is affected by the majority carrier density. The relationship between $C_{it}$ and defects density is mathematically demonstrated in Equation (3.8(f)) [8]. E.H. Nicollian et al. [8] suggested that it is easier to obtain $D_{it}$ in the depletion region. Since the mid gap conductance is influenced by the generation-recombination and weak inversion is influenced by the bulk loss, their physical model is more complicated to extract.

Conductance method relies on both CV and GV characteristics. In small area devices like 40 $\mu$m $\times$ 40 $\mu$m, the series resistance ($R_s$) introduces the frequency dispersion of both the measured capacitance and conductance. Applying a simple approach can exclude the effect of $R_s$ [58]. After obtaining the CV and GV data, the series resistance ($R_s$) can be calculated by using Equation (3.9) in the strong accumulation region at 1MHz. The original measured capacitance ($C_m$) and conductance ($G_m$) values can be subsequently corrected to exclude the effect of $R_s$ by using Equation (3.10). Additionally, without the knowledge of the impact of tunneling leakage current ($G_t$), the uncertainty of the conductance measurement can cause errors in $D_{it}$ estimation [59]. The new corrected conductance in the accumulation region can be obtained by subtracting $G_t$ from $G_C$. To obtain $G_t$, IV is measured at slow ramping voltage and calculated the derivative of IV as
shown in Figure 3.8.

\[ R_s = \frac{G_{ma}}{G_{ma}^2 + \omega^2 C_{ma}^2} \]  \hspace{1cm} (3.9)

\[ C_C = \frac{(G_{m}^2 + \omega^2 C_{m}^2) C_{m}}{a^2 + \omega^2 C_{m}^2} \]  \hspace{1cm} (3.10a)

\[ G_c = \frac{(G_{m}^2 + \omega^2 C_{m}^2) a}{a^2 + \omega^2 C_{m}^2} \]  \hspace{1cm} (3.10b)

\[ a = G_m - (G_m^2 + \omega^2 C_m^2) R_s \]  \hspace{1cm} (3.10c)

**Figure 3.8** dI/dV is the derivative from DC measurement of IV characteristics. This conductance is due to tunneling current. Small conductance signal is due to both tunneling current and interface states. When frequency is low enough, it converges to dI/dV.
Finally, the equivalent circuit shown in Figure 3.9(a) can be converted to the Figure 3.9(c) via an intermediate modeling step as shown in Figure 3.9(b) [60]. The relationship between $D_{it}$ and the corrected data (removed the impact of substrate resistance) are shown in Equation (3.11). As the circuit in Figure 3.9(a) converts to Figure 3.9(b), one can summarize it in Equation (3.11a). After taking into account of the time constant dispersion, Equation (3.11a) can be rewritten as Equation (3.11b) [8]. $D_{it}$ can then be calculated at the peak of $\frac{G_p}{\omega}$ as shown in Equation (3.11c), and $\frac{G_p}{\omega}$ can be calculated by relating Figure 3.9(b) to Figure 3.9(c) through Equation (3.11d).

![Diagram](image)

**Figure 3.9** (a) MOS-C model with the interface trap time constant ($\tau_{it} = R_{it}C_{it}$), where $C_{ox}$ is the oxide capacitance, $C_s$ is the substrate capacitance, $R_{it}$ and $C_{it}$ are the interface defects induced resistance and capacitance, respectively. (b) Simplified circuit of (a), $C_p$ and $G_p$ are the equivalent capacitance and conductance in the substrate. (c) Circuit used for measurement after the correction, $G_c$ is the corrected conductance and $C_c$ is the corrected capacitance. $D_{it}$ measurement by Labview program is shown in Figure 3.10.
$$\frac{G_p}{\omega} = \frac{q^2 \omega \tau_{it} D_{it}}{1 + (\omega \tau_{it})^2} \quad (3.11a)$$

$$\frac{G_p}{\omega} = \frac{q^2 D_{it}}{2 \omega \tau_{it}} \ln[1 + (\omega \tau_{it})^2] \quad (3.11b)$$

$$D_{it} \approx \frac{2.5 \left( \frac{G_p}{q^2} \right)_{\text{max}}}{\omega} \quad (3.11c)$$

$$\frac{G_p}{\omega} = \frac{\omega G_c C_{ox}^2}{G_c^2 + \omega^2 (C_{ox} - C_c)^2} \quad (3.11d)$$

**Figure 3.10** Screen shot of Labview program for $D_{it}$ measurement.

### 3.2.4 Deep Level Transient Spectroscopy (DLTS)

DLTS method is a technique based on capacitance transient over time (C-t) [9]. It is originally developed by D. V. Lang. And subsequently used by K. Yamasaki to study MOS interface state[61]. Transient as a function of time is evaluated by applying a voltage pulse on the device to estimate interface or bulk defects in system. This
phenomenon is also another version of measurement of CV hysteresis, and C-t is accurately measured the hysteresis at one specific voltage as Figure 3.11 shows.

![Figure 3.11](image)

**Figure 3.11** Two voltages, namely $V_a$ and $V_b$, are alternatively applied on the gate stack. Capacitance was sampled at two different time $t_1$ and $t_2$ after voltage was returned to $V_a$. Normally, an observation of increased capacitance will happen if there is majority carrier emission from the traps at the interface.

The amplitude of C-t transient ($\Delta C$) depends on voltage pulse, and an approximate interface defects density can be estimated by Equation (3.7), where $\Delta V$ is obtained from CV characteristics. The transient profile is following exponential, when there is only one type of defects in system (one specific energy level). DLTS measure emission time of trap or defect rather than emission, which means devices usually biased in depletion condition first, then a pulse voltage change the system to accumulation
condition (fill interface or bulk trap with majority traps), finally bias device back to depletion condition. The simulations of two biased conditions are plotted in Figure 3.12.

![Figure 3.12](image)

**Figure 3.12** Band diagram at two different bias conditions for p type substrate are plotted for both MOS structure (SiO$_2$/p type Si) and PN junction (n type CdS/p type CdTe): (a) MOS structure, bias voltage = 0 V, (b) MOS structure, bias voltage = -0.5 V (voltage difference between Si and SiO$_2$);

The emission rate of the majority traps therefore can be monitored as a C-t signal. Inset of Figure 3.11 shows a typical applied voltage pulse sequence and measured capacitance over time. Conventionally, device usually is filled with majority carrier rather than minority carrier [62, 63], since filling rate of majority carrier is in range of $\mu$s and it is easier for implementation when considering multiple C-t transients are to be averaged. Additionally, DLTS is only able to detect half band of interface since pulse has to be in one direction. Equation (3.12) shows the relationship between defect energy level relative to the valence band.

$$
\tau_p = \frac{1}{e_p} \frac{\exp \left( \frac{E}{kT} \right)}{T^2 \gamma_p \sigma_p} \tag{3.12}
$$

$E= E_T - E_V$ ($E_T$ is energy level of defect and $E_V$ is energy level of valence band), temperature $T$, and capture cross section $\sigma_p$, here subscript $p$ stands for hole. $\gamma_p$ is
material dependent signature. When \( C-t \) is measured over wide range temperatures, time constant of defects changes as well, and it is illustrated in Figure 3.13. If transient capacitances were sampled by two sampling time \( t_1, t_2 \), the difference of capacitance (\( \Delta C \)) will have maximum when time constant of defect is equals to \( \frac{t_1-t_2}{\ln(t_2/t_1)} \) as Equation (3.13) indicates.

\[
\tau = \frac{t_1-t_2}{\ln(t_2/t_1)} \text{ at } C_{\text{max}} \tag{3.13}
\]

Where \( t_1 \) and \( t_2 \) can be set to have serveal pairs of time constant and temperature (\( \tau, T \)). Finally, Arrhenius plot can be obtained via Equation (3.12), the slope of it is the activation energy \( E \) and intercept is \( (y_p \sigma_p)^{-1} \).
3.2.5 Flicker Noise

Flicker noise is a very critical parameter regarding noise performance of MOSFET [66]
and it is seen that flicker noise is a formation of discrete random telegraph noise (RTN) [67]. Though it is considered a reliability issue, it can also be used as diagnostic tool for quality of oxide and interface in terms of defects [68]. The origin of flicker noise is still a mystery since there is no microscopic picture of it. Two competing theories exist: McWhorter model and Hooge model. McWhorter model is based on assumption that carriers in channel are fluctuated by trap and detrap process via dielectrics as described by Figure 3.14 [69]. Hooge model is originally derived from study of bulk of metallic material, and it was applied on transistor but it cannot adequately explain the experimental data. The noise was attributed to mobility fluctuation of carriers in the channel [70]. To explain more experimental data by a physical model, a unified model was proposed by Prof. Hu. The unified model correlated them by a scattering coefficient [71]. Figure 3.15 shows the how three models are developed from conductance fluctuation theory: The defects in oxide fluctuate the conductance in channel by either scattering or trapping/detrapping.

![Figure 3.14](image)

**Figure 3.14** Carriers in channel are trapped and detrapped from dielectric.
First step of analyzing flicker noise is to find correct model of the device. Both of models predict a bias dependence of flicker noise level as Equation (3.14-3.15) show. If it is more than adequate to use unified model, then fitting values will be between the values predicted by the two other models (-2 for McWhorter and -1 for Hooge model).

$$N_t = \frac{S_{id}C_{EOT} W L f (V_g - V_{th})^2 \gamma}{q^2 k T I_d^2}$$  \hspace{1cm} (3.14)$$

$$\alpha_H = \frac{f W L C_{EOT} |V_g - V_{th}| S_{id}}{q I_d^2}$$  \hspace{1cm} (3.15)$$

The oxide trap density, $N_t$ and Hooge empirical parameter ($\alpha_H$) are given by [72], where $q$ is the elementary electron charge, $kT$ is the thermal energy, $\gamma$ is the attenuation coefficient ($10^8 \text{ cm}^{-1}$), $W$ is channel width, $L$ is channel length, $C_{EOT}$ is the gate dielectric capacitance per unit area, $S_{id}$ is measured noise level, $f$ is frequency point where noise
level was sampled, $I_d$ is drain current and $V_{g}\!-\!V_{th}$ is overdrive voltage.

### 3.2.6 IV Temperature Dependence

Current voltage measured at various high temperatures can give defect energy level in the oxide if IV characteristic is temperature dependent, and barrier height if is just a tunneling process assisted by electrical field. Therefore, measuring IV at various temperatures is necessary to evaluate oxide quality regarding oxide traps and conduction band offset of gate/dielectrics or dielectrics/substrate. The transport mechanism in dielectrics can be Poole-Frenkele mission (PF), hopping conduction (HC), Fowler-Nordeim tunneling (FN), direct tunneling (DT), and thermionic-field emission (TE), ohmic conduction [73]. DT current is significantly dependent on oxide thickness only if the thickness is below 4 nm [74]. Therefore, it is not considered in this study (above 4 nm physical thickness). Ohmic conduction is valid in very low field region or when dielectric is broken, and it is not critical as well. PF, HC, TE mechanisms predicts a temperature dependent behavior of IV since the electron can tunnel or jump over the barrier by thermal energy. On the other hand, FN is a mechanism dependent only by its barrier shape, and usually valid at high electric field range. All the models need electric field information present in the dielectrics. Equation (3.16) shows electric filed in typical bilayer structure of dielectrics, which has to meet boundary conditions.

$$
\varepsilon_{Al_2O_3} E_{Al_2O_3} = \varepsilon_{ZrO_2} E_{ZrO_2} \quad (3.16a)
$$

$$
E_{Al_2O_3} t_{Al_2O_3} + E_{ZrO_2} t_{ZrO_2} + V_{FB} = V_{Gate} \quad (3.16b)
$$

Where $E$ is electrical filed, $\varepsilon$ is dielectric constant, $t$ is dielectric thickness, $V_{Gate}$ is gate
voltage. The voltage consumed in substrate can be neglected when $V_{\text{Gate}}$ biases MOS-C into accumulation or inversion region. Equation (3.17-3.20) list FN, HC, PF, TE carrier transport mechanism in dielectrics, where $m$ stands for electron effective mass, $\hbar$ is Plant constant divided by $2\pi$. $A^*$ is effective Richardson constant. By using Equation (3.17), the barrier height ($\phi_b$) can be extracted. In Equation (3.18), HC model gives tunneling distance between two trap sites ($a$) and activation energy ($E_a$). PF can give trap energy level ($\phi_t$), which has the same meaning as $E_a$. The TE model as Equation (3.20) can extract barrier height $\phi_B$, this barrier is typical and lower than $\phi_b$ since electron can thermally overcome it.

$$J_{FN} = \frac{q^3 m_0}{16\pi^2 \hbar m_{ox} \phi_b} E^2 \exp \left( -\frac{4(2m_{ox})^{1/2}}{3q\hbar} \phi_b^{3/2} E^{-1} \right)$$ (3.17)

$$J_{HC} = q a n v \exp \left( \frac{qaE}{kT} - \frac{E_a}{kT} \right)$$ (3.18)

$$J_{PF} = q N_C \mu E \exp \left( \frac{\left( q^3 / \pi \varepsilon_r \varepsilon_o E \right)^{1/2} - q \phi_t}{kT} \right)$$ (3.19)

$$J_{TE} = A^* T^2 \exp \left( -q \left( \phi_B - \frac{\sqrt{qE / 4\pi \varepsilon_r \varepsilon_o}}{kT} \right) \right)$$ (3.20)

Detailed application of how to apply these models into real IV measurement will be discussed in Chapter 5.

