CHARACTERIZATION AND MODELING OF HOT-CARRIER DEGRADATION IN
SUB-MICRON NMOSFETS

By

Manish Prabhakar Pagey

Thesis
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE
in
Electrical Engineering

August, 2002

Nashville, Tennessee

Approved:
Sherra E. Kerns, Chair
Bharat L. Bhuva
ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Sherra E. Kerns, Dr. Bharat L. Bhuva, and Dr. Lloyd W. Massengill for their guidance, support and encouragement over the period of this work. I extend my gratitude to Eric Snyder, William Miller, and Reid Bennett for providing me with the opportunity to work at Sandia National Laboratories (Albuquerque) where all the experiments mentioned in this work were performed. I also thank Dr. Bui (AMD) for providing the test-structures for performing hot-carrier stressing experiments and the Semiconductor Research Corporation for funding this research. I am extremely thankful to Jean George for her help throughout this work.

I would also like to thank all the student members of the Space Electronics Research Group, past and present, for their help, criticisms, suggestions, and friendship which makes every day in the office a pleasant one.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Hot-Carrier Injection Phenomenon</td>
<td>2</td>
</tr>
<tr>
<td>Thesis Outline</td>
<td>4</td>
</tr>
<tr>
<td>II CHARACTERIZATION TECHNIQUES</td>
<td>6</td>
</tr>
<tr>
<td>Drain Current Characteristics</td>
<td>7</td>
</tr>
<tr>
<td>Threshold Voltage</td>
<td>9</td>
</tr>
<tr>
<td>Channel Transconductance</td>
<td>13</td>
</tr>
<tr>
<td>Effect of Hot-Carrier Injection on Drain Current Characteristics</td>
<td>13</td>
</tr>
<tr>
<td>Charge Pumping Measurement</td>
<td>18</td>
</tr>
<tr>
<td>Qualitative Analysis</td>
<td>18</td>
</tr>
<tr>
<td>Geometric Component of Charge Pumping Current</td>
<td>20</td>
</tr>
<tr>
<td>Surface State Component</td>
<td>22</td>
</tr>
<tr>
<td>Energy Distribution of Surface States</td>
<td>23</td>
</tr>
<tr>
<td>Implementation</td>
<td>26</td>
</tr>
<tr>
<td>III HOT-CARRIER DEGRADATION IN N-CHANNEL MOSFETS</td>
<td>27</td>
</tr>
<tr>
<td>Introduction</td>
<td>27</td>
</tr>
<tr>
<td>Carrier Injection and Gate Currents</td>
<td>28</td>
</tr>
<tr>
<td>Substrate Currents</td>
<td>30</td>
</tr>
<tr>
<td>Lucky-Electron Model</td>
<td>30</td>
</tr>
<tr>
<td>DC Stressing Experiments</td>
<td>32</td>
</tr>
<tr>
<td>Device Degradation Models</td>
<td>34</td>
</tr>
<tr>
<td>The Takeda and Suzuki Model</td>
<td>35</td>
</tr>
<tr>
<td>The Hu Model</td>
<td>35</td>
</tr>
<tr>
<td>The Woltjer Model</td>
<td>36</td>
</tr>
<tr>
<td>The Mistry et al. Approach</td>
<td>37</td>
</tr>
<tr>
<td>Model Comparison and Coupled Empirical Approach</td>
<td>38</td>
</tr>
<tr>
<td>Dynamic Degradation Considerations</td>
<td>46</td>
</tr>
<tr>
<td>IV CONCLUSIONS AND FUTURE WORK</td>
<td>50</td>
</tr>
</tbody>
</table>
Appendices

A ELECTRON SCATTERING IN IMAGE FORCE POTENTIAL WELL . . . . 51
REFERENCES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cross-section of an n-channel MOSFET illustrating the generation and injection</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>of hot-carriers into the gate-oxide</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Energy band diagram showing the barrier heights for injection of electron and</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>holes from Si to SiO&lt;sub&gt;2&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The device terminals used to measure “normal” and “reverse” mode parameters</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>relative to the terminals used during hot-carrier stressing experiments.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The extraction of linear extrapolated threshold voltage from measured I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;GS&lt;/sub&gt; characteristics for a long channel device.</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>The extraction of saturation extrapolated threshold voltage from I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;GS&lt;/sub&gt; measurements performed in saturation region for a long channel MOSFET.</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>The extraction of constant current threshold voltage from measured I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;GS&lt;/sub&gt; characteristics.</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>The I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;GS&lt;/sub&gt; characteristics of sub-micron MOSFET. The extrapolated threshold voltage is usually extracted by extrapolating from the point of maximum slope.</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>The normal and reverse mode I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;DS&lt;/sub&gt; characteristics measured before and after a typical hot-carrier stress in n-channel MOSFETs.</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>The normal and reverse mode I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;GS&lt;/sub&gt; characteristics measured before and after a typical hot-carrier stress in n-channel MOSFETs.</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>The damaged region is “masked” by the pinch-off region during I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;DS&lt;/sub&gt; characteristics in normal-mode while the whole region is “visible” during reverse-mode I&lt;sub&gt;DS&lt;/sub&gt;−V&lt;sub&gt;DS&lt;/sub&gt; measurements.</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>Basic MOS Charge Pumping Experiment</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>Typical Charge Pumping Characteristics</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>Annular MOS Capacitor</td>
<td>21</td>
</tr>
<tr>
<td>14</td>
<td>Typical charge pumping characteristics using constant amplitude pulse and</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>varying the base level.</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Gate Capacitance as a Function of Source/Drain Bias</td>
<td>25</td>
</tr>
</tbody>
</table>
The dependence of carrier injection and gate current on the gate bias in a conventional n-channel MOSFET. .................................................. 28

The substrate current vs. gate bias characteristics show a bell shaped curve. The initial increase in substrate current is due to increase in the carrier supply while the decrease in substrate current at high gate biases is due to a decrease in the lateral electric field. .................................................. 31

The dependence of injection and degradation modes on the gate bias in n-channel MOSFETs. .................................................. 33

The device degradation as a function of time under $I_{B,max}$ bias condition. The values obtained from the parameterized models are also plotted along with the measured data. .................................................. 39

The device degradation as a function of time under $E_{inj}$ bias condition. The parameterized conventional models are unable to correctly predict the device degradation under this condition. .................................................. 41

The device degradation as a function of time under $H_{inj}$ bias condition. The parameterized conventional models are unable to correctly predict the device degradation under this condition. .................................................. 42

Device degradation as a function of time for devices stressed using bias sequence 1. The combined effect of all three dominant degradation modes needs to be included at each bias condition to correctly predict the degradation. .................................................. 43

Device degradation for stressing under bias sequence 2. Conventional models deviate from the observed values while the coupled application of three models results in an improved prediction. .................................................. 44

Dynamic degradation shows a stronger time dependence than static degradation. .................................................. 48

Potential as a function of distance inside a dielectric close to the conductor/dielectric interface. .................................................. 52

Electron trajectory in the dielectric after emission at an angle $\theta$ to the normal to the interface. .................................................. 53
LIST OF TABLES

Table | Page
--- | ---
1. Mechanisms associated with performance degradation due to high electric fields in MOSFETs. | 2

vii
CHAPTER I

INTRODUCTION

The dimensions of metal-oxide-silicon field effect transistors (MOSFETs) have been decreasing due to the continuous demand for higher packing densities and faster circuit speeds. A logical choice for device scaling scheme, based on keeping the internal electric fields constant[1, 2], had to be abandoned due to several practical reasons. Some of the disadvantages associated with constant-field scaling include:

- Loss of compatibility with TTL power-supply voltage,
- Decrease in noise margins because the threshold voltage and the subthreshold slope do not scale, and
- Decrease in operating speeds in sub-micron devices due to the non-scaling of parasitic capacitances.

The use of alternate scaling schemes has produced increased electric fields in these devices during circuit operation. Within some ranges, these high electric fields result in increased carrier velocities and hence a higher operating speed. However, as the electric field increases, the carrier velocity saturates at a certain critical field, \( E_c \)[3]. Electric fields in excess of this value have no beneficial effects on the device performance. On the contrary, under such high electric fields, the mobile carriers in the silicon substrate can attain relatively high energies and result in incorrect circuit operation through a variety of mechanisms[2](Table 1). While most of these problems can be minimized, even eliminated in some cases, through appropriate circuit and device design, the injection of energetic carriers into the gate oxide and subsequent parameter shift through carrier trapping and interface trap generation poses one of the most significant long-term reliability concerns in extensively scaled MOSFETs.

This thesis presents a study of the degradation of MOSFET parameters, which affect circuit operation, due to processes initiated by injection of high energy carriers into the gate oxide, the characterization techniques used to measure and attribute these parameter shifts.
Table 1: Mechanisms associated with performance degradation due to high electric fields in MOSFETs.

