AN INVESTIGATION OF CARRIER TRANSPORT IN HAFNIUM OXIDE/SILICON DIOXIDE MOS GATE DIELECTRIC STACKS FROM 5.6-400K

By

Richard G. Southwick III

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Electrical and Computer Engineering

Boise State University

December 2010

© 2010

Richard G. Southwick III

ALL RIGHTS RESERVED

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Richard G. Southwick III

Dissertation Title: An Investigation of Carrier Transport in Hafnium Oxide/Silicon Dioxide MOS Gate Dielectric Stacks from 5.6-400K

Date of Final Oral Examination: 13 August 2010

The following individuals read and discussed the dissertation submitted by student Richard G. Southwick III, and they also evaluated his presentation and response to questions during the final oral examination. They found that the student passed the final oral examination, and that the dissertation was satisfactory for a doctoral degree and ready for any final modifications that they explicitly required.

William B. Knowlton, Ph.D.	Chair, Supervisory Committee
R. Jacob Baker, Ph.D.	Member, Supervisory Committee
Wan Kuang, Ph.D.	Member, Supervisory Committee
Maria Mitkova, Ph.D.	Member, Supervisory Committee
Eric M. Vogel, Ph.D.	External Examiner

The final reading approval of the dissertation was granted by William B. Knowlton, Ph.D., Chair of the Supervisory Committee. The dissertation was approved for the Graduate College by John R. Pelton, Ph.D., Dean of the Graduate College.

ACKNOWLEDGEMENTS

The completion of this dissertation would not be possible without the encouragement, support, and help of many individuals. I would like to thank my advisor Dr. William Knowlton for taking me on as an undergraduate in his research group. Dr. Knowlton graciously allowed me to continue my graduate studies working in his lab offering countless avenues to learn and grow. I have had the privilege of benefiting from Dr. Knowlton's mentorship and his encouragement has been valued and edifying during this undertaking. My appreciation is extended to Dr. Gennadi Bersuker, who has provided samples for this study, been very supportive of this work, and provided excellent discussions and feedback. I am grateful to Dr. Wan Kuang, Dr. R. Jacob Baker, Dr. Maria Mitkova, Dr. Elton Graugnard, Dr. Bernard Yurke, Dr. William Hughes, and Dr. Eric Vogel for all the support amidst the many demands on their time.

I am thankful to the many fellow students that I have had the privilege to work with. In particular, I would like to thank Ross Butler and Justin Reed. I am also grateful for Hieu Bui, Chris Buu, Shem Purnell, Blake Rapp, and Ryan Thompson who have helped with the measurement process.

To my family, whom I cherish the most, your support is most needed and appreciated. In addition, I recognize and honor the Almighty who sustains us all.

ABSTRACT

Hafnium oxide (HfO_2) is replacing silicon dioxide (SiO_2) as the gate dielectric in metal oxide semiconductor (MOS) structures driven mainly by need to reduce high leakage currents observed in sub-2nm SiO_2. The high dielectric constant of HfO_2 (~25) compared to SiO_2 (3.9 bulk) allows a thicker HfO_2 layer to be used in place of the thinner SiO_2 layer thereby reducing the gate leakage current in MOS devices while maintaining the same capacitive coupling provided by the thinner SiO_2. However, incorporating HfO_2 into MOS devices produces a SiO_2 interfacial layer between the Si substrate and HfO_2 interface. The increased complexity of the multilayer dielectric gate stack and introduction of new materials requires knowledge of the carrier transport mechanisms for accurate modeling and process improvement.

A large temperature dependence of the leakage current in HfO_2 gate dielectrics are observed compared to SiO_2, indicating temperature dependent leakage current measurements maybe well suited to understand the transport mechanism of HfO_2-based gate dielectrics. The leakage currents are measured for two different titanium nitride (TiN) metal gate stacks composed of either 3nm or 5nm HfO_2 on 1.1nm SiO_2 interfacial layer over temperatures ranging from 6K to 400K. For gate biases that yield equivalent electron energy barriers for the 3nm and 5nm HfO_2 gate stacks, the 5nm stack shows orders of magnitude less current and an order of magnitude larger increase in the gate leakage current with respect to temperature from 5.6K to 400K.

Knowledge of the energy band structure is crucial in determining what carrier transport mechanisms are plausible in multilayer dielectric stacks. Important parameters, necessary for modeling different transport mechanisms, can be extracted from accurately constructed energy band diagrams such as electric fields and barrier heights. An existing program developed by the author is further modified to incorporate image charge effects, multilayer dielectrics, and transmission coefficient calculations for use in this study.

Results indicate that the widely used Poole-Frenkel and Schottky conduction mechanisms for HfO_2 dielectrics can only explain a narrow electric field and temperature range and fail to explain the observed thickness dependence. Modeling the temperature dependence of 3nm and 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs with a combination of a temperature independent term, variable range hopping conduction, and Arrhenius expression (e.g., nearest neighbor hopping) describes the entire measured temperature range (6K to 400K). Additionally, HfO_2 defect densities can be extracted using the proposed model and provide densities in the range of ~10^{19} to ~10^{21} $cm^{-3}eV^{-1}$, which correlate well with defect densities reported in the literature. Defects in the HfO_2 are likely a result of oxygen vacancies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Hafnium oxide (HfO_2) is replacing silicon dioxide (SiO_2) as the gate dielectric of choice in metal oxide semiconductor (MOS) structures [1] due to the high dielectric constant (high-*k*) value of HfO_2 (~25) compared to SiO_2 (3.9 bulk). The dielectric constant of HfO_2 has reduced the gate leakage current in MOS devices by replacing thinner $SiO_2$with thicker HfO_2 while retaining desired device performance. However, the growth of HfO_2 on silicon (Si) generates a sub-nanometer thick SiO_2 interfacial layer (IL) leading to a dielectric bilayer (i.e., HfO_2-IL-Si) [2] that increases 1) the complexity of the energy bands, and 2) the electrically active point defect concentration as compared to SiO_2 MOS devices. Due to the increased complexity of MOS devices incorporating HfO_2, new questions arise. One of these fundamental questions is: what are the carrier transport mechanisms through the gate stack? This question does not only apply to HfO_2 but also to other dielectric materials that are quickly being integrated into other novel devices, such as multilayer dielectrics [3] and non-volatile memory [4-6]. The answer to this question could result in lower gate leakage current, providing: cooler devices, lower power consumption, longer battery life, and increased portability.

For the case of transport properties in HfO_2, a strong temperature dependence has been observed [7-9]. Figure 1.1 shows the temperature dependence of a HfO_2 gate stack (red data) compared to that composed of only SiO_2 (black data) as the gate dielectric [10]. Increased temperature dependence of the gate leakage current in the HfO_2 device is observed as compared to the SiO_2 device (Figure 1.1). Comprehending the conduction

mechanism in high-*k* materials has become of interest to the scientific community not only to understand the materials performance, but to understand the underlying physics and thus the material's limitations [9, 11-13]. Conduction or transport mechanisms (i.e., the kinetics) in materials are generally thermally activated, thus temperature dependent studies remain a key tool to understanding transport mechanisms.

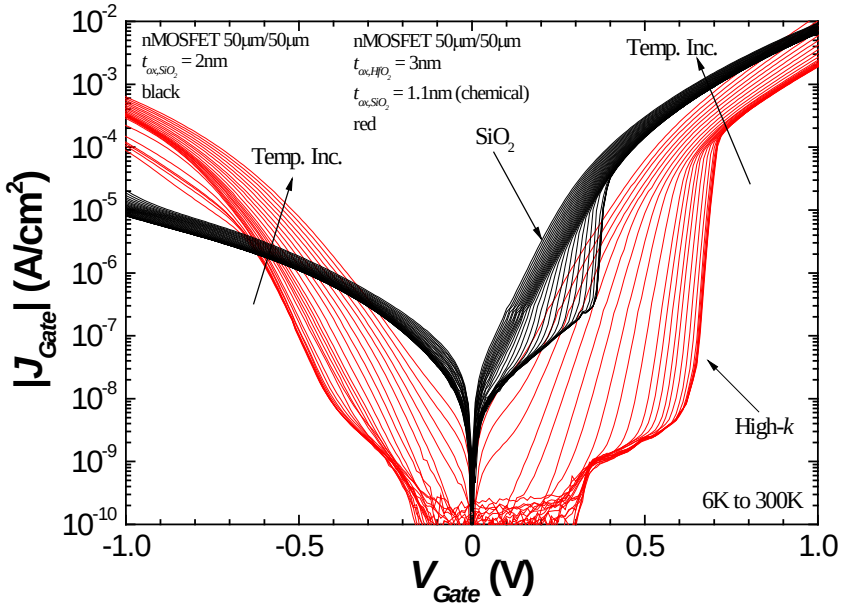


Figure 1.1: Gate leakage current for nMOSFETs composed of a HfO_2/SiO_2 gate dielectric stack and a SiO_2 gate dielectric for various temperatures ranging from 6 - 300K. The HfO_2/SiO_2 gate dielectric shows larger temperature dependence (red) than SiO_2 (black) at both positive and negative gate biases. Similar to [10].

For the general case of current transport in high-*k* dielectrics, many temperature dependent studies have been performed on silicon nitride [14, 15], titanium oxide [16], aluminum oxide [17], europium oxide [18], zirconium oxide [19], hafnium oxide [9, 11, 20-22], and others to determine the transport mechanism. Traditional temperature dependent transport mechanisms in high-*k* dielectrics have focused primarily on Poole-Frenkel emission [23-26], Schottky emission [7, 23], ohmic, and ionic conduction [27]. Since high permittivity materials are being used as a gate dielectric replacement material,

it has been suggested by theory that polarons, electron-phonon interactions, could play a significant role in current conduction [28, 29]. As electron-phonon interactions are thermally activated, temperature dependent measurements are appropriate for examining potential polaron mediated transport. For more information regarding polarons, see Appendix A.

Despite the large amount of temperature dependent work performed to understand carrier transport in HfO_2 (the high-*k* material that has received the most attention), the majority of the work has focused on above room temperature measurements with few studies analyzing transport below room temperature and even fewer at cryogenic temperatures below 77K. This leaves a significant portion of temperatures from near 0K to room temperature that has yet to be thoroughly examined and utilized in determining the carrier transport mechanism. Measurements performed near 0K have the added benefit of reduced thermal smearing of the Fermi energy level (E_f) relative to room temperature or even 77K, providing an appreciably sharper energy probe (Figure 1.2).

For work performed on HfO_2/SiO_2 gate dielectrics, many transport mechanisms have been suggested and include: Poole-Frenkel transport [7, 30-32], Schottky emission [7, 33, 34], polaron transport [35-37], quantum mechanical tunneling [30], and trap assisted tunneling [30, 38] among others. These studies highlight the lack of consensus in the field. It is evident that a comprehensive study has yet to be performed that has identified carrier transport regimes over large temperature and voltage regimes. The aim of this study is to identify carrier transport mechanisms over a broad temperature and voltage regime.

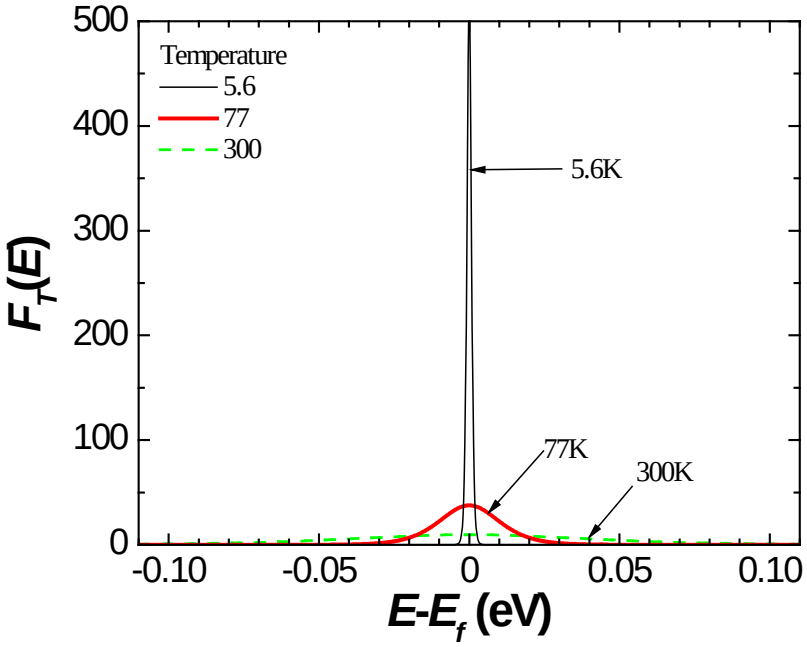


Figure 1.2: Thermal broadening function [39], $F_T(E) = \partial_E f(E)$, as a function of E-E_f at 5.6K, 77K, and 300K. The energy distribution at 5.6K is significantly narrower allowing the E_f position to be known much more accurately.

Employing a variable temperature probe station, the carrier transport in 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 MOS dielectric stacks from 5.6-400K is investigated. Results reveal regimes where carrier transport is dominated by either carrier-limited transport or conduction path-limited transport. For conduction path-limited transport, an approach suggested by the most common mechanisms used in the literature to explain the temperature dependence of the gate leakage current is taken to analyze the wide temperature and voltage range data measured in this study. First Poole-Frenkel analysis is performed followed by Schottky analysis. Results indicate that Poole-Frenkel transport is capable of describing the data in a limited voltage and temperature range and Schottky emission does not describe the temperature and voltage dependence of the gate leakage current. Using mechanisms suggested by temperature dependent carrier transport studies of disordered materials by the physics community, hopping

analysis is performed. Results indicate that the full temperature and voltage regime for conduction path-limited transport can be described using a variable range hopping term, Arrhenius term, and temperature independent term. Inconsistencies in the fit parameters of the variable range hopping term and the Arrhenius term are identified and proposed as further areas of research.

This study is presented in nine chapters. In the following chapter, an introduction to various transport mechanisms is presented. Chapter 3 details the instrumentation, software, test devices, and experimental procedures used to measure and analyze carrier transport. Chapter 4 describes the development of an energy band simulation tool useful in analyzing multilayer dielectric stacks in metal-insulator-metal and MOS stacks. Chapter 5 gives experimental results and observations to the behavior of the temperature dependence of the gate leakage current. Chapters 6, 7, and 8 explore three transport mechanisms (Poole-Frenkel, Schottky, and hopping conduction) as possible explanations of the gate leakage current and its temperature dependence. Finally, Chapter 9 summarizes important points and offers some directions for future work.

CHAPTER 2: TRANSPORT MECHANISMS

There are many carrier transport mechanisms all of which are influenced by temperature to some degree. The following subsections address the most common transport mechanisms.

2.1: Direct Tunneling (DT) and Fowler-Nordheim Tunneling (FNT)

The most basic carrier transport mechanisms, which account for the majority of the leakage current in SiO_2 based dielectrics, are purely quantum mechanical in nature and include direct tunneling (DT) and Fowler-Nordheim tunneling (FNT). Figure 1.1 shows the energy band diagrams for a SiO_2 MOSFET at flat-band (a), in DT (b), and in FNT (c).

Tunneling through the full width of the dielectric (i.e., trapezoidal barrier) is DT (Figure 2.1b), while tunneling through a triangular barrier is FNT (Figure 2.1c). Although quantum mechanical tunneling is not temperature dependent to first order, higher order temperature dependent factors do exist. These parameters include 1) the effective mass in the gate electrode, dielectric, and semiconductor; 2) the number of minority and majority carriers in the semiconductor; and 3) the energy of the minority and majority carriers. These parameters result in weakly temperature dependent carrier transport via quantum mechanical tunneling.

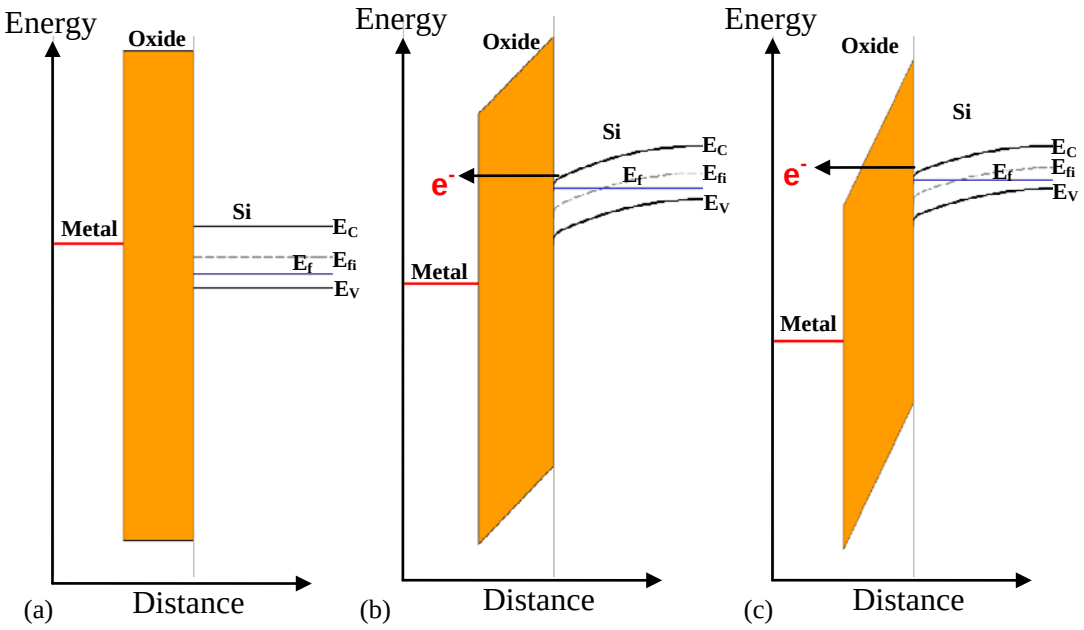


Figure 2.1: Energy band diagram showing: (a) flat-band condition, (b) DT – tunneling through the gate dielectric takes place at the full width of the gate dielectric compared to (c) FNT – tunneling through a triangular barrier, which reduces the tunneling distance, created using [40].

For a single barrier, the mathematical description of FNT can be simplified as [41, 42]:

$$J_{FN} = B_{FN} E_{ox}^2 \exp\left(-\frac{C_{FN}}{E_{ox}}\right) \tag{2.1}$$

where

$$B_{FN} = \frac{q^3}{16\pi^2 \hbar \phi_B}\left(\frac{m^*}{m_e}\right),\ C_{FN} = \frac{4\sqrt{2m^* \phi_B^3}}{3q\hbar}, \tag{2.2}$$

and q, $\hbar$, ϕ_B, E_{ox}, and m^* are the charge of an electron, reduced Planck's constant, barrier height between the oxide and semiconductor, the electric field in the oxide, and the effective mass of the electron in the oxide. Examining the simplified FNT equation in

(2.1) and (2.2), no direct temperature term is observed. However, E_{ox}, m^*, and ϕ_B all are slightly temperature dependent.

2.2: Variable-Range and Nearest-Neighbor Hopping Conduction

If defects exist in the oxide, carriers can move from one defect to another using the defects as "stepping stones" to pass through the oxide. If a carrier tunnels through the oxide using just one defect "step", the transport mechanism is typically called trap assisted tunneling (Figure 2.2) and it is considered a two-step tunneling process [43].

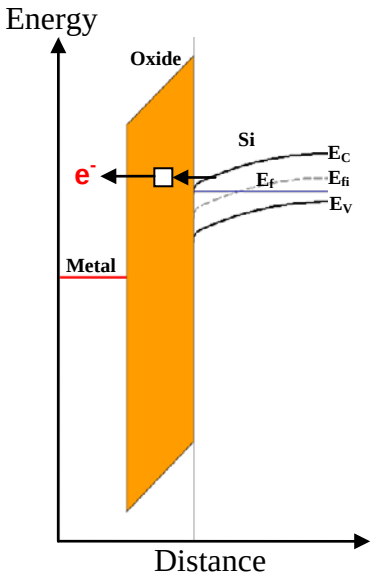


Figure 2.2: Energy band diagram [40] showing trap-assisted tunneling, using the presence of a defect in the oxide as a "stepping stone" for tunneling.

Conversely, if many defect traps exist, an impurity defect band is formed within the oxide and conduction takes place similar to transport in the conduction band (i.e., unbound or delocalized states) with an effective mass associated with the impurity band. Between these two carrier transport extremes exists the regime known as hopping conduction, which was first described by P.W. Anderson [44]. At electronic energy states

near the Fermi energy, thermally activated electrons hop from one localized state to another. The probability of an electron hopping from a state below the Fermi energy to a state above the Fermi energy level depends on three factors [45]:

a. The Boltzmann's factor $\exp(-\Delta E_{trap}/k_bT)$, where ΔE_{trap} is the energy difference between the two states, k_b is Boltzmann's constant, and T is temperature.

b. The probability of absorbing a phonon, v_{ph}, of the proper energy, which depends on the phonon spectrum.

c. A probability factor that depends on the overlap of the wave functions of the two states, $\exp(-2\alpha R)$, where R is the distance between the states and α is the inverse decay length of the localized state.

The distance to which the carrier must hop increases as the temperature decreases. This conduction mechanism is also referred to as phonon-assisted hopping and the probability of hopping is given as [45, 46]:

$$v_{ph} \exp\left(-2\alpha R - \frac{\Delta E_{trap}}{k_b T}\right). \tag{2.3}$$

At high temperatures, the temperature term dominates and carriers jump to the nearest localized state since the $\exp(-2\alpha R)$ term falls off quickly. This type of hopping conduction is called nearest neighbor hopping (NNH) conduction [47, 48] (Figure 2.3a). As the temperature decreases, insufficient energy is present to cause a transition from a lower to a higher energy state. Known as variable range hopping (VRH) conduction, carriers can then only hop to energy states of similar energies (Figure 2.3b).

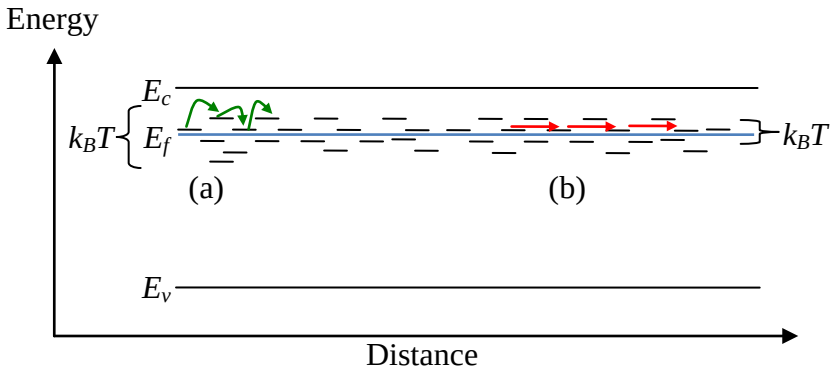


Figure 2.3: Energy band diagram[1] showing localized states around the Fermi energy, E_f. The energy of the conduction band, E_c, and the valance band energy level, E_v, are also shown. Nearest neighbor hopping conduction is shown in part (a) while variable range hopping conduction is shown in part (b).

The distance to which the carrier must hop increases as the temperature decreases. An increase in hopping distance as a function of decreasing temperature was first identified by Mott [47, 48]. Mott's treatment of identifying the optimal hopping distance resulted in a $T^{-1/4}$ dependence [45]. Many other approaches have been considered in modeling hopping conduction, including the application of percolation theory and for the three dimensional case all predict a $T^{-1/4}$ temperature dependence [45, 49]. It has been suggested that an arbitrary *d*-dimensional case for variable range hopping conductivity exists and is given by [46]:

$$\sigma = \sigma_{hop} \exp\left[-\left(\frac{T_0}{T}\right)^{\frac{1}{d+1}}\right] \tag{2.4}$$

where T_0 , the characteristic temperature, is given by:

[1] In crystalline materials, the density of states for valence electrons (valance band) and excited electrons (conduction band) are sharply defined. The energy difference between the top of the valance band and the bottom of the conduction band is defined as the band gap. In highly disordered semiconductors and insulators (i.e., amorphous or poly-crystalline), the density of states for valence electrons and excited electrons are not well defined (due to disorder in the bonding) and result in localized tail states. Rather than band edges and gaps, highly disordered materials use the analogy of mobility edges and a mobility gap. Figure 2.3, depicting hopping conduction, has been draw as a crystalline insulator for simplicity but is also applicable to highly disordered insulators.

$$T_0 = \frac{Ad}{2\alpha k_b N(E_F)}\left(2\alpha\left(\frac{d+1}{d}\right)\right)^{d+1} \tag{2.5}$$

and A and σ_{hop} are constants with $N(E_F)$ the density of states about the Fermi energy. For two dimensional hopping conduction, the temperature dependence would be $T^{1/3}$.

Hopping conduction can occur in both crystalline materials (as shown in Figure 2.3 for simplicity) and highly disordered materials. In crystalline materials, defects are introduced into the band gap by impurities or structural defects in the lattice. In addition to impurities, amorphous materials have defects in the mobility gap (analogous to band gap in crystalline materials), due to the random orientation of the atoms and unsatisfied bonds. Much of the work leading to descriptions of variable range hopping conduction and nearest neighbor hopping conduction was done on amorphous materials. Variable range hopping and nearest neighbor hopping are especially applicable to the system of dielectrics studied here (SiO_2 and HfO_2) as they are both disordered systems.

2.3: Poole-Frenkel Conduction

Poole-Frenkel (P-F) conduction or emission [24-26] acts to modify the potential well of trapped carriers. This modification is a result of high electric fields, which lowers the ionization energy [45] or the energy for thermal stimulation of electrons into the conduction band [27]. Frenkel, building on the previous work of Poole [24-26], showed that the trap barrier height (ϕ_{PF}) is lowered by [50]:

$$\Delta\phi_{PF} = \beta_{PF} E_{ox}^{1/2} \tag{2.6}$$

where

$$\beta_{PF} = \left(\frac{q^3}{\pi \varepsilon_0 \varepsilon_r} \right)^{1/2}, \tag{2.7}$$

ε_0 is the permittivity of free space, and ε_r and is the high frequency (i.e., dynamic) dielectric constant [51, 52]. Figure 2.4 shows a diagram illustrating the reduction in the potential well of a trap in the presence of an electric field.

The derivation of the energy potential lowering effect (2.6) depends on the hydrogenic potential. That is, the standard P-F mechanism incorporates the concept of a hydrogenic impurity for which the ionization energy potential is determined using the effective mass approximation [53-55]. The hydrogenic impurity includes both a charged ion impurity and a charged trapped carrier, which interact with each other. In his model, Frenkel only considered electrons and donor traps [24]. However, in the effective mass approximation, the difference between the hydrogenic ionization energy potential for an electron and hole is simply the effective mass of the respective carriers. Hence, there is no reason why the P-F model cannot be applied to holes as it is applied to electrons. A widely accepted mathematical expression for the standard P-F conduction (J_{PF}) can be written as:

$$J_{PF} = CE_{ox} e^{-\frac{q\phi_{PF} - \beta_{PF}\sqrt{E_{ox}}}{\xi k_b T}}, \tag{2.8}$$

where C, k_b, T, and ξ, are a constant, Boltzmann's constant, temperature, and a factor that depends on acceptor compensation [52, 56], respectively. The value of ξ is usually between 1 and 2. In Frenkel's original paper, $\xi = 2$. In much of the literature, $\xi = 1$ (e.g., [20, 23, 33, 38, 41, 57-59]), which corresponds to heavily compensated traps [52].

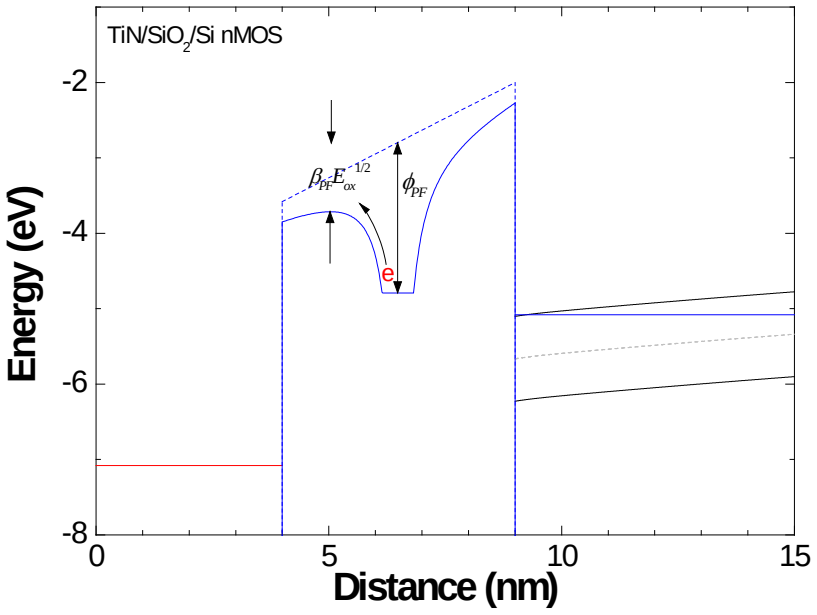


Figure 2.4: Energy band diagram showing Poole-Frenkel emission. As the electric field in the dielectric increases, the trap barrier is lowered. As the temperature increases, the trapped carrier has more energy and can more easily overcome the barrier. Created using [40].

The potential lowering effect of Poole-Frenkel emission has been applied to variable range hopping conduction by adding the $E_{ox}^{1/2}$ dependence in the exponential [45, 46]. Poole-Frenkel conduction dominates at high temperatures, where enough thermal energy is available to stimulate the carrier over the barrier, and high fields, where the barrier is significantly lowered [27]. At extremely high electric fields, the electron can hop downward without thermal activation yielding [45, 60]:

$$\sigma \propto \sigma_{hop2} \exp\left(-B / E_{ox}^{1/4}\right) \tag{2.9}$$

where σ_{hop2} and B are constants.

2.4: Schottky Emission

Emission of the carrier over the dielectric barrier (ϕ_B) and into the dielectric conduction band is known as thermionic emission. The barrier the charged carrier overcomes is lowered in the presence of an electric field due to the image charge theorem (discussed in Section 4.1.4). Thermionic emission over an electric field assisted lower barrier is known as Schottky emission. Figure 2.5 shows an energy band diagram of Schottky emission in a MOS gate stack composed of a single dielectric. The dashed line in Figure 2.5 represents the barrier to electrons without considering image charge effects while the solid line considers image charge effects due to the metal gate.

Schottky emission can be modeled using the following equation [27, 41]:

$$J_{SE} = A^* T^2 \exp\left(\frac{-q(\phi_B - \sqrt{qE_{ox}/4\pi\varepsilon_0\varepsilon_r})}{k_b T}\right) \tag{2.10}$$

where ε_r is the high frequency dielectric constant and A^* is the effective Richardson constant given by [41]:

$$A^* = \frac{4\pi q m^* k_b^{\,2}}{h^3} \tag{2.11}.$$

The parameters h and ϕ_B are Planck's constant and the barrier height between the semiconductor conduction band and the dielectric conduction band, respectively.

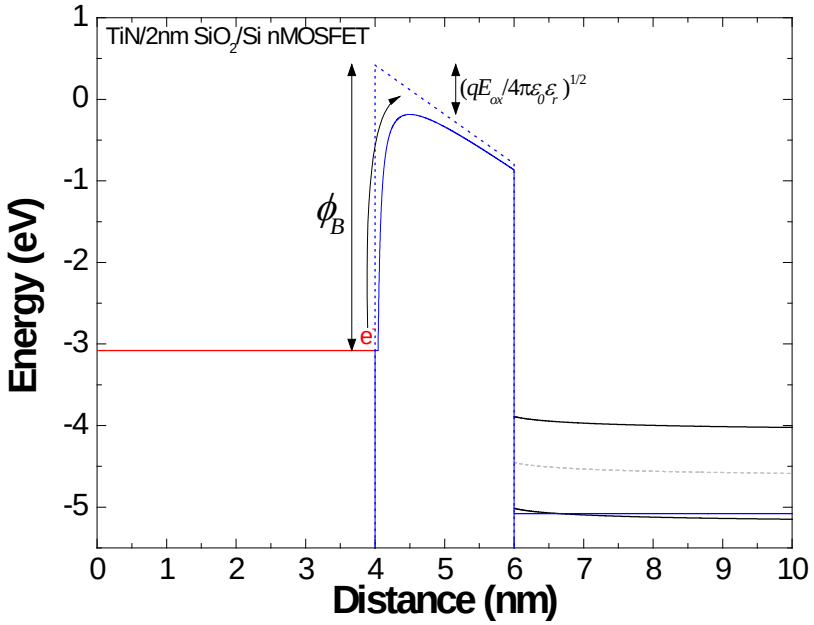


Figure 2.5: Energy band diagram depicting Schottky emission of an electron moving from the metal gate to the SiO_2 conduction band.[2] The Schottky emission process is similar to thermionic emission with the charged carrier barrier lowered by the presence of an electric field and the image charge theorem. Created using [40].

2.5: Determining the Conduction Mechanism

Determining which conduction mechanism is present in a dielectric is primarily determined by fitting experimental data to the various conduction mechanism equations shown above. Establishing which transport mechanisms are present is difficult as several transport mechanisms can occur simultaneously, which can distort experimental results. To deconvolute multiple transport mechanisms, it becomes necessary to characterize the dielectric under a wide variety of conditions where one transport mechanism can dominate over the others thereby simplifying the identification of the dominant mechanism.

[2] As will be shown later in Section 4.1.4 and Chapter 7, the effects of image charge from both dielectric interfaces (not shown in Figure 2.5) should be considered in ultra thin dielectrics.

The simplest method for measuring the conduction current is via a current-voltage (I-V) sweep. There is a wide array of equipment available to perform such a measurement and the measurement is relatively quick. Many of the conduction mechanisms discussed have a voltage or electric field dependence. These transport mechanisms include: FNT, Poole-Frenkel, and Schottky conduction. Understanding the behavior of the conduction current as a function of voltage is crucial to identifying which transport mechanism is responsible. Many transport mechanisms dominate under certain voltage conditions. For example, FNT dominates DT at high electric fields.

In addition to a voltage dependence, most of the transport mechanisms discussed have a temperature dependence (i.e., T in (2.3), (2.4), and (2.10)). Hence, most carrier transport studies are performed at room temperatures and above (e.g., [7, 22, 23, 61]). This study focuses on deconvoluting multiple transport mechanisms in a complex dielectric system by incorporating 1) a temperature range from cryogenic to above room temperature, and 2) a broad range of voltages.

CHAPTER 3: EXPERIMENTAL PROCEDURE

In this section, the software and instrumentation used in this study are described. Briefly discussed are: 1) an analytical software package in which significant scripts were coded to analyze large experimental data sets, and 2) the development of a one-dimensional multilayer: metal/oxide/semiconductor simulation software to understand the underlying physics of multilayer dielectric MOS structures. The latter will be described in greater detail in Chapter 4. Additionally, various hardware components used are presented as well as the test devices. Finally, the experimental procedure is discussed.

3.1: Software

Analysis, manipulation, and visualization of the experimental results primarily occurred using OriginPro 8.0[3]. Custom fitting routines using OriginPro's non-linear curve fitting toolbox were written to compare experimental results to various transport mechanisms described in Chapter 2. Custom data manipulation code was also written to minimize human error and increase the amount of data that can be examined. The amount of custom code written, leveraging the power of OriginPro, results in over 14,000 lines of code, 80% of which has been written by the author.

A primary means of understanding and explaining the various transport mechanisms that can occur in advanced devices is facilitated through energy band diagrams (e.g., Figure 2.1 - Figure 2.5). Calculating the energy band diagram as a result

[3] OriginPro is a registered trademark of OriginLab.

of various gate voltages requires extensive knowledge of device physics. Even using simple analytical solutions for modeling MOS capacitor operation requires the simultaneous solution of many interdependent equations. The complexity of the energy band diagram increases as additional dielectric layers are added to the gate stack of MOS devices. To address these needs, a program was developed capable of constructing a gate stack with multiple dielectric layers and identifying 1) tunneling mechanisms, 2) voltage drops, 3) electric fields, 4) charge distributions, 5) potential, 6) image charge effects, and 7) quantum mechanical tunneling transmission coefficients among other physical parameters (discussed in Chapter 4), expanding on a previous version of the program also written by the author [40].

3.2: Instrumentation

3.2.1: Electrical Measurement Instrumentation

Electrical characterization of the high-*k* gate MOSFETs were performed with the Keithley 4200 Semiconductor Characterization System (SCS) shown in Figure 3.1. The Keithley 4200SCS is a precision DC voltage-current source and measurement unit capable of either 1) sourcing voltage and measuring current, or 2) sourcing current and measuring voltage. Four medium power source monitoring units (SMU) are included in the Keithley 4200 capable of sourcing up to 20V with 100mA of current. Four pre-amps are attached to the SMUs for additional current sensitivity down to 100aA resolution with 10fA accuracy. The four SMUs allow all four MOSFET terminals (gate, drain, source, well/substrate) to be measured simultaneously. For capacitance-voltage (C-V) measurements, the Agilent 4284 LCR meter was used.

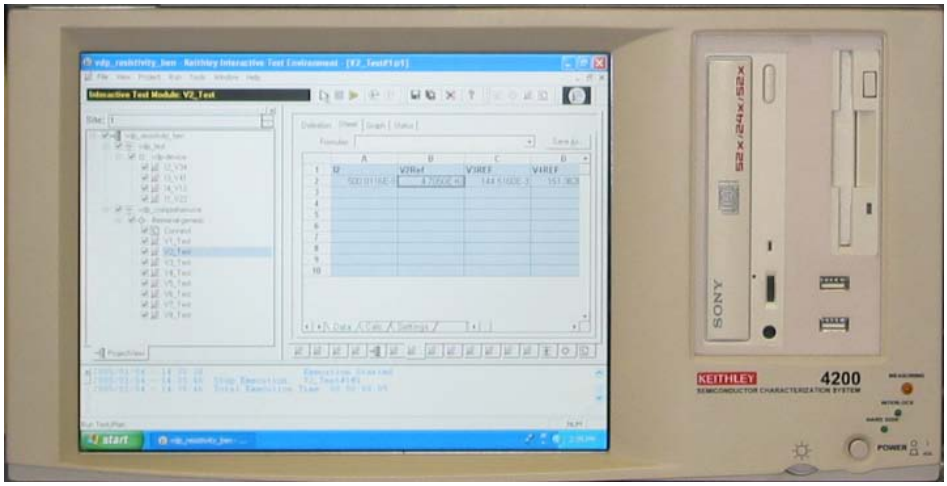


Figure 3.1: Keithley 4200 Semiconductor Characterization System used to measure device characteristics in this study.

3.2.2: Probe Station

A Janis Research custom built variable temperature probe station (5.6-450K) with four actively cooled (to limit heat injection into the devices) triaxial Kelvin probes was used for all measurements. The probe station is designed for low current sensitivity with a noise floor ~10fA. The low noise floor of the station is ideally suited for measuring the low current levels of gate leakage currents. The excellent DC characteristics have trade-offs with the AC characteristics of the system, which has a 3dB roll-off at ~10MHz. Experiments were performed under vacuum, necessary to reach the temperatures studied here, and under light tight conditions. The variable temperature probe station with supporting electrical instrumentation, including the Keithley 4200SCS, is shown in Figure 3.2. The small 5 cm sample chuck (Figure 3.2 inset) required that the wafer containing the MOS devices to be cleaved. Select devices were tested on a room temperature probe station (Micromanipulater Model 6200) before and after wafer cleaving to verify that the cleaving process did not damage the devices. Grounding mats

and electro-static discharge (ESD) boots straps were worn to protect the devices from ESD damage.

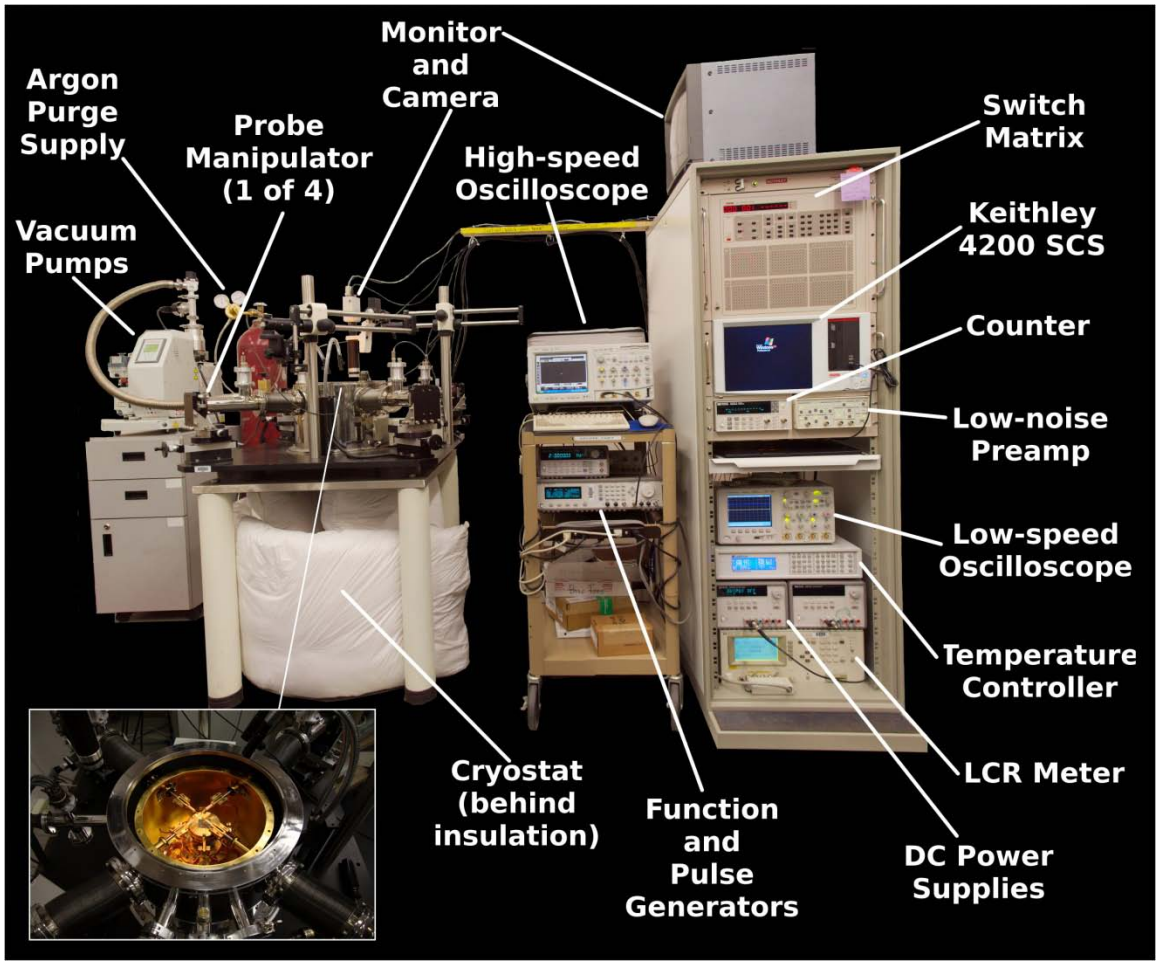


Figure 3.2: Variable temperature probe station and supporting equipment. Inset shows where the samples are placed and probed.

3.3: Devices

The MOSFET test devices used in this study were fabricated by SEMATECH[4] using a standard self-align process on epitaxial Si-substrate. A chemical SiO_2 IL was grown by treating Si with O_3. HfO_2 was deposited using atomic layer deposition for thicknesses of 3 and 5nm. Using chemical vapor deposition, 10nm of titanium nitride

[4]SEMATECH is an international semiconductor manufacturing consortium which performs advanced research and development on semiconductor devices based in Austin, TX, USA.

(TiN) was used as a metal gate capped by 150nm of amorphous Si. Source/drain junctions were activated using temperatures of 1000ºC for 10 seconds followed by a post metal forming gas anneal for 30 minutes at 480ºC [62].

A SiO_2 control wafer was also fabricated using the exact same wafer processing described above with the exception of the dielectric stack, which is composed of a 2.0nm thermal SiO_2 *in situ* steam generated oxide at 950ºC. The control wafer was used to establish gate leakage current comparisons between high-*k* and SiO_2, as well as associated temperature dependencies. Table 3.1 shows wafer split information.

