FABRICATION AND CHARACTERIZATION OF GALLIUM NITRIDE ELECTRONIC DEVICES

By

JERRY WAYNE JOHNSON

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2001
To DSJ and MSJ, with loving memories of the past and dreams for a bright future.
ACKNOWLEDGMENTS

The completion of such a manuscript is in no way the sole accomplishment of the author. I would first like to thank my research advisor, Fan Ren, for technical support and guidance throughout every aspect of this work. I would also like to extend my deepest gratitude to him for accepting me into his group and for the permanent mark he has made on my life both professionally and personally. He has taught me more than he could ever imagine. I am also indebted to the other members of my advisory committee: Tim Anderson, Steve Pearton, and Cammy Abernathy. Their significant contributions to this work are greatly appreciated. I am honored to have been associated with such an eminent committee.

Collaboration in scientific research has become the norm, not the exception. Certainly this work is a testament to the joint efforts of many individuals. First I would like to thank my colleagues and co-workers in the Department of Chemical Engineering: Ben Luo, Anping Zhang, Gerard Dang, Jeff LaRoche, Rishabh Mehandru, and Jihun Kim. I specifically thank Ben for many enlightening discussions and Friday meetings. May each of these continue in my absence. The excellent crystal growth, vacuum system expertise, and friendship of Brent Gila are gratefully acknowledged. Thanks are also extended to Kyu-Pil Lee for always finding time to assist me with ICP etching.

Several others have helped make the past 4½ years more enjoyable and/or more rewarding, and I thank them for it. Jay Lewis is recognized for his friendship, golf tournaments, and 6-string wizardry. Mike Mastro has endured sharing office space with
me during our concurrent graduate careers. Olga Kryliouk has been a friend, collaborator, and at times, a surrogate Mother. She is indeed one of the nicest people I have ever met. Nancy and Shirley in the ChE office deserve special recognition for their ever-smiling faces and friendly dispositions. I'd like to thank Mark George at Deposition Sciences for lessons learned in the finer points of both LP-CVD and Red Zinfandel.

There truly is a necessary balance in life. I am grateful to Albert Baca for the opportunity to spend 4 months at Sandia National Laboratories. Professionally, my time there was one of the most productive of my life. Acknowledgement is also warranted for fellow Sandians Ron Briggs, Cédric Monier, Randy Shul, Melissa Cavaliere, Gerry Lopez, Andrea Ongstad, Joel Wendt, and Marce Armendariz.

To my family, I am eternally indebted. Since my earliest memories, they have been a pillar of support and a source of confidence. They have motivated and allowed me to strive for things I could have never accomplished on my own. I thank them for their hard work, dedication to family, and their unwavering love. I will strive to model my fathering style after the parenting skills they have always demonstrated. If I am successful, Max will be very lucky—as I have been.

Finally, I thank my wife Traci. She is truly the most special person I have ever met and the impact she has made on my life is immeasurable. She is my wife, the mother of our son, and my very best friend. I feel this work is as much her achievement as it is my own. I wish there was a way for me to express the deepness of my love and appreciation.
# TABLE OF CONTENTS

| ACKNOWLEDGMENTS | iii |
| ABSTRACT | viii |
| **CHAPTERS** | |
| **1 INTRODUCTION** | 1 |
| 1.1 Motivation | 1 |
| 1.2 Dissertation Outline | 10 |
| **2 BACKGROUND AND LITERATURE REVIEW** | 12 |
| 2.1 Historical Review | 12 |
| 2.2 Crystal Structure | 17 |
| 2.3 Thin Film Growth | 18 |
| 2.3.1 Substrates for III-N Epitaxy | 18 |
| 2.3.2 Doping | 23 |
| 2.4 Etching | 27 |
| 2.4.1 Photoenhanced Wet Chemical Etching | 27 |
| 2.4.2 Dry Etching | 29 |
| 2.5 Metallization | 33 |
| 2.5.1 Ohmic Contacts | 37 |
| 2.5.1.1 n-type | 37 |
| 2.5.1.2 p-type | 40 |
| 2.5.2 Rectifying Contacts | 42 |
| 2.5.2.1 n-type | 42 |
| 2.5.2.2 p-type | 44 |
| 2.5.3 Surface Treatment | 44 |
| 2.6 III-N Optoelectronics | 45 |
| 2.6.1 Light Emitting Diodes | 46 |
| 2.6.2 Laser Diodes | 53 |
| 2.6.3 Detectors | 57 |
| 2.7 III-N Electronics | 58 |
| 2.7.1 Field Effect Transistors | 59 |
| 2.7.2 Bipolar Transistors | 65 |
| 2.7.3 Rectifiers | 69 |
# AlGaN / GaN HIGH ELECTRON MOBILITY TRANSISTORS

3.1 Introduction ................................. 72
3.2 MOCVD-Grown HEMTs on Al₂O₃ Substrates ....... 75
   3.2.1 Device Processing .......................... 75
   3.2.2 Optical Gate Devices ....................... 81
   3.2.3 Submicron Gate Devices .................... 86
   3.2.4 Small Signal Modeling ..................... 104
3.3 Direct Comparison of HEMTs Grown on Al₂O₃ or AlN/SiC ...... 111
3.4 AlGaN/GaN HEMTs on RF-assisted MBE-grown Epilayers ...... 143

# GROWTH AND ELECTRICAL CHARACTERIZATION OF NOVEL GATE OXIDES FOR GaN MOS DEVICES

4.1 Introduction .................................. 151
4.2 Molecular Beam Epitaxial Growth .................. 152
4.3 Quantification of Oxide Quality .................... 154
4.4 The MOS Capacitor ................................ 157
   4.4.1 Collection of Capacitance-Voltage Data .... 159
   4.4.2 Ideal MOS Capacitor .......................... 160
      4.4.2.1 Accumulation ............................ 161
      4.4.2.2 Depletion ................................. 164
      4.4.2.3 Inversion ................................ 166
      4.4.2.4 Flatband ................................ 168
      4.4.2.5 Deep Depletion .......................... 170
   4.4.3 Non-Ideal MOS ................................ 171
      4.4.3.1 Effective Oxide Charge .................. 171
      4.4.3.2 Interface Trapped Charge ............... 173
4.5 Gadolinium Oxide ................................ 176
   4.5.1 High Temperature Gd₂O₃ ....................... 179
   4.5.2 Low Temperature Gd₂O₃ ....................... 183
4.6 Scandium Oxide .................................. 191

# GaN METAL OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTORS

5.1 Introduction .................................. 194
5.2 MOSFET Processing ................................ 196
5.3 Results and Discussion ........................... 200
   5.3.1 Gd₂O₃ / GaN Devices ....................... 200
   5.3.2 SiO₂ / Gd₂O₃ / GaN Stacked Dielectric Devices ....... 203
6 GaN SCHOTTKY RECTIFIERS ON BULK GaN SUBSTRATES.......................... 208

6.1 Introduction.................................................................................. 208
6.2 Material Characterization and Diode Fabrication........................ 211
6.3 Current-Voltage Results and Discussion: Schottky Diodes.............. 217
6.4 Large- and Small-Area Bulk GaN Rectifiers with Implanted Guard Rings .......... 224
6.5 Switching Behavior...................................................................... 237

7 SUMMARY, CONCLUSIONS, AND FUTURE WORK .......................... 240

7.1. AlGaN / GaN HEMTs ................................................................. 240
7.2. Gate Oxide Growth and Processing........................................... 243
7.3. MOSFETs and MOSHEMTs...................................................... 246
7.4. GaN Rectifiers on Bulk Substrates........................................... 248

REFERENCES ................................................................................... 250

BIOGRAPHICAL SKETCH ................................................................. 268
MRCATION AND CHARACTERIZATION OF GALLIUM NITRIDE ELECTRONIC DEVICES

By

Jerry Wayne Johnson

December 2001

Chairman: Fan Ren
Major Department: Chemical Engineering

Gallium nitride (GaN)-based high electron mobility transistors (HEMTs), metal oxide semiconductor field effect transistors (MOSFETs), and Schottky rectifiers were fabricated and characterized. Novel dielectric materials Gd2O3 and ScO were evaluated as potential gate dielectrics for GaN MOS applications. The devices presented herein show tremendous potential for elevated temperature, high frequency, and/or high voltage operation.

AlGaN/GaN HEMTs were grown by MOCVD on sapphire and SiC substrates and by RF-MBE on sapphire substrates. Devices were fabricated with gate lengths from 100 nm to 1.2 μm. Drain current density approached 1 A/mm and extrinsic transconductance exceeded 200 mS/mm for small gate periphery devices. For the shortest gate length, a unity-gain cutoff frequency (fT) of 59 GHz and a maximum frequency of oscillation (fmax) of 90 GHz were extracted from measured scattering parameters. The experimental s-parameters were in excellent agreement with simulated results from small-signal linear
modeling. Large signal characterization of $0.25 \times 150 \ \mu m^2$ devices produced 2.75 W/mm at 3 GHz and 1.7 W/mm at 10 GHz. Devices fabricated on high thermal conductivity SiC substrates exhibited superior high temperature performance and a reduced density of threading dislocations.

Novel gate dielectrics Gd$_2$O$_3$ and ScO were grown by gas source molecular beam epitaxy (GSMBE). Current-voltage (I-V) and capacitance-voltage (C-V) data were collected from MOS capacitors to evaluate the bulk and interfacial electrical properties of the insulators. Single crystal Gd$_2$O$_3$ was demonstrated on GaN, but the resultant MOSFET exhibited a large gate leakage attributed to defects and dislocations in the oxide. MOSFETs with a stacked gate dielectric of Gd$_2$O$_3$/SiO$_2$ were operational at a drain source bias of 80 V and a gate bias of +7 V.

Bulk GaN templates grown by hydride vapor phase epitaxy (HVPE) were used to fabricate vertical geometry Schottky rectifiers. Size- and temperature-dependent I-V characteristics are reported. These devices show significant improvements in forward turn-on voltage, on-state resistance, and reverse recovery characteristics relative to previously reported devices fabricated on GaN layers grown on sapphire.
CHAPTER 1
INTRODUCTION

One might say, with tongue in cheek, perhaps, that once we have developed diodes, transistors, tubes, resistors, capacitors, and insulators to tolerate 500 [°]C, we will no longer have thermal design problems.


1.1 Motivation

The bulk of modern microelectronics rests upon the technology of silicon-based transistors. Silicon is the 2nd most abundant element in the Earth’s crust, and its availability, attractive material properties, and stable native oxide have all contributed to its position of prominence. Since shortly after the advent of the transfer resistor, or “transistor,” in the mid-20th century, the Si-based semiconductor electronics industry has grown at an unprecedented pace. Moore’s Law scaling has led to an exponential rate of development never before witnessed in any other industry (Figure 1-1). Today, the Si transistor is believed to be the most numerous man-made object on our planet [1], totaling well over 1 million Si MOSFETs for each person in the Western world. While silicon devices dominate over 90% of the ~$250 billion solid state market and continue to experience strong growth, alternative material systems are becoming increasingly important.
As technology continues to advance, heightened performance demands are being made on semiconductor devices for use in both niche and mainstream applications. While silicon has proven to be the primary contestant in the integrated circuit (IC)

![Graph showing the number of transistors per chip for Intel CPUs from 4004 Microprocessor (1971) to Pentium® III (late 1999). Note the log scale of the vertical axis. Moore's Law has held as the empirical economic and forecast basis for the entire silicon electronics industry for over 25 years [2].](image)

market, there is an ever-growing need for devices operating at conditions beyond the limits of silicon. Satellite and radar communication systems, high temperature (>200°C) electronics, and high voltage solid state switching are examples of the numerous areas of potential for appropriately designed non-silicon devices and circuitry. In addition, traditional devices operating in harsh (e.g., chemical, radiation) environments are desirable for many military and aerospace applications. For such devices, the limitations
of silicon are largely due to its inherent material properties, such as its narrow indirect band gap.