### 3.2.7 Time Dependent Dielectric Breakdown (TDDB)

Time dependent dielectric breakdown is a useful technique to study quality of dielectrics, since it is able to predict the oxide lifetime or total charge to breakdown dielectrics ($Q_{BD}$)
(\(Q_{BD}\) can be calculated by integrating current over time). The value \(Q_{BD}\) was used instead of breakdown time \(t_{BD}\) since its straightforward to evaluate oxide quality. A stress voltage is applied on MOS-C over time, and current is sampled until an observation of significant change occurs, which is an indication of oxide breakdown. The stress voltage usually is larger than the working voltage to accelerate breakdown process in order to have breakdown time \(t_{BD}\) in \(10^{-2}\) s to \(10^3\) s. It has to be noticed that the maximum \(t_{BD}\) extend to a longer value, that is time consuming. The minimum value of \(t_{BD}\) is dependent on sampling rate of specific instrument. At specific stress voltage, a certain amount of devices are necessary to be measured for its statistical information, which can be explained by Weibull distribution as shown in Equation (3.21) [75], where \(\beta\) and \(\eta\) are shape parameter and they are constant. The final \(Q_{BD}\) can be selected once 90% of devices are broken down. The charge to breakdown \(Q_{BD}\) and voltage acceleration factor \((AF)\) are two parameters used to evaluate the oxide layer quality, and \(AF\) can be extracted by measuring \(Q_{BD}\) at several stress voltages as Equation (3.22) shows. \(\beta\) in Equation (3.21a) is dependent on oxide thickness and trap size \(a_o\) (Equation (3.23a)) [76], where \(\alpha\) is a parameter describing the correlation between \(t_{ox}\) and \(\beta\). Moreover, \(\beta\) value is demonstrated to be linearly related to oxide thickness (Equation (3.23b)), where \(\gamma\) is the coefficient and \(t_{INT}\) is thickness of IL [77].

\[
F(Q_{BD}) = 1 - \exp\left(-\frac{Q_{BD}^\beta}{\eta}\right) \quad (3.21a)
\]

\[
W \equiv \ln(-\ln(1 - F)) = \beta \ln(Q_{BD}) - \ln(\eta) \quad (3.21b)
\]

\[
Q_{BD} = Q_o \exp(AF \cdot V) \quad (3.22)
\]
\[ \beta = \alpha \frac{t_{ox}}{a_o} \quad (3.23a) \]

\[ \beta = \gamma (t_{ox} + t_{INT}) \quad (3.23b) \]

The measurement TDDB could be significantly different when gate voltage is in different polarity since electron can tunnel through dielectrics by different mechanism, especially for bilayer structure.
CHAPTER 4

OXIDE STRUCTURE-DEPENDENT INTERFACIAL LAYER DEFECTS OF HfAlO/SiO\textsubscript{2}/Si STACK ANALYZED BY CONDUCTANCE METHOD

4.1 Research Motivation for Al Doped HfO\textsubscript{2} as high-k Dielectrics

Recently HfO\textsubscript{2}-based high-k gate dielectrics have become more popular in the advanced semiconductor devices because of their large energy gap and thermodynamic stability with the Si substrate [78]. However, aggressive downscaling of effective oxide thickness (EOT) requires improvement when dealing with gate leakage current and dielectric-substrate interfacial characteristics. At present, the fixed negative charge and interfacial layer defects of HfO\textsubscript{2}/Si gate stacks with an interfacial layer (IL) are not as robust as that of the SiO\textsubscript{2}/Si interface in earlier technologies. This degradation led to a reduction in carrier mobility because of the higher (\(\sim \times 10^{12}\) cm\textsuperscript{-2}) interface state density as compared to SiO\textsubscript{2}/Si (\(<\times 10^{11}\) cm\textsuperscript{-2}) due to additional scattering of carriers at the interface [3]. SiO\textsubscript{2}/Si interface constitutes a high-quality interfacial condition due to the amorphous structure and stoichiometric properties of SiO\textsubscript{2} even after 1100 °C annealing. It is well known that the interfacial layer defects are attributed to the Si dangling bonds at the interface and can be minimized by post metallization annealing (PMA) in forming gas (NH\textsubscript{3}). This observation was extensively studied by various groups and several quantitative assessments were reported [79, 80]. In the case of HfO\textsubscript{2}/Si, there is always a presence of an interfacial layer and the crystallization state of HfO\textsubscript{2} modulates the interface and its quality. Since HfO\textsubscript{2} tend to crystallize around 500 °C, the phase change process can highly impact the IL and the interface. Incorporation of Zr and Al in HfO\textsubscript{2} has been tried by various process conditions to enhance the dielectric quality and EOT downscaling.
[81, 82]. It is reported that the amorphous to polycrystalline phase transition temperature can be increased by 400 °C-500 °C with the addition of Al [78]. This characteristic helps eliminate electrical and mass transport along the grain boundaries that reduces the leakage current. There are several other advantages when Al is incorporated into HfO₂: an increase in dielectric constant, higher band gap, lower hysteresis, and reduction of electrical defects [83, 84]. It was reported by several groups through experiments and theoretical calculations that Al incorporated HfO₂ tends to crystallize to the tetragonal phase (t-HfO₂) instead of the monoclinic phase (m-HfO₂) [85-89]. Figure 4.1 shows schematic of the HfO₂–Al₂O₃ phase diagram.
Figure 4.1 Schematic of the HfO$_2$–Al$_2$O$_3$ phase diagram, the symbols represent the crystalline phases identified by electron diffraction in the films with different amounts of Al$_2$O$_3$ after annealing at 1400 °C.

Source:[85].

With the addition of Al$_2$O$_3$ into HfO$_2$, Park et al. [87] demonstrated an increased dielectric constant of the in ALD Hf aluminate films with (002) oriented tetragonal phase stabilization. For HfO$_2$/Al$_2$O$_3$ bilayer structures, Cho et al. [90] and Nishimura et al. [91] showed improved thermal stability for dielectrics annealed up to 900 °C. Incorporation of Al in alloy form provides better EOT downscaling potential and a reduced gate leakage current in comparison to bilayer form. But the impact of phase transition and Al diffusion
on the interface has not been extensively studied. N. Zhan et al. used deep level transient spectroscopy (DLTS) to study the HfAlO–Si interface under different PDA temperatures [92]. However, few reports are available about the impact on interface state density of Al incorporated HfO$_2$-based oxides under specific annealing conditions leading to phase transition. Since the $D_{it}$ is a critical parameter for advanced metal oxide semiconductor field effect transistors (MOSFET), its correlation with Al needs to be sufficiently explored. This chapter studies the role of a low percentage of Al, incorporated by HfAlO$_x$ cap layer or by HfO$_2$–HfAlO$_x$–HfO$_2$ multilayer, on the crystallinity phase of HfO$_2$ and its impact on the interface state density ($D_{it}$). It was observed that the effect of the incorporation of lower permittivity aluminum prevented the oxide from crystallizing. This increased the annealing temperature and formed a high quality of oxide (lower EOT). At 800 °C annealing temperature, the oxide phase varied from amorphous to crystalline (tetragonal) or partial crystallization (tetragonal) by incorporating different Al concentrations [93]. It was also observed that $D_{it}$ values are proportional to dielectric constant.

4.2 Device Information and Experimental Procedure
MOS-C samples, prepared using 13 different process conditions, are listed in the Table 4.1. Nine samples on the left with four additional samples on the right. T1 implies no post deposition annealing (PDA), T2 implies PDA at 680 °C, T3 implies PDA at 800 °C. PDA is done in N$_2$ environment. A5% means Al/(Al+Hf) equals to 5 %. Oxide deposition layers with cycles are shown. Cy stands for cycles. The 4 additional samples have same temperature and oxide deposition configuration corresponding to the sample in left
column in the same row, otherwise stated. The differences are listed in the comments column.

All the samples have 8 angstroms of interfacial layer SiO$_x$, intentionally deposited on the Si substrate by ozone treatment prior to depositing the high-k layer by atomic layer deposition (ALD) in a 300 mm TEL Trias™ cleanroom tool at TEL Technology Center. A set of the samples were annealed in N$_2$ environment at temperatures ranging from 680 °C to 800 °C for 60 s in a clustered rapid thermal annealing chamber without a vacuum break. 5 nm metal gate was then deposited by ALD TiN followed by 50 nm physical vapor deposition (PVD) TiN. The substrate was of p-type Si with (001) orientation. The doping concentration was approximately 10$^{15}$ cm$^{-3}$. More details of the process conditions can be found in the report by Tapily et al. [93]. Choi et al.’s work [94] describes the effect of the annealing environment N$_2$ or O$_2$ on EOT, crystallinity, leakage current, interface trap density and hysteresis. Figure 4.2 shows gate stack structure information.

![Figure 4.2 Typical MOS structure in this study. (a) cap structure, (b) sandwich structure.](image)

At least six MOS capacitors (MOS-C) from each of sample type with an area of 40 µm×40 µm were investigated using high precision LCR meter HP4284. The capacitance-voltage (CV) and conductance-voltage (GV) characteristics were obtained
from 5 KHz to 1 MHz. Because of the highly instability of capacitance characteristics at low frequencies, 5 KHz was selected as the lowest frequency for the measurements. It was expected that a frequency of 5 KHz would prevent any errors in estimation of $D_{it}$ results. Gate voltage was swept from inversion to accumulation (1 V to -1.5 V) with a 0.1 V step voltage. The amplitude of the small AC signal was kept at 5 mV. It is known that smaller step voltage during measurements contributes to accurate $D_{it}$ values in the depletion region and the lower frequency limit extends the probing region to mid gap for $D_{it}$ estimation. In this work, the selection of step voltage and the lowest frequency range can accurately predict the $D_{it}$ values in the depletion region. After obtaining the CV and GV data, the series resistance ($R_s$) was calculated by using Equation (3.9) in the strong accumulation region at 1MHz [58]. $R_s$ was found to be in the range of 1KΩ to 2KΩ. The original measured capacitance ($C_m$) and conductance ($G_m$) values were subsequently corrected to exclude the effect of $R_s$ by using Equation (3.10) [58]. The leakage current density was also measured at $V_{FB} – 1$ V to monitor whether HfAlO layer is crystallized or not. Figure 4.3 shows overall flow of the measurement.
Figure 4.3 Steps calculating \( C_c \), \( G_c \) and \( D_{it} \).
Table 4.1 List of Process Information of 13 HfAlO Samples

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Oxide deposition configuration</th>
<th>Sample ID</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>40 Cy HfO</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>T1-A0%a</td>
<td>40 Cy HfAlO</td>
<td>S-T2-A0%c</td>
<td>SPAN(^d) instead of PDA</td>
</tr>
<tr>
<td>T2-A0%</td>
<td>40 Cy HfAlO</td>
<td>S-T2-A4.8%</td>
<td>SPAN before PDA</td>
</tr>
<tr>
<td>T2-A5%</td>
<td>20 Cy HfO + 20 Cy HfAlO</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>T2-A4.5%b</td>
<td>20 Cy HfO + 20 Cy HfAlO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>40 Cy HfO</td>
<td>S-T3-A0.6%</td>
<td>10 Cy HfO + 10 Cy HfAlO</td>
</tr>
<tr>
<td>T3-A2.4%</td>
<td>30 Cy HfO + 10 Cy HfAlO</td>
<td>S-T3-A2.6%</td>
<td>10 Cy HfO + 20 Cy HfAlO</td>
</tr>
<tr>
<td>T3-A4.2%</td>
<td>20 Cy HfO + 20 Cy HfAlO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3-A6.6%</td>
<td>10 Cy HfO + 30 Cy HfAlO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(a\) A0% means no Al (less than 0.2 %) is detected by X-ray photoelectron spectroscopy (XPS) with open Al source.

\(b\) This sample actually did PDA at 700 °C, but the electrical property shows no difference from other 680 °C PDA sample.

\(c\) S stands for special case. The reason of they are categorized into special is either annealing method or no HfAlO cap.

\(d\) SPAN is Slot Plane Antenna in \(N_2\) environment.

4.3 CV, GV, EOT, Dielectric Constant, Leakage Current Density, \(D_H\) and \(V_{FB}\)

Figure 4.4 and Figure 4.5 show the comparison between the corrected and measured values of capacitance and conductance, respectively. In the accumulation region large dispersion was observed in capacitance (Figure 4.4). After the correction, the corrected capacitance, \(C_c\) shows less dispersion than the measured capacitance, \(C_m\). Similarly, the corrected conductance, \(G_c\) shows no dispersion and converged to zero in the accumulation region because there was no significant leakage conductance (\(G_l\)). The conductance bump in the depletion region is mainly due to the interfacial defects (Figure 4.5). Therefore, it is
clear that the effect of $G_t$ can be neglected in the measurement. It was further suggested that if the bump in the depletion region was impacted by the dispersion, the correction to series resistance was required and the tunneling or leakage current should be monitored in the accumulation region. If they were comparable to the conductance in the depletion region, then it is necessary to subtract the leakage conductance from the corrected conductance [59].

**Figure 4.4** Capacitance data of the sample fabricated by depositing 10 cycles of HfAlO as a cap layer on the 30 cycles of HfO (T3-A2.4%), where the corrected capacitance, $C_c$, show less frequency dispersion as compared to the measured capacitance, $C_m$. Inset figure is the corresponding band diagram when gate voltage equals to -1 V.
Figure 4.5 Conductance data of the sample fabricated by depositing 10 cycles of HfAlO as a cap layer on the 30 cycles of HfO (T3-A2.4%), where the corrected conductance, $G_c$, in the accumulation region is much smaller than in the depletion region. However, the measured conductance, $G_m$, show frequency dispersion in the accumulation region. Inset figure is the corresponding band diagram when gate voltage equals to -0.5 V.

The EOT and $V_{FB}$ values were calculated by the software CVC-7.0 developed by Hauser et al. [95] using the corrected capacitance, $C_c$. Without any capacitance corrections one can use the capacitance measured at 100 KHz for better accuracy. The results are listed in Table 4.2. Physical thickness and leakage current information were supplied by the TEL Technology Center. The $D_{it}$ values were calculated by the conductance method [8]. $\frac{G_p}{\omega}$ as a function of frequency at different bias voltages is plotted in Figure 4.6. Four peaks can be observed when the voltage step is 0.05 V. If the voltage step is changed to 0.1 V at least one of the peaks observed in Figure 4.6 can still be
visible in the present frequency range (not shown). The surface band bending ($\psi_s$) can be generated by the software CVC-7.0.$^{21}$ A small error can be introduced since the CVC software assumed a perfect interface without any $D_{it}$. The frequency range used here can only depict that the $\frac{G_p}{\omega}$ peaks around the middle of intrinsic Fermi level and the Fermi level of the substrate at the interface in the depletion region. This demonstrated that our $D_{it}$ was valid in the surface potential around $\phi_s = -0.2$ V, where $\phi_s = (E_F - E_i)/K_T$ at the interface (Figure 4.6 inset). The bulk potential ($\phi_B$) of the devices were found to be around -0.3 V. Figure 4.7 shows the corrected conductance ($G_c$) of 13 different samples as a function of the gate voltage. The peaks observed here are consistent with the $D_{it}$ calculated in Table 4.2. The higher the conductance peak the higher the $D_{it}$. Therefore, Figure 4.7 further validates the measured $D_{it}$ results.
Figure 4.6 $\frac{G_p}{\omega}$ is plotted by sweeping both frequency and gate voltage. The inset shows band diagram. $E_c$, $E_i$, $E_F$ and $E_v$ are the conduction band, intrinsic level, Fermi-level and valence band, respectively.