Table 3.1: Wafer Split Information

Split	Interfacial Layer Material	High-*k* Material	Gate Material
1	1.1nm chemical SiO_2	3nm HfO_2	TiN
2	1.1nm chemical SiO_2	5nm HfO_2	TiN
3	2.0nm thermal SiO_2	-	TiN

In addition to wafer fabrication, SEMATECH has performed some characterization tests to quantify the dielectric constant of the HfO_2 and IL using high resolution-electron transmission microscopy (HR-TEM) micrographs and C-V tests. Flatband and EOT were extracted using the North Carolina State University (NCSU) methodology [63]. The data (shown in [64]) gives the extracted EOT from the capacitance-voltage tests for each HfO_2 layer thickness. The equivalent oxide thickness (EOT)[5] of the dual layer dielectric can be described by:

$$EOT_{HfO_2/SiO_2} = \frac{3.9}{\varepsilon_{r,HfO_2}} t_{ox,HfO_2} + EOT_{SiO_2} \tag{3.1}$$

[5] Equivalent oxide thickness, or EOT, is a normalized thickness value based on the capacitance of the dielectric(s). The EOT normalizes the thickness of the dielectric(s) to that of bulk SiO_2, which has a relative dielectric constant of 3.9.

where ε_{r,HfO_2} is the relative dielectric constant of the HfO_2 high-*k* layer, t_{ox,HfO_2} is the physical thickness of the HfO_2 layer (extracted using HR-TEM), and EOT_{SiO_2} is the EOT of the SiO_2 IL. From (3.1), the y-axis intercept and the slope can be used to determine the EOT of the IL and the dielectric constant of the HfO_2, respectively. For the devices used in this study, a 0.7nm EOT of the IL and a dielectric constant of 20 for the HfO_2 layer were obtained. HR-TEM micrographs show the SiO_2 IL to be between 0.9nm and 1.1nm (Figure 3.3). Using the EOT of 0.7nm and the physical thickness of 1.1nm, a relative static dielectric constant (ε_{r,SiO_2}) of 6.13 is extracted using equation (3.2), where t_{ox} is the physical oxide thickness. The increased value of the static relative dielectric constant of the interfacial layer (6.13) compared to bulk (3.9) is attributed to oxygen deficiencies related to HfO_2 deposition and related processing [64]. Table 4.3 shows a compiled list of the different material parameters for each layer.

$$\varepsilon_r = \frac{3.9 t_{ox}}{EOT} \tag{3.2}$$

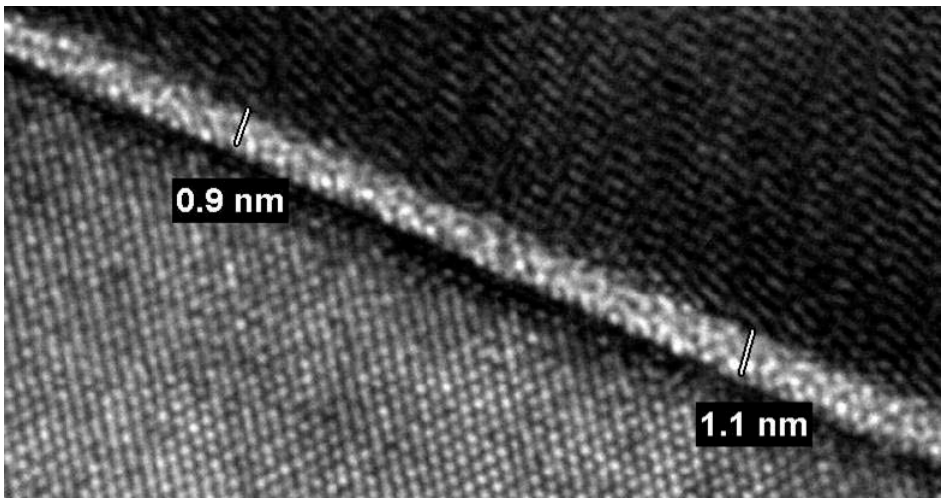


Figure 3.3: High resolution transmission electron microscope (HR-TEM) micrograph showing the interfacial layer thickness ranging from 0.9nm to 1.1nm. The darker top material is the HfO_2 high-*k* layer and the lower layer the crystalline Si substrate. HR-TEM micrograph courtesy of Dr. Du Li of Micron Technology.

Table 3.2: Materials Properties

	Gate (TiN)	HfO_2	SiO_2	Si
Work Function (eV)	4.45	-	-	-
Electron Affinity (eV)	-	2.65	0.95	4.15
Dielectric Constant	-	20	6.13	11.7
Band Gap (eV)	-	5.7	8.9	1.11
Thickness (nm)	-	3 and 5	1.1	-
Doping (cm^{-3})	-	-	-	$1x10^{18}$(n-type) $2x10^{15}$(p-type)

Although each wafer die contains many devices, the devices used in this study were large area devices in order to increase the signal to noise ratio of the gate leakage current to system noise, which is around 10fA. The largest n/pMOSFETs available for this study had widths and lengths (W/L) of 30μm/30μm and 50μm/50μm.

3.4: Measurement Procedure

As was discussed in Section 2.5, a simple method to study carrier transport mechanisms is to analyze the DC I-V characteristics of the gate leakage current at various temperatures. Developing analytical techniques for DC I-V characteristics of multi-layer stacks leverages existing common measurement techniques that are routinely performed for single layer dielectric gate stacks. DC I-V techniques are also well suited for use in the variable range probe station where the excellent low-noise DC characteristics can be utilized and the frequency response limitations of the system can be avoided.

For the I-V sweep of the gate leakage current (I_G-V_G), a technique called carrier separation [65-67] was used whereby all four MOSFET terminals are monitored, which allows the carrier type that is transported through the dielectric stack to be identified. To measure the inversion charge density, a slight bias on the drain current is applied during one of the tests. A 60s hold time was integrated into all tests to allow the device to reach steady state. A sweep delay of 1s was also used to allow the device to reach steady state

during the measurement sweep. Small sweep steps of 0.01V provide fine resolution of the gate leakage current with respect to the gate voltage, which is especially useful for large changes in the current with respect to voltage. Measurements were performed using the Keithely 4200's "*quiet*" setting to minimize noise. A single I-V sweep results in a test time of about 20mins. Table 3.3 describes the voltage range used and the resulting current measurement range for the various tests discussed. The voltage source range for all four terminals was set to "*best fixed*" where the Keithley 4200SCS determines what source range to use for the entire voltage sweep (i.e., the source range is not changed during the test).

Table 3.3: I-V sweep parameters

Test	V_{Gate}	V_{Drain}	$V_{Source,Sub/Well}$	I_{Gate}	$I_{Drain,Source,Sub/Well}$	Temperature (K)
1†	-1V*** to 1V 0.01V steps	0V	0V	Auto	Not Monitored	5.6, 7, 8, 9, 10, 15, 20 to 300 in 10K increments
2†	-1V*** to 1V 0.01V steps	$3K_bT$*	0V	100pA Limited Auto	100pA Limited Auto	5.6, 7, 8, 9, 10, 15, 20 to 300 in 10K increments
3‡	-2V to 2V 0.01V steps	0V	0V	Auto	100pA Limited Auto	6, 10, 15, 20 to 100 in 10K increments 100 to 400 in 20K increments
4‡	-1V to 2.5V** 0.01V steps	0V	0V	Auto	100pA Limited Auto	300

*pMOSFETs: bias is negative polarity. **for 5nm HfO_2 devices stop voltage was 3V. ***start voltage of high-*k* devices was -1.25V. †not performed on 5nm HfO_2 devices. ‡only performed on high-*k* devices

The temperatures at which measurements were performed are also shown in Table 3.3. Very fine temperature steps were used to accurately track the gate leakage current with respect to temperature. This study resulted in the one of the lowest temperature

studies conducted to date on HfO_2 and the most finely-spaced-temperature characterization of HfO_2 to date according to the author's knowledge. For the temperature range examined (5.6-400K), measurements began at the lowest temperature and increased until the entire temperature range of interest was covered. The station was allowed to sit for a minimum of 20 minutes between temperatures to allow the probe station to reach a thermal steady state.

A limited number of devices were used in taking measurements described in Table 3.3, they are:

a. for 2nm SiO_2, there were 4 devices: 2 nMOS and 2 pMOS,

b. for 3nm HfO_2, there were 7 devices: 3 nMOS and 4 pMOS, and

c. for 5nm HfO_2, there were 4 devices: 2 nMOS and 2 pMOS

A total 15 devices were used in creating the finely-spaced-in-temperature results discussed this study. Results correspond to trends observed in the temperature dependence of the gate leakage current in many other MOSFETs of the same gate stack, which were performed either using a larger temperature step size or differently sized MOSFETs.

Subsequent C-V measurements were also performed to determine the freeze-out behavior of carriers in n/pMOSFETs. A select number n/pMOSFETs composed of 3nm HfO_2/1.1nm SiO_2, 5nm HfO_2/1.1nm SiO_2, and 2nm SiO_2 devices were used to perform C-V using the Agilent 4284. For the C-V measurements, an AC probe frequency of 100kHz and amplitude of 45mV_{rms} was used. The C-V measurements were taken over a temperature range of 5.6K to 300K.

CHAPTER 4: ENERGY BAND SIMULATION TOOL

As is demonstrated in Figure 2.1-Figure 2.5, energy band diagrams are useful tools to explain solid state principles and to understand device behavior including metal-oxide-semiconductor (MOS) structures. Many device parameters can also be extracted from accurately constructed energy band diagrams. The parameters include the flatband voltage (V_{FB}), threshold voltage (V_t), and electric field. The latter is required for calculating the current from Fowler-Nordhiem tunneling [68, 69], Poole-Frenkel transport [70, 71], and Schottky emission [69, 72]. Accurate construction of energy band diagrams is time consuming and tedious, which can limit their usefulness. Quick hand sketches can lead to over exaggeration of features, which 1) limits analysis or 2) leads to incorrect assumptions. Accurate construction of energy band diagrams in modern MOS devices is further complicated by the recent introduction of multilayer dielectrics. MOS multilayer dielectric devices include high-*k* oxides in MOS devices [73] and trap-based flash nonvolatile memory devices (NVM) [74], such as SONOS [75] and TANOS [76], that incorporate trapped charge within the multilayer dielectric stack.

Device structure parameter calculations can be performed by hand, using spreadsheets, mathematical programs (e.g., Matlab, Mathematica) or with Poisson-Schrodinger equation solvers (PSES), which utilize finite-element approaches. Hand calculations can be quick and useful but typically employ simplifying assumptions, which limit their effectiveness. A quick visual method to understand these devices is not available. Although an improvement over hand calculations, employing spreadsheets or

mathematical programs for complex modeling can be challenging and time consuming. Assuming spreadsheets or mathematical programs accurately calculate device parameters for a MOS structure, changing the existing calculations from a single dielectric system to a more complex system (e.g., a three dielectric layer system) requires additional non-trivial programming. Conversely, extremely powerful finite element PSES are available to perform many device parameter calculations and to graphically represent a variety of device structures. The cost of software licenses and the effort required to learn the program interface is substantial. These barriers leave many without the ability to easily and quickly calculate the energy band diagram and associated parameters of multilayer dielectric device structures.

To address this dilemma, software was developed that quickly allows visualization of diagrams, including energy bands, potentials, and electric fields for dual layer high-*k* dielectrics on semiconductors with the ability to calculate and display voltages, electric fields, and capacitances as a function of applied voltages [77]. The approach is unique in that it does not rely on purely numerical techniques, such as finite-element PSES, but rather an iterative approach based on the underlying physics. The software is freely available to the public [78] and has received positive reviews from individual scientists and engineers at institutions and companies, including Penn State University, National Institute of Standards and Technology (NIST), Oregon State University, University of Florida, SEMATECH, Intel, Texas Instruments, ST Microelectronics, Freescale, IBM, and others. The author knows of no other software with equivalent capabilities, simplicity, speed, and cost (free).

A new version of the program is discussed here where the previous version developed by the author is expanded to support an unlimited number of dielectric layers with the added capability of adding metal layers between dielectric layers. The software is also expanded to display the potential through the dielectric stack, charge distributions, as well as the effects trap charge in the dielectric play on the energy band diagram and various device parameters. The improvements to the program minimize both the setup time and the learning curve compared to finite-element PSES. Parameter calculations and the representation of energy, electric field, or charge distributions are nearly instantaneous, and easier to initiate and execute than either spreadsheets, mathematical programs, or finite element PSES approaches.

4.1: Model

Accurate construction of an energy band diagram requires solving Poisson's equation:

$$\frac{d^2\phi}{dx^2} = -\frac{dE_{Field}}{dx} = -\frac{\rho}{\varepsilon}, \tag{4.1}$$

where x is distance, ϕ is potential, E_{Field} is electric field, ρ is charge density, and ε is the permittivity. The simulation program uses a modular approach to calculate the charge, electric field, potential, and energy of a structure composed of metals, dielectrics, and a semiconductor. In this approach, Poisson's equation, (4.1), is solved for each material independently when given a starting charge and potential. Using the individual solutions to (4.1), the entire structure is then solved. Due to the modular approach, an arbitrary number of materials may be used in any combination limited by the following rules.

Metals may be placed at any location in the gate stack as long as they are separated by dielectrics. The structure must have a metal as a starting material and end with either a metal or semiconductor material. A semiconductor can only be placed as an end material since an infinite thickness is assumed when solving the model. These few constraints leave a large number of structures that can be analyzed in one-dimension using the simulation tool, making it extremely useful in analyzing multilayer novel dielectrics, which are being used in advanced metal-insulator-semiconductor devices, such as trap-based non-volatile memory.

The next three subsections describe the model used for each material type in determining the solution to the entire structure. Following these models, additions to the simulation tool are described.

4.1.1: Metal Model

Metals in the model are assumed to be perfect conductors. This implies the electric field in the metal is zero or:

$$E_{metal} = 0, \tag{4.2}$$

which also implies that the potential drop in the metal is also zero or:

$$\Delta\phi_{metal} = 0. \tag{4.3}$$

Hence, the potential in the metal is equal to the potential in the previous material at the boundary. For the first metal in the structure (i.e., considered at the top of a stack of multilayers of materials), the starting charge is applied as an infinitely thin line charge to the bottom of the metal. For a metal inside the structure, the negative of the sum of the charges above the metal (ρ_{total}) are placed on the top of the metal ($\rho_{metal,top}$) as an infinitely thin line charge (due to charge neutrality), and the sum of the charges plus any trapped

charge assigned to the metal ($\rho_{metal,trapped}$) are placed on the bottom part of the metal as an infinitely thin line charge. These relationships are described by the following equations:

$$\rho_{metal,top} = -\rho_{total} \quad \text{and} \tag{4.4}$$

$$\rho_{metal,bottom} = \rho_{total} + \rho_{metal,trapped}\,. \tag{4.5}$$

For a metal that is placed as an end material (the bottom of the stack of materials), the sum of the charges above the metal are placed on the top of the metal as an infinitely thin line charge. Regardless of the thickness of the metal material, charge in the metal is modeled using infinitely thin line charges and does not consider quantum mechanical effects.

4.1.2: Dielectric Model

For each dielectric, an arbitrary number of line charges can be inserted within the material. Following Poisson's equation, the electric field (E) in the dielectric at a given line charge (j) is given by:

$$E_j = \frac{1}{\varepsilon}\left(\rho_{total} + \sum_{0}^{j}\rho_j\right), \tag{4.6}$$

where ρ_{total} is the sum of the charge in the structure up to the interface of the dielectric of interest and ρ_j is the jth line charge in the said dielectric. After the electric field is calculated, the potential (ϕ) can then be calculated at a given line charge by:

$$\phi_j = \phi_{boundary} + \sum_{1}^{j} -E_{j-1}\left(x_j - x_{j-1}\right), \tag{4.7}$$

where $\phi_{boundary}$ is the potential at the boundary of the dielectric of interest and x_j is the location of the *j*th line charge. Using the potential, the energy band is easily calculated by applying the appropriate electron affinities and band gaps.

Equations (4.6) and (4.7) are applicable at any location inside the dielectric and are not limited to a specified line charge.

4.1.3: Semiconductor Model

As mentioned above, a semiconductor can be placed at the bottom of the stack and the equations used for solving the charge, electric field, and potential of the semiconductor are taken from [79]. Using [79], the potential drop in the semiconductor (surface potential, ψ_s) for a given charge is given by:

$$Q'_C = \mp\sqrt{2q\varepsilon_s N_A}\sqrt{\phi_t e^{-\psi_s/\phi_t} + \psi_s - \phi_t + e^{-2\phi_F/\phi_t}\left(\phi_t e^{\psi_s/\phi_t} - \psi_s - \phi_t\right)} \tag{4.8}$$

where Q'_c, q, N_A, ε_s, ϕ_t, and ϕ_F are the semiconductor charge per unit area, electron charge, doping concentration, semiconductor dielectric constant, thermal voltage, and semiconductor body doping potential, respectively. To obtain the potential drop in the semiconductor as a function of distance (x), the integral of the inverse of the electric field yields the distance as follows [79],

$$\int_{\psi(x)}^{\psi_s} \frac{1}{E(\hat{\psi})} d\hat{\psi} = x - x_{surface} \tag{4.9}$$

,

where $E(\psi)$ is the electric field of the semiconductor as a function of potential and is calculated by [79]:

$$E(\psi) = \pm\frac{\sqrt{2q\varepsilon_s N_A}}{\kappa_s}\sqrt{\phi_t e^{-\psi(x)/\phi_t} + \psi(x) - \phi_t + e^{-2\phi_F/\phi_t}\left(\phi_t e^{\psi(x)/\phi_t} - \psi(x) - \phi_t\right)}. \tag{4.10}$$

Using (4.8)-(4.10), the potential, charge, and electric field in the semiconductor can be calculated. The potential of the entire stack is solved by choosing a starting gate charge and determining if the resultant potential matches the desired input voltage. A solution is obtained using this iterative approach. The result of the analytical algorithm built into the program and the assumptions described in Table 4.1 provides quick calculation speeds.

Table 4.1: Simulation Tool Assumptions

Fermi-Dirac approximated with Maxwell-Boltzmann statistics
Semiconductor doping >> intrinsic carrier concentration
Complete ionization of dopant atoms (Temperature > 100K)
Non-degeneracy in the semiconductor
Uniformly doped Si
Charge sheet approximation
Metal gate
Infinitely thin line charges
Quantum mechanical effects not considered
Device in equilibrium/no current flow

4.1.4: Image Charge Model

A charged particle at distance x from a perfect conductor plate will have its electric field lines terminate perpendicular to the interface between a conductor and nonconductor plate as shown in Figure 4.1. Opposite charged particles at a distance of $2x$ will have their electric fields perpendicular to an imaginary plane exactly half way (or distance x) and normal to the two charged particles. The force between a charged particle and conductor plate at a distance x from one another is therefore equal to the force between two oppositely charged particles at distance $2x$; this is known as the *image charge theorem*.

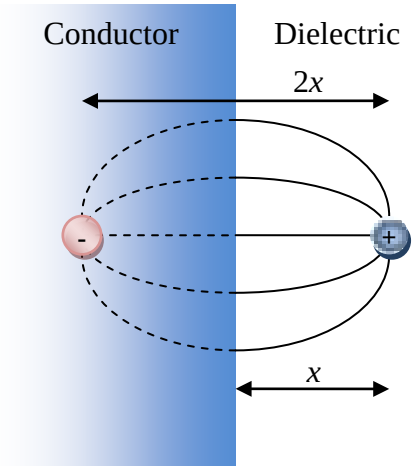


Figure 4.1: Diagram demonstrating the image charge theorem. A charged particle at distance x from a conductor experiences a force equal to the force between two opposite charged particles at distance $2x$.

The potential energy for two oppositely charged particles with electric charge q (the charge of an electron) at a distance $2x$ in units of eV is given by Coulomb's potential:

$$PE(x) = \frac{-q}{16\pi\varepsilon x} \tag{4.11}$$

where q is the charge of an electron and ε is the high frequency dielectric constant of the medium surrounding the electron. The potential energy of an electron[6] at a distance x from a conductor can also be described by (4.11). The potential energy of an electron in the dielectric of a MOS device modifies the barrier it sees due to its own charge interacting with the nearby conductor. If an electron is placed between parallel conductor plates (as is the case in metal-insulator-metal structures), image charges are formed across both conductor plates. The image charges are imaged again across each conductor plate and the pattern repeats itself infinitely. The potential of each of these image charges acting on the electron is summed using the superposition principle to give the potential energy of the electron acting with the two conductor plates. An example of how this

[6] An analogous case can also be made for holes.

occurs for the first three sets of image charges is shown in Figure 4.2. The infinite sum of these charges acting on the electron at *x* is given by:

$$PE(x) = \frac{q}{16\pi\varepsilon}\sum_{n=0}^{\infty}\left(-\frac{1}{nd+x}-\frac{1}{(n+1)d-x}+\frac{2}{(n+1)d}\right), \tag{4.12}$$

where *n* a variable of the summation, which takes on integer values, *x* is the distance to one conductor, and *d* is the distance between conductors.

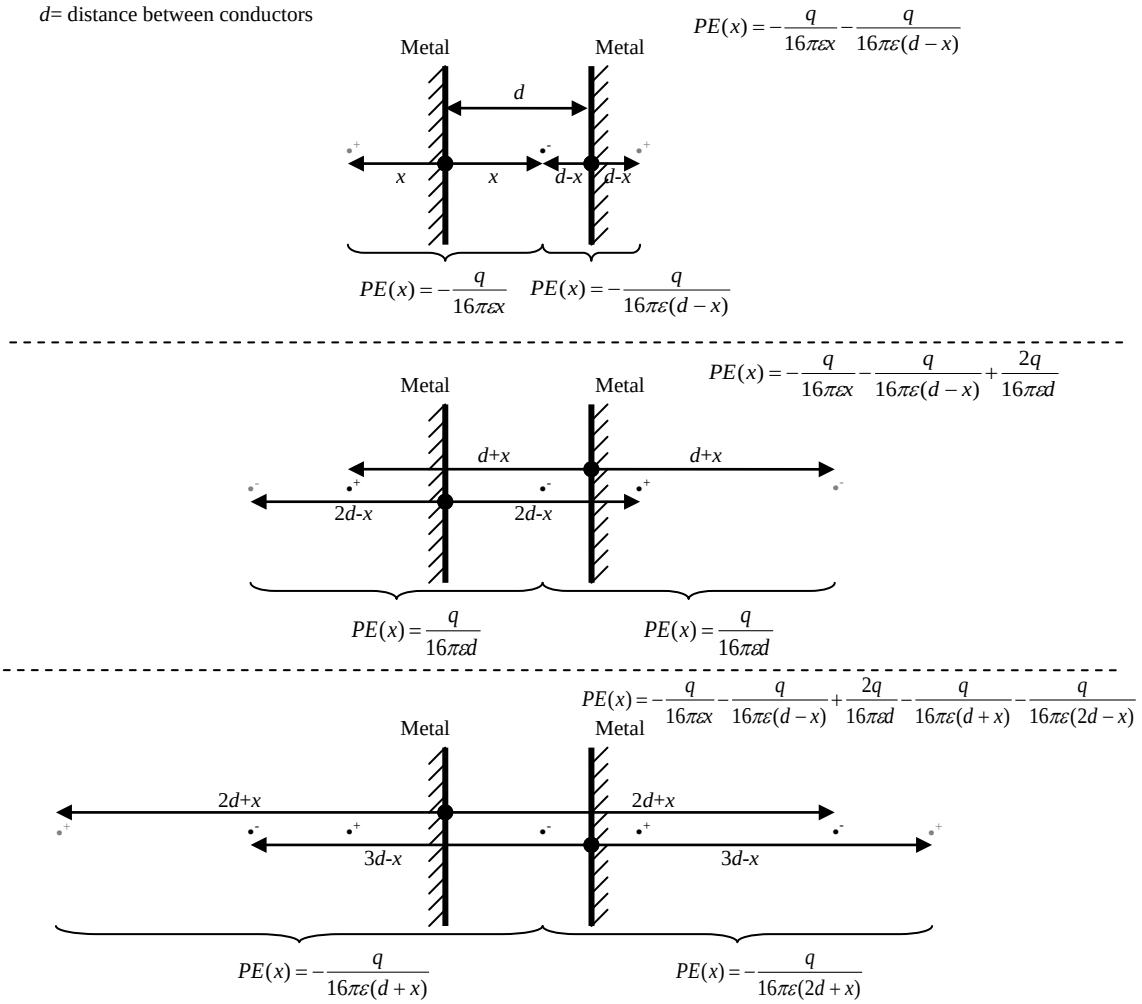


Figure 4.2: Image charges due to an electron between two conductors. Each imaged charge is imaged again across the other conductor resulting in an infinite series of image charges. The first three image charge sets are shown here (top, middle, and bottom) along with the resulting potential.

For multiple dielectrics, the image potential at the dielectric-dielectric interface must also be considered. The potential energy of image charge due to the dielectric-dielectric interface (a charge free interface) is given in units of eV as [80, 81]:

$$PE(x) = \frac{-q}{16\pi\varepsilon x} k \tag{4.13}$$

where

$$k = \frac{\varepsilon_a - \varepsilon}{\varepsilon_a + \varepsilon}, \tag{4.14}$$

ε is the high frequency dielectric constant where the charge is located, and ε_a is the high frequency dielectric constant of the adjacent dielectric. Examining (4.13) and (4.14) indicate that if $\varepsilon > \varepsilon_a$, then the charge is repelled from the boundary, and if $\varepsilon < \varepsilon_a$, the charge is attracted to the boundary. For thin multilayer dielectrics, the effects of boundary to the right and left of the dielectric must be considered. The solution to the image charge potential of a thin dielectric with a dielectric on either side (similar to Figure 4.2 where the metals are replaced with dielectrics) was solved by Kleefstra and Herman and is given as [81]:

$$PE(x) = \frac{q}{16\pi\varepsilon x} \sum_{n=0}^{\infty} (k_l k_r)^n \left(\frac{k_l}{nd + x} + \frac{k_r}{d(n+1) - x} + \frac{2k_l k_r}{(n+1)d} \right) \tag{4.15}$$

where

$$k_l = \frac{\varepsilon - \varepsilon_l}{\varepsilon + \varepsilon_l}, \quad k_r = \frac{\varepsilon - \varepsilon_r}{\varepsilon + \varepsilon_r}, \tag{4.16}$$

ε is the high-frequency dielectric constant of the material for which the potential is solved, ε_l is the high-frequency dielectric constant of the material to the left, and ε_r is the high-frequency dielectric constant of the material to the right. The potential in (4.15) can

also be used if either material to the side of the dielectric is a metal by setting k_l or k_r to -1. For the case in which the dielectric has a metal to the left and right (i.e., k_l and k_r equal -1), (4.15) reduces to (4.13). Within atomic distances of the interface, the potential image charge which is based on classical theory no longer applies [81]. The failure of classical theory within atomic distances of the interface is also observed by examining the potential in (4.11), (4.12), (4.13), and (4.15) as x goes to zero, which results in a potential energy the moves toward infinity.

Using (4.15) and (4.16), the image potential of the electron can be applied to the energy band diagram by subtracting the image potential from the potential calculated in Sections 4.1.1-4.1.2. Both metals and semiconductors can be considered as perfect conductors in the simulation tool and have the effects of the image charge potential applied to the energy band diagrams. An example showing the image barrier lowering for electrons due to the metal gate and Si substrate is shown in Figure 4.3. The incorporation of image charge effects is important in considering Schottky emission as was shown in Section 2.4. The extrapolation of image charge effects to multiple dielectrics is important in this study in which the bidielectric HfO_2/SiO_2 is studied. This model will be used when carrier transport due to image charge effects is considered in Chapter 7.

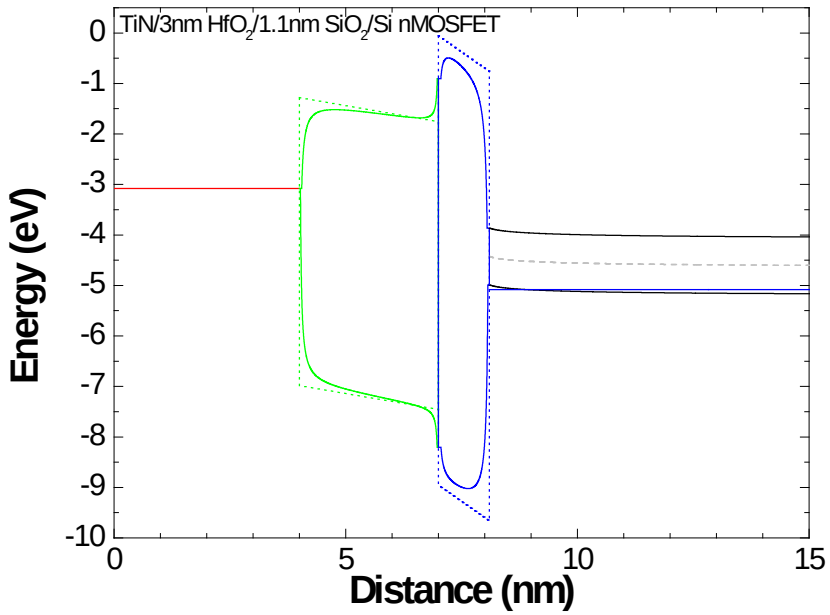


Figure 4.3: An example of a TiN/HfO_2/SiO_2/Si nMOS stack energy band diagram with image charge effects (solid lines) and without image charge effects (dotted lines) for electrons and holes.

4.1.5: Tunneling Distance

The simulation program also supports a geometrically calculated tunneling distance through the dielectrics. The tunneling distance, useful for determining the Fowler-Nordheim tunneling (FNT) distance or direct tunneling (DT) distance for a specified applied bias and the transition from DT to FNT as a function of applied bias, is calculated using the resultant energy barriers. The approach to determine tunneling distance is only a first-order approximation; however, it is of practical value. The tunneling distance is calculated through a given dielectric layer from an adjacent layer's conduction band (CB) or valence band (VB) into its own CB or VB or into another adjacent layer's CB or VB. Shown in Figure 4.4 is an example of the calculated tunneling distance in each dielectric layer for the nMOS device shown in Figure 4.3 without image charge effects.

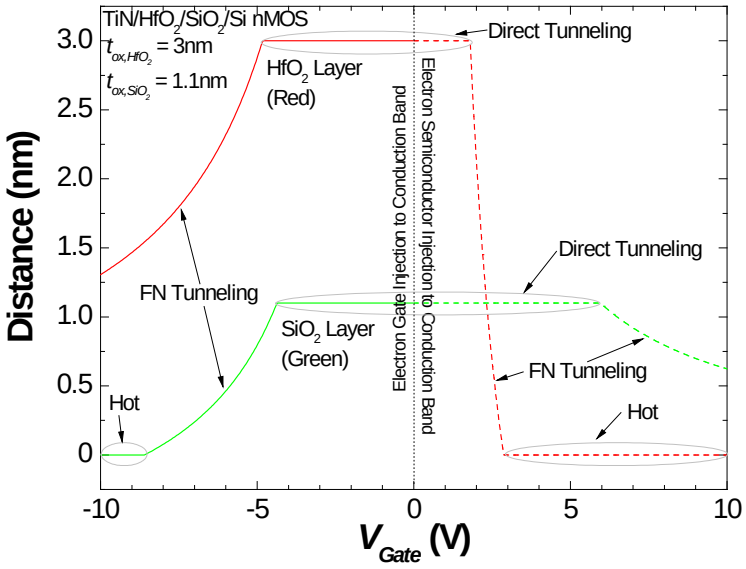


Figure 4.4: Tunneling distance calculations versus gate voltage using the simulation program for a TiN/3nm HfO_2/1.1nm SiO_2/Si nMOS device. Tunneling distance for the HfO_2 and SiO_2 layers are shown in red and green, respectively. The right side of the plot (dashed lines) corresponds to substrate electron injection and the left side to gate electron injection (solid lines). The direct tunneling, Fowler-Nordheim tunneling, and hot injection regimes are labeled.

4.1.6: Quantum Mechanical Transmission Coefficient

The simulation tool supports the calculation of the quantum mechanical transmission coefficient through multilayer dielectrics in the gate stack. To determine the transmission coefficient, the time-independent Schrödinger wave equation must be solved for the stack potential. The one-dimensional time-independent wave equation is given as:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+\phi\psi=E\psi \tag{4.17}$$

where $\hbar$ is the reduced Planck's constant, m is the mass of the particle, x is distance, E is energy of the particle, ϕ is the potential, and ψ is the wave function. The potential to solve is the trapezoidal potential (Figure 4.5), which represents the energy band diagram of the dielectric. Upon finding a solution to the trapezoidal potential, solutions to an arbitrary number of trapezoidal barriers can be determined using transfer matrices.

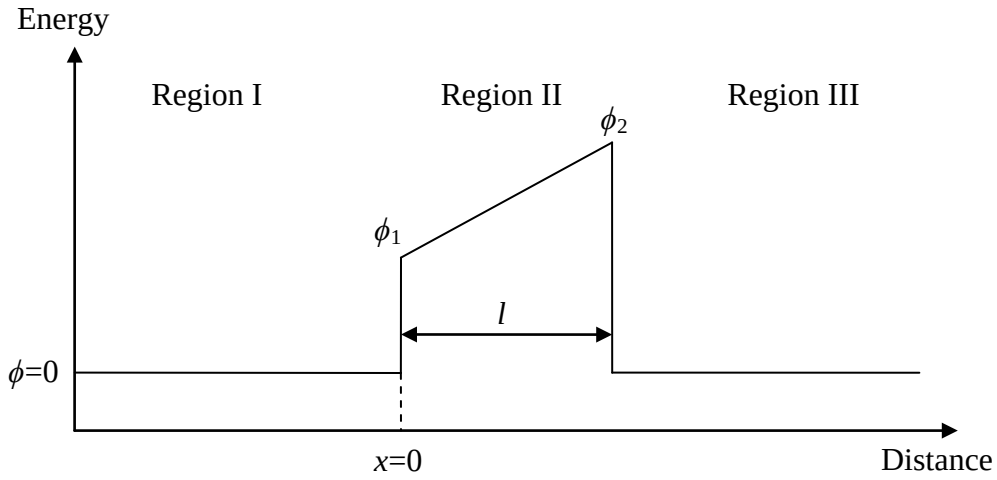


Figure 4.5: Diagram of a trapezoidal barrier to a particle.

To find a solution to the wave equation shown in Figure 4.5, (4.17) must be solved for each region shown. The general solutions for region I and III are the same since they have the same potential [i.e., $\phi_{I,III}(x) = 0$] and take the form of:

$$\psi_I(x) = A_r e^{jkx} + A_l e^{-jkx}, \tag{4.18}$$

$$\psi_{III}(x) = C_r e^{jkx} + C_l e^{-jkx}, \tag{4.19}$$

where the subscripts *r* and *l* of the constants *A* and *C* denote a right and left traveling wave respectively, *j* is the imaginary number, and

$$k = \frac{\sqrt{2mE}}{\hbar}. \tag{4.20}$$

For region II, (4.17) must be solved for the potential:

$$\phi_{II}(x) = \frac{\phi_2 - \phi_1}{l} x + \phi_1, \tag{4.21}$$

where ϕ_1 and ϕ_2 are the left potential and right potential of the trapizodial barrier respectively and *l* is the barrier length (Figure 4.5). The general solution to the wave equation for region II is in the form of Airy functions [82]:

$$\psi_{\mathrm{II}}(x) = B_1 AiryAi(\gamma + x\eta) + B_2 AiryBi(\gamma + x\eta)\,. \tag{4.22}$$

where

$$\gamma = \frac{\beta(\phi_1 - E)}{\left(\frac{\beta(\phi_2 - \phi_1)}{l}\right)^{2/3}}, \tag{4.23}$$

$$\eta = \left(\frac{\beta(\phi_2 - \phi_1)}{l}\right)^{1/3}, \tag{4.24}$$

and

$$\beta = \frac{2m}{\hbar^2}\,. \tag{4.25}$$

The solution to the wave equation for all three regions can now be determined by solving the boundary equations, which states the wave function is continuous and smooth at the boundaries, which is mathematically expressed as:

$$\begin{aligned} \psi_{\mathrm{I}}(0) &= \psi_{\mathrm{II}}(0) \\ \psi_{\mathrm{I}}'(0) &= \psi_{\mathrm{II}}'(0) \\ \psi_{\mathrm{II}}(l) &= \psi_{\mathrm{III}}(l) \\ \psi_{\mathrm{II}}'(l) &= \psi_{\mathrm{III}}'(l) \end{aligned}\,. \tag{4.26}$$

Solving the boundary equations and formulating the result in transmission matrix (***T***) form [83, 84] yields:

$$\begin{bmatrix} C_r \\ C_l \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} A_r \\ A_l \end{bmatrix}, \tag{4.27}$$

where

$$\begin{aligned} T_{11} &= e^{-jkl}\left\{\alpha_1\alpha_2 + \alpha_3\alpha_4 + \alpha_5\alpha_6\right\}/\alpha_7 \\ T_{12} &= e^{j2kd}e^{-jkl}\left\{\alpha_1\alpha_8 + \alpha_3\alpha_9 + \alpha_{10}\alpha_6\right\}/\alpha_7 \\ T_{21} &= e^{-j2kd}e^{jkl}\left\{\alpha_1\alpha_2 - \alpha_3\alpha_4 + \alpha_5\alpha_{11}\right\}/\alpha_7 \\ T_{22} &= e^{jkl}\left\{\alpha_1\alpha_8 + \alpha_3\alpha_9 + \alpha_{10}\alpha_{11}\right\}/\alpha_7 \end{aligned}, \tag{4.28}$$

$$
\begin{aligned}
\alpha_1 &= k \cdot AiryAi(\gamma + l\eta) \\
\alpha_2 &= jk \cdot AiryBi(\gamma) - \eta \cdot AiryBi'(\gamma) \\
\alpha_3 &= \eta \cdot AiryAi'(\gamma + l\eta) \\
\alpha_4 &= k \cdot AiryBi(\gamma) + j\eta \cdot AiryBi'(\gamma) \\
\alpha_5 &= -jk \cdot AiryAi(\gamma) + \eta \cdot AiryAi'(\gamma) \\
\alpha_6 &= k \cdot AiryBi(\gamma + l\eta) - j\eta \cdot AiryBi'(\gamma + l\eta) \\
\alpha_7 &= 2k\eta\left[AiryAi'(\gamma)AiryBi(\gamma) - AiryAi(\gamma)AiryBi'(\gamma)\right] \\
\alpha_8 &= -jk \cdot AiryBi(\gamma) - \eta \cdot AiryBi'(\gamma) \\
\alpha_9 &= -k \cdot AiryBi(\gamma) + j\eta \cdot AiryBi'(\gamma) \\
\alpha_{10} &= jk \cdot AiryAi(\gamma) + \eta \cdot AiryAi'(\gamma) \\
\alpha_{11} &= k \cdot AiryBi(\gamma + l\eta) + j\eta \cdot AiryBi'(\gamma + l\eta)
\end{aligned}
\quad , \tag{4.29}
$$

and *d* is a lateral distance shift in the barrier. The transmission matrix for multiple barriers at various distances *d* can be found by using (4.27). The transmission matrix through the entire stack of barriers is found by matrix multiplication of the individual barrier transmission matrices. The quantum mechanical transmission coeffecient (*T*) is found by:

$$
T = \left| \frac{T_{11}T_{22} - T_{12}T_{21}}{T_{22}} \right|^2 . \tag{4.30}
$$

An example of calculating the transmission coefficient for a TiN/3nm HfO_2/1.1nm SiO_2/p-type Si stack at V_{Gate} = 2V is shown in Figure 4.6.

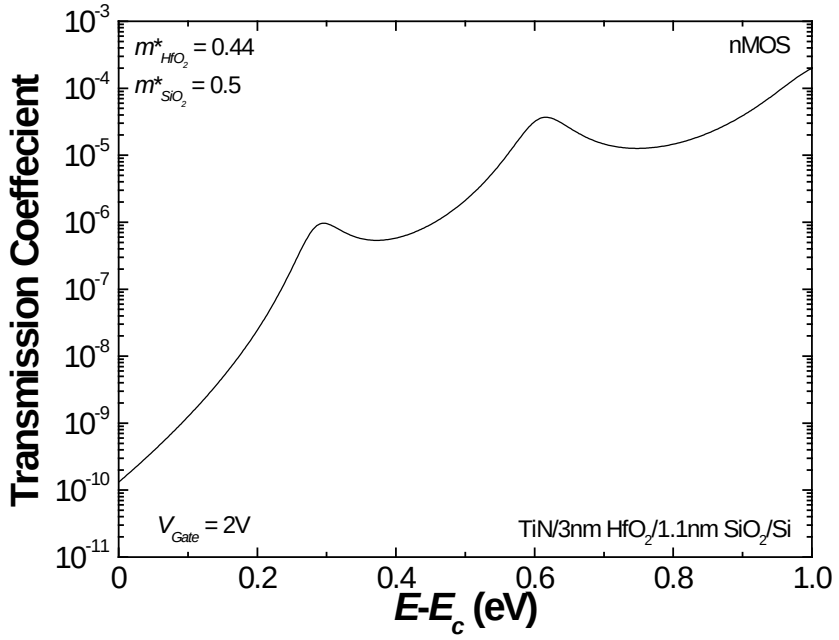


Figure 4.6: Transmission coefficient calculation for electron substrate injection. The transmission coefficient is shown for V_{Gate} = 2V at various energy levels above the conduction band (E_c) of the Si (gate stack in inset).

4.2: Capabilities

The simulation tool provides the user with the ability to change any of the materials parameters used in the model to increase accuracy. Interfacial layers, common in high-*k* dielectrics, can be simulated by adding additional layers permitting further refinement to the simulation. An unlimited amount of fixed sheet charge can be added to any dielectric to simulate trapped fixed charge. The simulation tool also supports temperature dependent energy band gaps and intrinsic carrier concentrations of semiconductors. An electric field dependent dielectric constant is also supported for potential use in MOS ferroelectric devices [85] and multilayer ferroelectric materials [86, 87]. Image charge effects and the quantum mechanical transmission coefficient through multilayer dielectrics are also modeled.

With these capabilities, complex metal-insulator semiconductor devices whose configurations can include $m\text{-}I_j\text{-}X$, $M\text{-}I_i\text{-}M\text{-}I_j\text{-}X$, and $M\text{-}I_i\text{-}(M\text{-}I)_j\text{-}X$ where *m* and *I* signify metal and insulator, *i* and *j* are integer numbers, and *X* denotes a metal or a semiconductor, can be readily examined and exported for further investigation. The current version of the program contains over 18,000 lines of code.

4.3: Validation of the Model

Because of the advanced capabilities of finite-element PSES programs, the PSES program SILVACO [88] was used to assess the accuracy of the simulation tool. The energy band diagram and electric field generated by the simulation tool described in the previous section is directly compared to the output of SILVACO. A gate stack composed of TiN/5nm HfO_2/1.1nm SiO_2/p-Si is used for the comparison. This stack corresponds to one of the stack types described in Section 3.3 and is of interest in this study. The input parameters used in the SILVACO simulations are given in Appendix B.

The energy band diagram calculated by the simulation tool and SILVACO with and without quantum mechanical (QM) correction is shown in Figure 4.7 for a gate bias of 2V and in Figure 4.8 for a gate bias of -2V. Both Figure 4.7 and Figure 4.8 show the simulation tool produces nearly identical results to the energy bands calculated by SILVACO without quantum mechanical effects. Appling quantum mechanical correction factors to the Si semiconductor using SILVACO results in a slight modification to the energy band diagram that is well approximated by the simulation tool. In particular, the SiO_2 IL should change the most with QM corrections as it is adjacent to the Si, yet the change is insignificant.

