Fabrication, Characterization, and Simulation of Highly-Scaled III-V MOS Devices

by

Vallen Gary Rezazadeh

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Solid-State Electronics

Department of Electrical and Computer Engineering
University of Alberta

© Vallen Gary Rezazadeh, 2017

Abstract

To enable scalable MOSFET technology in III-V semiconductor platforms, high quality semiconductor-oxide interfaces are essential. In this work, the role of surface reactions in the oxide deposition process is examined, with the objective of optimizing the thermodynamics of the semiconductor-oxide interface. A novel low-temperature plasma-enhanced atomic layer deposition (PEALD) technique was applied to deposit nanoscale high-k dielectrics on several III-V substrates, including InP, GaAs, InAs, and GaN. Approximately 7 nm of ZrO_2 was grown and patterned to form MOSCAP structures, which were subsequently analyzed through electrical characterization to evaluate dielectric and interface quality. The oxide films fabricated were found to have interface trap densities ranging from 10^{10}-10^{13} $eV^{-1}cm^{-2}$, and showed high capacitance densities (~2.5 $\mu F/cm^2$). GaN and InP MOSCAPs with ZrO_2 dielectric layers were found to have gate currents in line with direct tunneling phenomena and MOS mobilities approaching that of doped bulk semiconductors. Scaled InP MOSFET devices using these experimental oxide results were also simulated using improved device structures, with the aim of demonstrating optimal parameters for effective reduction of short-channel effects. These devices demonstrated high performance for various scaled gate lengths, with FinFET structures showing excellent frequency response at the 7 nm process node.

Preface

Significant portions of Chapters 3 and 5 of this thesis, along with excerpts of Chapters 4 and 6 have been previously published as Rezazadeh, Vallen, et al. "Defect Characterization of PEALD High-k ZrO_2 Films Fabricated on III-V Materials." *IEEE Transactions on Semiconductor Manufacturing* 29.4 (2016): 355. The author of this thesis was responsible for data collection, analysis, and manuscript composition in the aforementioned work. Several other individuals contributed in part to this previous work: K. M. Bothe assisted with data collection and production of scientific figures, A. Afshar assisted with the deposition of ALD high-k films used in experiments, while D. W. Barlage and K. C. Cadien were supervisory authors and assisted with concept formation and manuscript composition.

“Innovation is everything.”

Robert N. Noyce (1927-1990)

Acknowledgements

This work would not be possible without the contributions of several individuals.

Firstly, I would like to thank both Dr. Kyle Bothe and Dr. Douglas Barlage for their mentorship and technical guidance throughout this study. Their boundless expertise in device physics and design helped me through the toughest of mental gymnastics along the way.

In addition, I would like to thank my lab colleagues for their valuable discussions and assistance, including Mourad Benlamri, Christopher Cerioci, Alex Ma, Mei Shen, and Gem Shoute. I would also like to thank our collaborators Dr. Ken Cadien, Dr. Amir Afshar, and Dr. Triratna Muneshwar for running the ALD processes used in this work.

Last, but not least, I would like to thank my family, namely my mother Joanne, my father Nader, and my brother Mason for their unwavering support during my academic endeavours.

Table of Contents

List of Tables

List of Figures

Chapter 1.

Introduction

Since the invention of the transistor by Bardeen and Brattain in 1947 [1], semiconductor engineers have sought out increasingly innovative solutions to continue the frenetic improvements in performance and density that we now take for granted in electronic devices. Beginning with the observation of Moore's Law as a projection of exponential growth in integrated circuits [2] and the subsequent introduction of self-aligned gate MOSFETs [3], process engineers and device designers alike have continuously driven forward ground-breaking solutions to the seemingly insurmountable physical roadblocks faced along the way. As a result of this dependable progress, countless derivative technologies have been made possible, including staples of contemporary semiconductor applications such as multicore CPUs, monolithic microwave integrated circuits (MMICs), and high-speed mixed-signal ICs for optical communications. With the total revenues of the semiconductor industry reaching $335 billion in 2015 [4], the continued development of transistor technology is clearly critical. In recent years however, the viability of scaled transistors with sustained performance in both high-speed digital and RF applications has been called into question. This section will discuss the problems faced, and analyze how III-V materials in combination with high-k dielectrics could provide a practical solution to drive forward highly-scaled MOS devices with suitable performance against existing technologies such as Si MOSFETs and III-V HEMTs.

1.1. MOSFET Devices and Scaling

The metal-oxide semiconductor field effect transistor, or MOSFET, is by far the most common piece of electronics technology in existence today. As the foundation for modern digital logic in CMOS ICs, MOSFETs are fabricated by the billions in CPUs for many consumer electronic devices. In addition to their digital applications, MOSFETs also play an important role in analog circuits, mainly in the form of low-cost Si RFICs for cell phones, where they are often monolithically integrated with SiGe heterojunction bipolar transistors for improved amplifier

performance. Structurally, the MOSFET in its simplest form consists of three major components: a semiconducting substrate, doped source and drain regions, as well as an insulating gate dielectric, as shown in Figure 1.1. While alternative semiconductor substrates have been investigated for use as MOSFET IC substrates, silicon continues to remain the only truly viable choice due to a variety of technical challenges.

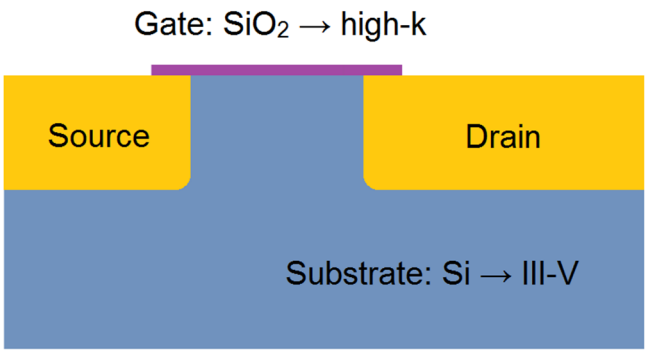


Figure 1.1. Schematic showing the primary components of a MOSFET. The modification of gate oxide and substrate materials in this study are indicated.

In addition to the economic scaling advantage foreseen by Moore's Law, several critical device performance figures are also improved by reducing device feature size. One such benchmark is power consumption, which is inversely correlated to the reduction in geometry. This reduction is extremely important in digital applications, where high density low-power devices have become possible through continued process scaling. Another critical benchmark improved by scaling is the cutoff frequency f_T, an important figure governing the maximum frequency response of a transistor and useful for high-frequency analog and mixed-signal applications. In the process of providing the scaling required to achieve these performance gains however, engineers have had to periodically introduce new technologies into MOSFET fabrication to mitigate the various negative effects encountered at short gate lengths. These technologies have included substantial overhauls of key MOSFET structural components, such as the introduction of the high-k gate dielectric, strain-engineered channels, and most recently the FinFET and Tri-Gate transistors. While these improvements have maintained the high performance of CMOS devices, they have also stretched the limits of sustained geometry scaling. This poses a serious problem for dense applications of CMOS, where scaling has continuously driven down both costs and power budgets for ICs. In a similar fashion, while RF applications have benefitted greatly from increased f_T through scaling, serious issues have surfaced that threaten the use of nanoscale MOSFETs for RFICs. These include

the reduction of intrinsic voltage gain (g_m/g_{ds}) [5], reduction of operating voltages [5], reduction of the maximum oscillation frequency f_{max} due to increased parasitics [6], as well as reduction of the quality of passive devices due to increased substrate doping [5]. With all these issues coalescing, engineers are now exploring alternative materials and device structures to continue to drive MOSFET technology forward in both digital and RF. One such solution is the introduction of III-V semiconductor substrates [7], which could potentially improve characteristics over a wide range of gate lengths and provide a viable platform for scaled MOSFETs. In this work, we will analyze the optimization of III-V MOS devices to enable suitable performance in short-channel scenarios.

1.2. III-V Semiconductors

While the commercial development of MOSFETs in non-silicon platforms has been lacking, III-V semiconductors are attractive for a multitude of applications, owing to their superior material properties compared to silicon. As opposed to single-element semiconductors such as silicon and germanium, III-V semiconductors are compound semiconductor materials, consisting of one element from Group III and one from Group V of the periodic table. As seen in Table 1.1, one of the main advantages of III-V materials is their higher electron mobilities compared to Si, with improvements of upwards of 1-2 orders of magnitude. Alongside improved mobility figures, III-V semiconductors such as GaN, GaAs, and InP have improved breakdown fields compared to Si, enabling higher operating voltages. These characteristics have made III-V transistors the devices of choice for power amplifiers operating at high frequencies [8].

Table 1.1. Comparison of Semiconductor Properties [9]

Material	Bandgap (eV)	Electron Mobility ($cm^2V^{-1}s^{-1}$)	Hole Mobility ($cm^2V^{-1}s^{-1}$)	Breakdown Field (Vcm^{-1})
InSb	0.170	77000	850	$1x10^3$
GaSb	0.726	3000	1000	$5x10^4$
InAs	0.354	40000	500	$4x10^4$
GaAs	1.424	8500	400	$4x10^5$
InP	1.344	5400	200	$5x10^5$
GaN	3.200	1000	350	$5x10^6$
Si	1.120	1400	450	$3x10^5$
Ge	0.661	3900	1900	$1x10^5$

As a result of the high mobilities in III-V semiconductors, electronic devices made from these materials also have reduced noise profiles compared to silicon [10]. For even further improved performance, III-V devices can be constructed using a high electron-mobility transistor structure (HEMT). Exploiting the ability of III-V materials to form tertiary compounds, these devices use the difference in bandgap between two epitaxially-grown lattice-matched III-V compounds to form a quantum well for electron transport. The well is then filled by a technique called modulation doping, which enables the dopants to be physically separated from the channel and thus prevents scattering [11]. As a result, HEMTs have been shown to provide dramatic increases in electron mobility, with peak mobility figures reaching 23500 cm^2/Vs for devices on GaAs wafers [12]. All of these factors, coupled with the high resistivity semi-insulating substrates offered by the GaAs, InP, and GaN material platforms, have made III-V transistors standard in many microwave applications [13]. Even with these benefits however, III-V transistors have not been immune from the substantial barriers confronted during device scaling. Although reduction of L_g has been shown to produce significant increases in f_T figures for HEMT devices, it has come at the expense of f_{max} and overall device power handling (owing to low operating voltages and only small increases in I_{ds}) [14]. In addition, depletion mode devices such as HEMTs suffer from a lack of interoperability with CMOS logic, significantly inhibiting their use in mixed-signal applications.

While III-V devices such as MESFETs and HEMTs have long been the industry standard for high-performance microwave-frequency applications, it has not been until recently that III-V materials have been studied for use in highly scaled MOSFETs. Although III-V MOSFETs should have major advantages over Si MOSFETs such as higher drive current (since $I \alpha \mu$) and lower gate delay due to improved carrier transport, their implementation also presents major setbacks in device operation. One of the most significant issues is the increased effect of short-channel effects and band-to-band tunnelling compared to silicon owing to their inherent material properties (described in further detail in Chapter 6). As a result, MOSFETs constructed from III-V materials suffer from substantially higher off-state current, particularly in those materials with small bandgaps.

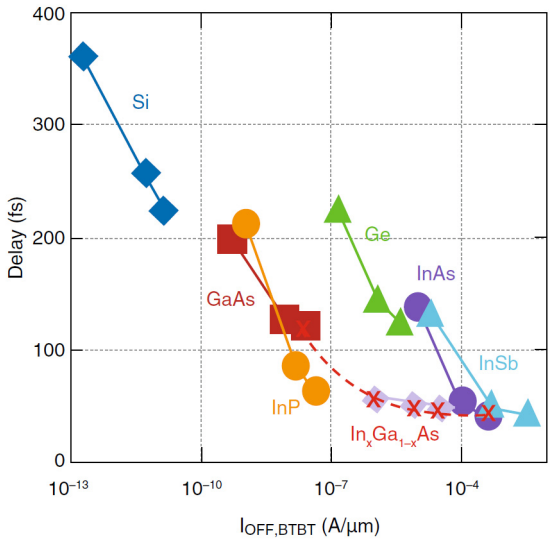


Figure 1.2. Plot of gate delay versus OFF-state band-to-band tunnelling current for various semiconductors in a 15 nm L_g MOSFET from Krishnamohan [15].

Although these negative effects severely constrain the use of InAs and InSb in highly-scaled low-power MOSFETs, materials such as InP and GaAs show much more manageable characteristics. In addition, the issues caused by short-channel effects can be reduced through the use of thinner semiconductor bodies, reducing leakage current and improving performance [15]. This may be achieved through the use of a multi-gate device such as a FinFET, or the use of III-V on insulator substrates as being developed by several groups [16] [17]. Alongside these approaches, other techniques in device design can also be used to mitigate these effects, as will be discussed in Chapter 6. Even without drastic corrective action to reduce off-current, highly-scaled III-V MOSFETs show significant promise for high-performance applications where power consumption is a less significant concern.

1.3. High-k Oxides and Their Characteristics

One of the most significant updates to the MOSFET process flow was the substitution of silicon dioxide gate dielectrics with that of high-k oxides. In order to effectively scale a MOSFET between process generations, it is necessary to reduce the thickness of gate oxide, a layer usually formed through thermal oxidation of the silicon substrate. With gate oxides approaching 1 nm at the 65 nm process node [18], gate leakage via tunnelling was quickly becoming unsustainable for future process generations. As a result, engineers introduced gate oxides based on transition

elements into the gate stack, removing the well-established SiO_2 layer [19]. This was accomplished through the introduction of a novel technique called atomic layer deposition (ALD), a process which enables the nanoscale deposition of the high-k dielectrics required (discussed in detail in Chapter 2). These materials offer dielectric constants several times higher than that of SiO_2, and can thus be used in a thicker film to offer an equal oxide capacitance density. From the formula for the gate oxide capacitance (identical to that of a parallel-plate capacitator):

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

As seen in Equation 1, the only two ways to increase the capacitance density are by decreasing the thickness of the dielectric (untenable as discussed earlier) or by using a material with a higher dielectric constant. In this case, the high-k dielectric produces an identical oxide capacitance to that of an equivalent oxide thickness (EOT) of SiO_2:

$$EOT = \frac{\kappa_{Si}}{\kappa_{high-k}}\, t_{high-k} \tag{2}$$

where $\kappa_{Si} = 3.9$ is the dielectric constant of SiO_2. This parameter is most commonly used as a figure of merit when comparing high-k process nodes to previous generations of processes.

Table 1.2. Comparison of Bulk Properties of Gate Dielectrics [20]

Material	Dielectric Constant κ	Band Gap (eV)	Cond. Band Offset on Si (eV)
Al_2O_3	9	8.8	2.8
Y_2O_3	15	6	2.3
Ta_2O_5	22	4.4	0.35
HfO_2	25	5.8	1.4
ZrO_2	25	5.8	1.5
La_2O_3	30	6	2.3
SiO_2	3.9	9	3.2
Si_3N_4	7	5.3	2.4

1.4. Historical Oxide Deposition on III-V

Although III-V semiconductors provide many tangible benefits for MOSFETs in various applications, the integration of high-quality MOS interfaces has proven difficult owing to the

undesirable chemical nature of many of these semiconductors [21]. Efforts that predate the development of ALD for high-k oxides were primarily focussed on the deposition of group III oxides, such as SiO_2 and Ga_2O_3 (gallium (III) oxide). Some examples of this work include deposition of gallium (III) oxide on GaAs through molecular beam epitaxy (MBE) [22], as well as the deposition of silicon dioxide on GaAs through liquid phase deposition [23]. While both of these studies achieved relatively promising results, the quality of the MOS interfaces was still short of that needed for production MOSFET devices. With the development of ALD as a viable oxide deposition method, the focus of research quickly shifted to high-k dielectrics. This presented a significant opportunity for exploration of MOS interfaces in III-V, as high-k oxides were essentially undeveloped in both Si and compound semiconductors. Since then, a large volume of work has been completed studying high-k on an array of III-V substrates, with a fairly wide range of results [7]. Investigations have explored different oxide materials (focussed primarily on Al_2O_3 and HfO_2), in addition to various surface pre-treatments and passivation techniques.

It is clear that with the rapid introduction of various ALD techniques for high-k oxides over the past decade, III-V semiconductors have become serious potential candidates for future generations of high-performance MOSFETs, in particular with respect to large-scale integration. With improved ALD processes and optimized parameters, high-quality MOS interfaces on III-V materials should be within close reach. In this work, we will examine a suitable high-k gate dielectric, ZrO_2, and subsequently analyze its chemical compatibility with the most prominent III-V semiconductors. As seen in Table 1.2, ZrO_2 offers one of the largest dielectric constants of commonly used high-k gate oxides, yet has been relatively underutilized in studies of high-k oxides on III-V materials. Subsequent chapters will focus on the deposition mechanisms and conditions of ALD ZrO_2 on III-V substrates, while making use of characterization of fabricated MOS devices to evaluate dielectric quality and the ability of these materials to satisfy requirements for highly-scaled devices.

Chapter 2. Atomic Layer Deposition

As discussed in Chapter 1, the proliferation of atomic layer deposition in microfabrication has been the critical enabling factor for deposition of high-k dielectrics on semiconductors. The ability to control film thickness to one monolayer of material differentiates ALD from earlier methods of chemical oxide deposition such as PECVD, where quality dielectric films can only be achieved in the range of 100 nm or greater. This feature of ALD is essential for highly scaled MOS gate oxide layers, where film thickness must be kept less than 10 nm. This chapter will introduce the mechanics of ALD deposition and discuss its application in III-V materials systems.

2.1. Operational Principles of ALD

Atomic layer deposition is a chemical vapour deposition technique that relies on the principle of so-called “half-reactions” to facilitate deposition of a desired material [24] [25]. This mechanism works by initially introducing a pulse of a gas-phase chemical precursor into the ALD chamber. The precursor molecules adsorb to the surface of the substrate, and the chamber is subsequently purged. A pulse of the next precursor is then pumped into the chamber, enabling a reaction to occur with the first precursor at the surface and thus forming a layer of the thin-film desired. The chamber is once again purged to remove excess reactants, and the process is repeated until the desired thickness of film is achieved. A unique aspect of this process is the fact that the thin-film formation reaction is completely self-limited, as the second precursor molecules introduced to the chamber are only able to react with the limited amount of the first precursor already adsorbed to the surface of the substrate. This ensures that the film deposited is highly conformal and uniform [26].

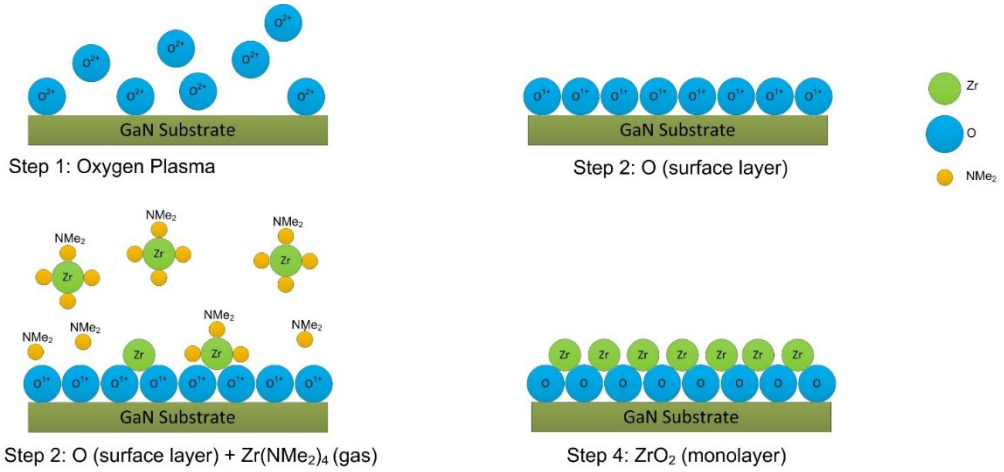


Figure 2.1. Half-reactions of the ZrO_2 PEALD deposition process illustrating adsorption and monolayer formation from von Hauff [27].