Figure 4.7 The corrected conductance ($G_c$) values of 13 different samples are plotted as a function of gate voltage.
The dielectric constant of the HfAlO was estimated by Equation (4.1). Interfacial layer is SiO$_2$ and hence was assumed to be 3.9. The interfacial layer thickness was measured by the X-ray photoelectron spectroscopy (XPS) and is found to be around 0.6 nm after post-treatment as shown in. The structure of the insulator can be speculated by relating the dielectric constant to the known and well calculated structure (dielectric constant 21 as amorphous phase, 29 as cubic phase, 70 as tetragonal phase) [88]. The leakage current density ($J_g$) that represents the formation of grain boundaries during initial crystallization process is close or above 0.01 A·cm$^{-2}$. Otherwise, it can be considered that dielectric is either partially crystallized or amorphous. Tapily et al. reported the details of the physical characterization of these samples [93]. HfAlO crystallization temperature as a function of the Al concentration was investigated by an in-situ XRD (X-Ray Diffraction) characterization unit. Around 6 % Al/(Al+Hf) concentration kept the HfAlO partially crystallized when the annealing temperature was 800 °C. The phase change in the oxide was interpreted by grazing incidence X-ray diffraction (GIIXRD) spectra [93]. As-deposited oxides were amorphous regardless of Al concentration. On the other hand, HfAlO tends to crystallize to a mix of tetragonal and cubic phase rather than the monoclinic phase when the low percentage of Al is incorporated into the HfO$_2$. The combined use of the near edge X-ray absorption fine structure (NEXAFS) of O absorption and the pair distribution function determined from grazing incidence extended X-ray absorption fine structure (GI-EXAFS) of Hf absorption were used for determining the phases present in the ultra-thin Al-doped HfO films [93]. Tetragonal phase data fits better than cubic phase. Therefore the dielectric constant of samples could be of any value between 21 and 70 if the possibility of mixed phase is
considered.

\[
\frac{1}{C_{\text{total}}} = \frac{1}{C_{\text{SiO}_2}} + \frac{1}{C_{\text{oxide}}} \quad (4.1a)
\]

\[
\frac{EOT}{K_{\text{SiO}_2}} = \frac{t_{\text{oxide}}}{K_{\text{oxide}}} + \frac{t_{\text{IL}}}{K_{\text{SiO}_2}} \quad (4.1b)
\]

\[
K_{\text{oxide}} = \frac{t_{\text{oxide}}}{EOT} - \frac{t_{\text{IL}}}{K_{\text{SiO}_2}} \quad (4.1c)
\]

**Table 4.2** List of Results of 13 HfAlO Samples

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>(V_{\text{FB}}) (V)</th>
<th>EOT (nm)</th>
<th>(J_g) (A/cm(^2)) by xps</th>
<th>HfO(_2) thickness (Å)</th>
<th>Al/(Al+Hf) (^a) %</th>
<th>(K)</th>
<th>(D(_k)) (cm(^2)eV(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>-0.22</td>
<td>1.10</td>
<td>0.0049</td>
<td>5.02</td>
<td>32.42</td>
<td>0.0</td>
<td>21</td>
</tr>
<tr>
<td>T1-A0%</td>
<td>-0.23</td>
<td>1.08</td>
<td>0.0099</td>
<td>4.86</td>
<td>32.47</td>
<td>0.0</td>
<td>21</td>
</tr>
<tr>
<td>T2-A0%</td>
<td>-0.46</td>
<td>1.03</td>
<td>0.0193</td>
<td>5.76</td>
<td>32.40</td>
<td>0.0</td>
<td>28</td>
</tr>
<tr>
<td>T2-A5%</td>
<td>-0.38</td>
<td>1.09</td>
<td>0.0011</td>
<td>6.03</td>
<td>32.54</td>
<td>5.0</td>
<td>26</td>
</tr>
<tr>
<td>T2-A4.5%</td>
<td>-0.37</td>
<td>1.08</td>
<td>0.0016</td>
<td>6.11</td>
<td>31.91</td>
<td>4.5</td>
<td>27</td>
</tr>
<tr>
<td>T3</td>
<td>-0.48</td>
<td>1.03</td>
<td>0.0114</td>
<td>6.09(^a)</td>
<td>28.54(^a)</td>
<td>N/A</td>
<td>26</td>
</tr>
<tr>
<td>T3-A2.4%</td>
<td>-0.58</td>
<td>0.87</td>
<td>0.0086</td>
<td>6.95</td>
<td>31.59</td>
<td>2.4</td>
<td>70</td>
</tr>
<tr>
<td>T3-A4.2%</td>
<td>-0.53</td>
<td>0.90</td>
<td>0.0036</td>
<td>7.15</td>
<td>31.03</td>
<td>4.2</td>
<td>65</td>
</tr>
<tr>
<td>T3-A6.6%</td>
<td>-0.50</td>
<td>1.02</td>
<td>0.0015</td>
<td>6.95</td>
<td>31.77</td>
<td>6.7</td>
<td>38</td>
</tr>
<tr>
<td>S-T2-A0%</td>
<td>-0.35</td>
<td>0.98</td>
<td>0.0092</td>
<td>6.86</td>
<td>32.38</td>
<td>0.1</td>
<td>42</td>
</tr>
<tr>
<td>S-T2-A4.8%</td>
<td>-0.34</td>
<td>1.06</td>
<td>0.0015</td>
<td>6.61</td>
<td>32.15</td>
<td>4.8</td>
<td>31</td>
</tr>
<tr>
<td>S-T3-A0.6%</td>
<td>-0.52</td>
<td>0.89</td>
<td>0.0101</td>
<td>6.45</td>
<td>31.22</td>
<td>0.6</td>
<td>50</td>
</tr>
<tr>
<td>S-T3-A2.6%</td>
<td>-0.48</td>
<td>0.91</td>
<td>0.0038</td>
<td>7.25</td>
<td>31.24</td>
<td>2.6</td>
<td>66</td>
</tr>
</tbody>
</table>

\(^a\) Thickness were measured by X-ray Reflectivity(XRR) instead of XPS.

\(^b\) Aluminum concentration was measured by XPS.
4.4 Discussion on the Role of Aluminum in HfO$_2$

To investigate the role of the aluminum on EOT, the dielectric constant of all the samples were estimated. The dielectric constant of both the 40 Cy HfO and 40 Cy HfAlO samples without the PDA anneal (T1 and T1-A0%, respectively, (Table 4.1)) is ~21. This is close to both theoretical values and the reported experimental results of amorphous HfO$_2$ [96, 97]. The observed flatband voltage, $V_{FB}$ for both the devices is approximately -0.22 V. The $D_{it}$ was found to be the lowest of all the samples ($\sim 3 \times 10^{11} \text{eV}^{-1}\text{cm}^{-2}$) because of the reduced interface roughness and a good stoichiometric interface. Next group of samples, 40 Cy HfAlO annealed at 680 $^\circ$C and 40 Cy HfO annealed at 800 $^\circ$C (T2-A0% and T3, respectively), have a dielectric constant of 28 and 26, respectively. It was reported by C. Zhao et al. that a 3 nm HfO$_2$ stack crystallizes at 600 $^\circ$C [98]. It is possible that these two samples were crystallized. This can be further confirmed by the leakage current measurements listed in Table 4.2.

To evaluate the possible crystallization, the leakage current density of annealed samples were plotted as a function dielectric constant as shown in Figure 4.8. The samples above the blue line are considered to be crystallized, since a high leakage current density was induced by formation of grain boundaries. The samples that were not crystallized or only partially crystallized follow a linear trend. Oxides with higher dielectric constant have smaller band gap, which can also cause a relatively higher leakage current (Figure 4.8).
Figure 4.8 Leakage current density measured at gate voltage ($V_{FB}$ -1 V). The results of 13 samples which were annealed at the temperatures 680 °C and 800 °C, are plotted as a function of dielectric constant.

The $V_{FB}$ of devices anneal at 800°C were found to be around -0.47 V. It is believed that negative fixed charge existed in as-deposited dielectrics. Chlorine from the precursor is a possible candidate to contribute to the oxide charge [99, 100]. However, the dielectric constants of these two sample types are not as low as 16 as in the case of monoclinic HfO$_2$ (m-HfO$_2$). It is, therefore, possible that some residual aluminum atoms (not detected by XPS) contributed to the formation of tetragonal phase that increased the dielectric constants (Figure 4.8), which is in between the range of amorphous (~21) and tetragonal phase (70) [88]. The $D_{it}$ values of these devices are close to $5 \times 10^{11}$ eV$^{-1}$cm$^{-2}$ showing a marginal increase compared to T1 and T1-A0% as described earlier.

When the samples 20 Cy HfO + 20 Cy HfAlO (T2-A5%) and 20 Cy HfO + 20 Cy HfAlO (T2-A4.5%) were subjected to a 680 °C PDA and a 700 °C PDA, respectively, no
significant variation in the electrical measurements were observed. About 5% and 4.5% Al was detected in the samples T2-A5% and T2-A4.5%, respectively. The dielectric constants of these two samples simply represent the impact of Al presence in HfO₂. Even though the dielectric constant of these samples is in the same range as that of T2-A0% and T3 as mentioned above, the partial crystallization process with higher Al concentration does not decrease the dielectric constant significantly. But the observed leakage current in the partially crystallized structures T2-A5% and T2-A4.5% is approximately one tenth of the crystallized T2-A0% and T3 samples (Figure 4.8 and Table 4.2). The observed $V_{FB}$ shifts for both the samples were about 0.1 V in the positive direction, possibly due to the negative fixed charge contributed by the Al accumulation at the HfAlO-Si interface [101, 102]. The $D_{it}$ moderately decreased to a level of $4 \times 10^{11}$ eV$^{-1}$cm$^{-2}$.

While the PDA temperature was maintained at 800 °C the samples 40 Cy HfO, 30 Cy HfO + 10 Cy HfAlO, 20 Cy HfO + 20 Cy HfAlO, and 10 Cy HfO + 30 Cy HfAlO (T3, T3-A2.4%, T3-A4.2%, and T3-A6.6%, respectively, (Table 4.1)) with different percentage of Al concentrations were incorporated via different stack structures. Dielectric constant of the sample T3-A2.4% is found to be ~70. It is clear that this sample was crystallized to the tetragonal phase giving a very high dielectric constant [88]. The leakage current for this sample was found to be $> 0.01$ A·cm$^{-2}$ which further confirms maximum crystallization (Figure 4.8). The $D_{it}$ is highest among all the 13 samples ($1.2 \times 10^{12}$ eV$^{-1}$cm$^{-2}$). This high $D_{it}$ value possibly is the reason of shifting $V_{FB}$ to negative direction about 0.06 V, if the interface defect is below the intrinsic Fermi level ($E_i$) are positively charged. With further increase in the Al concentration, the samples
T3-A4.2% and T3-A6.6% tend to limit the crystallization process. The values of $V_{FB}$, dielectric constant, leakage current and $D_{it}$ clearly represent partial crystallization (Figure 4.8 and Table 4.2). The $D_{it}$ of the sample T3-A6.6% is reduced to the level of $5 \times 10^{11}$ eV$^{-1}$cm$^{-2}$ to representing a more amorphous structure of the dielectric.

The behavior and the structural properties of the additional four samples need to be addressed individually. Some of the 40 Cy HfAlO samples (S-T2-A0%) were subjected to SPAN instead of PDA. Typically SPAN was carried out at a much lower temperature than PDA. SPAN densifies the dielectric layer and prevents the regrowth of interfacial layer. Due to low temperature processing, SPAN treatment partially reduced the fixed oxide charge as observed by the flat band voltage value, $V_{FB} = -0.35$ V, which is less than any of samples annealed at 800 °C (i.e. T3: $V_{FB} = -0.46$ V). The 20 Cy HfO +20 Cy HfAlO sample was subjected to SPAN prior to PDA at 680 °C (S-T2-A4.8%). The SPAN exposure moderately increases the dielectric constant, while no significant difference was noticed in other characteristics.

Two of the samples, 10 Cy HfO + 10Cy HfAlO + 20 cy HfO (S-T3-A0.6%) and 10 Cy HfO + 20 Cy HfAlO +10 cy HfO (S-T3-A2.6%) were prepared when the HfAlO layer was sandwiched in the middle instead of at the top as a cap layer. When subjected to 800 °C annealing, possible non-uniform distribution of Al in the dielectric layer led to the lower Al concentration at the top of the dielectric whereas moderately higher concentration at the interfacial layer is expected. This is possibly due to partial out diffusion of Al through HfO$_2$ layer. Because of possible non-uniform distribution of Al, it is difficult to correlate the dielectric constant with observed interface state.
Figure 4.9 The interface state density ($D_{it}$) values are plotted as a function of the dielectric constant of 13 samples.

Figure 4.9 shows the $D_{it}$ values as a function of dielectric constant for different samples. As discussed earlier, the dielectric constant increases with crystallization towards the complete tetragonal phase, favored by Al. Increased dielectric constant of the dielectric adjacent to the interface tends to increase the interface state density, $D_{it}$. As can be seen from Figure 4.9, the $D_{it}$ values of all 800 °C-annealed samples are in between the level of sample T3 (No annealing) and T3-A2.4% ($5 \times 10^{11} \sim 1.2 \times 10^{12}$ eV$^{-1}$cm$^{-2}$). The $D_{it}$ values increase with the crystallization of the dielectric, as the amorphous film shows the lowest $D_{it}$ values while the $D_{it}$ values are the highest at maximum crystallization. The $D_{it}$ of the samples that were not completely crystallized shows an almost linear function of dielectric constant as they plotted along the dashed line (Figure 4.9). The samples T2-A0%, T3, S-T3-A0.6% and T3-A2.4% were crystallized and demonstrate a higher leakage current and a comparatively higher $D_{it}$ values. S-T2-A0% and S-T2-A4.8% are the samples that were exposed to the SPAN process.
show a dielectric constant greater than the samples with only subjected to the PDA process.

![Figure 4.10](image.png)

**Figure 4.10** The interface state density ($D_{it}$) as a function of the Al/(Al+Hf) % of 13 samples are plotted. Data points within the dashed line are the 3 samples where Al was introduced by HfAlO as the cap layer on HfO$_2$. Data points in solid line are the 2 samples where Al was introduced by sandwiching HfAlO inside the HfO$_2$ stack.

The $D_{it}$ and $J_g$ values as a function of Al/(Al+Hf) at different PDA temperatures are shown in Figure 4.10 and Figure 4.11, respectively. $D_{it}$ as a function of the Al concentration shows (Figure 4.10) that the as-deposited sample without any PDA with 0 % percent Al has the lowest value of $D_{it}$. All devices subjected to 680 °C PDA have the $D_{it}$ values around $4 \times 10^{11}$ eV$^{-1}$cm$^{-2}$ irrespective of the Al concentration. Most of these devices have an Al/(Al+Hf)% either 0 % to 4 %. The dielectric constant of samples annealed at 680 °C PDA is $\sim 28$ regardless of the Al doping level (Table 4.2). Since it lacks the XPS data of Al/(Al+Hf)%%, which is between 0 % to 4 %, a more detailed investigation, is, therefore, required for Al concentration spanning from 0 % to 4 %.
Considering sample T3 (no Al) as a reference, samples with an 800 °C annealing show an initial increase in the $D_{it}$ with Al incorporation but $D_{it}$ values decrease with the increase of the Al concentration regardless of the deposited stack configuration (sandwich or cap). It can be observed from that the dielectric constant initially increases with the Al concentration but subsequently decreases after going through a peak at 2 %. The $D_{it}$ values, therefore, can possibly be reduced to the level that of the amorphous oxide if Al concentration is high enough. This phenomenon further suggest that it is preferred to relate $D_{it}$ to the dielectric constant rather Al concentration as shown in Figure 4.9. The $D_{it}$ difference, caused by gate stack configuration (cap or sandwich) during oxide deposition as shown in Figure 4.10, clearly suggest Al distribution in the dielectric can impact the interface state density. This property could be utilized to produce high quality devices. Sample S-T3-A2.6% is a typical example out of 13 samples: it has low EOT, high dielectric constant, low leakage current and moderate $D_{it}$ value.
Figure 4.11 The leakage current density ($J_g$) measured at the gate voltage ($V_{FB}-1$ V) as a function of the Al/(Al+Hf) % of 13 samples is shown as a function of Al concentration. Dashed dot line and dashed line follow the trend of $J_g$ varies with different Al doping level at 680 °C PDA and 800 °C PDA, respectively.

Figure 4.11 shows a comparison of the leakage current density, at a gate voltage ($V_{FB} - 1$ V), as a function of Al concentration for 13 sets of samples. Incorporation of higher concentration of Al could exponentially reduce the $J_g$. Moreover, samples formed by SPAN process show lower $J_g$. It was mentioned earlier that SPAN could significantly increase the dielectric constant even with an increase of high fixed charge density. It is worth to point out again, Al doping can enhance the quality of the oxide (lower EOT or higher dielectric constant), as Al increased the crystallized temperature. The fixed charge density could be reduced due to the benefit of higher annealing temperature. HfAlO prefer to remain amorphous at 800 °C PDA if the doping level of Al is above 6 % as shown by the dash-line in Figure 4.11.