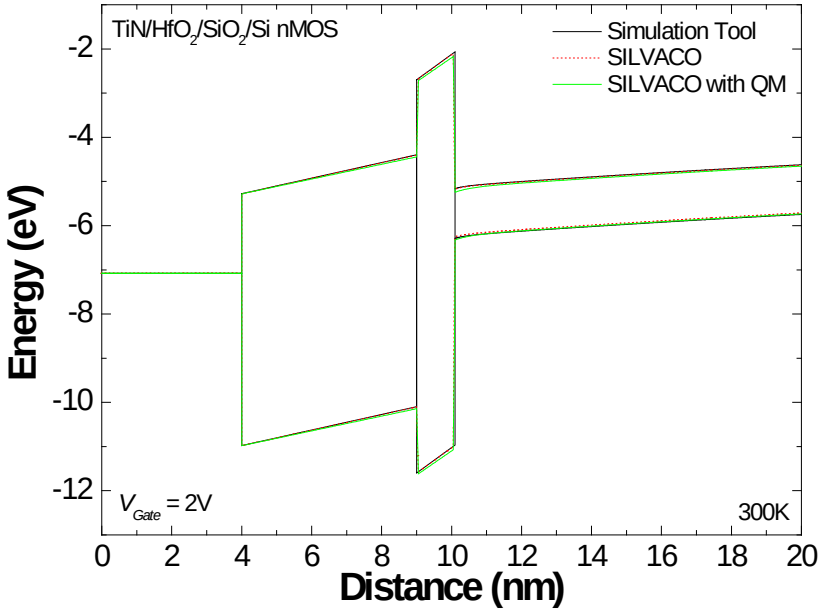


Figure 4.7: Energy band diagram of a TiN/5nm HfO_2/1.1nm SiO_2/p-Si stack calculated using the simulation tool, SILVACO, and SILVACO with quantum mechanical (QM) corrections for a gate bias of 2V. Minimal differences are observed.

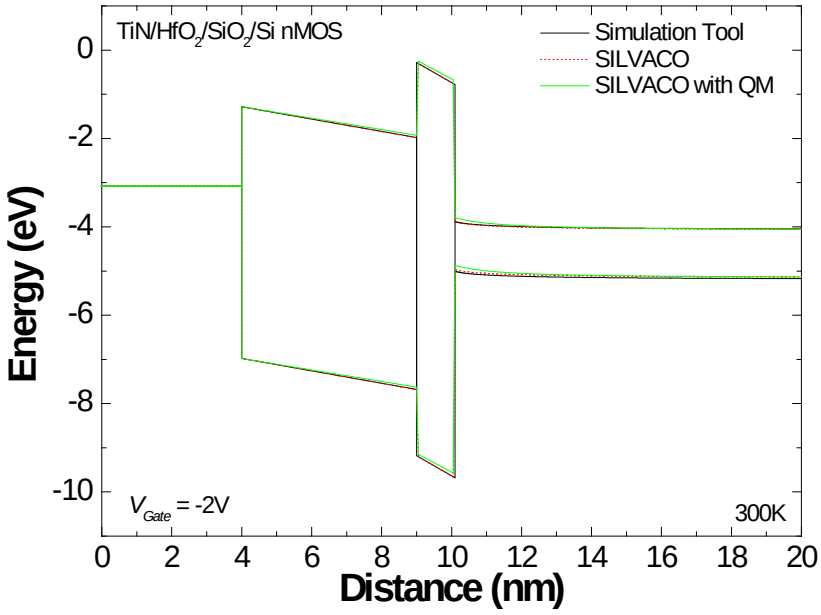


Figure 4.8: Energy band diagram of a TiN/5nm HfO_2/1.1nm SiO_2/p-Si stack calculated using the simulation tool, SILVACO, and SILVACO with quantum mechanical corrections for a gate bias of -2V. Minimal differences are observed.

A comparison of the calculated electric field for the dielectrics in the $TiN/HfO_2/SiO_2/p\text{-}Si$ stack is shown in Figure 4.9 as the gate voltage is stepped from -2V to 2V. The data are generated three different ways as were the energy band diagrams discussed above: 1) using the simulation tool, 2) using SILVACO without QM corrections, and 3) using SILVACO with QM corrections. The data symbols in Figure 4.9 are spaced in such a manner as to aid in the comparison. All three calculation methods yield similar results. The electric field calculated using the simulation tool (solid lines) and SILVACO without QM corrections (symbols) correspond nearly exactly. Comparing the output of the simulation tool to that of SILVACO with QM corrections (dashed lines) produces very similar results. Hence, Figure 4.9 demonstrates that the electric field of the dielectric layers changes very little with the quantum corrections performed by SILVACO. The electric field calculated using SILVACO with QM correction is slightly lower compared to the simulation tool due to an increase in the potential drop in the semiconductor. The simulation tool therefore overestimates the electric field of the dielectrics and underestimates the potential drop in the semiconductor. Additionally, this simulation tool over estimates the capacitance[7].

[7] Note that this discussion only applies when the material is a semiconductor and not a perfect metal.

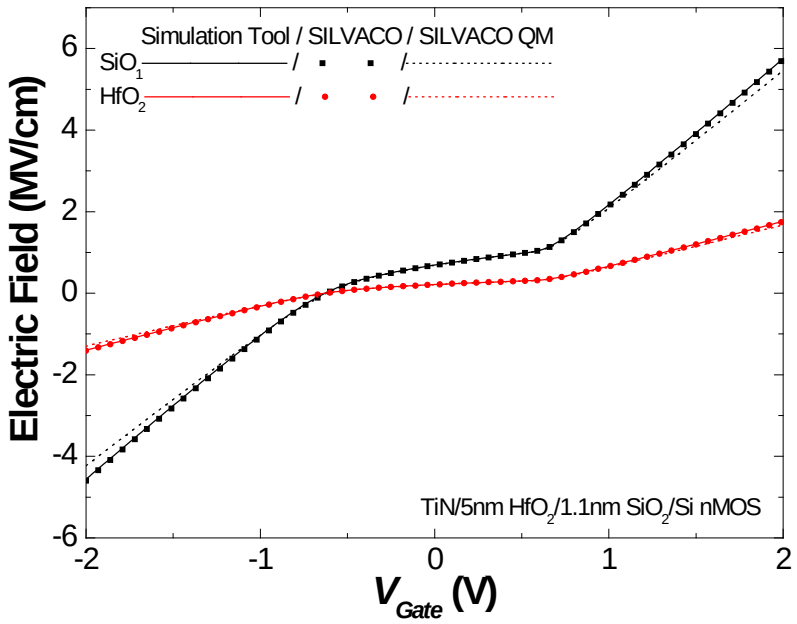


Figure 4.9: Electric field in the SiO_2 and HfO_2 layer as a function of gate voltage calculated by the simulation tool and SILVACO with and without quantum mechanical corrections.

These calculations and comparisons demonstrate that the simulation program provides results that are reasonable and very similar to much more sophisticated calculations. The time required to learn and then use the freeware program is considerably less thereby increasing the accessibility for those that are not accustomed to involved simulations on complex multilayer dielectric structures.

CHAPTER 5: EXPERIMENTAL RESULTS AND DISCUSSION

Representative gate leakage current measurements for the 3 different gate stacks investigated (3nm HfO_2/1.1nm HfO_2, 5nm HfO_2/1.1nm SiO_2, and 2nm SiO_2) as a function of gate voltage and temperature are shown in Figure 5.1 and Figure 5.2 for nMOSFETs and pMOSFETs, respectively. As was shown in Figure 1.1, Figure 5.1 and Figure 5.2 indicate that the temperature dependence of the gate leakage current for both n- and pMOSFET devices is much larger for the high-*k* devices compared to the SiO_2 control device.

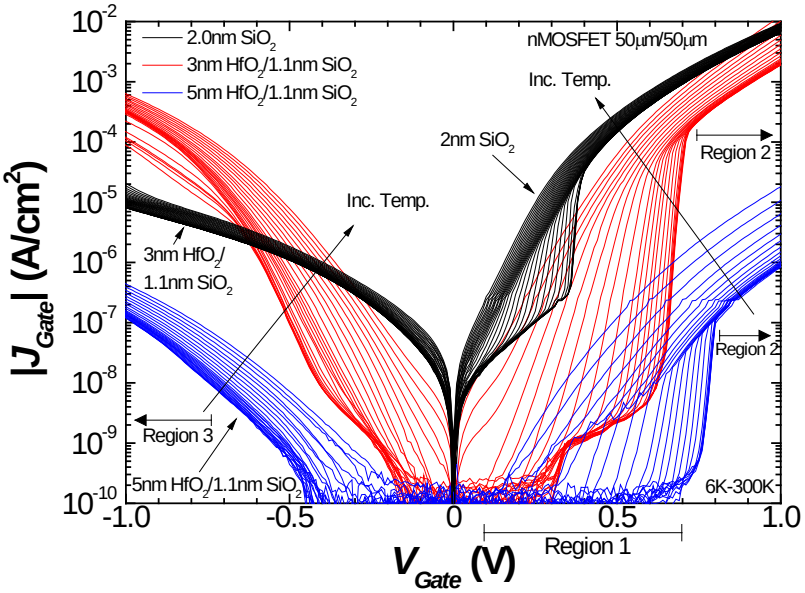


Figure 5.1: Gate leakage current for nMOSFETs composed of a 3nm or 5nm HfO_2/1.1 SiO_2 gate dielectric and a 2nm SiO_2 gate dielectric for temperatures ranging from 6K to 300K.

For positive gate biases of 1V, the 3nm HfO_2 nMOSFETs show an order of magnitude increase in gate leakage current with the pMOSFETs showing over three

orders of magnitude increase. The 5nm HfO_2 stack shows over an order of magnitude increase in the gate leakage current at a gate bias of 1V for both n- and pMOSFETs. Meanwhile, the SiO_2 control device displays less than an order of magnitude increase in the gate leakage current. For gate leakage currents near -1V, the temperature dependence of the gate leakage current is in general small compared to +1V. At -1V, the 3nm HfO_2 nMOSFET device shown in Figure 5.1 displays a large increase in the gate leakage current with respect to temperature similar to 3nm HfO_2 pMOSFET device at +1V shown in Figure 5.2. While the increase in gate leakage current for -1V gate biases is lower than 1V gate biases for the high-*k* devices, it is nonetheless higher than SiO_2.

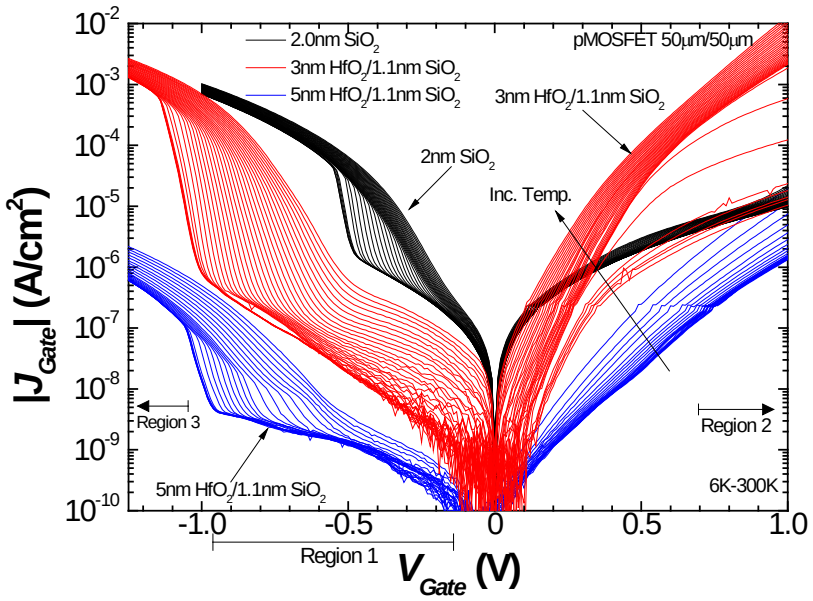


Figure 5.2: Gate leakage current for pMOSFETs composed of a 3nm or 5nm HfO_2/1.1 SiO_2 gate dielectric and a 2nm SiO_2 gate dielectric for temperatures ranging from 6K to 300K.

Since the EOTs for the 5nm HfO_2 and 2nm SiO_2 devices are relatively similar ($EOT_{5nm\ HfO2/1.1nm\ SiO2}$ = 1.7nm and EOT_{SiO2} = 2nm), the HfO_2/SiO_2 temperature dependence of the gate leakage current data differ from the SiO_2 data most likely because of the presence of HfO_2 defects, phonon modes, and perhaps carrier transport modes.

There is also a significant difference in the gate leakage current of the 5nm HfO_2 devices compared to the 3nm HfO_2 device. While the magnitude of the gate leakage current of the 5nm HfO_2 device is much smaller than the 3nm HfO_2, the temperature dependence is larger. The exception to the previous statement is the large increase in the gate leakage current observed a +1V for the 3nm HfO_2 nMOSFET devices and -1V for the 3nm HfO_2 pMOSFET devices. The majority of the gate leakage current increase occurs at low temperatures T < ~50K and the incremental increase in the gate leakage current at higher temperature T > 100K is larger in the 5nm HfO_2 device compared the 3nm HfO_2 device. This signifies strong thickness dependence in the transport mechanism in the HfO_2 layer or a transition between bulk transport and thin layer transport.

5.1: General Temperature Observations

The gate current versus gate voltage (I_G-V_G) temperature dependent curves for both n- and pMOSFETs show similar trends in three different regimes. These regimes are: 1) weak inversion (labeled Region 1 in Figure 5.1 and Figure 5.2 and the encircled regions in Figure 5.3), 2) relatively large positive gate biases (0.7V to 1V, labeled Region 2 in Figure 5.1 and Figure 5.2), and 3) relatively large negative biases (~ -1V labeled Region 3 in Figure 5.1 and Figure 5.2). The following subsections propose qualitative explanations for the temperature increases for the three regimes [10].

5.1.1: Weak Inversion – Region 1

For both n- and pMOSFETs composed of HfO_2/SiO_2 and SiO_2 gate dielectrics, a large increase in the gate leakage current density, J_{Gate}, versus gate voltage is observed (the encircled regions in Figure 5.3) during weak inversion of the MOSFET channel. As

the temperature increases, the slope of J_{Gate} vs V_{Gate} decreases following temperature dependent trends in sub-threshold slope [89]. Applying a $3k_BT$ bias (i.e., large enough that thermal noise is not a factor) between the source and drain [89] to measure the presence of inversion charge (Test 2 in Table 3.3), reveals the sudden increase in J_{Gate} (encircled regions in Figure 5.3) corresponds to the sudden increase of minority carriers in the channel (steep slope region of I_{Source} in Figure 5.4).

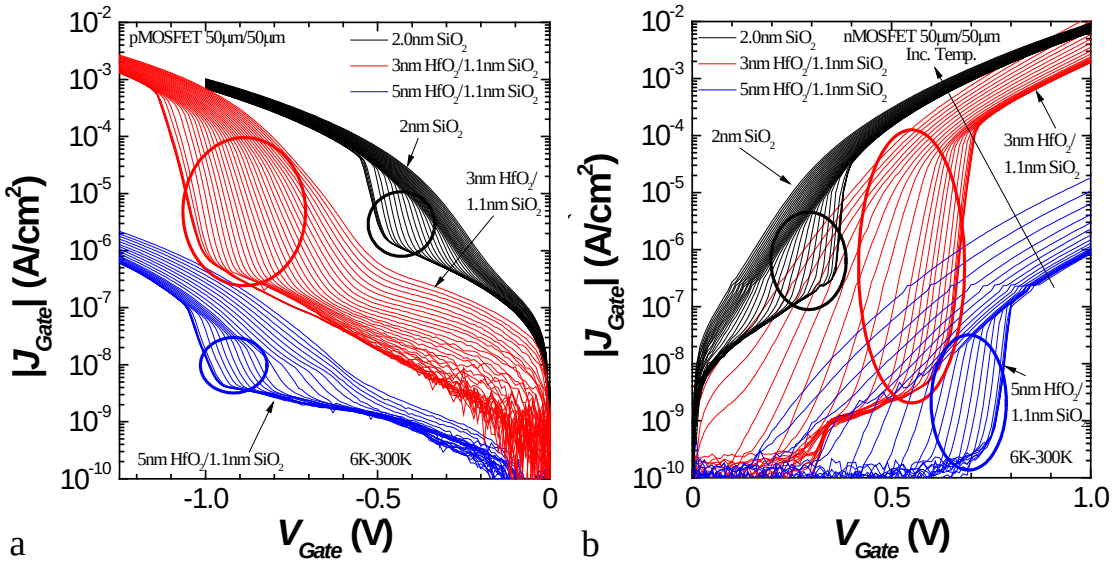


Figure 5.3: Gate leakage current density of (a) pMOSFETs and (b) nMOSFETs composed of 2nm SiO_2, 3nm HfO_2/1.1nm SiO_2, and 5nm HfO_2/1.1nm SiO_2 for temperatures ranging from 6K – 300K. The encircled regions indicate a common temperature response in the gate leakage current, which occurs in the weak inversion region.

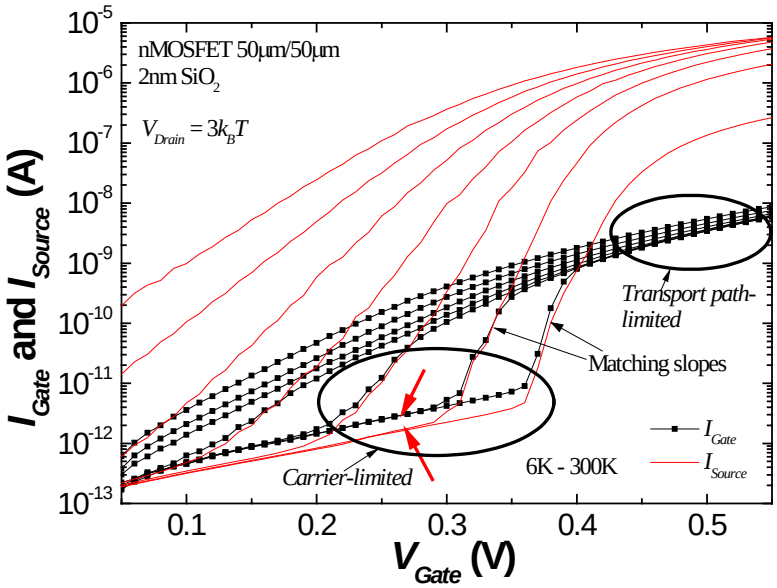


Figure 5.4: A comparison of source current, (also a measure of the inversion charge) to the gate leakage current. Current is transport path-limited above the threshold voltage (V_{TH}) and carrier rate limited below ~V_{TH} at low temperatures. Thick red arrows indicate source to gate overlap transport. Only select temperatures are shown for clarity.

The data of Figure 5.4 suggest *carrier-limited* and *transport path-limited* regimes. Consider Figure 5.5 where the currents from all four MOSFET terminals are plotted. In Figure 5.5, the drain and source current are asymmetric across the voltage axis. At voltages near -1.05V, the gate current is nearly symmetric with both the drain and source currents. In the perfect scenario, the drain current comes from the source current and hence the source and drain current would be symmetric across the voltage axis. The asymmetry in the drain current with respect to the source current and the symmetry in the gate current with respect to both the drain and source current near the threshold voltage indicate that the first carriers that enter the channel from the source and drain do not flow from the source to the drain (i.e., source and drain current are asymmetric across the voltage axis), but flow from the source and drain through the gate dielectric to the metal gate (i.e., the gate and drain/source currents are symmetric near the threshold voltage).

The increase in I_{Gate} at the threshold voltage is therefore due to minority carriers from the source and drain terminals entering the channel during inversion formation. Hence, prior to the threshold voltage (i.e., depletion/weak inversion), the gate leakage current is *carrier-limited,* labeled in Figure 5.4, due to the absence of minority carriers. The slope of the gate leakage current matches the sub-threshold slope of the source current (labeled "Matching slopes" in Figure 5.4), which indicates that the gate leakage current follows the number of carriers in the channel and is therefore carrier-limited in this regime. It appears that I_{Gate} then saturates even as the inversion charge continues to increase exponentially (as indicated by the exponential increase in the source current) before losing its exponential dependence with the gate bias. The continued exponential increases in the carrier concentration while the gate leakage current saturates as a function of gate voltage indicates a *transport path-limited* mechanism, labeled in Figure 5.4. As the temperature increases, the transition between *carrier-limited* and *transport path-limited* conduction occurs at lower voltages due to the increased number of minority carriers as temperatures increase. Above the threshold voltage however, a sufficient amount of minority carriers are present so that the leakage current is *transport path-limited* over the temperature range investigated.

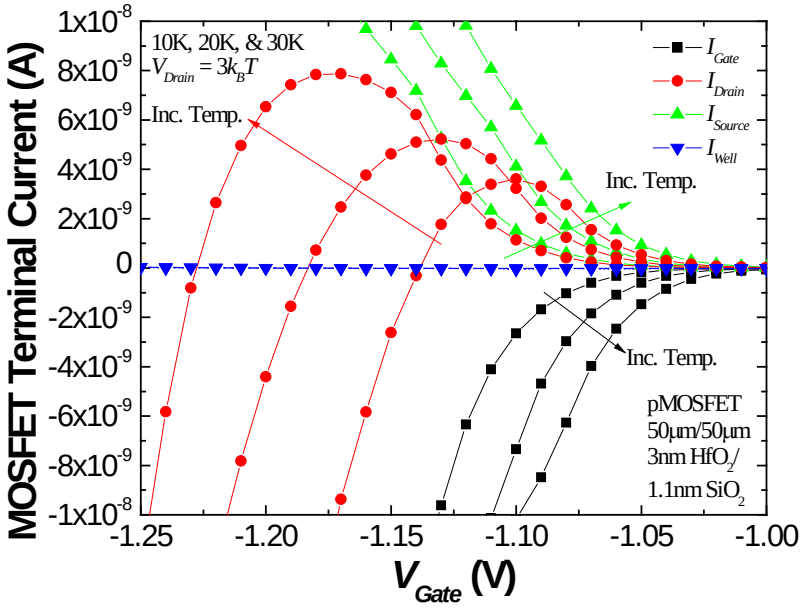


Figure 5.5: A plot of all four terminal currents at 10K, 20K, and 30K in a 3nm HfO_2/1.1nm SiO_2 pMOSFET. Tunneling current (i.e., gate current) is carrier starved as observed by the increase in the gate leakage current corresponding to the presence of minority carriers flowing from the source and drain to the gate.

Similar slopes for both the gate leakage and source currents with respect to gate voltage are observed in 2nm SiO_2 devices prior to channel inversion (region of interest is indicated by thick red arrows in Figure 5.4). The similar slopes between the gate and source currents in 2nm SiO_2 devices indicates that, in that regime (prior to weak inversion), the gate leakage current corresponds to the gate to source/drain overlap region. This suggestion is further confirmed by analyzing the gate leakage currents of two differently sized MOSFETs.

Information concerning the physical location of carrier transport through the gate dielectric stack can be obtained by examining the gate leakage currents of differently sized MOSFETs. If carrier transport is dominated by the conduction mechanisms of the gate dielectric stack, the gate leakage current is expected to scale with the area of the MOSFET (WxL). If the gate leakage current is dominated by interface effects at the edge

of the channel area, then the gate leakage current is expected to scale with the length of the leaky edge. So, if interface effects of the entire edge of the MOSFET channel area dominate the gate leakage current, then the gate leakage current is expected to scale with the perimeter of the device ($2L + 2W$). Hence, determining how the gate leakage current scales can indicate the location of where the dominate current transport is occurring.

To access the physical location of the gate leakage current, the gate leakage current density (scaled by area, J_{Gate}) of 30μm/30μm and 50μm/50μm 2nm SiO_2 devices is plotted in Figure 5.6. The gate leakage current densities scaled by *area* for the two differently sized devices are the same in the *transport path-limited* regime indicating that the leakage current is dominated by transport from the Si channel, which is *area* dependent.

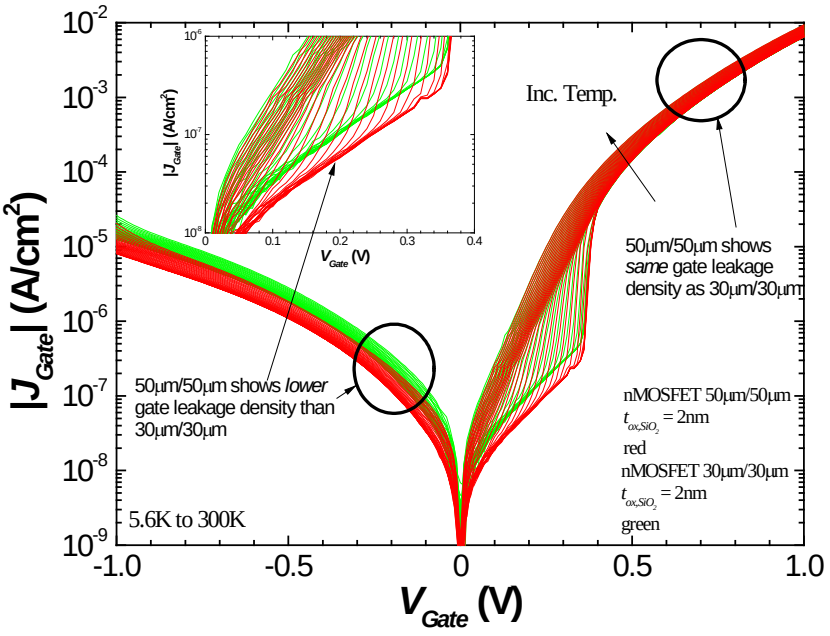


Figure 5.6: Comparison of the gate leakage current densities of two different area 2nm SiO_2 nMOSFETs, 30μm/30μm and 50μm/50μm. Areas of device operation where the gate leakage current densities are the same and differ are indicated. Inset shows the depletion/weak inversion regime.

An exception to the preceding argument is made in the *carrier-limited* regime and in accumulation where the 50μm/50μm device shows a lower gate leakage current density than the 30μm/30μm device (Figure 5.6). This may be explained by scaling the gate leakage current according to the MOSFET's *width* (*W*), $W_{Gate} = I_{Gate}/W$ (since the gate to source/drain overlap region scales with device width[8]), rather than the MOSFET's *area* as shown in Figure 5.7. The gate leakage current scaled by width (W_{Gate}) for the 30μm/30μm device and the 50μm/50μm device is the same (i.e., superimposed) in the accumulation and depletion regimes, indicating that the gate leakage current scales according to *width* in these regimes. Current transport in the accumulation and depletion regimes is therefore dominated by the area overlap of the gate and the source/drain junctions, which scales by *width*. Note that this only applies to the 2nm SiO_2 control devices. For the 3nm HfO_2 and 5nm HfO_2 devices, the gate leakage current scales well with area in all operating regimes, as indicated by the overlap of both J_{Gate}'s of the 30μm/30μm and 50μm/50μm devices in Figure 5.8.

[8] The area of the gate to drain/source overlap region is the length of the gate to drain/source overlap region multiplied by the MOSFETs width. The length of the gate to drain/source region is controlled by the manufacturing process in a self-aligned process and can be assumed constant between various MOSFET sizes. The width of the MOSFET is primarily controlled by the device layout and to a much lesser extent by the manufacturing process (especially in large width devices). The gate to drain/source overlap region therefore scales with the MOSFET's width.

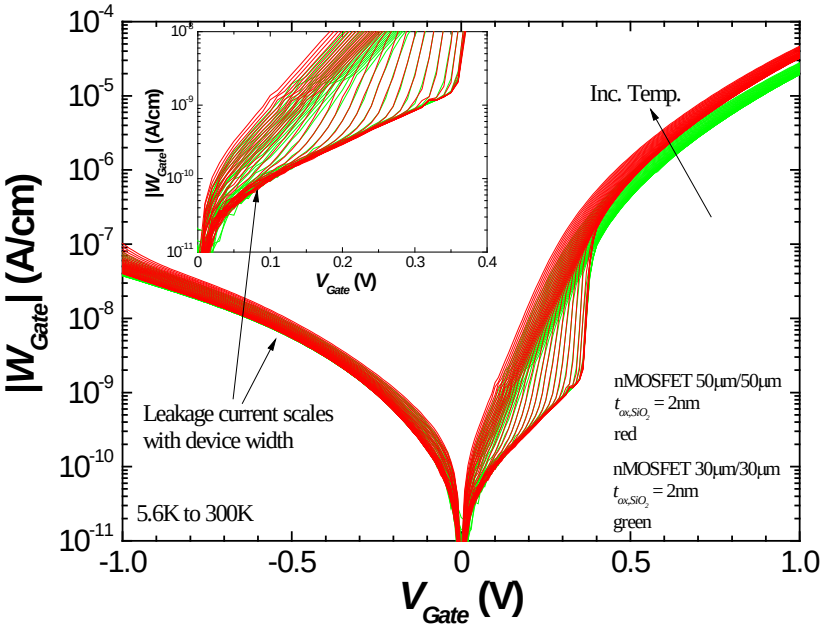


Figure 5.7: Comparison of the gate leakage current of 2nm SiO_2 nMOSFETs when scaled by the MOSFETs width. Overlap of currents in the accumulation and depletion regions indicate that current transport is occurring in the gate to source/drain overlap regions. Inset shows the depletion/weak inversion regime.

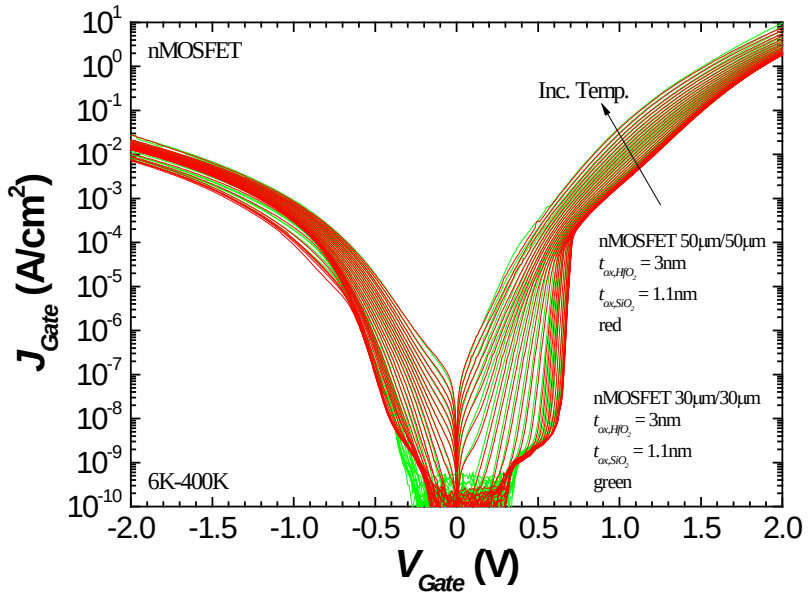


Figure 5.8: Gate leakage current density for 3nm HfO_2/1.1nm SiO_2 devices having areas 30µm/30µm and 50µm/50µm. The gate leakage current scales well with area, indicating current transport is dominated in the channel area of the device.

5.1.2: Relatively Large Positive Gate Biases – Region 2

For relatively large positive gate biases, gate biases above the threshold voltage in nMOS devices, the temperature dependence of the gate leakage current differs significantly between the three gate stacks examined. Figure 5.9 shows the gate leakage current density for nMOSFETs (Figure 5.9a) and pMOSFETs (Figure 5.9b) for the three gate stacks examined. The temperature range shown in Figure 5.9 extends from 6K to 300K except for the 3nm HfO_2/1.1nm SiO_2 pMOSFET device, where the temperature range is from 50K to 300K. Data for temperatures below 50K are not shown for the 3nm HfO_2/1.1nm SiO_2 pMOSFETs because another temperature dependent transport regime dominates in the accumulation regime (Figure 5.2). An explanation of the temperature dependent transport regime for $T < 50K$ for devices in accumulation is discussed below in Section 5.3.

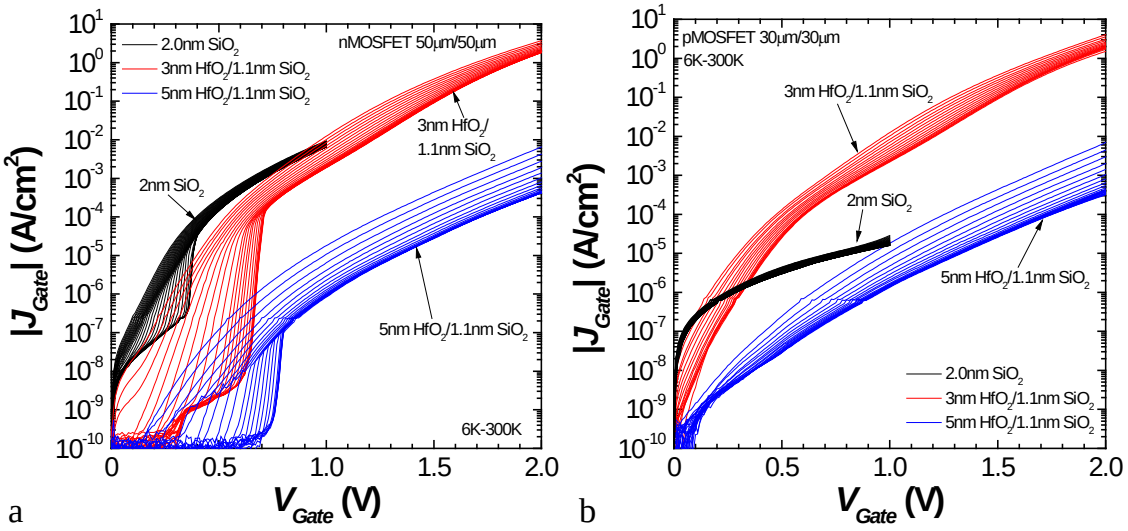


Figure 5.9: Gate leakage current density plots for (a) nMOS and (b) pMOS devices composed 2nm SiO_2, 3nm HfO_2/1.1nm SiO_2, and 5nm HfO_2/1.1nm SiO_2. Currents are shown for positive gate biases and for temperatures ranging from 6K to 300K except for the 3nm HfO_2/1.1nm SiO_2 pMOSFET where only temperatures ranging from 50K to 300K are shown.

Trends in the gate leakage current are consistent for n- and pMOSFET devices as shown in Figure 5.9 except for the depletion/weak inversion regime as was discussed in the section above. The gate leakage current increase with temperature of the 2nm SiO_2 control devices show relatively little increase in contrast with the large change in the gate leakage current of HfO_2 based gate stacks with temperature. The 5nm HfO_2 device has a greater temperature dependence in the gate leakage current compared to the 3nm HfO_2 devices. The greater temperature dependence of J_{Gate} in the thicker 5nm HfO_2 devices, compared to the 3nm HfO_2 devices, suggests that the temperature dependence of the gate leakage current is dominated by the HfO_2 layer and not the HfO_2/SiO_2 interface since the interface remains unchanged between the 3nm and 5nm HfO_2 devices.

A bulge in J_{Gate} over the large temperature range investigated is seen in the 3nm HfO_2/1.1nm SiO_2 devices in Figure 5.9 and has maximum at a gate voltage of ~1.1V. The temperature dependence of the gate leakage current increases until a gate voltage of ~1.1V (i.e., the maximum) is reached after which the temperature dependence of the gate leakage current decreases. This trend is observed in both n- and p-MOSFET devices but not in the 5nm HfO_2/1.1nm SiO_2 devices suggesting a SiO_2 effect.

A possible qualitative argument for the bulge seen in Figure 5.9 is as the gate bias increases, the Fermi energy level in the Si moves closer to the conduction band of the HfO_2. Nevertheless, at low temperatures (i.e., minimal thermal broadening –Figure 1.2) and low voltages, the Fermi level is not close enough to access defect states in the SiO_2 layer located in energy above the Si conduction band. As the temperature increases, the thermal energy distribution broadens (Figure 1.2), allowing electrons to occupy higher energy levels in the Si that have a greater probability to coincide with defect levels in the

SiO_2. The combined qualitative result is the temperature dependence of the gate leakage current increases as the gate voltage increases. As the gate bias continues to increase beyond ~1.1V, less temperature is required (resonance conditions are met between the defects and the Si Fermi level by further band bending) to promote electrons to energy levels in resonance with defects in SiO_2 layer, resulting in a decrease in the temperature dependence of J_{Gate} for V_{Gate} = ~1.1V-2V.

The suggestion that electrons from the Si contribute to the bulge in the gate leakage current (Figure 5.9) may be substantiated by examining the energy band diagram of an nMOS device for gate biases of 0V, 1V, and 2V shown in Figure 5.10. As the gate voltage increases, the barrier for electrons is reduced while the barrier for holes remains large for all three biases. The energy band diagram suggests an electron substrate dominated transport mechanism for the relatively large positive gate bias regime. This suggestion is confirmed using carrier separation analysis below in Section 5.2.1.

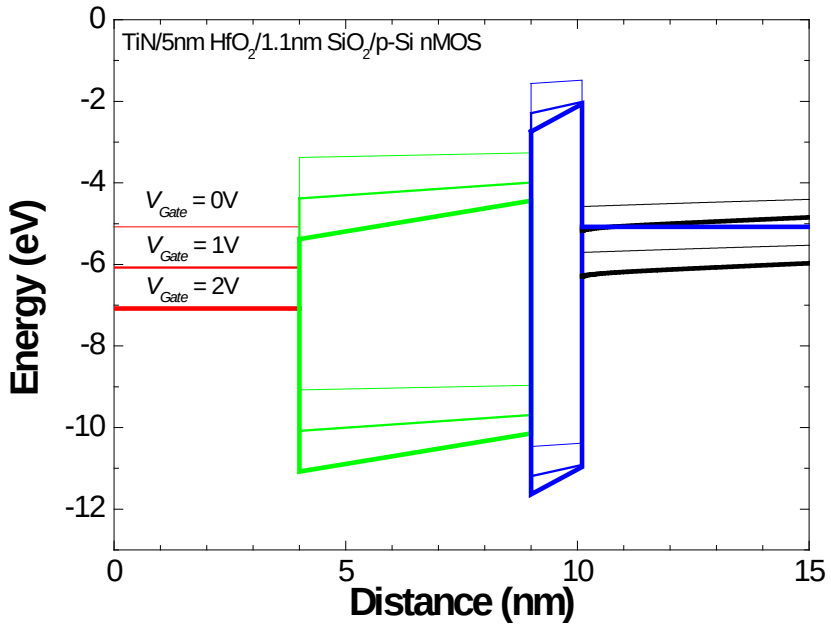


Figure 5.10: Energy band diagram of a TiN/5nm HfO_2/1.1nm SiO_2/p-Si nMOS gate stack at gate biases of 0V, 1V, and 2V. The barrier for substrate electrons decreases as the gate voltage increases. The barrier for holes remains appreciably large for all three gate voltages shown.

5.1.3: Relatively Large Negative Gate Biases – Region 3

Figure 5.11 shows the gate leakage current density for nMOSFETs (Figure 5.11a) and pMOSFETs (Figure 5.11b) for the three gate stacks examined at negative gate biases. The temperature range shown in Figure 5.11 ranges from 6K to 300K, except for the 3nm HfO_2/1.1nm SiO_2 nMOSFET device, where the temperature range is from 50K to 300K. Similar to 3nm HfO_2 pMOSFETs in accumulation, temperatures below 50K are not plotted for the 3nm HfO_2 nMOSFET due to another temperature dependent mechanism dominating in the accumulation regime for $T < 50$K. An explanation of the preceding mechanism is discussed below in Section 5.3.

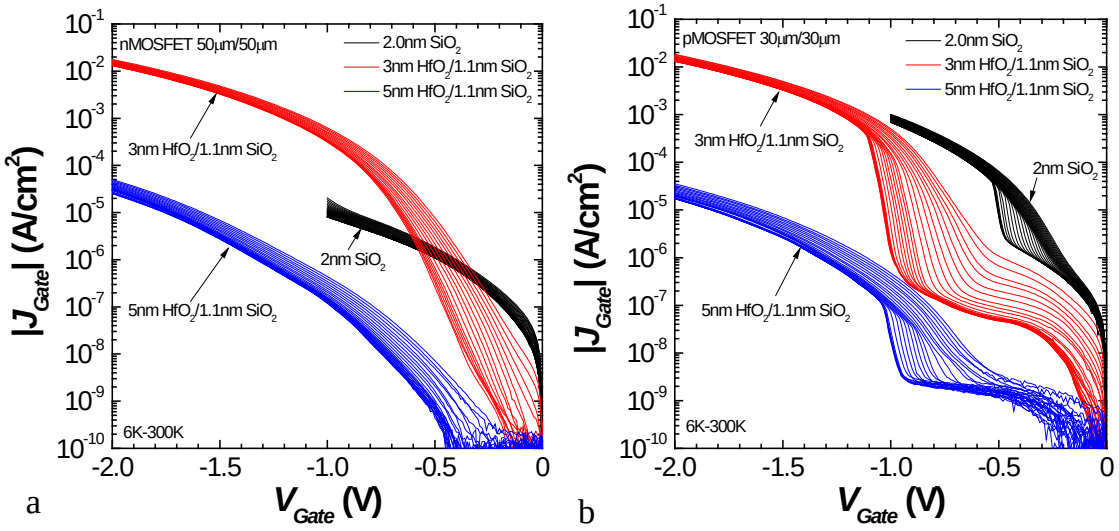


Figure 5.11: Gate leakage current density plots for (a) nMOS and (b) pMOS devices composed 2nm SiO_2, 3nm HfO_2/1.1nm SiO_2, and 5nm HfO_2/1.1nm SiO_2. Currents are shown for negative gate biases and for temperatures ranging from 6K to 300K, except for the 3nm HfO_2/1.1nm SiO_2 nMOSFET, where only temperatures ranging from 50K to 300K are shown.

Trends in the gate leakage current are consistent for n- and pMOSFET devices as shown in Figure 5.11, except for the depletion/weak inversion regime as was discussed in Section 5.1.1. The difference in the temperature dependence of the three gate stacks examined in the negative gate voltage regime is not large, which includes the temperature dependence difference of the 2nm SiO_2 control devices. The 5nm HfO_2 devices have a slightly stronger temperature dependence of J_{Gate} than the 3nm HfO_2 devices. The temperature dependence of J_{Gate} decreases as V_{Gate} approaches -2V, which may suggest an electric field dependence. A bulge is not seen in the 3nm HfO_2 temperature dependence of J_{Gate} in the negative bias regime as was observed at positive gate voltages in Figure 5.9.

As was done for positive gate biases in Figure 5.10, the energy band diagram of a TiN/5nm HfO_2/1.1nm SiO_2/p-Si nMOS device biased at 0V, -1V, and -2V is shown in

Figure 5.12 for completeness. Figure 5.12 indicates the barrier to gate electrons and substrate holes is large over the negative gate voltage regime of interest.

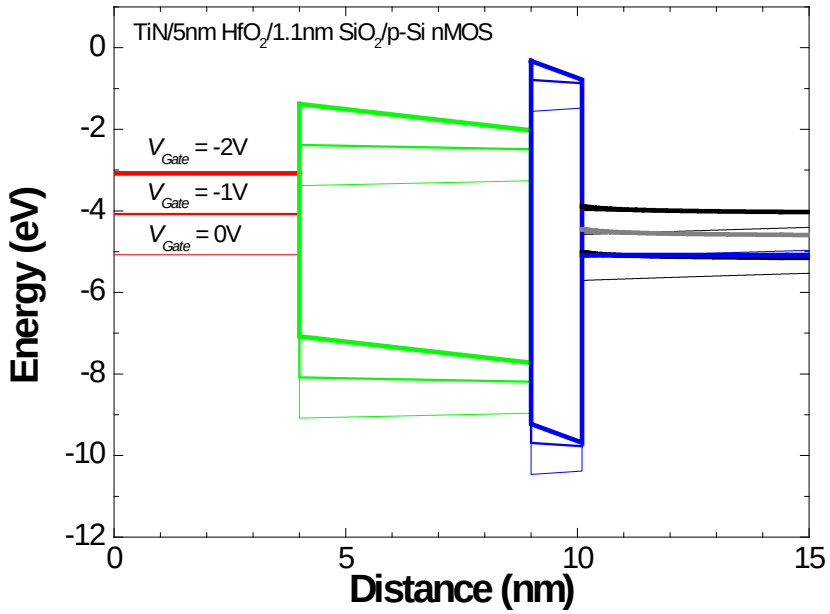


Figure 5.12: Energy band diagram of a TiN/5nm HfO_2/1.1nm SiO_2/p-Si nMOS gate stack at gate biases of 0V, -1V, and -2V. While the barrier to metal electrons decreases, the barrier remains formidable. The barrier for holes remains appreciably large for all three gate voltages shown.

