Lund University
Department of Physics

Bachelor Thesis 15 hp

Leakage current and breakdown of HfO_2/InGaAs MOS capacitors

Author:
Edvin Winqvist

Supervisor:
Guntrade Roll

Spring semester 2015
Submission date: May 20, 2015
Examination date: June 4, 2015

Lund University

Abstract

With the constant downscaling of transistors, silicon as a production material is falling out of favour because of increasing power consumption when the size of devices becomes smaller. Compound materials from group III-V in the table of elements are promising candidates to replace silicon. The aim of this work was to study current-voltage characteristics of a MOS capacitor made of the III-V compound InGaAs. Three samples were produced using atomic layer deposition (ALD) to apply an oxide layer of the high-κ material HfO_2 on the InGaAs surface. The thicknesses of these oxide layers were 4 nm, 6 nm and 6 nm where one of the 6 nm samples underwent post-metallisation annealing (PMA).

The electric field required to cause a hard breakdown through the oxide was found to be ~ 0.81 GV/m for the 6 nm annealed sample, ~ 0.90 GV/m for the 6 nm as deposited sample and ~ 1.13 GV/m for the 4 nm sample. In all three samples, the breakdown field was widely distributed which indicates an InGaAs-HfO_2 interface with a large variation in density of interface traps across the layer. The breakdown field was found to decrease with oxide thickness and PMA treatment, which might be attributed to a percolation path through the oxide being created more easily due to higher polycrystallinity in the thicker oxides.

The dominant leakage mechanism at higher biases was determined to be Fowler-Nordheim tunneling and was assumed to be direct tunneling or trap assisted tunneling at low bias. Temperature dependency measurements suggested that trap assisted tunneling gained significance at low bias when the temperature increased.

Acknowledgement

I would like to thank my supervisor Guntrade Roll for invaluable assistance and input during this work. I would also like to thank the Department of Electrical and Information Technology at the Faculty of Engineering at Lund University for allowing me to use their lab and measurement tools, along with the Lund Nano Lab for allowing me to use their lab to fabricate my samples.

Contents

Abstract 1

Acknowledgement 1

Contents 2

List of abbreviations 3

List of symbols 4

1 Introduction 5

2 Theory 6
- 2.1 Material properties . . . 6
- 2.2 Energy band structure . . . 7
- 2.3 Defects . . . 8
- 2.4 Tunneling . . . 9
  - 2.4.1 Trap assisted tunneling . . . 10
  - 2.4.2 Direct tunneling . . . 10
  - 2.4.3 Fowler-Nordheim tunneling . . . 11
- 2.5 Dielectric breakdown . . . 11

3 Experiments 13
- 3.1 Processing . . . 13
  - 3.1.1 Surface preparation . . . 13
  - 3.1.2 Atomic Layer Deposition . . . 14
  - 3.1.3 Photolithography . . . 14
  - 3.1.4 Metallisation . . . 15
- 3.2 Measurements . . . 16

4 Results and discussion 17
- 4.1 Breakdown voltage . . . 17
- 4.2 Temperature dependence . . . 21
- 4.3 Time dependent dielectric breakdown (TDDB) . . . 23

5 Summary and outlook 24

Bibliography 25

List of abbreviations

ALD	Atomic Layer Deposition
BD	Breakdown
CB	Conduction Band
CV	Capacitance Voltage
CVS	Constant Voltage Stress
DT	Direct Tunneling
EOT	Equivalent Oxide Thickness
FN	Fowler-Nordheim
IV	Current Voltage
MOS	Metal Oxide Semiconductor
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
MTAT	Multi-Trap Assisted Tunneling
PMA	Post Metal Annealing
SBD	Soft Breakdown
SILC	Stress Induced Leakage Current
STAT	Single-Trap Assisted Tunneling
TAT	Trap Assisted Tunneling
TDDB	Time Dependent Dielectric Breakdown
UV	Ultraviolet
VB	Valence Band
XPS	X-ray Photoelectron Spectroscopy

List of symbols

A	Area
c	Constant
C_{ox}	Capacitance of oxide
D_{it}	Density of interface traps
E_c	Conduction band energy
E_f	Fermi energy
E_i	Intrinsic energy
E_v	Valence band energy
ϵ_0	Permittivity of free space
F_{ox}	Electric field across oxide
$\hbar$	Reduced Planck constant
I_{DT}	Direct tunneling current
I_{FN}	Fowler-Nordheim tunneling current
I_O	Oxygen interstitial
κ	Dielectric constant
κ_{ox}	Dielectric constant of oxide
κ_{SiO_2}	Dielectric constant of silicon dioxide
m_e	Electron mass
m^*	Effective mass
m^*_{ox}	Effective mass in oxide
ϕ_b	Barrier height
q	Elementary charge
t_{bd}	Time to breakdown
t_{ox}	Oxide thickness
t_{EOT}	Equivalent oxide thickness
V_{FB}	Flatband voltage
V_G	Gate voltage
V_O	Oxygen vacancy
V_{ox}	Voltage across oxide

1 Introduction

The first transistor was invented in 1947, and it was perhaps the most important electronic invention of the century. It allowed the production of integrated circuits and processors that are the building stones of modern electronics. The development of transistors took off rapidly and in 1965, Gordon E. Moore formulated the famous Moore's law which states that the density of transistors in an integrated circuit will double every two years. The claim proves surprisingly accurate even today[1].

A metal oxide semiconductor (MOS) capacitor is the basis for the metal oxide semiconductor field effect transistor (MOSFET) which is a common type of transistor. As the name suggests, the MOS and MOSFET are constructed of three layers of a semiconductor, an oxide, and a metal. The MOSFET behaves as a switch with on and off states, and as a result, the switching speed and the switching energy are important properties. With the reduction in transistor size, the switching speed has significantly increased while the switching energy has decreased[2].

Among the many advantages of the downscaling of transistors, there are also disadvantages. The dominant semiconductor material throughout history has been silicon (Si) with an oxide layer of silicon dioxide (SiO_2), and when the SiO_2 layer becomes thinner, the effect of quantum mechanical tunneling increases and gives rise to a gate leakage current. The gate leakage current can result in an unacceptably large current loss that significantly increases power consumption. In addition, high leakage currents reduce the reliability of the oxide layer which raises the risk of an undesired dielectric breakdown[3].

If progress in transistor technology is to continue, the power consumption issue has to be overcome. One way is to introduce other materials than Si and SiO_2 to use for the transistors. A semiconductor with a lower effective electron mass than Si should have a higher carrier mobility, and would obtain the same switching speed as Si but at a lower operating voltage, reducing energy losses. An oxide with a high dielectric constant could be built thicker than SiO_2 but retain the electrical properties to reduce the leakage current, and subsequently, the power consumption[2].

Promising candidates to replace Si are group III-V compound materials. The electron mobility of compounds such as indium gallium arsenide (InGaAs) proves higher than for most other known substances[2]. A big advantage that Si has over semiconductors of these materials however, is that its native oxide SiO_2 is a decent insulator and can be effortlessly grown on top of the Si layer. The native oxides of group III-V semiconductors are poor, and more advanced methods are required for the selection and application of a different oxide, typically a dielectric material with a high dielectric constant[3].

The aim of this work was to investigate various electrical properties of a MOS device constructed of an InGaAs semiconductor and a hafnium dioxide (HfO_2) insulator. Three samples of varying thicknesses and preparation methods were to be processed and their breakdown voltage, time dependent dielectric breakdown (TDDB) and temperature dependence of the leakage current measured. The results should then be used to increase understanding of what leakage and breakdown mechanisms are active in the samples.

2 Theory

A MOS consists of a semiconductor layer, an oxide layer and a metal layer. The semiconductor layer can be n-doped (p-MOS) or p-doped (n-MOS). Above the semiconductor is the oxide layer which functions as an insulator. At the top is the metal layer, also known as the gate metal. In this chapter, relevant theory about the MOS capacitor will be discussed, including choice of materials, energy band structure and leakage mechanisms.

2.1 Material properties

As mentioned, InGaAs is an excellent semiconductor because of its low effective mass and high electron mobility[2]. Table 1 lists a number of important parameters as a comparison between Si and InGaAs.

Table 1: Electrical properties at 300 K for Si and $In_{0.55}Ga_{0.45}As$[4].