Group III-V compound semiconductors represent an important class of alternative materials for various electronic and optoelectronic applications. GaAs is by far the most mature of the III-Vs, but still lags well behind Si in terms of development. GaAs technology is attractive for many reasons, including a small electron effective mass, high drift velocity, and the availability of AlGaAs/GaAs heterostructures. GaAs substrates can be grown by typical crystal pulling techniques in large diameters at relatively low cost. An additional benefit for device applications is the availability of semi-insulating substrates, which reduce parasitic effects and enable high frequency devices. The wider bandgap of GaAs (1.42 eV versus 1.12 eV for silicon) greatly enhances radiation tolerance and high temperature operation. In addition, GaAs and most other III-V materials have direct energy gaps, which allows fabrication of efficient light-emitting devices such as light emitting diodes (LEDs) and laser diodes (LDs). Commercially important III-As electronic devices include GaAs MESFETs for RF power amplification. GaAs monolithic microwave integrated circuits (MMICs) are helping to fuel the booming market for wireless communications. Impressively, the number of GaAs FETs on a chip has reached into the millions [3]. One of the most important optoelectronic devices is the AlGaAs/GaAs LD, used extensively in the read/write heads of CD- and DVD-ROM drives. Several other III-V materials, such as InP, are useful in niche markets but play significantly lesser commercial roles than GaAs.

Wide bandgap semiconductors are--in many ways--a progression beyond GaAs, just as GaAs evolved to overcome some of the limitations of Si. Wide bandgap materials
simultaneously provide attractive properties for electronic devices as well as the potential for short wavelength (visible to deep UV) light emission and detection. The Group III-nitrides (AlN, GaN, InN, and their alloys) were initially researched for their promise to fill the void for a blue solid state light emitter. GaN, in particular, received considerable attention because of its direct energy gap of 3.39 eV (365 nm).

![Figure 1-2. Energy bandgaps of several III-V, II-VI, and elemental semiconductors as a function of lattice constant. Note the limits of the visible portion of the electromagnetic spectrum.](image)

The energy gaps of several common semiconductors are given in Figure 1-2 as a function of lattice constant. The approximate boundaries of the visible spectrum are
shown as is the nature of the energy transition (direct or indirect) for each material. From the figure, it is noted that most III-V materials have bandgaps from red to infrared. Certain alloys, such as AlInGaP, are important for the 1.3 and 1.55 μm telecommunications wavelengths, but none of the ‘traditional’ III-V materials can reach the shorter green and blue wavelengths. InGaN/GaN-based blue LEDs and LDs have recently been demonstrated and commercialized by Nichia Corporation. These light-emitting devices remain an active area of worldwide research.

Electronic devices from III-nitrides have been a more recent phenomenon. Although their potential has been realized for several decades, wide band gap semiconductor electronic research is only beginning to flourish. The large energy gaps responsible for short wavelength light emission in these materials also give rise to extremely low thermal carrier generation rates. Figure 1-3 shows the dramatic decrease in intrinsic carrier concentration ($n_i$) for GaN relative to Si and GaAs. At room temperature, $n_i$ for GaN is approximately 20 orders of magnitude lower than that of Si, leading to significantly reduced leakage currents. This is further illustrated in Figure 1-4, where the intrinsic temperature ($T_i$) is given as a function of semiconductor impurity concentration. The intrinsic temperature is the temperature at which the concentration of thermally generated intrinsic carriers equals the concentration of electrically active impurities in the semiconductor (intentional + unintentional doping). This value can be used as a metric of the high temperature operation limit for a given semiconductor. As illustrated, GaN electronic devices are able to operate without cooling at temperatures well in excess of the limits of Si or GaAs. This not only reduces the cost of a given
Figure 1-3. Semilog plot of intrinsic carrier concentration versus inverse temperature for Si, GaAs, and GaN.

Figure 1-4. Intrinsic temperature of Si, GaAs, and GaN as a function of semiconductor impurity concentration.
subsystem, but also makes bulky cooling equipment unnecessary, decreasing weight and size.

The thermal conductivity of GaN is three times that of GaAs, as listed in Table 1-1. For high power or high temperature applications, good thermal conductivity is imperative for heat removal or sustained operation at elevated temperatures. The development of III-N and other wide bandgap technologies for high temperature applications will likely take place at the expense of competing technologies, such as silicon-on-insulator (SOI), at moderate temperatures [4]. At higher temperatures (>300°C), novel devices and components will become possible. The automotive industry will likely be one of the largest markets for such high temperature electronics. Automotive control components fabricated with wide bandgap materials could be mounted directly to an engine block, reducing signal delay. Well logging equipment, military combat systems, and aerospace components are just a few of the additional sectors in the global market for these devices. High temperature applications promise to play an increasingly important role for electronic devices of the future, with a total world market projected to reach over $1 billion by 2005 [4].

One of the most noteworthy advantages for III-N materials over other wide bandgap semiconductors is the availability of AlGaN/GaN and InGaN/GaN heterostructures. A 2-dimensional electron gas (2DEG) has been shown to exist at the AlGaN/GaN interface, and heterostructure field effect transistors (HFETs) from these materials can exhibit 2DEG mobilities approaching 2000 cm²/V·s at 300K. Combined with the polarization-enhanced sheet charge density of ~10¹³ cm⁻² and large peak and saturation velocities, AlGaN/GaN HFETs are capable of handling large DC and RF
Table 1-1. Selected 300 K material properties relevant to electronic device applications for Si, GaAs, and wide bandgap semiconductors [5-7]

<table>
<thead>
<tr>
<th>Property</th>
<th>Si</th>
<th>GaAs</th>
<th>4H-SiC</th>
<th>Diamond</th>
<th>InN</th>
<th>α-GaN (AlGaN/GaN)</th>
<th>AIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Bandgap, $E_g$ (eV)</td>
<td>1.12</td>
<td>1.42</td>
<td>3.25</td>
<td>5.45</td>
<td>1.89</td>
<td>3.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Saturated Electron Velocity, $v_{sat}$ ($\times 10^7$ cm/s)</td>
<td>1.0</td>
<td>1.0</td>
<td>2.0</td>
<td>2.7</td>
<td>2.5</td>
<td>2.5 (2.7)</td>
<td>1.4</td>
</tr>
<tr>
<td>Peak Electron Velocity, $v_{max}$ ($\times 10^7$ cm/s)</td>
<td>--</td>
<td>2.1</td>
<td>--</td>
<td>--</td>
<td>4.3</td>
<td>3.1</td>
<td>1.7</td>
</tr>
<tr>
<td>Breakdown Field, $E_B$ (MV/cm)</td>
<td>0.3</td>
<td>0.6</td>
<td>3</td>
<td>10</td>
<td>--</td>
<td>3</td>
<td>--</td>
</tr>
<tr>
<td>Electron Mobility, $\mu_n$ (cm$^2$/V·s)</td>
<td>1400</td>
<td>8000</td>
<td>800</td>
<td>2200</td>
<td>3200</td>
<td>900 (2000)</td>
<td>135</td>
</tr>
<tr>
<td>Hole Mobility, $\mu_p$ (cm$^2$/V·s)</td>
<td>500</td>
<td>400</td>
<td>50</td>
<td>1600</td>
<td>--</td>
<td>50</td>
<td>14</td>
</tr>
<tr>
<td>Static Dielectric Constant, $\varepsilon_r$</td>
<td>11.8</td>
<td>12.8</td>
<td>9.7</td>
<td>5.5</td>
<td>15.3</td>
<td>9.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Thermal Conductivity, $\kappa$ (W/cm·K)</td>
<td>1.5</td>
<td>0.5</td>
<td>4.9</td>
<td>20-30</td>
<td>--</td>
<td>1.3</td>
<td>2.85</td>
</tr>
</tbody>
</table>

currents. HFETs have been by far the most heavily investigated III-N electronic device, largely due to their promise in the radio frequency semiconductor market. Power handling capabilities of ~12 W/mm appear feasible, and extraordinary large signal performance has already been demonstrated, with a current state-of-the-art of >10 W/mm at X-band and 6.6 W/mm at 20 GHz for AlGaN/GaN HEMTs on SiC substrates. These values are already far superior to the best power density achieved with GaAs devices, which is on the order of 1.5 W/mm at X-band [8]. The microwave market for III-N devices is expected to lie at moderate frequencies (X-band to Ku-band) and RF power levels greater than can be achieved with competing technologies. At higher frequencies
(and lower powers) materials such as InP are expected to dominate because of an extremely low electron effective mass.

Since the critical electric field for avalanche breakdown increases with the semiconductor bandgap, the predicted breakdown field of GaN is very high (~3 × 10^6 V/cm). This presents exciting possibilities for devices required to block large voltages, such as switches found in electrical transmission and distribution systems. Traditionally these devices are Si-based, with limitations on the maximum voltage ratings and current handling capabilities [9]. Wide bandgap solid state devices are expected to provide "cleaner" switching, without the voltage or inductance spikes typically associated with mechanical switches or p-i-n rectifiers. This will, in turn, allow the electric power grid to operate closer to its rated value, increasing efficiency. Due to the large power dissipation associated with switching, thermal management becomes a central issue in design and implementation of such devices. In this regard, III-nitrides also offer advantages because of their ability to operate at elevated temperatures without the need for external cooling. A major short-term target for wide bandgap switches is 13.8 kV, a common residential distribution mode. Future goals are 25 kV blocking voltage, 2 kA forward current, on-resistance < 2% of the rated voltage, and switching speeds of ~50 kHz [10]. Other applications include control circuitry for electronic motor drives, hybrid-electric automobiles and military vehicles, and more-electric airplanes and nautical vessels.

To summarize many of the attractive features of the III-nitrides, selected material properties of AlN, GaN, and InN relevant to electronic device applications are listed in Table 1-1, along with those of Si, GaAs, SiC, and diamond. Properties of InN are included for reference, although this material has not yet been applied extensively in
electronic devices. SiC and diamond are the primary competitors of the III-nitrides for many of the electronic device applications previously mentioned in this chapter. Currently, SiC technology is much more mature than that of GaN, with some RF devices already having reached the market [11]. Power devices from SiC will almost certainly make the first progress toward the commercial importance of wide bandgap electronics. However, GaN offers a larger breakdown field than SiC, as well as the greatly increased mobility afforded by heterostructures. It is believed that III-N devices will ultimately provide performance superior to that achievable with SiC. Diamond offers an extremely large bandgap and near-ideal material properties, but lags well behind due to difficulties associated with crystal growth and n-type doping.