5.2: Carrier Separation: Results and Discussion

An important element in determining the carrier transport mechanism(s) is determining the carrier type. Knowing the carrier type of the gate leakage current can eliminate consideration of various transport mechanisms as the cause of the gate leakage current. Identifying the type of carrier (i.e., electron or hole) that is contributing to the gate leakage current can be performed by carrier separation analysis [65-67]. Carrier separation analysis is performed by examining the currents in the source/drain and bulk terminals of the MOSFET, requiring the current through all four terminals of the MOSFET to be measured. In an nMOSFET device, the current through the source/drain

(n-type region) corresponds to electrons while the bulk current corresponds to holes (p-type region). The opposite is true in pMOSFETs where the source/drain corresponds to hole currents (p-type region) and the bulk current corresponds to electron current (n-type region). Carrier separation analysis is primarily performed with the MOSFET in inversion where recombination is minimal.

With the MOSFET in inversion, minority carriers from the gate are captured by the channel and collected by the source/drain terminals. Majority carriers from the gate quickly drift away from the channel to the bulk due to high electric fields in the semiconductor channel and are collected by the bulk terminal. With the MOSFET in accumulation, minority carriers from the gate are not collected by the inversion channel (an inversion channel is not present in accumulation) and diffuse to the drain/source terminals. Minority carrier diffusion to the source/drain may take considerable time, especially if the device is large (as is the case in this study). During diffusion, the minority carrier may recombine with majority carriers, giving rise to an increase in the bulk current rather than the source/drain. It therefore becomes difficult to know if the bulk current is due to majority carriers from the gate or from recombination with minority carriers. Thus, it is important to measure both n- and p-MOSFETs to characterize carrier transport for positive and negative voltages.

The terminal currents for the gate (black lines), source/drain (green lines, which have been added together mathematically to plot the total minority carrier current), and bulk (red lines) are shown for 3nm HfO_2/1.1nm SiO_2 n/pMOSFET devices in Figure 5.13 and 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs devices in Figure 5.14.

5.2.1: Positive Gate Biases

For both 3nm and 5nm HfO_2 nMOSFET devices in inversion (Figure 5.13a and Figure 5.14a), the gate current overlaps with to the source/drain (n-type regions) current implying the gate leakage current is due to substrate electron injection. Within the accumulation operating region of 3nm and 5nm pMOSFET devices (Figure 5.13b and Figure 5.14b), where recombination is occurring, the gate and bulk current (n-type region) are superimposed, indicating electron substrate injection. Hence, the gate leakage current for both the pMOSFET and nMOSFET devices is due to substrate electron injection.

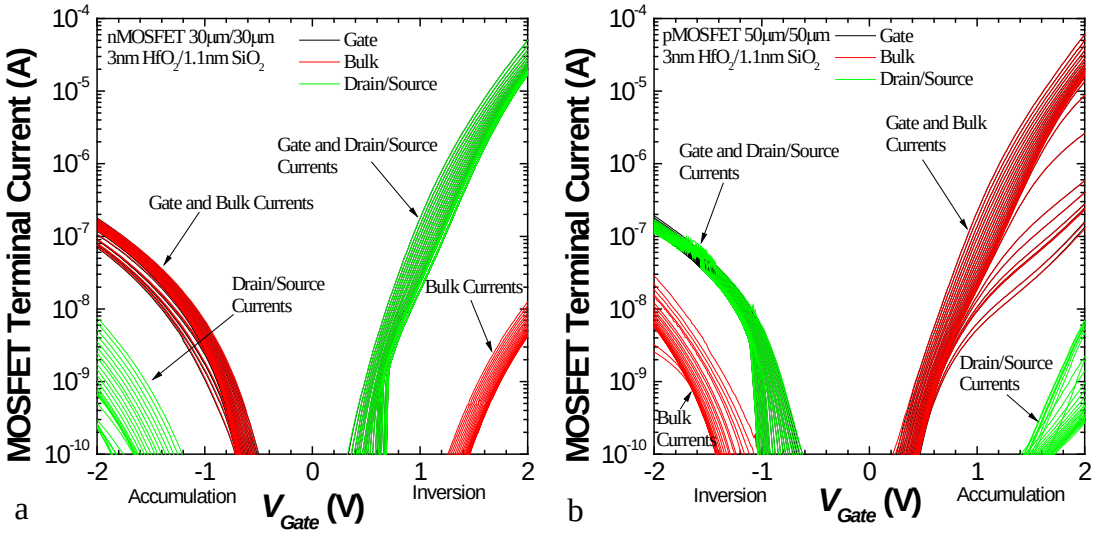


Figure 5.13: Currents monitored from each terminal for (a) an nMOSFET and (b) pMOSFET composed of 3nm HfO_2/1.1nm SiO_2. Note that many of the currents are superimposed.

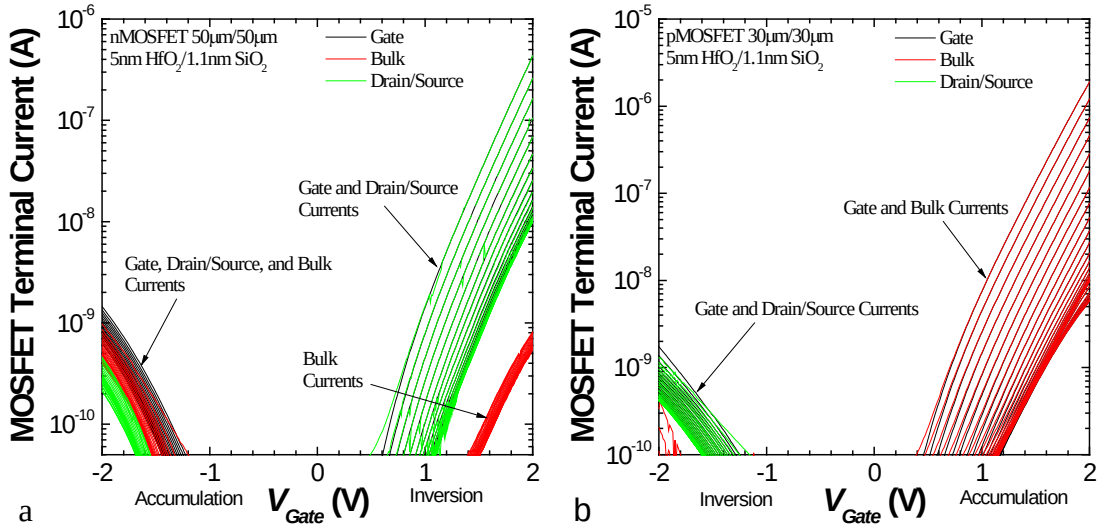


Figure 5.14: Currents monitored from each terminal for an (a) nMOSFET and (b) pMOSFET composed of 5nm HfO_2/1.1nm SiO_2. Note that many of the currents are superimposed.

5.2.2: Negative Gate Biases

The gate current and source/drain (p-type region) currents of 3nm and 5nm pMOSFETs in inversion (Figure 5.13b and Figure 5.14b) correspond very well, indicating that substrate hole injection is occurring at negative gate biases. Analyzing the 3nm HfO_2 nMOSFET currents in accumulation (Figure 5.13a), where electron-hole recombination is occurring, shows the gate and bulk (p-type region) currents are superimposed, suggesting substrate hole injection is occurring. For the 5nm HfO_2 nMOSFET device in accumulation (Figure 5.14a), the gate current is more equivalent to the bulk current than the source/drain current, which also suggests substrate hole injection. To a lesser extent in the 5nm HfO_2 nMOSFET (Figure 5.14a) but still noticeable is the large source/drain current in relation to the gate current, which indicates that perhaps electrons play a role in the gate leakage currents at negative gate biases in 5nm HfO_2 devices. Some evidence of this is also observed in the 5nm HfO_2 pMOSFET

(Figure 5.14b) in inversion, where the electron current (bulk current) is less than a decade smaller than the gate leakage current.

In summary, for positive gate biases of 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs, the gate leakage current is dominated by substrate electron injection and for negative gate biases substrate hole injection. In the 5nm HfO_2 devices in the negative gate bias regime, the gate electron current begins to have an increasing role in the gate leakage current as evidenced by Figure 5.14a, where the electron current nearly corresponds to the gate current. This analysis differs from the work of Shaimeev *et al.*, on HfO_2/SiO_2 gate stacks where carrier separation analysis results in electron current for both positive gate biases and negative gate biases [90]. It is important to note that the metal gate used by Shaimeev *et al.* was aluminum (Al), which has a work function of 4.1eV while the metal gate used in this study is TiN with a work function of 4.45eV. The smaller work function of Al reduces the conduction band offset of the metal gate with HfO_2, which would increase the access of metal electrons to defects near the HfO_2 conduction band edge or to the conduction band itself through FN tunneling. These two examples highlight the effect that the difference in the work function of the metal gate has on the gate leakage current and the type of carrier involved in the carrier transport.

5.3: Arrhenius Results and Discussion

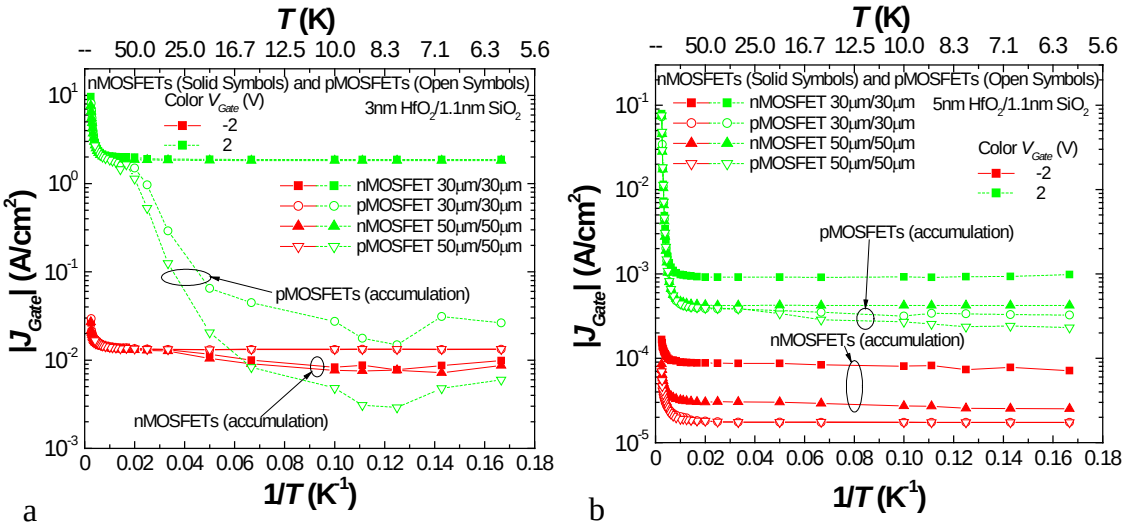
Many temperature dependent mechanisms follow an Arrhenius expression of the form (expressed here in terms of the gate current density):

$$J_{Gate}(T) = Ae^{-E_A/k_BT}, \tag{5.1}$$

where A and E_A represent a constant and activation energy. Determining the activation energy from (5.1) is performed by linearizing the equation (shown in the form of $y = mx + b$) as:

$$\ln(J_{Gate}(T)) = -\frac{E_A}{k_B}\frac{1}{T} + \ln(A), \tag{5.2}$$

Equation (5.2) illustrates that the activation energy can be derived from the slope of the data plotted as the natural log of J_{Gate} versus $1/T$, which is called an Arrhenius plot. Arrhenius plots[9] of the 3nm and 5nm HfO_2 devices for both nMOSFETs and pMOSFETs are shown in Figure 5.15 while Arrhenius plots of the 2nm SiO_2 control n/pMOSFETs are shown in Figure 5.16.



a b

Figure 5.15: Arrhenius plot of the gate leakage current density at gate biases of -2V and 2V for (a) 3nm HfO_2/1.1nm SiO_2 devices and (b) 5nm HfO_2/1.1nm SiO_2 devices.

[9] Notice that the Arrhenius plots shown (e.g., Figure 5.15) are plotted on a log base 10 scale and not on a natural log scale as described in (5.2). The log base 10 scale is used to better identify the gate current density values. Plotting the data on a log base 10 scale maintains the linearity of Arrhenius data but changes the offset and slope by a multiple of log(e). These differences are accounted for when extracting the activation energy.

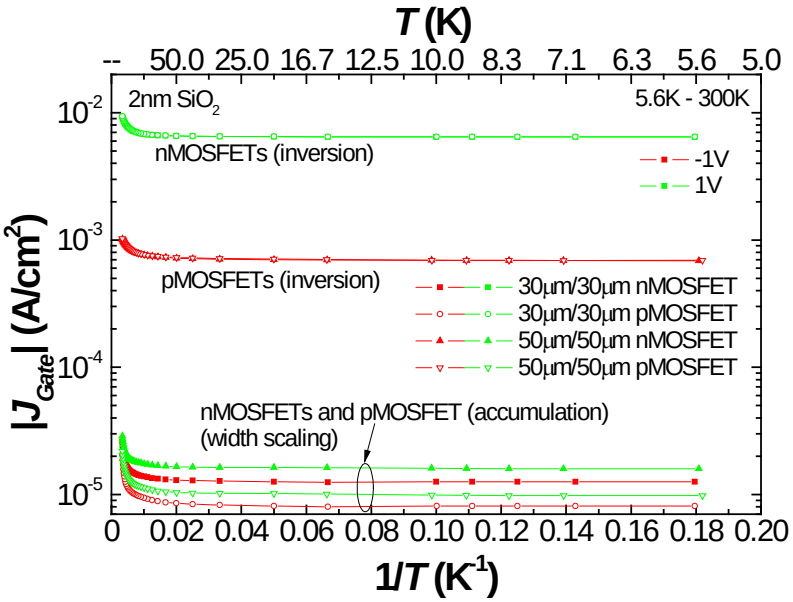


Figure 5.16: Arrhenius plot of the gate leakage current density of 2nm SiO_2 control device for n/pMOSFETs for gate biases of -1V and 1V.

Observed in Figure 5.15 and Figure 5.16 are two regimes that appear linear at first: 1) a weak temperature dependent regime is observed for $T < 50$K with μeV activation energies, and 2) a strong temperature regime is observed to $T > 150$K. The presence of multiple slopes in an Arrhenius plot is an indication that multiple processes are occurring [45]. The Arrhenius behavior of J_{Gate} (Figure 5.15 and Figure 5.16) reveals that the two ranges in temperature in which the slopes are significantly dissimilar suggests that two different transport regimes are present. This behavior is observed for all three gate stacks examined: 3nm HfO_2/1.1nm SiO_2, 5nm HfO_2/1.1nm SiO_2, and SiO_2 MOSFETs. Two transport regimes were also observed by Compagnoni *et al.* [11] in which they examined transport over a temperature range of 77-300K in Al/5nm HfO_2/0.87nm SiO_2/Si MOS capacitors. They report activation energies ~87meV for temperatures higher than 165K and 5meV for the lower temperature regime of 77-165K. In this study, a lower activation energy of ~1μeV is found, indicating a very weak

temperature dependence and most likely a non-Arrhenius behavior because the energy is unrealistically low. Of interest is to determine the activation energy in the low temperature range of Compagnoni *et al*. (77-165K). Figure 5.17 shows the 5nm HfO_2 data of this study in a temperature regime down to which Compagnoni *et al*. investigated (i.e., 77K). In this regime, an activation energy of 2.7meV is determined for the low temperature regime, which is consistent with the activation energy reported by Compagnoni *et al*. For the high temperature regime investigated by Compagnoni *et al.*, an activation energy of 66meV is obtain, similar to the 87meV reported by Compagnoni *et al*.. The much larger activation energy of 2.7meV is explained by realizing that the range of temperatures used down to 77K to calculate the activation energy is in a transition region, which is not linear (compare temperature regions of Figure 5.15b and Figure 5.17).

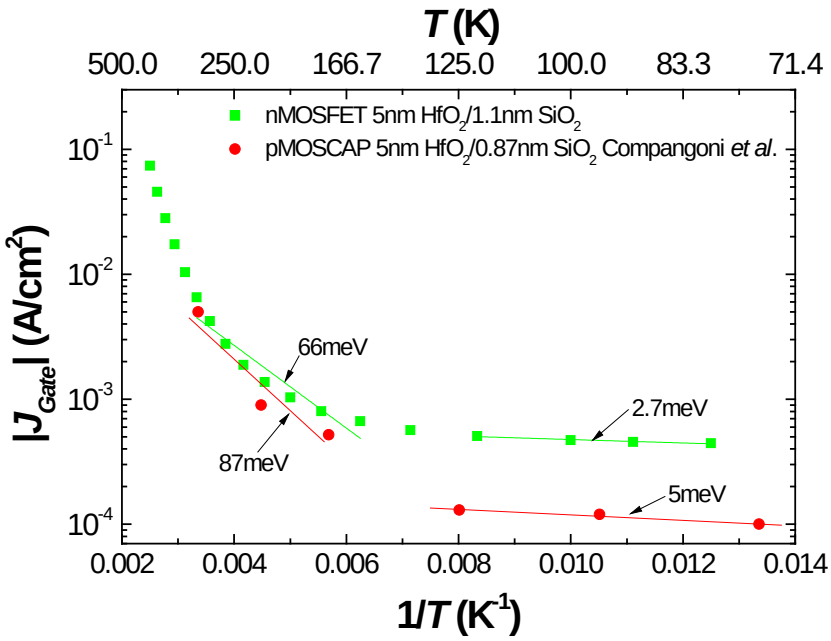


Figure 5.17: Arrhenius plot of the gate leakage current density for HfO_2/SiO_2 nMOSFETs. For gate leakage currents from ~125K to 77K, which appear linear, the weak temperature regime has an activation energy of 2.7meV. As the temperature range in this work continue down to 5.6K, ~125K to 77K is in a transition region and not linear.

The increments of temperature used in the study by Compagnoni *et al*. for the wide transition regime are large, thereby leaving the temperature range not well defined. The data in Figure 5.17 suggests that room temperature is part of the transition regime as it has significant curvature. The derivative of $\ln(J_{Gate})$ versus $1/T$ shows an ever increasing slope as the temperature increases, which indicates that the data does not follow a strict Arrhenius law in the temperature regime $T < 400$K. The activation energy extracted from the Arrhenius plot therefore depends on the temperature range used. To illustrate this effect, Figure 5.18 shows the activation energy versus gate voltage calculated using four different temperature ranges for 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs. Extracted activation energies in Figure 5.18 are only shown for large gate biases where MOSFET effects of depletion and weak inversion do not distort trends. The higher the temperature range used to extract the activation energy,

the larger the activation energy. The 5nm HfO_2 samples show a larger activation energy compared to the 3nm HfO_2 samples. In general, the activation energies of both n/pMOSFETs are similar for the various temperature ranges investigated (Figure 5.18).

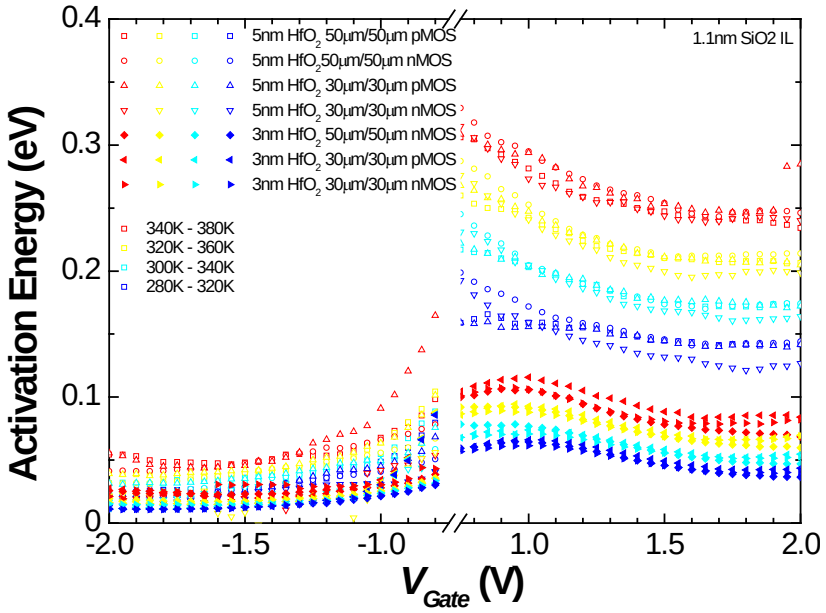


Figure 5.18: Activation energy versus gate voltage for various n/pMOSFETs composed of a bilayer of 3nm or 5nm HfO_2 on 1.1nm SiO_2. The activation energy is dependent on the temperature range used to extract it. The higher the temperature range used, the larger the calculated activation energy.

The Arrhenius plots of 3nm and 5nm HfO_2 devices in accumulation (Figure 5.15) show additional slope changes and thus indicate additional processes are occurring for temperatures less than 50K [10]. The 3nm HfO_2 samples clearly show both n/pMOSFETs having nearly the same temperature dependence until temperatures fall below 50K where devices in accumulation have low leakage currents. It is believed that the lack of thermally generated majority carriers from the donor/acceptor energy levels at temperatures below ~50K limit the amount of carriers available for transport through the gate stack. This idea is similar to the *carrier-limited* and *transport path-limited* regimes discussed in Section 5.1.1. Note that for devices in accumulation, majority carriers are the

dominate contribution to J_{Gate}, according to carrier separation analysis (Section 5.2), which indicates that substrate electrons dominate transport at positive gate biases (pMOSFETs in accumulation) and substrate holes dominate carrier transport at negative gate biases (nMOSFETs in accumulation). The largest change in the gate leakage current as temperature decreases below 50K is observed for 3nm HfO_2 pMOSFETs in accumulation. C-V data on a 3nm HfO_2 pMOSFET to determine the response of majority carriers as a function of temperature are shown in Figure 5.19. Similar to the gate leakage current (Figure 5.2), the capacitance in accumulation decreases as the temperature decreases below 50K (Figure 5.19). For the measured temperatures of 25K and 5.6K, a significant reduction in the capacitance signifies that the majority carriers are not responding to the capacitance AC probe frequency. The absence of a majority carrier response in the C-V measurements at temperatures below 50K further justifies the idea that a limited number of majority carriers in accumulation are limiting the gate leakage current in accumulation.

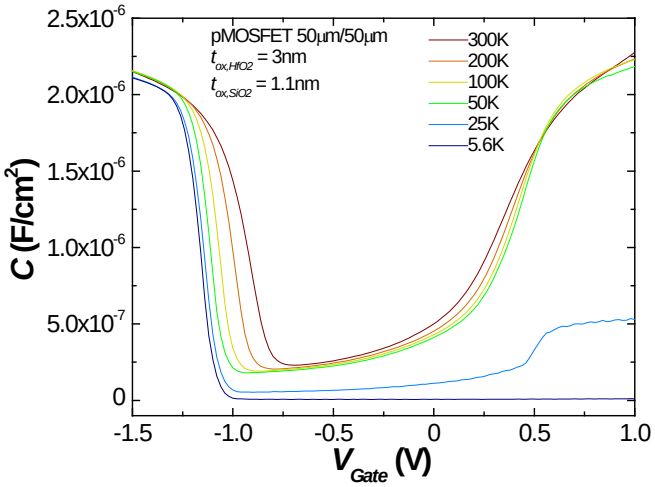


Figure 5.19: Capacitance-Voltage response of a 3nm HfO_2/1.1nm SiO_2 pMOSFET for temperatures ranging from 5.6K to 300K. A large decrease in the capacitance is observed in the accumulation regime as temperature decreases below 50K.

In addition to the temperature dependence of extracted activation energies, the activation energy also varies with gate voltage for both the HfO_2 samples as well as the SiO_2 control samples; see Figure 5.18 and Figure 11 in [10]. The gate voltage dependence of the activation energy (Figure 5.18) suggests an electric field dependence. Typical electric field correction factors to the activation energy are Poole-Frenkel and Schottky emission as was discussed in Chapter 2. Due to the extensive use of Pool-Frenkel and Schottky emission as explanations to the temperature dependence of the gate leakage current in high-*k* devices, Chapters 6 and 7 focus on examining the data using Poole-Frenkel and Schottky analysis.

CHAPTER 6: POOLE-FRENKEL ANALYSIS

[10]Poole-Frenkel (P-F) conduction has been a popular explanation of the temperature dependence of the gate leakage current of HfO_2 in the literature for high temperature measurements (300K to ~500K) [7, 21, 92, 93]. Through the use of P-F analysis, various trap energy barrier heights have been reported (e.g., 0.35eV [21], 0.68eV [20], 1.11-1.36eV [33], 1.5eV [7]) for HfO_2. The reported trap barrier heights may differ due to a variety of reasons, which may include different growth methods, pre- and post-growth processing, and device technologies used. In this chapter, P-F analysis is applied to the gate leakage currents measured in the high temperature region (T > 150K). Methods of extracting the electric field in the HfO_2 layer are discussed and the effect that various electric field calculation techniques have on the extracted trap barrier height are presented. Using the wide range of temperatures available and two different HfO_2 thicknesses, the standard P-F model is used to examine the trap barrier height versus both temperature range and HfO_2 thickness. In short, this chapter examines whether or not the standard P-F model can describe charge transport over a wide temperature range in HfO_2-SiO_2 bilayer dielectric stacks (6K – 400K).

6.1: Poole-Frenkel Analysis Techniques

The standard P-F conduction mechanism was described earlier in Section 2.3. P-F conduction does not describe the kinetics of carrier capture. It only considers carrier

[10] The main points and some content discussed in this chapter are taken from work published in [91].

emission. In order P-F conduction to dominate, thermal stimulation of the carrier over the reduced barrier height must be the limiting mechanism. Traps are usually considered to be filled through quantum mechanical tunneling [32]. P-F conduction also stipulates the nature of the traps involved in the conduction process due to its derivation of a hydrogenic-like state. That is, according to Frenkel, for trapping to occur, the impurity must be ionized in the non-trapped state and neutral in the trapped state [51]. P-F also assumes negligible interaction between traps, hence the trap concentration is considered low. An example depicting the P-F conduction path in 5nm HfO_2 with a 1.1nm SiO_2 interfacial layer is shown in Figure 6.1.

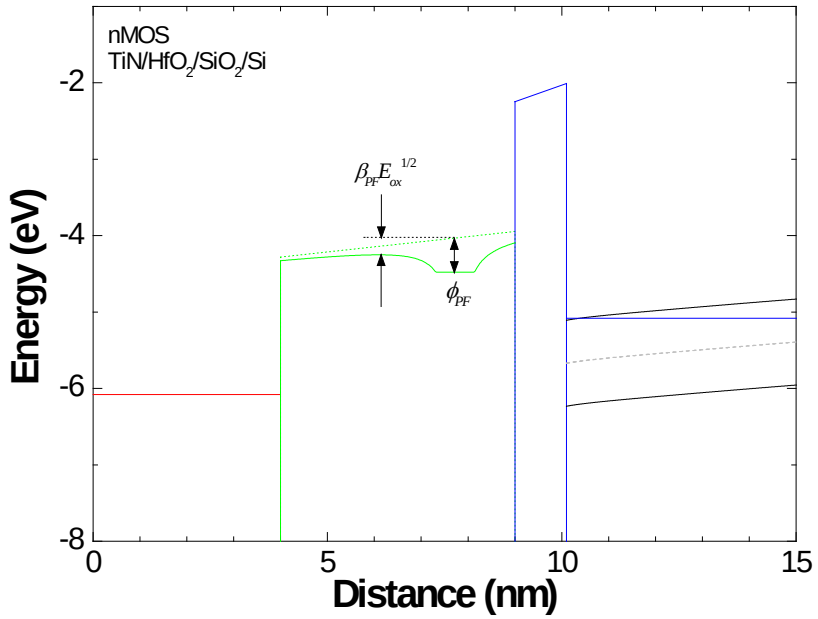


Figure 6.1: Energy band diagram of 5nm HfO_2/1.1nm SiO_2 nMOS device with (solid green line) and without (dotted green line) a hydrogenic trap. The diagram was created using [40], a ϕ_{PF} of 0.45eV, and a relative high frequency dielectric constant of 7.8; both values were extract from the data**.** After [91].

A common method to determine if a model describes data is by linearizing the model. New input and output parameters are derived using the linearized model and plotted; if the result is linear, a case of the model can be made. This type of plot for P-F

transport is called a P-F plot where (2.8) is linearized by plotting $\ln(J/E_{ox})$ vs. $E_{ox}^{1/2}$. The high frequency dielectric constant can be extracted from slope of the line, which yields $\beta_{PF}/\xi k_b T$ if ξ is known; however, usually it is assumed to be 1. The trap barrier height is extracted from the *y*-intercept:

$$\ln(C) - \frac{q\phi_{PF}}{\xi k_b T}. \tag{6.1}$$

To eliminate *C*, data from more than one temperature can be used. An alternative approach to extracting ϕ_{PF} is achieved by following [20, 57] in which (2.8) is linearized against the inverse of temperature instead of $E_{ox}^{1/2}$ by:

$$\ln\left(\frac{J_{PF}}{E_{ox}}\right) = -\frac{q\phi_{PF} - \beta_{PF}\sqrt{E_{ox}}}{\xi k_b}\frac{1}{T} + \ln(C). \tag{6.2}$$

Equation (6.2) predicts Arrhenius behavior. The slope of the resulting line is proportional to the reduced trap barrier height (ϕ_r), $q\phi_{PF}$-$\beta_{PF}E_{ox}^{1/2}$, which decreases as the electric field increases. The trap barrier height is found by extracting ϕ_r at various electric field strengths and then extrapolating ϕ_r to an electric field of zero, which yields ϕ_{PF}. This latter approach to finding ϕ_{PF} is the method used in this study.

6.2: Electric Field Calculation

Both of the methods discussed above for calculating ϕ_{PF} requires knowing the electric field. To extract the trap barrier height in the HfO_2[11], the electric field in the HfO_2 ($E_{ox,HfO2}$) should be used [10]. Various methods for calculating the electric field have been used in the literature (e.g., [16, 20, 32, 38, 41]). Simple methods used for calculating the

[11] Assumes the traps are located in the HfO_2 layer which is a reasonable assumption.

electric field have the advantage of being quick and easy. However, simple approximations to the electric field compared to more accurate calculations using CV modeling programs, finite element analysis tools, or the energy band simulation tool discussed in Chapter 4, can lead to significant differences [10, 91]. The large differences in the calculated values of the electric field has a substantial impact on the calculated ϕ_{PF} [91].

The electric field in a MOS device at a particular bias is not constant with temperature. The temperature dependence of the electric field depends on multiple factors, a primary one being the metal-semiconductor work function difference. However, for large doping concentrations[12], minimal variation in the electric field with temperature (10K to 400K) is observed in simulations [10]. This can be predominantly attributed to the negligible change in the semiconductor work function with temperature for large doping concentrations. Because the electric field remains nominally unchanged with respect to temperature, the temperature dependence of the electric field is assumed constant in this work.

As was mentioned above, a variety of methods exist to accurately calculate the electric field in the oxide of MOS devices. Many of these tools are proprietary, costly, and/or support only single layer dielectric oxides. The energy band simulation tool (Chapter 4) supports modeling the electric field in multiple dielectrics, is free [94], and is used in this study to extract the electric field versus gate bias. The energy band simulation tool was calibrated to the devices used in this study using the flatband voltage and materials parameters extracted by SEMATECH [64].

[12] In this work, a doping concentration of $1x10^{18}$ cm^{-3} is used for all simulations

6.3: Poole-Frenkel Analysis Results

Following the second method for calculating ϕ_{PF} described in Section 6.1, in which (6.2) is used, the gate leakage current (e.g., Figure 5.1 and Figure 5.2) is examined for relatively large positive and negative gate biases to examine the gate leakage current in the transport-limited regime (see Section 5.1.1). Recall from Section 5.2 that the gate leakage current corresponds to substrate electron injection for positive gate biases and to substrate hole injection for negative gate biases. In the following subsections, P-F analysis is first applied to electron substrate injection and then to hole substrate injection.

6.3.1: Electron Substrate Injection (Positive Gate Biases)

A representative P-F plot for electron substrate injection is shown in Figure 6.2. As was discussed in Section 6.1, (2.8) predicts P-F data should be linear on a P-F plot, $\ln(J_{Gate}/E_{ox,HfO2})$ vs $E_{ox,HfO2}^{1/2}$. The plot of the data in Figure 6.2 is not linear over the entire electric field range shown as it has a slight curvature to it. The data in Figure 6.2 can approach linearity if a smaller electric field range is specified. However, by doing so, the range P-F emission can explain the temperature dependence of the gate leakage current is severely limited, which brings into question the validity of the P-F model.

One method to examine the deviation from linearity in Figure 6.2 (i.e., the soundness of the P-F model) is to evaluate approximations used to derive the classic P-F equation (2.8). Classic P-F was derived by approximating Fermi-Dirac statistics with Boltzmann statistics. For large electric fields where the reduced trap barrier height is small, Boltzmann statistics fail to describe the thermal statistics causing non-linearities in the P-F plot; this occurs near the saturation voltage. A saturation P-F analysis approach

has been proposed in [95] and developed in [56, 96] in which Maxwell-Boltzmann statistics are replaced by Fermi-Dirac statistics to describe non-linear (e.g., saturated) P-F data. At the saturation voltage, the reduction of the barrier trap height is equal to the trap barrier height. Mathematically, this occurs when $\Delta\phi_{PF}$ in (2.6) equals ϕ_{PF}. When $\Delta\phi_{PF}$ is less than ϕ_{PF}, carriers are thermally ionized over the reduced trap barrier, which gives rise to a temperature dependence. When $\Delta\phi_{PF} = \phi_{PF}$, no barrier exists - thermal assistance is not needed for the carrier to leave the trap and the gate leakage current should not change with temperature at the saturation voltage. Saturation P-F therefore dominates when the conductivity ($J_{Gate}/E_{ox,HfO2}$) for various temperatures approaches a crossing point. In this study, the conductivity data in Figure 6.2 is not observed to cross or approach crossing, indicating that the P-F saturation is not occurring. Therefore, saturation P-F analysis for the non-linear data observed in Figure 6.2 is not applicable.

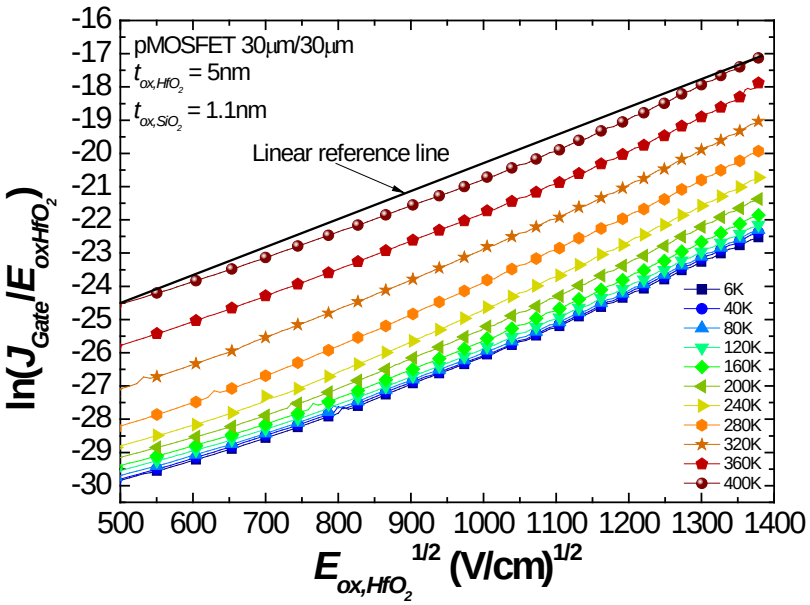


Figure 6.2: P-F plot of the gate leakage current of a 5nm HfO_2 1.1nm SiO_2 pMOSFET for temperatures ranging from 6K to 400K. Most temperatures are not plotted to increase clarity of the data. A linear reference line has been added to the plot to demonstrate curvature in the data. Similar to [91].

The deviation from linearity could be due to the other temperature dependent mechanisms that do not dominate but influence the temperature dependence of the gate leakage current.[13] P-F conduction has been cited many times in the literature as the reason for the temperature dependence of the gate leakage current (e.g., see Chapter 6 introduction) and merits an in-depth discussion of its application to the temperature data in this study. Therefore, continuing the analysis to verify whether or not the P-F transport follows the Arrhenius-type behavior predicted by (6.2), the conductivity ($J_{Gate}/E_{ox,HfO2}$) is plotted versus the inverse temperature at various electric fields as shown in Figure 6.3. For three temperature ranges in the strong temperature dependent regime, ϕ_r at various electric field strengths is extracted, assuming $\xi = 1$, and plotted against the square root of the electric field (Figure 6.4) in order to extract ϕ_{PF}. The data in Figure 6.3 are not completely linear and thus is another indication that the behavior is not standard P-F. The slope of the data is observed to increase with increasing temperature to the highest temperature (400K) for which measurements were performed. Therefore, ϕ_r extracted from the data will be dependent on the temperature range used as demonstrated in Figure 6.4. Note that a nonlinear Arrhenius plot (Figure 5.17) was observed and discussed in Section 5.3.

[13] Other temperature dependent mechanisms might include Schottky emission, trap assisted tunneling, etc. As will be shown later, it turns out that the standard P-F mechanism does not adequately explain the data for reasons other than a non-linear P-F plot.

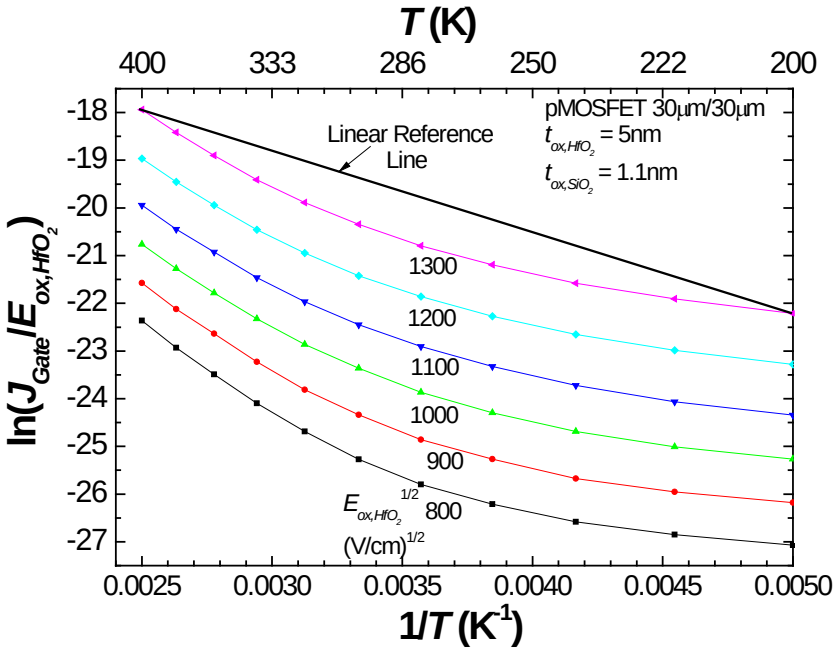


Figure 6.3: Arrhenius P-F plot at various electric field strengths of a 5nm HfO_2 1.1nm SiO_2 pMOSFET for temperatures 200K to 400K. The data is nonlinear as demonstrated by the linear reference line. Similar to [91].

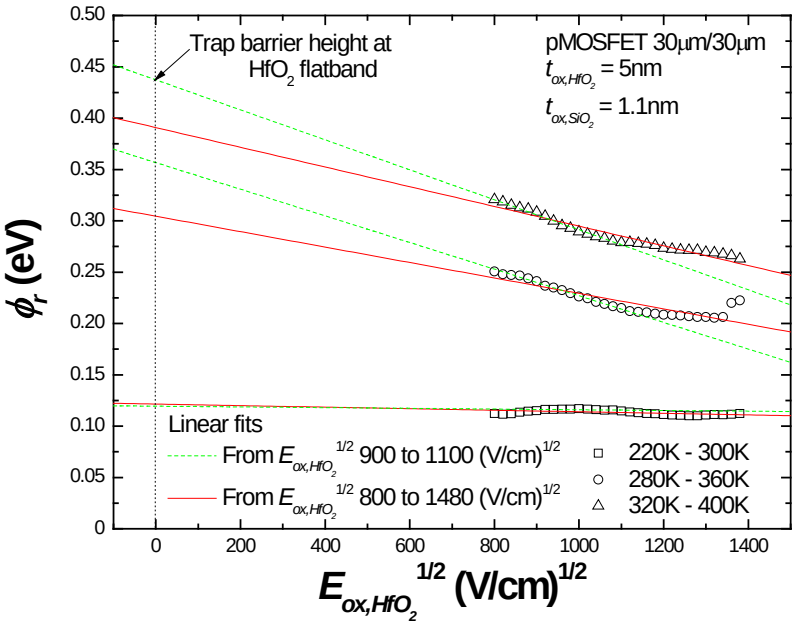


Figure 6.4: Reduced trap barrier height (ϕ_r) at various electric field strengths using three different temperature ranges from e.g. Figure 6.3. ϕ_r at flatband (vertical dashed line), or ϕ_{PF}, is determined by linear extrapolation, e.g. red (solid) and green (dashed) lines. Similar to [91].

Since the data in the Arrhenius P-F plot in Figure 6.3 show a slight curvature with the slope increasing as temperature increases, it is evident that (6.2), obtained by linearizing (2.8), cannot be applied to the data. Equation (2.8) was derived assuming a single trap energy level; hence, the data does not exhibit single trap energy level behavior. The observed nonlinearity of the data may stem from several possibilities. One possibility is that several series of linear data with dissimilar slopes (i.e., various trap energy levels) may exist and combine to form non-linear data. Another possible explanation may be that the wave functions of the traps overlap due to a high defect density. Overlapping of the wave functions would lead to the potentials of the traps overlapping, which may broaden the distribution of trap energy barrier heights (ϕ_{PF}'s) [97]. An additional possibility might be attributed to the presence of charged traps (not considered in P-F conduction) taking part in carrier transport, which is indicative of multivalent traps, which is supported by theoretical calculations showing the five charge states of oxygen vacancies in HfO_2 [98].

Due to the curvature of the data, only a limited range of data can be analyzed and explained using the standard P-F model. This finding highlights two uncertainties with the P-F model. First, if the data are linear over only a narrow range of temperatures, the extracted ϕ_{PF} seems physically tenuous. Second, when a narrow range of temperatures is selected for analysis, then the temperatures outside the narrow temperature range are not being considered, thereby indicating that other ϕ_{PF}*'s* exist.

Regardless of the cause, the curvature in the data in Figure 6.3 influences the extracted reduced trap energy barrier height (ϕ_r) depending on the temperature range used to perform the linear fit. Figure 6.4 shows ϕ_r extracted for a 5nm HfO_2 on 1.1nm SiO_2

nMOSFET using three different temperature ranges. To obtain ϕ_{PF}, ϕ_r is extrapolated to $E_{ox,HfO2} = 0$ shown by the dashed vertical line in Figure 6.4. As depicted in Figure 6.4, the temperature range used for calculating ϕ_r, as well as the range of electric field strengths to perform the linear regression fit to extrapolate ϕ_{PF}, yields significantly different results. The green (dashed) lines in Figure 6.4 correspond to extracting ϕ_{PF} using an electric field range of 900 to 1100 $(V/cm)^{1/2}$. The red (solid) lines in Figure 6.4 correspond to a linear fit using electric field strengths ranging from 800 to 1480 $(V/cm)^{1/2}$. Both red and green lines illustrate the wide distribution of ϕ_{PF}'s that can be extracted. The standard P-F model predicts a constant slope for data plotted in the manner of Figure 6.4. The heterogeneity of the slope in Figure 6.4 suggests that P-F conduction is not an appropriate model or is limited to a small electric field range.

As was shown in Figure 5.1 and Figure 5.2, the temperature dependence of the 5nm HfO_2 devices differed considerably from the 3nm HfO_2 devices. This temperature dependence difference is also observed in the P-F ϕ_{PF} extraction (Figure 6.5). Figure 6.5 shows the ϕ_r's for MOSFETs with 5nm HfO_2 and 3nm HfO_2. The extracted ϕ_{PF} for the n/pMOSFETs composed of 5nm HfO_2 is ~0.45eV below the HfO_2 conduction band for a low electric field fit and ~0.35eV below for a high electric field fit. The P-F analysis for n/pMOSFETs composed of 3nm HfO_2 results in a ϕ_{PF} of ~0.23eV below the HfO_2 conduction band. Therefore, the HfO_2 thickness affects the measured ϕ_{PF}. The similarly calculated ϕ_{PF}'s for n and pMOSFETs verify the same traps in the HfO_2 are responsible for the gate leakage current whether in nMOS or pMOS configuration and thereby independent of the Si doping and related processing. However, the P-F model does not

incorporate a dielectric thickness dependence and therefore cannot explain the data in Figure 6.5.