In addition to the ability to control the deposition of films to the atomic level, the ALD process also provides certain other advantages over other traditional chemical deposition techniques. One such advantage is the comparatively low deposition temperature possible with ALD deposition. Whereas previous techniques require temperatures of 300 °C (PECVD) to upwards of 700 °C (LPCVD) for effective deposition, ALD can be used with a wide range of deposition temperatures ranging from room temperature to 300 °C. As a result, the ALD has a fairly low thermal budget, and can be used at various points in the fabrication of devices.

2.2. High-k Oxides and the PEALD Process

In order to deposit high-k oxides via the ALD process, the most common process used is a thermal ALD (TALD) process. In this class of process, the high-k precursor is introduced into the chamber and reacts directly with the surface of the heated substrate, leaving only a portion of the molecule remaining. The chamber is purged and an oxygen source is subsequently introduced, most commonly water (H_2O) or ozone (O_3). The oxygen source proceeds to react with the metal oxide molecule on the surface of the substrate, forming the first layer of the high-k oxide thin film.

Alternatively, a plasma-enhanced ALD (PEALD) process can be also used to deposit high-k oxides. The PEALD process works in a similar but slightly different fashion to the TALD approach. The high-k precursor is introduced into the ALD chamber and adsorbs to the substrate as previously described. Instead of using a gas-phase oxygen source however, the oxygen is

introduced through use of a plasma species. This plasma can be of three different types; namely a direct plasma generated at the surface of the substrate, a remote plasma generated separately from the substrate, or a radical-enhanced plasma where ions and electrons from the plasma are prevented from interacting with the substrate [28]. Use of plasma O_2 versus gas-phase oxygen sources provides several benefits, including precision control over deposition exposure parameters as well as the ability to lower the reaction temperature required [28].

In common with both deposition methods however are the selection of precursors to act as the metal source for the high-k oxide. These precursors are divided into two main groups, namely metal halides and organometallic compounds. The metal halides are most often chlorinated species such as $HfCl_4$ and $ZrCl_4$, while the organometallics are organic-chain molecules such as $Al_2(CH_3)_6$ (tetramethylammonium or TMA). The primary difference separating most of these precursors are their individual gas phase characteristics such as vapour pressures and boiling/sublimation temperatures. In addition to process-specific considerations, organometallic precursors may be preferred in some applications due to the highly corrosive nature of metal halide reaction by-products [29]. These chemicals can cause degradation of reactor chambers, as well as any metal structures already formed on the deposition substrate [30].

2.3. Previous Results using PEALD

Of the high-k oxides, the most commonly studied materials for MOS applications have been by far Al_2O_3 and HfO_2 deposited through TALD processes. Although PEALD methods are interesting for deposition of oxide materials from a process engineering perspective, use of these processes in high-k studies is comparatively scarce. Despite this, the Cadien Nanofabrication Group in the Department of Chemical and Materials Engineering at the University of Alberta is actively studying the deposition of a variety of metal-oxide compounds using a highly-optimized PEALD process. Some of the recipes used include deposition of HfO_2 and ZrO_2 using organometallic precursors for high-k dielectric applications. These films have been previously deposited on GaN substrates to form extremely high-quality MOS interfaces, achieving interface trap densities (D_{it}) in the range of 10^{11} $cm^{-2}eV^{-1}$ and oxide capacitance densities (C_{ox}) in excess of 2.7 μFcm^{-2} [31]. Of these oxides, the ZrO_2 films deposited were of particularly high quality, with a peak effective MOS mobility of 375 $cm^2V^{-1}s^{-1}$, nearly twice that of the HfO_2 PEALD film. While deposition of Al_2O_3 is possible with the same process, its properties are overall less desirable than that of HfO_2 and

ZrO_2. Although Al_2O_3 ALD MOS structures have demonstrated low D_{it} figures, they tend to suffer from C_{ox} values several multiples lower than equivalent HfO_2 and ZrO_2 films. In addition, there is evidence that the superior band gap of Al_2O_3 is diminished significantly in ALD films, potentially eliminating any effective leakage current advantage [32]. Taking all this evidence into account, PEALD ZrO_2 appear to show the most promise for higher quality MOS interfaces on III-V semiconductors.

Chapter 3. Surface Chemistry of III-V Materials

N.B.: A significant portion of this chapter has been adapted from Rezazadeh et al. [33]

In order to effectively optimize the deposition of high-k oxides on III-V materials, a comprehensive understanding of the surface chemistry of III-V semiconductors is required. When attempting to create a high-quality and low defect MOS interface on III-V semiconductor substrates, several issues come into play including the orientation of bonds at the surface, the chemical nature of the native oxide layer, as well as the thermodynamics and mechanism of deposition. All of these competing factors must be considered in full when designing an optimal oxide deposition for novel semiconductor materials.

3.1. Native Oxides and Bonding on III-V

A significant component in the formation of traps at any MOS interface is the chemical nature of the native oxide layer formed on the surface of the semiconductor in equilibrium conditions. On Si substrates, this layer consists solely of SiO_2, with no other oxide components viable in normal conditions. Conversely, on III-V substrates, the native oxide layer will almost always contain several constituent compounds owing to the binary nature of the semiconductor. While this layer usually contains mostly phases of both pure group III oxides and group V oxides, certain materials such as InP have additional phases such as phosphates that are also possible. The complex interplay of the elements at the MOS surface is shown in the Zr-O-III-V quaternary phase diagram in Figure 3.1, where the large number of tie lines demonstrates the difficulty in isolating a pure semiconductor-oxide interface. Previous work on fabricating III-V MOS interfaces has included utilizing the native oxide layer through thermal and anodic oxidation of the underlying III-V substrate, however results using these methods have generally remained inferior to external deposition of oxide films [34]. For most III-V materials examined, the temperature of deposition greatly influences the composition and stoichiometry of this native oxide layer, and may precipitate other reactions such as substrate decomposition as well.

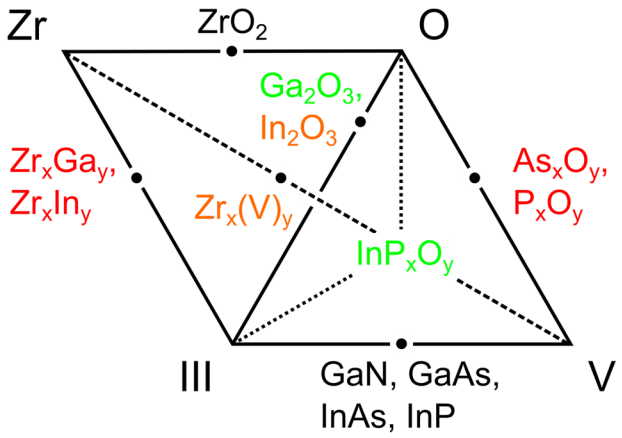


Figure 3.1. **The chemical phase diagram between group III elements, group V elements, elemental zirconium, and elemental oxygen. Formed compounds are shown at points along lines between vertices of the triangles [21]. Green labels indicate desired insulators, orange labels indicate mediocre insulators, and red labels indicate compounds with significant negative effects on MOS device operation.**

One useful method for analyzing the constituent parts of the main oxidation reaction is through an examination of thermodynamic quantities. A particularly suitable figure in this regard is the enthalpy of formation ΔH_f, which measures the relative chemical energy of a compound compared to its elemental precursors. By tabulating the various stable compounds contained within the phase diagram of the oxidation reaction according to ΔH_f, it is possible to determine which chemical compounds or phases will be favoured in the resulting products. Table 3.1 shows the most energetically preferred compounds for several III-V semiconductors (omitting pure Zr oxides) with ΔH_f figures obtained through density functional theory (DFT) calculations [35]. Known undesirable or poor insulators on each interface are italicized. For all semiconductors, Zr-III intermetallics have ΔH_f values close to that of the substrate and are not thermodynamically viable, and have thus been omitted.

Table 3.1. Enthalpy of Formation ΔH_f (eV/atom) of Various Stable Compounds in the Quaternary Zr-O-III-V System (*undesirable compounds in italics*)

	In	**Ga**	**Al**
P	$Zr_2P_2O_9$ (-3.186), ZrP_2O_7 (-3.013), InP_3O_9 (-2.419), $InPO_4$ (-2.386), *P_2O_5* (-2.369), In_2PO_5 (-2.199), *In_2O_3* (-1.882), *ZrP (-1.525), Zr_7P_4 (-1.273), ZrP_2 (-1.155), Zr_3P (-0.935)*		
As	*In_2O_3 (-1.882), As_2O_3 (-1.427), As_2O_5 (-1.398), $ZrAs$ (-1.149), Zr_3As_2 (-1.058), Zr_7As_4 (-1.003), $ZrAs_2$ (-0.845)*	Ga_2O_3 (-2.147), *$GaAsO_4$ (-1.796), As_2O_3 (-1.427), As_2O_5 (-1.398), $ZrAs$ (-1.149), Zr_3As_2 (-1.058), Zr_7As_4 (-1.003), $ZrAs_2$ (-0.845)*	Al_2O_3 (-3.300), *$AlAsO_4$* (-2.298), *As_2O_3 (-1.427), As_2O_5 (-1.398), $ZrAs$ (-1.149), Zr_3As_2 (-1.058), Zr_7As_4 (-1.003), $ZrAs_2$ (-0.845)*
N		Ga_2O_3 (-2.147), *ZrN (-1.800), Zr_3N_4 (-1.610), Zr_3N_2 (-1.484)*	

As seen in Table 3.1, both InP and GaN have proportionally higher numbers of desirable interfacial compounds as opposed to the other semiconductors in this system. In addition, it is clear that the composition of the interface presents serious issues in semiconductors such as GaAs and InAs, where unstable As oxide compounds such as As_2O_3 are formed during oxide deposition. Once formed, the As_2O_3 will spontaneously react with GaAs [36], resulting in elemental As which acts as a metallic mid-level interface state and diminishes device performance. This reaction is also thermodynamically favourable in InAs, however elemental As has been found to accompany As_2O_x and the reaction kinetics are less clear [37] [38]. In addition, In_2O_3 (a semiconducting oxide) is substantially present at the InAs surface, creating a poorer MOS interface. On the other hand in InP, indium-based phosphates are in a lower energy state on the basis of ΔH_f and are more chemically inert than As oxides [39]. Preparation of a surface with these phosphate compounds while suppressing the native In_2O_3 and P_2O_5 components has been shown to produce improved

MOS interfaces on InP [39] [40]. Another factor which may account for a lower D_{it} interface in InP is that P dimer bonds have been shown in theoretical works to lie above the conduction band of InP, in contrast to As dimers which lie within the bandgap of GaAs and act as trap states [41]. For GaN, the only viable compounds that are not ZrO_2 are the various Zr_xN_y phases, as the N-O compounds are volatile and disperse upon formation. This suggests that these compounds may play an important role in the formation of traps at the GaN-ZrO_2 interface.

3.2. Thermodynamics of Substrate Reactions

As mentioned previously in Section 3.1, the prevalence of negative substrate reactions can greatly affect the quality of the final semiconductor-oxide interface. While examining the thermodynamic quantities of each semiconductor, it was observed that higher molar entropy values seem to correlate with higher D_{it} values in high-k oxide interfaces, as shown in Figure 3.2. This effect was observed until reaching the As-based compounds, after which the molar entropy remained relatively constant while D_{it} continued to increase. One potential explanation for this is that while all three As-based semiconductors are similar in chemical structure and hence in molar entropy, different rates of As build-up at each interface due to substrate reactions causes variation in the value of D_{it}. For instance, the significant presence of As_2O_x alongside elemental As in InAs native oxidation suggests a more limited As segregation effect than that in GaAs, and thus a lower D_{it} due to reduced metallic interface states. Conversely, the enthalpy of formation of Al_2O_3 (ΔH_f = -3.300 eV/atom) is substantially more negative than that of both In_2O_3 and Ga_2O_3, and would suggest increased reaction of As in native oxides to form highly stable Al_2O_3 at the AlAs surface, resulting in a greater concentration of elemental As and higher D_{it}. Although not yet quantified, the buildup of As deposits has been observed in studies of the AlAs surface oxide [42].

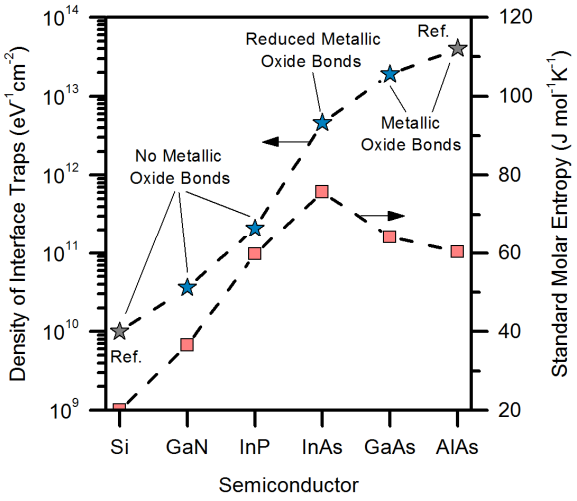
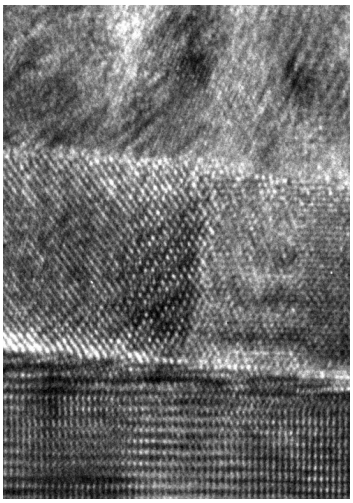


Figure 3.2. **D_{it} and molar entropy values for several semiconductors. A correlation in both quantities is noticed until reaching As-based materials.**

3.3. Improving Surface Chemistry via PEALD

It is clear from the previous discussions that the energy of the deposition reaction is a critical factor in the proliferation of native oxides on III-V semiconductors. In particular, the best insulators in the quaternary system are highly stable and will form in lower energy reactions, while the poorer compounds have less negative ΔH_f and will thus form in reactions with higher thermal energy. While several studies have investigated the use of surface treatments to passivate unwanted native metallic and oxide compounds [31] [43], these processes may suffer from the reintroduction of undesirable interface compounds due to the higher temperatures used during deposition. On the other hand, the O_2 PEALD process uses a far lower deposition temperature, bringing the thermodynamic state of the III-V substrate much close to equilibrium, room-temperature conditions. The improved nature of the PEALD oxide film is visible in previous results on GaN, where it contributed to a much lower D_{it} figure than previous approaches [31]. As seen in Figure 3.3, the TEM image of the GaN-ZrO_2 stack shows a highly ordered atomic interface, with essentially no major physical dislocations. XRD analysis of this MOS interface also indicated that the compounds present were mainly GaN and ZrO_2. In addition to keeping maintaining a lower energy interface, the ZrO_2 PEALD process applied to GaN resulted in a mainly amorphous dielectric film, reducing the effect of current leakage through the film.

a.)



b.)

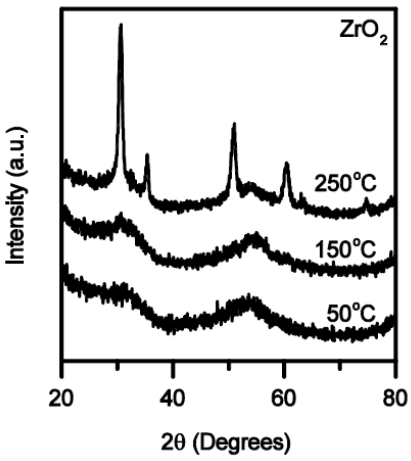


Figure 3.3. **a.) A TEM image of the ZrO_2-GaN interface, showing an ordered and periodic atomic structure, and b.) The XRD spectrum of PEALD ZrO_2-GaN interfaces at various deposition temperatures, showing a clear reduction in crystallinity. Both figures are adopted from von Hauff [27].**

Chapter 4. Characterization of MOSCAP Devices

While basic materials studies can provide valuable information about the characteristics of the MOS interface after oxide deposition, electrical characterization methods are needed to gain in-depth information about the quality of the MOS device in actual operation. These methods can gather a wide variety of data, including the density of interface traps, the effective mobility in the interfacial region, as well as the physical nature of the tunnelling current through the oxide layer. Useful characterization methods include IV, CV, and Cf profiling using an electrical analyzer in conjunction with requisite data processing. These techniques will be discussed in detail in this chapter, alongside relevant conclusions for MOS device operation.

4.1. MOSCAP Overview and Theory

The MOSCAP, or MOS capacitor, is a fundamental component in the fabrication of full MOSFET devices. Essentially acting as the MOSFET gate in isolation, the MOSCAP structure provides valuable insights into the performance of a particular gate stack on a certain semiconductor substrate. The MOSCAP consists of only a few primary components: a gate metal that serves as the "gate contact", an oxide layer, the underlying semiconductor substrate, and a "body contact" to the semiconductor bulk. The resulting layer stack is shown in Figure 4.1.

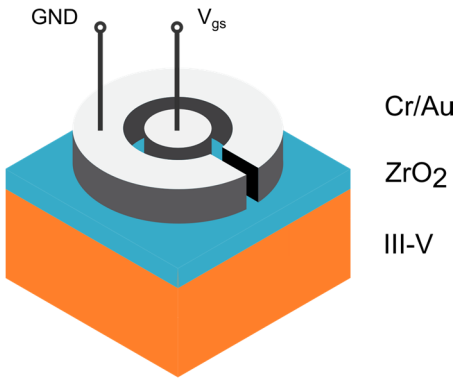


Figure 4.1. Physical layout of planar MOSCAP structures fabricated on III-V substrates, showing the layer stack of each device.

As the central modulating component of the MOSFET, the MOSCAP also allows for examination of the various regimes of charge control in the semiconductor such as inversion, depletion, and accumulation. As a result of these common phenomena, both devices share similar capacitance-

voltage profiles and can thus be studied with identical techniques. The capacitance output resulting from a sweep across the various regions of MOS device operation is shown in Figure 4.2. A wide gamut of device characteristics can be discerned from analysis of the CV output, including basic parameters such as flatband voltage as well as more complex features such as D_{it} or substrate doping density.

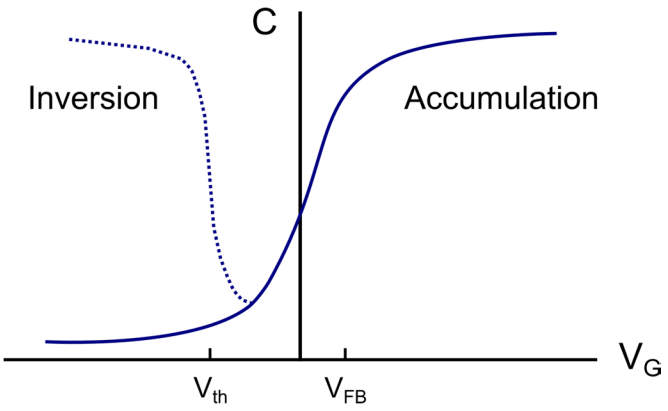


Figure 4.2. **Characteristic CV output of an n-body MOS capacitor, as fabricated in this study. The dotted line symbolizes the inversion CV peak, which is generally not seen for f > 10 kHz [44], as well as certain materials (discussed in Ch. 5). A higher quality MOS interface will exhibit a clearer and steeper transition into accumulation, as well as a higher peak capacitance (C_{ox}).**

4.2. Interface Traps and D_{it} Characterization

As discussed previously, one of the crucial issues during fabrication of MOS devices is the existence of interface traps. These traps are generated by a wide array of different sources, including crystal defects, poor bonding at the semiconductor-oxide interface, as well as the presence of impurities and metallic compounds. While the causes of these traps are fairly disparate, they share an identical method of reducing MOS device performance, forming a density of states at an energy somewhere in-between the valence and conduction bands. When the Fermi level of the gate is modulated, the traps form an intermediate stage for electrons moving towards the valence bands, and thus cause the charge to be "trapped". If the magnitude of D_{it} is sufficiently high, this results in a transistor that cannot effectively switch, with the Fermi level "pinned" to the surface.