To place the tremendous potential of this material system into proper perspective, Strategies Unlimited, a market research firm, recently released a market forecast predicting an annual growth rate of 28% for GaN electronic and optoelectronic devices for the period 1999 - 2009 [12]. Annual revenue is expected to skyrocket from $420 million in 1999 to >$4.8 billion by 2009. A majority of this revenue will stem from LED and LD sales, since the III-N electronics market will not begin until 2002. However, after 2002 the electronics sector is predicted to grow nearly 100% annually.

1.2 Dissertation Outline

This work deals exclusively with III-N electronic devices, with the only discussion of optoelectronic or photonic devices occurring in the background information of Chapter 2.

Chapter 2 reviews the technological progression of III-N materials and devices, with particular emphasis on past challenges, key processing issues, and current status of
understanding and development. Chapter 3 presents details of device fabrication and performance of AlGaN/GaN HEMTs on sapphire and SiC substrates. DC, small signal, and large signal characteristics are given. Chapter 4 deals with thin film growth and characterization of novel gate oxides for III-N MOS applications. The capacitance-voltage technique is used extensively to investigate the bulk and interfacial electrical characteristics of the gate dielectrics. This is followed in Chapter 5 by a demonstration of Gd$_2$O$_3$ as the gate insulator of GaN depletion mode MOSFETs. Chapter 6 discusses the fabrication and device performance of vertically-depleting Schottky diodes on high quality free-standing GaN templates. Chapter 7 gives a brief summary, concluding remarks, and suggestions for future work.
CHAPTER 2
BACKGROUND AND LITERATURE REVIEW

Group III-nitrides have been the most heavily researched compound semiconductor material system of the past decade. Entire conferences or symposia [13], books [14-16], and numerous review papers [17-20] have recently been devoted exclusively to the III-nitrides. The promise of these materials, particularly GaN, for short wavelength light emission was well understood more than 30 years ago. However, the III-nitrides have proven to be very different from ‘traditional’ III-V semiconductors, and several challenges unique to (Al,Ga,In)N were addressed before useful devices could be realized. Presently, the nitrides remain a relatively immature technology compared to other III-V semiconductors such as GaAs or InP. This chapter will introduce the pertinent chemical and physical properties of III-nitrides, chronicle their technical development, and present the current state of material growth, processing, and optoelectronic and electronic devices.

2.1 Historical Review

Early reports of GaN synthesis in various forms were given by Johnson et al. [21], Juza and Hahn [22], Grimeiss and Kowlmans [23], and Lorenz and Binkowski [24]. The resulting material was usually small in size and polycrystalline. Marushka and Tietjen in 1969 [25] were the first to report thin film growth of GaN over large areas (~2 cm² sapphire substrates). They used a hydride vapor phase technique with pure HCl and NH₃ gases and a H₂ carrier. The HCl reacted with metallic Ga to transport Ga-subchlorides to
the substrate for reaction with NH₃, forming electronic quality hexagonal GaN. Room
temperature optical absorption measurements correctly estimated the GaN direct bandgap
to be 3.39 eV at room temperature. All as-grown films were characterized by strong
n-type conductivity, which was speculated to be caused by nitrogen vacancies. Initial
attempts at p-type doping with Zn, Mg, Hg, Si, and Ge proved unsuccessful, and n-type
behavior was retained in all cases.

Although the inability to produce p-type doping in GaN would continue for nearly
2 decades, successful light emitting devices were demonstrated by Pankove et al. in the
early 1970s using metal-(Zn-doped insulator)-(n-type semiconductor) structures [26].
Hot electron injection from the n-type GaN layer was speculated to cause impact
ionization of luminescent centers in the insulating film. These devices could be tuned
over a broad spectral range by varying the Zn content in the cathode. Other reports
demonstrated luminescence from GaN/GaN:Mg and metal-Si₃N₄-GaN structures [27,28].
However, since these were MIS structures, very large electric fields were necessary to
produce radiative recombination, leading to devices much less efficient than conventional
p-n junction light emitters. Deep levels often introduced a yellow component (~550 nm)
to the light output, making the emission a greenish-blue color.

Despite the initial flurry of activity associated with the results of Marushka and
Pankove, the difficulty in achieving reproducible p-type doping led many groups to
abandon research in III-nitrides completely. Furthermore, the material quality was
generally very poor and adequate Ohmic contacts had not yet been developed for GaN.
Wide bandgap II-VI materials such as ZnS and ZnSe moved to the forefront of blue,
green, and UV optoelectronic device research. Much of the success achieved by today’s
III-N researchers is due to the work of Professor I. Akasaki at Noyaga University in Japan. Their group was one of the few that persisted with wide bandgap nitride research, and in the 1980s made two profoundly important contributions. The first of these was the development of a low temperature (LT) AlN buffer layer for GaN growth by metalorganic chemical vapor deposition (MOCVD). The resulting 2-step growth technique (Figure 2-1) greatly improved GaN epilayer quality, despite the 3-dimensional island growth of the AlN buffer films. The large nucleation surface area was shown to enhance lateral growth of the subsequent GaN, improving crystal quality, increasing uniformity, enhancing luminescence, and decreasing the background doping to $\sim 10^{17}$ cm$^{-3}$. The success of the AlN buffer is highlighted by noting that some variation of this technique is still employed by most modern III-N crystal growers.

![Temperature profile during MOCVD growth illustrating use of LT-AlN strain relief layer](image)

Figure 2-1. Temperature profile during MOCVD growth illustrating use of LT-AlN strain relief layer [29].

The donorlike nature of native defects and the lack of a shallow acceptor precluded the growth of reproducible p-type material until 1988, nearly 20 years after the first demonstration of thin GaN films. The “discovery” of p-type conductivity was made
somewhat accidentally by Amano and co-workers [30] while investigating cathodoluminescence of MOCVD-grown GaN:Mg in a scanning electron microscope (SEM). Their as-grown films were highly resistive, but showed a dramatic increase in luminescence efficiency and decrease in resistivity after exposure to the low energy electron beam (a technique they called low energy electron beam irradiation, or LEEBI). Post-LEEBI Hall measurements indicated distinct p-type behavior in these films to a depth corresponding to the penetration depth of the electron beam. With the availability of high quality epilayers and both n- and p-type material, the first p-n junction LED was demonstrated in 1989 [30]. This device employed 5000 Å of LEEBI p-GaN grown on 3 μm of UID (2 × 10^{17} cm^{-3}) n-GaN with aluminum top and side contacts. The p-n junction device showed a reduced turn-on voltage and a much greater near-bandedge electroluminescence efficiency compared to similar devices fabricated with insulating GaN:Mg layers.

In the 1990s, Nakamura and colleagues at Nichia Chemical Company advanced the III-N light-emitting technology at a pace unmatched before or since by any academic, industrial, or research group in the world. Although Akasaki’s LEEBI technique was successful in activating Mg-doped GaN, Nakamura et al. demonstrated that the same effect could be realized by thermal annealing at temperatures >750°C in a hydrogen-free ambient—a much simpler process [31]. Similar anneals in NH₃ showed the activation to be reversible, confirming the role of hydrogen as the passivating agent. In 1994, Nichia became the first company to take III-N LEDs to the commercial market, later followed by Cree and Hewlett Packard. The Nichia devices had remarkably high brightness despite a dislocation density >10^{10} cm⁻³ in the active region. Such a dislocation density would
easily preclude any operational light emitting devices from conventional III-V materials. The mechanism for optical emission from such highly defective III-N material has still not yet been explained in detail.

Nitride laser diodes (LDs) were also pioneered at Nichia by leveraging the experience base attained during the development of III-N LEDs. These devices always utilized an InGaN active layer due to certain difficulties associated with GaN active region devices. Since Nakamura et al. demonstrated the first pulsed operation of a III-N LD in 1996 [32], several other groups have reported fabrication of similar devices. Again, commercialization was first achieved by Nichia. Continuous wave (CW) operation has since been demonstrated, with recent estimated room temperature CW operation of >10,000 hours at an output power of 2 mW [33]. It is interesting to note that the Nichia structures are normally grown on a GaN buffer, unlike the LT buffer proposed by Akasaki and co-workers.

The most recent evolution of III-N technology has been its application to electronic devices. The amount of effort currently devoted to research-level exploration of III-N electronics is staggering. Demonstrations of GaN-based BJTs, HBTs, Schottky rectifiers, MESFETs, JFETs, HFETs, and MISFETs have appeared in the literature. Although there are currently few (if any) commercially available products, this will undoubtedly change in the near future. Silicon carbide MESFETs from Cree Research, Inc. are beginning to infiltrate the low frequency RF market with 12 W output power and 11 dB gain at 48 V and ~3 GHz. It is expected that III-N devices will eventually surpass the power handling capabilities of SiC, and extend the frequency of high power solid-state RF modules to Ku-band or beyond.
2.2 Crystal Structure

The III-nitrides can exist in a hexagonal (wurtzite) or a cubic (zincblende) polytype, each shown in Figure 2-2. The wurtzite structure (α-GaN) consists of interpenetrating hexagonal close packed (HCP) sublattices, each consisting solely or gallium or nitrogen atoms. Likewise, the zincblende crystal structure (β-GaN) is formed by interpenetrating body-centered cubic (BCC) cation and anion sublattices. The bandgap of β-GaN is slightly smaller than that of hexagonal GaN. Wurtzite GaN is by far the most common of the 2 polytypes, although the metastable zincblende phase has been grown on certain cubic substrates [34]. Zincblende AlN has also been reported from growth on β-SiC [35]. All subsequent discussions of GaN, AlN, or Al<sub>x</sub>Ga<sub>1-x</sub>N layers in this work specifically refer to wurtzite material due to the profusion of this polytype compared to β-GaN.

Table 2-1. Material Properties of AlN, GaN, and InN [14]

<table>
<thead>
<tr>
<th></th>
<th>GaN</th>
<th>AlN</th>
<th>InN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lattice Constant (Å)</td>
<td>a = 3.189</td>
<td>4.52</td>
<td>a = 3.112</td>
</tr>
<tr>
<td></td>
<td>c = 5.185</td>
<td></td>
<td>c = 4.982</td>
</tr>
<tr>
<td>Energy Gap (eV)</td>
<td>3.39</td>
<td>3.2</td>
<td>6.28</td>
</tr>
</tbody>
</table>

* theoretical
Figure 2-2. Crystal structures of III-nitrides: a) zincblende and b) wurtzite. The hexagonal wurtzite structure is the most common for electronic quality (Al,Ga,In)N.

2.3 Thin Film Growth

2.3.1 Substrates for III-N Epitaxy

Without question, the factor most responsible for hindering device development is the unavailability of an appropriate substrate material with thermal compatibility and a close lattice match to GaN. The lattice constants of the III-nitrides are, in general, much smaller than typical III-V semiconductors such as phosphides or arsenides, as illustrated in Figure 1-2 and Table 2-2. The hexagonal crystal structure of α-GaN serves to exacerbate this problem. The ideal substrate would be GaN itself, cut from a bulk sample, which could be used for homoepitaxy with perfect lattice and thermal expansion matching. However, bulk GaN growth is difficult to achieve in large scale due to the
Table 2-2. Material properties of GaN and candidate substrate materials for heteroepitaxial III-N growth.