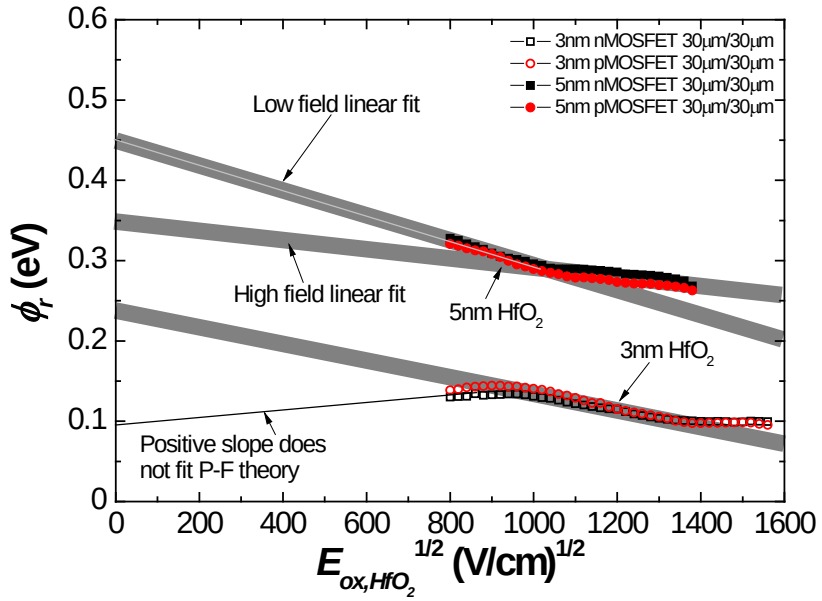


Figure 6.5: Electric field dependence of the ϕ_r (extracted from temperatures ranging from 320K to 400K) for 5nm HfO_2 and 3nm HfO_2 nMOSFETs. The thick grey lines represent linear fits to the data. For the 5nm HfO_2 devices, two linear regimes are observed and give different trap energy barriers. The thin black linear fit line for 3nm HfO_2 devices show P-F conduction is not valid in the low field regime as ϕ_r increases with the electric field instead of being reduced. Similar to [91].

A reason for the thickness dependence of calculated ϕ_{PF} may best be understood through Figure 6.1. Figure 6.1 shows the trap potential of a 0.45eV trap energy barrier with a relative high frequency dielectric constant of 7.8 (a high value) located in the 5nm HfO_2 layer. The value of the trap energy barrier was extrapolated from Figure 6.5 while the high frequency dielectric constant was extracted by calculating β from Figure 6.2 using 400K data and assuming $\xi = 1$. Because the P-F potential is based on the concept of a hydrogenic impurity, the Columbic potential between the electron and positive trap, shown in Figure 6.1, extends over a large distance that is comparable to the HfO_2 layer thickness. Depending on the location of the trap and thickness of the HfO_2, standard P-F

conduction may not dominate and the resulting transport mechanism becomes thickness dependent. P-F analysis, shown in Figure 6.5, indicates that the 3nm HfO_2 layer contains traps with lower ϕ_{PF}'s than the ϕ_{PF}'s extracted for the 5nm HfO_2. For the high field regime, the difference observed in ϕ_{PF}'s between 3nm and 5nm HfO_2 may result from the large delocalized trap potential (Figure 6.1) assumed by the P-F model via an effective mass approximation. A large delocalized trap potential similar in diameter to the thickness of the HfO_2 may reduce ϕ_{PF} more than what is predicted in (2.6). This supposition assumes that P-F conduction is prevalent. Conversely, the differences in observed ϕ_{PF}'s between 3nm and 5nm HfO_2 may be due to the presence of another carrier transport mechanism. Regardless of the possible explanations to describe the observed differences, the P-F model does not consider a dielectric thickness dependence and thus is inadequate to explain the differences observed in Figure 6.5.

For the 3nm HfO_2 devices shown in Figure 6.4 and Figure 6.5, ϕ_r for electric fields between ~700 and ~900 $(V/cm)^{1/2}$ increases with increasing field strength. If the governing carrier transport mechanism is P-F conduction, ϕ_r should decrease rather than increase with increasing electric field according to (2.6). Hence, for the 3nm HfO_2 devices in this study, it is evident that P-F transport does not explain the temperature dependence of the gate leakage current density for electric field strengths below ~900 $(V/cm)^{1/2}$ or gate voltages below 1V. Consequently, for typical MOSFET operation conditions, (V_{Gate} < ~1V), P-F transport is not dominant in 3nm HfO_2 and a different thermally activated carrier transport mechanism dominates.

Given the above analysis on electron substrate injection (positive gate biases), P-F conduction does not explain the temperature dependence of the gate leakage current.

6.3.2: Hole Substrate Injection (Negative Gate Biases)

P-F analysis is performed on 3nm and 5nm HfO_2 n/pMOSFETs at negative gate biases in much the same way as it was performed for positive gate biases. The only difference is in nomenclature; for negative gate biases, the electric field has a minus sign to indicate the polarity. This minus sign is removed before taking the square root of the electric field in (6.2) (i.e., no imaginary numbers are used) and replaced after calculating the square root, signifying the polarity (i.e., negative gate biases). This preservation of the polarity of the electric field is evident in the figures discussed below.

A P-F plot for negative gate biases is shown in Figure 6.6 that is typical for data obtained from 5nm HfO_2/1.1nm pMOSFETs. The P-F plot for negative gate biases show more linearity (see linear reference line in Figure 6.6) than exhibited by PF plots for positive gate biases (e.g., Figure 6.2). However, the Arrhenius P-F plot (Figure 6.7) is as nonlinear as the Arrhenius P-F plot for positive gate biases (Figure 6.3). The nonlinear Arrhenius P-F plot indicates that P-F conduction does not describe the temperature dependence of the gate leakage current over the temperature range investigated. Similar arguments given for the nonlinear Arrhenius P-F plot for positive gate biases can also be presented here (see discussion of Figure 6.3).

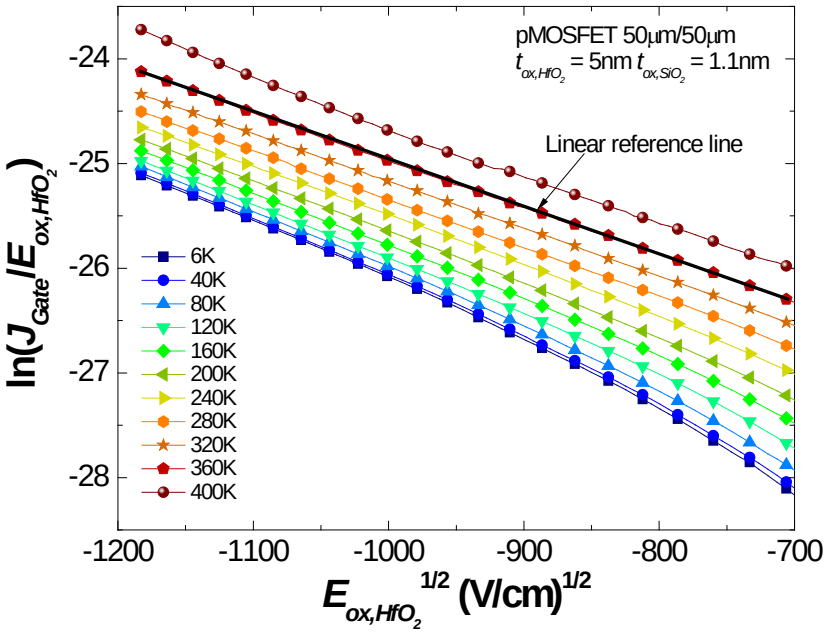


Figure 6.6: P-F plot of a 5nm HfO_2/1.1nm SiO_2 pMOSFET for negative gate biases. The negative $E_{ox,HfO2}$ values on the *x*-axis indicate the polarity of the electric field. Most temperatures are not plotted to increase clarity of the data. A linear reference line has been added to the plot to demonstrate curvature in the data.

The curvature of the Arrhenius P-F plot (Figure 6.7) leads to multiple ϕ_{PF}'s depending on the temperature range used as was demonstrated for positive gate biases, shown in Figure 6.4. Figure 6.8 shows ϕ_r as a function of the square root of electric field for three temperature ranges: 320K to 400K, 280K to 360K, and 220K to 300K. The lower the temperature range used to extract ϕ_r results in a lower calculated ϕ_{PF}, as demonstrated by the linear extrapolation lines for the various temperature ranges shown in Figure 6.8. The extraction of multiple ϕ_{PF}'s depending on the temperature range used indicates that P-F does not describe the temperature dependence of the gate leakage current or at least does not dominate. The electric field range, over which the linear fit to extrapolate ϕ_{PF} is performed, influences the result since ϕ_r versus $E_{ox,HfO2}^{1/2}$ is nonlinear. An interesting feature in Figure 6.8 is that as $E_{ox,HfO2}^{1/2}$ increase above 1,000 (V/cm)$^{1/2}$ for the 5nm HfO_2/1.1nm SiO_2 devices, ϕ_r increases rather than decreases as is predicted by P-

F conduction. The increase in ϕ_r as $E_{ox,HfO2}$ increases is another indication that P-F conduction does not describe the gate leakage current.

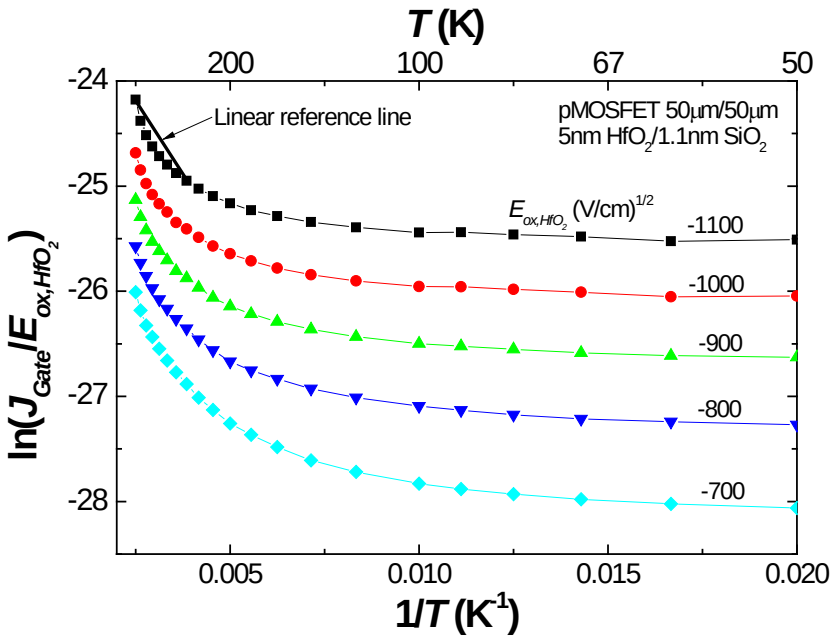


Figure 6.7: Arrhenius P-F plot at select electric field strengths of a 5nm HfO_2/1.1nm SiO_2 pMOSFET for temperatures 50K to 400K. The data is nonlinear as demonstrated by the linear reference line. The negative $E_{ox,HfO2}$ values indicate the polarity of the electric field.

The extracted values for ϕ_{PF} is ~0.11eV for the 5nm HfO_2/1.1nm SiO_2 devices extracted from a temperature range of 320K to 400K and ~50meV for the 3nm HfO_2/1.1nm SiO_2 devices over the same temperature range. The difference in the extracted ϕ_{PF} between 5nm and 3nm HfO_2 devices illustrates a limitation with P-F conduction as an explanation of the gate leakage current that also occurred when analyzing positive gate voltage currents. That is, the standard P-F conduction mechanism does not predict a thickness dependence; it is independent of thickness.

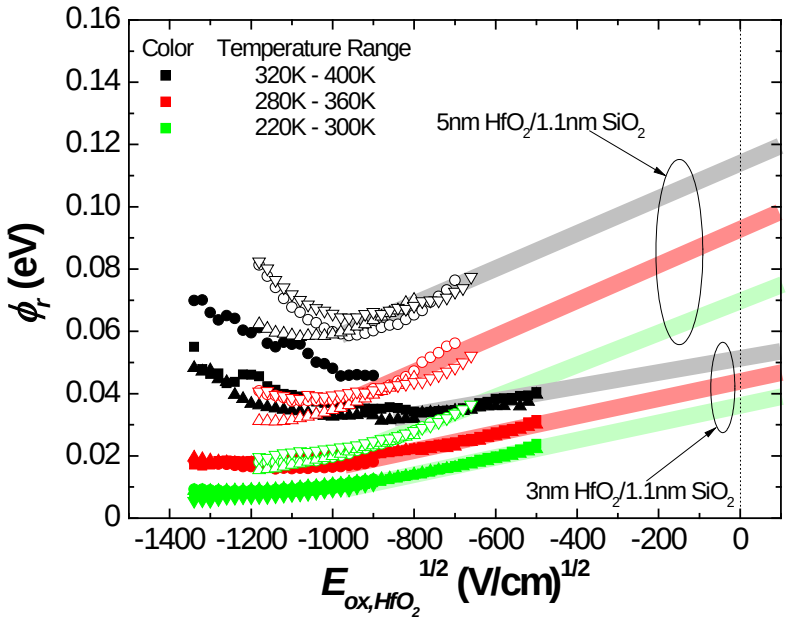


Figure 6.8: Electric field dependence of the ϕ_r (extracted from three temperature ranges) for 5nm HfO_2 (open symbols) and 3nm HfO_2 (closed symbols) n/pMOSFETs. The thick grey represent linear extrapolations to $E_{ox,HfO2}^{1/2} = 0$ for which ϕ_{PF}'s are extracted. For high electric fields, ϕ_r increases as $E_{ox,HfO2}$ increases instead of decreasing, indicating that P-F conduction does not describe the gate leakage data.

Additional evidence that suggests the data does not conform to the framework of the P-F model concerns the magnitude of the extracted value ϕ_{PF} (a maximum of ~0.11eV was extracted). Given the location of the trap (~0.11eV above the HfO_2 valance band), it is unlikely that the trap is filled over the voltage range of interest ($|V_{Gate}| < |-2V|$) due to the energy difference between the Fermi energy level in the Si (source of the hole carriers) compared to the trap energy level extracted by the P-F analysis performed here. The smallest energy level difference occurs when $V_{Gate} = -2V$ for which an energy band diagram is shown in Figure 6.9. An energy level difference of ~2.5eV and ~1.9eV exists between the defects and Si Fermi level at the HfO_2/SiO_2 and TiN/HfO_2 interface respectively at $V_{Gate} = -2V$, indicated by the arrows in Figure 6.9. The large energy difference between the extract trap energy and Si Fermi energy level (Figure 6.9)

indicates that capture of holes is unlikely for which thermal emission of the holes from the traps described by P-F is also unlikely. The large energy difference also calls into question whether thermal emission of holes from the trap (a small energy difference, Figure 6.9) is the limiting process for current conduction, which P-F conduction assumes. It would be more likely that hole capture by traps (Figure 6.9) is the limiting process for such a scenario. This is another indication that P-F conduction does not explain the temperature dependence of the gate leakage current.

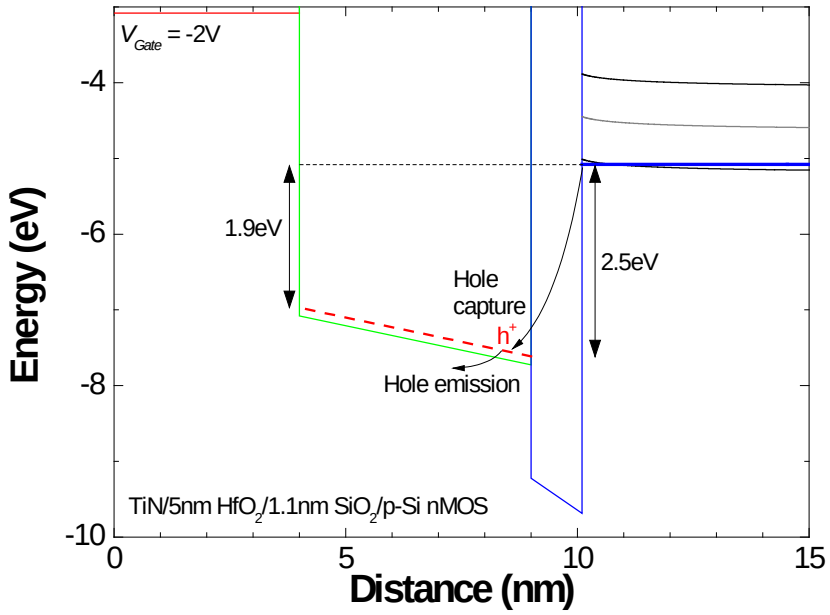


Figure 6.9: Energy band diagram showing the deepest extracted defect level (dashed red line, 0.11eV above the HfO_2 valance band) from the P-F analysis in comparison to the Fermi energy level in the Si (thick blue line). A large energy difference exists between the defect and Si Fermi energy level.

While P-F conduction has been used in the literature to explain the temperature dependence of the gate leakage current [7, 21, 92, 93], a thorough analysis of P-F conduction on the gate leakage temperature dependence indicates that for the devices and temperature range used, P-F conduction does not explain the gate leakage current for the

positive or negative gate bias regime. A variety of reasons were given for this conclusion, including:

1. For the positive gate bias regimes, the P-F plot is nonlinear over a large electric field range for positive gate biases (Figure 6.2).
2. The Arrhenius P-F plots are nonlinear (Figure 6.3 and Figure 6.7) for which many different trap barrier heights can be extracted depending on the temperature range examined. The standard P-F model was developed for mono-energetic defect levels for which a linear Arrhenius P-F plot is expected. The standard P-F model is therefore only able to describe a small temperature range.
3. Nonlinear ϕ_r's with respect to $E_{ox,HfO_2}^{1/2}$ were obtained for which different ϕ_{PF}'s can be extracted depending on the choice of electric field range used to perform the linear extrapolation (Figure 6.5). The small electric field range (which can then be approximated as linear) is evidence that the P-F conduction mechanism provides an inappropriate explanation to the gate leakage dependence.
4. A thickness dependence exists between the 3nm and 5nm HfO_2 devices for which the P-F model does not predict.
5. For the negative gate bias regime, the extracted trap energy level is non-sensical, yielding a shallow defect energy requiring a large thermal emission of holes from the Si to populate the energy level. This would result in a defect filling-limited transport mechanism and not thermal emission from the defect (P-F) as the limiting mechanism.

6. For 5nm HfO_2 devices at large negative biases ($|E_{ox,HfO_2}^{1/2}| > |-1000|$ $(V/cm)^{1/2}$) and for 3nm HfO_2 devices in the positive biases regime at small negative biases ($E_{ox,HfO_2} < 700$ $(V/cm)^{1/2}$), ϕ_r *increases* as the magnitude of the electric field increases. According to the P-F model ϕ_r should *decrease* as the magnitude of the electric field increases not *increase*.

Based on reasons summarized above, P-F does not explain the temperature dependence of the gate leakage current and other temperature dependent mechanism should be explored.

CHAPTER 7: SCHOTTKY ANALYSIS

Another popular carrier transport mechanism that may be used to explain the temperature dependence of the gate leakage current in HfO_2 devices is Schottky emission (e.g., [7, 23, 33, 34, 41, 61, 99]). However, Schottky emission discussed in the literature relative to HfO_2 only considers carrier emission into the HfO_2 dielectric conduction band from the gate electrode or substrate injection assuming an interfacial layer does not exists between the HfO_2 and substrate. For the devices used in this study, a 1.1nm SiO_2 interfacial layer is present between the HfO_2 and Si substrate, which must be considered when modeling thermionic charge injection from the substrate into the HfO_2 conduction band. In this chapter, a simple model taking into account the presence of the interfacial layer is developed. Application of the model is then applied to the gate leakage data of this study for both electron substrate injection and hole substrate injection.

7.1: Image Charge in multilayer dielectrics

An introduction to Schottky emission was given in Section 2.4 for a single dielectric. Schottky emission is the process in which the carriers are thermally stimulated into the conduction band of the dielectric. The energy required to reach the conduction band of the dielectric is decreased (reduced barrier height, ϕ_r) in the presence of an electric field due to the image charge theorem. The derivation of Schottky emission can be performed following [100] where the potential energy (PE) of the barrier (in units of Joules) as a function of distance is written as:

$$PE(x) = q\phi_B - qxE_{ox} - \frac{q^2}{16\pi\varepsilon_r\varepsilon_0 x}. \tag{7.1}$$

The first term in (7.1) is the barrier height (ϕ_B, in units of eV), which corresponds to the conduction band offset (for electrons) or valance band offset (for holes) if thermal stimulation is from a semiconductor. The first term could also refer to the metal work function dielectric conduction band offset (for electrons) or metal work function dielectric valance band offset (for holes) if thermal stimulation is from the Fermi energy of a metal. The second term considers the change in the potential energy in the presences of an electric field. The third term is the Coulomb potential due to the image charge theorem; see (4.11). Identifying the reduced potential barrier height as a function of the electric field is determined by finding the distance where the maximum of (7.1) occurs (denoted here as x_{max}) and then solving $PE(x_{max})$. The value x_{max} is determined by setting the derivative of (7.1) to zero and solving for x, which yields:

$$x_{max} = \frac{1}{4}\sqrt{\frac{q}{\pi\varepsilon_0\varepsilon_r E_{ox}}}. \tag{7.2}$$

The reduced trap barrier height, $PE(x_{max})$, can then be written as:

$$PE(x_{max}) = q(\phi_B - \sqrt{qE_{ox}/4\pi\varepsilon_r\varepsilon_0}). \tag{7.3}$$

Carriers are stimulated over the barrier by thermionic emission described by the Richardson equation [101, 102]:

$$J = A^* T^2 \exp\left(\frac{-q\phi}{k_b T}\right), \tag{7.4}$$

where A^* is the effective Richardson constant given in Section 2.4 and ϕ is the barrier height. The Schottky emission equation is found by replacing ϕ in (7.4) with the reduced barrier height (7.3), written here again for convenience:

$$J_{SE} = A^* T^2 \exp\left(\frac{-q(\phi_B - \sqrt{qE_{ox} / 4\pi\varepsilon_r\varepsilon_0})}{k_b T} \right). \tag{7.5}$$

For thick single layer dielectrics, (7.3) is a good approximation to the barrier height. However, as the layer thickness decreases, the image charge effects of both conductors becomes more important. An example of a thin 1.2nm SiO_2 barrier considering the image charge effects of only the semiconductor and both the semiconductor and metal is shown by the energy band diagram in Figure 7.1 using the Image Charge Model outlined in Section 4.1.4 and a relative high frequency dielectric constant of 2.38 for the SiO_2 layer [103]. Figure 7.1 shows a noticeable decrease in the barrier height (~128meV) when image charge effects of both the metal and semiconductor are considered compared to only the semiconductor.

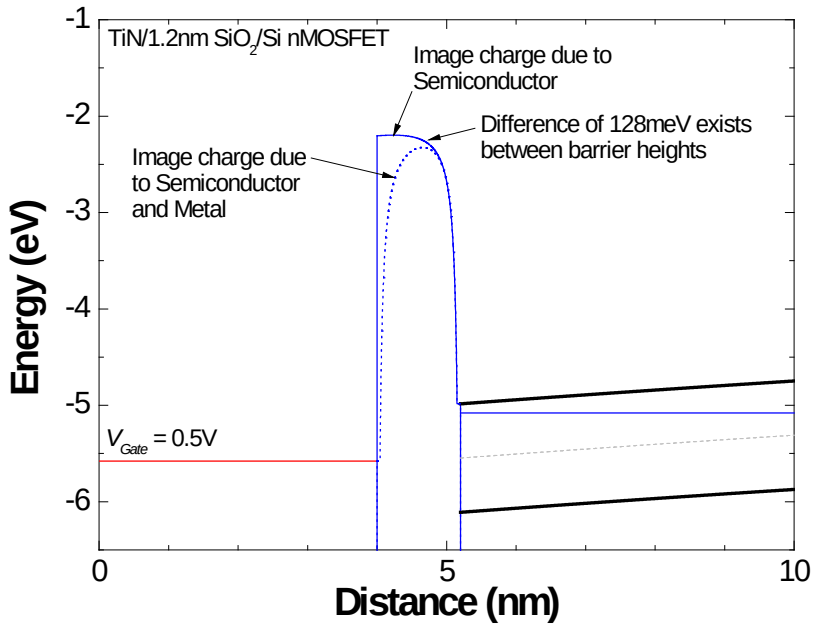


Figure 7.1: Energy band diagram of a TiN/1.2nm SiO_2/Si nMOSFET considering image charge effects of only the semiconductor and both the semiconductor and metal. A difference of ~128meV exists between the barrier heights of the two calculations. Created using [40].

Similar to ultra thin dielectrics, where the effects two conductors must be considered, multilayer dielectrics also present situations for which Schottky emission does not describe thermionic emission. For example, if another dielectric has a larger conduction band offset than the dielectric adjacent to the conductor from which carrier injection is occurring (e.g., metal-SiO_2 band offset is greater than that of metal-HfO_2), (7.5) may not apply. The maximum barrier height of the gate stack, for which thermionic conduction should be considered, may not be due to the image charge interaction with the first dielectric (i.e., the dielectric adjacent to the conductor). While image charge effects on the first dielectric does modify the maximum barrier height of the first dielectric, the band offset of another dielectric with respect to the conductor may be large enough that the maximum barrier height of the gate *stack* is controlled by the barrier height of the non-adjacent dielectric. An example of the second dielectric being the maximum barrier

height is illustrated using a HfO_2 (with a high frequency dielectric constant of 4, [104]) SiO_2 bilayer shown in Figure 7.2. The SiO_2 layer in Figure 7.2 controls the barrier height for thermionic emission due to the larger conduction band offset of SiO_2 with TiN compared to the conduction band offset of HfO_2 to TiN. Because the SiO_2 has the largest band offset of any of the dielectrics in the gate stack, its band offset with the metal Fermi energy level is the maximum barrier height for an electron injected from the gate metal. However, the SiO_2 layer is not adjacent to the gate metal and hence cannot not described by (7.3). Therefore, Schottky emission, expressed by (7.5), does not model electron gate thermionic emission in Figure 7.2. This is one example of the limited use of (7.5) when considering multilayer dielectrics and illustrates the attention that is required when using (7.5) to analyze a multilayer dielectric gate stack.

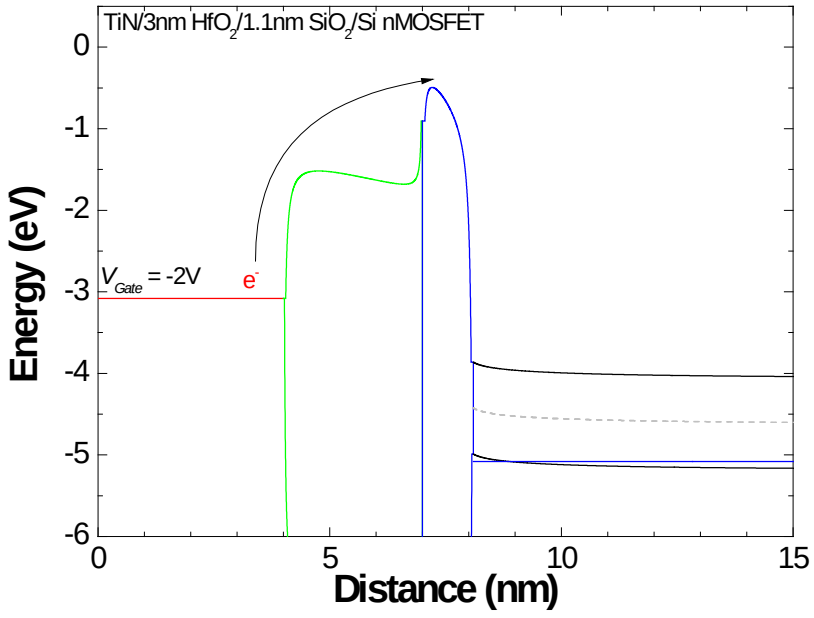


Figure 7.2: Energy band diagram of a TiN/3nm HfO_2/1nm SiO_2/Si nMOSFET biased at -2V. Image charges are used in the calculation of the dielectric conduction band. The barrier height for thermionic emission of electrons from the metal gate is dictated by the barrier height of the second dielectric (SiO_2) not the first dielectric adjacent to the metal gate (HfO_2). Created using [40].

Another example in which Schottky emission does not describe thermionic emission in multilayer dielectrics is illustrated by examining x_{max} in (7.2). For Schottky emission in a multilayer dielectric stack, (7.5) is only valid if x_{max} as described by (7.2) resides within the layer adjacent to the conductor from which thermionic emission is considered.[14] Equation (7.2) indicates that as the electric field approaches zero x_{max} approaches infinity. One can envision that at a low enough electric field an x_{max} that is greater than the thickness of the dielectric adjacent to the conductor could result. This situation is highlighted by considering thermionic emission at positive gate biases for the 3 or 5nm HfO_2/1.1nm SiO_2 dielectric gate stacks measured in this study. Figure 7.3 shows the energy band diagram for a 5nm HfO_2/1.1nm SiO_2 dielectric stack for substrate electron injection. The dielectric barriers are calculated three different ways in Figure 7.3: 1) without image charge effects (dashed lines), 2) with image charge effects only due to the semiconductor mimicking the barrier lowering described by Schottky emission (thin lines labeled "Schottky image" in Figure 7.3), and 3) with image charge effects due to both the conductor-dielectric interfaces and dielectric-dielectric interfaces (thick lines labeled "image charge" in Figure 7.3). A slight difference in the maximum barrier height is observed between the "Schottky image" and the "image charge" due to the SiO_2 dielectric – HfO_2 dielectric interaction. For the maximum barrier height calculated using the "Schottky image", (7.3), to approximate the maximum barrier height calculated considering the dielectric-dielectric interface ("image charge"), x_{max} must be slightly less than the dielectric thickness. Figure 7.4 plots the location of x_{max} as a function of electric

[14] The argument of the importance of the location of x_{max} also applies to thin single layer dielectrics. In multilayer dielectrics however, layers can be very thin without risk of large leakage currents due to direct tunneling if additional layers are present. Indeed, this situation is more common.

field for the SiO_2 IL in this study. For electric fields below ~1.25MV/cm, the calculated peak of the barrier height resides at or beyond the SiO_2 IL, which is 1.1nm thick and (7.3) fails to describe the maximum barrier height. Conventional Schottky analysis should therefore be performed at electric fields above ~1.25MV/cm for the devices in this study.

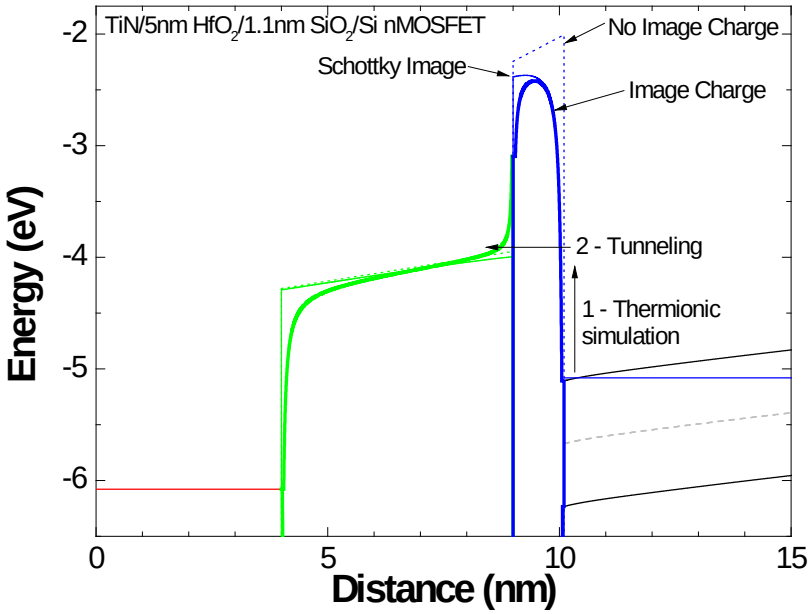


Figure 7.3: Energy band diagram of a 5nm HfO_2/1.1nm SiO_2 nMOSFET. The dielectric barriers are shown calculated (dashed line) with image charge effects, (solid thin lines) image charge due only to the semiconductor, and (solid thick lines) image charge due to the semiconductor, dielectric interfaces, and metal gate. The majority of the barrier lowering occurs in the SiO_2 IL while no appreciable barrier lowering occurs in the HfO_2 layer. Created using [40].

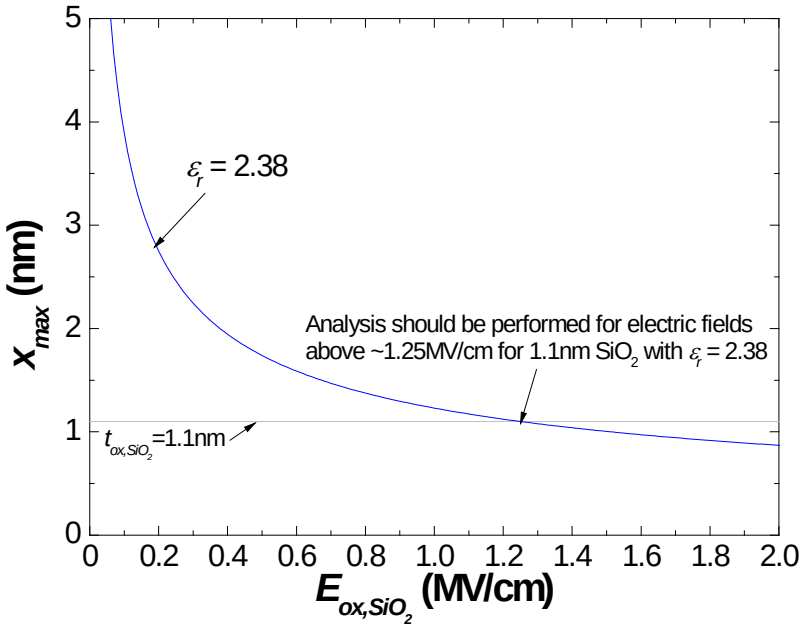


Figure 7.4: Plot of the distance where the maximum barrier height occurs, x_{max}, as a function of electric field in the SiO_2 interfacial layer for the high frequency dielectric constant of SiO_2, ε_r = 2.38. The plot indicates that electric fields above ~1.25MV/cm can be analyzed using (7.5) for the SiO_2 IL 1.1nm thick.

In contrast to the examples given above where Schottky emission did not describe thermionic emission, for the HfO_2/SiO_2 dielectric stack in this study, Schottky conduction as expressed by (7.5) does describe thermionic emission from the substrate over the SiO_2 IL. However, according to Figure 7.4, this is only true when the electric fields in the SiO_2 layer are above ~1.25MV/cm. The reasons that Schottky emission from the substrate over the SiO_2 can be safely examined are two-fold. First, carrier separation analysis, section 5.2, reveals that the gate leakage current corresponds to substrate injection for positive gate biases (electron injection) and negative gate biases (hole injection). The maximum barrier height in the HfO_2/SiO_2 bilayer is dictated by the SiO_2 layer as is shown in Figure 7.2 and described in the associated discussion. Since carrier transport originates from the substrate and the adjacent dielectric (SiO_2 layer) dictates the maximum barrier height for both electrons and holes, Schottky emission correctly describes thermionic emission over

the HfO_2/SiO_2 bilayer barrier. Second, as was described using Figure 7.4, for analysis performed for a SiO_2 electric field above ~1.25MV/cm, the location where the maximum barrier occurs resides in the SiO_2 layer, which is necessary for Schottky emission to apply. Schottky analysis can therefore be applied to the HfO_2/SiO_2 stacks used in this study to determine if Schottky emission describes the temperature dependence of the gate leakage current.

Schottky analysis is performed in much the same way as P-F conduction analysis was performed in Section 6.1 since (7.5) is an Arrhenius-like equation. The activation energy of the Arrhenius equation, which corresponds to the reduced barrier height given by (7.3), is derived from the slope at various $E_{ox,SiO2}^{1/2}$ points taken from the inverse temperature linearized form of (7.5):

$$\ln\left(\frac{J_{SE}}{T^2}\right) = -\frac{q(\phi_B - \sqrt{qE_{ox}/4\pi\varepsilon_0\varepsilon_r})}{k_b}\frac{1}{T} + \ln(A^*). \tag{7.6}$$

The barrier height at flatband, which should correspond to the SiO_2/Si conduction/valance band offset, is then determined by linear extrapolation of the activation energy versus $E_{ox,SiO2}^{1/2}$ [E_{ox} in (7.6)] to the point at which $E_{ox,SiO2}^{1/2}$ is equal to zero. The electric field in the SiO_2 layer was extracted using the energy band simulation tool described in Chapter 4. Performing Schottky analysis of the temperature dependent gate leakage data in this study however, reveals an extracted barrier height of ~0.4eV, which is considerably smaller than the SiO_2/Si conduction band offset of 3.1eV for positive gate biases. Schottky analysis for nMOSFETs biased at negative voltages is physically non-sensical as the extracted activation energy is negative. Schottky analysis for 5nm HfO_2 pMOSFETs at negative gate biases results in a barrier height of ~0.2eV,

which is significantly smaller than the SiO_2/Si valance band offset of 4.68eV. Since the extracted barrier height does not match the band offset, thermionic emission of carriers over the SiO_2 IL is not occurring.

This perhaps is not a surprising result given that the highest temperature for which devices were measured was 400K and the conduction and valance band offset of the SiO_2/Si interface is very large. Consideration should be given to emission through the SiO_2 (via tunneling, assumed in this argument to be a non-limiting mechanism) into the HfO_2 conduction band since HfO_2 has a band offset of 1.4eV, a much lower value than the SiO_2/Si conduction band offset of 3.1eV. Inspection of the reduced barrier height of the HfO_2 conduction band with respect to the Femi energy level of the Si in a 3nm HfO_2/1.1nm SiO_2 nMOSFET at V_{Gate} = 2V reveals a barrier of only 0.48eV. The reduced barrier height of 0.48eV is small enough that if tunneling through the SiO_2 layer can be considered as non-limiting, then thermionic emission into the HfO_2 conduction band may explain the temperature dependence of the gate leakage current.

Figure 7.3 shows the energy band diagram of thermionic stimulation of carriers above the HfO_2 conduction band where tunneling can take place through the thin SiO_2 layer. As observed in Figure 7.3, the maximum barrier height of the HfO_2 is not affected significantly by image charge effects. The maximum barrier height of the HfO_2 in relation to the Si band edge at the SiO_2/Si interface (neglecting image charge effects) is instead reduced by the potential drop in the in SiO_2 layer $V_{ox,SiO2}$, or:

$$V_{ox,SiO_2} = E_{ox,SiO_2} t_{ox,SiO_2} \tag{7.7}$$

where $E_{ox,SiO2}$ is the electric field in the SiO_2 layer (without image charge effects) and $t_{ox,SiO2}$ is the thickness of the SiO_2 layer. The HfO_2 barrier height is lowered by $V_{ox,SiO2}$ and

following Schottky in modifying Richardson's thermionic emission equation, the thermionic assisted tunneling (TAT) current is written as:

$$J_{TAT} = A^{*}T^{2}\exp\left(\frac{-q(\phi_B - E_{ox,SiO_2}t_{ox,SiO_2})}{k_bT}\right). \quad (7.8)$$

Determining if (7.8) describes the temperature dependence of the gate leakage current is performed by extracting ϕ_B using the relationship described in (7.8) and comparing it to the HfO_2/Si conduction and valance band offsets, which are ~1.4eV and ~3.18eV, respectively.[15]

Extracting ϕ_B using (7.8) is done using the same method that was used for Schottky emission and P-F conduction in Chapter 6, which is to linearize (7.8) versus the inverse of temperature as follows:

$$\ln\left(\frac{J_{TAT}}{T^2}\right) = -\frac{q(\phi_B - E_{ox,SiO_2}t_{ox,SiO_2})}{k_b}\frac{1}{T} + \ln(A^{*}). \quad (7.9)$$

Equation (7.9) reveals that data described by (7.8) plotted as $\ln(J_{Gate}/T^2)$ versus $1/T$ should result in a straight line whose slope is proportional to the activation energy, (ϕ_B-$E_{ox,SiO2}t_{ox,SiO2}$). Plotting the activation energy versus $E_{ox,SiO2}$ should result in a *y*-intercept of ϕ_B and a slope of $t_{ox,SiO2}$. In the following sections, this method of extracting ϕ_B is applied to the temperature dependence of the gate leakage current for both positive and negative gate biases to determine if (7.8) describes the data. The method of extracting ϕ_B for TAT, (7.8), is very similar to the method for extracting ϕ_B for Schottky emission, (7.5), except $E_{ox,SiO2}$ is used in TAT instead of $E_{ox,SiO2}^{1/2}$,which is used for Schottky. Due to these similarities, much of the commentary and discussion of the data in relation to

[15] Another method to determine if (7.8) describes the temperature dependence of the data is by examining if the slope of the activation energy versus $E_{ox,SiO2}$ plot results in the SiO_2 thickness of 1.1nm.

TAT can also be applied to Schottky emission. Where appropriate, discussion of the data in relation to Schottky emission is included as a supplemental discussion to the brief discussion of Schottky emission above.

7.2: Analysis of Thermionic Assisted Electron Tunneling Current

To examine the validity of the TAT model, a representative TAT characteristic plot, $\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}$, for positive gate biases is shown in Figure 7.5. Notice that the characteristic TAT plot shown in Figure 7.5 is similar to a characteristic Schottky plot except for the *x*-axis, which plotted as $E_{ox,SiO2}$ (TAT) instead of $E_{ox,SiO2}^{1/2}$ (Schottky). The data in Figure 7.5 is observed to have a slight curvature (as shown by comparison to the linear reference line) over the large electric field range shown (~2MV/cm to ~7MV/cm), which limits the validity of TAT as an explanation to the temperature dependence of the gate leakage current. In order for the data to be considered linear, as predicted by (7.9), only a relatively small electric field range can be used.

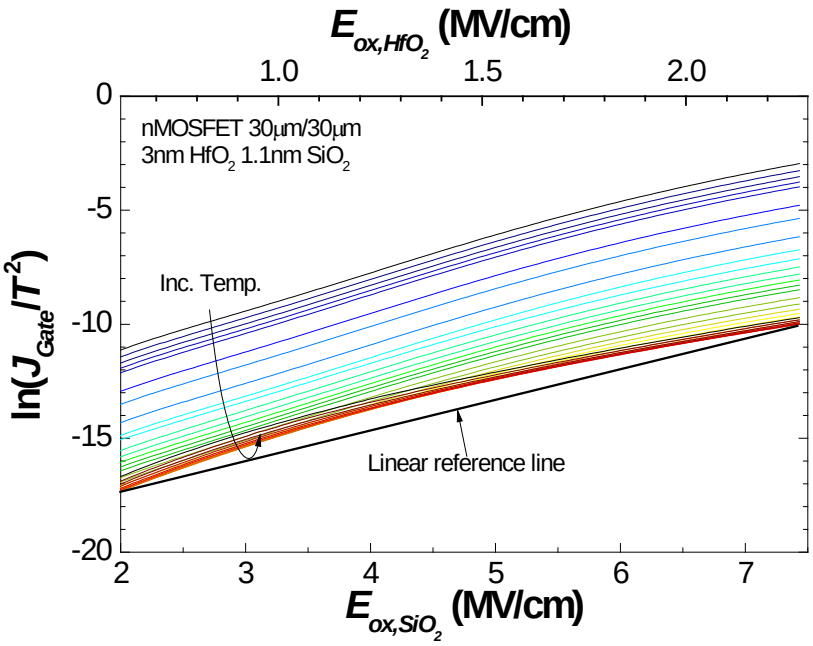


Figure 7.5: A characteristic thermionic stimulated tunneling current plot, $\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}$, for a 3nm HfO_2/1.1nm SiO_2 nMOSFET. A linear reference line has been added for comparison.