<table>
<thead>
<tr>
<th>Carrier Energy (E)</th>
<th>Mechanism</th>
<th>Effect on Device Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E &gt; 1.12eV$</td>
<td>Light emission</td>
<td>Leakage currents</td>
</tr>
<tr>
<td>$1.3eV &lt; E &lt; 1.8eV$</td>
<td>Impact ionization</td>
<td>Snap-back, latch-up, leakage.</td>
</tr>
<tr>
<td>$E &gt; 3.2eV$</td>
<td>Hot-electron injection</td>
<td>Parameter shift due to carrier trapping and interface trap generation</td>
</tr>
<tr>
<td>$E &gt; 4.8eV$</td>
<td>Hot-hole injection</td>
<td></td>
</tr>
</tbody>
</table>

to the underlying processes, and the use of empirical, semi-empirical, and physical models to predict the time dependence of the parameter degradation during circuit operation.

**Hot-Carrier Injection Phenomenon**

A brief overview of the hot-carrier injection phenomenon and the resulting device degradation will be provided in this section. The cross-section of a typical n-channel MOSFET operating in saturation is shown in Fig. 1. The large voltage drop across the pinch-off region results in a high lateral electric field close to the drain region. The carriers traversing this high field region reach energies which are considerably higher than the equilibrium thermal energy in the semiconductor lattice. These high energy carriers are called *hot-carriers*.

Hot-carriers with energies above the impact-ionization threshold ($\approx 1.6eV$)[2] can generate electron-hole pairs in this region through impact-ionization. Some of these carriers, with energies large enough to overcome the potential barrier between Si and SiO$_2$ and their momentum directed towards the Si-SiO$_2$ interface, can get injected into the gate oxide[4]. As shown in Fig. 2, the energy barrier for injection of electrons ($\approx 3.1eV$) is considerably smaller than that for holes ($\approx 4.8eV$) making hole-injection a less probable event as compared to electron-injection. The exact barrier at any given point along the channel is affected by the transverse electric field at that point due to the Schottky effect[3].

A large proportion of the injected electrons reach the gate terminal and contribute to the
Figure 1: Cross-section of an n-channel MOSFET illustrating the generation and injection of hot-carriers into the gate-oxide

Figure 2: Energy band diagram showing the barrier heights for injection of electron and holes from Si to SiO₂
gate current. However, some of the injected electrons can also get trapped at certain defects present in the gate oxide[5, 6]. Similarly, the injected holes can reach the gate and contribute to the gate current. However, as the hole mobility in SiO$_2$ is considerably smaller than that of electrons[7], holes have a higher probability of getting trapped. In addition to getting trapped in the gate-oxide, the injected carriers can also result in increase in the density of interface traps present at the Si-SiO$_2$ interface[8, 9, 10]. The presence of charge in the gate-oxide and at the Si-SiO$_2$ interface of a MOSFET results in modulation of the surface potential and carrier mobilities at the surface of the semiconductor. These phenomena alter the device current characteristics which can significantly reduce the operating lifetime of these devices[11].

As mentioned earlier, the injection of carriers into the gate-oxide is a significant concern for assuring the long-term reliability of modern digital CMOS circuits. As these degradation processes are relatively slow during circuit operation, in order to characterize the hot-carrier response of devices within short times they are often subjected to much larger biases than those present in real circuits. The presence of higher biases results in an acceleration of the degradation processes and hence such experiments are called accelerated stressing experiments. The extrapolation of the results obtained from the accelerated stressing experiments to real-life circuit operation is a subject of other chapters later in this thesis.

**Thesis Outline**

The various characterization techniques which are used to monitor the device degradation during accelerated stressing experiments and attribute it to the underlying physical mechanisms will be presented in Chapter II. Chapters III presents the issues specific to hot-carrier injection in n-channel MOSFETs. As a part of this work, a set of stressing experiments are suggested to study the various aspects of device degradation in n-channel MOSFETs comprehensively. In order to predict the impact of hot-carrier induced device degradation on the circuit operation as well as to optimize the parameter shifts due to hot-carrier injection, we need to accurately model the time dependence of these parameter shifts. Chapter III also presents an overview of some of the popular models and discusses their applicability to deep-sub-micron technologies. The results of some of the experiments during which the
suggested experiments were applied to monitor and model the device response to hot-carrier injection in a commercial technology are also presented. Finally, Chapter IV summarizes the significant results of this work and suggests the future work required to advance the understanding of the physical mechanisms involved in hot-carrier degradation processes.
CHAPTER II

CHARACTERIZATION TECHNIQUES

The injection of hot-carriers into the gate-oxide of MOSFETs triggers carrier trapping and interface trap generation processes. The presence of interfacial and bulk charge in the gate-oxide affects the current characteristics of these devices. The alteration of the current characteristics effectively results in variations in some of the parameters extracted from them such as the threshold voltage, subthreshold slope and the transconductance\([11, 12, 13]\). These parameter shifts can be used as measures of the degradation as well as a key to understand the underlying physical mechanisms. The drain current characteristics (\(I_{DS} - V_{GS}\) and \(I_{DS} - V_{DS}\)) can be utilized to provide accurate information about the degradation processes when the hot-carriers are injected uniformly along the channel of the device, such as during substrate hot-carrier injection experiments\([14, 15]\). However, during operation in a digital CMOS circuit, channel hot-carriers are injected into the gate-oxide in a localized region close to the drain. The interpretation of the parameter variations during non-uniform channel hot-carrier injection based on the results obtained from uniform injection experiments can be highly inaccurate. For example,

- The variations in the \(I_{DS} - V_{GS}\) characteristics measured in saturation can be used to obtain the contributions due to interface traps and fixed charge in the oxide using techniques such as the midgap method\([16]\). These techniques assume that the change in subthreshold slope is entirely due to interface traps while a parallel shift in the subthreshold characteristics is entirely due to fixed charge in the oxide. However, during channel hot-carrier injection, a localized fixed charge in the oxide can also result in a change in the subthreshold slope rendering the midgap technique useless for hot-carrier stressing experiments.

- Under uniform injection conditions, it can be assumed that the threshold voltage at each point along the channel of the device shifts by the same amount. The threshold voltage of the complete device is equivalent to that of any point along the channel under
this condition. However, during channel hot-carrier injection, the threshold voltage at each point in the region of injection can vary at a different rate. Thus, the definition of a single threshold voltage for the complete MOSFET loses its physical meaning[2].

At the same time, it is sometimes possible to explain the observed degradation under non-uniform carrier injection by different combinations and spatial distributions of interface trap density and fixed oxide charge[17]. Due to the limitations in the correct interpretation of the variations in drain-current characteristics, certain other characterization techniques, such as charge pumping and substrate current characteristics, have been used in the recent literature. Some of the most commonly used device characterization techniques which are used to monitor and understand the device degradation under channel hot-carrier injection will be described in this chapter.

**Drain Current Characteristics**

The drain current characteristics as a function of the drain bias as well as the gate bias have been used extensively in literature to extract parameters which can be used as degradation monitors. The commonly used parameters include $V_T$, $I_{DS}$, $G_m$, and $S$. The physical meaning of each of these parameters can be defined on the basis of a simplified theory of operation of MOSFETs[3]. However, while performing experiments, the drain current characteristics are obtained in terms of two-dimensional arrays of numbers. The above parameters need to be extracted numerically from these data and the extracted values may not directly correlate to the physical definitions of these parameters. The localized non-uniform nature of channel hot-carrier induced degradation further separates the physical meaning of these parameters from the values extracted numerically from the measured data. In certain cases, the “same” parameter has been extracted by different techniques in literature and may yield completely different information.

The different ways of experimentally measuring the drain current characteristics adds extra complexity to the interpretation process. Each of the above parameters, for example, can be measured from drain current characteristics obtained in either linear or saturation regions of operation. Similarly, the drain current characteristics in each of these cases can be
measured either in “normal” mode or “reverse” mode. In the “normal” or “forward” mode, the drain and source terminals of the device are the same as those used during the stressing experiments while they are interchanged in the “reverse” mode (Fig. 3). In view of these facts, it is advisable to take extreme care in interpreting the results of hot-carrier stressing experiments on the basis of parameters extracted from drain current characteristics. *It is usually essential to extract the same information about the physical mechanisms based on two independent characterization techniques to obtain a consistent understanding of the phenomena involved.* The two most important parameters which are extracted numerically from measured data are the threshold voltage and the channel transconductance. The definitions of these parameters and the techniques used to extract them numerically from measured data will be presented in the next two sub-sections.

![Diagram of device terminals](image)

**Figure 3:** The device terminals used to measure “normal” and “reverse” mode parameters relative to the terminals used during hot-carrier stressing experiments.
Threshold Voltage

The classical definition of the MOSFET threshold voltage is based on the one-dimensional analysis of a MOS capacitor[3]. According to this analysis the threshold voltage is defined as the gate bias which results in a surface potential, \( \psi_s \), at the SiO\(_2\) interface which is equal to \( 2\psi_B \), where \( \psi_B \) is the potential difference between the bulk Fermi-level, \( E_{F,Bulk} \), and the intrinsic Fermi-level, \( E_i \).