<table>
<thead>
<tr>
<th>Material</th>
<th>Crystal Structure</th>
<th>Lattice Parameters (Å)</th>
<th>Coeff. of Thermal Expansion $(\Delta a/a, \Delta c/c \times 10^6 \text{ K}^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-GaN</td>
<td>Hexagonal</td>
<td>a = 3.189, c = 5.185</td>
<td>5.59, 3.17</td>
</tr>
<tr>
<td>β-GaN</td>
<td>Cubic</td>
<td>4.52</td>
<td>5.2</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>Hexagonal</td>
<td>a = 4.758, c = 12.991</td>
<td>7.5, 8.5</td>
</tr>
<tr>
<td>4H-SiC</td>
<td>Hexagonal</td>
<td>a = 3.08, c = 10.07</td>
<td>--</td>
</tr>
<tr>
<td>6H-SiC</td>
<td>Hexagonal</td>
<td>a = 3.08, c = 15.12</td>
<td>4.2, 4.7</td>
</tr>
<tr>
<td>3C-SiC</td>
<td>Cubic</td>
<td>4.3</td>
<td>--</td>
</tr>
<tr>
<td>AlN</td>
<td>Hexagonal</td>
<td>a = 3.104, c = 4.966</td>
<td>4.2, 5.3</td>
</tr>
<tr>
<td>MgAl$_2$O$_4$</td>
<td>Cubic</td>
<td>8.083</td>
<td>7.45</td>
</tr>
<tr>
<td>ZnO</td>
<td>Hexagonal</td>
<td>a = 3.2496, c = 5.2065</td>
<td>2.9, 4.8</td>
</tr>
<tr>
<td>Si</td>
<td>Cubic</td>
<td>5.4301</td>
<td>3.59</td>
</tr>
<tr>
<td>GaAs</td>
<td>Cubic</td>
<td>5.6533</td>
<td>6.0</td>
</tr>
<tr>
<td>LiAlO$_2$</td>
<td>Tetragonal</td>
<td>a = 5.1687, c = 6.2679</td>
<td>7.1, 7.5</td>
</tr>
<tr>
<td>LiGaO$_2$</td>
<td>Orthorhombic</td>
<td>a = 5.402, b = 6.372, c = 65.407</td>
<td>6.0, 7.0</td>
</tr>
</tbody>
</table>
high dissociation pressure of nitrogen and necessary growth temperatures from 1200-1600°C. Sapphire (0001) is widely used as the substrate for heteroepitaxial GaN growth due to its relatively low cost, availability in large diameters, and stability at typical process temperatures [36]. However, GaN growth on Al₂O₃ suffers from a lattice mismatch of about 14%, resulting in a high density (10⁸-10¹⁰ cm⁻²) of threading dislocations in the epitaxial films, as shown in Figure 2-3. Also, the low thermal conductivity of sapphire (κ = 0.5 W/cm·K) is highly undesirable for devices operating at elevated temperatures since heat cannot be effectively dissipated through the substrate.

Alternative substrate materials such as Si, GaAs, InP, GaP, SiC, ZnO, MgO, MgAl₂O₄, NdAl₂O₄, NdGaO₃, ScAlMgO₄, LiGaO₂, and LiAlO₂ have been investigated [37-51], with varying levels of success. The lattice and thermal mismatch data for several of these potential substrate materials are summarized in Table 2-2. Of these, SiC has shown perhaps the most promise for electronic device applications. SiC, itself a wide bandgap semiconductor, has a better lattice match to GaN than sapphire (0001). In addition, SiC is an exceptional thermal conductor (κ ~ 3.3 - 4.9 W/cm·K), with the exact κ value depending on the polytype and doping. The most common SiC polytypes for nitride growth are 4H and 6H. N-type, p-type, and semi-insulating substrates are commercially available, with diameters up to 3 inches.

LiGaO₂ (001) and LiAlO₂ (100) are intriguing substrate candidates due to their small lattice mismatch to GaN. LiGaO₂ (LGO) is an orthorhombic crystal with hexagonal mismatch of <1%, while LiAlO₂ (LAO) is tetragonal with −1.4% mismatch between (11 2 0) GaN and (100) LAO [48]. Although these are promising features, very little work has yet been done to characterize nitride growth on these materials.
Bulk GaN growth by Czochralski or Bridgeman techniques from stoichiometric melts has been largely unsuccessful due to the very high melting temperature and high vapor pressure of N$_2$ above GaN. Such growths have been conducted with a specialized growth apparatus at temperatures of ~1600°C and nitrogen overpressures of ~15 kbar [52,53]. Small single crystals of GaN with dislocation densities ~10$^5$ cm$^{-3}$ have been produced, but the dimensions of the largest samples are still < 1 cm$^2$. The high background doping in these samples has been decreased by Mg incorporation. Interestingly, this has also been shown to improve crystal quality as measured by x-ray diffraction [54]. The fundamental drawbacks to this technique are cost and sample size. Both issues must be addressed for this method to gain widespread use for bulk III-N growth.

An interesting approach to bulk GaN growth involves hydride vapor-phase epitaxy (HVPE) of GaN on Al$_2$O$_3$ and post-growth removal of the sapphire to create a free-standing quasi-substrate [55-57]. This approach was used to grow bulk templates described in greater detail in Chapter 6. For near-term availability of bulk GaN samples, such VPE techniques appear promising. However, hydride growth techniques are typically characterized by higher background impurity concentration and lower crystal quality than films grown by MOCVD or MBE. Both of these aspects will need to be improved before device-quality material is produced by this technique. Recent reports [58,59], as well as the results presented in Chapter 6 of this work, indicate that the quality of HVPE nitride layers is continually progressing.

It is far from clear if there is a specific “substrate of the future” for III-N electronic and optoelectronic devices. Much work remains for those involved in the
crystal growth of these materials. Many creative approaches to dislocation reduction, including epitaxial lateral overgrowth (ELO), pendeoepitaxy, and cantilever epitaxy, have been demonstrated in the literature [60-68]. In GaN ELO, a patterned SiO₂ mask is used to regrow GaN material in regions above oxide window. The mechanism of coalescence of the overgrown nitride is the topic of active research [69], and is a function of the mask pattern size and shape as well as growth conditions. Threading dislocations tend to propagate to the surface of the overgrown nitride in regions above the window, but are greatly reduced in regions above the mask. Upon convergence of the lateral growth fronts, material of improved electrical and optical quality can be obtained.

Figure 2-3. Transmission electron microscopy (TEM) image of GaN film on sapphire illustrating high density of threading and mixed dislocations.
Growth by conventional or specialized techniques on common, non-lattice matched substrates (e.g. silicon) should not be discounted, since this would decrease substrate cost, increase substrate size, leverage the silicon processing technology, and allow direct integration of nitride devices with silicon ICs or other components.

2.3.2 Doping

The impetus for most of the early work in GaN was its potential as a short wavelength light-emitter. Typically, semiconductor light-emitting devices rely on a forward-biased p-n junction to create radiative electron-hole recombination, resulting in the emission of a photon. Obviously, for such devices to be fabricated, both n- and p-type conductivity must be possible from the host material. Bipolar electronic devices, such as bipolar junction transistors (BJTs) or heterostructure bipolar transistors (HBTs) also require (by definition) both n- and p-type layers.

In the earliest work on electronic quality GaN thin films, Marushka and Tietjen noted strong n-type conductivity, with electron concentrations from $1 - 5 \times 10^{19}$ cm$^{-3}$ [25]. They speculated that nitrogen vacancies were probably responsible for the unintentional doping. Later reports challenged this hypothesis, suggesting that the nitrogen vacancy (while donorlike) possesses a rather high energy of formation, precluding a sufficient concentration to account for the observed n-type conductivity [70]. Gallium interstitials [71] as well as silicon substituting on a gallium site [72] and oxygen substituting on a nitrogen site [72,73] have been suggested to play a role in this behavior. Improvements in crystal quality and decreased impurity incorporation have led to material with greatly reduced background doping.
Silicon doping has been demonstrated to provide reproducible intentionally n-type films. The solubility of Si in GaN is excellent, various MOCVD precursors are readily available, and films with electron concentrations from $\sim 10^{15}$ (UID) to $\sim 10^{20}$ cm$^{-3}$ have been demonstrated [74-76]. The Si concentration in MOCVD-grown films appears to be proportional to the Si/Ga ratio of the source gases [77]. The activation energy for this shallow donor is still under investigation, but appears to be in the range 10 - 30 meV [77,78]. Germanium doping with GeH$_4$ has also been investigated [79], but the incorporation efficiency is much lower than for silicon.

P-type doping has proven especially difficult for the III-nitrides, partially due the donorlike nature of native defects. Divalent species such as Zn, Cd, Mg, Be, and Ca [80-83], as well as Group IV elements such as carbon [84], have been investigated as potential p-dopants. Many of these effectively compensate the background electron concentration, but low resistivity p-type behavior is usually not obtained. It has been speculated that the strong ionic nature of the GaN crystal plays some role in the difficulty with p-doping. As with ionic wide bandgap II-VI semiconductors such as ZnS, valence band hybridization effects may lead to repulsion of certain acceptor species [85].

Magnesium has been the shallowest acceptor level and the most successful p-type dopant in GaN. However, the properties of p-GaN are still far from ideal, and achieving high hole concentrations with Mg as the dopant has proven difficult. The Mg acceptor is located rather deep in the energy gap, approximately 170 meV above the valence band edge. This leads to films with only $\sim 1\%$ of the incorporated acceptors actually contributing to the p-type conduction (i.e. Mg incorporation needs to be approximately 2 orders of magnitude higher than the desired hole concentration). This effectively places
an upper limit on the obtainable hole concentrations in Mg-doped films. For highly
doped films, the Mg content can approach 1% of the concentration of the host species,
well above the typical dopant content of other III-V semiconductors. In addition, Mg-
doped GaN films are extremely sensitive to hydrogen exposure. It was shown that the
LEEBI technique resulted in the dissociation of hydrogen from Mg:H complexes in the
as-grown material [86]. Without the compensating hydrogen, the Mg acceptors are able
to contribute to the hole conduction. Nakamura, using activation by thermal annealing,
demonstrated the reversibility of this process by re-compensating GaN:Mg films upon
annealing in hydrogen ambients [87].

Aside from the deep acceptor level, one problem specific to magnesium doping
by MOCVD is the "memory" effect. This effect is characterized by the inability to
abruptly "turn off" the amount of Mg present in the vicinity of the growing film. In
MOCVD growth, it is typically possible to abruptly increase or decrease doping levels by
modulating the flow of the dopant source to the chamber. This allows the crystal grower
to accurately place p-n junctions, δ-doped layers, or other features required by the device
structure. In the case of MOCVD Mg-doping, cyclopentadienly magnesium (Cp₂Mg) is
typically used as the Mg source. Cp₂Mg is a low vapor pressure crystalline solid at room
temperature, complicating its transport to the reaction chamber. Turn-on is usually
abrupt, but the turn-off is not well controlled, as illustrated in Figure 2-4. This is due to
the relatively low vapor pressure of Mg, making it easily "stick" to chamber walls during
growth. When Mg is desorbed from the chamber walls, it can be incorporated into the
film despite the fact that the Mg source may have already been turned off. The memory
effect is particularly troublesome in the growth of p-n heterostructures, such as for
AlGaN/GaN heterojunction bipolar transistors (HBTs). For an npn HBT, the memory effect can place the p-n junction in the wide bandgap emitter, instead of at the AlGaN/GaN interface. The resulting emitter homojunction eliminates the advantages associated with the conduction band offset of npn HBTs. Growth techniques such as RF-MBE are attractive alternatives for device structures where accurate p-n junction placement is critical.