While there are a wide variety of physical methods to characterize traps including deep-level transient spectroscopy and electron spin resonance measurements, the most accessible and preferred method to quantify interface traps is through electrical characterization. Two methods of

interest that can be achieved using an electrical parameter analyser are the forward-reverse voltage sweep (or hysteresis method) and the peak conductance method [45]. In the forward-reverse sweep method, the capacitance is measured while the voltage is swept from reverse bias to forward bias, and then back again. The subsequent shift in voltage between the two capacitance curves is indicative of the capacitive hysteresis of that particular MOS device, produced by trap charge carrier storage. The formula representing the resultant D_{it} is as follows:

$$D_{it} = \frac{C_{ox}}{qE_g}\,\Delta V \tag{3}$$

where C_{ox} is the peak oxide capacitance density, ΔV is the magnitude of hysteresis, E_g is the bandgap of the semiconductor, and q is the electronic charge. Alternatively, the more sensitive [46] conductance method can be used to measure D_{it}. This technique relies on the effect of both the capture and subsequent emission of charge by interface traps, in the form of an RC circuit with parameters G_{it} and C_{it}. The conductance is measured while the frequency is changed across a range of voltages. The approximation of the relation of probed conductance G_P to D_{it} is:

$$D_{it} = \frac{2.5}{qA}\left(\frac{G_P}{\omega}\right)_{max} \tag{4}$$

where $\omega=2\pi f$ is the angular frequency at which the peak conductance occurs. This approximation using the conductance method takes into account the broadness of the range of energies of interface traps, and can thus calculate D_{it} with higher accuracy. A more in-depth analysis of these methods is included in Schroder [45].

4.3. MOSCAP Current Analysis

An issue of major importance for the operation of MOS devices is the degree of leakage current between the gate and semiconductor substrate. As oxides scale down in thickness, the effects of gate leakage current become substantially more pronounced, and must be considered in device design due to their highly negative effect on power consumption. In highly-scaled MOS junctions, gate leakage is primarily caused by charge carriers quantum-mechanically tunnelling through the thin oxide layer. As the thickness of the oxide is reduced to less than 10 nm, the effective barrier thickness seen by these charge carriers approaches the quantum regime, increasing the probability of tunnelling dramatically.

While leakage current in thin oxide layers is primarily dominated by tunnelling phenomena, there are in fact a wide number of different physical mechanisms that enable charge carriers to "leak" from the semiconductor to the gate contact. These mechanisms are summarized through their energy band mechanisms in Figure 4.3 and are described in further detail below.

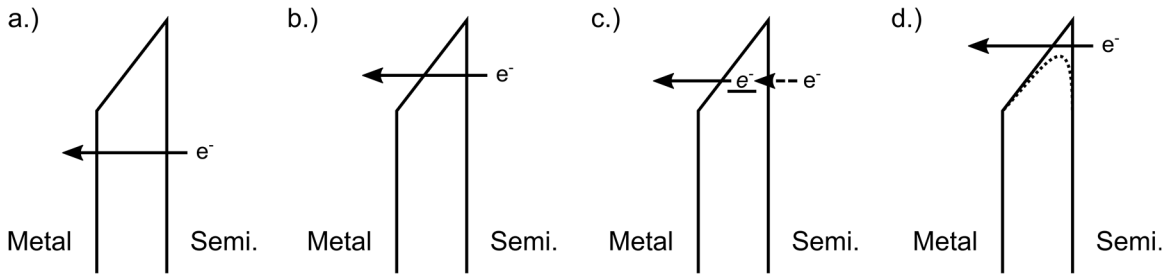


Figure 4.3. Schematic representation of the various MOS tunneling mechanisms, including a.) direct tunneling, b.) Fowler-Nordheim tunnelling, c.) Frenkel-Poole emission, and d.) Schottky emission.

In the direct tunnelling (DT) regime, charge carriers tunnel directly through the oxide energy barrier. This is most commonly seen in ultra-thin oxides of thicknesses less than 5 nm [47]. A "standard" analytic expression does not exist for DT current as tunnelling is highly dependent on the physical nature of the barrier, however assumptions can be used to obtain a simplified expression for MOS interfaces. From [48], the value of the electron DT current is given by:

$$J_T = \frac{q^3}{16\pi^2\hbar\phi_b} F_{ox}^2 \exp\left\{-\frac{4}{3}\frac{\sqrt{2m_{ox}}\phi_b^{3/2}}{\hbar q F_{ox}}\left[1-\left(1-\frac{qV_{ox}}{\phi_b}\right)^{3/2}\right]\right\} \quad (5)$$

where ħ is the reduced Planck constant, φ_b is the barrier height seen by electrons in the semiconductor (i.e. the conduction band offset between the oxide and the semiconductor), F_{ox} is the electric field on the oxide, m_{ox} is the effective mass of electrons in the oxide, and V_{ox} is the applied voltage to the oxide.

In the Fowler-Nordheim (FN) regime, the electric field on the oxide layer is high enough that the majority of carriers see a potential barrier that is triangular in shape, as opposed to rectangular. This can only occur if the applied electric field is very high, usually in excess of 10^6 Vcm^{-1}. The Fowler-Nordheim effect is a well understood phenomena, originally observed through

electrons tunnelling directly from a high-voltage metal tip to vacuum. While sometimes modified to account for different geometries, the relationship for MOS FN current [49] is given by:

$$J_{F-N} = \frac{q^3}{16\pi^2 \hbar \phi_b} F_{ox}^2 \exp\left(-\frac{\frac{4}{3}\sqrt{2m_{ox}}\phi_b^{3/2}}{\hbar q}\frac{1}{F_{ox}}\right) \tag{6}$$

where the constants are the same as those defined previously.

In the Frenkel-Poole (FP) emission regime, traps at the semiconductor-oxide interface provide lower energy states for electrons to cross the oxide potential barrier by means of thermal excitation. The relationship for FP current [50] is given by:

$$J_{F-P}\ \alpha\ E \exp\left[-\frac{q}{k_B T}\left(\phi_{trap} - \sqrt{\frac{qF_{ox}}{\pi\epsilon_{trap}}}\right)\right] \tag{7}$$

where φ_{trap} is the potential barrier height seen by electrons in the trap state and ε_{trap} is the dielectric constant of the trap.

Finally, in the Schottky emission regime, the applied field is not high enough to cause FN tunnelling, but instead lowers the barrier height seen by an electron in the semiconductor. This is caused by an image charge force on electrons inside of the semiconductor which acts in the opposite direction of the applied field. As the applied field is increased, the distribution of electrical potential shifts closer to the semiconductor, creating a higher reduction in the barrier by the image charge field which diverges at the semiconductor-oxide interface. The relation for Schottky current [50] is defined by the formula:

$$J_S = AT^2 \exp\left[-\frac{q}{k_B T}\left(\phi_B - \sqrt{\frac{qF_{ox}}{4\pi\epsilon_0}}\right)\right] \tag{8}$$

where T is the temperature of the material and A = 1.20173x10^6 $Am^{-2}K^{-2}$ is Richardson's constant. In metal electron emission studies, Richardson's constant is sometimes multiplied by a material-specific correction factor.

With the analytical relationships of the major oxide current emission modes established, it is possible to determine the physical phenomena occurring at the MOS interface over a range of

applied voltage. This can be performed using curve fitting techniques in a wide array of data analysis software suites, and can provide essential information about the quality of the interface and the gate stack itself. For instance, the presence of a direct-tunnelling relationship in the current output of a device indicates a higher-quality MOS interface when compared to a device with Frenkel-Poole tunnelling, which requires the presence of trap states to enable tunnelling action across the interface.

4.4. CV Modelling and Mobility

In addition to D_{it} quantification, CV profiling and analysis can be used to extract basic device parameters from the MOSCAP as well as more complex values such as the effective MOS mobility at the interface. One valuable device parameter that can be determined is the flatband voltage of the device, or V_{fb}. Since MOS devices are usually fabricated on doped substrates, the semiconductor bulk will have a non-zero electric potential. As a result, if the gate voltage is held at zero, an electric field will be present between the gate and body, including across the oxide. For the energy band diagram to remain continuous in all regions of device, the bands of the semiconductor must therefore bend. To flatten the bands and thus ensure a totally equal distribution of charge across the device, a voltage must be applied to the metal gate, shown in Figure 4.4. This voltage value is the flatband voltage of the device.

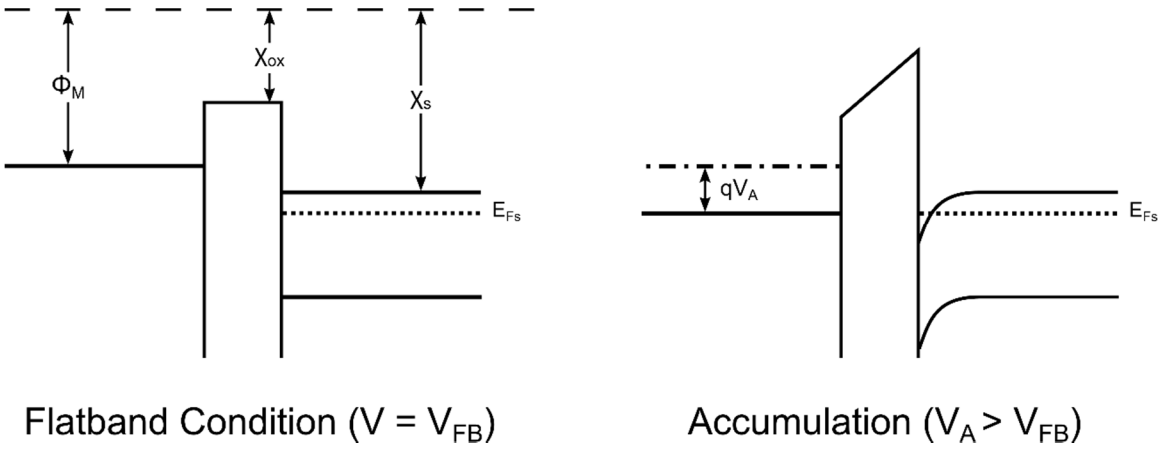


Figure 4.4. Orientation of the bands of the n-body MOS capacitor in both the flatband and accumulation device states.

The flatband voltage is an important parameter for operation of a MOS device as it represents the transition point from accumulation to depletion, and is thus a valuable reference point when interpreting CV curves. Since it is highly dependent on the gate metal used and its barrier height over the semiconductor, the flatland voltage is also useful as a parameter for gate stack workfunction engineering [51].

To begin an analysis, the theoretical flatband voltage can be calculated in a straightforward manner from basic material parameters. In the ideal case, the flatband voltage is equal to the difference in workfunctions between the metal and the semiconductor, namely:

$$V_{FB} = \Phi_M - \Phi_S = \Phi_M - \left[\chi + \frac{E_g}{2} - V_T \ln\left(\frac{N_D}{n_i}\right)\right] \tag{9}$$

where Φ_M is the metal workfunction, χ is the semiconductor electron affinity, E_g is the semiconductor bandgap, $V_T = 0.026$ V is the thermal voltage, and n_i is the intrinsic carrier density. In practice however, the metal-semiconductor barrier height is highly dependent on the characteristics of the material interface and often varies considerably from the ideal value. As such, the true flatband voltage of the device can be calculated using experimental CV data. This is accomplished by first calculating the flatband capacitance of the device under investigation, then interpolating between CV data points to find the associated flatband voltage. The flatband capacitance is found from the following relationship:

$$C_{FB} = \left(\frac{1}{C_{ox}} + \frac{L_D}{\epsilon_s}\right)^{-1} = \left(\frac{1}{C_{ox}} + \sqrt{\frac{V_T}{qN_A\epsilon_s}}\right)^{-1} \tag{10}$$

where C_{ox} is the physical parallel-plate capacitance generated by the oxide film, L_D is the Debye length of the semiconductor, and ε_s is the semiconductor permittivity.

Another quantity of interest with respect to the MOS interface is the effective MOS electron mobility μ_{MOS}. This parameter is a valuable measure of the quality of the oxide-semiconductor interface as it incorporates the various scattering mechanisms that occur at the surface of the semiconductor, and thus can directly describe the mobility of charge carriers in the channel region [52]. Like the quantities discussed previously in this section, the mobility can also be calculated through use of experimental CV data in conjunction with conductance-voltage measurements. The relation describing the effective mobility of a MOSCAP is as follows [53]:

$$\mu_{MOS} = \frac{1}{R_s Q_n} \tag{11}$$

where R_s is the sheet resistance of the MOSCAP "channel" (proportional to the measured conductance of the device) and Q_n is the mobile channel charge density which must be determined from available data. The most accurate [45] technique to calculate Q_n involves integrating the CV curve from the flatband voltage to the gate bias voltage [53]:

$$Q_n = \int_{V_{FB}}^{V_G} C dV \tag{12}$$

where V_G is the applied gate bias at which the mobility is being measured. Upon examination of the equation for μ_{MOS} however, it is clear that the presence of R_s introduces a geometric dependence, complicating analysis of non-traditional MOSCAP layouts. In addition, another major issue in completing this calculation is the selection of the CV measurement frequency. The frequency should be selected to minimize the capacitive effect of interface traps on the results, however selections may not always be optimal or even correct. One promising method for de-embedding the effects of both geometry and traps on the mobility measurements of MOS devices is through use of a comprehensive lumped element model as proposed by Bothe et al [53]. This model enables incorporation of individual components of the MOS interface in accumulation as circuit elements in an AC transmission line, allowing R_s to be fit as an independent model parameter using only the experimental conductance and capacitance measurements obtained. By fitting the full set of parameters to the complete spectrum of Cf data for each measured voltage, a modelled CV curve can be reconstructed as well, enabling a robust calculation of the MOS mobility without selection of a single measurement frequency.

4.5. Charge Centroid and Carrier Scattering

While the effective MOS mobility is a useful figure in its own right to compare between different MOS fabrication processes, it is desirable to compare the effective mobility against bulk mobilities in similar conditions to examine the magnitude of scattering occurring near the interface. To accomplish this, the relevant "doping concentration" for the effective MOS mobility must be determined for comparison to bulk scattering figures. In this sense, the "doping concentration" for MOS mobility can be thought of as the volumetric density of charge carriers near the interface

during accumulation. This density can be found by dividing the sheet charge density of accumulation electrons, Q_n, by the charge centroid or average depth position of those electrons. The charge centroid x_n is determined by treating the accumulation 2D electron gas as a triangular potential well, an approximation with analytic eigenfunctions in the form of the Airy functions, Ai(x) [54]. The closed form of these eigenfunctions is:

$$\psi(\mathrm{x}) = C\ Ai\left[\left(\frac{2m^*}{\hbar^2 q^2 E_{ox}^2}\right)^{\frac{1}{3}}(qF_{ox}x - E_n)\right] \tag{13}$$

where C is the eigenfunction normalization constant, F_{ox} is the oxide electric field in the accumulation region, and E_n are the energy eigenvalues of the solution, given by:

$$E_n = \left(\frac{\hbar^2}{2m^*}\right)^{1/3}\left[\frac{3\pi q F_{ox}}{2}\left(n - \frac{1}{4}\right)\right]^{2/3} \text{where } n = 1,2,\ldots \tag{14}$$

In this case, m^* refers to the effective mass of the semiconductor. By the postulates of quantum mechanics, the expectation value of the electron's position in the potential well can be determined through the use of the position operator, x. Before attempting this, the eigenfunctions must be normalized to obtain the value of the proportionality constant C. Following that, the charge centroid can be found by applying the position operator and integrating over all space. Since the interface is defined as x=0, the formula is as follows:

$$\langle x \rangle = \int_0^\infty \psi^*(x) x \psi(x) dx = \int_0^\infty |\psi(x)|^2 x dx \tag{15}$$

While fairly complicated, this integral is in fact analytic and can be derived with the aid of a symbolic math program. The resulting expression gives the charge centroid x_n:

$$x_n = C\left[\frac{2b^2 Ai(-ab)}{3} + \frac{2bAi'^2(-ab)}{3} - \frac{Ai(-ab)Ai'(-ab)}{3a^2}\right] \tag{16}$$

where Ai′(x) is the first-derivative of the Airy function, and a and b are placeholder variables defined as follows:

$$a = \left(\frac{2qm^*}{\hbar^2 F_{ox}}\right)^{1/3}, \qquad b = \frac{E_n}{qF_{ox}} \tag{17}$$

Crucially, these expressions contain dependence on both the fundamental material parameter m^*, as well as the oxide field F_{ox} corresponding to the onset of accumulation in a real MOS device. This enables the incorporation of experimental values for increased accuracy in estimating the centroid x_n. This approach will be used to compare the accumulation layers of the various III-V materials examined in this work against existing data for Si, and subsequently evaluate the impact of these parameters on scaled devices.

Chapter 5. Results and Materials Conclusions

N.B.: A significant portion of this chapter has been adapted from Rezazadeh et al. [33]

With the fabrication and testing methods needed to evaluate III-V MOSCAP performance established, real devices were fabricated with high-k dielectric layers. The MOSCAPs were electrically characterized using IV, CV, and Cf profiling, enabling pertinent device values to be extracted using the methods outlined in Chapter 4.

5.1. Materials Studied and Fabrication Methods

Using n-type substrates including silicon-doped (100) GaAs (N_D = 1x10^{18} cm^{-3}), silicon-doped (100) InAs (N_D = 1x10^{18} cm^{-3}), and sulfur-doped (100) InP (N_D = 1x10^{18} cm^{-3}), planar MOSCAP devices were fabricated. These III-V materials were selected on the basis of substrate availability and overall prevalence in III-V electronic devices. Before processing, wafers were individually cleaned using a two-step acetone and IPA rinse. ZrO_2 was then deposited on the aforementioned substrates using an O_2 remote plasma-enhanced process at a temperature of 100 °C in a Kurt J. Lesker ALD reactor, with a chamber pressure of 1 Torr. A film consisting of 40 cycles of oxide was deposited, alternating between tetrakis(dimethylamido)-zirconium and oxygen pulses. After oxide deposition, annular metal contacts were formed by ultrasonic acetone lift-off of RF sputtered chrome and gold films defined through contact lithography. The final fabricated device is illustrated in Figure 4.1.

5.2. In-Situ Ellipsometry

In-situ ellipsometry was performed by use of an ellipsometry module attached to the ALD reaction chamber to monitor the deposited thickness of ALD films in real-time. The thickness of the ZrO_2 film was continuously measured throughout the deposition process using a Si control piece. Using the data obtained, the thickness of the ZrO_2 as-deposited was found to be 6.87 nm, with the full deposition plot shown in Figure 5.1. As can be seen in the plot, the individual pulses of the ALD process are clear, with cyclic motion between each half-reaction. The ALD process begins with an initial O_2 plasma pulse to clean the substrate, followed by the chamber purge and

an organometallic precursor pulse, increasing the thickness. The chamber is purged yet again, then an O_2 pulse follows, causing a reduction in thickness due to the removal of ligands from the organometallic compounds following the surface reaction.

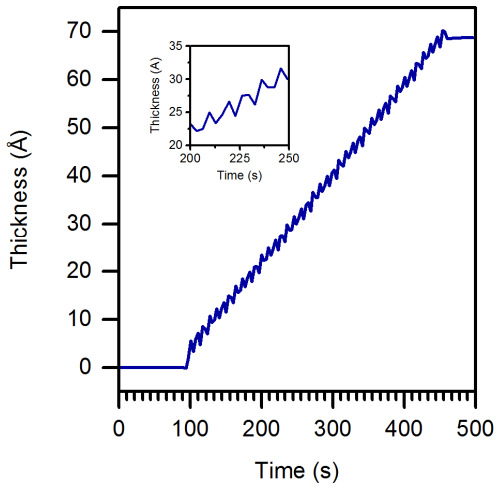


Figure 5.1. In-situ ellipsometry of oxide film thickness while performing PEALD deposition of ZrO_2. Data collected is from a Si (100) control wafer.