In summary, it was observed that the $D_{it}$ value is a function of the dielectric
constant in Al incorporated HfO$_2$. The $D_{it}$ is also directly related to the structure of dielectric. This is in contrast to the SiO$_2$/Si where SiO$_2$ is always amorphous under the conventional fabrication process. Even though the interfacial layer of our samples is chemically prepared, it does not act as a perfect insulator as the thickness is close to the atomic level [80]. The carriers in the Si substrate can respond to the HfAlO, and the trapping and de-trapping of carrier process is affected by the HfAlO structure. In addition, small percentage of Al is mainly responsible to change the structure of HfO$_2$ from amorphous to crystalline. The amorphous structure has lowest $D_{it}$ ($2.76 \times 10^{11}$ eV$^{-1}$cm$^{-2}$) whereas t-HfO$_2$ has the highest $D_{it}$ ($1.27 \times 10^{12}$ eV$^{-1}$cm$^{-2}$). The $D_{it}$ results of other structures are within the range from lowest to highest value. It is concluded that increasing the dielectric constant of high-k gate dielectrics comes with a price, higher $D_{it}$ that decreases the channel mobility.
CHAPTER 5

ELECTRICAL CHARACTERIZATION OF DRY AND WET PROCESSED INTERFACE LAYER IN Ge/HIGH-K DEVICES

5.1 Research Motivation for Interface Treatment of high-k/Ge

It is well known that the motivation to study Ge devices is due to high hole and electron mobility, in comparison with silicon, besides the process and integration compatibility compared to that of III-V materials. Ge devices with different high-k gate dielectrics like HfO$_2$, ZrO$_2$, Al$_2$O$_3$, GeON have been demonstrated. Mobility above 300 cm$^2$V$^{-1}$s$^{-1}$ has been reported for Ge PMOS [103]. On the other hand, Ge NMOS has exhibited poor drive current and low mobility [104]. One possible reason is due to the quality of the gate oxide/substrate interface. Large interface defect density may pin the Fermi level, capacitance-voltage (CV) and conductance-voltage (GV) data, show that they no longer behave like traditional MOS capacitor [105]. In addition, implementing the process in 300 mm wafers for manufacturing adds to the complexity. Even though Ge/high-k interface has been extensively studied, the high leakage current associated with these gate stacks continue to introduce frequency dispersion and hysteresis in CV and GV characteristics [14, 106]. This dispersion severely limits the understanding the interface and accurate estimation of interface state density, $D_{it}$ and equivalent oxide thickness (EOT). Low temperature measurement offers a solution to the frequency dispersion issue, which is extensively studied here.

Conventionally, device information like EOT, $V_{FB}$, bulk doping and surface potential as a function of gate voltage can be obtained by CV measurement at a specific frequency, which is usually 100 KHz. However, this is carried out under a few important
assumptions: (i) substrate resistance is zero; (ii) no or minimal interface defects exist; (iii) minority carrier generation rate can’t follow this specific frequency; and (iv) leakage current is small enough not to disturb the CV measurements at this specific frequency. Even Si substrate is not able to satisfy all these assumptions. Due technology scaling, substrate resistance became larger, varied from few ohms to kilo ohms, which causes frequency dispersion in the accumulation region [58, 107]. Interface defects cause frequency dispersion in the depletion region (Silicon device) [108]. Thirdly, due to the advanced process technology (there are less bulk defects existing in Si substrate compared to Ge) [109] and large bandgap (1.12 eV), typical majority carrier response time is range from 0.01s to 1s [110]. Minority carrier cannot follow 1 Hz frequency if substrate is undoped [52]. Therefore there is no frequency response in the inversion region. The DC leakage current can further disturb the CV measurement, and it is necessary for it to be effectively monitored before further calculation. As discussed, reciprocally, the dispersion of capacitance and conductance at wide range of frequency (100 Hz to 1 MHz) can be utilized to calculate the $R_s$, the substrate resistance, and $D_it$, the interface defect density, as long as the last two assumptions are appropriate. This process is typically followed to extract the information from silicon devices.

Ge devices, however, are not yet ready for this simple evaluation process. There are two main reasons: (i) Large interface density (larger than $10^{12}$ cm$^{-2}$eV$^{-1}$) [11] that causes Fermi level more or less pinned [111] and (ii) Small bandgap (0.67 eV) that allows fast minority carrier generation and large leakage current [112]. The first flaw of high-k/Ge devices is a paradox when study the interface state density in these devices. In other words, larger interface defect density changes device attributes. Both low density
and high density of defects are difficult to be measured correctly. It is, therefore, important to make sure that there exist three distinctive regions (accumulation, depletion, and inversion) that can be observed in the CV measurement. The second flaw is quite significant for Ge devices, which is not as relevant in silicon. The data, as shown later, is easy to obtain in inversion region for Ge devices at moderate frequencies due to the small bandgap and low bulk defects in Ge substrate. It is, therefore, imperative to study the Ge devices at lower temperatures such that the second flaw can be more or less addressed. More importantly, a temperature scan means one additional dimension to the measurement that definitely benefits the data analysis.

5.2 Device Information and Experimental Procedure

MOS capacitors were fabricated on 300mm p-type Ge epitaxially grown on Si with a graded SiGe buffer layer. After Ge epi, the samples were subjected to three interface formation: (i) chemical oxidation (Chemox); (ii) The native oxide on the epi Ge wafers was removed using a TEL Certas™ chemical oxide removal (COR) process followed by 1 nm oxide by TEL slot-plane-antenna (SPA) plasma oxidation COR+SPAOx; and (iii) COR followed by vapor O₃ treatment (COR+O₃). A bilayer ALD 1nm Al₂O₃/3.5nm ZrO₂ was deposited in-situ in a 300 mm TEL Trias™ cleanroom tool. Al₂O₃ was selected as it is known to have nucleate effectively and forms better interface with germanium. Figure 5.1 shows structure of MOS-C. Since the dielectric constant Al₂O₃ is not high enough ZrO₂ was used to achieve better EOT scaling and electrical performance. The ALD Al₂O₃ layer was grown using trimethyl Aluminum as Al precursor and tetrakis (ethylmethylamido) zirconium as the Zr precursor. H₂O was used as the oxidant in both
cases at a deposition temperature of 250°C. Low temperature ALD deposition has achieved better characteristics. In our case 250°C deposition temperature was used for better high-k/Ge interface. ALD TiN was used as the metal gate. The electrical characterization (CV and GV) was performed using HP4284 LCR meter, capacitance transient (C-t) were performed by Booton 7200 and the low temperature measurements were performed by using LakeShore Model 335 temperature controller. The device size is 100 μm×100 μm if not specified.

![Off-Specification ALD Capacitor Structure](image)

**Figure 5.1** Structure of MOS capacitor of high-k/Ge.

In this work, for the first time CV and GV characteristics of Ge/ALD 1nm-Al₂O₃/ALD 3.5nm-ZrO₂/ALD TiN MOS capacitors were measured on 300 mm wafers with three different interface treatments. HP4284 LCR meter was used for the measurement at frequencies range from 100 Hz to 1 MHz and at five different temperatures (100 K, 150 K, 200 K, 250 K, 300 K). The interface treatments are
(i) simple chemical oxidation (Chemox); (ii) chemical oxide removal (COR) followed by 1 nm oxide by slot-plane-antenna (SPA) plasma (COR&SPA); and (iii) COR followed by vapor O$_3$ treatment (COR&O$_3$). The Chemox is wet process. The other two types of samples are dry processed. The EOT, $V_{FB}$, bulk doping, surface potential, and interface quality are calculated and the results were discussed with reference to the processing conditions after correcting the CV and GV data. Additionally, $D_{it}$ was estimated by conductance method, capacitance spectroscopy and deep level transient spectroscopy (DLTS) to study the interface treatment and its impact on defects.

5.3 EOT, $V_{FB}$, $N_A$, Low Temperature CV, Surface Potential and $D_{it}$ Estimation

Figure 5.2(a) shows the CV plots of the Chemox sample at ten different frequencies measured at room temperature. Since 4284 LCR meter was programmed to conduct capacitance measurement at 10 different frequencies per voltage, the sweep rate is very low (approximately 5 s). The sweep was done from inversion to accumulation region. The frequency dispersion in the negative region (accumulation region) is due to the substrate resistance. However, it is not all about $R_s$. It is because before any evidence, one can easily argue about capacitance values in the accumulation region. After a simple $R_s$ correction (Figure 5.2(b)), the dispersion is reduced but is still present unlike silicon devices [35]. This dispersion is not acceptable for further analysis since it didn’t give a robust value in the accumulation region. Furthermore, the dispersion here is mainly due to the interface states and large interface defects density, which causes Fermi level pinning before entering the accumulation region. The interface can, therefore, respond to the frequency below or equal to 10 KHz and gives a pseudo accumulation region. But can
the value be used here for calculating $C_{ox}$ or EOT? The answer is yes, since there is no major difference between interface capacitance ($C_{it}$) or bulk capacitance ($C_B$), and both of them are added to the substrate capacitance ($C_s$). Moreover, the measured capacitance in the accumulation region should never be above oxide capacitance ($C_{ox}$) unless affected by the DC leakage current [57]. The low frequency data, therefore, can be used here to calculate EOT if consider low frequency capacitance in accumulation region ($C_{lf,acc}$), which is approximately equals to $C_{ox}$. Alternatively, Maserjian method [113] can be used to estimate an accurate $C_{ox}$ after using Equation (5.1). Nevertheless, the results should be comparable with $C_{lf,acc}$. EOT results of three different samples are listed in Table 5.1.

$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{2kT}{qC_{ox}V_G - V_{FB}}$$  \hspace{1cm} (5.1)

<table>
<thead>
<tr>
<th>Sample name</th>
<th>EOT (nm)</th>
<th>$V_{FB}$ (V)</th>
<th>Minimum $D_{it}$ (cm$^{-2}$eV$^{-1}$)</th>
<th>$J_g$ at -1V (mA/cm$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemox (wet)</td>
<td>0.93</td>
<td>0.45</td>
<td>$\sim$6x10$^{12}$</td>
<td>0.1</td>
</tr>
<tr>
<td>COR+SPAO$_x$ (dry)</td>
<td>1.02</td>
<td>0.24</td>
<td>$\sim$2x10$^{12}$</td>
<td>0.025</td>
</tr>
<tr>
<td>COR+O$_3$ (dry)</td>
<td>0.88</td>
<td>0.19</td>
<td>$\sim$6x10$^{12}$</td>
<td>0.2</td>
</tr>
</tbody>
</table>
As shown in Equation (5.1), it is necessary to know $V_{FB}$ before $C_{ox}$ can be obtained. A standard way, therefore, to estimate $V_{FB}$ is by using Equation (5.2) below, where $L_D$ is Debye length, $\varepsilon_s$ is electrical permittivity of Ge substrate, $V_t$ is thermal energy (26 mV), $N_A$ is bulk doping concentration, and $C_{ox}$ is the oxide capacitance per unit area. However, it is necessary to know the value of $N_A$ before calculating $L_D$. Moreover, error in estimation of $N_A$ will cause uncertainty in $C_{FB}$ result. This method is not robust to estimate $V_{FB}$. Therefore, method reported by R.J. Hillard et al. [114] was used to obtain $V_{FB}$. (Table 5.1). Only high frequency CV ($C_{hf}$) and $C_{ox}$ are involved to calculate to $V_{th}$. Differentiate $(1/(C_{hf}/C_{ox}))^2$ respect to $V_g$ and find corresponding gate voltage of the maxim value of the derivative, which is the threshold voltage.

\[
L_D = \frac{\varepsilon_s V_t}{q N_A} \tag{5.2a}
\]
\[
C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\varepsilon_s}} \tag{5.2b}
\]
Bulk concentration was calculated by Equation (5.3). It was under assumption that there was no interference from bulk defects and the interface was in the deep depletion region (1 MHz capacitance data at low temperature 100 K). The bulk concentration values of all three samples were around $10^{16}$ cm$^{-3}$.

$$N_A(W) = \frac{2}{q\varepsilon_s A^2 \frac{d(1/C^2)}{dV}}$$ \hspace{1cm} (5.3)

The original purpose of measuring CV at low temperature is to remove or decrease the interference of the minority carrier generation and reduce the impact of interface defects in the inversion region. It is obvious that capacitance significantly decreases in the inversion region when temperature is lowered (Figure 5.3). However, in the range from -1.5 V to 0 V, both frequency and temperature dispersion is also observed for all three samples. The reason of dispersion in accumulation region were extensively studied and explained by two models: border trap model [115] and disorder induced gap state (DIGS) model [116]. Kuzum et al. measured Ge MOSFET with GeON dielectric. The dispersion in accumulation region is not significant at 250 K due to a cap layer of SiO$_2$[117]. Significant accumulation region dispersion is reported by Herbert et al. [118] in p type III-V GaAs devices but not in n type. It is believed DIGS model is more reasonable for our device since the maxim $C_{acc}$ is still close to what is expected [116]. Another possible reason of dispersion in our samples is that Fermi level of our sample is pinned at flatband voltage (the device is not biased in the accumulation region due to large interface state density). This may explain why the mobility in the p substrate of Ge is lower than expected. The calculated $C_{FB}$ of Chemox sample using the Equation (5.2), is
about 0.005 pF/um². At 100 K, the capacitance measured at 1 MHz in Figure 5.4 shows a constant value around 0.0043 pF/um² when the gate voltage is decreasing below around 0.5 V ($V_{FB}$ of the Chemox sample). This proves that our sample is pinned at flatband due to the interface states.

![Figure 5.3](image)

**Figure 5.3** Corrected capacitance ($C_c$) at different frequencies and different temperatures is plotted as a function of gate bias for sample Chemox. (a) at five different temperature, (b) at 100 K and 300 K. Both frequency dispersion and temperature dispersion can be observed.

Another important observation in Figure 5.4 is that the capacitance value cannot reach the level 0.03 pF/um² at 100 K whereas at 300 K it can. There are two possible reasons: Fermi level statistics are changing when temperature is changing, and lower temperature causes Fermi statistics to be steeper instead of flat [119]; Alternatively, as Kuzum *et al.* explained [117], this change is due to the change of emission rate of defects as stated in Equation (5.4).

$$e_p = \frac{1}{r_p} = \sigma_p v_p N_v \text{ex} p \left( \frac{E_V - E_T}{kT} \right) = \sigma_p v_p n_i \text{ex} p \left( \frac{E_i - E_T}{kT} \right)$$  \hspace{1cm} (5.4)
where $\sigma_p$ is the defect cross section, $v_p$ is hole thermal velocity, $n_i$ is intrinsic carrier concentration, $N_v$ is effective density of states of hole at valence band, $E_v$ is valence band energy level, $E_i$ and $E_T$ is intrinsic energy level and defects energy level, respectively. It assumes that interface states only respond to valence band. Also, it is further assumed that keeping all the parameters same, the time constant, $\tau_p$ will increase when temperature decreases. In other words, for a specific frequency window like 100 Hz to 1 MHz, by varying temperature, interface state information in the bandgap can be obtained if Fermi level is not pinned.

![Figure 5.4](#) Corrected capacitance ($C_c$) of Chemox sample measured by different frequencies at 100 K is plotted as a function of gate bias. Frequency dispersion can be observed.

However, the second explanation is not reliable if Fermi level doesn’t move effectively as a function of temperature when the gate voltage is varying, especially when one considers $D_{it}$ as a function of bandgap. One could easily observe that Fermi level in
the bandgap near the flat band rarely moved under different temperatures (Figure 5.3). Because of this reason, it is observed no $V_{FB}$ shift due to capacitance interference (Figure 5.3) by varying the gate voltage at slow rate (5 s), therefore most interface states can response to the change in DC voltage. Presence of large interface states pinned the Fermi level in the device before it entered the accumulation region.