![Diagram of Mg incorporation during MOCVD growth of p-GaN](image)

Figure 2-4. "Memory" effect of Mg-dopant during MOCVD growth of p-GaN.

Highly doped p-type films are necessary for Ohmic contact layers on devices such as LEDs, LDs, and HBTs. Although Mg doping has allowed the fabrication of commercially viable p-n light emitting III-N devices, the poor Ohmic contacts that result from low concentrations of activated Mg acceptors has been perhaps the most limiting factor in the development of III-N bipolar electronic devices. Continued efforts to obtain a shallow acceptor level, increase Mg activation, or to demonstrate novel techniques such as superlattice doping or co-doping will hopefully help overcome many of the limitations currently associated with p-GaN films.
2.4 Etching

A commonly cited advantageous property of the III-N material system is chemical stability. The ability of the nitrides to withstand harsh chemical environments is attractive, for example, in many military applications. This property is due to the large bond strength of the nitrides, which is approximately 11.5, 8.9, and 7.7 eV/atom for AlN, GaN, and InN, respectively [9]. Unfortunately, this chemical stability also greatly complicates wet etching. The most systematic study of wet chemical etching was reported in 1997 [88]. In this work, the etch behavior of ~30 potential etchants (acidic and basic) was investigated at room temperature for AlN, GaN, InN, InAlN, and InGaN. No wet etch was found for GaN, InN, AlGaN, or InGaN. It was shown that KOH-based solutions were effective at etching AlN, with etch rates strongly dependent on crystal quality and solution temperature [89]. Other reports indicate a very slow GaN etch rate in ~50% NaOH/H2O (~20 Å/min.) although these samples were grown on GaAs substrates [90]. Molten KOH or H3PO4 has been shown to preferentially etch defects and dislocations in GaN [59,91-93]. While this is useful for materials characterization, it does very little to assist with device processing, where smooth etched morphology and materials selectivity are needed and the formation of etch pits is undesirable. The general lack of a suitable wet etch for the III-nitrides has led to a great deal of work to develop dry etch techniques for electronic and optoelectronic device processing steps such as mesa steps, recess etches, and facet formation. Alternative methods, such as photoenhanced wet chemical etching and laser ablation have also been investigated.

2.4.1 Photoenhanced Wet Chemical Etching

The dissolution kinetics of an etching process may be affected by illumination with light above the bandgap of the semiconductor. This type of illumination creates
electron hole pairs (EHP), which can participate in the overall etch mechanism. The first report of photoelectrochemical (PEC) wet etching of GaN was given by Minsky et al. [94]. Their experimental setup consisted of an electrochemical cell with front contact to the GaN surface and 325 nm illumination with a He-Cd laser. Dilute KOH and HCl solutions were investigated, with etch rates as high as several μm/min. for KOH. The authors speculated that the hole-assisted oxidation of the GaN surface contributed to the enhanced etching behavior, in a manner analogous to GaAs [95]. Similar reports using different etchants and illumination sources were given by other groups [96-98]. Youtsey et al. extended the work of Minsky and postulated the overall oxidation reaction:

$$2\text{GaN} + 6\Theta \rightarrow 2\text{Ga}^+ + \text{N}_2\uparrow$$ (2.1)

where Θ represents a photogenerated hole. They support this mechanism by noting the presence of bubbles during the etch process, consistent with the release of nitrogen gas. Cho et al. investigated the UV illumination intensity, KOH concentration, and solution temperature during PEC etching [99]. From their results, a maximum was found in the etch rate vs. KOH molarity curve for 450 W Hg lamp illumination. An activation energy of 0.8 ± 0.3 kcal/mol was extracted from an Arrhenius plot, indicating diffusion-limited etching. Also, Auger electron spectroscopy (AES) indicated that the etched films remained near-stoichiometric as compared to control samples.

Nearly all the reported PEC etch studies have investigated n-type GaN, with little work on intrinsic or p-type material. For p-GaN, Youtsey et al. reported no photoenhanced etching [100]. This is usually explained by the inability to confine holes at the surface of the semiconductor. Yang et al. have reported effective photoenhanced etching of p-GaN by applying a negative external bias to the substrate during etching.
The bias voltage effectively causes downward band bending at the semiconductor surface, allowing hole accumulation and photoenhanced etching in a manner analogous to n-type material.

Other recent work involves the development of PEC techniques that do not require physical contact to an external electrochemical cell. Bardwell et al. have described a process using $S_2O_8^-$, added to conventional KOH solutions [101,102]. The peroxydisulfate ion absorbs photons with $\lambda < 310$ nm, and produces highly oxidizing ions or radicals, depending on the pH of the solution. Photogenerated electrons are necessary to sustain the reaction, and for this reason dislocations are not readily etched, since these act as electron-hole recombination centers. Nonetheless, various device structures have been fabricated using this technique, including HFETs [103].

2.4.2 Dry Etching

No wet etch (conventional or PEC) has proven capable of fabricating smooth, well-controlled, anisotropic features with sufficiently fast etch rates for III-N device processing. In addition, while wet chemical etching has excellent selectivity for etching one material over another, it is generally unacceptable for maintaining good pattern transfer at dimensions below ~2 μm. These difficulties have led to extensive research in the development of III-N dry etch processing. Dry etch processes are characterized by chemical or physical mechanisms, or by a combination of both. For chemical etches, the reactant gases typically adsorb onto the semiconductor surface and undergo a surface reaction. The volatile etch products desorb and are swept away by the vacuum system. This type of etch is many ways analogous to wet etch processes, and material selectivity is often possible by an appropriate choice of reactant gases. However, as with wet etches,
anisotropic profiles are difficult to attain. Physical etches occur when energetic ions bombard the surface of the semiconductor and transfer enough energy to cause bond breaking. The sample is physically sputtered away, making selectivity more difficult to attain. Furthermore, such etches can cause significant damage to the semiconductor via the generation of point defects or preferential loss of certain species from the surface.

There are 3 major dry etch techniques that have been explored for the nitrides: reactive ion etching (RIE), electron cyclotron resonance etching (ECR), and inductively coupled plasma etching (ICP). The fundamental difference between RIE, ECR, and ICP etching is the method of plasma generation. RIE is the simplest of the 3, with a simple 13.56 MHz RF power applied between parallel electrodes. The substrate is placed directly on the powered electrode, and energized ions from the plasma gas(es) are accelerated toward the etching surface. Typical ion energies are in the hundreds of eV.

Anisotropic profiles can be obtained with RIE due to the physical component of the etch, but significantly more etch damage can also occur, for the same reason. Another problem with RIE is the low ion density (typically <10⁹ cm⁻³) and the associated difficulty with initial bond breaking in III-N materials [104].

High-density plasma sources such as ECR and ICP are attractive due to their effective decoupling of plasma density and ion energy. The high plasma density (2-4 orders of magnitude higher than typical RIE) assists with bond breaking and etch product desorption [105]. ECR plasmas are formed by passing current through permanent magnets to induce a strong magnetic field and create electron resonance at 2.45 GHz. Inductively coupled plasmas are generated by applying an RF bias to inductive coils encircling a dielectric vessel. The alternating electric field creates a magnetic field in the
vertical plane and confines electrons in a high-density plasma near the center of the chamber. In both ECR and ICP, an RF bias is superimposed on the sample electrode to control ion energy. Low pressures are typically used to increase the mean free path and enhance anisotropic etching. Both ECR and ICP platforms have been extensively investigated for III-N etch development. Chlorine-based chemistries have been most widely studied, due to the higher volatility of III-chlorides compared to other III-halides and hydrocarbons.

The earliest work on ECR etching of III-nitrides was conducted by Pearton and co-workers using BCl3, CCl2F2, and CH4/H2 chemistries [106]. Reasonable etch rates were achieved for AlN, GaN, and InN at low pressures and moderate bias levels. Higher substrate bias was necessary for AlN in CCl2F2, likely due to the formation of AlFx. For the Cl-containing chemistries, the Group III etch products were determined to be the respective III-chlorides, while the nitrogen was speculated to form NCl3 and CClN.

Vartuli et al. investigated alternative ICl/Ar and IBr/Ar chemistries in both ECR and ICP plasmas and reported very smooth etch morphologies and little preferential loss of nitrogen, but with etch rates slower than those achieved with Cl-based chemistries [107,108]. Selectivities of 5-10 for GaN over InN, AlN, or InAlN were reported for ICl/Ar, with selectivities much lower for IBr/Ar.

The effects of chamber pressure and Ar and BCl3 additions on Cl2 ICP etch characteristics were studied by Lee and co-workers [109]. They used 600W ICP power, 120 V DC self-bias, 70°C substrate temperature, and chamber pressures from 10-30 mtorr. The etch rate of GaN increased for Ar and BCl3 levels to up 10%, but decreased for higher additive content. An increase in chlorine radical density was speculated as the
cause for the increased etch rates, which were as high as 8500 Å/min with selectivity of 3.7 over SiO₂.

At present, Cl-based ICP seems to have prevailed as the preferred high-density III-N etch technique for reasons including easier scale-up, improved plasma uniformity over large areas, lower cost and power requirements, and lack of electromagnets and waveguides necessary with ECR [5,105]. Other techniques, including chemically-assisted ion beam etching (CAIBE) [110], reactive ion beam etching (RIBE), and laser ablation have been applied to the III-nitrides, with varying levels of success. In the near-term, ICP etching will likely remain the enabling technology for pattern transfer during III-N device processing. However, development of a simple, controllable wet etch with reasonable etch rate could greatly decrease the utilization of dry etch techniques for the nitrides.

Despite the success of high density plasma etch techniques, dry etch-induced lattice damage can severely degrade material and device properties. In GaN, the damage may occur as lattice defects, unintentional passivation, or preferential N₂ loss producing surface non-stoichiometry. This is an especially troublesome issue for III-N bipolar electronic devices, such as BJTs and HBTs, where typically a dry etch step is required to uncover the p-base. Etch damage can decrease the hole concentration or completely reverse the surface conductivity type, depending on the extent of the damage. A reduced physical component of the etch, consistent with lower RF chuck power (ion energy), is needed to help minimize damage. For this reason, the low ion energy and increased plasma density of ECR and ICP are more attractive than RIE for such steps.
2.5 Metallization

Metallization schemes, or contacts, are critical to the performance of semiconductor devices. A good contact not only provides the desired electrical characteristics, but must also be mechanically and thermally stable during device processing and operating conditions. There are 2 main categories of metallization schemes: Ohmic and rectifying (Schottky). Ohmic contacts allow electrons or holes to flow freely into or out of the underlying semiconductor--i.e. they should present a very low impedance to current flow. The I-V characteristics of an Ohmic contact are linear, thereby obeying Ohm’s law. Ideal Schottky contacts present a very low impedance when forward biased beyond the turn-on voltage (\(V_{on}\)) and present a very high impedance when reverse biased. Therefore, the I-V curve for an ideal Schottky contact is very steep (large slope) for \(V > V_{on}\), but very flat (near-zero slope) for reverse bias values up to breakdown. Reverse breakdown occurs at the onset of avalanche multiplication and is characterized by an abrupt increase in reverse current flow. Reverse breakdown is usually accompanied by irreversible damage to the device. The behavior of ideal Ohmic and Schottky contacts is illustrated in Figure 2-5.