Material	Band gap (eV)	Electron effective mass (m^*/m_e)	Electron mobility ($cm^2V^{-1}S^{-1}$)
Si	1.120	0.98	1400
$In_{0.55}Ga_{0.45}As$	0.73	0.04	13650

With the area of silicon transistors being downscaled, the oxide layer has to become thinner in order to retain the same capacitance. Eq. 1 shows the capacitance through the oxide layer, C_{ox}, where κ_{ox} is the dielectric constant of the oxide, ϵ_0 the permittivity of free space, A the area and t_{ox} the thickness of the oxide layer[5].

$$C_{ox} = \frac{\kappa_{ox}\epsilon_0 A}{t_{ox}} \tag{1}$$

As Eq. 1 suggests, an oxide with a κ value twice that of SiO_2 would retain the same capacitance as SiO_2 at double the thickness, significantly reducing leakage. A common technique used to quickly compare a dielectric material to SiO_2 is to calculate the equivalent oxide thickness (EOT), which shows the SiO_2 thickness required to achieve the same capacitance as the dielectric, see Eq. 2

$$t_{EOT} = \frac{\kappa_{SiO_2}}{\kappa_{ox}} t_{ox} \tag{2}$$

Here, t_{EOT} is the equivalent thickness of SiO_2 and κ_{SiO_2} is the dielectric constant of SiO_2[5].

Table 2: Dielectric constant, effective mass and band gap of SiO_2 and HfO_2[6][7].

Dielectric	κ	Electron effective mass (m^*_{ox}/m_e)	Band gap (eV)
SiO_2	3.9	0.40	9
HfO_2	25	0.22	5.8

Table 2 displays a set of important parameters of SiO_2 and HfO_2. While the dielectric constant κ is important, the leakage current depends on the tunneling effective mass of the oxide, and the height of the energy barrier between the semiconductor and oxide[7]. Atomic layer depositioned (ALD) as deposited HfO_2 that is used in this work is expected to be mostly amorphous, but become more polycrystalline when the thickness is built up[8]. However, the temperature used to deposit HfO_2 in this work, 120 °C, is lower than what is usually practised which can result in ligands from the precursor material remaining in the oxide layer and keeping it partially amorphous even with increasing thickness[9].

2.2 Energy band structure

Applying a bias V_G at the gate metal will result in a potential drop across the oxide. Depending on the applied voltage, electrons or holes in the semiconductor will be attracted to the semiconductor-oxide interface, inducing a bending of the energy bands in the region. At a certain flatband voltage V_{FB}, there is no net charge present at the interface and the energy band is flat, see Fig. 1a). The flatband voltage is typically close to zero but depends on parameters such as the number of trapped charges, which will be discussed in the next section[10].

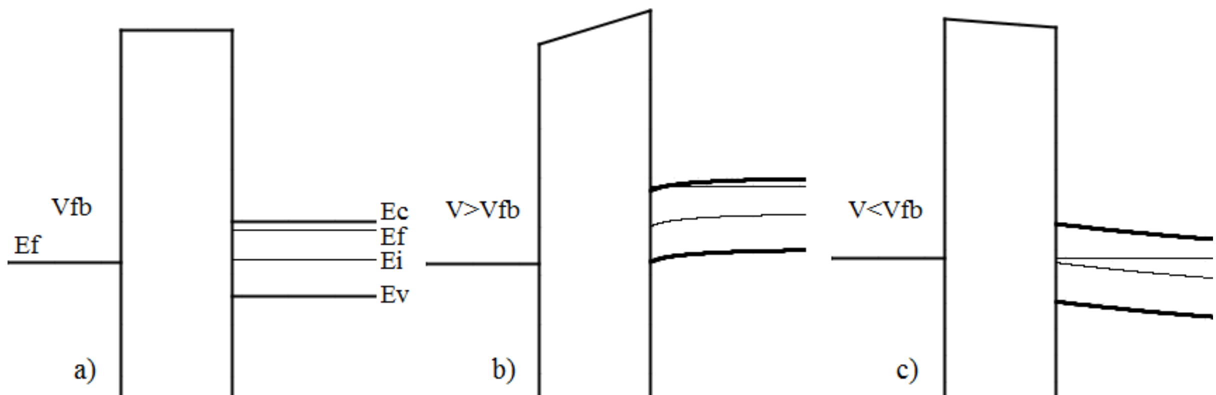

Figure 1: Energy band diagrams of a MOS capacitor with n-doped InGaAs. a) Flatband condition. b) Accumulation. c) Depletion. Images generated using a band diagram simulator from Boise State University[11].

When V_G is above the flatband voltage, there will be an abundance of electrons at the interface of the n-doped InGaAs. The bias leads to band bending and the Fermi energy E_f shifting closer to the conduction band, CB. In the quantum mechanical model, this results in a buildup of electrons near the interface, see Fig. 1b). This mode where the majority carrier in the semiconductor builds up near the oxide is called accumulation[10].

When V_G is lower than V_{FB}, the electrons in the semiconductor will repel from the semiconductor-oxide interface, which is called depletion. Fig. 1c) shows how the induced electric field starts to bend the potential near the oxide so that the Fermi level in this region moves away from the conduction band.

2.3 Defects

Defects in the oxide layer are sites with an excess or deficit of oxygen atoms, impurities of atoms that are not supposed to be in the material, or other types of faulty bonds. The defects may be electrically active and create states in the forbidden gap of the oxide that can attract electrons or holes from the semiconductor and trap them[12]. The number of available traps depend on applied voltage and the temperature[13].

An advantage of SiO_2 over high-κ materials is that the number of intrinsic defects it contains is notably lower. One reason for this is the way the silicon atoms bond in the crystal. SiO_2 is bound together with covalent bonds of a low coordination. The low coordination allows the SiO_2 crystal to relax and repair any dangling bonds, which is the dominant form of defect. High-κ oxides use ionic bonding with a less adjustable structure[6]. Their defects are primarily oxygen vacancies or oxygen deficiencies[14].

The existence of trapped charges in the oxide layer changes the electric field across it. As a result, with zero applied gate voltage, some bending of the potential will already be ongoing. It follows that V_{FB} is different from zero as it has to compensate for the trapped charges in order to obtain a zero field condition at the semiconductor-oxide interface[10].

Defects exist in the bulk of the material but are also created at the semiconductor-oxide interface where the difference in lattice parameters and bonding structure of the semiconductor and oxide may be distinctive[12]. The density of interface traps D_{it} is important as the trapping related electric field change is most likely to commence at the interface. There are various surface treatment techniques and deposition methods available that can be used to improve the semiconductor-oxide interface, but the result is unlikely to be ideal. Another way to reduce the amount of defects in a high-κ oxide is by annealing the MOS capacitor after fabrication, which is typically done by subjecting the sample to a heated gas of N_2/H_2. The annealing process may remove impurities which compresses the oxide layer[6]. Annealing can also close off dangling bonds in the oxide through bonding with hydrogen from the H_2 gas, which further reduces the number of active defects[12].

The defects in the oxide that are most likely to trap charges are those that reside in the energy range of the Fermi level E_f, which generally varies roughly between the valence band and conduction band edges of the semiconductor. The energy states of defects depend on the crystallinity of the oxide and in the case of interface traps, on the semiconductor-oxide interface texture. The energy levels can be measured or calculated theoretically, which can be very difficult for amorphous materials where the variation in atomic configuration is large[15].

Fig. 2 shows the energy levels of defects in bulk HfO_2 as obtained by [14], with the conduction band (CB) and valence band (VB) of InGaAs superimposed using [11]. Fig. 2 is meant as an illustration only and not as an absolute description of the energy levels of defects for the HfO_2 used in this project. The most likely defects in HfO_2 are oxygen vacancies, V_O, and oxygen interstitials, I_O, with the exponent in the figure indicating the charge of the relevant oxygen ion.

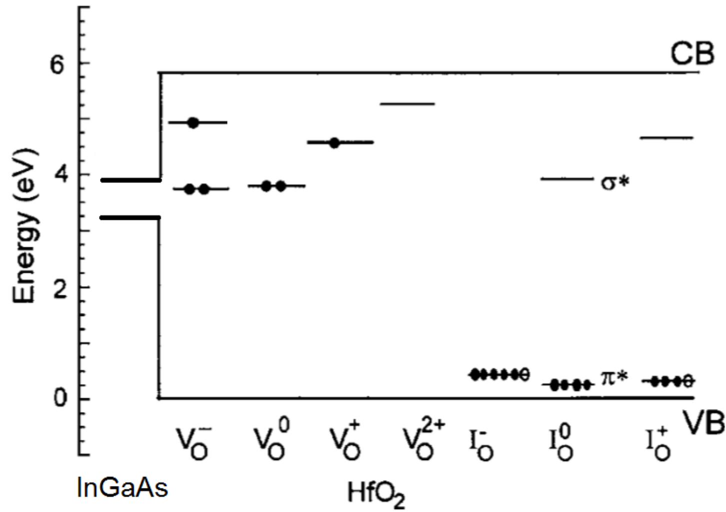

Figure 2: An illustration of defect energy levels in bulk HfO_2 with CB and VB of InGaAs drawn[11][14].