The above discussion suggests that the region of bandgap that was observed remained the same under different temperature. However, the time constant of those measured interface states is decreased when temperature is increased, therefore they can follow the ac signal after specific temperature. In Figure 5.3, it is observed that there is temperature dispersion at a specific frequency. The blue line up triangle (which is alternated between 0.01 pF/um$^2$ and 0.03 pF/um$^2$) is a good evidence for the first assumption. Figure 5.5 plots $C_{it}$ ($C_{it} = C_{acc} - C_{FB}$) as function of temperature ($1/kT$) at fixed frequency (100 KHz) to obtained the energy level of defects. Since $C_{it}$ is proportional to $e_p^2$ ($e=1/\tau$) (Equation 5.5), and Equation 5.4 shows the exponential dependence of $e_p$ on temperature ($1/kT$). The slope of plot in Figure 4.5 will have a value of $2E_a$. This defects energy level is very close to the value obtained by DLTS discussed later.

$$C = C_S + \frac{C_{it}}{1 + (\omega \tau_{it})^2}$$  \hspace{1cm} (5.5)
Figure 5.5 $C_{it}$ ($C_{it} = C_{acc} - C_{FB}$) as function of temperature ($1/kT$) at fixed frequency (100 KHz).

5.4 Results and Discussion

5.4.1 Impact of Interface Treatment

As discussed earlier, Ge/1nm-Al₂O₃/3.5nm-ZrO₂/TiN gate stacks (MOSCAPs) have three different interface treatments. The COR+SPAOₓ and COR+O₃ are dry treatments whereas Chemox is a wet processed interface. Figure 5.6 shows corrected the 1 MHz capacitances of three samples are plotted as a function of gate voltage at room temperature (device area is 40 $\mu$m × 40 $\mu$m). Considering $V_{FB}$, the dry treated interfaces exhibit more negative $V_{FB}$ shift compared to wet treated samples due to the presence of positive border traps for dry processed samples (see DLTS section). The EOT, on the other hand, for the dry processed interface (COR+O₃), shows a clear increase compared
to other two samples, as reported earlier (Table 5.1), COR+SPAOₙ shows highest EOT because of 1 nm additional SPAO. When plot the dry (COR+SPAOₓ) and dry (COR+O₃) processed CV at 100 K as a function of frequency (Figure 5.7), stark difference was observed between them. This indicates there exists certain bulk defects in the upper half bandgap that can follow low frequency at 100 K for dry (COR+O₃) processed sample. Even though the dry process interface exhibited excellent room temperature frequency dependent CV, in the depletion it indicates the existence of interface states near valence band for both wet and dry processed COR+O₃ samples. For COR+SPAOₓ samples, on the other hand, a reduced interface state was observed (Figure 5.6).

![Figure 5.6 Corrected 1 MHz capacitances (Cᵦ) of three samples are plotted as a function of gate voltage at room temperature. Device area is 40 μm × 40 μm.](image)

Figure 5.6 Corrected 1 MHz capacitances (Cᵦ) of three samples are plotted as a function of gate voltage at room temperature. Device area is 40 μm × 40 μm.
Figure 5.7 Corrected capacitance ($C_c$) is plotted as function of gate voltage at 100 K, (a) COR&SPAOn, (b) COR&O₃.

$D_{it}$ was calculated by conductance method [8] and further verified by capacitance spectroscopy method. It is important that the $D_{it}$ data are plotted as function of bandgap since it is necessary to relate the gate voltage to surface potential. Software CVC.2.0 was used to generate the surface potential [95] where $C_{ox}$, $V_{FB}$, and bulk concentration were the input.

In Figure 5.8(a), it was clearly observed that COR&SPAOn processed interface had lower interface states density at room temperature than the other two different processed samples. This further suggests that SPAO samples have improved interfacial layer quality in terms of interface defects density. At low temperature (100 K) SPAO samples also show a low mid-gap $D_{it}$. Figure 5.8(b) compares the $D_{it}$ values as a function of gate voltage measured by capacitance spectroscopy method and conductance method. Similar results were also observed by capacitance spectroscopy. Moreover, the difference of $D_{it}$ estimated at 100 K and 300 K (Figure 5.8(b)) is mainly due to their time constant variation with temperature. It is, therefore, imperative to understand the defect energy levels.
Figure 5.8 (a) Interface state density ($D_{it}$) is plotted as function of bandgap for three samples at two temperatures (100 K and 300 K), other $D_{it}$ as a function of temperature are within this range; (b) Calculated $D_{it}$ of three samples as a function of gate voltage by capacitance spectroscopy method and conductance method.

The low interface state density in COR+SPAO$_x$ samples indicates formation of an interfacial layer (IL) constituting GeO$_x$ with a possible unit cell of GeO$_2$ layer [120]. As mentioned by Zhang et al., dielectric constant decreases once IL changes from GeO$_x$ to GeO$_2$ layer because of plasma oxidation. That was clearly evident in CV measurement (Figure 5.6) and EOT estimation (Table 5.1) for COR+SPAO$_x$ samples. In addition, an increase in GeO$_x$ during SPAO treatment may be possible enhancing the EOT. In either way, interface treatment by SPAO plasma tends to passivate the interface further [120] by reducing the $D_{it}$ (Figure 5.8) for COR+SPAO$_x$ samples at the cost of an increase in EOT. It is believed former mechanism is more responsible for $D_{it}$ reduction. The chemical processes (both wet and dry) Chemox and COR+O$_3$ failed to passivate the interface with only formation of GeO$_x$ even though the EOT was decreased.

5.4.2 Deep Level Transient Spectroscopy (DLTS)

The interface of dry and wet processed Ge/high-k MOS structures was investigated by deep level transient spectroscopy (DLTS) [9] to further understand the nature of these
interface defects. The MOS devices were pulsed from mid-bandgap to accumulation region, then return to mid-bandgap to obtain the majority carrier trap emission information. Figure 5.9(a-c) shows the DLTS data for three corresponding interface treatments. The Arrhenius plots are shown in Figure 5.9(d). Comparing the DLTS spectrum, it can be concluded that dry processed devices have discrete border traps at the interface (H3 and H5). No such traps at the interface of wet processed MOS devices (Chemox) were observed. For all three samples, however, interface like traps were detected (H1, H2, H4). The defects density, $N_\text{t}$, was estimated by Equation (3.7), where $\Delta V$ is obtained from CV plot using measured $\Delta C$. The results suggest that it is around $10^{13}$ cm$^{-2}$eV$^{-1}$ for all three samples. This further confirms that the origins of interface state density, $D_{\text{it}}$, observed for these samples (Figure 5.8) earlier are mainly due to these traps. The time constants of these defects are in the order of milliseconds that were observed earlier (Figure 5.3). This can further explain the observed fluctuations in frequency dispersions curves at low temperatures in Figure 5.3. The relatively negative $V_{\text{th}}$ shift in dry processed samples compared to wet process samples can also be explained because of presence of additional hole traps at the interface. Table 5.2 summarized the energy levels and cross sections of these traps. The observed energy levels indicate the presence of a shallow level (H3) for COR+SPAO$_3$ sample as compared to a deep level (H5) for COR+O$_3$ samples. The impact of these levels on $D_{\text{it}}$ is clearly obvious as deep levels contribute significantly to interface state density.
Table 5.2 Arrhenius Plot Fitting Results

<table>
<thead>
<tr>
<th></th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_T-E_V$ (eV)</td>
<td>0.16</td>
<td>0.1</td>
<td>0.18</td>
<td>0.04</td>
<td>0.45</td>
</tr>
<tr>
<td>$\sigma$ (cm$^2$)</td>
<td>$6.3\times10^{-17}$</td>
<td>$1.3\times10^{-19}$</td>
<td>$3.6\times10^{-19}$</td>
<td>$5.8\times10^{-21}$</td>
<td>$8.6\times10^{-13}$</td>
</tr>
</tbody>
</table>

It is also necessary to discuss how to understand the impact of frequency on the measurement, since DLTS was conducted only at 1MHz, and conductance method was conducted at different frequencies. Since capacitance transient is to obtain capacitance value after around 200 hundred microseconds (Boonton 7200 capacitance meter), only 1 MHz frequency sine wave can be averaged out to obtain capacitance value in this range and to follow possible transient over time. CV measurement is usually under quasi static condition, where the data is more accurate than Boonton 7200. During the DLTS measurement, the time constant is more interesting than inaccurate capacitance value, as long as capacitance transient is robust. DLTS is more advanced to study the energy level than trap density since capacitance value is not as accurate as quasi static measurement.
Figure 5.9: Deep level transient spectrum for three samples, $C_{t1}$ and $C_{t2}$ are two capacitance values sampled at two different times ($t_1 < t_2$), $T$ is temperature, $\ln$ is natural logarithm, $\tau$ is time constant calculated as $\frac{t_2 - t_1}{\ln(t_2/t_1)}$. (a) simple chemical oxidation (Chemox), (b) chemical oxide removal (COR) followed by 1 nm oxide by slot-plane-antenna (SPA) plasma (COR&SPAOx), (c) COR followed by vapor O$_3$ treatment (COR&O$_3$). (d) is Arrhenius plot for all three samples.

5.4.3 I-V Characteristics

Figure 5.10 compares the gate leakage current density for different type of samples and it is listed in Table 5.1 as well. The SPAO processed interface shows lower current density (Figure 5.10(a)). As observed in Figure 5.10(b), the gate leakage current density increases mainly due to lower EOT. This is primarily due to reduced tunneling barrier caused by interfacial layer thinning, therefore increasing the tunneling constant. The samples with SPA plasma enhanced interfacial layer showed the lowest tunneling leakage current. It
was previously reported that SPA plasma helps better oxide growth with reduced impurities [7]. In addition, the SPA plasma makes atomically flat surface and interface, which helps the reduction in leakage current density [121, 122]. This further confirms that dry processed, especially COR+SPAO$_x$ interface is superior.

**Figure 5.10** Comparison of gate leakage current density ($J_g$), (a) $J_g$ as a function of gate voltage, (b) $J_g$ as a function of EOT for different splits.

### 5.4.4 Evidence of Edge Capacitance in Pseudo-Accumulation Region

The purpose of this section is to further study the source of capacitance obtained at high frequency. When capacitance is measured for device with different side length (10 μm, 20 μm, 30 μm, 40 μm, 50 μm, 100 μm), it is found that capacitance at high frequency (1 MHz) has a linear dependence of device perimeter (Figure 5.11(a)), this suggest that in this region, capacitance may be contributed from the edge. On the other hand, Figure 5.11(b) shows capacitance at low frequency has dependence on device area. Maxim of capacitance value obtained in pseudo-accumulation region cannot guarantee a large substrate capacitance. They maybe interface capacitance ($C_{it}$) or edge capacitance ($C_L$).
Figure 5.11 Corrected capacitance ($C_c$) measured at -1.5 V (pseudo-accumulation) were measured at two different frequencies, 1 MHz and 1 KHz, respectively. (a) At 1 MHz, the capacitance is contributed from edge and it has linear dependence of device perimeter (b) At 1 MHz, the capacitance is contributed from interface states in area and it has linear dependence of device area.

5.5 Conclusion

It is demonstrated an accurate parameter estimation method for Ge/ALD 1nm-Al$_2$O$_3$/ALD 3.5nm-ZrO$_2$/ALD TiN MOS capacitors with three different interface treatments. COR&SPAO$_x$ samples show superior interfacial characteristics at room temperature. After evaluating several parameters like EOT, flatband voltage, bulk doping, and interface defects density, COR&SPAO$_x$ (dry) has better interface quality than Chemox (wet) and COR&O$_3$ (dry) processed sample. However they all have higher $D_{it}$ values in the order of $10^{13}$ cm$^{-2}$eV$^{-1}$, which causes Fermi level more or less pinned, which is confirmed by low temperature measurements. Dry processed sample (COR&SPAO$_x$ and COR&O$_3$) has more negative $V_{th}$ shift due to existing border trap characterized using DLTS method. The levels of leakage current of three samples confirm EOT values. Therefore, COR&SPAO$_x$ (dry) has lowest trap assisted tunneling effect leading to lowest leakage current.
CHAPTER 6
IMPACT OF SLOT PLANE ANTENNA ANNEALING ON CARRIER
TRANSPORT MECHANISM AND RELIABILITY ON ZrO$_2$/Al$_2$O$_3$/Ge GATE STACK

6.1 Research Motivation of SPAO on High-k Layer

Ge has been extensively studied to replace Si due to its higher electron and hole mobility [103, 104], which can enhance metal oxide semiconductor field effect transistor (MOSFET) speed without any physical scaling. Before the development of high-k oxide in the industry, the hydroscopic property of GeO$_2$ impeded the implementation of Ge MOSFET since SiO$_2$ has a better stability as native oxide layer, grown on Si [42]. Nevertheless, after metal/high-k gate stack was introduced, the formation of appropriate thickness of GeO$_x$ or GeO$_2$ between high-k and Ge substrate was reported to have low interface state density ($D_{it}$) [123]. Growing a few Ångström GeO$_2$/GeO$_x$ intentionally is, therefore, necessary to obtain good interface quality for high-k/Ge gate stack [124]. The thickness of GeO$_2$/GeO$_x$ will impact the overall equivalent oxide thickness (EOT) because of its low dielectric constant compare to high-k layer [125]. Therefore the tradeoff between EOT and $D_{it}$ is inevitable issue. Besides, interface layer treatment of Ge and controlling its thickness are also critical to further scaling EOT below 1 nm [126]. Popular high-k dielectrics like HfO$_2$ and ZrO$_2$ have already been integrated into CMOS technology. It is reported that HfO$_2$/Ge gate stacks show a larger capacitance voltage (CV) hysteresis than Al$_2$O$_3$ [47], and Al$_2$O$_3$ can block electron injection from substrate effectively by large conduction band offset related to Ge substrate [47]. Moreover, ZrO$_2$ showed lower leakage current than HfO$_2$ with similar dielectric constant [127]. Therefore, a bilayer high-k stack (ZrO$_2$/Al$_2$O$_3$) was used in this study. The IL quality (interface state
density, $D_{n}$), EOT, and XPS analysis of IL atomic composition of these stacks was reported in a previous work [128], and they are summarized in the Table 6.1.