The wide bandgap of GaN makes contact formation somewhat challenging. To understand the fundamental mechanisms involved with contact formation, it is necessary to understand the band alignment of a metal/semiconductor junction, as shown in Figure 2-6. This figure is drawn for an n-type semiconductor and illustrates the band bending at the semiconductor surface that arises when the metal and semiconductor are brought into intimate contact and allowed to reach equilibrium (equality of Fermi levels, \(E_F\)). For the n-Ohmic contact, a metal with a small work function (\(\phi_m\)) is needed, where the work
function is defined as the energy from the metal Fermi level to vacuum. Depending on the value of $\phi_m$, the metal can induce a slight upward band bending or a very small Schottky barrier in the semiconductor, leaving majority electrons free to pass through the metal/semiconductor interface. An Ohmic contact is formed when the current-voltage characteristics are linear. A good Ohmic contact has a resistance much smaller than the intrinsic device resistance.

For the Schottky contact of Figure 2-6(b), a large metal work function induces negative band bending and creates a conduction band energy barrier, $\phi_B$, known as the Schottky barrier height of the contact. For an unpinned semiconductor Fermi level without image force lowering, the barrier height is given by [111]:

$$\phi_B = \phi_m - \chi$$

(2.2)

where $\chi$ is the electron affinity of the semiconductor. This barrier height is typically on the order of 1 eV. Hence, the current-voltage behavior of the ideal Schottky contact in Figure 2-6(b) can easily be understood as follows: For forward bias conditions (positive potential applied to the Schottky metal), electrons in the conduction band initially experience an energy barrier preventing them from entering the metal contact. As the forward bias voltage induces sufficient band bending to overcome the built-in potential of the Schottky contact, electrons are free to move from the
Figure 2-5. Current-voltage characteristics for (a) Ohmic and (b) Schottky contact. The Ohmic contact provides a low resistance path for electrons or holes to enter or exit the semiconductor. The Schottky contact allows large forward currents, but blocks reverse current flow up to the breakdown voltage.
Figure 2-6. Energy band diagrams for (a) n'-Ohmic, (b) n-Schottky, and (c.) n'-Ohmic contacts.
semiconductor to the metal. In the reverse bias regime, the band bending becomes more pronounced as the applied bias becomes more negative, and no current flows as the depletion region widens. Reverse breakdown occurs when avalanche multiplication takes place in the depletion region, leading to the onset of large current flow.

When the semiconductor doping density is sufficiently high, the $E_F$ lies very near the conduction band. In this case, equilibrium band bending takes place very near the semiconductor surface, leading to "thin" energy barriers as illustrated in Figure 2-6(c). Such structures allow quantum mechanical tunneling of electrons through the energy barrier and into the metal, providing the basis for a potential low-resistance Ohmic contact. This metallization strategy is common in AlGaAs/GaAs HEMTs, where a thin cap layer of $n^+$-GaAs is typically grown to facilitate Ohmic contact formation.

2.5.1 Ohmic Contacts

2.5.1.1 n-type

Aluminum was one of the first metals investigated for Ohmic contact formation to n-GaN, due to its low work function ($\phi_m = 4.3$ eV) and widespread use in semiconductor contacts and interconnects. While Al produced Ohmic behavior [112] at room temperature, a Ti/Al bilayer scheme was shown to provide superior contact resistance. Lin et al. reported a study of Au, Al, Ti/Au, and Ti/Al layers deposited on MBE-grown GaN and annealed in a temperature range from 500 - 900°C [113]. A minimum specific contact resistivity ($\rho_c$) of $8 \times 10^{-6}$ $\Omega \cdot \text{cm}^2$ was obtained for Ti/Al annealed at 900°C for 30 sec. These results suggested several features of the Ti/Al contact scheme. First, the work functions of Ti and Al are nearly the same, implying that a chemical mechanism is likely responsible for the improved contact resistance of the Ti/Al bilayer as compared to the
single-layer Al contact. Possible explanations are the gettering of nitrogen from the near-surface region of the semiconductor, removal of the native oxide, or a combination of both. For the former, Jenkins and Dow have shown that N vacancies exhibit donorlike behavior in GaN [114]. Thus, nitrogen gettering via TiN_x formation would serve to increase the electron concentration at the surface, thereby enhancing tunneling probability. In the latter case, thin native oxide layers at the metal/semiconductor interface have been shown to increase the contact resistance, since these oxides typically act as insulators. The Ti may diffuse through this native oxide to make intimate contact with the underlying GaN. Alternatively, the strong oxygen affinity of Ti may help pull the oxide away from the semiconductor surface during the formation of a discontinuous Ti- or mixed metal-oxide phase. The results of Lin and co-workers also suggest the importance of Al in the Ti/Al bilayer, since the Ti/Al contact exhibited electrical properties superior to the Ti/Au contact. This implies the formation of a Ti-Al intermetallic phase important to the electrical behavior of the contact. It should be noted that numerous metals other than Ti have used in bilayer schemes with Al, including Pd, Ta, Nd, Sc, and Hf [9]. Ohmic behavior typically results, but none have produced contact resistance values as low as Ti/Al. Numerical values obtained for annealed Ti/Al contacts deposited on properly cleaned n-GaN surfaces are \( R_c \sim 0.5 \, \Omega \cdot \text{mm} \) and \( \rho_c \sim 10^{-6} \, \Omega \cdot \text{cm}^2 \). Of course, these values depend on the doping density in the underlying semiconductor and the device structure.

It is desirable to cap the Ti/Al contacts with an oxidation-resistant metal such as Au to avoid excessive oxidation of the contact, especially during elevated temperature operation. To prevent excessive reaction between the Au layer and the Ti/Al bilayer, a
diffusion barrier is also included. Therefore, most modern n-ohmic metallization schemes employ annealed Ti/Al/x/Au, where x may be Ti, Ni, Pt, or Pd. However, the explicit roles of the diffusion barrier and cap layers are still not well understood. Cross-sectional transmission electron microscopy (XTEM) has shown the existence of Au at the GaN surface of Ti/Al/Pt/Au contacts, suggesting that the Ni diffusion barrier is not effective at blocking Au diffusion into the contact during annealing. In fact, Ni, Pd, and Pt have been shown to be permeable to Au during typical contact processing [115]. Their role in contact formation is not yet well-understood. Thermodynamic calculations may also be helpful in determining the phases present after high temperature annealing.

One issue unique to III-N n-ohmic contact formation is the extremely high annealing temperatures required to minimize contact resistance. Typical annealing conditions are ~800°C, 30 sec. in a N₂ ambient or overpressure. This very high annealing temperature is well above the melting point of Al (661°C), and typically leads to rough ohmic contact morphology. The edge definition of the patterned contact is also affected by the roughening, and edge acuity may become an issue for very small device features, such as the channel of an FET. A tradeoff between electrical properties and surface morphology seems to exist, with smoother morphologies available at lower annealing temperatures while sacrificing Rₑ. The Ti/Al ratio may play some role in this tradeoff, with reports of optimum Ti/Al ratios having been given in the literature [116]. TEM investigations of Ti/Al/Ni/Au contacts to AlGaN have shown that the mixed phases present at any given spatial position of the post-annealed contact are a strong function of the choice of materials (i.e. diffusion barrier) and the thermal history. Thus, a simultaneous optimization of Ti/Al thicknesses, Ti/Al ratios, diffusion barrier metal,
diffusion barrier thickness, annealing time, annealing temperature, and annealing ambient must be considered with respect to electrical properties and contact morphology.

2.5.1.2 p-type

Low resistance ohmic contacts to p-GaN have proven much more difficult to fabricate than those on n-type material. This difficulty is related to the wide bandgap of GaN and the large ionization energy of the Mg acceptor. For p-type semiconductors, the Fermi level is near the valence band, and the work function is much larger than for n-type. Alignment of metal and semiconductor Fermi levels therefore requires a metal with a larger \( \phi_m \). Due to the wide energy bandgap (\( \chi = 4.1 \) eV and \( E_g = 3.4 \) eV), typical p-GaN workfunctions are \( \approx 7 \) eV. There are no metal work functions this large, resulting in an equilibrium valence band energy barrier. The lack of highly-doped p-type films exacerbates this problem, since tunnel barriers are not available to increase hole conduction. In addition to these inherent materials-related limitations, certain device processing steps (such as dry etching) may cause preferential loss of N from the GaN surface. This is of particular importance in devices such as BJTs and HBTs, where a dry etch step is usually required to uncover thin p-base layers.

The most common p-ohmic metallization for GaN has been Ni/Au [117]. As with n-type contacts, a plethora of contact schemes have been investigated. A partial list of these includes Pd, Au, Ti, Al, Zn, Cu, Mo, Ta, Nb, Mo, V, W, Pt/Au [118], Ru/Au [119], Nb/Au [120], Pt/Ni/Au [121], Ni/Pd/Au [122], Ti/Pt/Au [123], Ni/AuZn [124], Cr/AuZn, Pd/Mg/Pd/Au [125], and Pd/Ag/Au/Ti/Au [126]. Ohmic behavior has been achieved, but p-contact resistance is still a limiting feature of both electronic and optoelectronic devices. In many cases, slight rectifying behavior is observed. Even if the
contact displays linear I-V characteristics, exceedingly high contact resistance can prove detrimental to device operation. For unipolar electronic devices such as FETs, the source and drain contact resistances are parasitic extrinsic circuit elements which increase self-heating and decrease the speed of the extrinsic device. For III-N photonic devices such as laser diodes (LDs), the situation can be even worse. Due to the resistive local heating and vertical nature of threading dislocations in GaN, thermally-enhanced migration of contact metal along dislocation lines can lead to device short-circuiting. For bipolar electronic devices such as npn HBTs, high base contact resistance (combined with the high sheet resistance) severely limits the output power of the device. The vertical current flow in these devices also makes them susceptible to the lifetime issues mentioned for LDs.

Although the Ni/Au contact has been widely discussed in the open literature, the mechanism of p-GaN/Ni/Au Ohmic contact formation is still far from clear. Generally, it is noted that annealing in an oxygen-containing ambient is necessary to achieve Ohmic behavior, implying the importance of a NiO phase. Ho et al. proposed a favorable band alignment between p-GaN and a thin interfacial layer of NiO [127]. Kiode and co-workers suggested a different mechanism, in which the oxygen anneal getters hydrogen and increases the hole concentration in the near-surface region of the semiconductor [128]. Qiao et al. recently reported a thorough study relating the electrical characteristics to the contact formation mechanism by depositing films of Ni, Ni/Au, Au, and NiO/Au [129]. They found that NiO/Au contacts (formed by annealing a thin Ni layer in air at 500°C for 5 min. then redepositing Au) did not produce Ohmic behavior, which seems to disprove the favorable band alignment mechanism proposed by Ho. Kiode’s hypothesis
of O gettering was supported by the work of Qiao et al., and it seems clear that oxygen plays some role in layer reversal of the as-deposited Ni/Au contact. Rutherford Backscattering Spectroscopy (RBS) data indicated a p-GaN/Au/NiO structure after annealing in air. The Au layer at the interface is not likely a continuous phase, with NiO, Ni-Ga-Au intermetallics and mixed oxides also present.