The traps in the oxide can facilitate the tunneling of electrons, and their build-up is the root of dielectric breakdown as explained in the following sections.

2.4 Tunneling

The leakage current through the oxide is ideally a result of the quantum mechanical effect of tunneling. Electrons and holes have a finite probability of penetrating through a potential barrier. In a MOS capacitor, the tunneling effect can take different approaches where some important variations are trap assisted tunneling, direct tunneling and Fowler-Nordheim tunneling, see Fig. 3, that will be explained in more detail in this section. Leakage current can also be a result of other processes such as Poole-Frenkel emission, thermionic emission, ohmic conduction or dielectric relaxation[16].

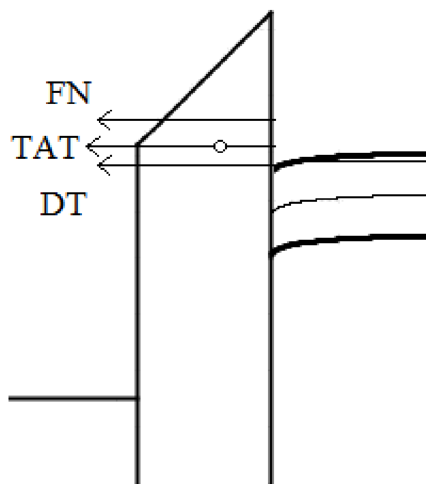

Figure 3: An illustration of direct tunneling, trap assisted tunneling and Fowler-Nordheim tunneling[17].

2.4.1 Trap assisted tunneling

Trap assisted tunneling (TAT) is the effect of carriers tunneling to electrically active defects in the oxide. As the electric field or temperature increases the probability of tunneling to a trap increases, after which the electron can tunnel further towards the gate metal. If the applied voltage is reduced, trapped charges may detrap back to the oxide. The TAT clearly depends on the density of traps in the oxide and is therefore generally more relevant in high-κ materials where the amount of intrinsic defects tend to be larger[18].

Trap assisted tunneling of electrons is not an elastic process. Electrons that tunnel to a defect release phonons of energy $\hbar\omega$ to reduce their energy to that of the trap. The TAT process through the oxide has two models. The process may be by single-trap assisted tunneling (STAT), which is a two-step process where the carrier passes through one trap on its way to the oxide, or by multi-trap assisted tunneling (MTAT), where if the oxide is highly degraded with defects, the carrier can hop between multiple traps to permeate the oxide[12]. An illustration of an STAT process can be seen in Fig. 3.

2.4.2 Direct tunneling

Direct tunneling is only detectable when the oxide thickness is very thin, in the order of a few nm. Direct tunneling of electrons or holes refers to the penetration of the carrier between bands in the semiconductor and gate metal directly through the oxide. The three types of tunneling electron tunneling from the conduction band (ECB), electron tunneling from the valence band (EVB) and hole tunneling from the valence band (HVB) are all variations of direct tunneling through the potential barrier caused by the insulating oxide. HVB is effectively the same thing as electron tunneling from valence band to valence band. An illustration of ECB is shown in Fig. 3[7][17].

Deriving an expression for the direct tunneling is a difficult process that requires numerous assumptions. Eq. 3 is a model for the direct tunneling gate current where the finite availability of carriers for tunneling is neglected. This result is an acceptable approximation except for very thin oxides (< 2 nm) at low voltages[17].

$$I_{DT} = c \cdot F_{ox}^2 \exp\left\{ -\frac{4}{3}\frac{\sqrt{2m_{ox}^*}\phi_b^{3/2}}{\hbar q}\frac{1}{F_{ox}}\left[1-(1-\frac{qV_{ox}}{\phi_b})^{3/2}\right]\right\} \tag{3}$$

In this expression, c is a constant, F_{ox} is the electric field across the oxide, m_{ox}^* is the effective electron mass in the oxide, ϕ_b is the barrier height between conduction bands at the semiconductor-oxide interface, $\hbar$ is the reduced Planck constant, q is the electron charge and V_{ox} is the voltage across the oxide. The electric field induced across the oxide is given by

$$F_{ox} = \frac{V_G}{t_{ox}} \tag{4}$$

2.4.3 Fowler-Nordheim tunneling

Fowler-Nordheim (FN) tunneling through the oxide can occur when a large enough bias is applied at the gate to shift the side of the potential barrier near the metal downwards. The electrons in the conduction band can tunnel through the then triangular shaped potential barrier, see Fig. 3. The triangular shape effectively makes the barrier thinner and the probability of tunneling higher[19].

By making the assumptions that the available electrons for tunneling can be described as a Fermi gas and their availability is not governed by temperature, and that the potential barrier shape is triangular, Eq. 5 can be deduced for the tunneling current caused by the FN effect. If FN tunneling is the dominant leakage mechanism, the logarithm of I_{FN}/F_{ox}^2 against $1/F_{ox}$ should yield a straight line[17].

$$I_{FN} = c \cdot F_{ox}^2 \exp\left[-\frac{4}{3}\frac{\sqrt{2m_{ox}^*}\phi_b^{3/2}}{\hbar q}\frac{1}{F_{ox}}\right] \quad (5)$$

2.5 Dielectric breakdown

Under high electrical stress, a large number of traps in the oxide will be filled and there may even be additional conductive spots generated if the applied bias is sufficiently high. These spots can be seen experimentally using various techniques such as AFM related measurements[20]. The additional traps filled or created causes an increased leakage current known as the stress induced leakage current (SILC)[12].

The position of a trap is area independent and it can form anywhere in the oxide. However, experiments show that the grain boundaries between crystallites in the oxide have a greater density of filled traps than other regions. With sufficient degradation of the oxide, a series of traps may be formed along the grain boundaries and create a conduction path that electrons can hop through from trap to trap, which greatly increases leakage current. The creation of a conduction path is referred to as a dielectric breakdown and the impact of it has two stages, a soft breakdown or a hard breakdown[13].

A soft breakdown is characterised by a large and abrupt increase in leakage current. The sudden escalation can be explained by a single percolation path being formed which allows more electrons to flow through the oxide, see Fig. 4. A hard breakdown is also identified by an abrupt increase in leakage current, but multiple order of magnitudes higher than in a soft breakdown, see Fig 4b). After a hard breakdown, the oxide is highly damaged and its structure destroyed, and it has effectively turned into an ohmic conductor[21].

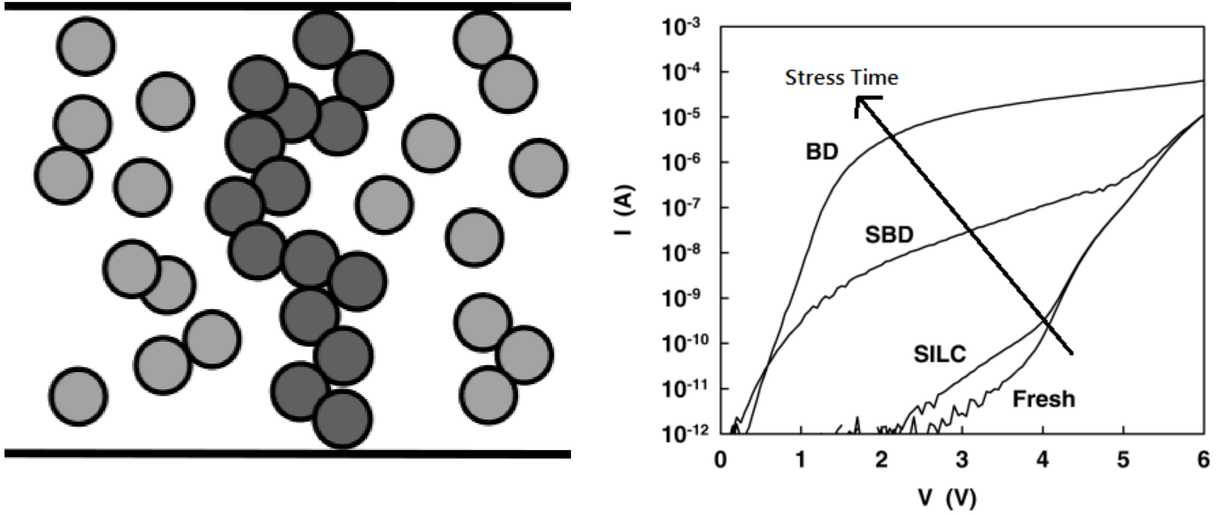

Figure 4: Illustration of dielectric breakdowns. a) A percolation path created through the oxide layer. b) Typical current characteristics for fresh samples, and samples under various degrees of degradation. Image b) is modified from [22].