The threshold voltage is extracted experimentally from the \( I_{DS} - V_{GS} \) characteristics. In the linear mode of operation, the \( I_{DS} - V_{GS} \) characteristics for long channel devices can be approximated using :

\[
I_{DS} = \frac{W}{L} \mu C_{ox} (V_{GS} - V_T)V_{DS}
\]

(1)

The threshold voltage can be extracted from the measured \( I_{DS} - V_{GS} \) characteristics by extrapolating the curve in the high \( V_{GS} \) region to \( I_{DS} = 0 \). The intersection of this extrapolated curve with the \( V_{GS} \)-axis gives the threshold voltage (Fig. 4). The threshold voltage obtained using this technique is called the \textit{linear extrapolated threshold voltage}, \( LVT_{ext} \). A similar extrapolation can be performed on \( I_{DS} - V_{GS} \) measurements in saturation using the approximate relation :

\[
\sqrt{I_{DS}} = \sqrt{\frac{W}{2L}} \mu C_{ox} (V_{GS} - V_T)
\]

(2)

Once again, the extrapolation of the \( \sqrt{I_{DS}} - V_{GS} \) curve to \( I_{DS} = 0 \) gives the \textit{saturation extrapolated threshold voltage}, \( SVT_{ext} \) (Fig. 5).

The threshold voltage of a MOSFET represents the gate bias at which the device turns “ON”. In other words, at any given drain bias, the drain current of an ideal device will be zero at gate biases below the threshold voltage and increase with the gate bias when it goes above the threshold voltage. As the threshold voltage defines the amount of drain current in a MOSFET, it is common to define the threshold voltage as the gate bias which results in a certain amount of drain current. The threshold voltage extracted using this definition is called the constant current threshold voltage, \( V_{T,ci} \) (Fig. 6). In the experimental results presented in the rest of this thesis, for example, \( V_{T,ci} \) is defined as the gate bias for which \( I_{DS} = (W/L) \times 1 \mu A \).
Figure 4: The extraction of linear extrapolated threshold voltage from measured $I_{DS} - V_{GS}$ characteristics for a long channel device.
Figure 5: The extraction of saturation extrapolated threshold voltage from $I_{DS}$–$V_{GS}$ measurements performed in saturation region for a long channel MOSFET.
Figure 6: The extraction of constant current threshold voltage from measured $I_{DS}-V_{GS}$ characteristics.
While Eqs. 1 and 2 work well for long channel devices, short channel effects tend to deviate the characteristics of modern devices from these approximate equations. The linear $I_{DS} - V_{GS}$ characteristics of a 0.5µm n-channel MOSFET are shown in Fig. 7. As can be seen from this figure, for large values of the gate bias, the drain current does not increase linearly with the gate bias. The extrapolated threshold voltage is usually extracted from such curves using the same approach as before assuming that Eq. 1 holds at the point of maximum slope along the curve (Fig. 7). A more precise method to extract the threshold voltage would be to fit an accurate model for the device characteristics of sub-micron devices, such as the BSIM3 model[18], to the measured characteristics. However, this approach is considered impractical while analyzing stressing experiments performed on a large volume of devices due to the high computational complexity of the curve fitting process. Obviously, no such problem exists in extracting the constant-current threshold voltage.

**Channel Transconductance**

The channel transconductance is defined as the rate of change of the drain current as a function of the gate bias at a given drain bias[3]:

$$G_m(V_{DS}) = \left[ \frac{\partial I_{DS}}{\partial V_{GS}} \right]_{V_{DS}=\text{const}} = \frac{W}{L} \mu C_{ox} V_{DS}$$  \hspace{1cm} (3)

As seen from Fig. 4, the transconductance can be easily extracted for long channel devices by taking the numerical derivative of the drain current with respect to the gate bias using the measured $I_{DS} - V_{GS}$ characteristics. In short channel devices, however, the derivative does not have a constant value and shows a non-monotonic nature with $V_{GS}$ (Fig. 7). The channel transconductance in short channel devices is usually taken as the maximum value of this derivative and sometimes referred to as the maximum transconductance for clarity.

**Effect of Hot-Carrier Injection on Drain Current Characteristics**

The typical $I_{DS} - V_{DS}$ and $I_{DS} - V_{GS}$ characteristics of an n-channel MOSFET before and after hot-carrier stressing experiment are shown in Figs. 8 and 9 respectively[2]. The “normal” and “reverse” characteristics after the stressing experiment have been plotted here.
Figure 7: The $I_{DS}$–$V_{GS}$ characteristics of sub-micron MOSFET. The extrapolated threshold voltage is usually extracted by extrapolating from the point of maximum slope.
Figure 8: The normal and reverse mode $I_{DS} - V_{DS}$ characteristics measured before and after a typical hot-carrier stress in n-channel MOSFETs.
Figure 9: The normal and reverse mode $I_{DS}-V_{GS}$ characteristics measured before and after a typical hot-carrier stress in n-channel MOSFETs.
The post-stress $I_{DS} - V_{DS}$ characteristics in the normal-mode join the pre-stress characteristics as the device goes into saturation. This can be explained by the fact that the device damage is localized near the drain region of the device. As the pinch-off region extends over the damaged region in the normal-mode $I_{DS} - V_{DS}$ characteristics, the damaged region stops affecting the device characteristics. However, in the reverse-mode of operation, the damage affects the device characteristics for all values of the “drain” bias (Fig. 10). A similar effect is observed in $I_{DS} - V_{GS}$ characteristics performed in saturation. The constant current threshold voltage extracted from reverse mode $I_{DS} - V_{GS}$ characteristics in saturation region is often used as a monitor of the device degradation as the whole of the damaged region contributes to the threshold voltage shift in this case.

"Normal" mode

"Reverse" mode

Figure 10: The damaged region is “masked” by the pinch-off region during $I_{DS} - V_{DS}$ characteristics in normal-mode while the whole region is “visible” during reverse-mode $I_{DS} - V_{DS}$ measurements.
Charge Pumping Measurement

The charge pumping technique is one of the most sensitive techniques for measuring the density of interface traps in MOS transistors. The basic charge pumping experiment will be described in this section. The technique has been a subject of many theoretical analysis and simulation studies. These studies have resulted in the development of numerous variations of the basic technique in order to obtain detailed information about the energy and spatial distribution of interface traps generated during hot-carrier stressing experiments\[19, 20, 21, 22\]. One of these variations has been used in the stressing experiments described in Chapter III and will be outlined at the end of this section.

The experimental setup for the basic charge-pumping experiment \[19\] is shown in Fig [11]. The substrate current is measured using a DC ammeter after being smoothed by a capacitor. The ammeter measures the negative reverse leakage current of the source-substrate and the drain-substrate junctions when no pulse is applied to the gate. On applying a gate pulse, the substrate current reverses sign indicating an effective charge flow from the source/drain to the substrate inspite of the reverse bias. This current is called the charge-pumping current, $I_{CP}$. The pulse base level is kept at a constant value in accumulation while the pulse amplitude is varied. It is also observed, by plotting the charge pumping current as a function of the pulse amplitude, that (Fig [12]):

- $I_{CP}$ is proportional to the frequency of the gate pulse for a rectangular pulse but drops at a rate greater than linear for a triangular pulse waveform,
- the current saturates at a gate pulse amplitude of $V_T + V_R$, where $V_T$ is the threshold voltage of the transistor,
- the charge pumping current scales proportional to the gate-area, and
- the magnitude of the saturating current decreases as the reverse bias, $V_R$, is increased.

**Qualitative Analysis**

The charge pumping phenomenon can be explained qualitatively as follows. On applying the gate pulse, corresponding depletion and inversion layers (if the pulse amplitude is greater
than \( V_T \) are formed in the substrate. The minority carriers in the inversion layer can come from one of the following sources:

- surface generation
- generation in the depletion region
- diffusion from bulk into the depletion region followed by drift to the surface.
- the source/drain regions.

The contribution of the first three sources can be neglected as compared to that from the source/drain regions. On removing the gate pulse, the inversion region charge flows back to the source/drain (in regions close to the junctions) while a part of it recombines with the majority carriers from the substrate (in regions far from the junctions) which forms a component of the charge-pumping current. This component is highly dependent on the geometry of the device and hence is termed the geometric component. The major component of the charge pumping current comes from the recombination of the fast surface states with
majority carriers from the bulk. This component is called the *surface state component* of the charge pumping current.

The reduction in $I_{CP}$ with increase in the reverse bias can be explained by the widening of the source/drain depletion regions. This reduces the effective channel length and hence the gate-area over which the interface traps contribute to the charge pumping current. As no inversion layer is formed when the gate pulse amplitude is below the threshold voltage, the charge pumping current drops to a very low value. This also means that this experiment can be used to obtain only the average trap density in the bandgap rather than the energy distribution of the surface states. Several variations of the basic experiment which can be used to obtain the energy distribution have been suggested and are discussed in Section ...