2.5.2 Rectifying Contacts

2.5.2.1 n-type

Rectifying contacts on n-GaN have also been extensively investigated, due to their importance in Schottky diodes and as FET gates. Experimental Schottky barrier heights show a clear dependence on metal work function, indicating an unpinned surface Fermi level. This is illustrated in Figure 2-7, where it is clear that metals with large work functions should be chosen as n-GaN Schottky contacts to maximize barrier height. These data are in sharp contrast to other common III-V semiconductors such as GaAs and InP, where the barrier height is largely independent of $\phi_m$ due to a large density of surface states. The degree of scatter in the Figure 2-7 data is likely due to variable GaN epilayer quality, since the growth and processing technology in the III-N material system is still in its early stages. Factors potentially responsible for the wide range of $\phi_B$ reported in the literature include surface roughness, surface preparation, local stoichiometry variations, or thermal history [130]. The presence of high concentrations of defects or dislocations has been shown to affect the I-V behavior of GaN Schottky diodes and rectifiers [131].

To date the largest barrier heights of $\sim$1.1 eV have been realized with Pt ($\phi_m = 5.6$ eV) contacts. Near-ideal Pt/GaN and Au/GaN Schottky diodes have been fabricated with
n = 1.04 [132] and 1.03 [133], respectively, defined from the forward current density expression for a Schottky contact:

\[
J_F = A^*T^2 \exp\left(\frac{q\phi_B}{kT}\right) \left[\exp\left(\frac{qV}{nkT}\right) - 1\right]
\]

(2.3)

where \( J_F \) is the forward current density, \( A^* \) is Richardson’s constant, \( q \) is the electronic charge, \( \phi_B \) is the barrier height, \( V \) is the applied bias, \( n \) is the ideality factor, \( k \) is Boltzmann’s constant, and \( T \) is the absolute temperature. For an ideal Schottky contact, \( n = 1 \). Reported values for GaN Schottky diodes range from near-ideal to > 4.

![Schottky barrier height versus metal workfunction](image)

Figure 2-7. Schottky barrier height versus metal workfunction for various contacts reported in the literature. The trend of increasing \( \phi_B \) with increasing \( \phi_m \) suggests the surface Fermi level is unpinned in GaN.

Although Pt has demonstrated a slightly larger, more consistent barrier height on n-GaN, many state-of-the-art AlGaN/GaN HEMTs use Ni/Au as the gate contact. Nickel
Schottky contacts exhibit comparable electrical properties, and it has been found that Ni adheres slightly better to the (Al)GaN surface than Pt. The improved adhesion facilitates easier lift-off of the patterned gates and improves yield.

2.5.2.2 p-type

There have been very few reports of rectifying contacts to p-GaN. The well-documented difficulties in achieving reproducible p-type crystal growth and Ohmic contact formation have precluded extensive study of p-type contacts and devices, especially FETs.

2.5.3 Surface Treatment

Even with optimized metallization recipes, proper surface cleaning is imperative to achieve intimate contact between metal and semiconductor. It is well known that the GaN surface forms a thin Ga$_2$O$_3$ native oxide when exposed to air. The presence of native oxides or residual organic residues can increase the contact resistance of Ohmic contacts and can lead to nonidealities in rectifying contacts. In addition, an atomically smooth, clean surface is imperative for devices requiring regrowth steps. Examples include LEDs and LDs employing epitaxial lateral overgrowth (ELO), HBTs with regrown emitter, or MISFETs. For all of these, the success of the metallization or regrown epilayers is largely determined by the quality and cleanliness or the semiconductor surface. A typical surface clean may include organic solvent treatments, acid etches, and rinsing in deionized water. It has been shown that ex-situ HCl and HF are successful in removing a large amount of the surface oxygen and that a UV/ozone clean can greatly decrease the amount of surface hydrocarbon [134]. The amount of oxygen reduction is much less for AlN surfaces, presumably due to the larger bond strength of the Al-O bond relative to Ga-O. This is of practical importance for Ohmic
contact formation to AlGaN/GaN HEMTs, where the surface layer typically contains ~20-30% Al. *In-situ* thermal or plasma treatments are also viable options for contaminant desorption prior to crystal growth, but are less feasible with typical metal deposition equipment.

Although a minor impact of surface treatments on the electrical properties of GaN contacts has been reported by some groups [135], other groups have reported dramatic improvements in contact properties, particularly p-type, with sulfur-based treatments [136]. Ammonium sulfide, (NH₄)₂S, has been credited with effective passivation of the p-GaN surface and modification of the Fermi level, reducing the valence band barrier height. Such surface treatments have previously been applied to AlGaAs/GaAs microdisk lasers [137]. Specific contact resistivities in the low- to mid-10⁻⁵ Ω⋅cm² have been reported for (NH₄)₂S-treated p-GaN samples [138]. These results are noteworthy, but at this point require further study.

### 2.6 III-N Optoelectronics

Direct bandgap semiconductors allow efficient radiative recombination of electrons and holes because their conduction band minimum and valence band maximum lie at the same position in momentum space (k-space). The photons emitted during a recombination process have an energy equal to the bandgap of the semiconductor and a wavelength $\lambda = E_g / (2\pi \cdot h)$, where $E_g$ is the semiconductor bandgap and $h$ is Planck’s constant. III-nitrides are wide direct bandgap semiconductors with energy gaps spanning the electromagnetic spectral range from visible (InN: 1.9 eV) to far-UV (AlN: 6.2 eV) (Figure 1-2). The wide direct bandgaps lead to exciting possibilities for short wavelength solid state light emitting devices, such as light emitting diodes (LEDs) and laser diodes
(LDs). The super-bright blue III-N LEDs can be incorporated with standard III-V red and yellow emitters for flat panel displays with color mixing across the visible spectrum. Traffic lights from LEDs deliver reduced power consumption and longer lifetimes than conventional bulbs. Also, by introducing a phosphor material above the LED, blue light from InGaN LEDs can be absorbed and re-emitted as white light. Such technologies promise to challenge incandescent sources for residential and outdoor lighting applications in the foreseeable future. The market for non-communication LEDs is currently ~$2.1 billion and is expected to reach over $3 billion in less than 2 years. Analogously, non-communication laser diodes currently enjoy a $1 billion market, which is expected to double in the next 2 years [139]. Communications, printing, and data storage are important commercial technologies that will benefit from continued development of III-N LDs. Current solid-state lasers for such applications are AlGaAs or InGaAsP alloys, emitting in the infrared. Since the minimum spot size for reading or writing data scales with $\lambda^2$, the resolution of laser printers and storage capacity of CDs or DVDs would be greatly increased by UV-emitting III-N LDs. This market alone is expected to generate $87$ million in 2003. Other optoelectronic applications for III-nitrides include solar-blind UV detectors, which take advantage of the wide bandgap of AlGaN alloys in detectors sensitive to wavelengths less than 300 nm while insensitive to visible and IR solar background. Such detectors are useful, for example, in military and defense applications.

2.6.1. Light Emitting Diodes

The demonstration, development, and commercialization of blue and ultraviolet p-n junction LEDs have all taken place within the past 12 years. This rapid progress is a
testament to the both the uniqueness of the material system and the massive research efforts that have been devoted to it by various academic, corporate, and governmental research labs. Although available from relatively few vendors, the GaN LED market has grown from $0 before 1994 to $420 million in 1999 (Figure 2-8). III-N LEDs offer tremendous advantages in output power, quantum efficiency, and luminous intensity over competing materials systems such as SiC, GaP, or (Zn,Cd,Mg)(S,Se,Te). The indirect bandgaps of SiC and GaP result in devices with significantly reduced brightness (~10 - 20 mcd) and quantum efficiency (~0.1%) compared to direct bandgap III-N devices.

Although green II-VI LEDs have been fabricated with reasonable brightness, the short lifetimes and propagation of crystal defects in II-VI semiconductors have precluded their widespread commercialization.

![GaN LED Market Revenue (1995 - 1999)](image)

Figure 2-8. GaN LED revenue for 5-year period ending 1999. The annual compound growth rate is nearly 80%.
The first demonstration of a InGaN p-n junction LED was given in 1989 [30]. Shortly thereafter, Nakamura and co-workers at Nichia Chemical Co. reported an LED emitting at 430 nm with a power level approximately one order of magnitude higher than comparable SiC devices [140,141]. In effect, this began Nichia’s dominance of nitride LED and LD research and development in the 1990s. The structure of a Nichia high-brightness blue LED, first fabricated in 1993, is given in Figure 2-9(a). The device contains an AlGaN/InGaN/AlGaN double heterostructure (DH), with 15% Al content in the cladding layers. The as-grown sample was annealed at 700°C in N₂ to activate the Mg acceptor. Ti/Al and Ni/Au metallizations were employed as n-ohmic and p-ohmic metallizations, respectively. The In₀.₀₆Ga₀.₉₄N active layer was co-doped with Si and Zn and resulted in emission near 450 nm, probably due to impurity-assisted recombination. The FWHM at peak wavelength was ~70 nm. For such devices, a blue shift of 10 - 15 nm was typically observed by increasing the forward current from 100 μA to 20 mA. The output power and external quantum efficiency of these devices were 3 mW and 5.4% at a forward current of 20 mA [142,143]. The brightness of the AlGaN/InGaN/AlGaN blue DH LEDs was ~2.5 cd, comparable to red AlGaAs LEDs.

Quantum well structures were developed by Nakamura and co-workers to extend the III-N LED wavelength range to >500 nm. In order to increase the emission wavelength of InGaN LEDs, increased In incorporation is necessary in the active region of the device. However, due to the lattice and thermal mismatch between InN and (Al,Ga)N, as well as thermodynamic considerations, high quality InGaN epilayers with high indium content have been difficult to achieve in practice. Thin InGaN quantum layers (~30Å) allow elastic strain accommodation between the well(s) and barriers,
reducing misfit dislocations due to strain relaxation. It is believed that such dislocations in the active region play a role in reducing the quantum efficiency of light-emitting devices, such as LEDs. Nakamura et al. demonstrated single quantum well (SQW) p-AlGaN/InGaN/n-InGaN/n-AlGaN LEDs and demonstrated electroluminescence (EL) from 450 nm (blue) to 590 nm (yellow). A difference of ~170 meV was observed between the peak emission wavelength and the unstrained value of the energy gap. This difference was attributed to quantum size and exciton effects of the active layer, although the details of the mechanism were not elucidated. To improve the output power and spectral width of these devices, a new p-AlGaN/InGaN/n-GaN SQW structure was investigated, as shown in Figure 2-9(b). Ultimately, this became the standard structure for high-brightness LEDs from blue to orange. Typical peak wavelengths and FWHM of the EL signal were 450 and 20 nm (blue) and 520 and 30 nm (green). Although the FWHM becomes larger at longer wavelengths, these devices offer much narrower emission bands than previously reported DH LEDs. The output power and external quantum efficiency of these devices at 20 mA forward current were 5.1 mW and 9.1% (blue) and 3 mW and 6.3% (green). The green device displayed a record luminous intensity of 12 cd with 10° cone viewing angle at 20 mA forward current. The output power, quantum efficiency, and luminous intensity of the green device are higher than any previous LED at a similar wavelength from AlInGaP, GaP, or ZnTeSe.