5.3. Electrical Characterization Data

After MOSCAPs were successfully fabricated, the various electrical measurements described were conducted using a Keithley 4200-SCS analyzer unit to extract information about the quality of the ZrO_2 films deposited.

5.3.1. D_{it} Characterization

Applying both the low-frequency capacitance-voltage hysteresis and capacitance-frequency conductivity extraction methods to the series of substrates processed, D_{it} values were obtained for GaAs, InAs, and InP (compiled in Table 5.1), with previous results on GaN included for comparison. The CV curves and hysteresis for each material are shown in Figure 5.2, while the D_{it} spectrum obtained from experimental Cf data is plotted in Figure 5.3. A side-by-side comparison of these results to prior work is found in Table 5.2.

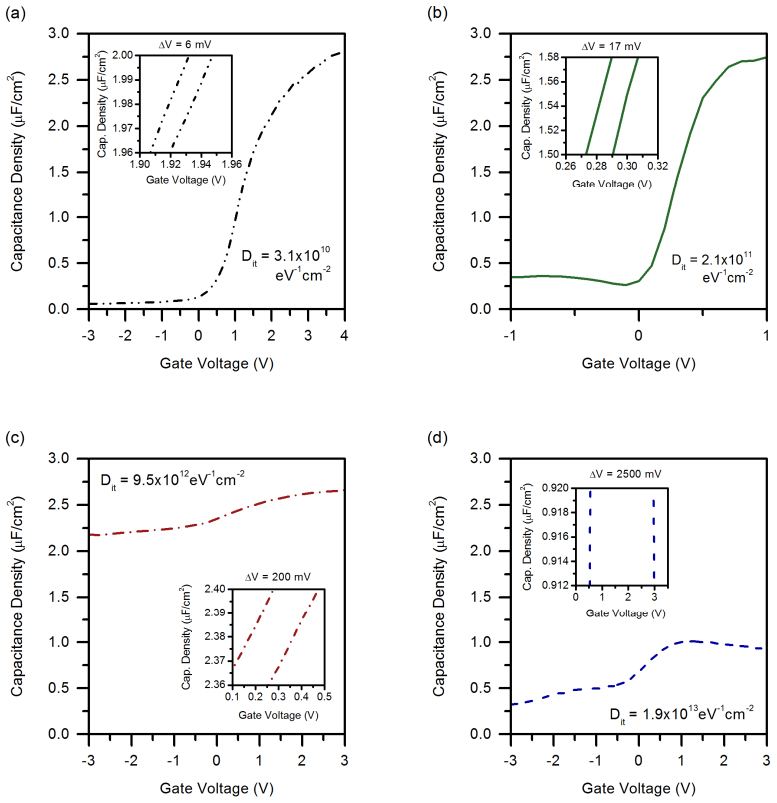


Figure 5.2. **The low frequency (10 kHz) CV characteristics for ZrO_2 MOSCAPs on different III-V semiconductors; (a) GaN, (b) InP, (c) InAs, and (d) GaAs. The corresponding density of interface traps were extracted from the forward/reverse biased CV hysteresis.**

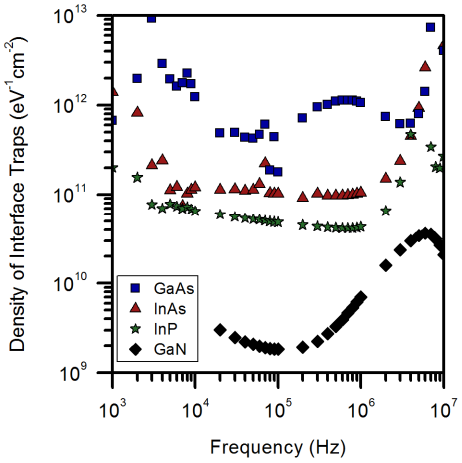


Figure 5.3. **The interface trap density for the ZrO_2 MOSCAPs obtained from the conductance-frequency spectrum.**

Table 5.1. D_{it} values of PEALD ZrO_2 on Various III-V Materials

Material	Conductance Method ($eV^{-1}cm^{-2}$)	Hysteresis Method ($eV^{-1}cm^{-2}$)	Peak Capacitance ($\mu F/cm^2$)
GaAs	$1.9x10^{13}$	$1.9x10^{13}$	1.02
InAs	$4.6x10^{12}$	$9.5x10^{12}$	2.66
InP	$4.5x10^{11}$	$2.1x10^{11}$	2.75
GaN	$3.6x10^{10}$	$3.1x10^{10}$	2.82

Table 5.2. Comparison of D_{it} to Selected Work on III-V

Material	Ref.	D_{it} ($eV^{-1}cm^{-2}$)	Peak Capacitance ($\mu F/cm^2$)
GaAs	This work	$1.9x10^{13}$	1.02
	[55] (MBE)	$5x10^{11}$	2.5
	[56]	$6.2x10^{12}$	2.3
	[57]	$2x10^{12}$	0.9
InAs	This work	$4.6x10^{12}$	2.66
	[58] (MBE)	$2.2x10^{11}$	1.2
	[59]	$6.6x10^{12}$	1.8
	[60]	$3x10^{13}$	1.6
InP	This work	$2.1x10^{11}$	2.75
	[61]	$3x10^{11}$	0.6
	[62]	$1.5x10^{12}$	2.7
	[63]	$2x10^{12}$	0.61
	[64]	$5x10^{12}$	1.4
GaN	This work	$3.1x10^{10}$	2.82
	[65]	$7x10^{10}$	0.25
	[66]	$2.2x10^{12}$	0.95
	[67]	$4.6x10^{12}$	1.05
	[68]	$1.2x10^{12}$	0.357

Of these materials, GaN showed the lowest overall D_{it} through both methods applied (on the order of 10^{10} $eV^{-1}cm^{-2}$), while GaAs showed the highest overall D_{it} (on the order of 10^{13} $eV^{-1}cm^{-2}$). By this measure, films on GaN and InP compare favorably to high-k dielectrics currently used in production Si devices, which have D_{it} figures on the order of 10^{11} $eV^{-1}cm^{-2}$ [69]. Both GaN and InP demonstrated efficient gate control using the ZrO_2 dielectric, with clear modulation of the capacitance with respect to voltage. All of the semiconductors with the exception of GaAs showed

a fairly high peak capacitance density (~2.5 $\mu F/cm^2$). With the combination of D_{it} and peak C_{ox}, these results exceed the benchmark of most previous records for high-k films on InP (~10^{12} $eV^{-1}cm^{-2}$) [61].

It is worth noting that the III-V semiconductor MOSCAPs studied did not demonstrate inversion in CV profiling, with the exception of InP which showed a small bump in the reverse bias region. In InP, GaAs, and GaN, this lack of substantial inversion is most likely due to increased recombination rates in these materials, owing to their lower intrinsic carrier concentrations compared to silicon. A similar dependence of inversion on material recombination parameters has also been demonstrated in other studies, where InGaAs n-body MOSCAPs demonstrated significant inversion owing to their lower bandgap and extremely high hole minority carrier lifetime (~10 s) [9], while similarly fabricated InP n-body MOSCAPs failed to show inversion [70]. The presence of a small inversion peak in the CVs obtained for InP MOSCAPs in this study may also suggest reduced surface recombination due to mitigation of recombination centers at a higher quality oxide-semiconductor interface. On the other hand, in InAs, the lack of inversion is most likely due to a completely "pinned" oxide-semiconductor surface preventing any significant charge modulation from occurring.

5.3.2. Current Characterization

Using IV profiling, the oxide leakage current was measured for each MOSCAP, with the results shown in Figure 5.4. While GaN/ZrO_2 offers an excellent D_{it} figure, it does not show as low of leakage current due to a smaller band offset and favorable effective mass ratios. In contrast, InP shows similarly low D_{it} but a lower leakage current because of a higher potential barrier formed by band offsets (due to a lower electron affinity). Despite this, both of these material systems show lower leakage current than InAs, which had significantly higher D_{it} and very little modulation of accumulation charge in the CVs collected.

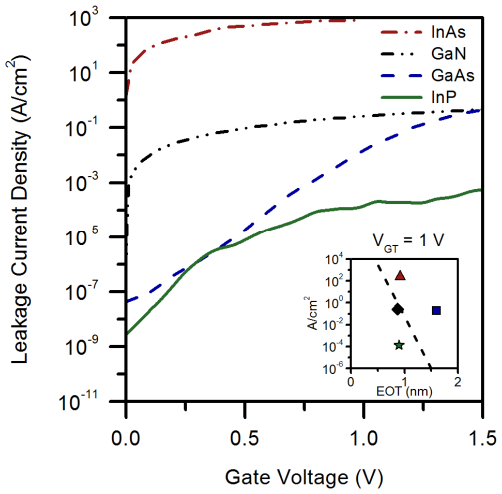


Figure 5.4. **The forward bias leakage current of the ZrO_2 MOSCAPs on InAs (red), GaN (black), GaAs (blue) and InP (green). The inset compares the MOSCAPs equivalent oxide thickness with respect to tunneling current at V_{GT} = 1 V. The dashed line represents the accepted HfO_2 EOT tunneling current observed in Si MOS.**

To analyze the current of each material system, several curve fits were performed, allowing for the relationship between electric field and current density to be determined by observing the statistically best fit. The resulting optimal fits are plotted alongside the measured data in Figure 5.5. For GaAs, it was found that Frenkel-Poole trap tunneling current matched the experimental results extremely well up to a field of approximately 1.5×10^6 V/cm, at which point Schottky emission (E = 1.5×10^6 to 1.8×10^6 V/cm) and subsequently Fowler-Nordheim emission (E greater than 2×10^6 V/cm) dominated the current output. For both InP and GaN, a direct tunneling regime was found to match the current output most closely. The presence of DT current in InP and GaN as opposed to the Frenkel-Poole emission observed in GaAs suggests a higher quality oxide-semiconductor interface in these materials, matching the observations made through D_{it} analysis. As expected, the oxide-InP barrier height in the model fit is higher than that found on GaN. As observed in all three of these material systems, the predictive power of the model fit breaks down as the field approaches zero, seen previously in other works [48]. For InAs, the experimental output was completely ohmic in both the reverse and forward directions, implying no potential barrier present for charge carriers. As a result, no attempt was made to analyze the current.

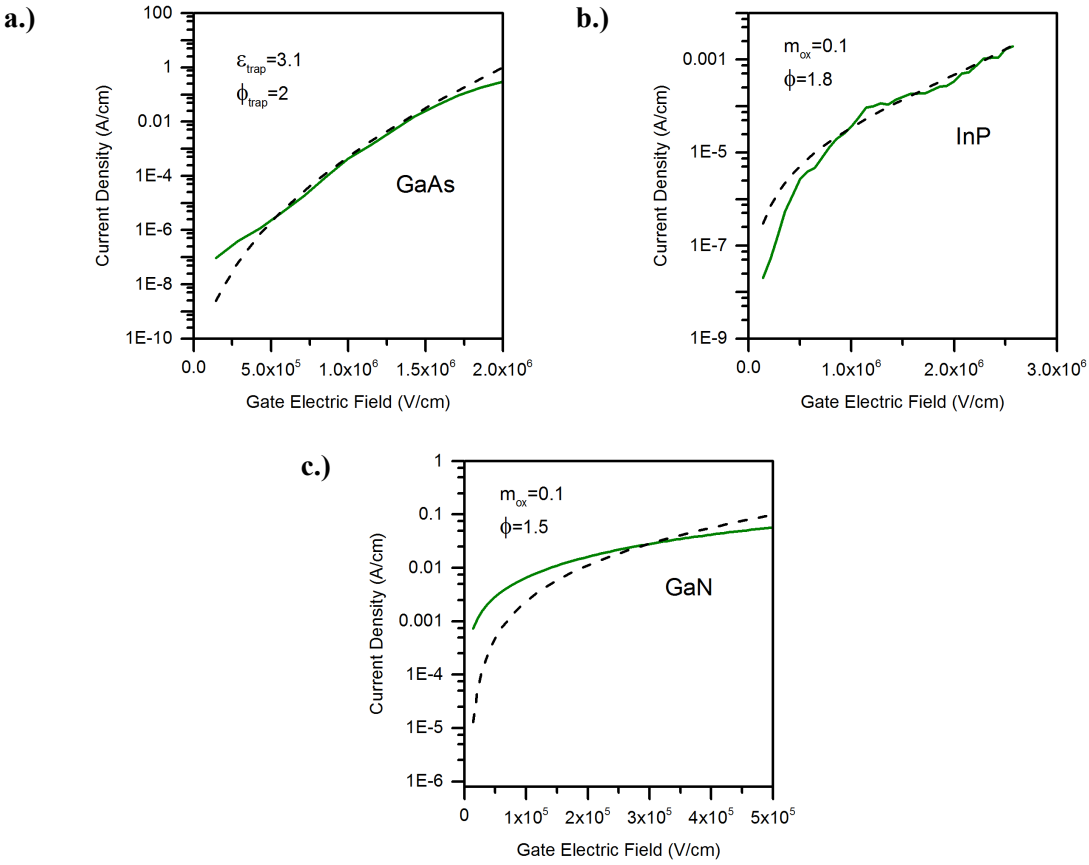


Figure 5.5. The experimental forward bias leakage current for GaAs, InP, and GaN respectively, with model fits overlaid and indicated by a dashed line. Optimal fitting parameters for Frenkel-Poole emission (GaAs) and direct tunneling (InP, GaN) are included in each plot.

5.3.3. Flatband Voltage and Mobility Extraction

Alongside the previous parameters, the flatband voltage of the MOSCAPs on each III-V material was calculated. These parameters have been tabulated and are shown in Table 5.3. For GaAs and InAs, the density of traps was too high and as a result the transition to the accumulation region of the MOSCAP was indistinct and V_{fb} could not be determined.

Table 5.3. Flatband Voltage of Fabricated MOSCAP Devices

Material	Theoretical V_{fb} (V)	Experimental V_{fb} (V)
GaN	0.46	0.67
InP	0.10	0.23
GaAs	0.42	-
InAs	-0.40	-

With these values, it is also possible to compute the effective MOS mobility of GaN and InP MOSCAPs using the transmission line model as described in Chapter 4. The model is calibrated using the various physical constants of each MOSCAP device, and the experimental Cf data is then entered to fit the relevant model parameters. The result is a frequency-independent set of capacitance-voltage data that can be used for further analysis of channel charge and mobility. A sample MOSCAP CV curve computed for InP by the Bothe model is shown below in Figure 5.6.

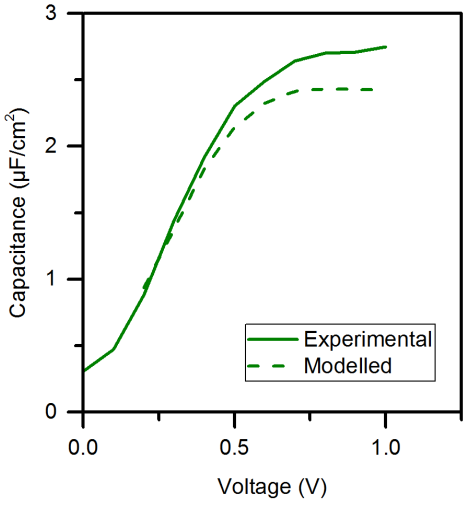


Figure 5.6. The model-computed CV curve for the ZrO_2-InP MOSCAP alongside experimentally-obtained results at 10 kHz. The model results show good agreement with the shape of the experimental CV, with an underestimation of peak C_{ox} due to the wider range of frequencies considered in the model.

The mobility for each gate voltage is then extracted by dividing the conductance by the accumulation charge integrated from V_{fb} to the applied voltage. In this case, both G_p and Q_n are found using the Bothe model. To provide comparability of the mobility figure, the voltage is substituted by the oxide electric field F_{ox}. This is accomplished by dividing the gate voltage V_g by the physical oxide thickness t_{ox}. The relationship of effective MOS electron mobility μ_{MOS} versus F_{ox} is shown in Figure 5.7 for GaN and InP MOSCAPs.

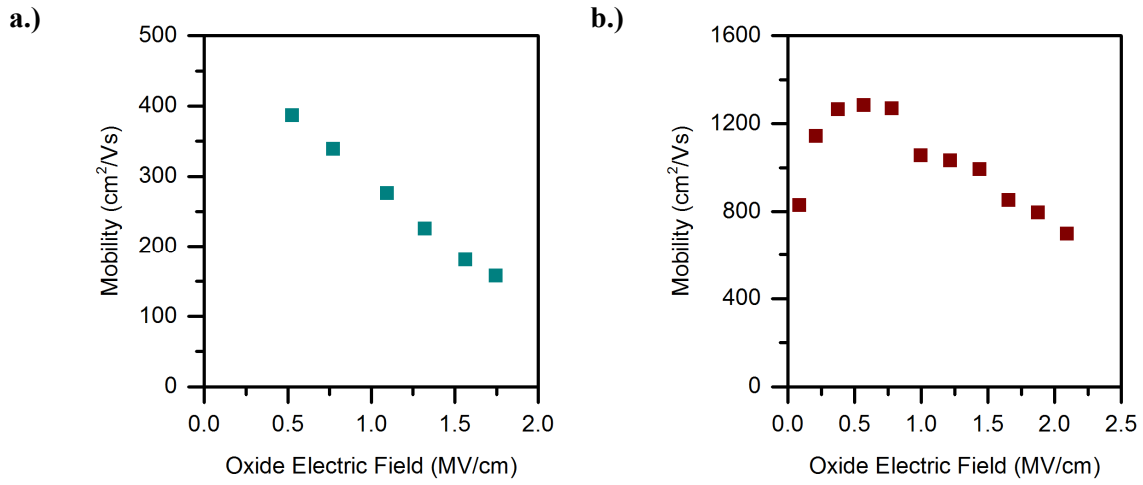


Figure 5.7. Experimental MOS mobility versus applied oxide electric field for a.) GaN and b.) InP MOSCAPs.

5.3.4. Charge Centroid Calculation

Finally, to enable comparisons to doped bulk semiconductor mobility, the volumetric charge density at the interface was computed for both GaN and InP. This was achieved by using accumulation sheet charge values (Q_n) from the Bothe model and calculated charge centroids for each semiconductor. Charge centroid values obtained via the expectation value method for all semiconductors studied are shown in Table 5.4, along with the effective masses used from [9].

Table 5.4. Accumulation Charge Centroids of Various Semiconductors

Material	Charge Centroid at Accumulation (nm)
GaN	1.05 (m^*=0.2m_0)
InP	2.29 (m^*=0.08m_0)
GaAs	2.48 (m^*=0.063m_0)
InAs	2.40 (m^*=0.023m_0)
Si [71]	0.65 (m^*=0.98m_0)

As seen in Table 5.4, the III-V compounds studied all have larger accumulation charge centroid values compared to silicon, ranging from a slight increase in GaN to a threefold increase in GaAs and InAs. As a result, the spatial extent of the accumulation layer is broader and penetrates further into the bulk of these materials [72]. This places a more stringent limit on the channel thickness or

FinFET fin width for MOSFETs using these materials, as devices with channel dimensions past a certain threshold will suffer from significantly reduced channel carrier densities.

Finally, the mobility values obtained in Section 5.2.3 for GaN and InP were plotted against the volumetric charge densities obtained for each gate voltage step in the model. The results of this analysis are shown in Figure 5.8, demonstrating the convergence of the MOS mobility to the bulk semiconductor mobility. It can thus be inferred that electron scattering is primarily due to the charge carriers in the accumulation layer and not the high-k gate stack, indicating the high quality of MOS structures using PEALD ZrO_2 films on both of these semiconductors. These mobility figures are also compared to results obtained in previous MOSCAP studies on InP and GaN bulk substrates in Table 5.5. As shown, the PEALD ZrO_2 films have higher mobility figures, suggesting improved oxide quality in combination with the D_{it} and C_{ox} results reported earlier in this chapter.