**Geometric Component of Charge Pumping Current**

After the gate pulse is removed, some of the minority carriers in the inversion layer recombine with the majority carriers from the substrate giving rise to the geometric component of the charge pumping current. This recombination occurs primarily in regions far from
the junctions where the field lines are perpendicular to the surface and is highly dependent on the rate at which the gate voltage is reduced. If the gate voltage is removed at a slow enough rate, minority carriers will get sufficient time to diffuse towards the source/drain junctions and contribute to the currents at the source/drain terminals thus reducing the geometric component of $I_{CP}$. This can be verified by using a sawtooth waveform instead of a rectangular pulse at the gate.

Another way to verify this is through the use of annular MOS capacitors as shown in Fig [13]. This geometry is highly unfavorable for charge removal through the junction and hence the charge pumping current here shows a non-saturating characteristics when a rectangular pulse is applied to the gate. The charge pumping current, however, saturates when the gate is excited with a triangular waveform. The magnitude of $I_{CP}$ also drops drastically [19].

Figure 13: Annular MOS Capacitor
**Surface State Component**

After the removal of the gate pulse, the pumped charge density can be expressed as:

\[ Q_{ss} = \alpha Q_{inv} + qN_{it} \]

where,

- \( Q_{ss} \): net pumped charge density per gate pulse.
- \( \alpha \): fraction of the inversion layer charge that recombines in the bulk.
- \( Q_{inv} \): free charge density in the inversion region.
- \( N_{it} \): surface state density per unit area.

The free charge density can be assumed to be zero for gate pulse amplitudes less than the threshold voltage and equal to \(-C_{ox}(V_g - V_t)\) for \(|V_g| > |V_t|\). Thus, for \(|V_g| > |V_t|\), one can write the charge pumping current as:

\[ I_{cp} = fA_G\left[ -\alpha C_{ox}(V_g - V_t) + qN_{it} \right] \]

where,

- \( C_{ox} \): oxide capacitance per unit area
- \( A_G \): gate area
- \( V_t \): MOS threshold voltage
- \( f \): frequency of the gate impulse

If the Fermi level at the surface goes from \( E_1 \) to \( E_2 \) during charge pumping measurements then the charge that will recombine with the majority carriers from the bulk is given by:

\[ Q_{ss} = qA_G \int_{E_1}^{E_2} D_{it}(E) dE \quad (4) \]

If the mean surface-state density is given by \( \overline{D_{it}} \) and the total sweep of the surface potential by \( \Delta \psi_s \), then:

\[ Q_{ss} = q^2 A_G \overline{D_{it}} \Delta \psi_s \quad (5) \]

\[ I_{cp} = fQ_{ss} = fq^2 A_G \overline{D_{it}} \Delta \psi_s \quad (6) \]
If the geometric component of the charge pumping current is small, the saturation of the charge pumping current can be explained by the pinning of the surface potential once strong inversion is reached. This is supported by the observation that the saturation occurs at the measured threshold voltage of the MOSFET. The charge pumping current is proportional to the region of the bandgap that is swept by the Fermi level when the gate pulse is applied and also to the surface trap density in this region. This region becomes constant after strong inversion is reached causing a saturation of the charge pumping current.

**Energy Distribution of Surface States**

As already mentioned, the basic charge pumping experiment as described by Brugler cannot be directly used to obtain the energy distribution of the surface states in the bandgap. An alternative approach [23] is to measure the charge pumping current as a function of the pulse base level keeping the pulse amplitude constant. Typical characteristics obtained from such measurements on a p-channel device are shown in Fig[14].

![Figure 14: Typical charge pumping characteristics using constant amplitude pulse and varying the base level.](image)

**Case A**  When the pulse base level is above the threshold voltage of the MOSFET, the
potential is always pinned at $2\phi_F$ and hence the occupancy of the surface states does not change when the pulse is applied at the gate. The resulting charge pumping current is very small and is primarily due to the geometric component (which is evident in the case when $V_R = 0V$).

**Case B** When the pulse base level is between 0 and $V_T$, some of the surface states contribute to the charge pumping current and $I_{CP}$ increases as the pulse base level approaches 0V.

**Case C** The charge pumping current goes on increasing even when the pulse base level goes beyond 0V as long as the pulse height is enough to invert the MOSFET at its maximum value.

**Case D** When the pulse base level reaches a value when the pulse is no longer able to invert the channel the charge pumping current drops to a very small value again.

The maximum reached by the charge pumping current decreases as the reverse bias is applied for reasons already mentioned. The falling edge of the characteristics shifts towards a lower voltage as the channel needs a larger gate voltage to be inverted now. This shift is equal to the sum of the reverse voltage and the shift in threshold voltage due to a voltage difference between the source and the substrate. The rising edge of the characteristics however remains unchanged on applying the reverse bias as the surface potential is not affected by the reverse bias $^1$.

In order to obtain the energy distribution from these measurements we need the effective gate area of the MOSFET device. The device width is unaffected by the reverse bias but the effective device length can vary as the reverse voltage on the source and drain is varied. The depletion width around the drain and source regions can be approximated using the following expression:

$$x_d = \sqrt{\frac{2\varepsilon_S}{qN_D}(V_R + \phi_{bs})}$$

$^1$Another variation of the approach is to apply a pulse with its base level fixed at a voltage in strong inversion and varying the pulse amplitude towards accumulation of the channel
where,
\[ \phi_{bi} \] : junction built-in potential.
\[ N_D \] : substrate doping concentration.

The effective channel width can be obtained by measuring the reverse bias at which the two depletion regions meet. This can be done by:

- applying a reverse bias at both the source and drain terminals and monitoring the gate capacitance. The gate capacitance will show a sudden dip when the two depletion regions meet (Fig [15]).
- increasing the reverse bias at the drain until punch-through is reached. Since punch-through cannot be distinguished clearly from avalanche injection, this method is not as accurate the previous one.

![Figure 15: Gate Capacitance as a Function of Source/Drain Bias](image)

The substrate doping concentration, \( N_D \), can be obtained by measuring the dependence of the threshold voltage on the substrate bias.
In order to obtain the energy distribution of the surface states, we need the relation between the applied gate voltage and the surface potential. This is usually obtained from quasi-static C-V measurements \(^2\). The measured C-V plot will contain parasitic capacitance (for example, due source/drain diffusions overlapping the gate electrode) which can be considered to be independent of the applied gate bias. The effect of these parasitic capacitances can be removed by normalizing the characteristics to have a maximum value of \(C_{ox}\). \(C_{ox}\) itself can be measured by three terminal measurement of the gate capacitance as a function of the gate-substrate bias\(^3\).

**Implementation**

The various characterization techniques discussed in this chapter and several variations of these techniques were implemented as a part of this work into the wafer-level reliability tools developed by Sandia National Laboratories in collaboration with Hewlett-Packard.

\(^2\)In order to maintain thermal equilibrium during quasi-static C-V measurements, a very slow sweep rate (\(\sim 1\) V/s) should be used

\(^3\)The time constant of a discrete trapping level is given by:

\[
\tau = \frac{1}{v_t \sigma} = \frac{1}{n_s + p_s + 2n_i \cosh \left[ \frac{q}{k_B T} (\psi_t - \psi_s) \right]}
\]

where,

\(\sigma\) : capture cross-section of the trap  
\(v_t\) : thermal velocity of the carriers  
\(n_s/p_s\) : surface concentration of electrons/holes  
\(\psi_t\) : trap potential  
\(\psi_s\) : surface potential
CHAPTER III

HOT-CARRIER DEGRADATION IN N-CHANNEL MOSFETS

Introduction

The degradation of n-channel MOSFETs due to hot-carrier injection has been studied for over 20 years[24, 11, 12, 25, 26]. A considerable amount of progress was made in the initial works in understanding the injection mechanisms and generation of interface traps in long channel devices[14, 27, 28, 29, 4]. However, as the technologies evolved and device sizes shrunk, the theories based on long channel devices were not sufficient to explain the observed device response to hot-carrier injection[30, 31, 32, 33, 34]. The exact nature of the processes involved in hot-carrier induced MOSFET degradation is still not well understood and a lot of disagreement exists between the results presented by different hot-carrier studies. The principal reason for the lack of understanding and agreement is the inability of any electrical characterization technique to provide a clear evaluation of the underlying phenomena. A combination of several independent characterization techniques can sometimes yield a clearer understanding but still requires some amount of speculation to explain all the observations. Other factors such as the two dimensional nature of the injection process[35], the coupling between different injection regimes as the gate and drain biases vary, and the strong dependence of the injection processes on the device structure and technology further complicate the analysis of data obtained from the hot-carrier stressing experiments[36]. This chapter provides a summary of the present understanding of the processes involved during hot-carrier degradation of n-channel MOSFETs. The stressing and modeling techniques which have been used in past literature have been evaluated in the light of experiments performed on a commercial sub-micron technology. This study provides a clear understanding of the shortcomings of the current techniques for evaluation of hot-carrier damage in n-channel devices. Based on the results of this analysis, a set of hot-carrier stressing experiments have been suggested to study comprehensively the response of n-channel devices to
injection mechanisms present during real-life circuit operation.