There is a publication of an x-ray photoelectron analysis of the semiconductor-oxide interface after breakdown. The study was performed on MOS capacitors made of InGaAs with Al_2O_3 as dielectric, but it can be assumed that an analogous result would be observed for HfO_2. After dielectric breakdown under positive bias, oxides of In, Ga and As were formed at the semiconductor-oxide interface. Under negative bias, the number of oxidized substrates was reduced. The result was explained by atoms diffusing into the dielectric by the induced electro-migration effect during the breakdown of the oxide[23].

3 Experiments

Three MOS samples of InGaAs and HfO_2 were fabricated in different ways and their measurement results compared. Current-voltage (IV) measurements were performed where the leakage current and breakdown points were measured at a varying voltage. Measurements where the samples were exposed to a constant voltage stress (CVS), measurements of the SILC, and IV measurements of unstressed devices under varying temperature were done.

3.1 Processing

The metal oxide semiconductor arrangement was processed on the surface of an indium phosphide (InP) wafer. The InP wafer was doped with iron (Fe) to make it non-conductive. Two samples with InGaAs of the composition $In_{0.55}Ga_{0.45}As$ were created. In both samples, HfO_2 of different thicknesses was used as the insulator layer. The metal layer was produced of a stack of nickel (Ni), palladium (Pd) and gold (Au). The Ni was administered first due to its good interface with HfO_2, and Au was added as the top layer because of its high conductivity. Pd was used as an intermediate to prevent the mobile gold atoms from piercing into the oxide, and for its good adhesion with Ni and Au.

3.1.1 Surface preparation

To optimize the interface between the InGaAs and HfO_2 layers, the InGaAs surface had to be thoroughly cleaned and deprived of native oxides. To eliminate undesired particles from the surface, the samples were immersed in acetone and positioned in an ultrasonic cleaner for 3 min. The ultrasonic cleaning process was then repeated twice where acetone was replaced with isopropanol and distilled water respectively. To clear out native oxides, the samples were placed in a 1:1 hydrochloric (HCl) solution for 30 s and rinsed in water.

One sample underwent an additional process that the other did not. A dummy gate was created using a technique similar to how the drain and source would be created in a MOSFET. The sample was baked on a hot plate at 200 °C for 5 min. Hydrogen silsesquioxane (HSQ) was then spin coated onto the surface at 3000 rpm for 1 min, after which the sample was positioned on the 200 °C hot plate again for 1 min. The sample was then cured in an oven at 300 °C for 60 min.

The cleaning process of the surface continued by the samples being placed in hydrofluoric acid (HF) for 3 min and then rinsed with water, which removed the HSQ applied to one of them. Previous experiments show that the dummy gate process has no significant effect on the results of measurements on the sample. After the HF treatment, the samples were processed in an ozone cleaner for 15 min, in which UV radiation is absorbed by oxygen (O_2) on the sample that turns into atomic oxygen (O) and ozone (O_3). The various oxygen molecules react with free radicals that were released in the UV process, to form molecules such as carbon dioxide (CO_2) or water vapour (H_2O)[24]. Finally, the samples were immersed into a solution of $(NH_4)_2S \cdot 20\%\ H_2O$ for 20 min.

3.1.2 Atomic Layer Deposition

To administer the HfO_2 layer on top of the cleaned InGaAs surfaces, atomic layer deposition (ALD) was used. ALD is a method commonly utilized to deposit thin films with precision in thickness down to monoatomic layers. Such a high level of control is achieved by releasing two precursor substances into the chamber where the sample is located, one for Hf and one for O, alternating between them. The precursors will attach to every available surface site and if both of them give rise to a self-limiting process where only a single layer is deposited during each sequence, very high precision is achieved[25]. As a precursor for Hf, tetrakis(dimethylamido)hafnium (TDMA-Hf) was used and H_2O vapour as a precursor for oxygen.

The ALD chamber was first ventilated with N_2 for 20 min. It was then heated to 120 °C, and the Hf precursor was heated to 75 °C and the temperatures were stabilized for 2 min before the process started. The ALD began with five 0.15 s pulses of TDMA-Hf, with 50 s wait time in between, being released into the chamber to achieve a self-cleaning effect. ALD self-cleaning is a phenomenon where the Hf precursor and native oxides on the InGaAs surface perform substitution reactions to cleanse the surface of native oxides and create HfO_2[26]. After the initial pulses of TDMA-Hf, 0.15 s pulses of TDMA-Hf and 0.015 s pulses of water vapour were alternated, ventilating the chamber with N_2 during 50 s downtime periods between each pulse. The cycle of varying TDMA-Hf and H_2O pulses was repeated 60 times for one sample, and 40 times for the other, where one cycle is expected to roughly correspond to a layer of 1 Å, for thicknesses of 6 nm and 4 nm.

3.1.3 Photolithography

To create a pattern of metal dice on the surface to measure on, rather than a complete layer, a photolithographic process was implemented. The photoresist LOR-10B was spun onto the HfO_2 surfaces at 6000 rpm for 90 s and the samples were baked on a hot plate at 190 °C for 5 min to improve the stiffness of the resist. A second resist, S1813, was spun on top of the first at 6000 rpm for 90 s and baked at 115 °C for 90 s, see Fig. 5a).

The samples were placed under a mask in a UV emitter and one cycle of 5 seconds at 0.05 MPa was irradiated on the surfaces in order to alter the chemical structure of the resists where the mask allowed the UV light to shine through. After the UV process, the samples were placed in developer MF319 for 1 min, and then in water for 1 min to dissolve the irradiated parts of the resists. Figure 5b) shows the result of the process.

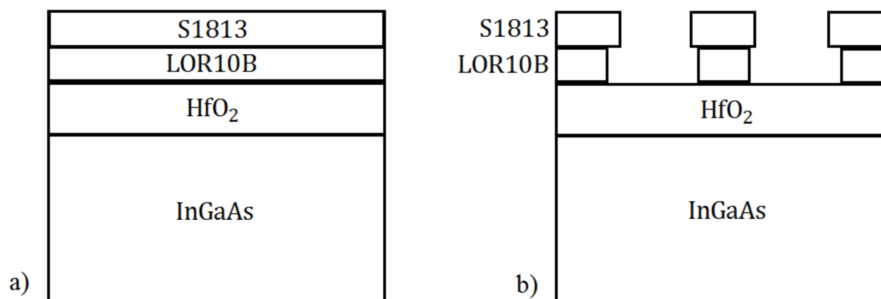

Figure 5: Photolithographic process. a) Sample with photoresists spun onto the surface. b) Appearance after irradiation and dissolution.

3.1.4 Metallisation

The metallisation was done by placing the sample in an evaporation chamber. The chamber was emptied of air after pieces of Ni, Pd and Au were positioned in the chamber. A detector in the ceiling measured the thicknesses of the metal layers created on the samples as the evaporated metals cooled down. The detector had a resolution down to the Å level.

The metallisation was done on the 6 nm sample first beginning with the evaporation of Ni until a thickness of 4 nm was achieved. Next a 10 nm layer of Pd and finally a 200 nm layer of gold. The measured layer thicknesses on the 4 nm sample were 6 nm Ni, 11 nm Pd and 170 nm Au.

The metallisation process produced metal pads in the holes created by the photolithography, and on top of the resists, see Fig 6a). To remove the resists, with the metal on top, the samples were submerged in acetone for 30 min and then placed in an ultrasonic bath for 20 s. To eliminate any remaining residuals, the samples were immersed in Remover 1165 and heated to 70 °C for 10 min. Figure 6b) is an illustration of the samples after the removal of the resists, and Figure 7 is an image taken of the surface on the 4 nm sample through a microscope. The metal dice were measured to have an area of approximately 2400 μm^2. After the dissolving process, the 6 nm sample was broken into two pieces where one was annealed at 350 °C under influence of N_2/H_2 for 5 min. Table 3 displays a summary of the processes that varied between the samples during the fabrication.