<table>
<thead>
<tr>
<th>Sample name</th>
<th>EOT (nm)</th>
<th>$V_{FB}$ (V)</th>
<th>$D_{n}$(cm$^{-2}$eV$^{-1}$) at $E_i$-0.15 eV</th>
<th>Interfacial Layer Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ge/SPAO/Al$_2$O$_3$/ZrO$_2$</td>
<td>0.98</td>
<td>0.015</td>
<td>2.14×10$^{12}$</td>
<td>GeO$_x$</td>
</tr>
<tr>
<td>Ge/Al$_2$O$_3$/SPAO/ZrO$_2$</td>
<td>1.29</td>
<td>-0.228</td>
<td>3.93×10$^{11}$</td>
<td>GeO$_x$</td>
</tr>
<tr>
<td>Ge/Al$_2$O$_3$/ZrO$_2$/SPAO</td>
<td>1.13</td>
<td>-0.362</td>
<td>6.87×10$^{11}$</td>
<td>GeO$_2$</td>
</tr>
</tbody>
</table>

### 6.2 Device Information and Experimental Procedure

MOS capacitors were fabricated on 300mm Ge epitaxially grown on Si with a graded SiGe buffer layer. After Ge epi, the samples were subjected to chemical oxide removal (COR) using a TEL Certas$^\text{TM}$. The samples were subjected to three SPAO annealing sequences: (i) Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ (SPAO before high-k), (ii) Ge/Al$_2$O$_3$/SPAO/ZrO$_2$ (SPAO between high-k), (iii) Ge/Al$_2$O$_3$/ZrO$_2$/SPAO (SPAO after high-k). The ALD Al$_2$O$_3$ layer was grown using trimethyl Aluminum as Al precursor and tetrakis (ethylmethylamido) zirconium as the Zr precursor. H$_2$O was used as the oxidant in both cases at a deposition temperature of 250°C. ALD TiN was used as the metal gate. The current-voltage (I-V) and TDDB measurement were performed using HP5156B semiconductor parameter analyzer. The devices under test are 100 μm× 100 μm. I-V characteristics were obtained at different temperatures (25 °C, 50 °C, 80 °C, 110 °C). To obtain the statistical information of TDDB, 30 devices were measured at each stress condition. The charge to break down value ($Q_{BD}$) is extrapolated by four stress points to evaluate the quality of oxide and IL.
The reliability of high-k layer depends on gate leakage current density level \((J_g)\) and charge to breakdown \((Q_{BD})\). The current density not only determines performance of memory and chip circuit, but also may affect how fast the oxide layer degrades. After certain amount of the injection of the carriers through oxide layer, it can be irreversibly broken down. Moreover, these two parameters are somehow correlated if both of them are only dependent on the oxide layer quality. The time dependent dielectric breakdown (TDDB) measurement is now showing polarity dependence on thin devices since I-V depends on barrier condition at the interface (gate/oxide and oxide/substrate) [129, 130] and it is expected that bilayer structure will further enhance this phenomenon since different electric field across the layers and variance in dielectric quality [131]. Moreover, as EOT is scaling down to below 1 nm, interfacial layer will have more impact on final TDDB performance [132] if a soft breakdown time is measured instead of hard breakdown. If the degradation is due to the IL, a lower current density cannot predict a longer lifetime of a dielectric layer in terms of device stability. In Section 6.3, the carrier transport mechanisms were evaluated for three samples, namely, before high-k ALD (Atomic Layer Deposition), in between two different ALD high-k layers, and after ALD.

**Figure 6.1** Application of SPAO at different stages of gate stack deposition.
high-k, to understand the effect of SPAO on the gate leakage current and carrier transport mechanism in dielectrics [73]. The trap distributions observed in this experiment can significantly impact the reliability of the dielectric. Therefore, samples were subjected to TDDB measurements [10] to further understand the reliability of p-Ge/Al₂O₃/ZrO₂/TiN gate stacks under different SPAO exposure. Charge to breakdown ($Q_{BD}$) was estimated from the TDDB measurements which were carried out at different voltage stress conditions. Subsequently, voltage acceleration factor ($AF$) was extracted. TDDB and I-V characteristics of three samples were studied by using both gate electron injection (GEI) and substrate electron injection (SEI) modes. This gives an overall evaluation of the high-k oxide and IL quality under different SPAO conditions.

### 6.3 Carrier Transport Mechanism

To understand the carrier transport mechanisms and/or prior to applying the physical carrier transport models, it is important to observe I-V characteristics at different temperatures. At negative voltage range (gate electron injection) Figure 6.2(a) and Figure 6.2(b) show I-V dispersion at different temperatures for sample Ge/SPAO/Al₂O₃/ZrO₂ and sample Ge/Al₂O₃/SPAO/ZrO₂ respectively. On the other hand, the temperature dependence of I-V is greatly reduced for sample Ge/Al₂O₃/ZrO₂/SPAO (Figure 6.2(c)). It is reported that SPAO can effectively reduce oxide traps [133]. Therefore, it is believed that reduced temperature dependence is mainly due to reduced traps density in ZrO₂ and Al₂O₃ layers. Figure 6.2(d) compares the leakage current after gate voltage is subtracted by flat band voltage ($V_{FB}$). It shows that sample Ge/Al₂O₃/ZrO₂/SPAO has lowest leakage current, followed by sample
Ge/Al$_2$O$_3$/SPAO/ZrO$_2$. Since the I-V is not significantly dependent on temperature (Figure 6.2(c)) at high electric field, FN tunneling is believed to be the dominant mechanism [134] for both GEI and SEI. Electric field is calculated by Equation (3.16) [135] assuming the dielectric constants of ZrO$_2$ and Al$_2$O$_3$ as 25 and 9, respectively. The barrier heights were then extracted by fitting $\ln(J_0 E^2)$ to $1/E$ as Equation (3.17) indicates, where $E$ is the electric field, $q$ is the electronic charge, $m_o, m_{ox}$, are the electron mass in free space and in the oxide, respectively; $\hbar$ is the Planck’s constant, and $\phi_b$ is the barrier height. The electrons mass used in this calculations for ZrO$_2$ and Al$_2$O$_3$ are 0.5$m_o$ and 0.3$m_o$, respectively [136, 137]. In Figure 6.3, $\ln(J_0 E^2)$ is plotted as function of $1/E$ and different barrier heights were obtained from the slopes in both GEI and SEI modes respectively. Metal (TiN) to dielectric barrier height, $\phi_{b(ZrO2)}$ at high field during gate injection (barrier height of ZrO$_2$) is calculated as 1.3 eV. The estimated $\phi_{b(ZrO2)}$ is lower than the theoretical calculation (1.6 eV) due to the breakdown of high-k before a sufficiently high electric field was applied. Substrate to dielectric barrier height $\phi_{b(AlO3)}$ at high field was estimated during substrate (Ge) injection (barrier height of Al$_2$O$_3$) is 2.2 eV assuming electrons tunnel through thin GeO$_x$. On the other hand, $\phi_{b(AlO2)}$ is very close to the expected theoretical value [47].
Figure 6.2 Current density ($J_g$) is plotted as a function of gate voltage ($V_g$) at four different temperatures (25 °C, 50 °C, 80 °C, 110 °C) for three different samples, (a) Ge/SPAO/Al$_2$O$_3$/ZrO$_2$, (b) Ge/Al$_2$O$_3$/SPAO/ZrO$_2$, (c) Ge/Al$_2$O$_3$/ZrO$_2$/SPAO. (d) $J_g$ as a function of $V_g$-$V_{FB}$ is plotted for the above three samples at 25 °C where $V_{FB}$ is the flat band voltage.

Figure 6.3 $\ln(J_gE^{-2})$ is plotted as function of $1/E$ as FN tunneling for sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO, $J_g$ is current density and $E$ is electric field. (a) gate electron injection mode (GEI), (b) substrate electron injection mode (SEI).
Note that the current voltage characteristics in SEI mode at low field, in the range from -0.5 V to 0.5 V, low to medium $E$ field range for sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO (Figure 6.2(c)) I-V is clearly a function of temperature. Therefore, Poole-Frankel (PF) and hopping conduction (HC) mechanism is possibly the dominant mechanism in this range and fits well I-V characteristics. The current expression for HC is given by Equation (3.18) [73].

In Equation (3.18), $a$ is the mean hopping distance (i.e., the mean spacing between trap sites), $n$ is the electron concentration in the conduction band of the dielectric, $v$ is the frequency of thermal vibration of electrons at trap sites, and $E_a$ is the activation energy, namely, the energy level from the trap states to the bottom of conduction band. As shown in Figure 6.4(a), HC model fits $J_g$-$T$ (current temperature function) characteristic well in medium $E$ field range. Subsequently, the slope, $S$, ($S = qaE$-$E_a$) obtained from Figure 6.4(a), can be plotted as function $E$ field. In Figure 6.4(b), the fitted slope is the product of mean hopping distance ($a$) and electronic charge ($q$), and intercept of slope is the activation energy. The calculated values $a$ and $E_a$ are 0.3 nm and 0.16 eV respectively in sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO.
Figure 6.4 (a) Current density $\ln(J_g)$ is plotted as function of $1/kT$ for sample Ge/$\text{Al}_2\text{O}_3$/ZrO$_2$/SPAO, in SEI mode. (b) Slope value ($S = qaE - E_a$), which is obtained from (a) is then plotted as function electric field $E$ (SEI mode).

As discussed above, sample Ge/$\text{Al}_2\text{O}_3$/ZrO$_2$/SPAO shows FN tunneling in the high $E$ field range since both ZrO$_2$ and Al$_2$O$_3$ were exposed to SPAO. On the other hand, ZrO$_2$ layers were not subjected to SPAO for both the samples Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ and Ge/$\text{Al}_2\text{O}_3$/SPAO/ZrO$_2$. Therefore, both HC and PF mechanisms seems to be dominated for the entire range of measured I-V characteristics (high and low electric field range). The PF current expression is given by Equation (3.19) [73], where $\mu$ is the electronic drift mobility, $N_C$ is the density of states in the conduction band, $q\phi_t$ is the trap energy level, and the other notations are the same as defined in Equation (3.18). The $\phi$ has the physical meaning similar to that of $E_a$ in HC model. The symbol difference is just to differentiate that the values that are obtained using different current models. The conduction mechanism, PF is different from that of HC. In PF, electrons in trap centers are thermally excited to conductance band of oxide, and subsequently relaxed to another trap center. In case of HC, electrons transit between trap sites by trap assisted tunneling.
Figure 6.5 Hopping Conduction and Poole-Frenkel emission (PF) were used to fit the I-V characteristics for sample Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ (solid symbol) and sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$ (open symbol) at both GEI mode (black symbol, squares) and SEI mode (red symbol, circles). (a) Slope value ($S = qaE - E_a$) is plotted as function electric field $E$. (b) Slope value ($S = \beta E^{1/2} - q\phi_t$) is plotted as function $E^{1/2}$, where $eta = q^3/\pi\varepsilon_r\varepsilon_0$. The slope value $S$ is the fitting value from ln$(J_g)$ versus 1/kT.

In Figure 6.5, both HC and PF model were applied to sample Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ and sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$. The slope value $S$ is the fitting value from ln$(J_g)$ versus 1/kT (not shown here). HC model (Figure 6.5(a)) gives same trap energy level $E_{a1}$ as 0.09 eV for both GEI mode (black open symbol) and SEI mode (red open symbol) in sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$. On the other hand, HC model gives $E_{a1}$ as 0.09 eV for GEI mode (black solid symbol) and $E_{a2}$ as 0.22 eV for SEI mode (red solid symbol) in sample Ge/SPAO/Al$_2$O$_3$/ZrO$_2$. In a brief, for GEI mode, electrons tunnel through ZrO$_2$ layer assisted by the same trap centers ($E_{a1}$), since ZrO$_2$ layer of both samples were not subjected to SPAO. For SEI mode, electrons tunnel via two different trap centers, $E_{a1}$ and $E_{a2}$, for sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$ and sample Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ respectively. As mentioned earlier, SPAO can significantly remove trap center $E_{a2}$ (0.22 eV) in GeO$_x$/Al$_2$O$_3$ layer for the Ge/Al$_2$O$_3$/SPAO/ZrO$_2$
device, when SPAO is processed after Al₂O₃ deposition. The both trap energy levels $E_{a2}$ and $E_{a1}$ were observed in Ge/SPAO/Al₂O₃/ZrO₂ sample.

When PF model (Figure 6.5(b)) was used similar behavior was observed as that of HC model. The trap energy level, $\phi_1$ of 0.13 eV calculated by PF model, was observed for GEI mode (black solid and open symbols) in both the samples Ge/SPAO/Al₂O₃/ZrO₂ and Ge/Al₂O₃/SPAO/ZrO₂. In SEI mode one trap energy level $\phi_1$ was observed with identical value of 0.13 eV. On the other hand, for SEI mode (red and black solid symbols) in Ge/SPAO/Al₂O₃/ZrO₂ a second trap energy level, $\phi_2$, was observed in addition to $\phi_1$ (0.13 eV). Different symbols were used as they are calculated by different model. Both HC model and PF model are similar in trap-rich oxide layers with the difference in carrier transit process between the trap sites. HC gives one more parameter, the hopping distance, $a$, of about 0.3 nm from the slope in Figure 6.5(a). Carrier transport mechanisms for GEI and SEI mode are marked in band diagram as Figure 6.6(a) and Figure 6.6(b) respectively, and mechanisms and trap energy levels are summarized in Table 6.2.

| Sample name                  | Transport Mechanism† | Trap energy Level $\phi$ (eV) | $Q_{BD}$ (C/µm²) at |1V| GEI/SEI | Acceleration factor (V⁻¹) GEI/SEI |
|-----------------------------|----------------------|--------------------------------|---------------------|---------|----------------------------------|
| Ge/SPAO/Al₂O₃/ZrO₂          | PF/HC                | 0.13eV/0.27eV                  | ~10⁰/10¹             | 10¹/10⁻⁴ | -6.1/-6.7                        |
| Ge/Al₂O₃/SPAO/ZrO₂          | PF/HC                | 0.13eV/0.13eV                  | ~10²/10⁴             | 10⁻⁷/3.8 | -10.7/-3.8                       |
| Ge/Al₂O₃/ZrO₂/SPAO          | FN                   | NA                             | ~10⁻³/10⁻⁴           | 17.9/-1.6 |                                  |

†Transport mechanisms were extracted at high electric field range.

It is, therefore, clear that SPAO contributes to reduction of a trap sites in the GeOₓ/Al₂O₃ layer when SPAO is processed after Al₂O₃ deposition. This reduction
process was not observed if SPAO is processed prior to Al$_2$O$_3$ deposition (Ge/SPAO/Al$_2$O$_3$/ZrO$_2$ sample). It is further observed that sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO has the best performance for GEI mode, since most of the oxide trap centers were removed from the dielectric by SPAO [133]. The trap distribution in the dielectric and the interfacial layer (IL) will have significant impact on dielectric degradation. We have also observed difference in as measured by the XPS and reported in our previous work [128]. Therefore, the performance of SEI mode is dependent on IL as will be more critical. The TDDB study was, therefore, employed to further understand the contribution of the trap distribution observed above to the degradation process, discussed in the following section.

![Diagram of carrier transport mechanisms](image)

**Figure 6.6** Carrier transport mechanisms are explained in band diagram for both GEI mode and SEI mode. (a) GEI mode, band diagram is simulated under $V_g = -1.5$ V. (b) SEI mode, band diagram is simulated under $V_g = 1$ V.

### 6.4 Time Dependent Dielectric Breakdown Performance

The TDDB characteristics of all three samples were studied. The charge to breakdown, $Q_{BD}$, and voltage acceleration factor ($AF$) were the two main parameters used in both GEI
and SEI modes to evaluate the dielectric quality and any contribution of the interfacial layer. The oxide breakdown in this work is defined as the abrupt change of sampling gate current ($I_g$). Although secondary abrupt changes of $I_g$ exist (not shown here) that were not considered. $Q_{BD}$ was used instead of time to breakdown ($T_{BD}$) to further understand the trap distribution observed in the I-V characteristics. Weibull distribution was used to explain TDDB statistic shown in Equation (3.21) [75], where $\beta$ and $\eta$ are shape parameter and they are constant. $\beta$ measured in Equation 3.21(a) is about 0.7 for all samples and it is stress voltage independent (not shown here). But it depends on oxide thickness and trap sphere size ($a_o$) theoretically as Equation (3.23a) indicates [76] where $\alpha$ is a parameter describing the correlation between $t_{ox}$ and $\beta$.