**Carrier Injection and Gate Currents**

The hot-carrier induced degradation processes are initiated by the injection of high-energy carriers from the channel of the device into the gate oxide. The physical mechanisms which are involved in the device degradation are strongly dependent on the relative concentration of electrons and holes injected at any given location along the channel. This section provides a qualitative analysis of the dependence of electron and hole injection currents on the gate bias of a conventional n-channel MOSFET.

![Figure 16: The dependence of carrier injection and gate current on the gate bias in a conventional n-channel MOSFET.](image-url)

The qualitative nature of electron and hole injection currents as well as the gate current
in n-channel MOSFETs is shown in Fig. 16. The injection of carriers is determined by two factors: the concentration of carriers in the channel and the accelerating lateral electric field near the point of injection.

In the subthreshold region, the concentration of electrons and holes in the channel is relatively small which results in a negligible injection of both the carriers. As the gate bias is increased, the carrier supply in the channel increases exponentially resulting in the increase of both electron and hole injection currents. At relatively low gate biases, the transverse electric field and the Schottky barrier lowering favor the injection of holes\[2, 17\] over electrons. The oxide field under this condition also favors the transport of the injected holes to the gate terminal resulting in a gate current comprised primarily of holes.

As the gate bias increases, the barrier lowering effect for holes levels off while the electron injection current continues to increase due to a relatively smaller barrier height for electrons. However, as long as a repulsive field is present in the oxide, the injected electrons are scattered back to the interface. Thus, the gate current under these gate biases is negligible as compared to the electron injection current. At higher gate biases, the repulsive oxide field decreases and a larger proportion of injected electrons contribute to the gate current resulting in a peak gate current around \( V_{GS} = V_{DS} \). For \( V_{GS} > V_{DS} \), the lateral electric field close to the drain decreases with increasing gate bias and hence the electron injection current also decreases. However, the presence of an attractive oxide field results in a gate current which is almost equal to the electron injection current at these biases.

Even though one can only measure the carriers which reach the gate terminal of the device, the presence of large concentration of injected electrons in the mid-gate-bias region and holes in the low-gate-bias region has been confirmed using split-gate transistors\[37\] and two dimensional simulations\[38\]. Some of the main features of the gate and injection current characteristics that should be noticed here are:

- In n-channel devices, electrons are injected into the gate oxide under all gate biases.

- The maximum electron injection current occurs when the measured gate current magnitude is close to its minimum value (in the mid-gate-bias region).

- The peak of the hole-injection current is considerably smaller than that of the electron-
injection current due to the relatively large interfacial barrier for hole-injection.

**Substrate Currents**

The substrate current in n-channel devices consists mainly of the holes generated through impact ionization in the channel region. The impact ionization rate is dependent on both the concentration of carriers in the channel and the lateral electric field. Under low gate biases, the supply of channel carriers increases with the gate bias resulting in an increase in the substrate current (Fig. 17). However, as the gate bias reaches close to the drain bias, the lateral electric field decreases resulting in a decrease in the substrate current. Thus, the variation of the relative equilibrium between the carrier supply and the lateral electric field results in the well-known bell-shaped substrate current characteristics in n-channel devices.

**Lucky-Electron Model**

The injection of hot-electrons into the gate-oxide has been popularly modeled using the “lucky-electron” model[4, 35]. The lucky-electron concept models the probability of a channel electron reaching the gate terminal as a combination of the probabilities of the following events:

1. the electron gains sufficient energy in the lateral electric field to overcome the interfacial potential barrier and retains this energy after a collision directs its momentum towards the interface,
2. the electron reaches the interface without suffering any more collisions, and
3. the electron is not scattered back into the semiconductor in the image-force potential well present close the interface[28](Appendix A).

The lucky-electron model in its original form as well as with some modifications suggested by other authors[12, 39] suffers from certain serious problems[2]. Despite the objections raised against the model by several studies, the lucky-electron model remains the most extensively used approach to estimate the injection currents due to its simplicity and reasonable agreement with experimental results. On the basis of the lucky-electron concept, it can be shown
Figure 17: The substrate current vs. gate bias characteristics show a bell shaped curve. The initial increase in substrate current is due to increase in the carrier supply while the decrease in substrate current at high gate biases is due to a decrease in the lateral electric field.
that the fraction of the channel carrier supply which is injected into the gate oxide is given by:

$$\frac{I_{inj}}{I_{supply}} \propto \exp \left( -\frac{\phi}{q\lambda E} \right)$$

(7)

where, $\phi$ is the interfacial potential barrier, $\lambda$ is the mean free path of the carriers and $E$ is the effective lateral electric field. As the substrate current is generated due to impact ionization in the lateral electric field close to the drain, it is linearly proportional to the drain current. Based on the lucky-electron concept, it can be shown that the ratio of the substrate and drain currents can be approximated by:

$$\frac{I_B}{I_{DS}} = C \exp \left( -\frac{\phi_i}{q\lambda E} \right)$$

(8)

where, $\phi_i$ is the impact ionization energy. As seen from this equation, the ratio of the substrate and drain currents provides a direct measure of the average lateral electric field. This ratio, also called the multiplication factor is often used as a monitor for the device degradation.

**DC Stressing Experiments**

As shown in Section, the relative concentration of electrons and holes injected into the gate oxide is a strong function of the gate bias. In the earlier studies of n-channel devices, it was observed that the degradation in n-channel devices is predominantly due to increase in interface trap density[35, 40]. A strong correlation was noticed between the device degradation due to interface trap formation and the substrate current[40] with the maximum degradation observed at the gate bias which resulted in the maximum substrate current. In most of the earlier technologies, it was also shown that the results of hot-carrier stressing experiments performed under the condition of maximum substrate current can be extrapolated to predict the device lifetimes under real-life circuit operation[41, 42, 43, 44]. Due to the success of the lucky-electron approach to model the increase in interface trap density during these experiments, hot-carrier stressing under maximum substrate current bias condition, $V_{GS, ib_{max}}$, became a standard test for monitoring n-channel device degradation[45, 35].
some of the more recent studies, it was observed that the device degradation models based on this stressing technique are not sufficient to correctly predict the time dependence of device degradation under dynamic operation in certain technologies[36].

Hot-carrier stressing experiments in some of the recent works, however, show that hole injection under low gate-biases and electron injection under high gate biases can also result in significant device degradation under circuit operation[6, 46, 47, 48]. Mistry et al. identified three dominant degradation modes in n-channel MOSFETs and incorporated their effects into an AC lifetime model[33]. According to their study, the device degradation is dominated by increase in interface traps at mid-gate biases ($V_{GS} \approx V_{DS}/2$), close to $I_{B,\text{max}}$ condition, as expected. At the same time, interface traps as well as oxide traps are generated at low gate biases ($V_{GS} \approx V_{DS}/5$) under hole injection ($H_{\text{inj}}$) while the device degradation is dominated by carrier trapping during electron injection ($E_{\text{inj}}$) at high gate biases ($V_{GS} \geq V_{DS}$) along with a small increase in interface trap density. The expected degradation modes under each of the three injection conditions along with the type of carriers injected are listed in Fig. 18.

<table>
<thead>
<tr>
<th>Name</th>
<th>Injected Carrier</th>
<th>Degradation Process</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>elec</td>
<td>hole</td>
</tr>
<tr>
<td>$E_{\text{inj}}$</td>
<td>y</td>
<td>n</td>
</tr>
<tr>
<td>$H_{\text{inj}}$</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>$I_{B,\text{max}}$</td>
<td>y</td>
<td>y</td>
</tr>
</tbody>
</table>

Figure 18: The dependence of injection and degradation modes on the gate bias in n-channel MOSFETs.

In order to monitor the device behavior under each of these degradation modes, we performed three different set of single-bias DC stressing experiments on a set of test structures. The test structures used in the results presented in the rest of this chapter were commercial LDD n-channel MOSFETs with $L \times W = 0.5\mu m \times 10\mu m$ with an oxide thickness of 10nm. The drain was biased at 5.2V during all of these stressing experiments. The $I_{B,\text{max}}$ condition for these test structures corresponds to $V_{GS} = 1.9V$ while the $E_{\text{inj}}$ experiments were per-
formed at $V_{GS} = 5.2V$. Each of these stressing experiments were performed for a duration of 3600 seconds.

The $H_{inj}$ condition corresponds to $V_{DS} = 5.2V$ and $V_{GS} = V_{DS}/5 = 1.04V$. Under this bias condition, the trapped holes mask the effect of increase in interface trap density and hence the degradation cannot be observed in drain characteristics of the devices. In order to expose the interfacial damage in these devices, hole-injection steps of 180 seconds each were followed by electron injection for 15 seconds to neutralize the trapped holes[33]. Twenty such hole injection steps were performed to give a total hole injection period of 3600 seconds.

In a real circuit, a device will be subjected to each of the three degradation modes as the applied gate and drain biases change as a function of time. The combined effect of all the three degradation modes and the coupling between the modes were studied by subjecting the same device to the above $E_{inj}$, $I_{B_{max}}$, and $H_{inj}$ bias conditions, in that order(Sequence 1), for 1200 seconds at each bias condition. In another set of experiments, the devices were stressed at the same bias sets in the reversed order(Sequence 2), i.e. $H_{inj}$, $I_{B_{max}}$ and $E_{inj}$, for 1200 seconds at each bias.