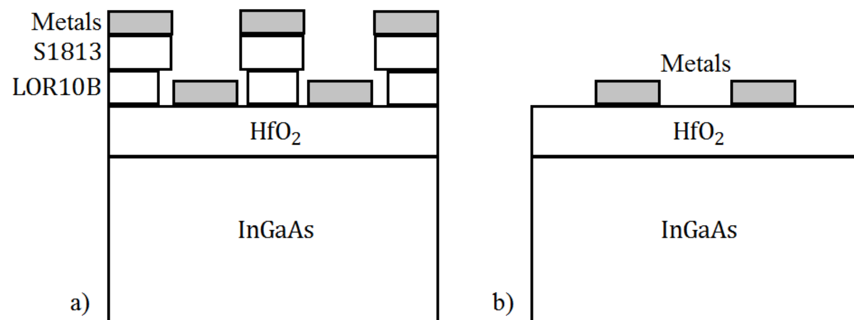

Figure 6: Metallisation process. a) Sample with metals evaporated onto the surface of the resists and in the holes created by the photolithography. b) The same sample after removal of the resists.

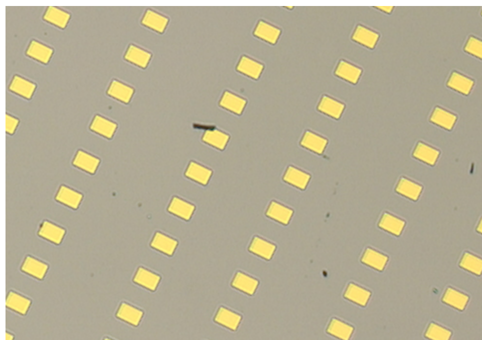

Figure 7: Microscope image of the surface of sample *b*. The area of the metal dice is around 2400 μm^2.

Table 3: Sample thicknesses and processes done during the fabrication.

Sample	Oxide thickness (nm)	Dummy gate pre ALD	PMA
1	6	Yes	Yes
2	6	Yes	No
3	4	No	No

3.2 Measurements

A needle probe was used to apply a voltage and measure the current directly on the wafer. The two needles were connected to two dice on the surface of each sample and a hard breakdown was forced by applying a high voltage. One of the broken dice was then used as back contact as the second needle was attached to an unbroken die for the measurements.

An IV measurement was performed on 20 dice on each sample where the leakage current through the oxide at an increasing bias was analysed. The voltage was linearly increased from 1 V and up to 6 V in steps of 0.05 V, to ensure that a hard breakdown occurred in every measurement. Another measurement was performed where the SILC was investigated by doing 10 cycles of 30 s measurements between 0 and 3 V with a high stress voltage being applied between each sweep.

A third measurement was carried through where the samples were exposed to a CVS at a voltage slightly below the average breakdown voltage of the sample, in order to measure the time to breakdown, t_{bd}. Each sample was analysed up to 20 times at 3 different biases.

A final measurement was performed where the temperature dependence of the current was analysed. The samples were placed on a thermoelectric plate and measurements were performed at -25, 0, 25, 50 and 75 °C in sweeps that went from -3 V to 3 V and back in steps of 0.1 V. Two dice on each of the non-annealed 6 nm and the 4 nm samples were measured.

4 Results and discussion

The breakdown voltage, breakdown time and SILC measurements were performed on all samples while the temperature dependency was only done on the 6 nm as deposited and the 4 nm samples. For the sake of consistency, when comparing the outcomes of more than one sample, the results from the 6 nm as deposited sample are displayed in black, the 6 nm annealed sample in red, and the 4 nm sample in blue.

4.1 Breakdown voltage

Fig. 8a) shows the distribution of the breakdown voltage obtained for the three samples. Fig. 8b) shows a cumulative distribution plot of the breakdowns, plotted against electric field. The 4 nm sample clearly displays the highest breakdown field, while the annealed 6 nm sample appears to be the worst of the three.

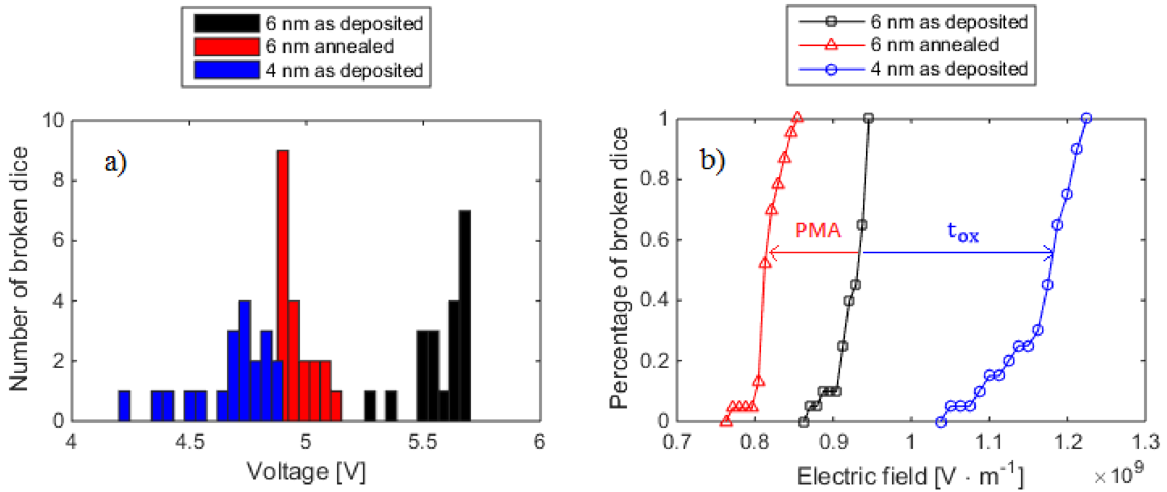

Figure 8: Distribution of the breakdown voltage in the three samples. a) A histogram showing the number of dice that broke within each range of 0.05 V. b) A cumulative distribution plot of the breakdown points.

For all three samples, the distribution covers a wide bias range which indicates a large variation in defect density and placement in the oxide. The diversity could stem from a variety of sources. The clean room used for the fabrication belongs to the university where the work is done by hand, and may not be at industrial standards where processes are done automatically without human error and contamination. The samples were exposed to air for a period of time between the steps in the cleaning process and before the ALD, which could make an important impact on the quality of the InGaAs surface and semiconductor-oxide interface. Native oxides could have formed during the exposure that the ALD self cleaning process did not clear out.

The ALD process was done at the relatively low temperature of 120 °C where the amount of defects created should be higher than at temperatures around 300 °C. The oxide properties should also be slightly worse, however the oxide should become more crystalline at higher temperature[9].

The thin 4 nm sample could withstand a higher electrical field before breaking. The reason for this could be a lower grade of crystallinity. As the HfO_2 layer grows thicker in the ALD process, the layer should become increasingly polycrystalline from the more amorphous state of the initial layers[8]. The 6 nm samples should therefore have more grain boundaries than the 4 nm sample. As previously mentioned, the defect paths tend to prefer to follow grain boundaries and the leakage current should then be higher in the 6 nm samples than in the 4 nm sample at the same electric field[13].

The 6 nm annealed sample breaks at a lower electric field than the as deposited 6 nm sample. The reason why the annealed sample shows worse properties than the non-annealed could also be connected to the crystallinity of the samples. The annealing process is supposed to reduce the number of defects and may compress the oxide which also tends to align the crystals in the layer. If the grain boundaries are more aligned in the annealed sample, creating a conduction path through the oxide in the annealed sample should be more achievable than in the non-annealed sample.

A plot of the leakage current leading up to the breakdown can be seen in Fig. 9. The measurement was performed first on the 6 nm annealed sample with a poor resolution at the lower currents, and these parts are thus chosen not to be displayed in the graph. The lower resolution also explains the high amount of noise that can be seen in the plot just before the breakdown of the annealed sample. To make sure that the leakage current of this sample acted in the same way as the other samples, two additional measurements were performed using the higher resolution. The results of these two measurements can be observed in the figure to behave in a comparable style at low applied bias.