Additionally, Figure 7.5 shows that for the majority of the temperatures examined, $\ln(J_{Gate}/T^2)$ decreases as the temperature increases as indicated by the arrow. A clearer representation of this trend is shown in Figure 7.6 where $\ln(J_{Gate}/T^2)$ is plotted versus $1/T$ for select $E_{ox,SiO2}$ values. For temperatures below ~240K, $\ln(J_{Gate}/T^2)$ increases as the temperature decreases and results in a negative activation energy, which is physically implausible for thermionic emission. Only for temperatures above ~240K does the data have negative slope, which results in a positive activation energy. Due to the negative activation energy for temperature below ~240K, TAT and Schottky emission fail to explain the temperature dependence of the gate leakage current in this temperature range.

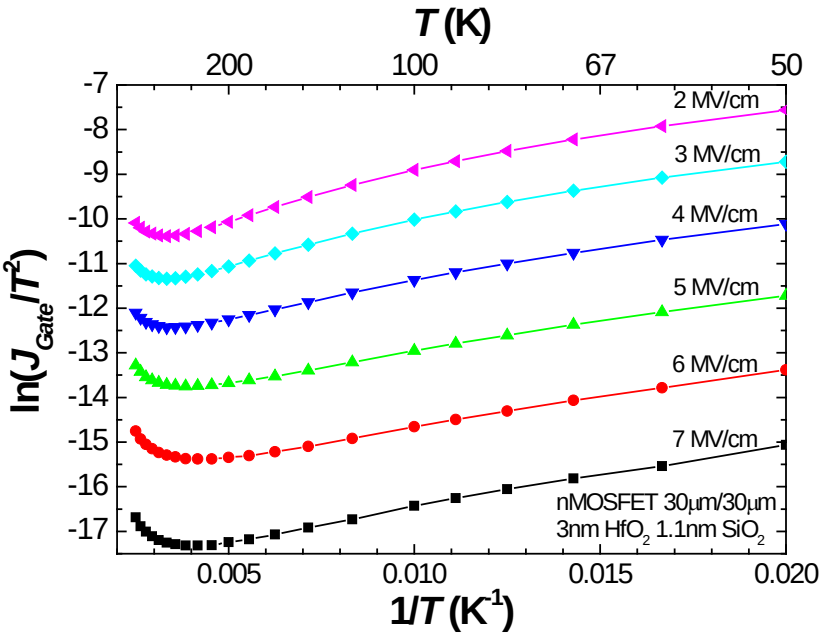


Figure 7.6: Characteristic thermionic stimulated tunneling current plot, $\ln(J_{Gate}/T^2)$ versus $1/T$ of a 3nm HfO_2/1.1nm SiO_2 nMOSFET for select SiO_2 electric fields. For temperatures below ~240K, the increasing $\ln(J_{Gate}/T^2)$ values for decreasing temperatures results in a negative activation energy, which is physically non-sensical for thermionic emission.

The data for T > ~240K is nonlinear, which results in an activation energy that is dependent on the temperature ranges chosen to perform the extraction, indicating that the behavior of the data is not Arrhenius. Similar observations of a temperature range-dependent activation energy were made during P-F analysis (Section 6.3) and the Arrhenius analysis (Section 5.3). Extracting the activation energy using temperatures 360K to 400K results in Figure 7.7. The activation energy is nonlinear with respect to $E_{ox,SiO2}$ and results in a barrier height extrapolation that depends on the electric field range over which the linear fit is performed. A significant difference in the range of activation energies is also observed between devices composed of 3nm and 5nm HfO_2. Since both of the stack types have the same SiO_2 IL thickness, the activation energy at a given $E_{ox,SiO2}$ (or $E_{ox,SiO2}^{1/2}$ for Schottky emission) should be the same according to (7.9) and

(7.6). Hence, TAT (or Schottky emission) does not explain the temperature dependence of the gate leakage current. A linear extrapolation of the activation energy (Figure 7.7) to $E_{ox,SiO2} = 0$ results in a ϕ_B ranging from ~0.15eV to ~0.4eV, which is significantly lower than the HfO_2/Si conduction band offset of ~1.4eV. This disparity is another indication that TAT does not explain the temperature dependence of the gate leakage current for the positive gate bias regime.

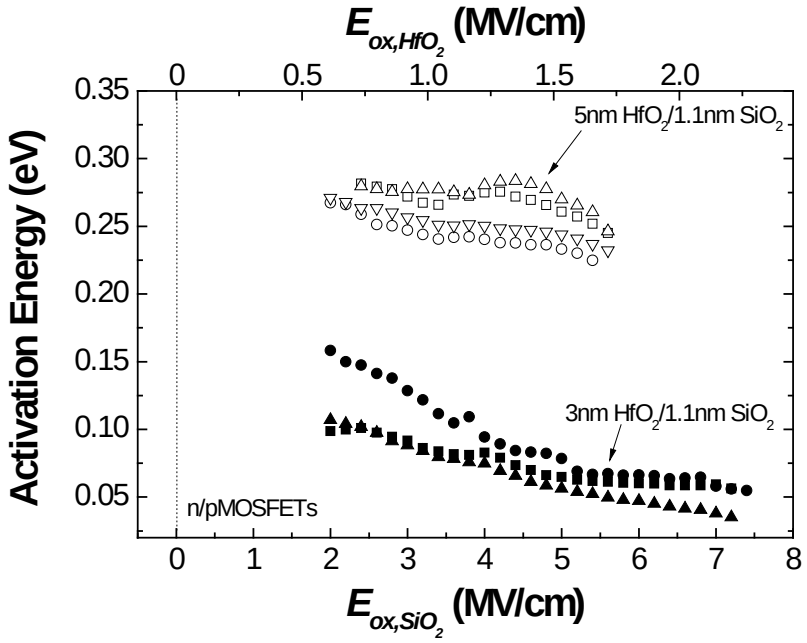


Figure 7.7: Activation energy (or reduced barrier height) versus SiO_2 electric field for n/pMOSFETs composed of 5nm HfO_2/1.1nm SiO_2 and 3nm HfO_2/1.1nm SiO_2. The barrier height at flatband is derived using linear extrapolation to zero electric field (dashed vertical line). Various barrier heights from ~0.15eV to ~0.4eV can be extracted depending on the device, stack type, and electric field ranged used.

7.3: Analysis of Thermionic Assisted Hole Tunneling Current

The validity of the TAT model is analyzed using the gate leakage current at negative gate biases, which corresponds to substrate hole injection. A characteristic $\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}$ plot is shown in Figure 7.8. The negative sign on $E_{ox,SiO2}$ indicates the polarity of the electric field. A slight curvature in the data is observed in the

data as shown by comparison to the linear reference line. According to (7.9), the data should be linear. As temperature increases, $\ln(J_{Gate}/T^2)$ decreases for most of the temperature range. Plots of $\ln(J_{Gate}/T^2)$ versus $1/T$ are shown in Figure 7.9 and Figure 7.10 for several HfO_2 gate stacks.

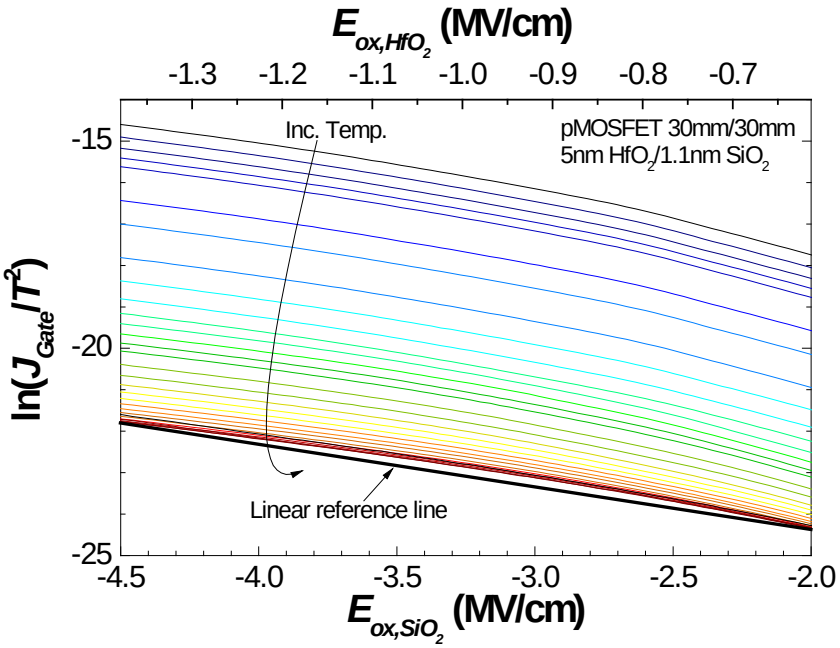


Figure 7.8: A characteristic thermionic stimulated tunneling current plot, $\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}$, for a 5nm HfO_2/1.1nm SiO_2 pMOSFET. A linear reference line has been added for comparison.

Figure 7.9 is a characteristic thermionic Arrhenius plot for a 3nm HfO_2 nMOSFET. Similar behavior were also observed for data from 3nm HfO_2 pMOSFETs and 5nm HfO_2 nMOSFETs and are not shown for brevity. Figure 7.10 is a representative plot of 5nm HfO_2 pMOSFETs. Except for 5nm HfO_2 pMOSFETs (Figure 7.10), the slope for the inverse temperature plots (e.g., Figure 7.9) is positive, which results in a negative activation energy. The negative activation energy indicates that the TAT/Schottky model does not fit the data.

For 5nm HfO_2 pMOSFETs (Figure 7.10), a negative slope is shown in the $1/T$ plot for temperatures above ~360K. Activation energies extracted from these slopes are very small however; less than ~0.1eV. In addition, for $E_{ox,SiO2}$ = -3MV/cm to -4.5MV/cm the activation energy increases rather than decreases as the electric field increases (data not shown). Equation (7.9) predicts that the activation energy, (ϕ_B-$E_{ox,SiO2}t_{ox,SiO2}$), should decrease as $E_{ox,SiO2}$ increases.[16] The incorrect change in the activation energy with electric field indicates that the gate leakage current of the 5nm HfO_2 pMOSFETs is not explained by TAT/Schottky conduction. A linear fit to the point at which the activation energy decreases as the electric field increases ($E_{ox,SiO2}$ = -2MV/cm to -3MV/cm) results in an extracted ϕ_B = ~70meV, which is significantly smaller than the HfO_2/SiO_2 valance band offset of ~3.18eV. The large discrepancy between the valance band offset and the extracted ϕ_B indicates that TAT does not model the data.

[16] Note in (7.5), (7.7), (7.8) and (7.9) and other similar equations that the electric field term refers to the magnitude of the electric field which is a positive value. The minus sign in a negative electric field value refers to its polarity. The activation energy should decrease for larger magnitude values regardless of its polarity.

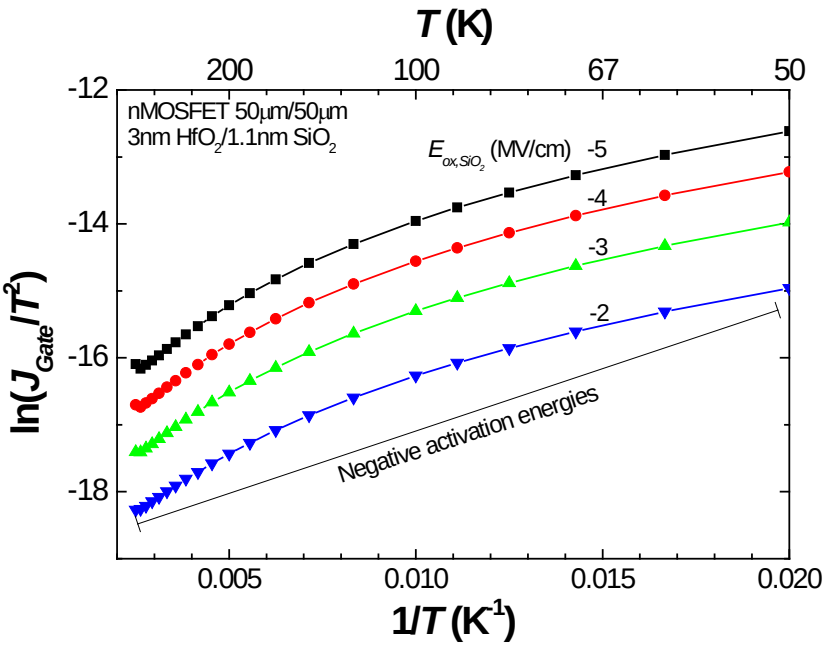


Figure 7.9: Characteristic thermionic stimulated tunneling current plot, $\ln(J_{Gate}/T^2)$ versus $1/T$ of a 3nm HfO_2/1.1nm SiO_2 nMOSFET for select SiO_2 electric fields. For temperatures below ~400K, the increasing $\ln(J_{Gate}/T^2)$ values for decreasing temperatures results in a negative activation energy, which is physically invalid for thermionic emission.

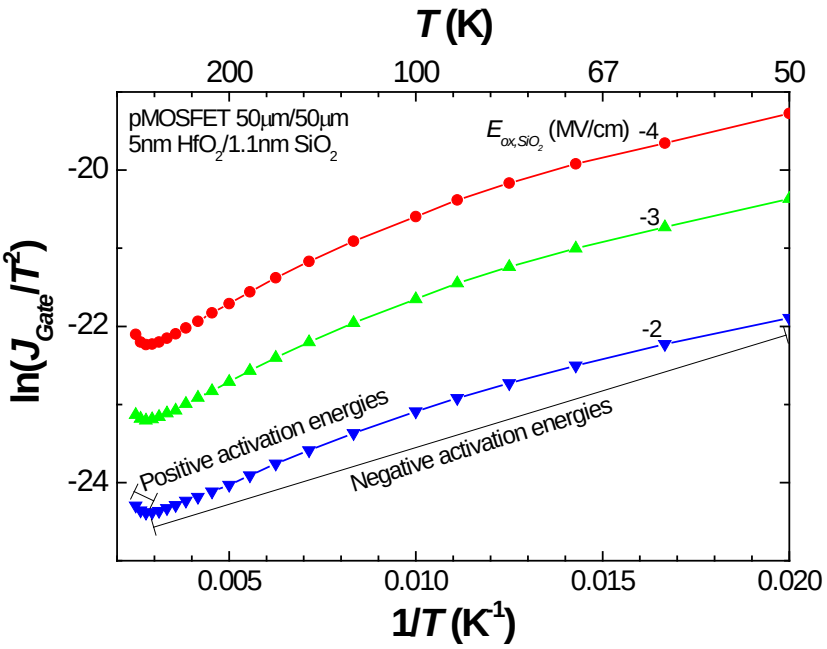


Figure 7.10: Characteristic thermionic stimulated tunneling current plot, $\ln(J_{Gate}/T^2)$ versus $1/T$ of a 5nm HfO_2/1.1nm SiO_2 MOSFET for select SiO_2 electric fields. For temperatures below ~360K, the increasing $\ln(J_{Gate}/T^2)$ values for decreasing temperatures results in a negative activation energy, which is physically implausible for thermionic emission.

In summary, models using thermionic emission over the SiO_2 layer (Schottky emission) or into the HfO_2 conduction/valance band, via tunneling through the SiO_2 (TAT), does not explain the temperature dependence of the gate leakage current for positive or negative gate biases. The data fail to follow the Schottky and TAT models based on the following observations:

1. Nonlinear (i.e., non-Arrhenius) behavior of the characteristic plot of $\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}$, which is predicted by TAT, (7.9). Nonlinear behavior of the characteristic plots for Schottky emission [$\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}^{1/2}$] were also observed (not shown).
2. Negative activation energies were obtained for temperatures less than ~300K (activation energies should be positive).

3. Extracted barrier heights (ϕ_B) that significantly differ from the reported band offsets with Si.

To summarize the findings in Chapters 6 and 7, two of the most commonly reported temperature dependent conduction mechanisms for HfO_2, P-F and thermionic emission, fail to explain the strong temperature dependence of the gate leakage current for the devices used in this study. It is important to note that both P-F and thermionic emission studies reported in the literature in regards to HfO_2 MOS devices used temperatures greater than what was used in this study. While for greater temperature ranges, P-F and thermionic emission may explain the temperature dependence of the gate leakage current, they do not explain the temperature dependence for temperatures ranging from 6K to 400K in these devices leaving other possible mechanisms to be explored.

CHAPTER 8: HOPPING CONDUCTION ANALYSIS

In the analysis of the temperature dependence of the gate leakage current in 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs, the data has been compared with several models. Two of the most popular models for the temperature dependence of the gate leakage current in HfO_2 MOS devices, P-F (Chapter 6) and Schottky (Chapter 7), fail to explain the temperature dependence of the gate leakage current of the devices in this study over the temperature range 6K to 400K. Notable problems with P-F model were 1) a range of defect levels extracted between 3nm and 5nm HfO_2 samples yielding a thickness dependence for which the P-F model does not incorporate, 2) the extracted defect energy level is highly sensitive to the chosen temperature range, and 3) the inability of the P-F model to explain the transport mechanism trend for gate voltages less than ~1V (i.e., device operation voltage range) in 3nm HfO_2 samples. Schottky conduction also fails to explain the strong temperature dependence of devices studied for the temperature range of 6K to 400K. Notable problems with Schottky emission were 1) Schottky emission is thickness independent while a thickness dependence for the extracted barrier heights was found between 3nm and 5nm HfO_2 devices, 2) the extracted barrier heights differed significantly from the band offsets of SiO_2 or HfO_2 with Si, and 3) negative activation energies were found for temperature less than ~300K, leaving a wide temperature range where the trend of the gate leakage current with temperature cannot be explained. Analysis of one-dimensional quantum mechanical transmission coefficient calculations for electron substrate injection (positive gate biases), described in

Appendix C, show nearly the same transmission coefficient for the equivalent distance of Fowler-Nordheim tunneling into the HfO_2 conduction band for the 3nm and 5nm HfO_2 gate stacks. The gate leakage current of 3nm and 5nm HfO_2 devices at equivalent tunneling distances show orders of magnitude difference, indicating that quantum mechanical tunneling from the Si to HfO_2 conduction band is not responsible for the temperature dependence of the gate leakage current.

In short summary, the following is concluded about the conduction path of the carriers as they move from the Si substrate to the TiN gate given the analysis shown thus far.

1. Carriers are not emitted from a single defect level in HfO_2 to the HfO_2 conduction band (for electron substrate injection) or the valance band (for hole substrate injection), (P-F analysis).
2. The dominate transport mechanism is not due to carriers going over the barrier (Schottky analysis).
3. The dominate transport mechanism is not due to quantum mechanical tunneling from the Si to HfO_2 conduction band (transmission coefficient analysis, Appendix C).

Due to points 2 and 3, defect mediated transport is suggested as the dominate transport mechanism over the temperature range investigated, 6K to 400K. Point 1 may suggest many defect levels exist in the HfO_2 layer, which would be consistent with the poly-crystalline/amorphous nature of the HfO_2. The thickness dependence of the gate leakage current could be due to a longer percolation path in 5nm HfO_2 devices compared to the 3nm HfO_2 devices for defect mediated transport. The literature also suggests a strong

defect presence in HfO_2 based dielectric stacks (e.g., [105, 106]). Oxygen vacancies in the HfO_2 have been suggested as the origin of traps as well as impurity metals and polaronic traps (e.g., [28, 29, 107]).

In the semiconductor device community, trap assisted tunneling via electrically active deep level point defects have been used to describe leakage currents in thin MOS dielectrics (e.g., [108, 109]). However, another approach to defect mediated transport is known as hopping conduction (see Section 2.2). Hopping transport has been used extensively in disordered systems to explain the temperature dependence of the conductivity (e.g., [45, 110-113]). This chapter explores the application of hopping conduction models to the temperature dependence of the gate leakage current of HfO_2/SiO_2 bilayer dielectric stacks.

8.1: Analysis of Hopping Conduction

Hopping conduction was described in Section 2.2 and for three-dimensional variable range hopping predicts a $T^{-1/4}$ dependence in the conductivity for a constant density of states with respect to energy around the Fermi level. For the two-dimensional case applicable to the thin dielectrics investigated in this study, a $T^{-1/3}$ dependence in the conductivity is predicted and the density of states around the Fermi level, $N(E_F)$ is given by [see (2.5)] [114]:

$$N(E_F) = C_2^{\,3}\,\frac{\alpha^2}{k_b t(T_0)} \tag{8.1}$$

where C_2 is a constant (derived to be approximately 2). The pre-exponential constant σ_{hop} in (2.4) has also been derived in the literature and depends on various material parameters

(e.g., defect density, velocity of sound in the medium, deformation potential) [115, 116]. Traditional analysis of variable hopping conduction is performed by linearizing (2.4) as:

$$\ln(\sigma) = -T_0 T^{-\frac{1}{d+1}} + \ln(\sigma_{hop}), \tag{8.2}$$

which indicates that a plot of $\ln(\sigma)$ versus $T^{-1/d+1}$ for data that exhibit variable range hopping should result in a straight line. The conductivity through the gate stack of the devices in this study are not linear when plotted as suggested by (8.2) with *d* equal to three or two, indicating that variable range hopping conduction cannot describe the temperature dependence of the conductivity by itself.

Arrhenius analysis performed in Section 5.3 indicates a weak temperature sensitive regime (6K-50K), a strong temperature regime (near 400K), and a very wide temperature transition regime between the two. An Arrhenius plot of the gate leakage current density at various gate voltages above V_{th} in the transport limited regime (0.85V to 2V) is shown in Figure 8.1, illustrating the three general temperature regimes identified above. Figure 8.1 shows the wide temperature transition regime (labeled "Region 1") and the strong temperature sensitive regime (labeled "Region 2"). The inset in Figure 8.1 shows the weak temperature sensitive regime (labeled "Region 3").

Following the three general temperature regimes observed in the Arrhenius plot (Regions 1-3), a model for the conductivity through the HfO_2 (which assumes the presence of the SiO_2 IL in the HfO_2/SiO_2 bilayer is a non-limiting function of the conductivity of the HfO_2 layer or the HfO_2/SiO_2 stack for that matter) for the full temperature regime measured (6K-400K) is proposed:

$$\sigma_{HfO_2} = \sigma_{hop} e^{-\left(\frac{T_0}{T}\right)^{1/3}} + \sigma_{Arr} e^{-\frac{E_1}{k_b \cdot T}} + \sigma_0 . \quad (8.3)$$

The first conductivity term represents two-dimensional variable range hopping to describe Region 1, the second conductivity term represents an Arrhenius mechanism with activation energy E_1 to model Region 2, and the last conductivity term represents temperature independent conduction mechanisms (e.g. direct/FN tunneling, temperature independent hopping due to high electric fields) to model Region 3. The addition of the three conduction mechanisms as shown in (8.3) implies parallel conduction, where it is assumed that the conduction path of one mechanism does not significantly affect the conduction path of another, or at least if it does, does so when the affected mechanism is not the dominate conduction mechanism. The parallel conduction mechanism approach has been used in the literature in conjunction with describing temperature dependent conductivity that exhibit multiple mechanisms (e.g., [43, 45, 117]).

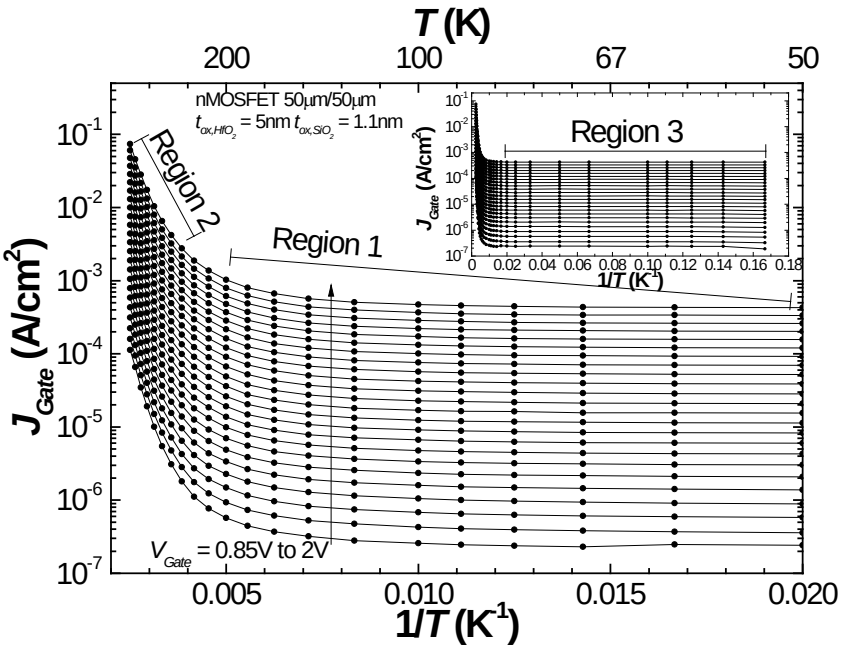


Figure 8.1: Representative Arrhenius plot of the gate leakage current density for gate voltages ranging from 0.85V to 2V of a 5nm HfO_2 nMOSFET. A wide nonlinear region is shown from 50K to 400K. The inset shows a very weak temperature region for $T <$ 50K.

Equation (8.3) cannot be linearized with respect to temperature, hence nonlinear curve fitting is used to determine if (9.3) describes the temperature dependence of the gate leakage current from the perspective of conductivity through the HfO_2, which is assumed to be the material limiting current flow. The conductivity through the HfO_2 layer at various gate voltages is calculated using $\sigma_{HfO2}=J_{Gate}/E_{ox,HfO2}$. Nonlinear curve fitting is performed by adjusting the fitting parameters until the error between the model and data reach a local minimum. The method used in this study to calculate the error between the model and data is Chi-Squared, which is:

$$\chi^2 = \sum w_i \left(y_i - f(x_i)\right)^2, \tag{8.4}$$

where y_i and x_i is the i-th data point, $f(x_i)$ is the model evaluated at x_i, and w_i is the weighting. Without weighting, the reduction of Chi-Squared is biased toward larger

values of y_i, where the difference between the model and data has a larger impact on Chi-Squared. A weighting function of $w_i = 1/y_i^2$ creates a relative distance error where all points are equally weighted regardless of their value. The weighting function $w_i = 1/y_i$ is used for performing the nonlinear curve fits in this study as a compromise between the actual error and relative error between the model and data [118].

In order for nonlinear curve fitting to be effective, the initial estimation of the parameters (i.e., guess or initialization parameters) must be similar enough to the best fit parameters that when the nonlinear curve tool performs its error minimization techniques the local global minimum (the best fit) results rather than a local non-global minimum. The initial guess parameters were determined by performing two nonlinear fits to two different temperature ranges of the data. The low temperature regime ($T < 200$) was fit to:

$$\sigma_{HfO_2,low-temp} = \sigma_{hop} e^{-\left(\frac{T_0}{T}\right)^{1/3}} + \sigma_0 \tag{8.5}$$

and the high temperature regime ($T > 200$) was fit to:

$$\sigma_{HfO_2,high-temp} = \sigma_{Arr} e^{-\frac{E_1}{k_b \cdot T}} + \sigma_0 . \tag{8.6}$$

where σ_0 was initialized to the minimum conductivity of the temperature range (6K). The fitting results of (8.5) and (8.6) were then used as the initialization parameters for performing the nonlinear curve fit using (8.4).

8.2: Electron Substrate Injection Results and Discussion

Fitting the HfO_2 conductivity data using (8.3) and device level shared fit parameters[17] T_0 and E_1 results in an adjusted *R*-squared value greater than 0.999. Using singular T_0 and E_1 parameters to describe the temperature dependence of the conductivity across multiple gate voltages with high accuracy highlights the effectiveness in using (8.3) to model the data. An example of the correlation between the experimental data (symbols) and the fitted model (lines) is shown in Figure 8.2 (3nm HfO_2) and Figure 8.3 (5nm HfO_2). As observed in Figure 8.2 and Figure 8.3, the model in (8.3) describes the temperature dependence of the gate leakage current for the entire temperature range studied (6K to 400K). For pMOSFETs where the device is in accumulation, fitting was performed for temperatures above 40K where carrier-limited transport is not occurring. The model describes the large non-linear region (~50K to ~400K) on the Arrhenius plot that was discussed in Section 5.3.

For the 3nm HfO_2 samples, the T_0 parameter was extracted providing a value ~$3x10^5$K.[18] Assuming a decay length (α^{-1}) of 1nm[19] results in a density of states around the Fermi level in the HfO_2, $N(E_F)$, of ~$1x10^{20}$ ($cm^{-3}eV^{-1}$) using (2). The extracted defect density is similar to those reported by Xiong *et al*. for HfO_2 also manufactured by SEMATECH, which ranged from ~10^{19} to ~10^{21} $cm^{-3}eV^{-1}$ [121]. The average spacing

[17] "Shared fit parameters" refers to the same parameter value used to fit across multiple data sets. In the context of this study, singular T_0 and E_1 values are used to fit data over multiple conductivities extracted at various gate voltages. Single values for T_0 and E_1 are only enforced on the conductivities of the device being fitted and not across multiple devices. Hence "device level" (the level at which the parameters are shared) is used to describe the fitting procedure used.

[18] As a reference, the characteristic temperature parameter, T_0, for many materials can be quite large. For example, TiO_2 – $2x10^7$K [119], Cu-vermiculite intercalation – ~$1x10^5$ to ~$1x10^6$ [46], Ge – ~$1x10^8$K [120], and Si – ~$1x10^8$K [120].

[19] As a reference, the defect decay length in Ge is ~1nm [114]. As will be shown later the assumed value here of 1nm is close to an extracted value using electric field dependent hopping, ~0.83nm.

between defects is then $N^{-1/3} \approx 2$nm, if one considers defects spread over 1eV. Calculating the average spacing between defects assuming an energy spread of 1eV is not unreasonable as a large distribution in energy can exist for defects in HfO_2 indicated by theoretical calculations of oxygen vacancies [98]. For the 5nm HfO_2 samples, T_0 was extracted to ~8×10^5K resulting in $N(E_F)$ of ~2×10^{19} ($cm^{-3}eV^{-1}$) and an average defect spacing of 3.7nm considering defects spread over 1eV. A significant difference in the defect density is observed between the 3nm and 5nm HfO_2 samples. This may be due to a large concentration of defects in the SiO_2 at the HfO_2/SiO_2 interface, which are averaged lower in the 5nm HfO_2 compared to the 3nm HfO_2 samples due to its greater thickness; the variable range hopping model used assumes a constant defect density about the Fermi level [46, 115]. A significant concentration of defects in the SiO_2 layer was suggested in Section 5.1.2.

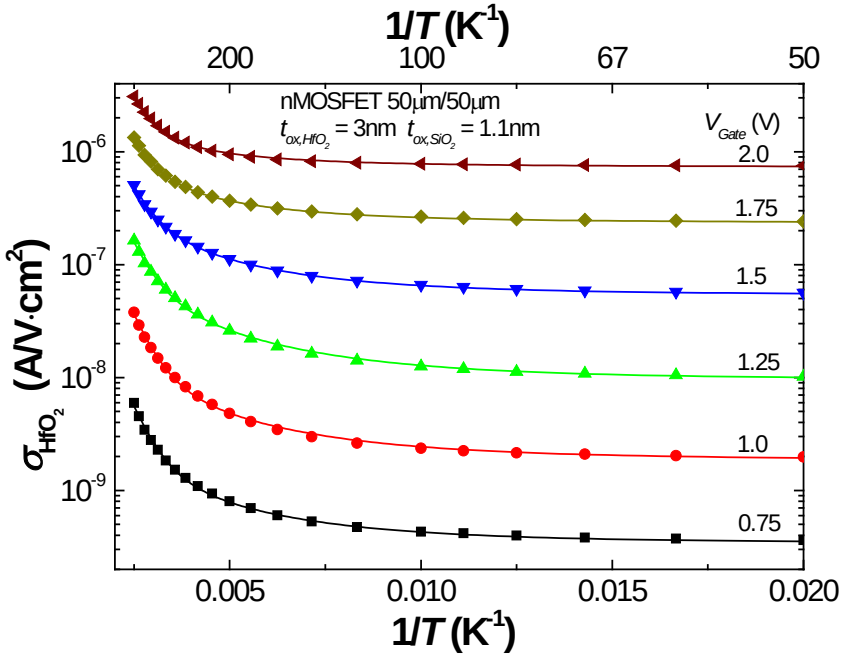


Figure 8.2: Arrhenius plot of experimental conductivity data (symbols) for gate voltages ranging from 0.75V to 2V of a 3nm HfO_2/1.1nm SiO_2 nMOSFET. Fit of variable range hopping and Arrhenius model described by (8.3) is also shown (lines). Model describes the nonlinear region (50K to 400K) very well in addition to the entire temperature region studied (T < 50K), not shown. Only gate voltages in 0.25V increments are shown for clarity.

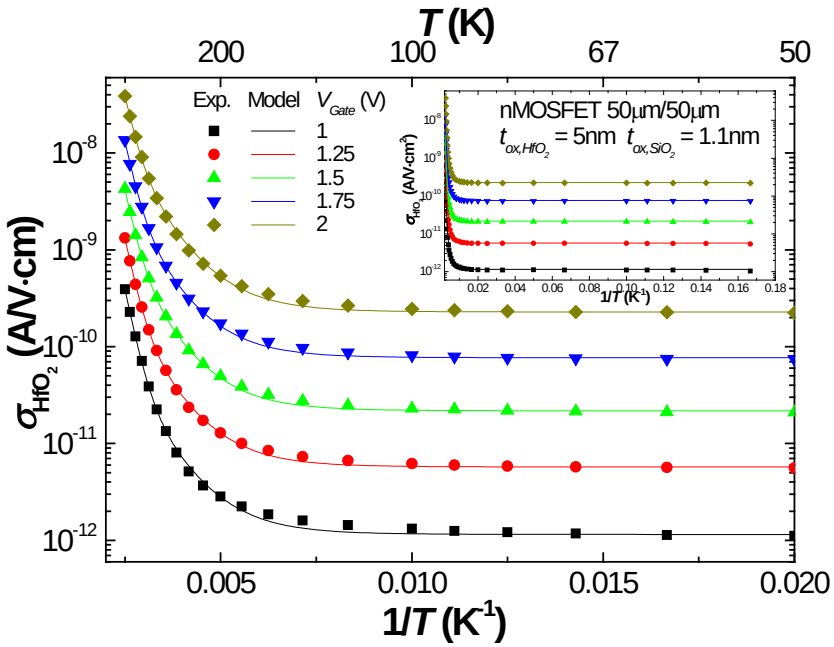


Figure 8.3: Arrhenius plot of experimental conductivity data (symbols) for gate voltages ranging from 1V to 2V for a 5nm HfO_2/1.1nm SiO_2 nMOSFET. Fit of variable range hopping and Arrhenius model described by (8.3) is also shown (lines). Model describes the nonlinear region (50K to 400K) very well in addition to the entire temperature region studied (inset). Only gate voltages in 0.25V increments are shown for clarity.

Fitting results for the Arrhenius expression provides an activation energy of ~0.19eV for the 3nm HfO_2 samples and ~0.42eV for the 5nm HfO_2 samples. If E_1 is allowed to vary for each gate voltage, the activation energy does not decrease as the gate voltage increases as predicted by the Poole-Frenkel conduction mechanism. There are many transport mechanisms that are represented by Arrhenius expressions. Below is a list of various kinetic processes that are described by an Arrhenius expression [45].

1. Thermal stimulation of from carriers from the defect energy level to the conduction band. The activation energy is the energy difference between the Fermi energy level and the conduction band.

2. Thermal stimulation of carriers from one defect state to another defect state where the activation energy is the difference between the two defect energy levels.
3. Hopping conduction within a finite energy band (energy bandwidth limited) where the activation energy corresponds to order of half the width of the energy band of defect states. The hopping energy may include a lattice deformation energy associated with the presence of the charged carrier (i.e., defect mediated polaron hopping).
4. Nearest neighbor hopping conduction where the activation energy is on the order of the energy band width of the defect states.

The activation energy can also be the sum of multiple processes, such as a hopping energy term and a polaron energy term [45]. As was discussed in the introduction, due to the high dielectric constant of HfO_2, polaron motion may have an effect on the transport mechanism.

Theoretical predictions of activation energies in monoclinic HfO_2 near the extracted activation energies reported here are: electron self-trapped polaron energy (0.32eV below conduction band) [29], electron polaron hopping transport energy (0.05eV) [29], thermal activation energy of an electron from the oxygen vacancy V^{-2} charge trap state (0.56eV) [28], and the energy difference of the oxygen vacancy V^{-1} and V^{-2} charge states (0.2eV) [28].

For the 3nm HfO_2 samples, the extracted activation energy of ~0.19eV is close to the energy difference between the oxygen vacancy V^{-1} and V^{-2} charge states (0.2eV). The similar energy levels suggest electron transport at large temperatures occurs by electron

hopping between these two charge states. For the 5nm HfO_2 samples, the extracted activation energy of ~0.42eV is similar to the electron self-trapped polaron energy (0.32eV) and the energy difference between the oxygen vacancy V^{-2} charge trap state and the HfO_2 conduction band (0.56eV). Additional studies are needed to determine the relationship between the Arrhenius activation energy and HfO_2 thickness and what mechanisms they model.

8.3: Hole Substrate Injection Results and Discussion

The HfO_2 conductivity data for the entire temperature range at negative gate voltages can also be described by (8.3) using device level shared fit parameters T_0 and E_1. Fitting the model to the data results in an adjusted *R*-squared value greater than 0.999. Figure 8.4 (3nm HfO_2) and Figure 8.5 (5nm HfO_2) are examples of the correlation between the experimental data (symbols) and the fitted model (lines). As observed in pMOSFETs, Figure 8.5, the model in (8.3) describes the temperature dependence of the gate leakage current for the entire temperature range studied (6K to 400K). For nMOSFETs (Figure 8.4), where at negative gate biases the device is in accumulation, fits were only performed for temperatures above 40K so that the gate leakage current modeled is dominated by transport-limited conduction (see Section 5.3). For both n- and pMOSFETs, the model describes the large nonlinear region (~50K to ~400K) on the Arrhenius plot that was discussed in Section 5.3.

The extracted characteristic temperature, T_0, for 3nm HfO_2 samples is ~1.7×10^4K, which, assuming a localized defect decay length of 1nm (α^{-1}) as was done for positive gate biases, results in a defect density about the Fermi level of ~$1.5 \times 10^{21} cm^{-3} eV^{-1}$. For the

5nm HfO_2 samples, T_0 is ~4.3x10^{4}K, which, again assuming α^{-1}=1nm, results in a defect density about the Fermi level of ~3.9x10^{20}cm^{-3}eV^{-1}. The extracted defect density for hole variable range hopping transport is about an order of magnitude above the defect density extracted for electrons in the previous section. The large defect densities for electrons and holes suggests that there is room for significant improvement in HfO_2 layer, which would reduce the temperature dependence of the gate leakage current.

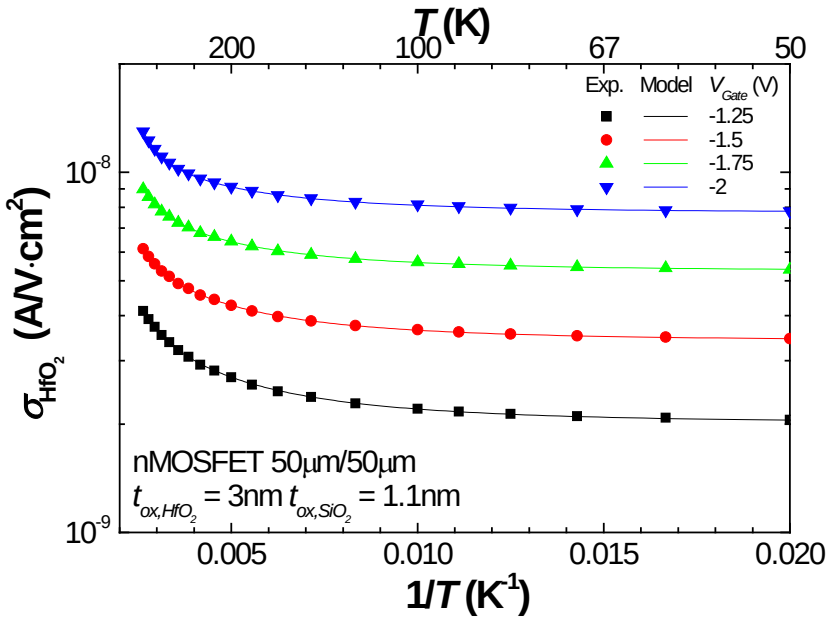


Figure 8.4: Arrhenius plot of experimental conductivity data (symbols) for gate voltages ranging from -0.75V to -2V of a 3nm HfO_2/1.1nm SiO_2 nMOSFET. Fit of variable range hopping and Arrhenius model described by (8.3) is also shown (lines). The model describes the nonlinear region (50K to 400K) very well. Temperatures below 40K were not used during the fit due to carrier rate limited conduction in accumulation (see Section 5.3).Only gate voltages in 0.25V increments are shown for clarity.

The extracted activation energy for the Arrhenius term in (8.3) is ~0.19eV for the 3nm HfO_2 samples and ~0.32eV for the 5nm HfO_2 samples. Theorized activation energies for holes include: polaron self-trap energy (0.57 and 0.14eV) [29], polaron hopping transport energy (0.08, 0.16, 0.19, and 0.56eV) [29], and the energy difference of the oxygen V^+ and V^0 states (0.01eV) [28]. The injection location of the carriers (near

the Si Fermi level) into the HfO_2 layer is near the middle of the HfO_2 band gap signifying that the Arrhenius expression does not correspond to carrier thermal transitions from defects near the middle of the HfO_2 band gap, [V^+ - E_C (3.39eV) and V^0 - E_C (3.19eV) oxygen vacancy charge states] to the HfO_2 band edge as the energy difference between the defect levels and HfO_2 valance band is too large compared to the extracted activation energy. The Arrhenius expression also is not likely to correspond to polaron hopping of holes in the HfO_2 valance band, as the energy difference between the Si Fermi energy level and self-trapped hole in the HfO_2 valance band is ~2.7 eV, which is significantly larger than the extracted activation energy. The Arrhenius expression might represent nearest neighbor hopping conduction or energy bandwidth limited conduction where the activation energy is proportional to the width of the defect energy band. Possible defects for which nearest neighbor hopping conduction could occur is the V^+ and V^0 oxygen vacancy charge states. For disordered layers or high defect densities, the distribution of energy levels for defects would be more broadly distributed than the mono-energetic levels calculated using monoclinic HfO_2. Additional studies are necessary to further determine the mechanism that the Arrhenius term represents.

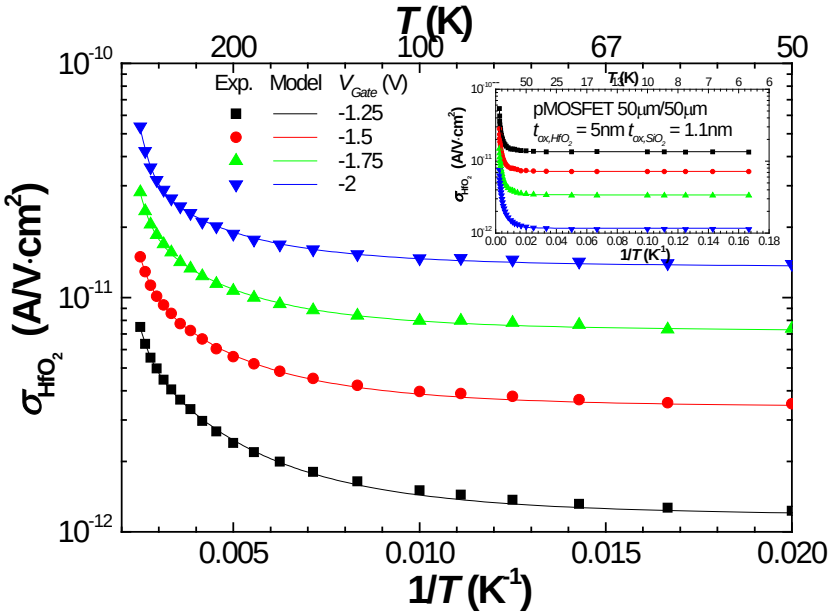


Figure 8.5: Arrhenius plot of experimental conductivity data (symbols) for negative gate voltages ranging from -1.25V to -2V for a 5nm HfO_2/1.1nm SiO_2 pMOSFET. Fit of variable range hopping and Arrhenius model described by (8.3) is also shown (lines). The model describes the nonlinear region (50K to 400K) very well as well as the entire temperature region studied (inset). Only gate voltages in 0.25V increments are shown for clarity.