Each stressing experiment was repeated on up to 13 devices from the same wafer. The results shown here represent typical measurements. Several device parameters were measured as a function of the stress time during each of these stressing experiments. The results shown here use the percentage decrease in maximum linear transconductance ($\% \Delta G_m$) as a monitor of the device degradation. The transconductance, $G_m$, is measured as the maximum slope of $I_{DS}$-$V_{GS}$ characteristics of the device measured at $V_{DS} = 0.1V$. The drain, gate, source, and substrate currents under the stress biases were also monitored as a function of stress time during each stressing experiment in order to parameterize some of the degradation models (Section ).

Device Degradation Models

In order to predict the device degradation due to hot-carrier injection under circuit operation, several empirical and semi-empirical models have been suggested in the past literature. This section provides an overview of some of the most popular models that are presently being used to model parameter shifts and device lifetimes.
The Takeda and Suzuki Model

Takeda and Suzuki[40] suggested a simple time dependent model for parameter shifts based on the experiments performed under the I_{B,max} condition. In their experiments, they observed that ΔV_T or %ΔG_m could be modeled empirically using expressions such as:

\[ \%ΔG_m = At^n \]  \hspace{1cm} (9)

where, A and n are empirical parameters extracted separately for each technology. The parameter n has a strong dependence on the gate bias used during the stressing experiments but little dependence on the drain bias. Similarly, the parameter A shows a strong dependence on the drain bias while its independent of the gate bias. In order to allow extrapolation from the drain bias used under stressing conditions to real-life drain biases, the dependence of A on the drain bias can be modeled as:

\[ A \propto \exp \left( -\frac{1}{V_{DS}} \right) \]  \hspace{1cm} (10)

The simplicity of this model allows quick and easy extrapolation of device lifetimes under stressing conditions to the real-life biases. This approach has been used extensively for fast first-order benchmarking of different technologies. In the past, a duty-cycle based extension of this model has been applied to evaluate device lifetime under AC operation[41]. However, this model finds little application in predicting the dynamic degradation of present generation MOSFETs.

The Hu Model

Hu et al.[35] extended the lucky electron approach to obtain a semi-empirical model for predicting device degradation under circuit operation. The physical basis of this model and the use of this model in circuit reliability simulators such as BERT[49] has made this the most popular hot-carrier degradation model for n-channel MOSFETs. A complete analysis and the assumptions involved in this model can be found elsewhere[4, 35]. For the purpose of this work, we will use a popular form of this model as given by:

35
\[ \% \Delta G_m(t) = \left[ A \int_0^t \left( \frac{I_B}{I_{DS}} \right)^m I_{DS} \, dt \right]^n \] (11)

where, A, m, and n are empirical parameters. As the degradation model is driven by substrate and drain currents, it can be directly applied to device degradation under dynamic operation once the empirical parameters have been obtained from single-bias DC stressing experiments. This model assumes that the device degradation is entirely due to increase in the interface trap density and is usually parameterized using stressing experiments performed under the \( I_{B,\text{max}} \) condition.

**The Woltjer Model**

The “lucky-electron” model predicts the interface trap generation at maximum-substrate current conditions but fails at other bias conditions. The model suggested by Woltjer et al.\cite{34} incorporated the dependence on the oxide electric field in this model to extend it to all biases under which the effect of interface traps dominates. The influence of oxide charge has been avoided in this work by defining the device lifetime at a large number of interface states in terms of the change in charge pumping current. The effect of the charge has, however, been observed if the lifetime is defined as a smaller value. As interface trap generation depends approximately exponentially on the oxide electric field, \( E_{\text{ox}} = (V_g - V_d - V_t)/t_{\text{ox}} \), the following correction to the “lucky-electron” model has been suggested:

\[ \frac{I_{cp}}{e} \propto E_{\text{ox}}/E_o \] (12)

where, \( \tau_{I_{cp}} \) is the charge-pump lifetime defined as \( \Delta I_{cp}/W/f = 100pA/m/Hz \) and \( E_o \) is an empirical parameter.

The degradation data obtained on MOSFETs with different dimensions and oxide-thicknesses has been fitted to the following model using the same set of fitting parameters for the technology used\cite{50}:

\[ \log_{10}(\tau_{I_{cp}}) = \left[ -A \cdot \log_{10} \left( \frac{I_b}{I_d} \right) \right]^n - \log_{10} \left( \frac{I_d}{W} \right) - \frac{E_{\text{ox}}}{E_o} + C_{W3} \] (13)

where, A, n, \( E_o \) and \( C_{W3} \) are the fitting parameters. This gives the following relationship for the degradation of the charge-pumping current as a function of time:

36
The empirical parameter, $A$, $B$, $C$, $m$, $n$, and $l$ are extracted from the stressing experiments performed under the corresponding bias conditions. For example, the parameters $A$ and $m$ are extracted by stressing the device under $I_{B,\text{max}}$ bias condition.

The device lifetime under dynamic operation can be obtained in terms of these models using the following approach. If the terminal currents of the device are known as a function of time during the dynamic operation of the device, the contributions of each of the three degradation mechanisms over one cycle of the AC waveform can be calculated by integrating the individual models listed above:

$$\frac{1}{\tau_{I_{B,\text{max}}}} = \frac{1}{AT} \int_0^T I_B^m dt$$

(18)
\[
\frac{1}{\tau_{H_{\text{inj}}}} = \frac{1}{BT} \int_0^T \left( \frac{I_B}{I_{DS}} \right)^n I_{DS} dt
\]

\[
\frac{1}{\tau_{E_{\text{inj}}}} = \frac{1}{CT} \int_0^T \left( \frac{I_G}{I_{DS}} \right)^l I_{DS} dt
\]

(19) (20)

If each of the integral above can be treated as damage functions, the dynamic lifetime, \(\tau_{AC}\), can be obtained using a Matthiessen-like rule as follows:

\[
\frac{1}{\tau_{AC}} = \frac{1}{\tau_{I_{B,\text{max}}}} + \frac{1}{\tau_{H_{\text{inj}}}} + \frac{1}{\tau_{E_{\text{inj}}}}
\]

(22)

This approach has been successfully applied to predict the device degradation under dynamic operation in simple digital CMOS circuits based on the three sets of single-bias DC stressing experiments. However, in order to be used in a circuit reliability simulator, a model for device degradation should predict the time dependence of the degradation rather than the lifetime. As the dynamic degradation of n-channel MOSFETs has been shown to have a much stronger time dependence as compared to static degradation[36], the definition of the lifetime itself can affect the validity of lifetime models for certain technologies.

**Model Comparison and Coupled Empirical Approach**

In order to assess the validity and applicability of the existing degradation models, each of the above models were parameterized and applied to predict the device degradation under the bias sequences.

The device degradation under the \(I_{B,\text{max}}\) as a function of time is shown in Fig. 19. The data obtained during this stressing experiment has been used to parameterize the Takeda and Hu models given by Eqs. 9 and 11 respectively. The values obtained from the parameterized models are also plotted in Fig. 19. As can be seen from this figure, both the models are able to accurately follow the device degradation under the \(I_{B,\text{max}}\) bias condition.

The device behavior under \(E_{\text{inj}}\) and \(H_{\text{inj}}\) bias conditions is shown as a function of time in Figs. 20 and 21, respectively. If the mechanism represented by the above models (by Takeda and Suzuki and Hu et al.) is the only dominant degradation mode under all bias combinations, these models should be able to predict the device degradation under \(E_{\text{inj}}\) and...
Figure 19: The device degradation as a function of time under $I_{B,max}$ bias condition. The values obtained from the parameterized models are also plotted along with the measured data.
as seen from Figs. 19, 20, and 21, these models fail to predict the device behavior correctly at biases other than the $I_{B,max}$ condition.

The device degradation under the DC biases sequences is shown in Figs. 22 and 23. Once again, the application of the conventional models to predict the device degradation under the bias sequences results in highly inaccurate prediction of the device behavior. These results clearly show that the application of models extracted from stressing experiments performed at a single bias condition to device operation under varying bias conditions, as present during circuit operation of the device, can result in incorrect prediction of the device behavior.