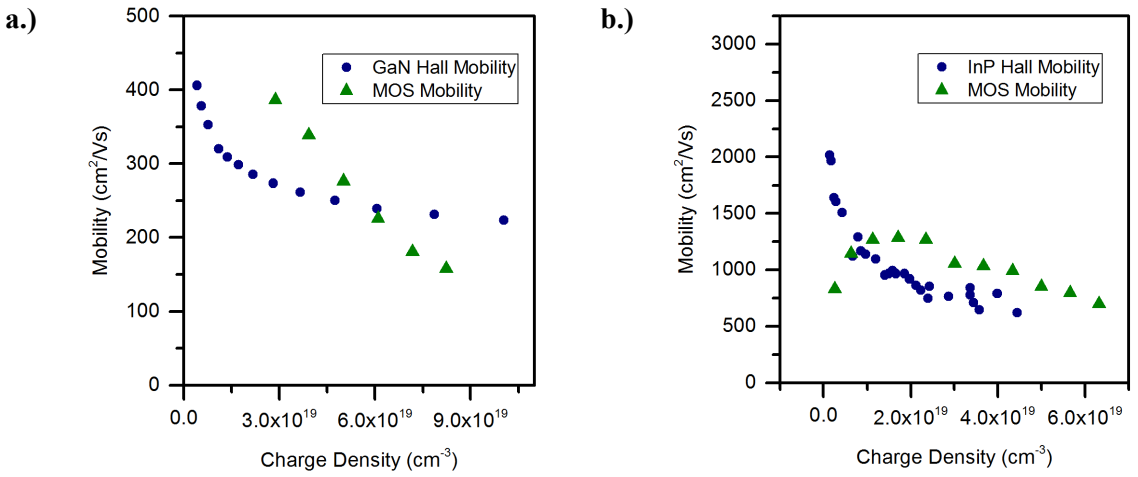


Figure 5.8. Experimental MOS mobility and accumulation charge density of GaN and InP MOSCAPs contrasted with ideal calculated Hall mobility of GaN [35] and experimentally determined Hall mobility of S-doped InP [36].

Table 5.5. Comparison of Peak MOS Mobility to Prior Work

Material	Reference	Oxide	MOS Mobility (cm^2/Vs)
InP	This work	ZrO_2	1283
	[73]	Al_2O_3	1100
	[74]	Al_2O_3/HfO_2	885
	[75]	$HfAlO_2/ZrO_2$	634
GaN	This work	ZrO_2	387
	[76]	$LaAlO_3/SiO_2$	201
	[77]	SiO_2/Al_2O_3	192
	[78]	SiO_2	167

Chapter 6. Simulations and Projections

While achieving a high-quality MOS interface is arguably the most fundamental task in fabricating an effective MOSFET device, studying the prototypical MOSCAP in isolation provides limited information of the operation of an actual MOSFET. In order to examine the hypothetical operation of a real III-V MOSFET, a series of simulation studies were undertaken using technology computer-aided design (TCAD) software. In particular, a sample InP MOSFET was constructed in Crosslight TCAD using the experimental dielectric interface characteristics obtained. This simulation was designed to analyze the performance of an entire high-mobility III-V MOSFET device using real oxide and mobility parameters for improved predictive power. Of the materials tested, InP was chosen for the simulation on the basis of its low interface trap density (compared to InAs and GaAs), as well as its beneficial properties at high frequencies. The device simulated incorporated a 50 nm thick regrown n-type $In_{0.53}Ga_{0.47}As$ source and drain ($N_D = 1x10^{20}$ cm^{-3}) to improve contact resistance in a 100 nm physical gate length structure with an L_{eff} of 90 nm, a gate potential barrier of 0.1 eV, and an EOT of 1.2 nm. While this EOT corresponds to that specified in the ITRS Roadmap for high-performance processes at the "90 nm node" [79] and not the physical L_g used, the smaller EOT value was used regardless to help mitigate the increased short-channel effects encountered in InP as described in the following sections. After simulation, a combination of electrostatic, DC, and AC device parameters were obtained through the TCAD software solver and analyzed. These results were used for a wide variety of analysis, including device optimization, comparison with Si, and to enable projections of highly scaled devices.

To simulate the various InP MOSFET devices desired, the APSYS (Advanced Physical Models of Semiconductor Devices) simulation program was used within the Crosslight TCAD software suite. Input is provided to the APSYS simulator in the form of a device layer stack with material definitions and mesh numbers, which is then discretized into a triangular mesh grid for solution of differential equations through the finite element method (FEM). To obtain the device electrical characteristics, the simulator uses Poisson's equation to obtain electrical potential and the drift-diffusion equation to obtain carrier concentrations, solving both in a self-consistent fashion. Electrical boundary conditions are defined in term of equipotential "contacts" that can be defined as either ohmic or Schottky in nature. These contacts are subsequently used to extract relevant device characteristics, including current-voltage curves as well as more complex values such as AC parameters. In addition to the base drift-diffusion equations, the APSYS simulator can capture a

wide array of ancillary physical phenomena. Of particular interest in these InP high-k MOSFET simulations are the models used for treating traps and carrier mobility within the semiconductor. Within the simulator, traps are treated as regions of charge (for Poisson's equation) with associated capture and release terms governed by Shockley-Hall-Read recombination (for the drift-diffusion equation). For mobility in InP, the simulator uses the transferred electron model [80] to account for field-induced degradation of mobility. Likewise, degradation of low-field mobility induced by impurity-scattering is accounted for through an inverse parameter-fitted model [81].

For very highly-scaled MOSFETs (especially at $L_g \precsim 25$ nm), several additional physical phenomena become important in the simulation of devices. One such phenomenon is the quantization of carrier occupancy in highly confined structures, such as the two-dimensional electron gas (2DEG) formed at the MOS interface [82] [83]. This is particularly critical in devices such as multiple-gate MOSFETs, where tight confinement in the fins of nanoscale FinFET and Tri-Gate transistors can cause considerable variation in the spatial carrier distributions of the device [84]. Another phenomenon of concern for nanoscale devices is the velocity overshoot effect or "ballisticity" of carriers in short channel length structures [85]. In devices with this effect present, the distance between the source and drain is short enough that charge carriers can traverse the channel without being scattered. As a result, the charge carriers can achieve a velocity higher than the saturation velocity of the material, leading to the "overshoot" or "ballistic" effect. While this effect can be partially captured using modified drift-diffusion simulations or other semiclassical methods [85] [86], at extremely small gate lengths (< 10 nm) it becomes necessary to use quantum methods such as the non-equilibrium Green's function [87] [88]. Although the APSYS simulator has the ability to perform both MOS quantum well and NEGF simulations, only the quantum solver was successfully implemented in these experiments. This was performed self-consistently with the standard APSYS drift-diffusion framework in simulations of FinFETs to enable examination of the carrier distributions in various iterations of these devices. While the lack of ballistic effects may slightly reduce the accuracy of these simulations, it did not affect the primary goal of demonstrating fundamental differences in SCE and operational output of devices with major structural changes. It is worth noting that the missing features have been shown to improve device characteristics, with velocity overshoot generally increasing ON-state current [85].

6.1. Device Structure and Optimization

Perhaps the most vital set of issues faced by the introduction of III-V substrates in MOSFET fabrication are the enhanced short-channel effects (SCE) encountered during device operation. Of particular note for III-V materials are the increased subthreshold leakage current and drain-induced barrier lowering (DIBL). These effects are caused by a wide variety of physical phenomena, including band-to-band tunnelling (for smaller bandgap materials), punchthrough, and gate-induced drain leakage (GIDL) [50]. At issue in the short-channel regime is the physical extent of the depletion region of the source/drain. To reduce the resistance of the S/D contacts and thus maintain a high drive current, it is desirable to have a very high doping concentration in the S/D regions. However, this design choice causes the depletion region of the source-substrate and drain-substrate junctions to extend further outwards to balance the increased charge within the S/D regions. For a small gate length MOSFET, this is problematic as the overall spatial extent of the device is reduced. In the most extreme case, this can cause the depletion region of both these junctions to coverage upon another and eventually merge, creating an exponential dependence on drain voltage in saturation and thus increasing the current dramatically. Even when full punchthrough does not occur, capacitive coupling of the depletion region to the gate enhances the forward bias of the substrate-source junction and thus exponentially increases the current by this factor [89]. The corresponding increase in voltage is given by:

$$\Delta V_c = \frac{C_{ox} + C_{dep}}{C_{ox}} \Delta V_G = \left(1 + \frac{\epsilon_s}{W_{dep} C_{ox}}\right) \Delta V_G \tag{18}$$

where ΔV_c is the change in the substrate or channel-source forward bias, and W_{dep} is the depletion region width. Critically, this increase depends on the dielectric constant of the semiconductor, implying that semiconductors with higher dielectric constants than Si (including InP) will suffer from increased SCE, as confirmed in previous studies [90]. The reduced band-edge density of states in these semiconductors may also contribute to reduced drive current and increased DIBL in strong inversion at small length scales, owing to a DOS "bottleneck" in the 2D electron gas at the channel region [90] [91].

Alongside subthreshold leakage current and DIBL, several other short-channel effects exist that may impede operation of the III-V MOSFET device. These are mainly due to the high electric fields encountered in the channel region of devices with short gate lengths and correspondingly

thin oxides [50]. One such effect is the degradation of carrier mobility in the channel owing to the high vertical or transverse electric fields on the gate oxide. With a thinner oxide layer, the increased electric field over the channel pulls charge carriers closer to the surface, and as a result their wavefunctions penetrate further into the oxide and are scattered more heavily. In addition to this, other, more destructive field effects can occur such as hot carrier injection, where high-energy electrons accelerated by the electric field collide with the oxide or semiconductor lattice. Likewise, the horizontal or lateral electric field component causes similar issues in the form of impact ionization in the semiconductor bulk. These effects can degrade the reliability of short-channel MOSFET devices, and may lead to critical failures such as time-dependant destructive breakdown of the gate oxide layer.

Although there has been a flurry of research into III-V MOSFETs in recent years, there has been a relative sparseness in studies of device structure optimization to minimize the short-channel effects described above. This has been influenced by both the primary experimental focus of demonstrating oxide-semiconductor compatibility and integration of quantum-well layer stacks in most new III-V MOSFETs, resulting in the adoption of standard HEMT-like device structures with minimal S/D engineering [92] [93]. As a result, there are significant opportunities for studies of the variation of key device structural parameters to mitigate SCE, including novel S/D and substrate geometries as well as modification of doping profiles in the individual regions. The effects of these variations on device performance will be studied in this section.

6.1.1. Source/Drain Optimization

One of the major advantages of the epitaxially regrown source/drain is the ability to create extremely shallow metallurgical junctions between the S/D and substrate (usually in the range of 5-10 nm). The shallow junction significantly reduces the extent of the depletion region in the channel, and as a result inhibits short-channel effects. While the usual design trade-off for implementing shallow S/D regions is a much higher contact resistance, epitaxial regrowth allows the S/D to keep its optimal size by extending it above the substrate in a “raised” fashion, as shown in Figure 6.1. This ensures that the low contact resistance is maintained, while still receiving the benefit of a shallow metallurgical junction. While the raised epitaxial S/D is one of the few process optimizations in widespread use in the fabrication of III-V MOSFETs, aspects of the structure may be sub-optimal. In particular, a rectangular source-drain geometry is used which forces the gate

oxide to grow along the sidewalls of the regrown layers to overlap the S/D regions. This alters the electric field patterns at these points, and may potentially reduce the gate electrostatic control of carriers induced into the channel.

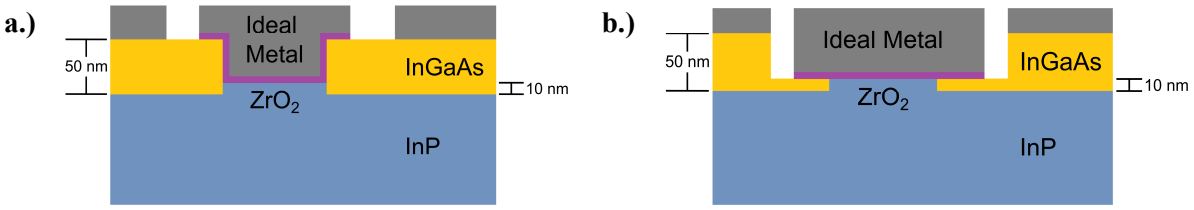


Figure 6.1. Schematic and layer stack of the simulated structures with a.) no planarization etch applied to the gate, and b.) total planarization of the gate and source/drain.

To examine the effects of varied source-drain regrowth geometries on device performance, a series of several structures were defined in the TCAD software. Between each structure, a source-drain recess etch of different depth was applied, aiming to show the effects of increased gate planarization. As the gate and source-drain regions became more planar with increasing etch depth, there was a noticeable increase in both cutoff frequency and drain current. This effect is likely due to an increase in channel control and inversion charge transport efficiency between the channel and source-drain. A fully-recessed source-drain was found to produce the most optimal results, with an ON current of 2.95×10^3 $mAmm^{-1}$ at V_{ds} = 3 V and a peak cut-off frequency of 183 GHz. On the other hand, the non-planarized approach utilized often with regrown InGaAs source-drains [93] showed decreased performance, with an ON current of 2.08×10^3 $mAmm^{-1}$ at V_{ds} = 3 V and a cut-off frequency of 85.7 GHz. The schematic of the optimal S/D geometry and the effect of this etch depth on drive current are both illustrated graphically in Figure 6.2. Likewise, the electric field profiles for both the non-planarized and totally planarized devices are shown in Figure 6.3. It is readily apparent that the electric field is more concentrated in the channel of the planarized device, confirming the increased carrier transport in these structures. As a result of the higher field magnitude in the channel region, these devices have higher drift current and increased channel transconductance (producing a higher f_T).

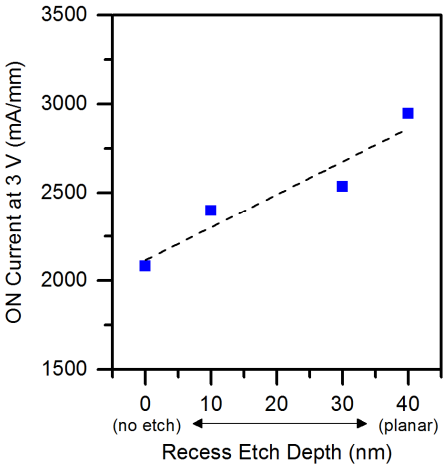


Figure 6.2. **Peak ON currents of simulated high-V_{th} InP MOSFETs with varying gate overlap recess etch depths in the source-drain regions.**

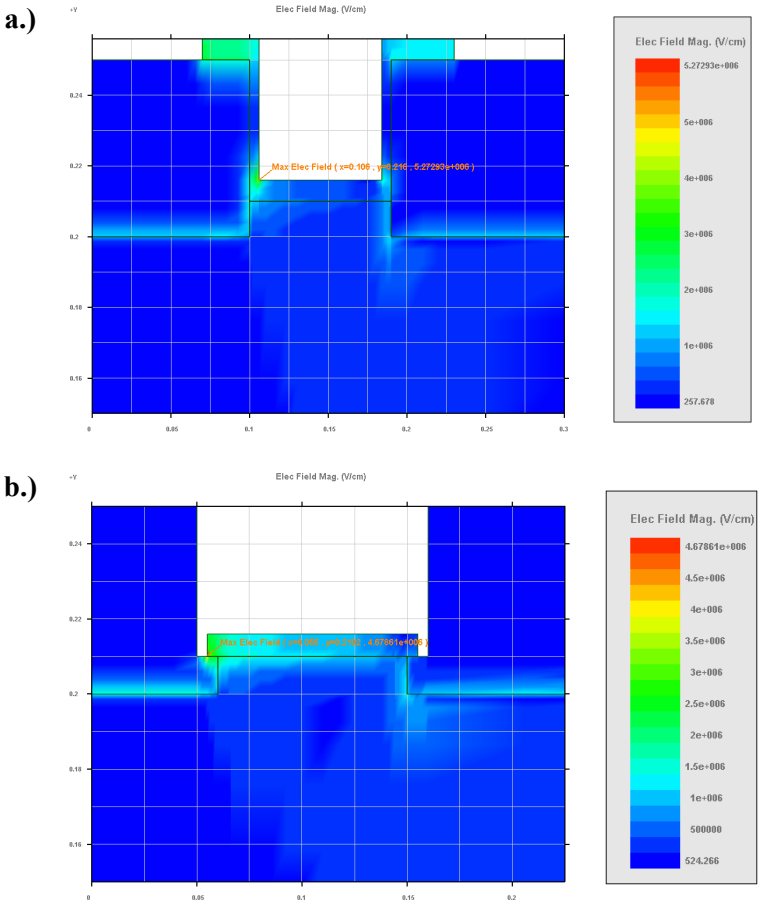


Figure 6.3. **Captured field magnitude profiles of the simulated InP MOSFET with both a.) no planarization etch, and b.) total planarization.**

6.1.2. Substrate Optimization

Perhaps most importantly in a scaled device however, the semiconductor substrate doping must be optimized for a given gate-length to ensure reduced SCE on the operation of the MOSFET device. In particular, the subthreshold characteristics such as subthreshold slope and leakage must be kept within acceptable limits. This is especially of importance in logic devices, where an excellent ON/OFF ratio with high swing between states is needed to ensure proper operation of digital circuits. As described in the preface of Section 6.1, the SCE that degrade such performance are mainly due to the effects of the S/D-bulk depletion region, and as such are directly controlled by the substrate doping. By increasing the substrate doping, the spatial extent of the depletion region can be reduced significantly, potentially unlocking significant improvements in MOSFET subthreshold characteristics. In III-V materials, this can also be accomplished through use of a highly doped epitaxial surface layer, enabling the advantages of semi-insulating substrates while mitigating potential SCE. To quantify these enhancements, the substrate doping was varied from $N_A = 1x10^{15}$ to $1x10^{18}$ cm^{-3} using the 100 nm L_g template, with previous results on semi-insulating InP ($N_A = 1x10^{10}$ cm^{-3}) included for comparison. For each case, V_{gs} was swept from 0 to 3 V, with V_{ds} held at 0.01 V. To calculate DIBL, the V_{th} was measured using the constant-current method at both low and high bias. In cases where the low-bias current was too high, the voltage value was extrapolated from the bottom of the curve. The resulting subthreshold current curves are shown in Figure 6.4, with extracted device data compiled in Table 6.1. From $N_A = 1x10^{10}$ to $1x10^{16}$ cm^{-3}, the subthreshold slopes (SS) and ON/OFF ratios are far too low for practical circuit applications. With $N_A = 1x10^{17}$ cm^{-3}, the ON/OFF ratio increases dramatically and the SS approaches ideal values. This is however accompanied by a reduction of drive current at both 1 and 3 V owing to an increase in threshold voltage, with the change especially significant at $N_A = 1x10^{18}$ cm^{-3}.