8.4: Electric Field Dependent Variable Range Hopping

While the parameters T_0 and E_1 were held constant over the various gate voltages investigated, parameters σ_{hop}, σ_{Arr}, and σ_0 were allowed to vary with gate voltage. Figure 8.6 gives an example of how σ_{hop}, σ_{Arr}, and σ_0 vary with V_{Gate}. As observed in Figure 8.6, the parameters σ_{hop} and σ_{Arr} span several orders of magnitude as V_{Gate} increases to 2V showing a strong V_{Gate} dependence. However, the variable range hopping model, (2.4), was derived for low electric fields and does not predict a V_{Gate} dependence [122].

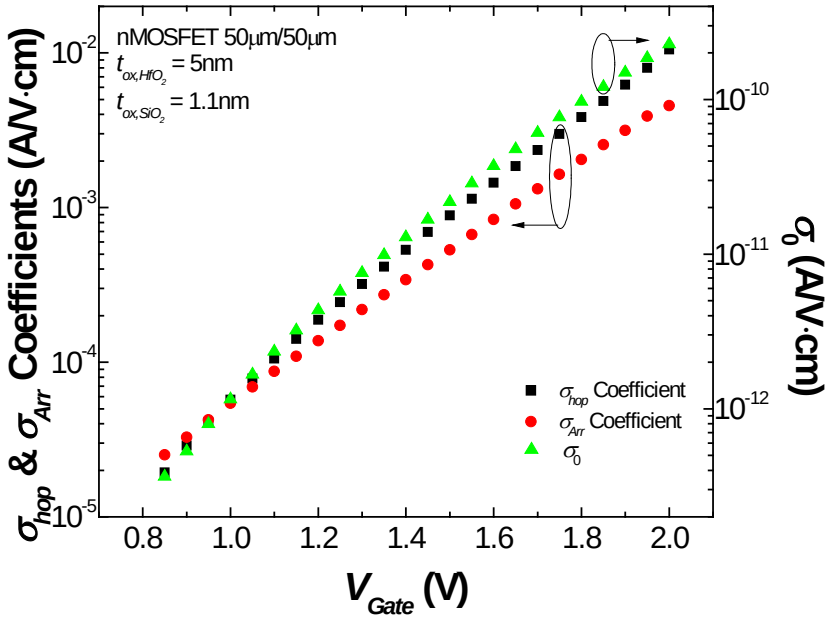


Figure 8.6: Example of the large change in the σ_{hop}, σ_{Arr}, and σ_0 fitting parameters with respect to V_{Gate}. While σ_0 is expected to change with V_{Gate}, σ_{hop} and σ_{Arr} are not expected to have such a strong dependence on V_{Gate}.

As was discussed in Section 2.2, at high electric fields hopping can occur without thermal assistance. Analogous to variable range hopping in three-dimensions, which has a temperature dependence of $T^{-1/4}$ for a constant density of state, electric field dependent variable range hopping has a dependence of $E_{ox}^{-1/4}$ as shown in (2.9). For two-dimensional variable range hopping the electric field dependence becomes $E_{ox}^{-1/3}$ similar to temperature dependent variable range hopping in two dimensions. The conductivity through the HfO_2 versus $E_{ox}^{-1/3}$ at positive gate biases is shown in Figure 8.7 for 6K, where temperature dependent hopping is minimized. The data in Figure 8.7 is linear over almost the entire electric field range above the threshold voltage where conduction path-limited transport occurs. A linear fit to the data (red line in Figure 8.7) demonstrates the linearity of the data when the log of the conductivity is plotted versus $E_{ox,HfO2}^{-1/3}$ and

indicates that non-thermally activated hopping due to the presence of high electric fields has an important role in the temperature dependence of the gate leakage current.

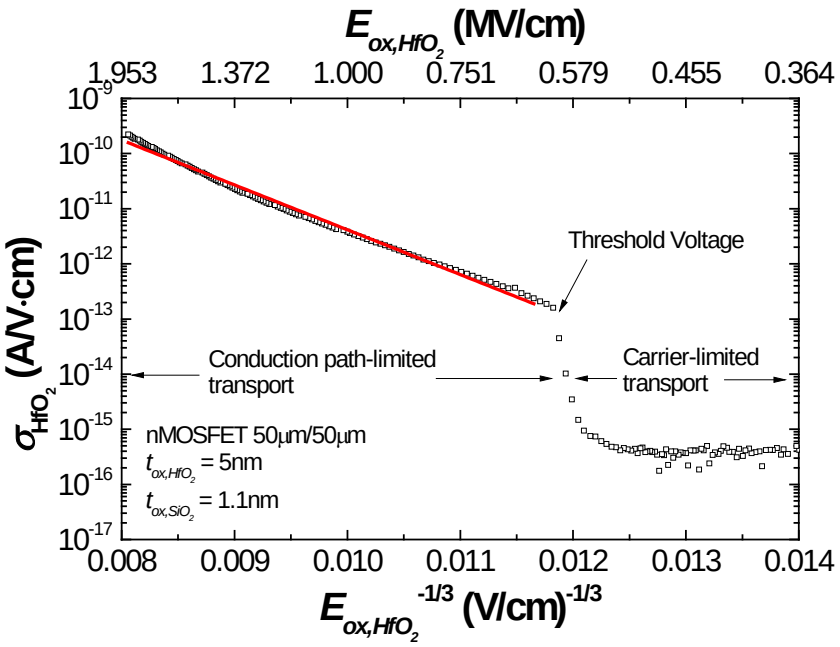


Figure 8.7: Two-dimensional electric field dependent hopping plot of 5nm HfO_2/1.1nm SiO_2 nMOSFET. The conductivity through the HfO_2 at 6K indicates that electric field dependent hopping participates in the gate leakage current. An adjusted R-Squared value greater than 0.99 exists between the data and the linear fit.

Models have been developed for hopping conduction in high electric fields [123, 124]. A model for two-dimensional variable hopping conductivity in an electric field can be written as [124]:

$$\sigma = Ae^{-\left(\frac{2T_0}{T(P+Q)}\right)^{1/3}} \tag{8.7}$$

where

$$P = \frac{1+f/2}{(1+f)^2}, \qquad Q = \frac{3f}{2}+1, \qquad f = \frac{E_{ox,HfO_2}q}{2\alpha k_b T},$$

q is the charge of an electron, and A is a constant weakly dependent on the electric field and temperature [122]. The density of states at the Fermi energy level is given by (8.1). The most probable hop distance is given by [124]:

$$R_{hop} = \frac{1}{3\alpha}\left(\frac{T_0}{T}\right)^{1/3} \tag{8.8}$$

To verify if this model explains the gate voltage dependence of the variable range hopping, the three conductivity terms of (8.3) are plotted in Figure 8.8 to determine where each term dominates.

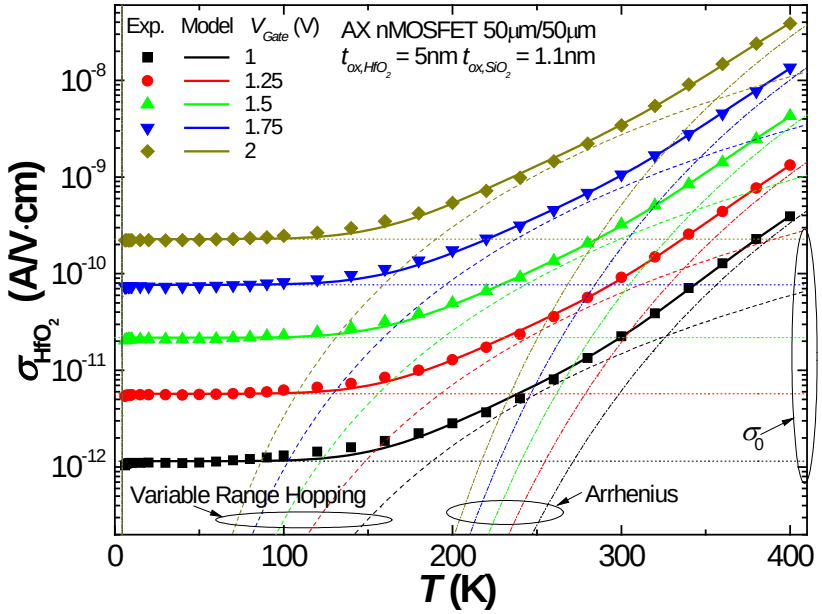


Figure 8.8: Temperature dependence of the conductivity through the HfO_2 for various gate voltages. Experimental data is shown by symbols while the model is shown in lines. The three terms of the model are also plotted separately demonstrating where each component dominates. Only gate voltages in 0.25V increments shown for clarity.

From Figure 8.8, the temperature independent conductivity term (σ_0) dominates for T < 120K. In Figure 8.8, the Arrhenius conductivity term appears to dominate at temperatures near and above 400K, which is the highest temperature measured. Higher temperature measurements are required to determine if the Arrhenius conductivity term

correctly models the high temperature region (T > 380K). The variable range hopping conductivity term dominates for temperatures around 240K.

To determine if high electric field variable range hopping models describe the variable range hopping dominated temperature range, (8.7) was used to fit the data over a temperature range of 200K-260K as a function of temperature and electric field. It is expected that as temperature increases, the density of states about the Fermi energy level in the HfO_2 and the decay length of the localized wave function of the defects do not change with temperature. Therefore, the fitting parameters of A, α, and T_0 should be the same for each temperature investigated. Figure 8.9 shows an example of the best fit obtained for shared parameters (i.e., A, α, and T_0) for each temperature while the inset illustrates the best fits for non-shared parameters (i.e., parameters allowed to vary for each temperature). As observed in Figure 8.9 (not the inset), (8.7) does not describe the complete electric field range investigated (corresponding to V_{Gate} = 0.85V to 2V). The electric field dependent variable range hopping expression correctly models the high field regime but fails to describe the relatively low electric field regime.

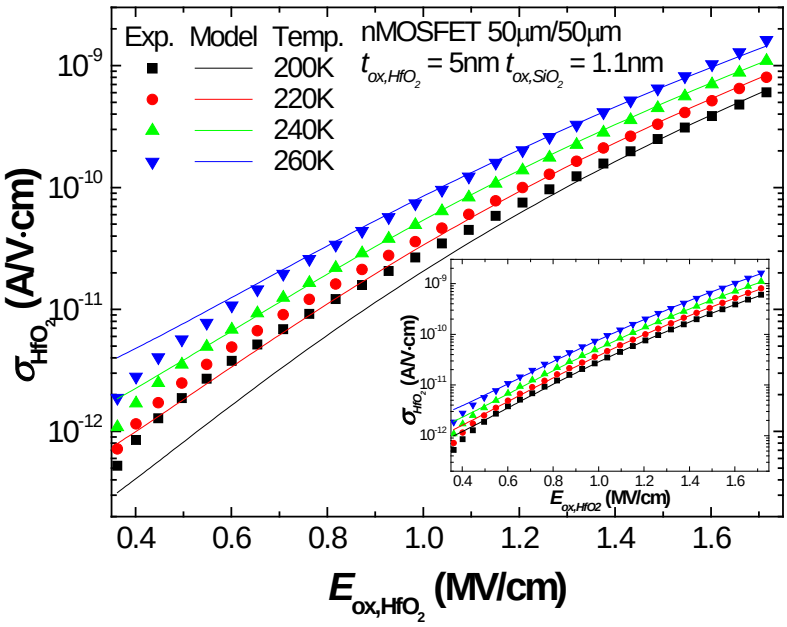


Figure 8.9: Example of conductivity through the HfO_2 as a function of electric field in the HfO_2 200K-260K. Best fit simulations using (8.7) are shown by lines while experimental results are shown by symbols. Only every fifth experimental data point is shown for clarity. The main plot shows best fit results if shared parameters are considered over the temperature range shown while the inset shows bet fit results for parameters not shared across temperatures.

Extracted parameters for the 3nm HfO_2 samples results in T_0 = ~5x10^{6}K and α = ~1.2x10^{7}cm^{-1} (recall 1x10^{7}cm^{-1} was assumed earlier), which using (8.1) and (8.8) corresponds to a defect density of ~7x10^{18}cm^{-3}eV^{-1} and hopping distance of ~7nm. The extracted hopping distance is larger than the physical distance of the dielectric stack (4.1nm in total). Average variable range hopping distances up to six times larger than the thickness of the material in two-dimensional hopping has also been observed in amorphous Si and germanium (Ge) [45]. Although in the presence of an electric field hopping is more likely to occur in the direction of the electric field, hopping is not restricted to be directionally normal to the gate stack plane. Hence, hops greater than the thickness of the dielectric are therefore expected, leading to average variable range hopping distances greater than the thickness of the material. The ratio of the hopping

distance with respect to the material thickness further suggests two-dimensional hopping is present. A rough estimate between two-dimensional and three-dimensional hopping occurs when the material thickness is three to four times larger than the hopping distance [45].

Extracted parameters for the 5nm HfO_2 samples results in a T_0 = ~1x10^{7}K and α = ~2.2x10^{7}cm^{-1}, which results after using (8.1) and (8.8) in $N(E_F)$ = ~9x10^{18}cm^{-3}eV^{-1} and R_{hop} = ~5nm. Notice that the extracted hopping distance is nearly the same as the HfO_2 thickness and just shy of the dielectric stack thickness (6.1nm). This indicates a one or two hop transport through the dielectric stack. Using (8.7) to model the gate leakage resulted in a larger T_0 and smaller defect density compared to using (8.3) for both the 3nm and 5nm HfO_2 stacks.

From Figure 8.9, the best fit achieved for shared parameters between temperatures results in poor fits for the lower E_{ox,HfO_2} values, which indicates the electric field dependent variable hopping model for the 5nm and 3nm samples is limited. The inset in Figure 8.9 shows excellent fits over the entire strong inversion regime investigated but was achieved by considering independent T_0, α, and A terms for each temperature, implying different defect density and defect localization lengths based on temperature, which is unlikely. So while (8.7) can correctly model a particular temperature, it fails when considering multiple temperatures. Fits performed of σ_{HfO_2} versus T show that adding Arrhenius or σ_0 conduction terms to (8.7) provides little benefit in improving fits when holding α constant over various temperature and E_{ox,HfO_2} ranges.

In this chapter, application of two-dimensional variable range hopping models were applied to nMOS HfO_2/SiO_2 bilayer gate stacks to determine if they could explain

the temperature dependence of the gate leakage current over a wide temperature range (6K – 400K). Using a three-conductivity term model that consists of 1) a temperature independent term, 2) variable range hopping term, and 3) Arrhenius term described by (8.3), the temperature dependence of the gate leakage current can be well described from 6K-400K with an adjusted R-squared value greater than 0.999. For variable range hopping models to completely describe the temperature dependence of the gate leakage current, some issues need to be addressed.

1. The physical meaning of the Arrhenius expression in (8.3) which results in an activation energy of ~0.19eV for 3nm HfO_2 samples and ~0.42eV for the 5nm HfO_2 samples. While at large gate biases the activation energy may correspond to thermal emission from defects near the Fermi energy level in the HfO_2 layer to the conduction band at high gate biases, for low positive gate biases and for negative gate biases it is presently unclear what the activation energy of the Arrhenius term might represent.
2. The gate voltage dependence of the fitting parameters σ_{hop} and σ_{Arr} span orders of magnitude for a gate bias range of ~1V to 2V. A link between the fitting parameters and the gate voltage is missing.
3. Using a hopping conduction model for high electric fields in an attempt to explain the variation of the pre-exponential factor with gate voltage falls short of describing the gate leakage current over many temperatures yielding defect densities and decay lengths that vary with temperature.

Despite these limitations, the model successfully describes the temperature dependence of the gate leakage current over a very broad temperature range for both positive and negative gate biases.

CHAPTER 9: SUMMARY AND FUTURE WORK

As new dielectrics, including multilayer dielectrics, continue to be introduced into semiconductor devices, an increased understanding of the materials performance and limitations is necessary. With respect to reliability and power consumption, understanding the charge carrier transport through the dielectric is of particular interest. Of immediate concern to the semiconductor community is the material of HfO_2, which is quickly being integrated into the production line. In this study, the gate leakage current of $TiN/HfO_2/SiO_2/Si$ n/pMOSFETs devices were measured over a temperature range of 6K to 400K using finely spaced temperature steps. The large temperature range over which the devices were measured give an accurate description of the gate leakage current with respect to temperature for which analysis can be performed to determine the carrier transport mechanism. In the next two sections, a brief summary of the important observations and analysis are presented followed by some suggestions for future work.

9.1: Summary

Multilayer dielectrics increase the complexity of carrier transport through the dielectric stack requiring new tools to aid in their understanding. An energy band simulation tool was developed to assist in visualizing complex MOS systems and perform analysis. The simulation tool allows an arbitrary number of dielectrics to be placed in the MOS gate stack and calculates the energy band, electric field, charge distribution, and potential diagrams. Additional analyses include flatband voltage, threshold voltage,

capacitance, tunneling distance, trap charge effects, transmission coefficient, and image charge effects. Versions of the program are available to the public free of charge [94] providing a simulation tool that is a valuable resource to the semiconductor community.

General observations of the temperature dependence of the gate leakage current in HfO_2/SiO_2 gate stacks included:

1. The temperature dependence of the 3 and 5nm HfO_2/1.1nm SiO_2 bilayer stacks is significantly greater than the 2nm SiO_2 control stack (Section 5.1).
2. Carrier transport is dominated by a carrier limited regime in the depletion/weak inversion regime and a transport path-limited regime in the strong inversion and accumulation regimes (Section 5.1.1).
3. Carrier separation analysis of the HfO_2/SiO_2 gate stacks indicates that at positive gate biases the gate leakage current is due to substrate electrons moving to the metal electrode and at negative gate biases substrate holes moving to the metal electrode (Section 5.2).
4. Arrhenius plots of the gate leakage current indicate multiple transport mechanisms are present. A weak temperature dependent regime is identified (T < 50K) that is nearly temperature insensitive and requires measurements at temperatures lower than 50K. Studies only performed to liquid nitrogen temperatures (77K) indicate an activation energy in the meV range for the weak temperature dependent regime (Section 5.3).
5. For devices biased in accumulation, the gate leakage current decreases (especially noticeable in 3nm HfO_2 pMOSFETs) for T < 50K while devices biased in inversion do not. Capacitance-voltage measurements indicate the

reduction of the gate leakage current is due to freeze-out of majority carriers creating a carrier-limited transport regime (Section 5.3).

6. For the strong temperature dependent regime, the Arrhenius slope never saturates (becomes linear) as the temperature increases over the temperature regime investigated. The extracted activation energy is therefore very dependent over the temperature regime from which it is extracted (Section 5.3).
7. The activation energy in the strong temperature dependent regime is dependent on the gate voltage (Section 5.3).

Common mechanisms used to explain the temperature dependence of the gate leakage current are P-F conduction and Schottky emission. The gate voltage dependence (i.e., an electric field dependence) of observation 7 listed above also suggests that P-F conduction or Schottky emission might apply to the temperature dependence of the gate leakage current in the HfO_2/SiO_2 stacks studied. P-F analysis and Schottky analysis, however, do not describe the temperature dependence of the gate leakage current for the HfO_2/SiO_2 bilayer stack studied here. Notable problems with the P-F conduction (Chapter 6) as applied to the bilayer stack are:

1. For the positive gate bias regimes, the P-F plot is nonlinear over a large electric field range for positive gate biases (Figure 6.2).
2. The Arrhenius P-F plots are nonlinear (Figure 6.3 and Figure 6.7) for which many different trap barrier heights can be extracted depending on the temperature range examined. The standard P-F model was developed for mono-energetic defect levels for which a linear Arrhenius P-F plot is

expected. The standard P-F model is therefore only able to describe a small temperature range.

3. Nonlinear ϕ_r's with respect to $E_{ox,HfO_2}^{1/2}$ were obtained for which different ϕ_{PF}'s can be extracted depending on the choice of electric field range used to perform the linear extrapolation (Figure 6.5). The small electric field range (which can then be approximated as linear) is evidence that the P-F conduction mechanism provides an inappropriate explanation to the gate leakage dependence.
4. Thicker 5nm HfO_2 samples resulted in larger extracted trap barrier heights than thin 3nm HfO_2 samples, resulting in a thickness dependence. The standard P-F model does not predict thickness dependence in the leakage current.
5. For 3nm HfO_2 MOSFETs, P-F transport could not explain the transport mechanism trend for gate voltages less than 1V. As the device voltage operating point is in this regime, ~1V and less, the dominate leakage current mechanism is not explained by the standard P-F analysis.
6. For the negative gate bias regime, the extracted trap energy level is non-sensical, yielding a shallow defect energy which requires a large thermal emission of holes from the Si to populate the energy level. This would result in a defect filling-limited transport mechanism and not thermal emission from the defect (P-F) as the limiting mechanism.
7. For 5nm HfO_2 devices at large negative biases ($|E_{ox,HfO_2}^{1/2}| > |-1000|$ $(V/cm)^{1/2}$) and for 3nm HfO_2 devices at small negative biases ($E_{ox,HfO_2} < 700$ $(V/cm)^{1/2}$),

ϕ_r *increases* as the magnitude of the electric field increases. According to the P-F model, ϕ_r should *decrease* as the magnitude of the electric field increases not *increase*.

In addition to examining Schottky emission, an expression for thermionic stimulated tunneling (TAT) current was developed for bilayers where the interfacial layer is sufficiently thin that thermionic stimulation is the limiting conduction mechanism. Notable problems with Schottky and TAT emission as an explanation to the temperature dependence of the gate leakage current in this study are (Chapter 7):

1. Nonlinear (i.e., non-Arrhenius) behavior of the characteristic plot of $\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}$, which is predicted by TAT, (7.9). Nonlinear behavior of the characteristic plots for Schottky emission [$\ln(J_{Gate}/T^2)$ versus $E_{ox,SiO2}{}^{1/2}$].
2. Negative activation energies for temperatures less than ~300K (should be positive).
3. Extracted barrier heights (ϕ_B) that significantly differ from the reported band offsets with Si.

The last transport mechanism that was investigated as an explanation to the temperature dependence of the gate leakage current was hopping conduction (defect mediated transport). Hopping conduction has long been a popular mechanism to explain the temperature dependence of the conductivity in disordered systems within the physics community. Variable range hopping conduction could not describe the transport mechanism over the entire temperature regime. Using a three-term expression, (8.3), composed of a temperature independent term, variable range hopping term, and an

Arrhenius term, the conductivity through both the 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 gate stacks can be described through the entire temperature regime investigated with a reduced *R*-squared value greater than 0.999. Below are some important observations from the fit parameters.

1. For 3nm HfO_2/1.1nm SiO_2 n/pMOSFETs at positive gate biases (substrate electron injection), a defect density of ~10^{20} $cm^{-3}eV^{-1}$ about the Fermi energy level was extracted from $T_0 \approx 3x10^5$K and $\alpha = 10^7$ cm^{-1}.
2. For 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs at positive gate biases (substrate electron injection), a defect density of ~$2x10^{19}$ $cm^{-3}eV^{-1}$ about the Fermi energy level was extracted from $T_0 \approx 8x10^5$K and $\alpha = 10^7$ cm^{-1}.
3. Activation energies of ~0.19eV for the 3nm HfO_2 samples and ~0.42eV for the 5nm HfO_2 samples were extracted from the Arrhenius term for positive gate biases.
4. For negative gate biases (hole substrate injection) of 3nm HfO_2/1.1nm SiO_2 n/pMOSFETs, a defect density of ~$1.5x10^{21}$ $cm^{-3}eV^{-1}$ about the Fermi energy level was extracted.
5. For negative gate biases (hole substrate injection) of 5nm HfO_2/1.1nm SiO_2 n/pMOSFETs, a defect density of ~$3.9x10^{20}$ $cm^{-3}eV^{-1}$ about the Fermi energy level was extracted.
6. Activation energies of ~0.19eV for the 3nm HfO_2 samples and ~0.32eV for the 5nm HfO_2 samples were extracted using the conductivity at negative gate biases (substrate hole injection).

The defect densities extracted correspond well to other values reported in the literature for HfO_2/SiO_2 gate stacks [121]. It is unclear what transport mechanism the activation energies of the Arrhenius term might represent as many different thermally activated mechanisms are expressed by an Arrhenius equation. Additional work is required to identify which Arrhenius mechanism is responsible for the behavior observed. The pre-exponential terms of the variable range hopping term and Arrhenius term have a very strong gate voltage dependence. A link between the pre-exponential terms and the gate voltage needs to be established. Application of high electric field variable range hopping models performed well when modeling a single temperature but failed when considering multiple temperatures. Despite the short comings of the hopping models in describing transport through the HfO_2/SiO_2 gate stack, they do describe the temperature dependence of the gate leakage stack very well. The excellent fit of the temperature dependence of the gate leakage current can be used to empirically model the phenomenon and promotes continued study of the application of hopping models to the temperature dependence of the gate leakage current in novel dielectric stacks.

A significant implication for substrate electron injection is suggested by examining the results of the average hopping distance extracted during hopping analysis (Section 8.4) and the bulge observed and discussed in relation to Figure 5.9 (Section 5.1.2). For the 3nm HfO_2/1.1nm SiO_2 devices, the extracted average hopping distance (~7nm) is greater than the stack thickness (4.1nm), while for the 5nm HfO_2/1.1nm SiO_2 devices it is not. For the 5nm HfO_2/1.1nm SiO_2 devices an average hopping distance of ~5nm was extracted compared to a physical thickness of 6.1nm. In relation to Figure 5.9, a bulge in the temperature dependence of the gate leakage current is observed in the 3nm

HfO_2/1.1nm SiO_2 devices and not in the 5nm HfO_2/1.1nm SiO_2 devices. The existence of a bulge in the 3nm HfO_2 and large hopping distance suggest that the transport mechanism in 3nm HfO_2/1.1nm SiO_2 devices is dominated by the SiO_2 layer, while for the 5nm HfO_2/1.1nm SiO_2 devices the transport mechanism is dominated by the HfO_2 layer. This observation is consistent with a study performed by Bersuker *et al.*'s [125] on TiN/3 or 4nm HfO_2/1.1nm SiO_2/Si gate stacks (which are very similar to the TiN/3nm HfO_2/1.1nm SiO_2/Si gate stacks used in this study), where defects above the Si conduction band edge in the SiO_2 IL were found responsible for increases in the stress induced leakage current. In describing a possible qualitative mechanism for the bulge observed in the temperature dependent leakage data of 3nm HfO_2/1.1nm SiO_2 MOSFETs, the same layer (SiO_2) and energy location of the defects (above the Si conduction band edge) was suggested (see discussion below Figure 5.9). This correlation further suggests, that for substrate electron injection in 3nm HfO_2/1.1nm SiO_2 MOSFETs, carrier transport is dominated by the SiO_2 layer and for 5nm HfO_2/1.1nm SiO_2 MOSFETs, carrier transport is dominated by the HfO_2 layer.

This study of carrier transport in HfO_2/SiO_2 MOS dielectric stacks has many important applications to current and future MOS dielectric stacks for the following reasons.

1. In this study, a bilayer dielectric stack is investigated. Novel dielectrics in MOS devices usually require the presence of more than one dielectric.
2. The presence of multiple dielectrics requires special attention to understand possible carrier transport mechanisms and how to model them. The approach

to analyzing the conduction mechanism(s) in the HfO_2/SiO_2 bilayer dielectric stacks in this study can be used as a template for other dielectric stacks.

3. Software capable of calculating the energy band structure and extracting important parameters developed for this study can be used on dielectric stacks composed of an arbitrary number of dielectrics.
4. Analysis techniques used in this study can also be applied to other novel dielectrics in MOS devices.

In addition to contributing to the understanding the carrier transport through the HfO_2/SiO_2 dielectric stack, this dissertation can be used as a model to analyze dielectric stacks of more than one dielectric and demonstrates the advantages of using characterization techniques over a wide temperature range.

9.2: Future Work

The work in this study provides a solid foundation for continued investigation into the temperature dependence of novel dielectrics and, in particular, the HfO_2/SiO_2 bilayer gate stack. Some particular directions for future work that this study inspires are:

1. Investigating the existence of a link between pre-exponential terms, σ_{hop} and σ_{Arr}, and the gate voltage. High electric field variable range hopping models, which relate hopping conductivity to gate voltage, exist and were primarily developed for *three*-dimensional electric field dependent variable range hopping. Further development of *two*-dimensional electric field dependent variable range hopping models may explain the temperature dependence of the gate leakage current.

2. A significant difference between the temperature dependence of 3nm HfO_2 and 5nm HfO_2 gate stacks were observed. Hopping models suggest a thickness independence in the conductivity. Hence, the SiO_2 interfacial layer and HfO_2/SiO_2 interface play an important part in the gate leakage current of thin HfO_2 samples. Examining the thickness dependence of the temperature dependence of the gate leakage current may further determine the role the SiO_2 layer and HfO_2/SiO_2 interface plays on the temperature dependence of the gate leakage current. Thicker HfO_2 samples may follow hopping models more accurately allowing for better determination of defect densities in the HfO_2 layer and could therefore be used for process optimization.
3. The success of applying DC hopping models to the temperature dependence of the gate leakage current suggest further studies be performed measuring and modeling the AC conductivity of the temperature dependence of the gate leakage current as further confirmation of hopping conduction.
4. The success of modeling the gate leakage current at both positive and negative gate biases using hopping conduction suggests multi-phonon trap assisted tunneling (MPTAT) [126, 127], which is also a defect mediated process that is strongly temperature dependent, may describe the temperature dependence of the gate leakage current. Description of the temperature dependence using MPTAT of the gate leakage current at positive gate biases for 3nm and 5nm HfO_2 nMOSFET used in this study have already been shown to capture important trends in the temperature dependence of the gate leakage current [128].

5. Arrhenius plots of the gate leakage current indicate non-Arrhenius behavior over the temperature regime examined. Modeling the temperature dependence of the gate leakage current using (8.3), however, showed that the Arrhenius equation coupled with variable range hopping equation fits the data very well. As the Arrhenius term dominates at high temperature (Figure 8.8), measuring the gate leakage current to higher temperatures (T > 400K) would help confirm the existence of the Arrhenius mechanism and may provide additional information on which Arrhenius transport mechanism is occurring.
6. In the application of image charge to multilayer dielectrics, a model for a thin layer dielectric sandwiched between two other materials of infinite thickness was used. In multilayer dielectric MOS devices, however, neighboring dielectrics can be quite thin yielding additional interfaces, which were not considered when deriving (4.15). The presence of additional interfaces near the thin dielectric layer will modify the potential slightly in the thin dielectric layer. Due to the inverse decay of the image charge potential with respect to distance, effects of additional interfaces were considered negligible. Further work is necessary to determine how much of an effect additional interfaces will have on the potential of a thin dielectric. Finite element physics solvers such as COMSOL® are suggested as an aid in this study.

REFERENCES

[1] K. Mistry, C. Auth, B. A. Beattie, D. A. Bergstrom, M. A. Bost, M. A. Brazier, M. A. Buehler, A. A. Cappellani, R. A. Chau, C. H. A. Choi, G. A. Ding, K. A. Fischer, T. A. Ghani, R. A. Grover, W. A. Han, D. A. Hanken, M. A. Hattendorf, J. A. He, J. A. Hicks, R. A. Huessner, D. A. Ingerly, P. A. Jain, R. A. James, L. A. Jong, S. A. Joshi, C. A. Kenyon, K. A. Kuhn, K. A. Lee, H. A. Liu, J. A. Maiz, B. A. McIntyre, P. A. Moon, J. A. Neirynck, S. A. Pae, C. A. Parker, D. A. Parsons, C. A. Prasad, L. A. Pipes, M. A. Prince, P. A. Ranade, T. A. Reynolds, J. A. Sandford, L. A. Shifren, J. A. Sebastian, J. A. Seiple, D. A. Simon, S. A. Sivakumar, P. A. Smith, C. A. Thomas, T. A. Troeger, P. A. Vandervoorn, S. A. Williams, and K. A. Zawadzki, "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," presented at Electron Devices Meeting, 2007. IEDM 2007. IEEE International, pp. 247-250 2007.

[2] J. F. Conley, Y. Ono, and D. J. Tweet, "Densification and improved electrical properties of pulse-deposited films via in situ modulated temperature annealing," *Applied Physics Letters*, vol. 84, 2004.

[3] K. Karakaya, A. Zinine, J. G. M. v. Berkum, M. A. Verheijen, Z. M. Rittersma, G. Rijnders, and D. H. A. Blank, "Characterization of Laminated CeO_2-HfO_2 High-*k* Gate Dielectrics Grown by Pulsed Laser Deposition," *Journal of The Electrochemical Society*, vol. 153, pp. F233-F236, 2006.

[4] S. Jeon, J. H. Han, J. H. Lee, S. Choi, H. Hwang, and C. Kim, "High work-function metal gate and high-κ dielectrics for charge trap flash memory device applications," *Electron Devices, IEEE Transactions on*, vol. 52, pp. 2654-2659, 2005.

[5] H.-T. Lue, S.-Y. Wang, E.-K. Lai, Y.-H. Shih, S.-C. Lai, L.-W. Yang, K.-C. Chen, J. Ku, K.-Y. Hsieh, R. Liu, and C.-Y. Lu, "BE-SONOS: A bandgap engineered SONOS with excellent performance and reliability," presented at Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International, pp. 547-550 2005.

[6] Y. Matsui, M. Hiratani, S. Kimura, and I. Asano, "Combining Ta_2O_5 and Nb_2O_5 in Bilayered Structures and Solid Solutions for Use in MIM Capacitors," *Journal of The Electrochemical Society*, vol. 152, pp. F54-F59, 2005.

[7] W. J. Zhu, T.-P. Ma, T. Tamagawa, J. Kim, and Y. Di, "Current transport in metal/hafnium oxide/silicon structure," *Electron Device Letters, IEEE*, vol. 23, pp. 97-99, 2002.

[8] R. G. Southwick, J. Reed, G. Bersuker, and W. Knowlton, "Preliminary Study of Temperature Dependence of TiN/3nm HfO_2/1.1nm SiO_2/Si Gated MOSFETs," presented at IIRW 2007.

[9] L.-F. Mao, "Modeling of temperature dependence of the leakage current through a hafnium silicate gate dielectric in a MOS device," *Semiconductor Science and Technology*, vol. 22, pp. 1203-1208, 2007.

[10] R. G. Southwick III, J. Reed, C. Buu, H. Bui, R. Butler, G. Bersuker, and W. B. Knowlton, "Temperature (5.6-300K) Dependence Comparison of Carrier Transport Mechanisms in HfO_2/SiO_2 and SiO_2 MOS Gate Stacks," *International Integrated Reliability Workshop Final Report*, pp. 48-54, 2008.

[11] C. M. Compagnoni, A. S. Spinelli, A. Bianchini, A. L. Lacaita, S. Spiga, G. Scarel, C. Wiemer, and M. Fanciulli, "Temperature dependence of transient and steady-state gate currents in HfO_2 capacitors," *Applied Physics Letters*, vol. 89, pp. 103504, 2006.

[12] B. Govoreanu, P. Blomme, M. Rosmeulen, J. V. Houdt, and K. D. Meyer, "A model for tunneling current in multi-layer tunnel dielectrics," *Solid-State Electronics*, vol. 47, pp. 1045-1053, 2003.

[13] D. S. Jeong and C. S. Hwang, "Tunneling-assisted Poole-Frenkel conduction mechanism in HfO_2 thin films," *Journal of Applied Physics*, vol. 98, pp. 113701-6, 2005.

[14] H. Tanaka, "Limitation current in Si_3N_4/SiO_2 stacked dielectric films," *Applied Surface Science*, vol. 147, pp. 222-227, 1999.

[15] Y. Takahashi and K. Ohnishi, "Estimation of insulation layer conductance in MNOS structure," *Electron Devices, IEEE Transactions on*, vol. 40, pp. 2006-2010, 1993.

[16] J. C. Tinoco, M. Estrada, B. Iñiguez, and A. Cerdeira, "Conduction mechanisms of silicon oxide/titanium oxide MOS stack structures," *Microelectronics Reliability*, vol. 48, pp. 370-381, 2008.

[17] J. Kolodzey, E. A. Chowdhury, T. N. Adam, Q. Guohua, I. Rau, J. O. Olowolafe, J. S. Suehle, and C. Yuan, "Electrical conduction and dielectric breakdown in aluminum oxide insulators on silicon," *Electron Devices, IEEE Transactions on*, vol. 47, pp. 121-128, 2000.

[18] A. A. Dakhel, "Poole-Frenkel electrical conduction in europium oxide films deposited on Si(100)," *Crystal Research and Technology*, vol. 38, pp. 968-973, 2003.

[19] F.-C. Chiu, Z.-H. Lin, C.-W. Chang, C.-C. Wang, K.-F. Chuang, C.-Y. Huang, J. Y.-m. Lee, and H.-L. Hwang, "Electron conduction mechanism and band diagram of sputter-deposited $Al/ZrO_2/Si$ structure," *Journal of Applied Physics*, vol. 97, pp. 034506-4, 2005.

[20] T. P. Ma, H. M. Bu, X. W. Wang, L. Y. Song, W. He, and M. M. Wang, "Special reliability features for Hf-based high-k gate dielectrics," *IEEE TDMR*, vol. 5, pp. 36-44, 2005.

[21] G. Ribes, S. Bruyere, D. Roy, C. Parthasarthy, M. Muller, M. Denais, V. Huard, T. Skotnicki, and G. Ghibaudo, "Physical origin of Vt instabilities in high-k dielectrics and process optimisation," *Integrated Reliability Workshop Final Report, 2005 IEEE International*, pp. 75-78, 2005.

[22] W. Zhu, T. P. Ma, T. Tamagawa, Y. Di, J. Kim, R. Carruthers, M. Gibson, and T. Furukawa, "HfO_2 and HfAlO for CMOS: thermal stability and current transport," presented at International Electron Device Meeting, pp. 20.4.1-4 2001.

[23] H. W. Chen, S. Y. Chen, K. C. Chen, H. S. Huang, C. H. Liu, F. C. Chiu, K. W. Liu, K. C. Lin, L. W. Cheng, C. T. Lin, G. H. Ma, and S. W. Sun, "Electrical characterization and carrier transportation in Hf-silicate dielectrics using ALD gate stacks for 90 nm node MOSFETs," *Applied Surface Science*, vol. 254, pp. 6127-6130, 2008.

[24] J. Frenkel, "On Pre-Breakdown Phenomena in Insulators and Electronic Semi-Conductors," *Physical Review*, vol. 54, pp. 647-648, 1938.

[25] H. H. Poole, "Dielectric constant of mica in intense fields," *Philosophical Magazine*, vol. 32, pp. 112-129, 1916.

[26] H. H. Poole, "Temperature variation of electrical conductivity of mica," *Philosophical Magazine*, vol. 34, pp. 195-204, 1917.

[27] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, 1981.

[28] D. M. Ramo, J. L. Gavartin, A. L. Shluger, and G. Bersuker, "Spectroscopic properties of oxygen vacancies in monoclinic HfO_2 calculated with periodic and embedded cluster density functional theory," *Physical Review B*, vol. 75, pp. 205336, 2007.

[29] D. M. Ramo, A. L. Shluger, J. L. Gavartin, and G. Bersuker, "Theoretical Prediction of Intrinsic Self-Trapping of Electrons and Holes in Monoclinic HfO_2," *Physical Review Letters*, vol. 99, pp. 155504-4, 2007.

[30] Z. Xu, M. Houssa, S. D. Gendt, and M. Heyns, "Polarity effect on the temperature dependence of leakage current through HfO_2/SiO_2 gate dielectric stacks," *Applied Physics Letters*, vol. 80, pp. 1975-1977, 2002.

[31] S. Dueñas and et al., "Electrical properties of high-pressure reactive sputtered thin hafnium oxide high-k gate dielectrics," *Semiconductor Science and Technology*, vol. 22, pp. 1344, 2007.

[32] D. S. Jeong, H. B. Park, and C. S. Hwang, "Reasons for obtaining an optical dielectric constant from the Poole--Frenkel conduction behavior of atomic-layer-deposited HfO_2 films," *Applied Physics Letters*, vol. 86, pp. 072903-3, 2005.

[33] S. Pan, S.-J. Ding, Y. Huang, Y.-J. Huang, D. W. Zhang, L.-K. Wang, and R. Liu, "High-temperature conduction behaviors of HfO_2/TaN-based metal-insulator-metal capacitors," *Journal of Applied Physics*, vol. 102, pp. 073706-5, 2007.

[34] H. W. Chen, F. C. Chiu, C. H. Liu, S. Y. Chen, H. S. Huang, P. C. Juan, and H. L. Hwang, "Interface characterization and current conduction in HfO2-gated MOS capacitors," *Applied Surface Science*, vol. 254, pp. 6112-6115, 2008.

[35] J. Gavartin and A. Shluger, "*Ab initio* modeling of electron-phonon coupling in high-k dielectrics," *physica status solidi (c)*, vol. 3, pp. 3382-3385, 2006.

[36] J. L. Gavartin, D. M. Ramo, A. Shluger, and G. Bersuker, "Polaron-like Charge Trapping in Oxygen Deficient and Disordered HfO2: Theoretical Insight," *ECS Transactions*, vol. 3, pp. 277-290, 2006.

[37] A. M. Stoneham, J. Gavartin, A. L. Shluger, A. V. Kimmel, D. M. Ramo, H. M. Rønnow, G. Aeppli, and C. Renner, "Trapping, self-trapping and the polaron family," *Journal of Physics: Condensed Matter*, vol. 19, pp. 255208, 2007.

[38] I. Z. Mitrovic, Y. Lu, O. Buiu, and S. Hall, "Current transport mechanisms in $(HfO_2)_x(SiO_2)_{1-x}/SiO_2$ gate stacks," *Microelectronic Engineering*, vol. 84, pp. 2306-2309, 2007.

[39] K. Chatterjee, M. C. Boyer, W. D. Wise, T. Kondo, T. Takeuchi, H. Ikuta, and E. W. Hudson, "Visualization of the interplay between high-temperature superconductivity, the pseudogap and impurity resonances," *Nature Physics*, vol. 4, pp. 108-112, 2008.

[40] R. G. Southwick and W. B. Knowlton, "Stacked dual oxide MOS energy band diagram visual representation program (IRW student paper)," *IEEE Transactions on Device and Materials Reliability*, vol. 6, pp. 136-145, 2006.

[41] K. Y. Cheong, J. H. Moon, H. J. Kim, W. Bahng, and N.-K. Kim, "Current conduction mechanisms in atomic-layer-deposited HfO_2/nitrided SiO_2 stacked gate on 4H silicon carbide," *Journal of Applied Physics*, vol. 103, pp. 084113-8, 2008.

[42] D. K. Schroder, *Semiconductor material and device characterization*, 3rd ed. Hoboken, New Jersey: Wiley, 2006.

[43] B. L. Yang, P. T. Lai, and H. Wong, "Conduction mechanisms in MOS gate dielectric films," *Microelectronics Reliability*, vol. 44, pp. 709-718, 2004.

[44] P. W. Anderson, "Absence of diffusion in certain random lattices," *Physical Review*, vol. 109, pp. 1492-1505, 1958.

[45] N. F. Mott and E. A. Davis, *Electronic processes in non-crystalline materials*, 2nd ed. Oxford: Clarendon Press, 1979.

[46] M. J. Kellicutt, I. S. Suzuki, C. R. Burr, M. Suzuki, M. Ohashi, and M. S. Whittingham, "Variable-range-hopping conduction and the Poole-Frenkel effect in a copper polyaniline vermiculite intercalation compound," *Physical Review B*, vol. 47, pp. 13664, 1993.

[47] N. F. Mott, "Conduction in glasses containing transition metal ions," *Journal of Non-Crystalline Solids*, vol. 1, pp. 1-17, 1968.

[48] N. F. Mott, "Conduction in non-crystalline materials. III. Localized states in a pseudogap and near extremities of conduction and valence bands," *Philosophical Magazine*, vol. 19, pp. 835-852, 1969.

[49] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors*. New York: Springer-Verlag, 1984.

[50] R. M. Hill, "Poole-Frenkel conduction in amorphous solids," *The philosophical magazine*, vol. 23, pp. 59-86, 1971.

[51] J. G. Simmons, "Poole-Frenkel Effect and Schottky Effect in Metal-Insulator-Metal Systems," *Physical Review*, vol. 155, pp. 657, 1967.

[52] J. R. Yeargan and H. L. Taylor, "The Poole-Frenkel Effect with Compensation Present," *Journal of Applied Physics*, vol. 39, pp. 5600-5604, 1968.