In another attempt at empirical modeling, device degradation under each of the three bias conditions was modeled using time dependent equations based on the lifetime models suggested by Mistry et al.[33] as given by:

$$I_{B,max}$$

$$\%\Delta G_m(I_{B,max}) = A I_B^m t^n$$  \hspace{1cm} (23)

$$H_{inj}$$

$$\%\Delta G_m(H_{inj}) = B \left( \frac{I_B}{I_{DS}} \right)^p (I_{DS})^q$$  \hspace{1cm} (24)

$$E_{inj}$$

$$\%\Delta G_m(E_{inj}) = C \left( \frac{I_G}{I_{DS}} \right)^r (I_{DS})^s$$  \hspace{1cm} (25)

These equations were fitted to the data obtained under the corresponding bias conditions in order to obtain the model parameters. For example, the parameters B, p, and q in Eq. 24 are obtained through empirical fit using the data of the type shown in Fig. 21. If a single degradation mechanism dominates at each of these bias conditions, the summation of the contributions from each of the models should give the total device degradation under the bias sequences. However, a simple addition $(\%\Delta G_m(I_{B,max}) + \%\Delta G_m(E_{inj}) + \%\Delta G_m(H_{inj}))$ of contributions from each of these models to predict the device behavior under the bias sequences results in an overestimation of the device degradation. The Matthiessen-like rule suggested by Mistry et al. can be used, with different values of $\%\Delta G_m$ to define the device lifetime, to obtain $\%\Delta G_m$ function of time. The application of this approach also results in an overestimation of the damage for stressing under the two bias sequences. These results
Figure 20: The device degradation as a function of time under $E_{\text{inj}}$ bias condition. The parameterized conventional models are unable to correctly predict the device degradation under this condition.
Figure 21: The device degradation as a function of time under $H_{\text{inj}}$ bias condition. The parameterized conventional models are unable to correctly predict the device degradation under this condition.
Figure 22: Device degradation as a function of time for devices stressed using bias sequence 1. The combined effect of all three dominant degradation modes needs to be included at each bias condition to correctly predict the degradation.
Figure 23: Device degradation for stressing under bias sequence 2. Conventional models deviate from the observed values while the coupled application of three models results in an improved prediction.
suggest that multiple mechanisms are responsible for the device degradation under each of the bias conditions. We attempted a novel approach to model the combined effect of the three mechanisms at each of the bias conditions using the following expression:

\[ \% \Delta G_m = \int \left[ A I_B^{m} I_m^{n} \left( \frac{m}{I_B} dI_B + \frac{n}{t} dt \right) \right] + \int \left[ B \left( \frac{I_B}{I_{DS}} \right)^p (I_{DS})^q \left( \frac{p}{I_B} dI_B + \frac{(q - p)}{I_{DS}} dI_{DS} + \frac{q}{t} dt \right) \right] + \int \left[ C \left( \frac{I_G}{I_{DS}} \right)^r (I_{DS})^s \left( \frac{r}{I_G} dI_G + \frac{(s - r)}{I_{DS}} dI_{DS} + \frac{s}{t} dt \right) \right] \] (26)

The three terms on the right hand side of this equation include the contributions due to each of the three degradation modes. These terms are obtained by adding the partial derivatives of the expressions in Eqs. 23, 24, and 25 with respect to time and each of the terminal currents. The model has been shown in this form to emphasize the fact that the variations in the device currents are used explicitly when obtaining the model parameters. This coupled equation has been fitted to the data shown in Figs. 19, 21, and 20 simultaneously in order to obtain the associated parameters. As seen from these figures, the values obtained from the fitted model closely follow the measured data. As the model includes the contributions due to all the three modes at any given bias condition, it is also expected to predict the device degradation under arbitrary combinations of drain and gate biases. In order to verify this, the parameterized model is used to predict the device degradation under the two bias sequences as shown in Figs. 22 and 23. As seen here, the coupled model is able to predict the device degradation under varying bias conditions more accurately than any of the other models. As the model predicts the time dependence of the device degradation, it is also adequate for use in circuit reliability simulators.

All of the above empirical models have had limited amount of success in predicting the hot-carrier response under dynamic operation. The models which are based on physical principles, however, make several assumptions about the field distribution and degradation processes in these MOSFETs[35]. These assumptions cannot be justified for deep sub-micron devices. On the other hand, completely empirical approaches (such as the coupled approach presented above) provide limited insight into the physical mechanisms for hot-carrier degra-
dation and may not work for all technologies. In view of these observations, we suggest that the focus of hot-carrier modeling efforts should shift from models based on macroscopic quantities, such as the terminal currents and biases, to microscopic parameters such as the variations in the electric fields in the semiconductor and the oxide due to carrier injection, trapping and interface trap generation processes. This effort is suggested as a future work for the continuation of this study in Chapter IV.

Dynamic Degradation Considerations

The dynamic hot-carrier degradation modes in MOS transistors have been a subject of considerable controversies. It has been reported[51] that “strong transient effects directly related to intrinsic time constants of the MOSFET are absent down to slopes in the sub-nanosecond regime”. Thus, the MOSFET intrinsic time constants will not significantly affect the dynamic degradation.

The dynamic response of the injected carriers is determined by their transport and trapping properties in the oxide and the interface. Electrons have been reported to have a drift mobility of about 20cm$^2/(V\cdot s)$ under an oxide-field of 1MV/cm. Under these conditions, an electron injected or generated into the oxide will move out of a 200 Å oxide in less than a picosecond. Holes, on the other hand, have drift mobilities which can be several orders of magnitude less than electron mobilities. A field-independent drift-mobility of $2 \times 10^{-5} cm^2/(V\cdot s)$ has been reported[52]. The presence of oxide defects in the vicinity of the hole can further decrease the effective mobility of the holes. This can be explained by the fact that the hole transport in oxides occurs through hoping between neighboring oxygen atoms with an activation energy of hoping of about 0.16eV and since the activation energy of hoping between defects (0.37eV) is much higher, hole transport is slowed down around these defects. A high mobility phase of holes within a few picoseconds after generation has also been reported. The drift mobility of holes could be as high as 1cm$^2/(V\cdot s)$ during this period[53]. The positively charged hole polarizes the SiO$_2$ lattice during this picosecond period in such a way that the hole gets localized on an oxygen atom of a Si-O-Si bond. Further transport of the hole occurs through the hoping mechanism mentioned above and the mobility decreases correspondingly. The high mobility period of holes, while observed
for holes generated inside the oxide, is not present for holes generated in the substrate and then injected into the gate-oxide.

Another phenomenon which significantly affects the hot-carrier device degradation during circuit operation is the detrapping of trapped charge. This mechanism can be active over periods ranging from a few nanoseconds to several days and hence becomes a major concern in evaluating the hot-carrier instabilities of devices in real-life operating conditions.

The trapping of electrons in the gate-oxide is the most significant mode of hot-carrier degradation in p-channel MOSFETs. Brox and Weber[52] studied the detrapping behavior of trapped holes in p-MOSFETs by applying a constant negative bias on the gate after hot-carrier stress. The device recovery during the post-stress period was observed to be strongly dependent on the oxide field and was logarithmic in time. This logarithmic nature of the recovery characteristics indicates that the detrapping could not be the result of field assisted thermal emission or trap-band impact ionization with Fowler-Nordheim injected electrons as the detrapping would be expected to be exponential with time under these mechanisms. A possible mechanism could be tunneling discharge of the trapped electrons out of the oxide. This detrapping of electrons during device operation can lead to device lifetimes in excess of those estimated using simple electron trapping models. The strong dependence of electron detrapping on the oxide field suggests that this phenomenon will become more important in future sub-μm devices as the oxide thickness also scales accordingly.

The comparison of dynamic hot-carrier stressing experiments with static experiments shows that the dynamic stress produces a steeper time dependence as compared to a single bias static stress[36]. Devices stressed under maximum substrate current conditions were compared with devices stressed under inverter like dynamic conditions. The dynamic degradation showed a much stronger time dependence as compared to the static conditions. The approximate nature of the curves as extracted from [36] is shown in Fig. 24. Extrapolation of the static stressing results to dynamic conditions by accounting for duty-cycle will only result in a shift of the curves along the time axis without any change in the slope of the curves. Investigation of any transient effects that might be causing this behavior shows that there is no change in the nature of the dynamic stressing characteristics for frequencies ranging from 25Hz ($t_r = t_f = 3$ms) to 25MHz ($t_r = t_f = 3$ns). This indicates that differences
between static and dynamic degradation behaviors are not due to transient effects\textsuperscript{1}. Alternating stress conditions at two different biases ($V_g = 1\text{V}, V_d = 8\text{V}$ and $V_g = 8\text{V}, V_d = 8\text{V}$) show a stronger time dependence as compared to the static stress conditions.

We conclude that the differences between static and dynamic behaviors are due to sequences of various stressing biases under dynamic condition as compared to a single bias static condition. These results clearly show that static stress under a single fixed bias cannot be directly correlated to dynamic stress conditions.

![Figure 24: Dynamic degradation shows a stronger time dependence than static degradation.](image)

At low gate biases, hole-injection is the dominant degradation mode while electron in-

\textsuperscript{1}The authors, however, mention that in some technologies, such as the conventional nitride passivated devices, strong transient effects were observed.
jection dominates at higher gate biases. In an alternating stress condition, hole injection and interface trap formation occur under low gate biases. The presence of positive charge in the gate oxide results in reduction of the channel electric field. A continued stress at this condition will result in a corresponding reduction in injected holes. Injection of electrons at higher gate biases neutralizes the positive charge causing the electric field to increase again. This leads to a stronger dependence on stressing time under the alternating static stressing conditions. This can be confirmed by monitoring the substrate current under these stressing conditions. The substrate current shows a much larger reduction under a single bias stress as compared to an alternating stress at two different biases.