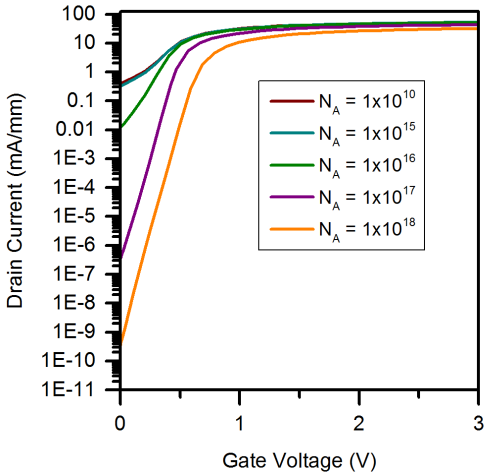


Figure 6.4. **I_{ds}-V_{gs} curves of the 100 nm InP MOSFET with a range of different substrate doping levels.**

Table 6.1. **Analysis of Semiconductor Body Doping Variation at L_g = 100 nm**

N_A (cm^{-3})	SS (mV/dec)	DIBL (mV/V)	ON/OFF Ratio at 1 V	I_{ON} at V_{DS} = 1 V (mA/mm)	I_{ON} at V_{DS} = 3 V (mA/mm)
$1x10^{10}$	483	$1.45x10^{3}$	13.9	882	$3.47x10^{3}$
$1x10^{15}$	442	$1.40x10^{3}$	14.5	877	$3.46x10^{3}$
$1x10^{16}$	183	$1.03x10^{3}$	25.1	842	$3.43x10^{3}$
$1x10^{17}$	70.3	133	$5.59x10^{5}$	676	$3.20x10^{3}$
$1x10^{18}$	65.3	25.5	$1.29x10^{11}$	218.3	$2.61x10^{3}$

Likewise, a similar set of variation experiments can be applied to the thickness of semiconductor bulk material. Reduction of the semiconductor body thickness is desirable for a number of different reasons, primarily due to the resulting reduction in parasitic diffusion capacitance and leakage current between the S/D and body. As the doping concentration of smaller gate-length devices is increased to counteract increased SCE, the S/D-body capacitance increases proportionally, leading to higher dynamic power consumption. As mentioned in Chapter 1, the semiconductor body can be reduced in size by the use of semiconductor-on-insulator technology, or through 3D transistor geometries such as FinFETs. In the case of a semiconductor-on-insulator (XOI) device, two possible implementations are possible, namely partially-depleted XOI or fully-depleted XOI [94]. In PD-XOI, the body thickness is large enough that the channel region is not entirely depleted, leading to certain drawbacks. Conversely, in FD-XOI, the entire thickness of semiconductor (including the channel) is depleted through. To examine the characteristics of varied InP body thicknesses in a single-gate planar configuration, the standard simulation procedure was

run with InP body thicknesses (maintaining the standard S/D-channel thickness of 10 nm) of 210 nm, 110 nm, 60 nm, 35 nm, and 10 nm. These iterations were performed using N_A = 1x10^{10} and 1x10^{17} at a gate length of 75 nm. The results are shown in Figure 6.5 and Table 6.2. As the body thickness approaches 10 nm, there is a substantial increase in the ON/OFF ratio, and the SS comes extremely close to its ideal value of 60 mV/dec, accompanied once again by a corresponding decrease in drive current. In addition, the drive currents converge for both substrate doping levels studied as the substrate thickness is reduced.

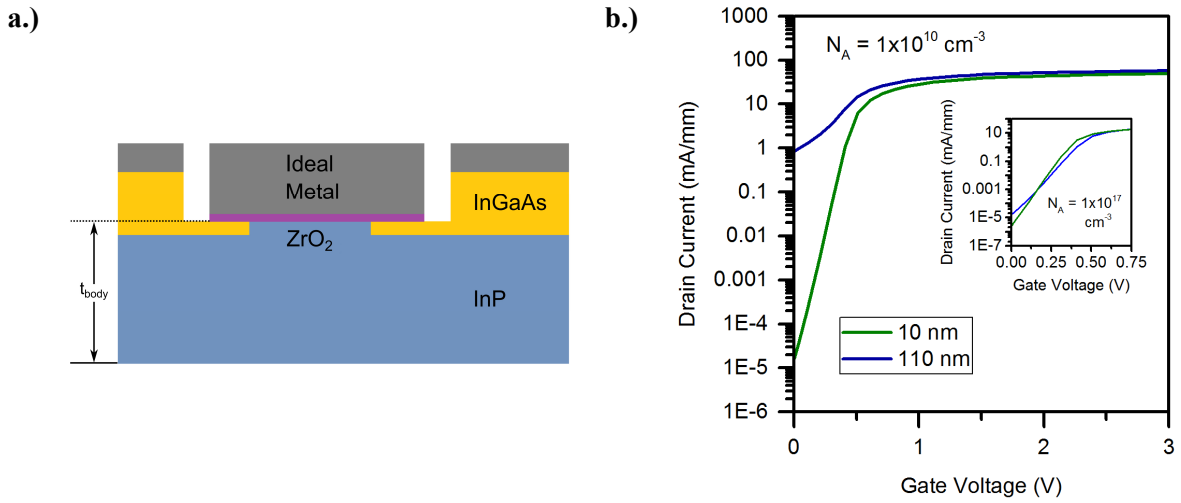


Figure 6.5. **a.) Schematic indicating the semiconductor body thickness, b.) I_{ds}-V_{gs} curves of the 75 nm L_g InP MOSFET with 10 nm and 110 nm t_{body} at both N_A = 1x10^{10} cm^{-3} and 1x10^{17} cm^{-3} (shown in inset).**

Table 6.2. **Analysis of Semiconductor Body Thickness Variation at L_g = 75 nm**

t_{body} (nm)	N_A (cm^{-3})	SS (mV/dec)	DIBL (mV/V)	ON/OFF Ratio at 1 V	I_{ON} at V_{DS} = 1 V (mA/mm)	I_{ON} at V_{DS} = 3 V (mA/mm)
210 nm	1x10^{10}	611	1160	9.96	1.06x10^3	3.81x10^3
	1x10^{17}	77.2	184.1	3.08x10^4	816	3.53x10^3
110 nm	1x10^{10}	602	1045	11.4	1.04x10^3	3.84x10^3
	1x10^{17}	86.0	307.3	1.92x10^3	859	3.56x10^3
60 nm	1x10^{10}	301	872.4	17.7	1.01x10^3	3.65x10^3
	1x10^{17}	142	520.4	95.2	918	3.56x10^3
35 nm	1x10^{10}	133	483.9	74.9	973	3.55x10^3
	1x10^{17}	122	386.3	305	922	3.50x10^3
10 nm	1x10^{10}	62.3	83.98	4.15x10^7	752	2.85x10^3
	1x10^{17}	62.9	84.11	5.76x10^7	735	2.85x10^3

6.2. Scaled Devices and Projections

Since MOSFET device characteristics can be improved substantially through the use of gate length scaling, it is important to examine the effects of a corresponding reduction in dimensions. Beginning with the 100 nm L_g device template and $N_A = 1x10^{17}$ cm^{-3}, InP MOSFET simulations were performed at L_g = 75 nm, 44 nm, and 25 nm. All lateral dimensions with the exception of the S/D overlap and gap (both maintained at 5 nm) were reduced by the corresponding gate length scaling factor L_g'/L_g, including the lateral S/D length and oxide gate length. The oxide thickness was scaled in predefined increments between "process nodes", using a 4.5 nm t_{ox} for L_g = 75 nm, a 3.0 nm t_{ox} for L_g = 44 nm, and a 1.7 nm t_{ox} for L_g = 25 nm. Both I_{ds}-V_{ds} and I_{ds}-V_{gs} curve families were obtained for each device. Follow-up simulations at 25 nm were performed using $N_A = 1x10^{18}$ and $5x10^{18}$ to determine the ability of device parameter changes to mitigate the extreme SCE encountered. The results are shown in Figure 6.6 and Table 6.3. It should be noted that the V_{th} roll-off curve shown in Figure 6.6 is most likely steeper than the actual V_{th} roll-off, owing to the lack of data points near the curve corner point as seen in other studies [50] [95]. Taking this into consideration, the devices show good scaling characteristics overall with acceptable V_{th} roll-off at fixed N_A until an L_g of around 55 nm. As seen in Table 6.3, the drive current increases in a linear fashion with gate length, however at 25 nm the 3 V drive current saturates at a significantly lower fixed value. This is demonstrated further in the I_{ds}-V_{ds} curves of the 25 nm devices (plotted in Figure 6.6 for $N_A = 5x10^{18}$), which show that V_{gs} values of 2 V and higher have essentially no effect on the current output of the device. This effect is most likely due to the saturation of the relatively small DOS at the short L_g MOS interface. Since the confining electric field increases in tandem with device dimensions, the depth of the potential well increases substantially in very highly scaled devices, reducing the spatial extent of the 2DEG underneath the oxide-semiconductor interface. However, since the spatial distribution of carrier states is further spread out in InP compared to Si, more of the electrons lie outside of the well, resulting in a limited density of states inside. As the bias is increased further, the well becomes narrow to the point that it captures all of the electrons in the limited states of the well, resulting in a constrained channel capacity for charge carriers and capping the drain current at a fixed value. Additionally, even a fairly high substrate doping of $5x10^{18}$ cm^{-3} is unable to correct the SCE found in the 25 nm device.

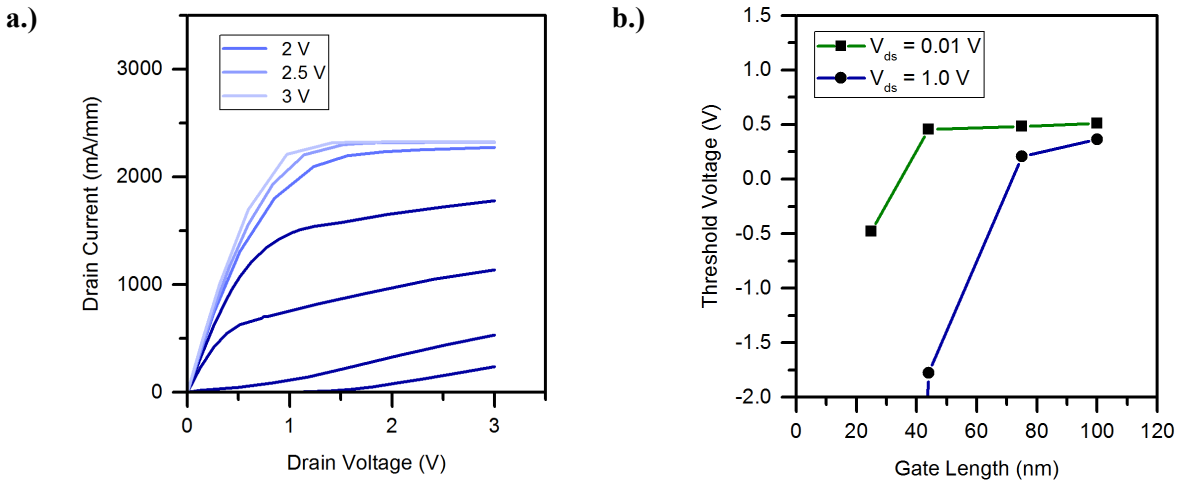


Figure 6.6. **a.) I_{ds}-V_{ds} characteristics of the 25 nm planar InP MOSFET with N_A = 5x10^{18} showing current ceiling effect, b.) V_{th} rolloff of L_g scaled devices.**

Table 6.3. **Analysis of MOSFET Characteristics with L_g Scaling**

L_g (nm)	SS (mV/dec)	DIBL (mV/V)	ON/OFF Ratio at 1 V	I_{ON} at V_{DS} = 1 V (mA/mm)	I_{ON} at V_{DS} = 3 V (mA/mm)
100	70.3	133	5.59x10^5	676	3.20x10^3
75	86.0	306	2.39x10^3	847	3.63x10^3
44	183	1.74x10^3	10.2	1.17x10^3	4.14x10^3
25 nm (N_A=1x10^{17})	1.32x10^3	6.73x10^3	2.25	1.59x10^3	2.33x10^3
25 nm (N_A=1x10^{18})	308	2.03x10^3	5.92	1.23x10^3	2.32x10^3
25 nm (N_A=5x10^{18})	120	426	331	755	2.33x10^3

In a similar fashion to Si MOSFETs [96], the use of non-traditional device structures must be used to counteract the "scaling wall" of SCE encountered at extremely small L_g. As mentioned previously, one such alternative is the dual-gate or FinFET device structure. This design has been previously investigated by several groups for use in III-V MOSFETs [7] [97], owing to its improved performance over planar devices. In the FinFET process, "fins" of semiconductor material are formed, usually by means of multiple lithographic patterning and etching. A gate oxide is then blanket deposited over the fin, enabling the gate to effective surround the fin on two sides of the semiconductor. This "multiple gate" structure enhances the electrostatic action of the gate over the channel, and as a result produces higher drive current along with the associated benefits of a smaller

semiconductor body. To examine the effects of a FinFET-style device structure on the prototype InP MOSFET, a "fin" body was formed and an additional gate was added to the backside of the L_g = 25 nm simulation device, using a combined S/D thickness of 30 nm. This device was simulated for fin widths of 20 nm, 15 nm, and 10 nm. I_{ds}-V_{ds} characteristics for the 20 nm w_{fin} device are shown in Figure 6.7, alongside plots of the channel carrier distribution in Figure 6.8 as well as various extracted device parameters in Table 6.4. Predictably, the FinFET structure essentially doubles the drive current at 1 V owing to the addition of a second gate in the device. However, devices with fin widths of 20 nm and 15 nm retain the current cap phenomena encountered in planar devices, and more critically show no improvement in SCE. Another observation of note is the apparent change in the shape of the I_{ds}-V_{ds} curves for the 20 nm width device from 1 V onwards. For values less than 1 V, the curves exhibit higher channel length modulation, whereas above 1 V the curves are substantially flatter. In contrast, reduction of the fin width to 10 nm results in a substantial improvement of most device characteristics, including a significant enhancement of subthreshold characteristics. As seen in the charge carrier plot, the 10 nm w_{fin} device shows a peak electron concentration nearly 1.5x higher than the 20 nm w_{fin} device, owing to improved electrostatic control of charge. In addition, the electron concentration in the fin bulk is increased, most likely due to a reduction of the surface well depth. These results indicate an optimal ratio of fin width to gate length somewhere around 0.4-0.5L_g, slightly less than that indicated by Si device design rules which suggest a fin width of at least 0.6L_g for optimal device performance [98].

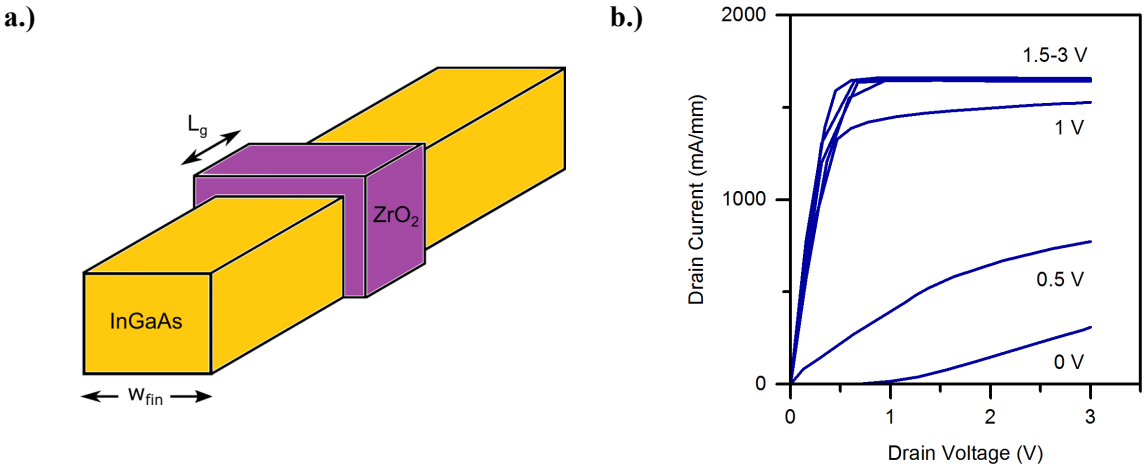
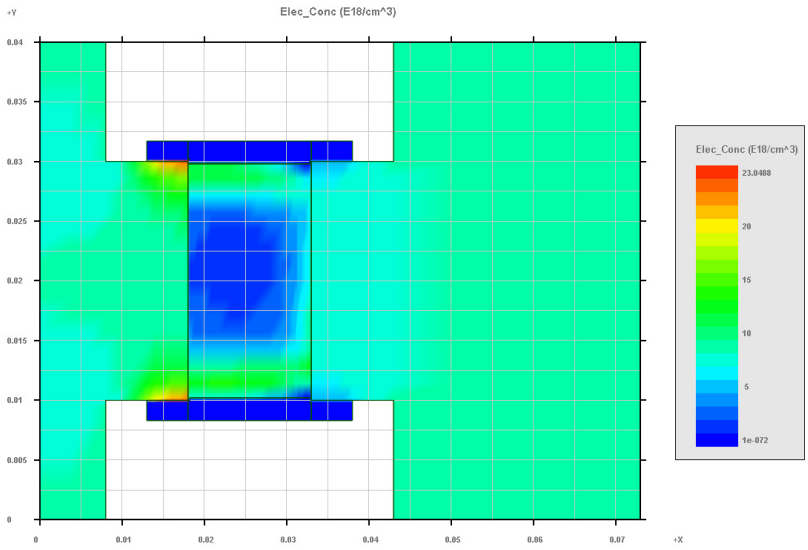


Figure 6.7. **a.) Side profile of the FinFET device showing the critical parameters of the device. Not seen in this schematic are the InP channel region and top hardmask, both hidden by the gate oxide. b.) I_{ds}-V_{ds} characteristics of the 20 nm w_{fin} InP FinFET, showing retention of the current cap phenomenon.**

a.)



b.)

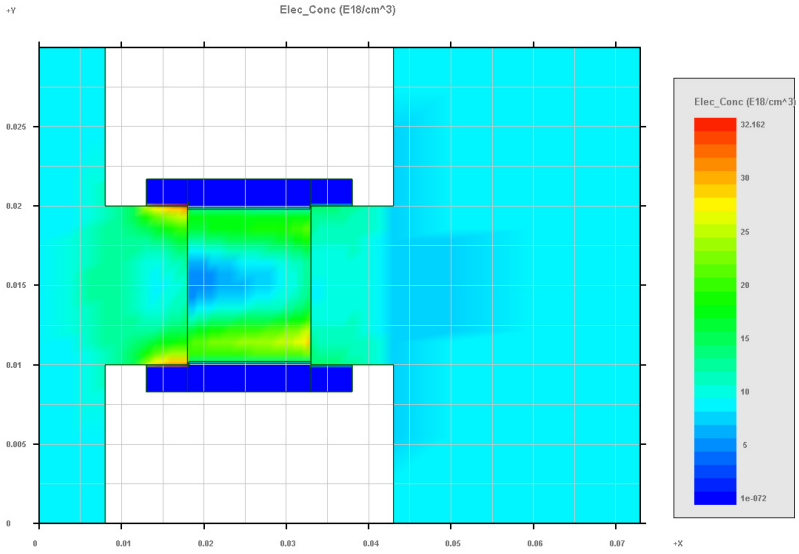


Figure 6.8. **Channel electron concentrations of the simulated for L_g = 20 nm FinFET device at V_{ds} = 3 V with a.) w_{fin} = 20 nm, and b.) w_{fin} = 10 nm. The apparent lack of symmetry is due to error introduced by meshing.**

Table 6.4. Analysis of FinFET Fin Width Scaling

w_{fin} (nm)	SS (mV/dec)	DIBL (mV/V)	ON/OFF Ratio at 1 V	I_{ON} at V_{DS} = 1 V (mA/mm)	I_{ON} at V_{DS} = 3 V (mA/mm)
20	147	535	90.8	$1.45x10^3$	$1.64x10^3$
15	104	301	$3.51x10^3$	$1.17x10^3$	$1.64x10^3$
10	76.9	122	$2.10x10^5$	704	$1.54x10^3$

It is important to note that the procedure to find the total current output of a FinFET device is slightly more involved than for a planar device. In this case, the simulated device current (in mA/mm) must be multiplied by both the height of the fin and the number of fins in the given device.