[53] P. Y. Yu and M. Cardona, "Fundamentals of Semiconductors - Physics & Materials Properties." Berlin: Springer, 1996, pp. 152-157.

[54] A. Hayes and A. M. Stoneham, *Defects and defect processes in nonmetallic solids*. NY: wiley & Sons, 1985.

[55] A. K. Jonscher, "Electronic properties of amorphous dielectric films," *Thin Solid Films*, vol. 1, pp. 213-234, 1967.

[56] W. R. Harrell and J. Frey, "Observation of Poole-Frenkel effect saturation in SiO_2 and other insulating films," *Thin Solid Films*, vol. 352, pp. 195-204, 1999.

[57] C.-C. Yeh, T. P. Ma, N. Ramaswamy, N. Rocklein, D. Gealy, T. Graettinger, and K. Min, "Frenkel-Poole trap energy extraction of atomic layer deposited Al_2O_3 and Hf_xAl_yO thin films," *Applied Physics Letters*, vol. 91, pp. 113521-3, 2007.

[58] Y. Lu, O. Buiu, S. Hall, I. Z. Mitrovic, W. Davey, R. J. Potter, and P. R. Chalker, "Tuneable electrical properties of hafnium aluminate gate dielectrics deposited by metal organic chemical vapour deposition," *Microelectronics Reliability*, vol. 47, pp. 722-725, 2007.

[59] Y. Lu, O. Buiu, S. Hall, and P. K. Hurley, "Optical and electrical characterization of hafnium oxide deposited by MOCVD," *Microelectronics and Reliability*, vol. 45, pp. 965-968, 2005.

[60] P. J. Elliott, A. D. Yoffe, and E. A. Davis, "Hopping conduction in amorphous germanium," *AIP Conference Proceedings*, vol. 20, pp. 311-319, 1974.

[61] F.-C. Chiu, "Interface characterization and carrier transportation in metal/HfO_2/silicon structure," *Journal of Applied Physics*, vol. 100, pp. 114102-5, 2006.

[62] G. Bersuker, J. Peterson, J. Barnett, A. Korkin, J. H. Sim, R. Choi, B. H. Lee, J. Greer, P. Lysaght, and H. R. Huff, "Properties of the interfacial layer in the high-k gate stack and transistor performance," *Proc. of ECS Spring Meeting*, vol. 2005-05, pp. 141, 2005.

[63] J. R. Hauser and K. Ahmed, "Characterization of ultra-thin oxides using electrical C-V and I-V measurements," *AIP Conference Proceedings*, vol. 449, pp. 235-239, 1998.

[64] G. Bersuker, J. Peterson, J. Barnett, A. Korkin, J. H. Sim, R. Choi, B. H. Lee, J. Greer, P. Lysaght, and H. R. Huff, "Properties of the interfacial layer in the high-k gate stack and transistor performance," presented at Proc. of ECS Spring Meeting, pp. 141-145 2005.

[65] W. Mizubayashi, N. Yasuda, H. Ota, H. Hisamatsu, K. Tominaga, K. Iwamoto, K. Yamamoto, T. Horikawa, T. Nabatame, and A. Toriumi, "Carrier separation analysis for clarifying leakage mechanism in unstressed and stressed $HfAlO_x/SiO_2$ stack dielectric layers," *Reliability Physics Symposium Proceedings, 2004. 42nd Annual. 2004 IEEE International*, pp. 188-193, 2004.

[66] S. Takagi, N. Yasuda, and A. Toriumi, "Experimental evidence of inelastic tunneling in stress-induced leakage current," *Electron Devices, IEEE Transactions on*, vol. 46, pp. 335-341, 1999.

[67] C. Chang, C. Hu, and R. W. Brodersen, "Quantum yield of electron impact ionization in silicon," *Journal of Applied Physics*, vol. 57, pp. 302-309, 1985.

[68] R. H. Fowler and L. W. Nordheim, "Electron emission in intense electric fields," *Roc. Royal Society London A*, vol. 119, pp. 173, 1928.

[69] E. L. Murphy and J. Good, R.H., "Thermionic emission, field emission, and the transition region," *Physical Review*, vol. 102, pp. 1464-1473, 1956.

[70] J. Frenkel, "On pre-breakdown phenomena in insulators and electronic semi-conductors," *Physical Review*, vol. 54, pp. 647-648, 1938.

[71] R. G. Southwick III, J. Reed, C. Buu, R. Butler, G. Bersuker, and W. B. Knowlton, "Limitations of Poole-Frenkel Conduction in Bilayer HfO_2/SiO_2 MOS Devices," *IEEE Transactions on Device and Materials Reliability*, vol. Accepted for publication, 2010.

[72] J. G. Simmons, "Richardson-Schottky effects in solids," *Physical Review Letters*, vol. 15, pp. 967-968, 1965.

[73] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-k gate dielectrics: Current status and materials properties considerations," *Journal of Applied Physics*, vol. 89, pp. 5243-5275, 2001.

[74] K. Prall, "Scaling non-volatile memory below 30nm," presented at Non-Volatile Semiconductor Memory Workshop, 2007 22nd IEEE, pp. 5-10 2007.

[75] M. H. White, D. A. Adams, and J. Bu, "On the go with SONOS," *Circuits and Devices Magazine, IEEE*, vol. 16, pp. 22-31, 2000.

[76] S. J. Rhee, F. Zhu, H. S. Kim, C. Y. Kang, C. H. Choi, M. H. Zhang, T. Lee, I. Ok, S. A. Krishnan, and J. C. Lee, "Systematic analysis of silicon oxynitride interfacial layer and its effects on electrical characteristics of high-k HfO2 transistor," *Applied Physics Letters*, vol. 88, 2006.

[77] R. G. Southwick III and W. B. Knowlton, "Stacked dual oxide MOS energy band diagram visual representation program (IRW Student Paper)," *IEEE Transactions on Device and Materials Reliability*, vol. 6, pp. 136-145, 2006.

[78] M. J. van Dort, P. H. Woerlee, and A. J. Walker, "A simple model for quantisation effects in heavily-doped silicon MOSFETs at inversion conditions," *Solid-State Electronics*, vol. 37, pp. 411-414, 1994.

[79] Y. Tsividis, *Operation and Modeling of The MOS Transistor*, 2nd ed. USA: Oxford University Press, 2003.

[80] R. Williams, "Image force interactions at the interface between an insulator and a semiconductor," *Journal of Applied Physics*, vol. 45, pp. 1239-1242, 1974.

[81] M. Kleefstra and G. C. Herman, "Influence of the image force on the band gap in semiconductors and insulators," *Journal of Applied Physics*, vol. 51, pp. 4923-4926, 1980.

[82] H. Iwamoto, V. M. Aquino, and V. C. Aguilera-Navarro, "Scattering coefficients for a trapezoidal potential," *International Journal of Theoretical Physics*, vol. 42, pp. 1795-1807, 2003.

[83] J. Walker and J. Gathright, "A transfer-matrix approach to one-dimensional quantum mechanics using Mathematica," *Computers in Physics*, vol. 6, pp. 393-399, 1992.

[84] H. Kroemer, *Quantum Mechanics for Engineering, Materials Science, and Applied Physics*. New Jersey: Prentice Hall, 1994.

[85] M. P. Warusawithana, C. Cen, C. R. Sleasman, J. C. Woicik, Y. Li, L. F. Kourkoutis, J. A. Klug, H. Li, P. Ryan, L.-P. Wang, M. Bedzyk, D. A. Muller, L.-Q. Chen, J. Levy, and D. G. Schlom, "A ferroelectric oxide made directly on silicon," *Science*, vol. 324, pp. 367-70, 2009.

[86] M. Stengel, D. Vanderbilt, and N. A. Spaldin, "Enhancement of ferroelectricity at metal-oxide interfaces," *Nature Materials*, vol. 8, pp. 392-397, 2009.

[87] C. Ang and Z. Yu, "dc electric-field dependence of the dielectric constant in polar dielectrics: Multipolarization mechanism model," *Physical Review B*, vol. 69, pp. 174109, 2004.

[88] H. Takagi, S. I. Rybchenko, Y. Fujishiro, and M. Awano, "Effect of interface passivation on the magnetoresistance of granular magnetite Fe_3O_4," *Applied Physics Letters*, vol. 89, pp. 132509, 2006.

[89] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*. New York: Cambridge University Press, 1998.

[90] S. Shaimeev, V. Gritsenko, K. Kukli, H. Wong, E. H. Lee, and C. Kim, "Single Band Electronic Conduction in Hafnium Oxide Prepared by Atomic Layer Deposition," presented at 2005 IEEE Conference on Electron Devices and Solid-State Circuits, pp. 703-706 2005.

[91] R. G. Southwick III, J. Reed, C. Buu, R. Butler, G. Bersuker, and W. B. Knowlton, "Limitations of Poole-Frenkel Conduction in Bilayer HfO_2/SiO_2 MOS Devices," *IEEE Transactions on Device and Materials Reliability*, vol. 10, pp. 201-207, 2010.

[92] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high-*k* dielectrics reliability issues," *IEEE Transactions on Device and Materials Reliability*, vol. 5, pp. 5-19, 2005.

[93] W. Zhu, T. P. Ma, T. Tamagawa, Y. Di, J. Kim, R. Carruthers, M. Gibson, and T. Furukawa, "HfO_2 and HfAlO for CMOS: thermal stability and current transport," *Electron Devices Meeting, 2001. IEDM Technical Digest. International*, pp. 20.4.1-20.4.4, 2001.

[94] "Boise State University Energy Band Diagram Program Webpage," http://coen.boisestate.edu/bknowlton/research/

[95] R. Ongaro and A. Pillonnet, "Poole-Frenkel (PF) effect high field saturation," *Revue de Physique Applique*, vol. 24, pp. 1085-1095, 1989.

[96] W. R. Harrell, T. H. Cordella, and K. F. Poole, "Implications of non-linear Poole-Frenkel plots on high-k dielectric leakage," *ECS Transactions*, vol. 1, pp. 705-716, 2006.

[97] N. F. Mott, "Metal-Insulator Transition," *Reviews of Modern Physics*, vol. 40, pp. 77-83, 1968.

[98] J. L. Gavartin, D. M. Ramo, A. L. Shluger, G. Bersuker, and B. H. Lee, "Negative oxygen vacancies in HfO_2 as charge traps in high-k stacks," *Applied Physics Letters*, vol. 89, pp. 082908-3, 2006.

[99] H. Wang, Y. Wang, J. Zhang, C. Ye, H. B. Wang, J. Feng, B. Y. Wang, Q. Li, and Y. Jiang, "Interface control and leakage current conduction mechanism in HfO_2 film prepared by pulsed laser deposition," *Applied Physics Letters*, vol. 93, pp. 202904-3, 2008.

[100] S. O. Kasap, *Principles of Electronic Materials and Devices*, 2nd ed. New York: McGraw-Hill, 2002.

[101] O. W. Richardson, "Radiation of ions from wire carrying oscillatory current," *Proceedings of the Cambridge Philosophical Society*, vol. 11, pp. 168-178, 1901.

[102] O. W. Richardson, *The emission of electricity from hot bodies*, 2nd ed. London: Longmans, Green and Co., 1921.

[103] Y. J. Cho, H. M. Cho, Y. W. Lee, H. Y. Lee, I. W. Lee, S. K. Lee, J. W. Sun, S. Y. Moon, H. K. Chung, H. Y. Pang, S. J. Kim, and S. Y. Kim, "Optical characterization of silicon dioxide layers grown on silicon under different growth conditions," *Thin Solid Films*, vol. 313-314, pp. 292-297, 1998.

[104] X. Liu, S. Ramanathan, A. Longdergan, A. Srivastava, E. Lee, T. E. Seidel, J. T. Barton, D. Pang, and R. G. Gordonb, "ALD of hafnium oxide thin films from tetrakis(ethylmethylamino) hafnium and ozone," *Journal of The Electrochemical Society*, vol. 152, pp. G213-G219, 2005.

[105] A. Kerber, E. Cartier, L. Pantisano, R. Degraeve, T. Kauerauf, Y. Kim, A. Hou, G. Groeseneken, H. E. Maes, and U. Schwalke, "Origin of the threshold voltage instability in SiO_2/HfO_2 dual layer gate dielectrics," *Electron Device Letters, IEEE*, vol. 24, pp. 87-89, 2003.

[106] A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, "Characterization of the V_T-instability in SiO_2/HfO_2 gate dielectrics," *IEEE Reliability Physics Symposium Proceedings*, pp. 41-45, 2003.

[107] D. M. Ramo, J. L. Gavartin, A. L. Shluger, and G. Bersuker, "Intrinsic and defect-assisted trapping of electrons and holes in HfO2: an ab initio study," *Microelectronic Engineering*, vol. 84, pp. 2362-2365, 2007.

[108] J. Wu, L. R. Register, and E. Rosenbaum, "Trap-assisted tunneling current through ultra-thin oxide," *IEEE International Reliability Physics Symposium*, pp. 389-395, 1999.

[109] O. Blank, H. Reisinger, R. Stengl, and M. Gutsche, "A model for multistep trap-assisted tunneling in thin high-k dielectrics," *Journal of Applied Physics*, vol. 97, pp. 0144107-1-7, 2005.

[110] M. Pollak and B. Shklovskii, *Hopping Transport in Solids (Modern Problems in Condensed Matter Sciences)*: North Holland, 1991.

[111] V. Ambegaokar, B. I. Halperin, and J. S. Langer, "Hopping Conductivity in Disordered Systems," *Physical Review B*, vol. 4, pp. 2612, 1971.

[112] K. Okada and K. Taniguchi, "Electrical stress-induced variable range hopping conduction in ultrathin silicon dioxides," *Applied Physics Letters*, vol. 70, pp. 351-353, 1997.

[113] S. Banerjee, B. Shen, I. Chen, J. Bohlman, G. Brown, and R. Doering, "Conduction mechanisms in sputtered Ta[sub 2]O[sub 5] on Si with an interfacial SiO[sub 2] layer," *Journal of Applied Physics*, vol. 65, pp. 1140-1146, 1989.

[114] M. L. Knotek, M. Pollak, T. M. Donovan, and H. Kurtzman, "Thickness Dependence of Hopping Transport in Amorphous-Ge Films," *Physical Review Letters*, vol. 30, pp. 853, 1973.

[115] F. R. Allen and C. J. Adkins, "Electrical Conduction in Heavily Doped Germanium," *Philosophical Magazine*, vol. 26, pp. 1027-1042, 1972.

[116] R. Mansfield, "Hopping Conduction in III-V Compunds," in *Hopping Transport in Solids*, vol. 28, M. Pollak and B. Shklovskii, Eds. Amsterdam: Elsevier Science, 1991, pp. 349-375.

[117] S. M. Sze, "Current Transport and Maximum Dielectric Strength of Silicon Nitride Films," *Journal of Applied Physics*, vol. 38, pp. 2951-2956, 1967.

[118] H. J. Motulsky and A. Christopoulos, *Fitting Models to Biological Data using Linear and Nonlinear Regression. A practical guide to curve fitting*. Sand Diago CA: GraphPad Software Inc., 2003.

[119] A. Yildiz, S. B. Lisesivdin, M. Kasap, and D. Mardare, "High temperature variable-range hopping conductivity in undoped TiO_2 thin film," *Optoelectronics and Advanced Materials*, vol. 1, pp. 531-533, 2007.

[120] A. J. Lewis, "Conductivity and thermoelectric power of amorphous germanium and amorphous silicon," *Physical Review B*, vol. 13, pp. 2565, 1976.

[121] H. D. Xiong, D. Heh, M. Gurfinkel, Q. Li, Y. Shapira, C. Richter, G. Bersuker, R. Choi, and J. S. Suehle, "Characterization of electrically active defects in high-k gate dielectrics by using low frequency noise and charge pumping measurements," *Microelectronic Engineering*, vol. 84, pp. 2230-2234, 2007.

[122] N. Apsley and H. P. Hughes, "Temperature- and field-dependence of hopping conduction in disordered systems, II," *Philosophical Magazine*, vol. 31, pp. 1327-1339, 1975.

[123] M. Singh, R. B. Thompson, and O. Dumas, "Electric-field-dependent variable-range hopping conductance in quasi-two-dimensional systems: Application to $PrBa_2Cu_3O_{7-y}$-based superconductor-normal-metal-superconductor junctions," *Physical Review B*, vol. 53, pp. 6806, 1996.

[124] U. Kabasawa, Y. Tarutani, M. Okamoto, T. Fukazawa, A. Tsukamoto, M. Hiratani, and K. Takagi, "Size effect on variable-range-hopping transport in $PrBa_2Cu_3O_{7-x}$," *Physical Review Letters*, vol. 70, pp. 1700, 1993.

[125] G. Bersuker, D. Heh, C. Young, H. Park, P. Khanal, L. Larcher, A. Padovani, P. Lenahan, J. Ryan, B. H. Lee, H. Tseng, and R. Jammy, "Breakdown in the metal/high-k gate stack: Identifying the "weak link" in the multilayer dielectric," presented at Electron Devices Meeting, 2008. IEDM 2008. IEEE International, pp. 1-4 2008.

[126] A. Schenk and H. Hermann, "A new model for long term charge loss in EPROM's," *Ext. Abstract Int. Conf. Solid State Devices Materials (SSDM)*, pp. 494-496, 1994.

[127] L. Larcher, "Statistical Simulation of Leakage Currents in MOS and Flash Memory Devices With a New Multiphonon Trap-Assisted Tunneling Model," *IEEE Transactions on Electron Devices*, vol. 50, pp. 1246-1253, 2003.

[128] L. Vandelli, A. Padovani, L. Larcher, R. G. Southwick III, W. B. Knowlton, and G. Bersuker, "Modeling Temperature Dependency (6 - 400K) of the leakage current through the SiO_2/high-K stacks," *ESSDERC*, vol. accepted, 2010.

[129] S. Elliott, *The Physics and Chemistry of Solid*. Chichester: Wiley, 1998.

[130] J. Ranninger, "Introduction to polaron physics: Basic concepts and models," in *Polarons in Bulk Systems*. Amsterdam: IOS Press, 2006, pp. 476.

[131] T. Holstein, "Studies of polaron motion. II. The "small" polaron," *Annals of Physics*, vol. 8, pp. 343-389, 1959.

[132] D. Emin and A. M. Kriman, "Transient small-polaron hopping motion," *Physical Review B*, vol. 34, pp. 7278-7289, 1986.

[133] L. D. Landau, "Über die Bewegung der Elektronen in Kristalgitter," *Phys. Z. Sowjetunion* vol. 3, pp. 644-645, 1933.

[134] F. B. McLean, H. E. Boesch Jr., and J. M. McGarrity, "Field-dependent hole transport in amorphous SiO_2," *Applications of Cryogenic Technology*, pp. 19-23, 1978.

[135] N. T. Fourches, "Charge buildup by irradiation in metal-oxide-semiconductor structures at cryogenic temperatures: Basic mechanisms and influence of dose and dose rate," *Physical Review B*, vol. 55, pp. 7641, 1997.

[136] H. E. Boesch, Jr. and F. B. McLean, "Hole transport and trapping in field oxides," *IEEE Transactions on Neclear Science*, vol. NS-32, pp. 390-5, 1985.

[137] Y. Liu, A. Shanware, L. Colombo, and R. Dutton, "Modeling of charge trapping induced threshold-voltage instability in high-*k* gate dielectric FETs," *Electron Device Letters, IEEE*, vol. 27, pp. 489-491, 2006.

[138] J. Reed, R. Southwick III, and W. B. Knowlton, "Experimental Evidence for Polaron Hopping Conduction in HfO2 - A Cryogenic Study," in *International Integrated Reliability Workshop poster presentation*: IEEE, 2008.

[139] Y. Zheng, C. Troadec, A. T. S. Wee, K. L. Pey, S. J. O'Shea, and N. Chandrasekhar, "BEEM studies on metal high *K*-dielectric HfO_2 interfaces," *Journal of Physics: Conference Series*, vol. 61, 2007.

[140] W. C. Wu, C. S. Lai, T. M. Wang, J. C. Wang, C. W. Hsu, M. W. Ma, and T. S. Chaoa, "Current transport mechanism for HfO_2 gate dielectrics with fluorine incorporation," *Electrochemical and Solid-State Letters*, vol. 11, pp. H15-H18, 2008.

[141] S. Monaghan, P. K. Hurley, K. Cherkaoui, M. A. Negara, and A. Schenk, "Determination of electron effective mass and electron affinity in HfO2 using MOS and MOSFET structures," *Solid-State Electronics*, vol. 53, pp. 438-444, 2009.

[142] N. Yang, W. K. Henson, and J. J. Wortman, "A Comparative Study of Gate Direct Tunneling and Drain Leakage Currents in N-MOSFET's with Sub-2-nm Gate Oxides," *IEEE Transactions on Electron Devices*, vol. 47, pp. 1636-1644, 2000.

[143] A. Ghetti, "Characterization and modeling of the tunneling current in Si-SiO2-Si structures with ultra-thin oxide layer," *Microelectronic Engineering*, vol. 59, pp. 127-136, 2001.

APPENDIX A

A.1 Introduction to Polarons

A.1 Introduction to Polarons

The presence of a charged carrier will distort its surroundings to some extent [45] leading to electron-phonon interactions. If the distortion is large, the rigid-band approximation fails and coupling between the lattice (phonons) and the electron must be considered [129]. An example of such an interaction is in polar ionic crystals. An electron in such a material will act to polarize the immediate surrounding lattice. To reduce the energy of the electron-lattice interaction, a quantized deformation of the lattice occurs resulting in phonons or a phonon cloud around the charged carrier; together they from a quantized quasi-particle called a polaron. As the electron moves, it will carry the phonon cloud with it and in some instances act like a free particle with a higher effective mass [129].

Polarons are divided into two categories: large polarons or Fröhlich polarons and small polarons or Holstein polarons [130, 131]. In large polarons, the radius of the polaron (the extent of the charge carrier phonon interaction) extends over many atoms. For small polarons, the phonon-charged carrier interaction is limited to the immediate surrounding atoms. Small polarons can sometimes be called “self-trapped” electrons or holes if the charge carrier “digs” a potential well that it cannot escape due to the rearrangement of the surrounding atoms. Extrinsic traps, such as oxygen vacancies in HfO_2, can mediate polaron transport in a material. The extrinsic trap can confine the carrier for sufficiently long periods of time increasing the probability of lattice relaxation and thus enhancing the electron-phonon interaction [132]. In HfO_2 a relaxation energy of ~0.5 to 0.7eV is theorized for the negative charge state [28].

Although small polarons have sparked immense interest in the physics community since the idea was first proposed by Landau [133] and later formalized by Holstein in 1959 [131], polaron conduction has not been extensively applied to carrier transport in MOSFET gate stack with perhaps the exception of irradiated SiO_2 [134-136]. With the introduction of high-*k* dielectrics into the gate stack, a renewed interest in the investigation of polarons is being observed. Currently the investigation of polaron conduction in high-*k* MOSFET stacks is limited to a few groups and a handful of studies [28, 29, 36, 37, 98, 137, 138]. Most of the work performed has been theoretical with very few experimental studies.

Since phonons are directly coupled to temperature, variable temperature studies are ideally suited to aid in the investigation of polaron conduction in high-*k* dielectrics. Of the experimental work performed in polaronic transport, only one study has been reported, to the author's knowledge, in which multiple temperatures were examined and the author took part in that study [138].

APPENDIX B

B.1 SILVACO Input Parameters Without Quantum Mechanical Correction

B.1 SILVACO Input Parameters Without Quantum Mechanical Correction

```
go atlas

mesh
x.m l=0 spac=1
x.m l=1 spac=1
y.m l=-0.02 spac=0.00005
y.m l=0 spac=0.00005
y.m l=0.1 spac=0.0001
y.m l=0.5 spac=0.02

region num=1 silicon y.min=0
region num=2 oxide y.max=0 y.min=-0.0011
region num=3 nitride y.max=-0.0011 y.min=-0.0061

electrode substrate
electrode name=gate x.min=0 x.max=1 y.max=-0.0061 y.min=-0.02

doping uniform conc=1e18 p.type

save outf=TiNHfO2SiO2Si.str

go atlas

mesh inf=TiNHfO2SiO2Si.str

MATERIAL material=oxide  eg300=8.9 affinity=0.95 permit=6.13
MATERIAL material=nitride eg300=5.7 affinity=2.65 permit=20
MATERIAL material=silicon affinity=4.05

contact name=gate  workfunction=4.45
models cvt srh

method carriers=0
output con.band val.band

solve init

solve vgate=0 vstep=-0.1 vfinal=-2 name=gate
save outf=TiNHfO2SiO2SiNeg2V.str
tonyplot TiNHfO2SiO2SiNeg2V.str
solve vgate=-2.01

log outf=TiNHfO2SiO2SiNeg2toPos2.log master

PROBE NAME=SurfPot X=0.1 Y=0.0 POTENTIAL
PROBE NAME=BulkPot X=0.1 Y=0.3 POTENTIAL
PROBE NAME=EoxSiO2 X=0.1 Y=-0.0005 FIELD DIR=90
PROBE NAME=EoxHfO2  X=0.1 Y=-0.004 FIELD DIR=90

solve vstep=0.01 vfinal=2.0 name=gate
save outf=TiNHfO2SiO2SiPos2V.str

### plot results
tonyplot TiNHfO2SiO2SiPos2V.str
tonyplot   TiNHfO2SiO2SiNeg2toPos2.log

quit
```

B.2 SILVACO Input Parameters With Quantum Mechanical Correction

B.2 SILVACO Input Parameters With Quantum Mechanical Correction

```
go atlas

mesh
x.m l=0 spac=1
x.m l=1 spac=1
y.m l=-0.02 spac=0.00005
y.m l=0 spac=0.00005
y.m l=0.1 spac=0.0001
y.m l=0.5 spac=0.02

region num=1 silicon y.min=0
region num=2 oxide y.max=0 y.min=-0.0011
region num=3 nitride y.max=-0.0011 y.min=-0.0061

electrode substrate
electrode name=gate x.min=0 x.max=1 y.max=-0.0061 y.min=-0.02

doping uniform conc=1e18 p.type

save outf=qm_TiNHfO2SiO2Si.str

go atlas

mesh inf=qm_TiNHfO2SiO2Si.str

MATERIAL material=oxide  eg300=8.9 affinity=0.95 permit=6.13
MATERIAL material=nitride eg300=5.7 affinity=2.65 permit=20
MATERIAL material=silicon affinity=4.05

contact name=gate  workfunction=4.45
models fermi schro new.eig ox.poisson qy.min=-0.0015 qy.max=0.5 qx.min=0
       qx.max=1 ox.schro fixed.fermi

method carriers=0
output con.band val.band t.quantum eigens=4

solve init

solve vgate=0 vstep=-0.1 vfinal=-2 name=gate
save outf=qm_TiNHfO2SiO2SiNeg2V.str
tonyplot qm_TiNHfO2SiO2SiNeg2V.str
solve vgate=-2.01

log outf=qm_TiNHfO2SiO2SiNeg2toPos2.log master

PROBE NAME=SurfPot X=0.1 Y=0.0 POTENTIAL
PROBE NAME=BulkPot X=0.1 Y=0.3 POTENTIAL
PROBE NAME=EoxSiO2 X=0.1 Y=-0.0005 FIELD DIR=90
PROBE NAME=EoxHfO2  X=0.1 Y=-0.004 FIELD DIR=90

solve vstep=0.01 vfinal=2.0 name=gate
save outf=qm_TiNHfO2SiO2SiPos2V.str

### plot results
tonyplot   qm_TiNHfO2SiO2SiNeg2toPos2.log
tonyplot qm_TiNHfO2SiO2SiPos2V.str
quit
```

APPENDIX C

C.1 Transmission Coefficient Analysis

C.1 Transmission Coefficient Analysis

In a study performed by Ling-Feng Mao on poly-Si/HfSiO /Si, the temperature dependence of the gate leakage current (100K-300K) was modeled using quantum mechanical tunneling (no defect mediated tunneling was considered) [9]. The temperature range used in Mao's study is similar to the temperature range in this study (6K-400K) and while performed on a different gate stack begs the question "can quantum mechanical tunneling explain the temperature dependence of the gate leakage current?" If the answer is no, it suggests defect mediated transport is responsible for the gate leakage current and the temperature dependence (recall that carrier emission over the barrier was shown not to be responsible for the gate leakage current in Chapter 7).

As was discussed in Section 2.1, to first order, tunneling is temperature independent but many secondary parameters depend on temperature, such as effective mass, carrier energy and population levels, band offsets, etc. The quantum model developed in Section 4.1.6 only calculates the transmission coefficient through the dielectric barriers; it does not calculate the carrier energy levels or population density quantum mechanically.[20] However, comparing the transmission coefficient between the 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 gate stacks can provide insight into whether quantum mechanical tunneling is responsible for the temperature dependence of the gate leakage current.

Consider Figure C.1 where the energy band diagram for a 3nm HfO_2 (thick lines)/1.1nm SiO_2 nMOSFET and 5nm HfO_2 (thin lines)/1.1nm nMOSFET are shown.

[20] The energy band simulation tool uses classical mechanics and Boltzmann's statistics as an approximation to Fermi-Dirac statistics to calculate the carrier density in the semiconductor.

For both the 3nm and 5nm HfO_2 devices, the stack is biased such that the electric field in the HfO_2 layer is 2MV/cm. Due to the same electric field bias and that the SiO_2 IL and Si substrate is the same for both layers, the band bending in the SiO_2 and Si substrate is the same as shown in Figure C.1. To first order, the quantized energy states in the Si inversion layer and population density are the same for the 3nm and 5nm HfO_2 devices at this bias condition. Also notice that the tunneling distance for electrons from the substrate through both the SiO_2 (direct tunneling) and HfO_2 (Fowler-Nordheim tunneling) at or above the Si Fermi energy level is the same for the 3nm and 5nm HfO_2 devices. Given these similarities, one should expect that, *if* the gate leakage current is due to quantum mechanical tunneling, *then* the gate leakage current for 3nm HfO_2 nMOSFETs biased at 1.8V ($E_{ox,HfO2}$ = 1.98MV/cm) should equal the gate leakage current of 5nm HfO_2 nMOSFETs biased at 2.2V ($E_{ox,HfO2}$ = 1.98MV/cm). Here the aforementioned premise is considered.

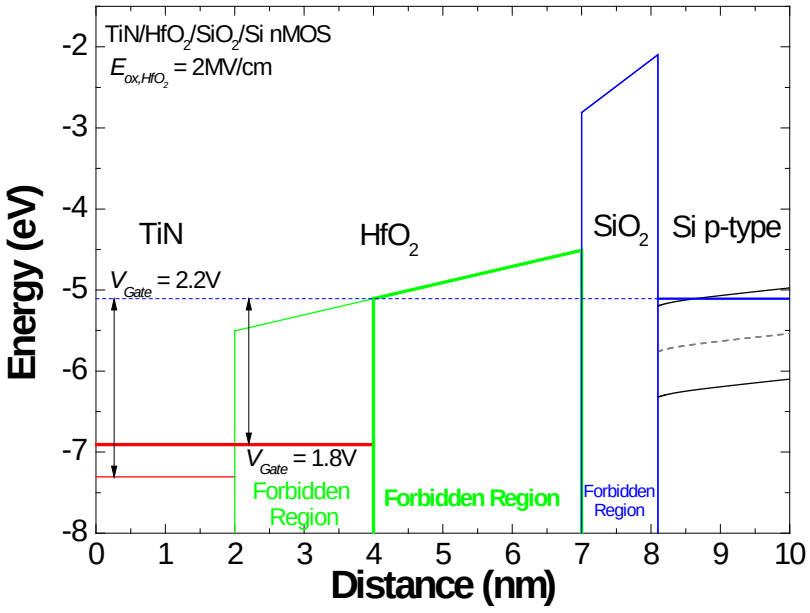


Figure C.1: Energy band diagram for 3nm HfO_2 (thick lines) and 5nm HfO_2 (thin lines) nMOSFETs with a 1.1nm SiO_2 IL biased at an electric field of 2MV/cm in the HfO_2 layer. The tunneling distance for both stacks is identical.

Consider the gate leakage current of 3nm HfO_2/1.1nm SiO_2 and 5nm HfO_2/1.1nm SiO_2 at the four $E_{ox,HfO2}$ biases shown in Figure C.2. A difference in the gate leakage current between the 3nm and 5nm HfO_2 devices that is close to two orders of magnitude is observed in Figure C.2 for equivalent $E_{ox,HfO2}$ biases. The forbidden region of the barrier (labeled in Figure C.1) for equivalent $E_{ox,HfO2}$ larger than 2MV/cm in the 3nm and 5nm HfO_2 stacks is the same. While the forbidden region of the barrier is the same for the 3nm and 5nm nMOSFETs, shown in Figure C.1, outside the forbidden region the barrier is not the same on the TiN/HfO_2 side. For the 3nm HfO_2 case, the electron exits into the TiN layer and for the 5nm HfO_2 case the electron exits into the HfO_2 conduction band. Even if higher electric fields are considered, the potential barrier outside the forbidden region is not identical. The transmission coefficient is strongly dependent on the thickness and height of the forbidden region of the barrier and to a lesser extent the shape

of the barrier outside the forbidden region. Does the shape of the barrier outside the forbidden region modify the transmission coefficient enough to account for a near two order magnitude difference in the gate leakage current? To quantify the effect this will have on the transmission coefficient, the energy band simulation tool is used to calculate the transmission coefficient for the 3nm and 5nm HfO_2 devices for comparison.

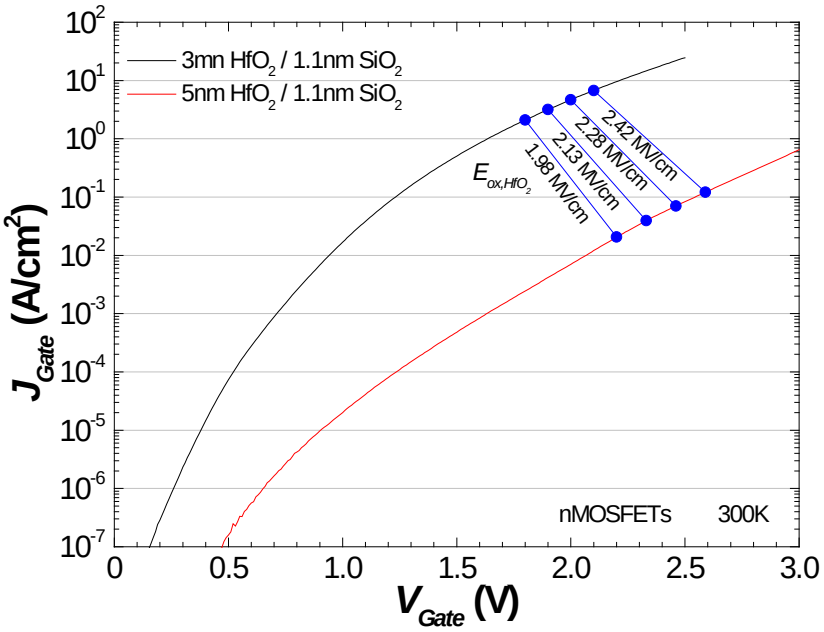


Figure C.2: Comparison of the gate leakage current of 3nm HfO_2/1.1nm SiO_2 to 5nm HfO_2/1.1nm SiO_2 at equivalent $E_{ox,HfO2}$ biases where the tunneling barrier over the tunneling distance is the same.

In order to calculate the transmission coefficient for the 3nm and 5nm HfO_2 devices, an effective mass for HfO_2 is required. However, a wide variety of effective masses of electrons in HfO_2 have been reported in the literature and range in value from ~$0.1m_0$ to ~$0.5m_0$ [7, 34, 61, 139-141], where m_0 is the mass of an electron in free space. Here a comparison of the transmission coefficient using an effective mass of $0.44m_0$ [139] and $0.11m_0$ [141] for HfO_2 layer is considered with an effect mass of $0.5m_0$ [142,

143] for the SiO_2 IL. The transmission coefficients for 3nm and 5nm HfO_2 devices at four bias conditions are shown in Figure C.3.[21] The transmission coefficient is shown for a range of electron energies (E) from the Si conduction band edge (E_c) at the SiO_2/Si interface to 1eV above E_c. Resonance peaks (i.e., local maxima) are observed in the transmission coefficient (Figure C.3) over this energy range. The energy location of the transmission coefficient resonance peaks changes as $E_{ox,HfO2}$ changes. The energy location of the transmission coefficient resonance peaks also depends on the HfO_2 thickness. While there are general similarities in the position and magnitude of the transmission coefficients between 3nm and 5nm (Figure C.3), the differences are substantial enough such that a direct comparison would be unrealistic.

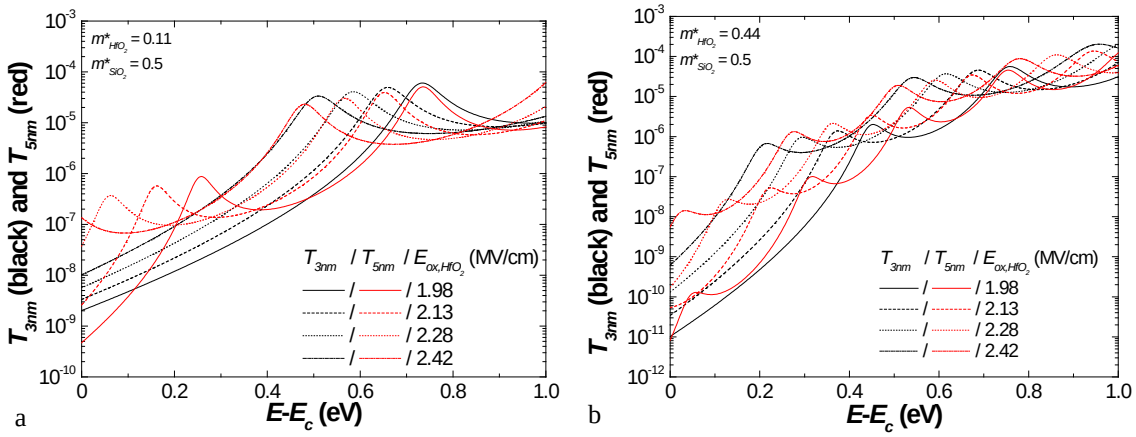


Figure C.3: Transmission coefficients as a function of energy above the Si conduction band for 3nm HfO_2/1.1nm SiO_2 (black lines) and 5nm HfO_2/1.1nm SiO_2 (red lines) barriers at four $E_{ox,HfO2}$ bias points using HfO_2 effective masses of (a) $0.11m_0$ and (b) $0.44m_0$.

A resonance peak is observed at lower energy levels in the 5nm HfO_2 device compared to the 3nm HfO_2 device. The resonance peak at lower energy in the 5nm HfO_2

[21] The effective mass of the electron in the Si substrate and TiN gate are not considered in the transmission coefficient calculation. The transmission coefficient calculation only considers the ratio of the incidence and transmitted waves through the dielectrics, which is independent of the gate and substrate effective masses.

device, which results in a larger transmission coefficient, suggests that the 5nm HfO_2 should have a larger tunneling current compared to the 3nm HfO_2 devices. Tunneling current does not solely depend on the transmission coefficient however, but also on the number of carriers available to tunnel (carrier concentration). The carrier concentration varies significantly with energy and requires solving the Poisson-Schrödinger equations to accurately determine the carrier concentration versus energy. Due to this complexity, the average of the transmission coefficients of 3nm and 5nm HfO_2 stacks are compared as a simple means to evaluate whether enough disparity between the transmission coefficients exist to account for the approximately two orders of magnitude difference in the measured leakage current.

The chosen method of comparing the average difference in the transmission coefficient is given as:

$$\begin{aligned}\%\Delta_{Avg}\overline{T}_{5nm} &= \frac{1}{E_0}\int_{E_C}^{E_C+E_0}\frac{T_{5nm}(E)-\dfrac{T_{5nm}(E)+T_{3nm}(E)}{2}}{\dfrac{T_{5nm}(E)+T_{3nm}(E)}{2}}\times 100dE \\ &= \frac{100}{E_0}\int_{E_C}^{E_C+E_0}\left(\frac{2T_{5nm}(E)}{T_{5nm}(E)+T_{3nm}(E)}+1\right)dE\end{aligned} \tag{C.1}$$

and

$$\begin{aligned}\%\Delta_{Avg}\overline{T}_{3nm} &= \frac{1}{E_0}\int_{E_C}^{E_C+E_0}\frac{T_{3nm}(E)-\dfrac{T_{5nm}(E)+T_{3nm}(E)}{2}}{\dfrac{T_{5nm}(E)+T_{3nm}(E)}{2}}\times 100dE \\ &= \frac{100}{E_0}\int_{E_C}^{E_C+E_0}\left(\frac{2T_{3nm}(E)}{T_{5nm}(E)+T_{3nm}(E)}-1\right)dE\end{aligned} \tag{C.2}$$

,

where E_0 is the energy interval above the conduction band over which to compare and average the transmission coefficients. In the formulation above for comparing the transmission coefficient, (C.1) and (C.2), the difference in the transmission coefficient of one stack and the average transmission coefficient of both HfO_2 stacks is taken and then referenced to the average transmission coefficient of both stacks and turned into a percent by multiplying by 100. By integrating over an energy interval of E_0 above the conduction band and dividing by E_0, the average difference in the transmission coefficients is computed. Two important advantages exist by using this method of comparison.

1. By comparing the difference in the transmission coefficient of one stack to the average transmission coefficient of both stacks [e.g. T_{5nm}-(T_{3nm}+T_{5nm})/2] and referencing that difference to the average transmission coefficient of both stacks [i.e., dividing by (T_{3nm}+T_{5nm})/2], a symmetrical comparison is made (i.e., $\%\Delta_{Avg}\overline{T}_{3nm}$ and $\%\Delta_{Avg}\overline{T}_{5nm}$ are equal but opposite values). By comparing the transmission coefficient to the average voids the possibility where both stacks indicate they have a higher transmission coefficient than the other.
2. By referencing the difference in the transmission coefficients [e.g., T_{5nm}-(T_{3nm}+T_{5nm})/2] to the average transmission coefficient [i.e., (T_{3nm}+T_{5nm})/2] before performing the integral, the difference in the transmission coefficients are weighted evenly. At large energy levels, the transmission coefficients are larger, creating larger differences between the transmission coefficient of one stack and the average transmission coefficient of both stacks. Without normalizing this value to the average transmission

coefficient of both stacks, the comparison would be skewed toward higher energy levels.

Using an energy interval, E_0, of 1eV a comparison of $\%\Delta_{Avg}\overline{T}_{3nm}$ and $\%\Delta_{Avg}\overline{T}_{5nm}$ for the four $E_{ox,HfO2}$ values shown in Figure C.3 is given in Table C.1. Table C.1 indicates that the transmission coefficient of the 5nm HfO_2 stack is slightly larger on average than the 3nm HfO_2 stack, regardless of the HfO_2 effective mass considered ($0.11m_0$ or $0.44m_0$). Allowing for variation between the modeled energy band structure and the actual energy band structure is not enough to close the gap between the nearly two decade difference between the gate leakage current between the 3nm and 5nm HfO_2 gate stacks. The discrepancy between the theoretical transmission coefficient and the experimental gate leakage current indicates that quantum mechanical tunneling (non-defect mediated) cannot explain the gate leakage current or its temperature dependence for the HfO_2 gate stacks in this study.

Table C.1: Comparison of Average Transmission Coeffecients

$E_{ox,HfO2}$ (MV/cm)	**3nm HfO_2**			**5nm HfO_2**		
	V_{Gate} (V)	$\%\Delta_{Avg}\overline{T}_{3nm}$ $m^*=0.11m_0$	$\%\Delta_{total}\overline{T}_{3nm}$ $m^*=0.44m_0$	V_{Gate} (V)	$\%\Delta_{Avg}\overline{T}_{5nm}$ $m^*=0.11m_0$	$\%\Delta_{Avg}\overline{T}_{5nm}$ $m^*=0.44m_0$
1.98	1.8	-6.52	-14.49	2.2	6.52	14.49
2.13	1.9	-10.52	-12.97	2.33	10.52	12.97
2.28	2	-9.64	-9.62	2.46	9.64	9.62
2.42	2.1	-5.28	-0.84	2.58	5.28	0.84
	Average (%)	**-7.99**	**-9.48**	**Average (%)**	**7.99**	**9.48**