It has been observed by Bellens et al. [54] that the experimental setup used during AC stressing measurements can significantly affect the results. In some cases, the purpose of the experiment cannot be fulfilled unless these effects are eliminated. An enhanced degradation and substrate current component has been observed during the falling edge of gate when the drain is held high. Bellens et al. observed this to be due to insufficient coupling of the source to the ground due to parasitic inductance in the wiring used. The enhanced substrate current can be eliminated by putting a resistor at the drain and grounding the source at the probe tip.
A comprehensive study of the hot-carrier degradation mechanisms and the currently available models for parameter degradation in n-channel MOSFETs has been presented in this thesis. The stressing experiments indicate that multiple degradation modes will be responsible for the device response under dynamic operation. The conventional models for hot-carrier degradation assume that a single degradation mechanism dominates device behavior. These models accurately predict device degradation under single-DC-bias stressing experiments but result in incorrect predictions when the devices are subjected to multiple degradation modes by varying the applied biases.

A coupled model based on the approach suggested by Mistry et al.[33] has been successfully used in this work to predict device behavior under bias sequences on the basis of stressing experiments performed at single-DC biases. In recent literature, other similar empirical modeling approaches have been used to predict hot-carrier response of n-channel MOSFETs. However, the highly empirical nature of these methodologies results in loss of insight into the underlying physical mechanisms. In extensively scaled device, these fundamental degradation mechanisms are also expected to be strongly dependent on the device structure, the underlying electric fields during device operation and the exact location of the hot-carrier induced damage. The localized nature of hot-carrier damage makes it impossible to attribute the shifts in macroscopic parameters, such as the threshold voltage and transconductance, to definite physical mechanisms unambiguously. The stressing experiments suggested in this work need to be supplemented by a rigorous analysis of the degradation mechanisms at the microscopic level and their dependence on the technology and fabrication process used to clearly understand the hot-carrier phenomena in deep submicron device.
Appendix A

ELECTRON SCATTERING IN IMAGE FORCE POTENTIAL WELL

The classical image force theory has been applied to obtain the potential around interfaces between conductors and insulators. It has been shown that even though the image force law holds asymptotically at large distances from the interface, it does not describe the potential close to the interface accurately. The assumption that the interface between the conductor and insulator is homogeneous also breaks down at distances comparable to inter-atomic distances. Under these conditions the concepts of image charge and dielectric constant have no meaning[28]. The image charge concept has been used to describe the Schottky lowering of potential barrier between Si and SiO$_2$ under an applied field. The barrier height can be accurately predicted as long as the potential maximum (highest value of the barrier) does not occur within atomic dimensions from the interface. This condition can occur under high electric fields which are present in modern sub-$\mu$m devices. Experimental results indicate that the image force concept should hold for distances greater than 10Å from the interface.

The various components of the potential inside a dielectric close to the (semi-)conductor-dielectric interface are shown in Fig. [25]. The electrostatic potential in the dielectric is given by Eq. (27). In this equation, $\phi_1$ is the interface potential step at $x=0$, the second term ($\int_0^x E(z)dz$) is the potential resulting from the electric field due to applied biases and the third term ($-q/16\pi\epsilon_i x$) is the classical image force potential.

$$\phi(x) = \phi_1 + \int_0^x E(z)dz - \frac{q}{16\pi\epsilon_i x} \tag{27}$$

The field in the oxide, $E(x)$, consists of two components—the field due to any applied biases at the two interfaces of the dielectric and the field due to any charge situated inside the dielectric. If $V_a$ is the applied voltage across the dielectric, $\phi_1$ and $\phi_2$ are the barrier energies at the two interfaces and $E_1(x)$ is the field due to the charge inside a dielectric of thickness $d$, then:
The distance, $x_0$, to the potential maximum can be obtained by differentiating Eq. (27) with respect to $x$ and equating it to zero. This gives:

$$x_0 = \left[ \frac{q}{16\pi\varepsilon_x E(x_0)} \right]^{\frac{1}{2}}$$

The potential at this point is:

$$\phi_{\text{max}} = \phi_1 - \int_0^{x_0} E(x)dx - \frac{q}{16\pi\varepsilon_i x_0}$$

Consider an electron injected into the dielectric with a momentum $p$ directed at an angle $\theta$ to the normal to the interface (Fig. 26) and energy $\mathcal{E}$. The electron will follow a non-linear trajectory due to the non-uniform field in the dielectric.

Let us first assume that the electron does not suffer any electron-phonon collisions before reaching $x=x_0$. In order for the electron to reach $x_0$ without being turned around by the electric field, it must have an initial momentum normal component $p_n$ greater than some critical value $p_c$ which corresponds to the potential barrier energy. This means that an
Figure 26: Electron trajectory in the dielectric after emission at an angle $\theta$ to the normal to the interface.

electron with a certain given momentum $p$ will be injected into the oxide only if the angle $\theta$ is less than a certain critical angle $\theta_c = \cos^{-1}(p_c/p)$. If the momentum direction is randomly distributed then the probability of escape of an electron is:

$$
P_{esc} = \begin{cases} 
1 - \frac{p_c}{p} & \text{if } p \geq p_c \\
0 & \text{if } p \leq p_c
\end{cases}
$$

(31)

The analysis presented above holds if the electron mean-free-path in the dielectric is large as compared to $x_0$. However, electron-phonon interactions can result in mean-free-paths comparable to $x_0$ and scattering effects need to be included in the escape probability. Let the mean-free path in the dielectric as a function of the depth inside the dielectric be given by $l(x)$. The probability that an electron injected at an angle $\theta$ will reach $x_0$ without scattering is given by:

$$
P(x_0) = \exp \left( -\int_{0}^{x_0} \frac{dx}{l(x) \cos(\theta)} \right)
$$

(32)

Assuming that the electron velocity angle is randomly distributed over the half sphere, the total probability that an electron injected into the dielectric with a momentum $p$ will reach $x_0$ without scattering is given by Eq. (32) averaged over the half sphere while $\theta$ is constrained to be less than $\theta_c$ (Eq. 33).
\[ P_0 = \frac{1}{2\pi} \int_0^{2\pi} d\psi \int_0^{\cos^{-1}(p_c/p)} \exp \left( -\int_0^{x_0} \frac{dx}{l(x)} \right) \sin(\theta) d\theta \]  

(33)

Putting \( y = (1 - \cos(\theta)) \) in Eq. (33) and assuming \( l(x) = l \), we get :

\[ P_0 = \int_0^{[1-(p_c/p)]} \exp \left( \frac{-a}{1-y} \right) dy \]

\[ a = \int_0^{x_0} \frac{dx}{l(x)} = \frac{x_0}{l} \]

(34)

Since the critical angle, \( \theta_c \), is small for energies encountered in the study of hot-carrier injection, we can assume that \([1 - (p_c/p)]\) is small compared to unity. This allows one to simplify the integration in Eq. (34) to give :

\[ P_0 \approx \int_0^{[1-(p_c/p)]} \exp(-a(1 + y)) dy \]

\[ = \frac{l}{x_0} \left\{ 1 - \exp \left[ -\frac{x_0}{l} \left( 1 - \frac{p_c}{p} \right) \right] \right\} \exp \left( -\frac{x_0}{l} \right) \]

(35)

Further assumption that \((x_0/l)[1 - (p_c/p)]\) is small as compared to unity gives :

\[ P_0 \approx \left[ 1 - \frac{p_c}{p} \right] \exp \left( -\frac{x_0}{l} \right) \]

(36)

Similarly, for \((x_0/l)[1 - (p_c/p)] >> 1\) :

\[ P_0 \approx \frac{l}{x_0} \exp \left( -\frac{x_0}{l} \right) \]

(37)

It was assumed here that the electron does not suffer any scattering before reaching \( x = x_0 \). The process of injection after one or more scattering events can also have comparable probability under certain conditions. The inclusion of this probability results in the following expression for the total probability of escape[28] :

\[ P_{esc} = \left( 1 - \frac{p_c}{p} \right) \left( 1 + g\left( \frac{x_0}{l} \right) \right) \exp \left( -\frac{x_0}{l} \right) ; \quad p > p_c \]

(38)

Here, \( g(x_0/l) \) vanishes as \( x_0/l \) approaches zero and while it can exceed unity for large \( x_0/l \), its variation with respect to \( x_0/l \) is small as compared to the exponential term. Thus, if we
neglect the variation in $g()$ and express the factor $(1 - (p_c/p))$ in terms of the corresponding electron energies, the escape probability is given by:

$$P_{esc}(E, x_0) = C \left[ 1 - \left( \frac{E_b}{E} \right)^{\frac{3}{2}} \right] \exp \left( -\frac{x_0}{l} \right); \quad E > E_b \quad (39)$$

In Eq. (39), $C$ is a constant and $E$ and $E_b$ are the electron energy and the interface barrier energy respectively.
REFERENCES


58