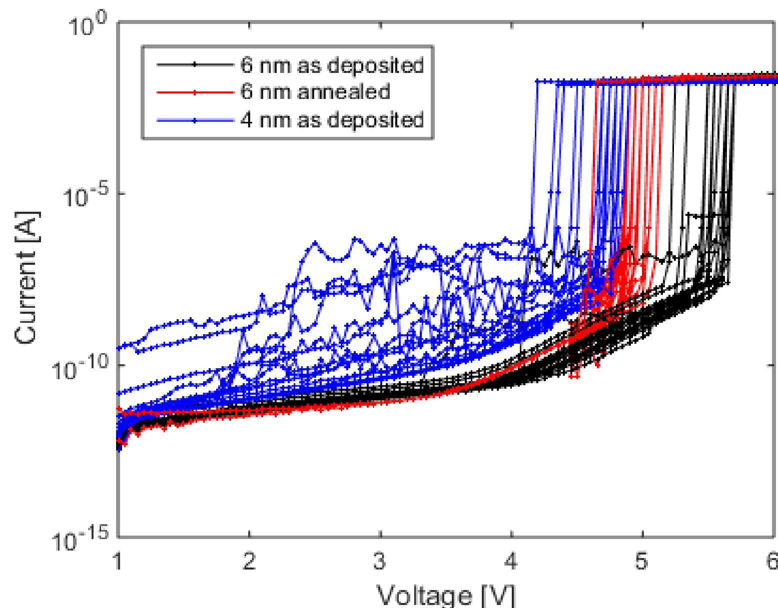

Figure 9: The leakage current measured at an increasing voltage for the three samples.

In Fig. 9, the measurements of the 4 nm sample appears more noisy than those of the 6 nm samples. The reason for the higher noise level is probably a result of the oxide layer being thinner, which suggests that the impact each trapped charge has on the leakage

current becomes more significant. This effect has been seen before, such as in [27]. As an illustration to allow for easier visualisation, the alignment of two traps in the thin oxide could allow the electrons to hop through the oxide more easily, while it may require three traps in a row to cause the same effect in the thicker samples. During the process of trapping more charges with increasing bias, the thinner sample should therefore see more noise as the influence of the alignment or disalignment of each electron becoming trapped is more important.

In all samples, the leakage current is observed to increase rapidly between the first measurement points when the measurement is begun at 1 V. This is likely a measurement artifact as no such effect was observed in other measurements where the bias range was larger. After that, the current increases more slowly up until a point around 4 V, depending on the sample, where the current begins to rise at a quicker rate. It can be assumed that a different leakage mechanism has become more significant. Finally, the current makes an abrupt increase indicating a hard breakdown. The leakage current in the 4 nm sample is higher than in the 6 nm samples which is consistent with tunneling equations like Eq. 4 in which at a constant bias, F_{ox} is higher if the oxide is thin.

A soft breakdown can be distinguished in some of the measurements but most progress directly from a low leakage current to a hard breakdown. A soft breakdown was only observed in roughly one tenth of the measurements. The low amount of soft breakdowns detected is likely a result of the area of the dice measured on being relatively large compared to other transistor devices. It has been shown that with a larger area the probability of a soft breakdown occurring is reduced[21].

The results from the SILC measurements have been opted not to be shown as they were not as fruitful as expected. The average measurement showed no clear signs of SILC. What could be construed from the results however was that they coincide with the results already discussed here, such as a soft breakdown only appearing in a small part of the measurements.

A band diagram simulation was performed in the program from Boise State University [11] where the band structure for different regions of the leakage current graph was investigated. In Fig. 10a) the band structure of the 6 nm sample at 2 V is shown, which is the region where the current is slowly increasing, and Fig. 10b) shows the band structure at 4.5 V which is the region where the current increases at a higher rate.

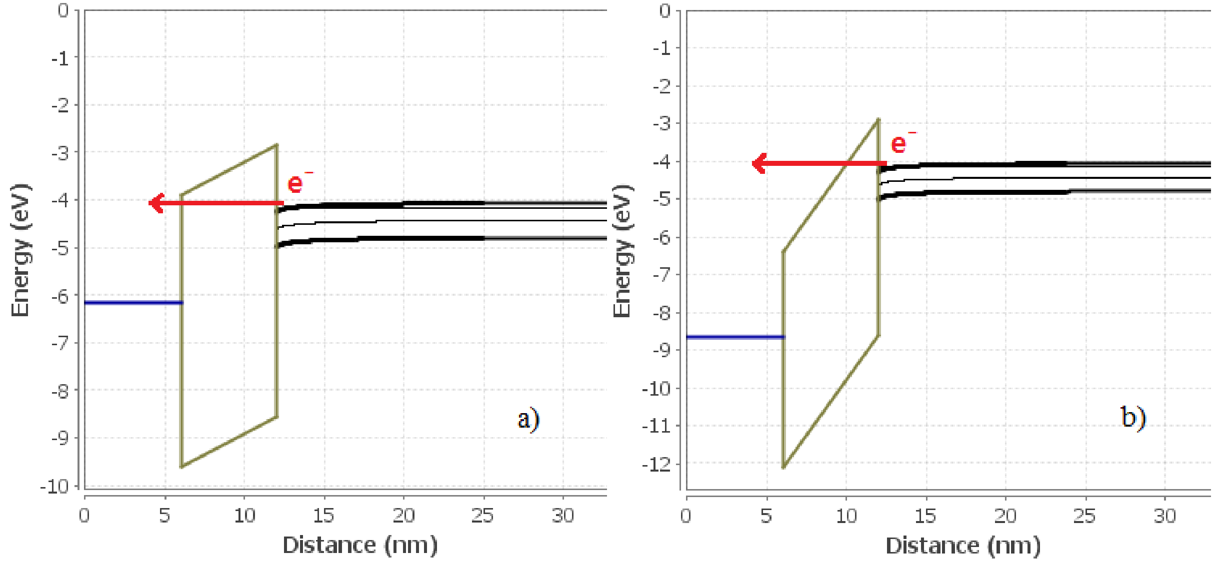

Figure 10: Energy band diagrams for the 6 nm sample. a) Band diagram at 2 V applied bias from the range of slow current increase. b) Band diagram at 4.5 V where the current increase happens at a quicker rate.

In Fig. 10a), the electrons in the conduction band of InGaAs have to tunnel through the complete structure of the potential barrier in order to cause a leakage current. The dominant tunneling process in this region could be either direct tunneling or temperature dependent trap assisted mechanisms. If the dominant process is TAT, the current generated from the trapped electrons that continue to tunnel all the way through the barrier outweighs that generated from DT. The amount of trapped electrons should increase with temperature, which will be analyzed in the next section.

In Fig. 10b), the width of the barrier that the electrons have to tunnel through is reduced as the electric field across the oxide increases. It can be assumed that the dominant leakage mechanism at these higher biases becomes FN tunneling.

A Fowler-Nordheim fit was performed where the logarithm of I_{FN}/F_{ox}^2 was plotted against $1/F_{ox}$, see Fig 11. Values from one measurement on the 6 nm as deposited sample in the region 4.5 V to 5.2 V where FN tunneling appears to be dominant were used for the fit. The plot is linear which indicates that FN tunneling is a good approximation. The slope of the fit is -15 GV/m which can be compared to the theoretical value of -5.6 GV/m calculated using data from Table 2 and a barrier height ϕ_b of 1.45 eV, as obtained from the band diagram program [11]. The difference in the measured and theoretical values can come e.g. from variations in the oxide thickness from the assumed 6 nm, but the theoretical result also changes depending on the source used to obtain m_{ox}^*.

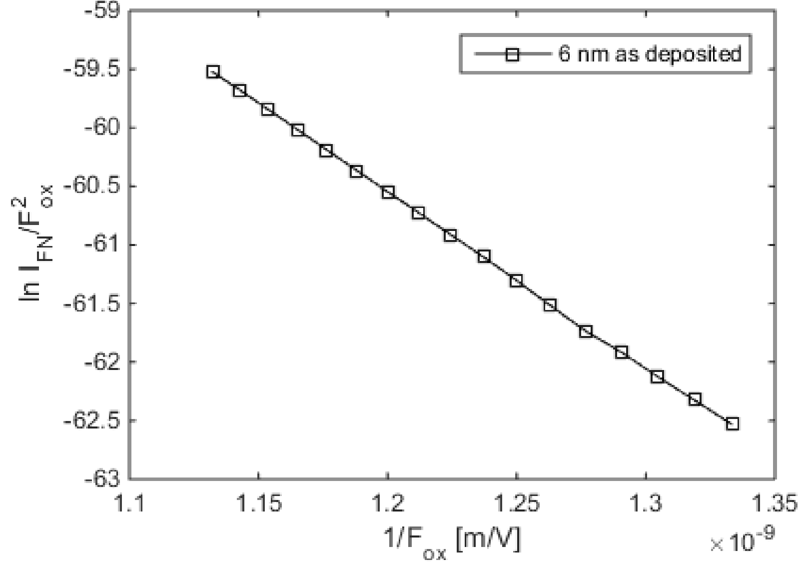

Figure 11: A Fowler-Nordheim plot of the 6 nm as deposited sample between 4.5 and 5.2 V.