Figure 6.7 shows $\beta$ values of all samples for both GEI and SEI mode. The same value of 0.7 for all samples indicates that $a_o$ is similar in all samples. Experimentally, the observed $\beta$ value is linearly related to oxide thickness (Equation (3.23(b)), where $\gamma$ is the coefficient and $t_{INT}$ is thickness of IL [77]. It is difficult to distinguish between intrinsic breakdown ($\beta \geq 1$) and extrinsic breakdown ($\beta < 1$), as $\beta$ value is decreased [10, 76, 138]. This low $\beta$ value measured in this work also indicated that $Q_{BD}$ has a broad statistical distribution (three orders). The final $Q_{BD}$ is selected at 90% of its value for devices breakdown.
Figure 6.7 $\beta$ values are around 0.7 and similar for all samples and stress conditions, which were obtained by Weibull plot.

Figure 6.8 shows that $Q_{BD}$ values were obtained at four different gate voltages in both GEI mode (solid symbol) and SEI mode (open symbol) for all samples. The stress voltages were selected based on voltage ramping breakdown measurement in previous work [128]. $Q_{BD}$ and $AF$ parameters were summarized in Table 6.2, where $AF$ values were obtained based on Equation (3.22). At GEI mode (Figure 6.6(a)), the amount of tunneling electrons dependent on the quality of ZrO$_2$, which is the first layer that electron has to tunnel through. The $AF$ value of -6.1 is the largest for sample Ge/SPAO/Al$_2$O$_3$/ZrO$_2$, followed by an $AF$ value of -10.7 for sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$, and sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO has the lowest $AF$ value of -17.9. Also, it was observed that Ge/Al$_2$O$_3$/ZrO$_2$/SPAO has the largest $Q_{BD}$. Therefore, it can be concluded that the SPAO can affect the TDDB characteristics GEI mode by significantly removing the traps from the dielectric layers[133] ZrO$_2$ and Al$_2$O$_3$. The
dielectric quality was enhanced by SPAO in terms of $AF$ and $Q_{BD}$. If traps are removed from only the $\text{Al}_2\text{O}_3$ layer the $AF$ value increased and $Q_{BD}$ was decreased.

Figure 6.8 Charge to breakdown value ($Q_{BD}$) is plotted as function of gate voltage ($V_g$) for both GEI mode (solid symbol) and SEI mode (open symbol).

The same conclusion cannot be applied to SEI mode, since TDDB degradation is also affected by the degradation of IL ($\text{GeO}_2$ or $\text{GeO}_x$). XPS measurement from earlier work on these devices concluded that sample $\text{Ge/Al}_2\text{O}_3/\text{ZrO}_2$/SPAO formed $\text{GeO}_2$ IL, while the other two samples formed $\text{GeO}_x$ IL [128]. Some studies reported that ALD process will decompose $\text{GeO}_2$ into $\text{GeO}_x$ [139] since $\text{GeO}_2$ is not thermally stable [140]. This explained why only sample $\text{Ge/Al}_2\text{O}_3/\text{ZrO}_2$/SPAO has $\text{GeO}_2$ as IL since subsequent ALD process decomposed the $\text{GeO}_2$ to $\text{GeO}_x$ for other two samples. As open symbols show in Figure 6.8, sample $\text{Ge/Al}_2\text{O}_3/\text{ZrO}_2$/SPAO has the largest $AF$ and lowest $Q_{BD}$ among three samples due to formation of $\text{GeO}_2$. On the other hand, samples
Ge/SPA0/Al2O3/ZrO2 and Ge/Al2O3/SPA0/ZrO2 show better $Q_{BD}$. The formation of unstable fragmented IL causes an increase in $AF$ values (Table 6.2), and only sample Ge/SPA0/Al2O3/ZrO2 has similar $AF$ values at both GEI and SEI mode, which indicates a similar breakdown process. Note that although GeO2 shows worst TDDB result among of three different SPA0 treatments. This suggests that GeO2 has the worst resistance to stress in terms of device stability compared to GeOx. Electrons transit through thinner GeOx IL, and TDDB measured the quality of Al2O3 rather than IL. After taking into account of $D_{it}$ and EOT (Table 6.2), it can be concluded that formation of GeOx or GeO2 is helpful for improvement of $D_{it}$ values at the cost of increased EOT values and formation of GeOx is better for device reliability rather than GeO2 because of rapid degradation of GeO2 can cause gate stack instability, although sample Ge/Al2O3/ZrO2/SPA0 has the best performance of I-V.

Figure 6.9 Gate leakage current ($I_g$) is plotted as function sampling time for all samples.
Figure 6.9 shows that typical stress induced leakage current (SILC) in TDDB measurement, which was not observed in GEI mode. This SILC is due to additional electric field caused by electrons trapped in the Al$_2$O$_3$ and GeO$_x$ layer. As more electrons are trapped, the increased electric field enhances the process of tunneling. Another phenomenon, observed at SEI mode, was that leakage current was decreased initially before SILC process for Ge/Al$_2$O$_3$/ZrO$_2$/SPAO. This is possibly due to GeO$_2$ instability, and initially GeO$_2$ transformed to GeO$_x$, which has larger conduction band offset to block electron from substrate [141, 142]. It was also observed that there exists secondary breakdown in sample Ge/Al$_2$O$_3$/ZrO$_2$/SPAO, which is another evidence of IL degradation (not shown here). Therefore, TDDB measurements at SEI mode were representing the quality of IL rather than overall oxide (Figure 6.8).

6.5 Conclusion

The TiN/ZrO$_2$/Al$_2$O$_3$/p-Ge gate stacks were studied for carrier transport mechanisms and reliability in terms of oxide breakdown. Different SPAO annealing conditions reveals that although SPAO can effective remove traps in high-k dielectrics and subsequently reduce the leakage current, the formation of GeO$_2$/GeO$_x$ is inevitably impact the reliability. Trap energy levels in ZrO$_2$ and Al$_2$O$_3$ are found to be 0.13 eV and 0.28 eV, respectively by fitting Poole-Frenkel emission model. TDDB found out that GeO$_2$ can be degraded faster than GeO$_x$ since GeO$_x$ as electron can transit through IL even though both of the layers can gives a similar $D_{it}$ values of around $5\times10^{11}$ cm$^{-2}$eV$^{-1}$. If the ALD process decomposes GeO$_2$ into a stable GeO$_x$ the overall performance of the gate stack is more stable as shown in case of sample Ge/Al$_2$O$_3$/SPAO/ZrO$_2$. 
CHAPTER 7
FLICKER NOISE MECHANISM ON THICK AND THIN OXIDE FINFET

7.1 Motivation of Noise Model Study on FinFET

Noise characterization in terms of reliability became more and more critical in continuously scaling CMOS technology, since working overdrive voltage margin is reduced as transistors shrink. Noise study (especially 1/f noise) is necessary to be well understood to prevent reliability issues [66, 143]. Additionally, 1/f noise can be served as a non-destructible diagnostic tool to investigate the quality of interfaces of oxide/substrate, gate/oxide and oxide itself [68, 144, 145]. FinFET structures were implemented in an industrial setting to replace original planar structure by FinFETs for 22 nm technology node to reduce short channel effect [146]. As far as flicker noise is concerned, FinFET transistor has to obtain appropriate noise performance to meet the scaling requirements, like noise level [147] and cutoff frequency [148].

To explain the noise behavior, there exists two physical models explaining 1/f noise that predict a gate bias dependence of noise spectra. McWhorter model explains that 1/f is originally due to the carrier number fluctuation (\(\Delta n\)) in channel caused by trapping and detrapping in the dielectric. Recently, it was reported that it is more prominent in small devices [69]. Hooge model, on the other hand, proposes that noise is from mobility fluctuation (\(\Delta u\)) scattered by phonons [149]. As both cases were supported by experimental data, the source of 1/f noise still remains unresolved. K.K. Hung et al. proposed a unified model to fit experimental data without \textit{ad hoc} assumption on noise mechanism [71]. In planar structures, noise mechanisms of n and p channel are well
studied and reported [67, 72, 143, 150-155], where noise in nMOSFET is mostly due to carrier number fluctuation (Δn), while noise in pMOSFET is due to mobility fluctuation (Δu). In FinFET structure, on the other hand, random telegraph noise (RTN) is more widely studied instead of flicker noise (mechanisms are rarely reported) since RTN is more prominent in small area devices indicating Δn [155]. It is reported that Δn dominates in both n and p FinFETs, however Δu only presents in pFinFET [156, 157]. Since EOT is shrinking to sub nm region, it is necessary to study how oxide thickness relates to noise mechanism besides the dependence of device area.

In this work, noise mechanisms were studied on thick and thin FinFET structures to understand the underlying origin of the flicker noise. 1/f noise was measured at both linear and saturation region for n and p FinFET devices to understand how oxide thickness relates to noise mechanisms. FinFET devices with two different equivalent oxide thickness (EOT) were characterized (namely, thick FinFET and thin FinFET), and noise mechanisms were determined by fitting the existing noise models, the McWhorter model (carrier number fluctuation model) and the Hooge model (mobility fluctuation model) at different temperatures (-40 °C, 25 °C, 125 °C). The observations were further confirmed by evaluating the noise spectrum slopes. The short channel effect on the noise mechanism and the advantages of FinFET structure over planer structure were also discussed.
Table 7.1 Information of Devices Under Test

<table>
<thead>
<tr>
<th>Device Information</th>
<th>MGHK</th>
<th>SiON</th>
<th>Thick FinFET</th>
<th>Thin FinFET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel Length</td>
<td>short</td>
<td>short</td>
<td>long</td>
<td>short</td>
</tr>
<tr>
<td>EOT</td>
<td>thin</td>
<td>thin</td>
<td>thick</td>
<td>thin</td>
</tr>
</tbody>
</table>

7.2 Device Information and Measurement Method

The devices under test (DUT) were listed in the Table 7.1. Planar structures with different oxides, high-k with metal gates (MGHK) and silicon oxynitride (SiON) were evaluated to compare with FinFET structures (Table 7.1).

The noise measurements were conducted by the combination of BTA9812 (signal amplifier, filter control units), B1500 (DC bias set up and IV characterization), and HP35665A (spectrum analyzer). All the hardware was controlled by software Noisepro via GPIB connection. Eight dies of each device types were measured in order to obtain the statistical information.

The 1/f noise characteristic was measured in both linear region and saturation region. The noise level at 25 Hz ($S_{id}$) was sampled at least three different bias conditions by setting the drain current ($I_d$). $I_d$ was pre-selected from $I_d$-$V_g$ plot by choosing $V_g$ values close to $V_{th}$. $V_{th}$ was determined by square law method in saturation region. To determine noise model effectively (more data points) and obtain oxide trap volume density as a function of gate voltage, each device was measured for 1/f noise at different bias conditions by sweeping $V_g$ (50 mV interval) at both linear and saturation regions. The averaged data from eight devices were fitted well into data from sweep, which ensured the validation of data before identifying the appropriate noise model.
7.3 Results and Discussion

7.3.1 Bias Dependence of Flicker Noise

As described earlier, two physical $1/f$ noise models predict the gate bias dependence of noise spectra. In 3D FinFET structure, the $1/f$ noise has bias dependence as shown in Figure 7.1 and Figure 7.2. As gate bias increases, the $1/f$ noise spectrum increases with a minimum change of slope of $S_{id}$ versus frequency (Figure 7.1). Furthermore, normalized $S_{id} (S_{id}/I_d^2)$ is decreased as bias conditions were alternated (Figure 7.2). To explain this behavior the two competing models McWhorter model [69] and Hooge model [149] were evaluated here. Because of trapping and detrapping in the dielectric, McWhorter model suggests that the $1/f$ noise is dependent on carrier number fluctuation ($\Delta n$) in channel and is more prominent in small devices [66, 143]. On the other hand, because of carrier scattering by phonons, Hooge model proposed that the source of noise is from mobility fluctuation ($\Delta u$) [149]. It is possible that when gate bias increases channel mobility decreases leading to increase in $1/f$ noise referring to Hooge’s model. As both models are supported by experimental data, K.K. Hung et al. proposed a unified model for the source of $1/f$ noise to explain the experimental data [158]. It was observed that the bias dependent data ($S_{id}/I_d^2$) changes as overdrive voltage ($V_g-V_{th}$) changes. The noise mechanism is determined by plotting normalized noise ($S_{id}/I_d^2$) against overdrive voltage ($V_g-V_{th}$) in both linear and saturation region. The noise mechanism is attributed to $\Delta n$ model if slope value is -2, or $\Delta u$ if slope value is -1 [72].
7.3.2 Short and Long Channel Effect on Flicker Noise

As shown in Figure 7.2(a), there is a difference in normalized $S_{id}$ between linear region and saturation region for thin nFinFET. This is believed to be due to short channel effect [158]. The discrepancy is even more prominent in thin oxide FinFET at high current ($\log(I_d) = -4.5(A)$) regions as shown in Figure 7.3. It is known that, with highly doped source and drain with steep doping density gradients, the effective length is
approximately equal to the physical length between source and drain. Therefore, the discrepancy in the thick oxide FinFETs is mainly due to the short channel effect. However, for lightly doped drain (LDD) structures, the effective length can be larger than the source/drain spacing, because the channel can extend into the lightly-doped source and drain especially for high gate voltages in linear bias condition [159]. In this case, thin oxide FinFETs with short channel length shows this phenomenon, which causes an even larger noise difference between linear and saturation regions.

Figure 7.3 The normalized $S_{id}$ ($S_{id}/I_d^{\frac{2}{3}}WL$), is plotted for all four type FinFET at both linear and saturation regions. At relative large gate bias, ~$V' + 0.2$ V, thin oxide FinFET shows larger discrepancy than thick oxide due to both short channel and LDD effect.

7.3.3 Noise Model Extraction at Room Temperature

The noise mechanism was initially extracted in linear region for all four type devices at
room temperature since minimum series resistance effect [150] and low electrical field in channel was similar to the ideal model. The noise characteristics (Figure 7.4) shows the slope values of -1 ($\Delta u$) for thick pFinFET and for n/p thin FinFETs. Whereas, only thick nFinFET shows a slope of -2 ($\Delta n$). For thick FinFET (Figure 7.4(a)), noise mechanisms for n and p are different, this is consistent with the results from L. K. J. Vandamme et al. [151] where 1/f noise in nMOSFET was dominated by carrier number fluctuation, $\Delta n$, while mobility fluctuation, $\Delta u$ was dominant in pMOSFET. In Figure 7.5, the oxide trap density, $N_t$ is plotted as a function of overdrive voltage, calculated using Equation (3.14) in linear region for thick nFinFET since it represents number fluctuation, $\Delta n$. On the other hand, Hooge empirical parameter ($\alpha_H$) for thick pFinFET and thin p/n FinFETs are plotted against overdrive voltage using Equation (3.15). The values of $N_t$ and $\alpha_H$ are calculated at low overdrive voltage to exclude series resistance effect.

![Figure 7.4](image)

**Figure 7.4** Noise mechanisms were extracted at low gate bias in linear region by fitting normalized $S_{id}$ versus $V_g - V_{th}$. (a) n/p thick FinFET, (b) n/p thin FinFET.
Figure 7.5 Left axis shows oxide trap density ($N_t$) against overdrive voltage in linear region for thick nFinFET, right axis shows Hooge empirical parameter ($\alpha_H$) against overdrive voltage ($V_g-V_{th}$) in linear region for thick pFinFET and thin n/p FinFET.