Previous studies of FinFET devices have also shown that these structures can benefit significantly from a source-drain underlap [99], as opposed to the customary S/D overlap found in planar MOSFETs. These works have demonstrated reduction of SCE, leading to lower SS and DIBL in underlapped FinFETs. In addition, reduction or elimination of the gate overlap can dramatically reduce the gate-source and gate-drain capacitances, potentially enabling extremely high cutoff frequencies for small L_g devices. Alongside the benefits of improved performance, use of an underlapped structure reduces the dimensional requirements of lithography to define the fin channel. To study the effects of these corresponding changes in the gate structure of InP MOS devices, 15 nm L_g FinFETs with fin widths of 10 nm were created using gate-S/D spacings (L_{ext}) of -5 nm (overlapped), 0 nm, and 5 nm (underlapped). The results of these simulations are shown in Figure 6.9 and Table 6.5, with the channel carrier distributions shown in Figure 6.10. This data confirms the anticipated benefits of the increased underlap, with the L_{ext} = 5 nm device showing minimal DIBL and a reduction of the SS to near-ideal values. In addition, the cutoff frequency of the device increases substantially as the S/D overlap is reduced, reaching a high of 557 GHz for L_{ext} = 5 nm. One observation of note is the low cutoff frequency observed in the overlapped 15 nm L_g FinFET, which was substantially lower than that of the 25 nm L_g planar device.

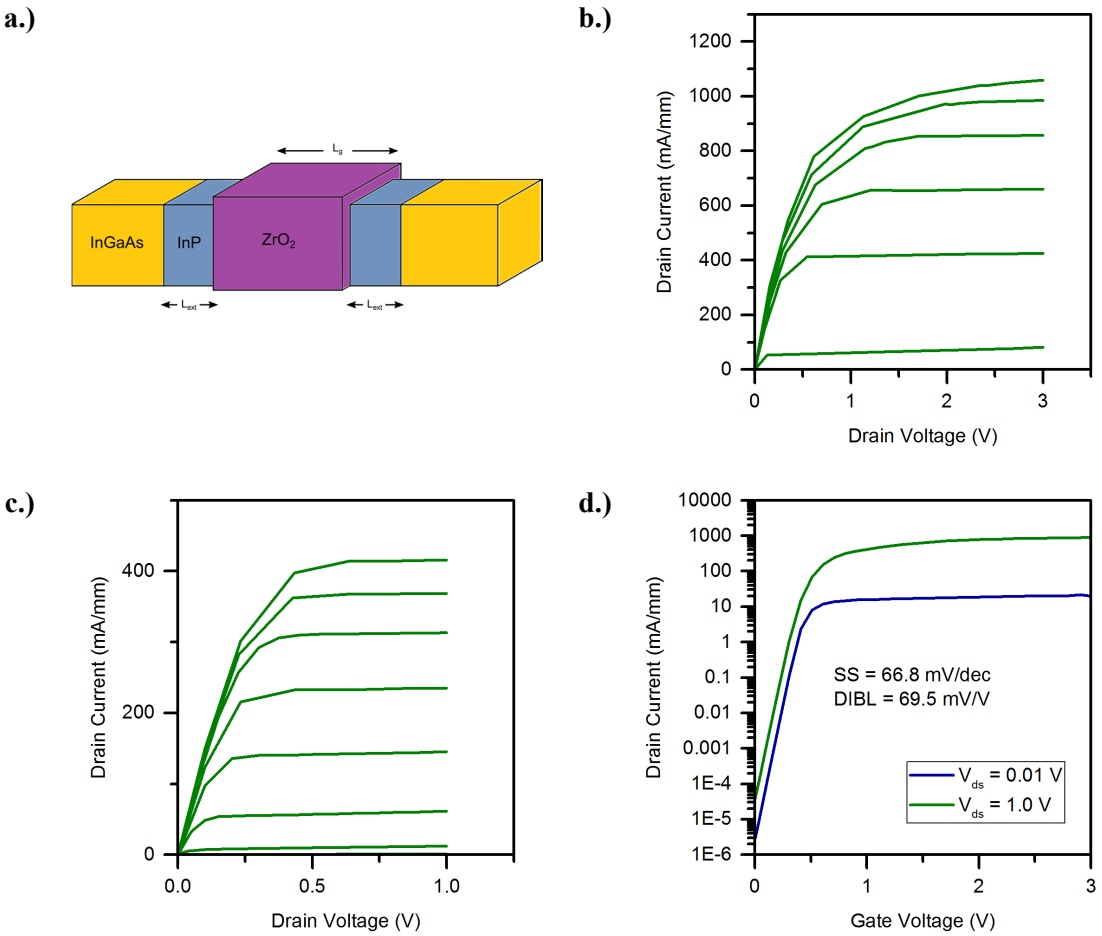
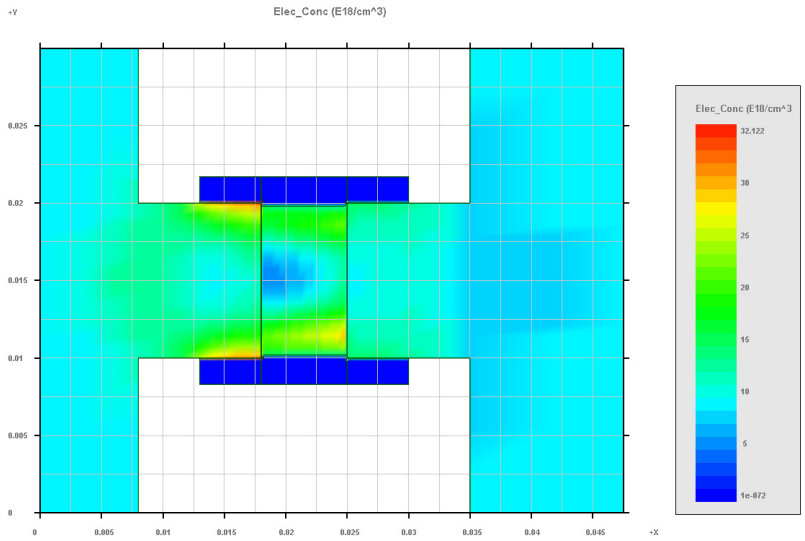


Figure 6.9. **a.) Schematic showing the device layout and indicating L_{ext} of an underlapped InP FinFET, b.) I_{ds}-V_{ds} characteristics of the L_{ext} = 5 nm FinFET device with V_{gs} increased in 0.5 V steps from 0-3 V, c.) I_{ds}-V_{ds} characteristics of the same device with V_{gs} increased in 0.1 V steps from 0-1 V, d.) I_{ds}-V_{gs} and subthreshold characteristics of the same device.**

a.)



b.)

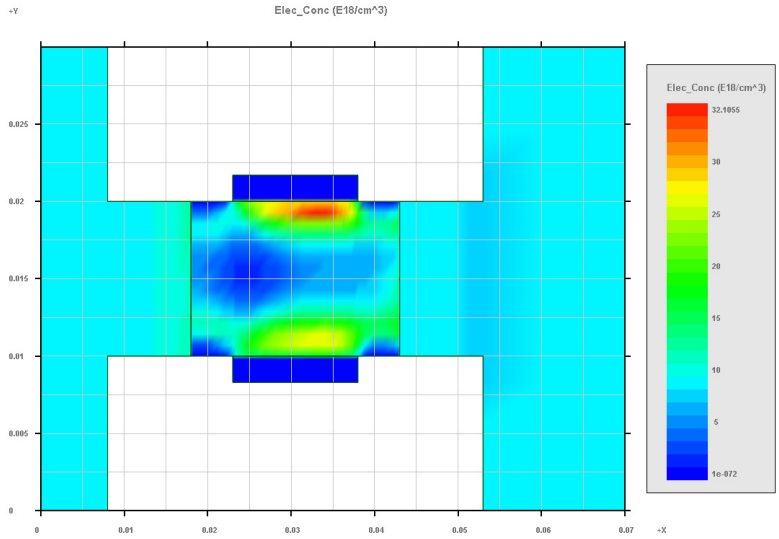


Figure 6.10. **Channel electron concentrations of the simulated L_g = 15 nm FinFET device at V_{ds} = 3 V for a.) L_{ext} = -5 nm, and b.) L_{ext} = 5 nm.**

Table 6.5. **Analysis of FinFET Source-Drain Spacing Variation**

L_{ext} (nm)	SS (mV/dec)	DIBL (mV/V)	ON/OFF Ratio at 1 V	I_{ON} at V_{DS} = 1 V (mA/mm)	I_{ON} at V_{DS} = 3 V (mA/mm)	f_T (GHz)
-5	155	558	16.9	867	1.54×10^3	261
0	79.0	142	2.10×10^5	705	1.26×10^3	520
5	66.8	69.5	1.19×10^7	417	1.06×10^3	557

As seen in the carrier concentration profiles of the simulated FinFETs, the device with underlap shows a more ring-like distribution of carriers with an absence of electrons at the corners of the channel region, as opposed to the devices with overlap that effectively only have half of this pattern. In addition to having this full pattern of carriers in the channel, the minimum overlap device also is able to pull carriers closer to the gate, creating a higher peak concentration of electrons under the gate and leaving a more significant absence of electrons in the middle of the channel. This pulls more of the channel electrons into the triangular energy wells at the MOS interface, producing increased carrier confinement. The result of this is a device with higher electrostatic integrity and more clearly defined current paths than the overlapped device, leading to substantially improved subthreshold and ON/OFF characteristics. It should be noted however that as a result of pushing the corner carrier concentration further out from the gate, the current path to the source and drain is restricted compared to the overlap device which allows a much larger area of the fin to participate in current transport. As a result, the device with minimum overlap shows a reduced drive current at both 1 and 3 V. While further simulations showed that the drive current at 1 V can be increased by lowering the barrier height of the gate further, this change had relatively little effect on the value of I_{ds} at 3 V, implying the value is close to a saturated level.

6.3. Comparison with HEMTs and Si MOSFETs

Examining the increase in cutoff frequencies obtained for InP MOSFETs against gate length, the scalability of the InP devices studied is fairly clear. These results are plotted in Figure 6.11 against comparable HEMTs in InP, GaAs, and GaN material platforms. While several InP and GaAs HEMTs outperform the frequency response of planar devices studied at small L_g, use of a 15 nm L_g FinFET structure with S/D underlap produces f_T comparable to the fastest HEMTs. In addition, the InP FinFET has substantially lower subthreshold current, lower SS, and improved saturation characteristics compared to high f_T HEMTs [14], all while maintaining high-frequency power handling capabilities approaching that of HBT devices [100]. This performance is achieved with comparable transconductance to the HEMT devices, ensuring low noise operation for extremely high-frequency applications.

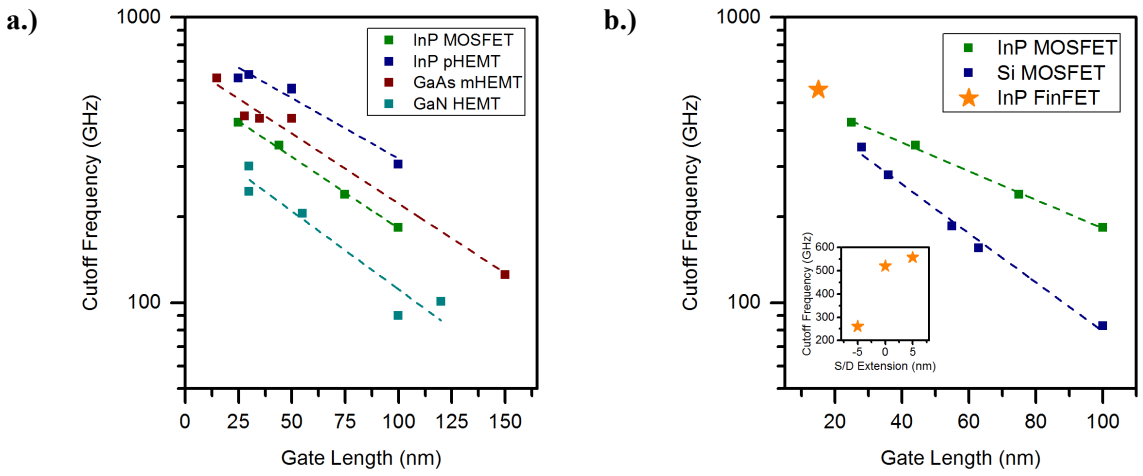


Figure 6.11. **f_T of the InP MOSFET at various L_g, compared against a.) InP [14][101-103], GaAs [104-107], and GaN [108-112] HEMTs, b.) Si MOSFETs [113-114] (with comparison of InP FinFETs and corresponding L_{ext}).**

Likewise, these results are competitive with the characteristics of highly-scaled Si MOSFETs currently found in industry. For instance, in Intel's 14 nm Tri-Gate production process, use of a 20 nm physical gate length with 8 nm fin widths results in good subthreshold characteristics, with a SS of 65 mV/dec and DIBL of 60 mV/V [115]. In comparison, the InP FinFET with L_{ext} = 5 nm achieves similar values at a smaller L_g of 15 nm, producing a SS of 66.8 mV/dec and DIBL of 69.5 mV/V. Of particular note are the fact that these results are obtained with the use of only 2 gates as opposed to the 3 gates found in the Intel device structure. In addition, devices show operability at high V_{ds} values of 3 V without breakdown, enabling high-power applications at frequencies approaching the terahertz domain. Figure 6.11 shows plots of f_T against gate length for both InP MOSFETs and Si MOSFETs. While planar InP MOSFETs show higher f_T for most gate lengths, Si MOSFETs demonstrate a larger increase in f_T for scaled devices, eventually surpassing planar InP devices for very small L_g. The InP FinFET described still far exceeds the Si devices examined however, reaching an f_T value nearly twice than that of the next best device.

Chapter 7. Conclusions

In this work, ALD ZrO_2 high-k oxide films were examined for suitability with various III-V semiconductor materials with the aim of demonstrating integration in functional MOS devices. In particular, the surface chemistry of these III-V compounds was analyzed to understand how native oxides can be supressed through optimal thermodynamics of oxide deposition. A plasma-enhanced atomic layer deposition technique using organometallic precursors was used to deposit 6.9 nm ZrO_2 high-k films on GaAs, GaN, InAs, and InP n-type substrates at low temperature. MOSCAP devices were fabricated using these oxide films and characterized using a parametric analyzer, demonstrating improved capacitance density, lower interface trap density, and higher effective MOS mobility for GaN and InP substrates. Simulations were then run to examine the device characteristics of entire InP nMOSFETs using the experimentally extracted MOS parameters obtained. The InP MOSFETs simulated showed excellent drive current and cutoff frequencies, as well as high scalability for future device nodes. Various device and structural parameters were varied and subsequently optimized to ensure controlled short-channel effects with sustained performance for small L_g InP devices. The results of these experiments produced optimal InP MOSFETs with excellent frequency response and subthreshold characteristics, while maintaining distinct advantages over alternative technologies such as III-V HEMTs and Si MOSFETs.

7.1. Future Work and Directions

While this work has addressed many of the central components necessary for constructing optimal InP MOS devices, several aspects have been left outstanding that could be addressed in future studies. In particular, a more in-depth analysis of the quantum components of these devices could be performed, including further analysis of the channel carrier mechanics alongside the incorporation of NEGF methods. Likewise, further characterization of the RF properties of these scaled MOS devices is warranted, including a discussion of figures of merit such as maximum power gain alongside analysis of device S-parameters. Combining all of this comprehensive output, SPICE simulations could be also used to evaluate the performance advantages of these InP MOSFETs in real circuit applications against competing technology platforms.

Another major point overlooked in this work is the implications of increased body or channel doping in integrated circuit manufacturing. As the overall size of devices is reduced, the number of atoms in the channel decreases with a cubic dependence. This causes significant issues in the manufacturing yield of small L_g MOSFETs, where the statistical variation of dopant atoms can introduce a significant variation in the threshold voltage owing to the relatively small number of atoms present in the channel [116]. This can however be rectified by use of either an epitaxial lower doped channel with a highly doped substrate or a delta doped layer placed underneath the channel [117]. Future work on these InP MOS devices could incorporate an analysis of the use of these structures to maintain performance while ensuring maximum manufacturability.

As is clear from current projections, Si devices are quickly running out of headroom for scaling in both high-speed digital and analog domains. With the significant economic repercussions of a slow-down in IC performance readily apparent, engineers must turn to feasible alternatives to continue the advancement of high-frequency deep submicron process nodes. III-V MOSFETs, particularly those constructed in the InP material platform, demonstrate excellent characteristics even when scaled to gate lengths of 15 nm (equivalent to the 7 nm manufacturing node). Future work in the physical integration of fabrication processes at the nanoscale will be necessary to create fully functional InP MOSFETs for integrated circuit applications.

References

[1] Bardeen, John, and Walter Hauser Brattain. "The transistor, a semi-conductor triode." *Physical Review* 74.2 (1948): 230.

[2] Moore, G. E. "Cramming More Components Onto Integrated Circuits." *Proceedings of the IEEE* 86.1 (1998): 82-85.

[3] Faggin, Federico, and Thomas Klein. "Silicon gate technology." *Solid-State Electronics* 13.8 (1970): 1125-1144.

[4] Semiconductor Industry Association. *Global Semiconductor Sales Top $335 Billion in 2015*. 1 Feb. 2016. Web.

[5] Ellinger, F., et al. "Review of advanced and beyond CMOS FET technologies for radio frequency circuit design." *2011 SBMO/IEEE MTT-S International Microwave & Optoelectronics Conference (IMOC)*, IEEE, 2011.

[6] Subramanian, Vaidyanathan, et al. "Identifying the bottlenecks to the RF performance of FinFETs." *2010 International Conference on VLSI Design*, IEEE, 2010.

[7] Del Alamo, Jesús A. "Nanometre-scale electronics with III-V compound semiconductors." *Nature* 479.7373 (2011): 317-323.

[8] Raab, Frederick H., et al. "Power amplifiers and transmitters for RF and microwave." *IEEE Transactions on Microwave Theory and Techniques* 50.3 (2002): 814-826.

[9] "NSM Archive – Physical Properties of Semiconductors." IOFFE Institute. Web.

[10] Baechtold, W., et al. "Si and GaAs 0.5 μm-gate Schottky-barrier field-effect transistors." *Electronics Letters* 9.10 (1973): 232-234.

[11] Liu, William. *Fundamentals of III-V Devices: HBTs, MESFETs, and HFETs/HEMTs*. Wiley, 1999.

[12] Papanicolaou, N. A., et al. "Sb-based HEMTs with InAlSb/InAs heterojunction." *Electronics Letters* 41.19 (2005): 1088-1089.

[13] Nakanisi, T., et al. "Growth of high-purity GaAs epilayers by MOCVD and their applications to microwave MESFET's." *Journal of Crystal Growth* 55.1 (1981): 255-262.

[14] Kim, Dae-Hyun, and Jesús A. Del Alamo. "30-nm InAs pseudomorphic HEMTs on an InP substrate with a current-gain cutoff frequency of 628 GHz." *IEEE Electron Device Letters* 29.8 (2008): 830-833.

[15] Krishnamohan, Tejas, Donghyun Kim, and Krishna C. Saraswat. "Properties and Trade-Offs of Compound Semiconductor MOSFETs." *Fundamentals of III-V Semiconductor MOSFETs*. Springer US, 2010, pp. 7-27.

[16] Ko, Hyunhyub, et al. "Ultrathin compound semiconductor on insulator layers for high-performance nanoscale transistors." *Nature* 468.7321 (2010): 286-289.

[17] Kim, Sang-Hyeon, et al. "High Performance Tri-Gate Extremely Thin-Body InAs-On-Insulator MOSFETs With High Short Channel Effect Immunity and Tunability." *IEEE Transactions on Electron Devices* 61.5 (2014): 1354-1360.

[18] Bai, Peng, et al. "A 65nm logic technology featuring 35nm gate lengths, enhanced channel strain, 8 Cu interconnect layers, low-k ILD and 0.57 μm^2 SRAM cell." *2004 International Electron Devices Meeting (IEDM)*, IEEE, 2004.