4.2 Temperature dependence

The results of the measurements at varying temperature is shown in Fig. 12a) for the 6 nm as deposited sample and Fig. 12b) for the 4 nm sample. The measurement began at -3 V and the arrows indicate how the current changed as the applied voltage increased from -3 V to 3 V and then back to -3 V.

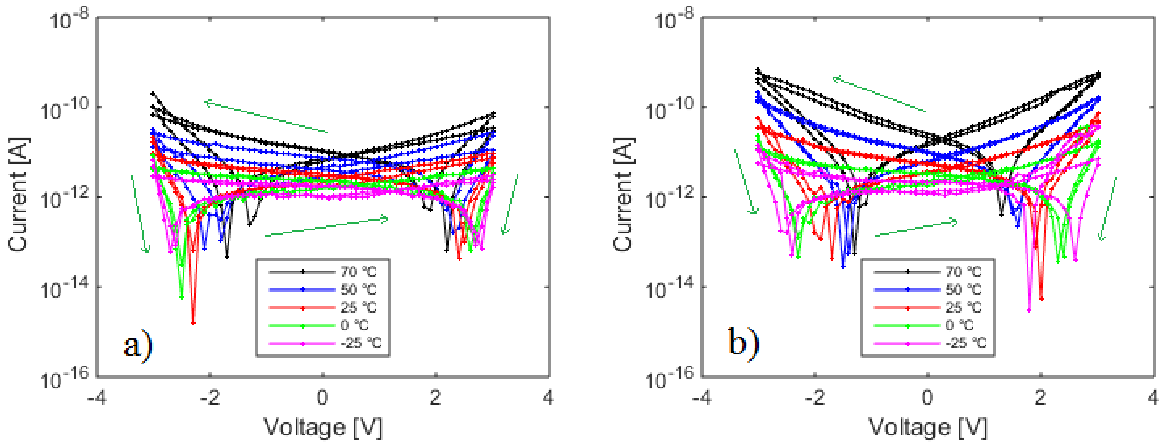

Figure 12: Leakage current at varying temperature of the 6 nm and 4 nm as deposited samples. The arrows indicate the progress of the current.

In Fig. 12 there is an obvious dependence on the temperature as the current exponentially increases with it. The current curves are butterfly shaped which can be explained by considering the behavior of trapping. As the measurement begins at the high negative bias of -3 V, holes are trapped in the oxide. As the applied voltage starts to move towards 0 from -3 V, the measured current quickly reaches a minimum. The minimum current

point represents the zero field condition, which has changed from 0 V because of the large number of trapped holes altering the electric field across the oxide.

When the applied current proceeds towards zero and positive values, the holes being trapped in the oxide are gradually becoming detrapped and replaced by trapped electrons, until the maximum field is reached at 3 V. As the bias is reduced again, the current quickly reaches a zero point similarly to what transpired at negative biases. This time the trapped charges are electrons and the zero field condition is relocated to occur at a positive bias. As the voltage continues back towards negative values, the trapped electrons are replaced by holes until the curve reaches its starting point at -3 V.

It appears that the zero field condition, or compensation field, for trapped charges approaches zero with increasing temperature, indicating fewer trapped charges. This behavior is opposite of the predicted, since the number of trapped charges is expected to increase with temperature. One theory that could explain the change in compensation field is that the probability of carriers being detrapped also increases with the temperature. If the theory is true, at 70 °C the trapped charges are already starting to detrap at a relatively fast rate when the bias changes from its maximum at -3 or 3 V towards zero. It can be distinguished in the figure that the slope of the current between its maximum and the zero field condition is more gentle at higher temperatures compared to the steep slope at low temperature. The slope difference becomes easier to see in Fig. 13 where the current against the theoretical, trapping independent electric field across the oxide is plotted for both the 4 nm and 6 nm as deposited samples. Fig. 13a) is a plot at 0 °C and Fig. 13b) at 70 °C.

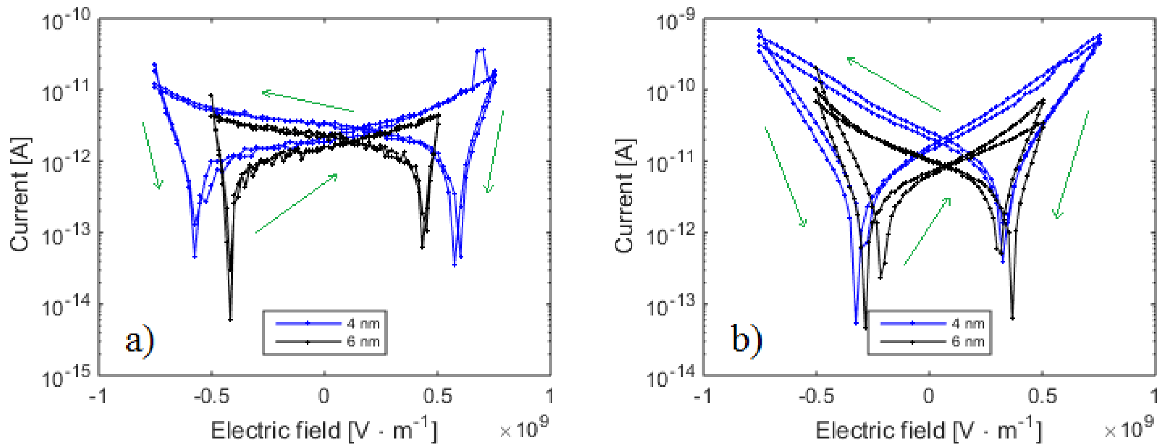

Figure 13: Measured current against the theoretical electric field, independent of trapped charges, for the 4 nm and 6 nm as deposited samples. a) Current measured at a temperature of 0 °C. b) 70 °C.

The number of interface traps in the 4 nm and 6 nm samples are expected to be the same, yet the zero field condition in Fig. 13a) is closer to zero for the thicker sample. This is easily explained by the maximum of the electric field being higher for the thin sample, allowing more charges to be trapped, which shifts the zero field condition away from zero. In Fig. 13b) the zero field condition approaches zero for both samples, possibly due to an increased rate of detrapping at higher temperatures as discussed.

It appears that the hysteresis between the two samples in Fig. 13 increases with the temperature. A reason for this could be that the carriers have a greater energy at higher temperatures which allows them penetrate deeper into the oxide than at low temperature. As a result of the thickness difference, the probability to tunnel through the 4 nm sample increases more with the higher energy than through the 6 nm sample, which would ultimately lead to more leakage.

4.3 Time dependent dielectric breakdown (TDDB)

Fig. 14 shows a measurement performed on 15 dice on the 6 nm annealed sample at a constant voltage stress of 4.4 V. The TDDB in this plot varies between 1 s, which is the time step between each measurement point, and up to 120 s. Measurements at 4.2 V and 4.5 V, and also those on the other samples, show a similar seemingly random distribution.

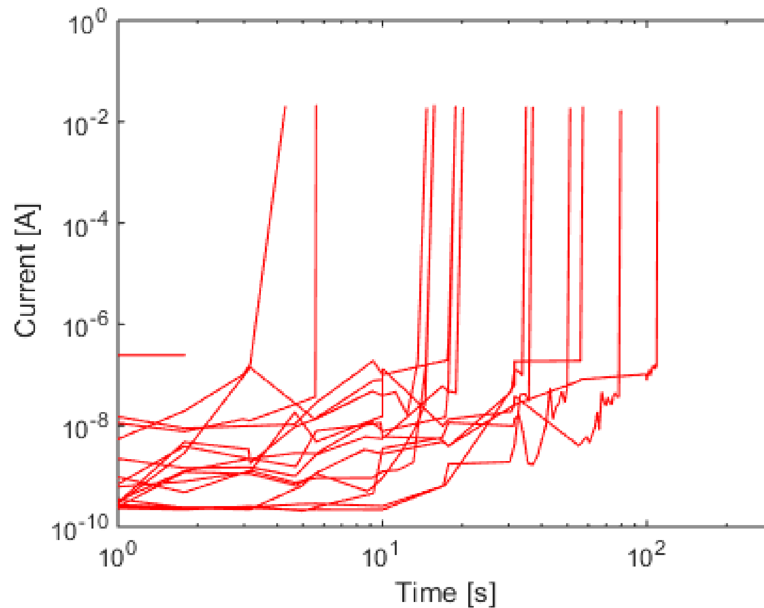

Figure 14: Leakage current measurement on the 6 nm annealed sample at a CVS of 4.4 V.