7.3.4 Temperature Effect on Flicker Noise Model

When EOT is decreased, both n and p FinFET show $\Delta u$ mechanism (Figure 7.4(b)). This raises questions: What if oxide thickness decreased to few Å? Intuitively, if there is no oxide layer at all, the $\Delta u$ will be prominent. Thin oxide will have less overall volume of traps and larger tunneling current. Theoretically, $\Delta n$ predicts $S_{id}I_d^2 \propto t^2$ and $\Delta u$ predicts $S_{id}I_d^2 \propto t$, therefore, scaled oxide thickness has a larger impact on $\Delta n$ rather than $\Delta u$ mechanism. Is it true that thinner oxides have the noise mechanism linked to surface phenomenon at room temperature? Yes, F. Crupi et al. also reported an interfacial layer thickness dependent of noise mechanism [152]. The same method discussed above can be applied to the devices under other temperatures (-40 °C, 125 °C). The purpose is to observe whether it is possible to alter the noise mechanism at a different temperature. Several groups have studied the 1/f noise at different temperatures. The motivation to
characterize $1/f$ noise under different temperatures is to study the origin of $1/f$ noise. The McWhorter model predicts no temperature dependence of $1/f$ noise, since tunneling of carrier depends less on temperature (although temperature will change Fermi level position in the bandgap and trap-assist tunneling is dependent on temperature) [160]. On the other hand, Hooge model is explained by phonon scattering [149], therefore, depends more on temperature [161]. J. Chang et al. studied $1/f$ noise behavior on MOSFET from room temperature to 5 K [153]. They reported input referred noise of nMOSFET has no dependence on gate voltage and temperature, and concluded that it is McWhorter model. On the contrary, pMOSFET has both gate voltage and temperature dependence on the input referred noise. The $1/f$ noise decreases as temperature decreases from room temperature to 150 K, and noise increases when temperature is lowered further. At 20 K, the noise spectrum shows Lorentzian spectrum, which they considered as G-R noise caused by implant freeze out. Low temperature (150 K) Lorentzian spectrum is also observed on nFinFET device by W. Guo et al., who concluded nFinFET is McWhorter model and Lorentzian spectrum is caused by defects in the depletion region [156]. Lartigau et al. has similar results on their SOI nMOSFET [154]. Achour et al. reported noise mechanism of pFinFET changes from Hooge model to McWhorter model when temperature is lowered from room temperature to 10 K [162]. The above reviews concluded that flicker noise mechanism changes to $\Delta n$ rather than $\Delta u$ when phonon scattering is not significant at low temperatures. The observation of Lorentzian spectrum is an evidence of single defects in oxides [67]. In our case, as shown in Figure 7.6, noise mechanism became dominate from $\Delta u$ to $\Delta n$ (slope value from -1 to -2) when temperature decreased to -40 °C (Figure 7.6(b-d)), while thick nFinFET kept the $\Delta n$ (-2) dependence.
As discussed earlier, this is mainly due to reduction in phonon scattering and defect density in the oxide.

Figure 7.6 Normalized $S_{id} (S_{id}/I_d^2)$ is plotted as function of overdrive voltage ($|V_g-V_{th}|$) at linear bias condition under three different temperatures (-40 °C, 25 °C, 125 °C) for four types of devices (a) thick nFinFET, (b) thick pFinFET, (c) thin nFinFET, (d) thin pFinFET. The data is an average of eight devices.

The 1/f noise spectrum can increase or decrease when temperature is changed [153-156, 162, 163]. Ohguro et al. found the 1/f noise is increased when temperature is decreased for the bulk FinFET, which is explained by correlating noise to the surface electrical field, and their simulation shows electrical field is lower at higher temperature.
Lee et al. report similar behavior for sub 100 nm pMOSFET but not for nMOFET [155]. Our devices showed a similar behavior as reported by Ohguro (not shown here).

Study of noise mechanism by directly fitting the model is straightforward as discussed above. The slope of noise spectrum can also provide some insight regarding underlying physics of noise mechanism. If both noise mechanisms exist in a device, to some extent, one of them can overwhelm the other (discuss later). During the measurement, when the mechanism seems to transformed, the slope of noise spectrum will change statistically. Either reducing the temperature or increasing horizontal electrical field (measured in saturation region), \( \Delta n \) phenomenon can be more prevailing than \( \Delta u \) as shown in Figure 7.6 and Figure 7.7, respectively. The former one is suppressing phonon scattering (smaller \( \alpha_H \)) [161], while the later one is enhancing charge traps by generating hot carriers in the channel (carriers gain energy to overcome barrier) [164]. In Figure 7.7, thick pFinFET shows a change of noise mechanism at saturation region (from \( \Delta u \) to \( \Delta n \)), while other devices keep same noise mechanism (not shown here). Although there is a mathematical model to describe the situation when both \( \Delta n \sim \Delta u \) by surface scattering coefficient [158], the time constant distributions of two mechanisms are expected to be different intrinsically [149, 158]. Therefore, it is expected that there is a change of flicker noise spectrum slope \( \gamma \) (\( S_{id} = \frac{1}{f_{\gamma}} \)).

In Figure 7.8, it is found that the slope of spectra (\( \gamma \)) tend to be same at linear and saturation bias conditions when temperature is at -40 °C (black symbols). The reason is \( \Delta n \) mechanism dominates \( \Delta u \) mechanism at low temperature. When \( \gamma \) of n/p FinFET are
similar under same processes (thick oxide and thin oxide) at -40 °C they have the same oxide defect profile. At higher temperature, it is difficult to explain the discrepancy of $\gamma$, however the overall trend plot remains similar under the same process conditions.

Figure 7.7 Normalized $S_{id}$ ($S_{id}/I_d^2$) is plotted as function of overdrive voltage ($V_g-V_{th}$) at both linear and saturation bias conditions for thick pFinFET.
Figure 7.8 Slope of flicker noise spectra, $\gamma (S_{id} = \frac{1}{f^\gamma})$ plotted for four devices at different temperatures (-40 °C, 25 °C, 125 °C) and bias conditions (linear and saturation). Slope values are average value obtained from eight devices at both the bias conditions.

7.3.5 Planar Structure VS FinFET

In the last part of this work, $1/f$ noise characteristics of FinFET is compared with planar structure (MGHK and SiON) after normalization with respect to different device areas and drain current levels at linear bias condition (Figure 7.9(a) trend plot). The 2D structure (MGHK and SION) show comparable $1/f$ noise level at 25Hz under different $I_d$ conditions. The $1/f$ noise level is quite comparable at low drain current conditions for 2D and 3D structure. However, the noise level of 3D structure (thin FinFET) is reduced significantly (about one decade) when the drain current is increased (nFET is similar to pFET). Additionally, variation of 3D structure is smaller than 2D structure (Figure 7.9(b)). The pFET have similar $1/f$ noise behavior as nFET (not shown here). Small variation was discussed by Fan et al. by TCAD simulation [165]. They concluded that
because of the stronger correlation between the charged trap and subthreshold current conduction the planar bulk devices with RDF (random doping fluctuation) exhibit broader RTN dispersion than the FinFET device. The variation in FinFET is mainly from line edge roughness (LER) and work function variation (WFV) [165]. Experimentally, as reported by Ohguro et al., the 3D structure shows a low 1/f noise level as well as low 1/f noise variation compared to 2D structure [148]. The possible reason is, because the bulk silicon of the planar devices was doped a relatively larger variation was observed compared to intrinsic silicon bulk of FinFET devices. The smaller normalized 1/f noise of the 3D structure may be attributed to the intrinsic silicon bulk (less carriers scatter with dopants in the bulk), smaller surface electrical field (larger depletion region) compared to doped planar structure and modified short channel length due to lightly-doped drain (LDD) at higher gate voltages [159].

![Figure 7.9](image-url)

**Figure 7.9** (a) Normalized $S_{id}$ ($S_{id}/I_d^{2*WL}$) is plotted against different structure at different current bias condition (linear region) for nFET; (b) Normalized $S_{id}$ ($S_{id}/I_d^{2*WL}$) is plotted against drain current ($I_d$) at linear region for nFET.
7.4 Conclusion

In this work, thick and thin oxide n/p FinFETs were compared to understand the impact of EOT on the $1/f$ noise mechanisms. pFinFET is considered as mobility fluctuation model regardless of its oxide thickness. However, the noise model of nFinFET deviates from $\Delta n$ model to $\Delta u$ model when EOT is decreased. The oxide volume trap density and Hooge empirical parameter were extracted based on existing model. For thin oxide, noise dispersion between linear and saturation region at high current bias condition is discovered. The short channel length and LDD structure are found to be very sensitive to the gate voltages in the linear region. Flicker noise behavior at three different temperature (-40 °C, 25 °C, 125 °C) was studied to enlighten the physical mechanism of flicker noise. Flicker noise mechanism deviated from $\Delta u$ to $\Delta n$ and noise spectrum slope measured at linear and saturation region tend to converge when temperature was decreased. Finally, 2D and 3D structures were compared to show FinFET has a more uniform $S_{id}$ distribution and a reduced normalized noise level when transistor is biased at high drain current in linear region.
CHAPTER 8
SUMMARY AND FUTURE WORK

8.1 Summary
This dissertation investigates the oxide/substrate interface quality and the dielectric quality of metal oxide semiconductor (MOS) gate stack structures as they are critical to future CMOS technology. Since high mobility substrates like Ge have attracted increasing attention to enhance the devices performance we have investigated the high-k dielectrics on both silicon and germanium substrates for EOT scaling and interface performance. In this study, we have characterized several MOS structures, prepared by advanced atomic layer deposition (ALD) process and with pre and post treatment by plasma generated by slot plane antenna (SPA).

Different electrical characterization methods like CV, GV, DLTS, IV, TDDB, and noise measurement were used throughout this work. They have been briefly introduced in Chapter 3. Here, a comparison is presented for conductance method, low frequency CV, DLTS, and flicker noise in terms of their capability and difficulty of implementation. It is listed in Table 8.1.

As can be expected, DLTS and flicker noise measurement are more difficult to implement as compared to the other two methods. Conductance and low frequency CV methods are part of the same model for device characterization, which can be implemented simultaneously for calibration of measurement set up. However, they are not able to provide true energy levels of the trap and it is usually fulfilled by simulation softwares like CVC 7.0. The detection of trap energy level is usually limited in depletion
region, and can be extended to inversion region if substrate is not doped (no minority carrier response). To extract the energy level directly DLTS measurements are required. Flicker noise is a reliability issue became serious issues when device is scaled. The level of noise sometimes is more important than their defects information.

**Table 8.1** Comparison of Characterization Method for $D_{it}$

<table>
<thead>
<tr>
<th>Methods</th>
<th>Difficult of implementation</th>
<th>Measuring time</th>
<th>Accuracy of $D_{it}$</th>
<th>Resolution of $D_{it}$</th>
<th>Energy level?</th>
<th>Detection range in bandgap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conductance</td>
<td>Easy</td>
<td>Minutes</td>
<td>High</td>
<td>High</td>
<td>No</td>
<td>Depletion region</td>
</tr>
<tr>
<td>Low frequency CV</td>
<td>Easy</td>
<td>Minutes</td>
<td>High</td>
<td>Medium</td>
<td>No</td>
<td>Depletion region</td>
</tr>
<tr>
<td>DLTS</td>
<td>Hard</td>
<td>Hours</td>
<td>Medium</td>
<td>Low</td>
<td>Yes</td>
<td>Accumulation &amp; depletion region</td>
</tr>
<tr>
<td>Flicker noise</td>
<td>Medium</td>
<td>Hours</td>
<td>Medium</td>
<td>Low</td>
<td>No</td>
<td>Inversion region</td>
</tr>
</tbody>
</table>

In Chapter 4, it was observed that the $D_{it}$ value is a function of the dielectric constant in Al incorporated HfO$_2$. The $D_{it}$ is also directly related to the structure of dielectric. This is in contrast to the SiO$_2$/Si where SiO$_2$ is always amorphous under the conventional fabrication process. Even though the interfacial layer of our samples is chemically prepared, it does not act as a perfect insulator as the thickness is close to the atomic level. The carriers in the Si substrate can respond to the HfAlO, and the trapping and de-trapping of carrier process is affected by the HfAlO structure. In addition, small percentage of Al is mainly responsible to change the structure of HfO$_2$ from amorphous to crystalline. The amorphous structure has lowest $D_{it}$ ($2.76 \times 10^{11}$ eV$^{-1}$cm$^{-2}$) whereas t-HfO$_2$ has the highest $D_{it}$ ($1.27 \times 10^{12}$ eV$^{-1}$cm$^{-2}$). The $D_{it}$ results of other structures are within the range from lowest to highest value. It is concluded that increasing the
dielectric constant of high-k gate dielectrics comes with a price, higher $D_{it}$ that decreases the channel mobility.

In Chapter 5, Ge/ALD 1nm-Al$_2$O$_3$/ALD 3.5nm-ZrO$_2$/ALD TiN MOS capacitors with three different interface treatments were demonstrated. COR&SPAO$_x$ samples show superior interfacial characteristics at room temperature. After evaluating several parameters like EOT, flatband voltage, bulk doping, and interface defects density, COR&SPAO$_x$ (dry) has better interface quality than Chemox (wet) and COR&O$_3$ (dry) processed sample. However they all have higher $D_{it}$ values in the order of $10^{13}$ cm$^{-2}$eV$^{-1}$, which causes Fermi level more or less pinned, which is confirmed by low temperature measurements. Dry processed sample (COR&SPAO$_x$ and COR&O$_3$) has more negative $V_{th}$ shift due to existing border trap characterized using DLTS method. The levels of leakage current of three samples confirm EOT values. Therefore, COR&SPAO$_x$ (dry) has lowest trap assisted tunneling effect leading to lowest leakage current.

In Chapter 6, the interfacial layer quality ($D_{it}$), and quality of high-k layer in terms of leakage current density and TDDB under different SPAO annealing were studied and discussed. It is found that SPAO forms GeO$_2$ at the interface, however subsequent ALD process decomposed the unstable GeO$_2$ into GeO$_x$. GeO$_x$ is more reliable than GeO$_2$ in terms of TDDB results with no significant different in $D_{it}$ values. SPAO removed high-k layer defects as confirmed by the temperature I-V measurements. Different SPAO strategy also formed different interfacial layer thickness. Oxygen has to diffuse through high-k layer when SPAO is performed after high-k deposition. Chapter 6 somewhat also revealed improvement and understanding of post annealing and ALD process, which is very critical to future implementation of sub-nm EOT on high-k layers regarding that
they impact the interfacial layer quality.

As FinFETs are being used in advanced technology nodes, Chapter 7 compared the different flicker noise behavior of FinFETs with that of planar structure in terms of noise level. FinFET shows lower noise level at high gate bias and in linear bias. Additionally, the impact of EOT on the noise mechanisms was studied in detail at different temperature. It was observed the flicker noise mechanism in nMOS is modified from McWhorter model to Hooge model as EOT is reduced while that in pMOS, Hooge model is relevant. When temperature decreases, flicker noise mechanism shows more dependence on McWhorter model, which is a consequence of reduced phonon scattering in the system.

### 8.2 Future Work

For the future work, the characterization methods will be correlated further. The following issues have to be resolved to further understand the high-k/substrate interface and to implement new processes in CMOS technology.

1. EOT high-k/Si can be scaled below 0.5 nm by preparing thinner interfacial layer, however $D_{it}$ has to be monitored.

2. Interface treatment of high-k is necessary to be studied under different ambience comprehensively to reduce $D_{it}$.

3. High-k/Ge device need more study on their reliability like flat band voltage shift under different stress polarities to confirm it is interfacial layer degradation.

4. Gate dielectrics of high-k/Ge can be studied for different gate dielectrics.
REFERENCES