[19] Mistry, Kaizad, et al. "A 45nm logic technology with high-k+ metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging." *2007 International Electron Devices Meeting (IEDM)*, IEEE, 2007.

[20] Robertson, John. "High dielectric constant gate oxides for metal oxide Si transistors." *Reports on Progress in Physics* 69.2 (2005): 327.

[21] Wilmsen, C. W. "Oxide layers on III–V compound semiconductors." *Thin Solid Films* 39 (1976): 105-117.

[22] Passlack, Matthias, et al. "In-situ Ga_2O_3 process for GaAs inversion/accumulation device and surface passivation applications." *1995 International Electron Devices Meeting (IEDM)*, IEEE, 1995.

[23] Houng, M. P., et al. "Extremely low temperature formation of silicon dioxide on gallium arsenide." *Journal of Applied Physics* 82 (1997): 5788-5792.

[24] George, Steven M. "Atomic layer deposition: an overview." *Chemical Reviews* 110.1 (2009): 111-131.

[25] Halls, Mathew D., and Krishnan Raghavachari. "Atomic layer deposition growth reactions of Al_2O_3 on Si (100)-2×1." *The Journal of Physical Chemistry B* 108.13 (2004): 4058-4062.

[26] Ritala, Mikko, et al. "Perfectly conformal TiN and Al_2O_3 films deposited by atomic layer deposition." *Chemical Vapor Deposition* 5.1 (1999): 7-9.

[27] Von Hauff, Peter. *Metal Oxide Processing on Gallium Nitride and Silicon*. Diss. University of Alberta, 2012.

[28] Van Hemmen, J. L., et al. "Plasma and thermal ALD of Al_2O_3 in a commercial 200 mm ALD reactor." *Journal of the Electrochemical Society* 154.7 (2007): G165-G169.

[29] Goldstein, D. N., and S. M. George. "Enhancing the nucleation of palladium atomic layer deposition on Al_2O_3 using trimethylaluminum to prevent surface poisoning by reaction products." *Applied Physics Letters* 95.14 (2009): 143106.

[30] Elers, Kai, and Wei-Min Li. "Metal nitride deposition by ALD with reduction pulse." U.S. Patent No. 6,986,914. 17 Jan. 2006.

[31] Bothe, Kyle M., et al. "Electrical Comparison of HfO_2 and ZrO_2 Gate Dielectrics on GaN." *IEEE Transactions on Electron Devices* 60.12 (2013): 4119-4124.

[32] Afanas'ev, V. V., et al. "Band alignments in metal–oxide–silicon structures with atomic-layer deposited Al_2O_3 and ZrO_2." *Journal of Applied Physics* 91.5 (2002): 3079-3084.

[33] Rezazadeh, Vallen, et al. "Defect Characterization of PEALD High-k ZrO_2 Films Fabricated on III-V Materials." *IEEE Transactions on Semiconductor Manufacturing* 29.4 (2016): 355.

[34] Ye, P. D., et al. "GaAs MOSFET with oxide gate dielectric grown by atomic layer deposition." *IEEE Electron Device Letters* 24.4 (2003): 209-211.

[35] Saal, James E., et al. "Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD)." *JOM* 65.11 (2013): 1501-1509.

[36] Thurmond, C. D., et al. "GaAs Oxidation and the Ga-As-O Equilibrium Phase Diagram." *Journal of the Electrochemical Society* 127.6 (1980): 1366-1371.

[37] Hollinger, G., R. Skheyta-Kabbani, and M. Gendry. "Oxides on GaAs and InAs surfaces: An x-ray-photoelectron-spectroscopy study of reference compounds and thin oxide layers." *Physical Review B* 49.16 (1994): 11159.

[38] Henegar, Alex J., and Theodosia Gougousi. "Native oxide transport and removal during the atomic layer deposition of Ta_2O_5 on InAs (100) surfaces." *Journal of Vacuum Science and Technology A* 34.3 (2016): 031101.

[39] Wilmsen, C. W., ed. *Physics and Chemistry of III-V Compound Semiconductor Interfaces*. Plenum Press, 1985.

[40] Robach, Y., et al. "Passivation of InP using $In(PO_3)_3$-condensed phosphates: From oxide growth properties to metal-insulator-semiconductor field-effect-transistor devices." *Journal of Applied Physics* 71.6 (1992): 2981-2992.

[41] Lin, L., and J. Robertson. "Defect states at III-V semiconductor oxide interfaces." *Applied Physics Letters* 98.8 (2011): 2903.

[42] Schwartz, G. P., B. V. Dutt, and G. J. Gualtieri. "Elemental arsenic in native oxide films on $Al_xGa_{1-x}As$." *Applied Physics Letters* 39.1 (1981): 52-54.

[43] Wang, Yanzhen, et al. "Improved electrical properties of HfO_2-based gate dielectrics on InP substrate using Al_2O_3/HfO_2 and SF_6 plasma treatment." *Electrochemical and Solid-State Letters* 14.7 (2011): H291-H293.

[44] Hu, Chenming. *Modern Semiconductor Devices for Integrated Circuits*. Prentice Hall, 2010.

[45] Schroder, Dieter K. *Semiconductor Material and Device Characterization*. Wiley-IEEE Press, 2006.

[46] Nicollian, E. H., and A. Goetzberger. "The Si-SiO_2 Interface — Electrical Properties as Determined by the Metal-Insulator-Silicon Conductance Technique," *Bell System Technical Journal* 46.6 (1967): 1055–1133.

[47] Depas, Michel, et al. "Determination of tunnelling parameters in ultra-thin oxide layer poly-$Si/SiO_2/Si$ structures." *Solid-State Electronics* 38.8 (1995): 1465-1471.

[48] Ranuárez, Juan C., M. Jamal Deen, and Chih-Hung Chen. "A review of gate tunneling current in MOS devices." *Microelectronics Reliability* 46.12 (2006): 1939-1956.

[49] Lenzlinger, M., and E. H. Snow. "Fowler-Nordheim Tunneling into Thermally Grown SiO_2." *Journal of Applied Physics* 40.1 (1969): 278-283.

[50] Sze, Simon M., and Kwok K. Ng. *Physics of Semiconductor Devices*. John Wiley & Sons, 2006.

[51] Chau, Robert, et al. "Application of high-κ gate dielectrics and metal gate electrodes to enable silicon and non-silicon logic nanotechnology." *Microelectronic Engineering* 80 (2005): 1-6.

[52] Pierret, R. F., and C. T. Sah. "An MOS-oriented investigation of effective mobility theory." *Solid-State Electronics* 11.3 (1968): 279-290.

[53] Bothe, Kyle M., et al. "Capacitance Modeling and Characterization of Planar MOSCAP Devices for Wideband-Gap Semiconductors With High-κ Dielectrics." *IEEE Transactions on Electron Devices* 59.10 (2012): 2662-2666.

[54] Hareland, Scott A., et al. "A computationally efficient model for inversion layer quantization effects in deep submicron N-channel MOSFETs." *IEEE Transactions on Electron Devices* 43.1 (1996): 90-96.

[55] Shahrjerdi, Davood, et al. "Unpinned metal gate/high-κ GaAs capacitors: Fabrication and characterization." *Applied Physics Letters* 89.4 (2006): 043501.

[56] Konda, R. B., et al. "Electrical characteristics of ZrO_2/GaAs MOS capacitor fabricated by atomic layer deposition." *Journal of Vacuum Science and Technology A* 31.4 (2013): 041505.

[57] Chang, Y. C., et al. "Effective reduction of interfacial traps in Al_2O_3/GaAs (001) gate stacks using surface engineering and thermal annealing." *Applied Physics Letters* 97.11 (2010): 112901.

[58] Wang, C. H., et al. "InAs hole inversion and bandgap interface state density of 2×10^{11} $cm^{-2}eV^{-1}$ at HfO_2/InAs interfaces." *Applied Physics Letters* 103.14 (2013): 143510.

[59] Wang, C. H., et al. "High-k dielectrics on (100) and (110) n-InAs: Physical and electrical characterizations." *AIP Advances* 4.4 (2014): 047108.

[60] Wheeler, D., et al. "Deposition of HfO_2 on InAs by atomic-layer deposition." *Microelectronic Engineering* 86.7 (2009): 1561-1563.

[61] Zhao, Han, et al. "Effects of InP barrier layer thicknesses and different ALD oxides on device performance of $In_{0.7}Ga_{0.3}As$ MOSFETs." *2010 Device Research Conference (DRC) Digest*, South Bend, IN, 2010, pp. 55-56.

[62] Chen, Yen-Ting, et al. "ZrO_2-Based InP MOSFETs (EOT≈1.2 nm) Using Various Interfacial Dielectric Layers." *ECS Transactions* 35.3 (2011): 375-382.

[63] Wu, Y. Q., et al. "Enhancement-mode InP n-channel metal-oxide-semiconductor field-effect transistors with atomic-layer-deposited Al_2O_3 dielectrics." *Applied Physics Letters* 91.2 (2007): 022108.

[64] Galatage, R. V., et al. "Effect of post deposition anneal on the characteristics of HfO_2/InP metal-oxide-semiconductor capacitors." *Applied Physics Letters* 99.17 (2011): 172901.

[65] Wu, Y. Q., et al. "Photo-assisted capacitance-voltage characterization of high-quality atomic-layer-deposited Al_2O_3/GaN metal-oxide-semiconductor structures." *Applied Physics Letters* 90.14 (2007): 143504.

[66] Son, Junwoo, et al. "Fixed charge in high-k/GaN metal-oxide-semiconductor capacitor structures." *Applied Physics Letters* 101.10 (2012): 102905.

[67] Esposto, Michele, et al. "Electrical properties of atomic layer deposited aluminum oxide on gallium nitride." *Applied Physics Letters* 99.13 (2011): 133503.

[68] Nepal, Neeraj, et al. "Assessment of GaN surface pretreatment for atomic layer deposited high-k dielectrics." *Applied Physics Express* 4.5 (2011): 055802.

[69] Lin, Y-S., R. Puthenkovilakam, and J. P. Chang., "Dielectric property and thermal stability of HfO_2 on silicon." *Applied Physics Letters* 81.11 (2002): 2041-2043.

[70] Lin, Dennis H. C., et al. "Electrical characterization of the MOS (Metal-oxide-semiconductor) system: High mobility substrates." *ECS Transactions* 34.1 (2011): 1065-1070.

[71] Suné, Jordi, Piero Olivo, and Bruno Riccó. "Quantum-mechanical modeling of accumulation layers in MOS structure." *IEEE Transactions on Electron Devices* 39.7 (1992): 1732-1739.

[72] Lubow, Abigail, Sohrab Ismail-Beigi, and T. P. Ma. "Comparison of drive currents in metal-oxide-semiconductor field-effect transistors made of Si, Ge, GaAs, InGaAs, and InAs channels." *Applied Physics Letters* 96.12 (2010): 122105.

[73] Xu, Min, et al. "New insights in the passivation of high-k/InP through interface characterization and metal–oxide–semiconductor field effect transistor demonstration: Impact of crystal orientation." *Journal of Applied Physics* 113.1 (2013): 013711.

[74] Wang, Yanzhen, et al. "Improved electrical properties of HfO_2-based gate dielectrics on InP substrate using Al_2O_3/HfO_2 and SF_6 plasma treatment." *Electrochemical and Solid-State Letters* 14.7 (2011): H291-H293.

[75] Chen, Yen-Ting, et al. "ZrO_2-Based InP MOSFETs (EOT≈1.2 nm) Using Various Interfacial Dielectric Layers." *ECS Transactions* 35.3 (2011): 375-382.

[76] Tsai, C. Y., T. L. Wu, and Albert Chin. "High-Performance GaN MOSFET With High-κ $LaAlO_3/SiO_2$ Gate Dielectric." *IEEE Electron Device Letters* 33.1 (2012): 35-37.

[77] Kambayashi, Hiroshi, et al. "High quality SiO_2/Al_2O_3 gate stack for GaN metal–oxide–semiconductor field-effect transistor." *Japanese Journal of Applied Physics* 52.4S (2013): 04CF09.

[78] Tang, K., W. Huang, and T. Paul Chow. "GaN MOS capacitors and FETs on plasma-etched GaN surfaces." *Journal of Electronic Materials* 38.4 (2009): 523-528.

[79] "Process Integration, Devices, and Structures". *The International Technology Roadmap for Semiconductors 2003 Edition*. International SEMATECH, 2003.

[80] Turin, Valentin O. "A modified transferred-electron high-field mobility model for GaN devices simulation." *Solid-State Electronics* 49.10 (2005): 1678-1682.

[81] *Crosslight Device Simulation Software - General Manual*. Crosslight Software Inc., 28 Sept. 2015. PDF.

[82] Jallepalli, S., et al. "Electron and hole quantization and their impact on deep submicron silicon p-and n-MOSFET characteristics." *IEEE Transactions on Electron Devices* 44.2 (1997): 297-303.

[83] Pirovano, Agostino, Andrea L. Lacaita, and Alessandro S. Spinelli. "Two-dimensional quantum effects in nanoscale MOSFETs." *IEEE Transactions on Electron Devices* 49.1 (2002): 25-31.

[84] Vardi, Alon, Xin Zhao, and Jesús A. del Alamo. "Quantum-size effects in sub 10-nm fin width InGaAs FinFETs." *2015 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2015.

[85] Granzner, Ralf, et al. "Simulation of nanoscale MOSFETs using modified drift-diffusion and hydrodynamic models and comparison with Monte Carlo results." *Microelectronic Engineering* 83.2 (2006): 241-246.

[86] Lundstrom, M., and Z. Ren. "Essential physics of carrier transport in nanoscale MOSFETs." *IEEE Transactions on Electron Devices* 49.1 (2002): 133-141.

[87] Naveh, Y., and K. K. Likharev. "Modeling of 10-nm-scale ballistic MOSFET's." *IEEE Electron Device Letters* 21.5 (2000): 242-244.

[88] Martinez, Antonio, et al. "A self-consistent full 3-D real-space NEGF simulator for studying nonperturbative effects in nano-MOSFETs." *IEEE Transactions on Electron Devices* 54.9 (2007): 2213-2222.

[89] Roy, Kaushik, Saibal Mukhopadhyay, and Hamid Mahmoodi-Meimand. "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits." *Proceedings of the IEEE* 91.2 (2003): 305-327.

[90] Dutta, T., et al. "Origins of the short channel effects increase in III-V nMOSFET technologies." *2012 13th International Conference on Ultimate Integration on Silicon (ULIS)*, IEEE, 2012.

[91] Fischetti, M. V., et al. "Simulation of electron transport in high-mobility MOSFETs: Density of states bottleneck and source starvation." *2007 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2007.

[92] Singisetti, Uttam, et al. "$In_{0.53}Ga_{0.47}As$ Channel MOSFETs with Self-Aligned InAs Source/Drain formed by MEE Regrowth." *IEEE Electron Device Letters* 30.11 (2009): 1128-1130.

[93] Li, Q., et al. "Inverted-type InAlAs/InGaAs MOSHEMT with Regrown Source/Drain Exhibiting High Current and Low On-resistance." *2012 International Conference on Compound Semiconductor Manufacturing Technology (CS MANTECH) Digest*, Boston, MA, 2012.

[94] Vitale, Steven A., et al. "FDSOI process technology for subthreshold-operation ultralow-power electronics." *Proceedings of the IEEE* 98.2 (2010): 333-342.

[95] Gwoziecki, Romain, et al. "Optimization of V_{th} roll-off in MOSFET's with advanced channel architecture-retrograde doping and pockets." *IEEE Transactions on Electron Devices* 46.7 (1999): 1551-1561.

[96] Yu, Bin, et al. "FinFET scaling to 10 nm gate length." *2002 International Electron Devices Meeting (IEDM)*, IEEE, 2002.

[97] Wu, Y. Q., et al. "First experimental demonstration of 100 nm inversion-mode InGaAs FinFET through damage-free sidewall etching." *2009 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2009.

[98] Kranti, Abhinav, Jean-Pierre Raskin, and G. Alastair Armstrong. "Optimizing FinFET geometry and parasitics for RF applications." *2008 SOI Conference*, IEEE, 2008.

[99] Trivedi, Vishal, Jerry G. Fossum, and Murshed M. Chowdhury. "Nanoscale FinFETs with gate-source/drain underlap." *IEEE Transactions on Electron Devices* 52.1 (2005): 56-62.

[100] Liu, H. G., S. P. Watkins, and C. R. Bolognesi. "15-nm base type-II InP/GaAsSb/InP DHBTs with F_T= 384 GHz and a 6-V BV CEO." *IEEE Transactions on Electron Devices* 53.3 (2006): 559-561.

[101] Wojtowicz, M., et al. "0.10 μm graded InGaAs channel InP HEMT with 305 GHz f_T and 340 GHz f_{max}," *IEEE Electron Device Letters* 15.11 (1994): 477-479.

[102] Yamashita, Yoshimi, et al. "Pseudomorphic $In_{0.52}Al_{0.48}As/In_{0.7}Ga_{0.3}As$ HEMT's with an ultrahigh f_T of 562 GHz." *IEEE Electron Device Letters* 23.10 (2002): 573-575.

[103] Mei, Xiaobing, et al. "First demonstration of amplification at 1 THz using 25-nm InP high electron mobility transistor process." *IEEE Electron Device Letters* 36.4 (2015): 327-329.

[104] Yeon, Seong-Jin, et al. "610 GHz InAlAs/$In_{0.75}$GaAs metamorphic HEMTs with an ultra-short 15-nm-gate." *2007 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2007.

[105] Elgaid, K., et al. "50-nm T-gate metamorphic GaAs HEMTs with f_T of 440 GHz and noise figure of 0.7 dB at 26 GHz." *IEEE Electron Device Letters* 26.11 (2005): 784-786.

[106] Lee, Kang-Sung, et al. "35-nm Zigzag T-Gate Metamorphic GaAs HEMTs With an Ultrahigh f_{max} of 520 GHz." *IEEE Electron Device Letters* 28.8 (2007): 672-675.

[107] Zaknoune, M., et al. "InAlAs/InGaAs metamorphic HEMT with high current density and high breakdown voltage." *IEEE Electron Device Letters* 19.9 (1998): 345-347.

[108] Lu, Wu, et al. "AlGaN/GaN HEMTs on SiC with over 100 GHz f_T and low microwave noise." *IEEE Transactions on Electron Devices* 48.3 (2001): 581-585.

[109] Sun, H. F., et al. "100 nm gate AlGaN/GaN HEMTs on silicon with f_T= 90 GHz." *Electronics Letters* 45.7 (2009): 376-377.

[110] Lee, Dong Seup, et al. "300-GHz InAlN/GaN HEMTs with InGaN back barrier." *IEEE Electron Device Letters* 32.11 (2011): 1525-1527.

[111] Sun, Haifeng, et al. "205-GHz (Al, In) N/GaN HEMTs." *IEEE Electron Device Letters* 31.9 (2010): 957-959.

[112] Lee, Dong Seup, et al. "245-GHz InAlN/GaN HEMTs with oxygen plasma treatment." *IEEE Electron Device Letters* 32.6 (2011): 755-757.

[113] Barlage, Doug, et al. "High-frequency response of 100 nm integrated CMOS transistors with high-K gate dielectrics." *2001 International Electron Devices Meeting (IEDM)*, IEEE, 2001.

[114] Li, Hongmei, et al. "Technology scaling and device design for 350 GHz RF performance in a 45nm bulk CMOS process." *2007 IEEE Symposium on VLSI Technology*, IEEE, 2007.

[115] Natarajan, S., et al. "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm^2 SRAM cell size." *2014 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2014.

[116] Li, Yiming, et al. "Discrete dopant fluctuations in 20-nm/15-nm-gate planar CMOS." *IEEE Transactions on Electron Devices* 55.6 (2008): 1449-1455.

[117] Asenov, Asen, and Subhash Saini. "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1-μm MOSFET's with epitaxial and δ-doped channels." *IEEE Transactions on Electron Devices* 46.8 (1999): 1718-1724.