It was clear that the TDDB took place faster as the stress voltage grew closer to the average breakdown voltage measured for each respective sample. The distribution region widened with lower bias or became more narrow with a bias closer to the breakdown voltage. Apart from this, no real conclusion could be drawn from these measurements regarding the breakdown time. After seeing the wide distribution in breakdown voltage for each sample, the result obtained for the TDDB was expected since it makes a significant difference if the TDDB measurement is performed at 4.4 V and the die would have had a breakdown at 4.7 V or 5.2 V.

5 Summary and outlook

Three MOS samples using InGaAs with HfO_2 layers of 4 nm, 6 nm and 6 nm, applied by ALD, were processed. One of the 6 nm samples was annealed. The electric field required to cause a hard breakdown through the oxide was $\sim$ 0.81 GV/m in the 6 nm annealed sample, $\sim$ 0.90 GV/m in the 6 nm as deposited sample and $\sim$ 1.13 GV/m in the 4 nm sample and thus it decreased with oxide thickness and PMA treatment. The lower breakdown field might be attributed to higher polycrystallinity in the thicker oxides that allow a percolation path to form more easily. The results from all three samples were widely distributed which likely stemmed from a variation in the density of interface traps across the oxide layer.

The dominant leakage mechanism at higher biases was determined to be Fowler-Nordheim tunneling. The results suggested that direct tunneling or trap assisted tunneling had the largest influence at low bias, with trap assisted tunneling gaining more significance as temperature increased.

The results obtained in this work could likely have been improved if the fabrication process was more tightly controlled. With a more uniform oxide layer having fewer defects distributed across it directly after the ALD process, the results should have been more centralized. Besides improving the process utilized in this work, there are numerous other parameters during the ALD process that can be modified in attempts to improve results, such as changing the temperature and pulse lengths.

The Hf precursor used, TDMA, is a common Hf precursor but there are many other compounds that include Hf and can be adopted in the ALD process. There could be a variation in how uniformly different precursors administer the Hf across the surface, and how much contamination of other particles from the precursor end up on the surface. There are also other precursors for oxygen available that can be tested. In addition, methods to reduce the amount of abundant oxygen such as plasma etching are available. Obviously, if interest is not of the results from HfO_2/InGaAs in particular, there are many other materials such as InAs or GaAs as semiconductor, and Al_2O_3 or ZrO_2 as oxide that can be used in various combinations.

The measurements suggested that the leakage current increase due to temperature is likely a result of TAT, but other mechanisms can not be discarded. Measurements of e.g. the relaxation current could have been performed had there been enough time, to understand its impact on the leakage current. On the subject of other measurements, CV measurements could have been done to identify the filling and creation of interface traps, and XPS measurements could have given an image of the oxide layer in the samples for an assessment of the amount of defects present.

Bibliography

[1] Robert R Schaller. Moore's law: past, present and future. *Spectrum, IEEE*, 34(6):52–59, 1997.

[2] Jesús A Del Alamo. Nanometre-scale electronics with III-V compound semiconductors. *Nature*, 479(7373):317–323, 2011.

[3] John Robertson. High dielectric constant gate oxides for metal oxide Si transistors. *Reports on Progress in Physics*, 69(2):327, 2006.

[4] Ioffe Physico-Technical Institute Website NSM Archive. [Online]. Available at: http://www.ioffe.rssi.ru/SVA/NSM/Semicond/. Acessed 2015-03.

[5] Chris Beer. *Fabrication and characterisation of novel Ge MOSFETs.* PhD thesis, University of Warwick, 2007.

[6] J Robertson. High dielectric constant oxides. *The European physical journal applied physics*, 28(03):265–291, 2004.

[7] Yee-Chia Yeo, Tsu-Jae King, and Chenming Hu. MOSFET gate leakage modeling and selection guide for alternative gate dielectrics based on leakage considerations. *Electron Devices, IEEE Transactions on*, 50(4):1027–1035, 2003.

[8] Patrick S Lysaght, Joseph C Woicik, M Alper Sahiner, Byoung-Hun Lee, and Raj Jammy. Characterizing crystalline polymorph transitions in HfO 2 by extended x-ray absorption fine-structure spectroscopy. *Applied Physics Letters*, 91(12):122910–122910, 2007.

[9] Claudia Richter, Tony Schenk, Uwe Schroeder, and Thomas Mikolajick. Film properties of low temperature HfO2 grown with H2O, O3, or remote O2-plasma. *Journal of Vacuum Science & Technology A*, 32(1):01A117, 2014.

[10] Bart Van Zeghbroeck. Principles of semiconductor devices. *Colarado University*, 2004.

[11] Boise State University Band Diagram Program Website. [Online]. Available at: http://nano.boisestate.edu/research-areas/multi-dielectric-energy-band-diagram-program/. Accessed 2015-04.

[12] Robert Entner. *Modeling and simulation of negative bias temperature instability.* 2007, PhD thesis.

[13] G Bersuker, J Yum, L Vandelli, A Padovani, L Larcher, V Iglesias, M Porti, M Nafria, K McKenna, A Shluger, et al. Grain boundary-driven leakage path formation in HfO2 dielectrics. *Solid-State Electronics*, 65:146–150, 2011.

[14] K Xiong, J Robertson, MC Gibson, and SJ Clark. Defect energy levels in HfO 2 high-dielectric-constant gate oxide. *Applied Physics Letters*, 87(18):183505–183505, 2005.

[15] JR Weber, A Janotti, and CG Van de Walle. Native defects in Al2O3 and their impact on III-V/Al2O3 metal-oxide-semiconductor-based devices. *Journal of Applied Physics*, 109(3):033715, 2011.

[16] Simon M Sze and Kwok K Ng. *Physics of semiconductor devices.* John Wiley & Sons, 2006.

[17] Juan C Ranuarez, M Jamal Deen, and Chih-Hung Chen. A review of gate tunneling current in MOS devices. *Microelectronics reliability*, 46(12):1939–1956, 2006.

[18] Michel Houssa, M Tuominen, M Naili, Valeri Afanasev, Andre Stesmans, S Haukka, and MM Heyns. Trap-assisted tunneling in high permittivity gate dielectric stacks. *Journal of Applied Physics*, 87(12):8615–8620, 2000.

[19] M Lenzlinger and EH Snow. Fowler-Nordheim Tunneling into Thermally Grown SiO2. *Journal of Applied physics*, 40(1):278–283, 1969.

[20] M Porti, S Gerardin, M Nafrı, X Aymerich, A Cester, A Paccagnella, P Schiavuta, R Pierobon, et al. Systematic characterization of soft-and hard-breakdown spots using techniques with nanometer resolution. *Microelectronic engineering*, 84(9):1956–1959, 2007.

[21] Muhammad A Alam, Bonnie Weir, Jeff Bude, P Silverman, and Don Monroe. Explanation of soft and hard breakdown and its consequences for area scaling. In *Electron Devices Meeting, 1999. IEDM'99. Technical Digest. International*, pages 449–452. IEEE, 1999.

[22] TP Chen, MS Tse, CQ Sun, S Fung, and KF Lo. Snapback behaviour and its similarity to the switching behaviour in ultra-thin silicon dioxide films after hard breakdown. *Journal of Physics D: Applied Physics*, 34(17):L95, 2001.

[23] P Shekhter, F Palumbo, K Cohen Weinfeld, and M Eizenberg. X ray photoelectron analysis of oxide-semiconductor interface after breakdown in Al2O3/InGaAs stacks. *Applied Physics Letters*, 105(10):102908, 2014.

[24] John R Vig. UV/ozone cleaning of surfaces. *Journal of Vacuum Science & Technology A*, 3(3):1027–1034, 1985.

[25] Steven M George. Atomic layer deposition: An overview. *Chemical Reviews*, 110(1):111–131, 2009.

[26] CH Chang, Y-K Chiou, YC Chang, KY Lee, TD Lin, TB Wu, M Hong, and J Kwo. Interfacial self-cleaning in atomic layer deposition of HfO 2 gate dielectric on In 0.15 Ga 0.85 As. *Applied physics letters*, 89(24):242911–242911, 2006.

[27] N Wrachien, A Cester, E Zanoni, G Meneghesso, YQ Wu, and PD Ye. Degradation of III–V inversion-type enhancement-mode MOSFETs. In *Reliability Physics Symposium (IRPS), 2010 IEEE International*, pages 536–542. IEEE, 2010.