Red InGaN LEDs have been fabricated, but still suffer from poor InGaN crystal quality and phase separation due to high In-content. Presently, AlGaAs LEDs offer superior performance. The incorporation of InGaN blue, InGaN or GaP green or yellow, and AlGaAs red LEDs allows near-complete coverage of the chromaticity diagram, as
illustrated in Figure 2-10. The Nasdaq MarketSite in Times Square, New York City—the largest flat panel display in the world—is an impressive manifestation of such technology (Figure 2-11). Continued advancements in InGaN blue and green LEDs are projected, and white LED packages for general lighting applications are eagerly anticipated in the near future. Modules capable of 100 lm/W luminous flux conversion efficacy with total output of 50 lm have been predicted [139].

Corporate research and development dominates the current activity in III-N LEDs, with Nichia, Cree, Toyota Gosei, OSRAM Opto, and LumiLeds among the major participants. The current state-of-the-art is still set by Nichia, offering a full product line of blue, green, red, and white LEDs, with output powers from 2 mW (red) to 6 mW (blue) and luminous intensities as high as 10 cd (green) for devices operational to 85°C.
Figure 2-9. Cross-sectional schematic of (a) AlGaN/InGaN/AlGaN double heterostructure blue LED and (b) single quantum well InGaN green LED.
Figure 2-10. Chromaticity diagram showing spectral coverage of currently available LEDs from various III-V materials systems. The circular regions in the center represent various shades of white light. Adapted from [144].
2.6.2 Laser Diodes

The announcement of the first current-injection blue laser diode in 1996 [145] was a monumental leap for III-N technology. Such a device has been coveted for nearly 30 years, and several reports of optically-pumped lasing from various III-N structures had appeared [146-149]. However, Nakamura’s InGaN multi-quantum well (MQW) structure was the first to achieve stimulated emission by pulsed current injection. These devices were grown by two-flow MOCVD and typical Ti/Al and Ni/Au ohmic metallizations were deposited on n- and p-type layers, respectively. Mirror facet formation was accomplished by RIE to define the stripe geometry lasers. Dry etching
was necessary due to the lack of cleavage planes along the (0001) sapphire face, exacerbated by the fact that the nitride columns are slightly rotated with respect to the hexagonal orientation of the substrate. Consequently, the morphology of the etched facet of the InGaN device was not particularly good (~50 nm roughness), necessitating the use of high reflection coatings. Still, the 417 nm output (FWHM = 1.6 nm) represented the shortest wavelength ever obtained from a semiconductor laser. The output power and threshold current density ($J_{th}$) of this device were 215 mW (at 2.3 A) and 4 kA·cm$^{-2}$, respectively, under pulsed current conditions at room temperature [145].

Subsequent reports described the fabrication of the first cleaved facet GaN-based laser [150] and growth and fabrication of LDs on closely-lattice matched MgAl$_2$O$_4$ substrates [151]. In a surprising announcement, room temperature continuous wave (CW) operation was reported less than 1 year after the first demonstration of pulsed current InGaN LDs [152-154]. The threshold current, turn-on voltage, and threshold current density were 80 mA, 5.5 V, and 3.6 kA·cm$^{-2}$, respectively. A CW output power of 1 mW was obtained at 405.83 nm. Estimated CW operation was >30 hours for a device with structure given schematically in Figure 2-12 [155]. The spontaneous and stimulated emission was attributed to excitons at deep levels originating from InGaN quantum dots [156].

These early LDs were typically characterized by extremely short lifetimes due to p-contact spiking along dislocation lines, shorting the p-n junction. This was, in part, caused by the poor p-ohmic contacts, which led to significant local heating and high turn-on voltage (34 V for the first Nichia device). The high threshold current density also contributed to significant heat generation in the active layer. Improvements to CW
lifetime and operating voltages were addressed by 2 design changes, namely the introduction of modulation-doped strained-layer superlattices (MD-SLS) and epitaxial lateral overgrowth (ELO) [157]. The SLS structure allowed for thermal and lattice mismatch stresses to be accommodated elastically, reducing cracking commonly associated with thick AlGaN cladding layers. The modulation doping helped reduce the operating voltage. Epitaxial lateral overgrowth (ELO) is a growth technique which uses selectively masked regions for subsequent regrowth of low dislocation density GaN (see Section 2.3.1). The regrown GaN nucleates homoepitaxially and grows laterally over the masked regions, which do not allow nucleation. The growth conditions are chosen such that lateral growth proceeds faster than vertical growth, and the film eventually coalesces above the mask regions. Threading dislocations can be reduced below $10^6$ cm$^{-2}$ in the regions above the mask. The window regions typically retain a relatively high dislocation density ($10^7$ - $10^9$ cm$^{-2}$) In addition, the coalescence front may contain lattice misregistry, tilt boundaries, and/or anti-phase domains. Namamura and co-workers applied this growth technique to fabricate devices above both masked and window regions. The devices above the window region had a threshold current density of 6-9 kA·cm$^{-2}$ while those above the mask exhibited values as low as 3 kA·cm$^{-2}$, clearly suggesting the role of threading dislocations as leakage paths, increasing $J_{th}$. The structure of a LD incorporating the InGaN MD-SLS structure and ELO substrate is given in Figure 2-13. Room temperature CW operation of 2,200 hours was verified, and estimates of >10,000 hours were extracted from the degradation speed, \( \frac{dI}{dt} \) (mA/100 h), where I is the operating current of the device [158].
Currently, Nichia retains its leadership role in commercially-available short wavelength LDs, with a product line consisting of 30 mW single and multiple transverse mode devices at $\lambda = 405$ nm. These high power lasers are expected to find applications in optical storage in the very near future. At least 20 groups have demonstrated functional blue LDs, with Cree and Toyota Gosei currently among the closest to challenging Nichia in the commercial arena.

Figure 2-12. Structure of CW InGaN laser diode, with detailed band diagram of active region quantum wells [155].
2.6.3 Detectors

In addition to light-emitting devices such as LEDs and LDs, III-N semiconductors present exciting possibilities for use in visible and ultraviolet (UV) photodetectors. Much of the interest for such devices stems from military and defense applications such as missile detection. Other potential uses include ozone monitoring, flame detection, chemical analysis, and aerospace communications. Many of these end-uses require operation in extreme radiation, high temperature, or chemical environments. The III-nitrides are well-suited for such harsh conditions, owing to their wide bandgap and chemical stability. The potential exists for detectors with a large spectral coverage (200 - 650 nm) utilizing the entire (In,Ga,Al)N system. Devices capable of operation near 280 nm are especially attractive, due to an ozone absorption band in the solar spectrum. The stray radiation at this wavelength (that would constitute background noise in a
photodetector) is effectively filtered out by the atmosphere. Devices operating in this energy range are known as “solar blind.” AlGaN alloys with ~50% Al content absorb in this region of the electromagnetic spectrum.

Numerous types of III-N photodetectors have been reported in the literature, including p-n junction, metal-semiconductor-metal, p-i-n, photoconductive, and transparent Schottky contact devices [159]. Avalanche photodiodes are not easily implemented in the III-N system due to the materials’ inherent high breakdown fields [9]. Of the reported devices, transmission cutoff is typically very abrupt, and noise and dark current characteristics are in some cases even lower than for Si detectors [160]. Detector response times have decreased to the nanosecond range for the best devices [161].

2.7 III-N Electronics

In addition to its 3.4 eV bandgap, GaN possesses attractive material properties which warrant the tremendous attention recently given to nitride-based electronic devices for high temperature, high power, and high frequency applications. Due to the early research emphasis on light-emitting devices, the technology of GaN electronics lags behind that of LEDs and LDs. However, a recent surge in interest has led to promising results and vigorous active research in GaN bipolar and unipolar devices. In addition, III-N electronics research has been able to leverage much of the processing technology developed for LEDs and LDs. Demonstration of GaN-based MESFETs, JFETs, HFETs, MISFETs, BJTs, HBTs, Schottky diodes, and p-i-n rectifiers have all appeared recently in the literature.
2.7.1 Field Effect Transistors

The first III-N electronic device was a GaN MESFET fabricated by Khan and co-workers in 1993 [162]. Although there have been numerous reports in the literature of MESFETs with reasonable performance [163,164], the most noteworthy advantage of the III-nitrides over other wide bandgap semiconductors, such as SiC, is the availability of AlGaN/GaN heterostructures. High quality heterostructures with up to 50% Al content have been demonstrated by a variety of growth techniques, including MOCVD and MBE. Remarkable progress has been made in recent years in high performance heterostructure field effect transistors (HFETs, also known as MODFETs or HEMTs†) grown on a variety of substrates including sapphire, SiC, and Si [165-172]. A typical HEMT structure is given in Figure 2-14.

The extremely high current density achievable with AlGaN/GaN heterostructures (up to 1.6 A/mm has been reported) is a result of the large polarization-induced field and large conduction band offset in the AlGaN/GaN system. This polarization field has both a piezoelectric, strain-induced, as well as a spontaneous polarization component. For the AlGaN HEMTs most widely studied to date, the spontaneous polarization field dominates. Even in the absence of modulation doping, this built-in field induces a two-dimensional electron gas (2DEG) that is linearly proportional to the Al-mole fraction, $x$ ($0 \leq x \leq 1$) as $n_s = x \cdot (5 \times 10^{13} \text{ cm}^{-2})$. For the ~30% Al-comparison most widely studied, a

† Modulation doped field effect transistors (MODFETs) are actually a subset of HFETs, utilizing selective barrier doping to spatially separate ionized donors from the electrons in the 2DEG. This spatial separation leads to an increase in channel mobility and, for this reason, these devices are also known as high electron mobility transistors (HEMTs). Due to polarization effects in the III-N system, undoped HEMTs are possible, and typically the terms HFET, MODFET, and HEMT are used interchangeably in the nitride literature. The convention adopted for this work will be to use the term ‘HEMT’ to describe all heterostructure FET devices.
channel sheet electron density of $\sim 1.5 \times 10^{13} \text{cm}^{-2}$ can be realized that is a factor of 5 - 10 times higher than typical GaAs or InP pHEMTs. The associated mobility at this high current density is in the range of 1000 - 1500 cm$^2$/V·s, which is well below that achieved in the conventional III-V material. However, the combined $\mu_e n_s$ product remains competitive. It is the $\mu_e n_s$ product that enables this material system to be applicable to low noise amplifiers as well as power amplifiers.

Figure 2-14. Generic HEMT device structure, illustrating modulation doping; 2-D electron gas; source, gate, and drain contacts; and mesa isolation.

The high voltage capability of AlGaN/GaN HEMTs is the result of a critical electric field of $\geq 3 \text{ MV/cm}$. This value is 10 times that of Si and 5 times that of GaAs. For typical AlGaN/GaN HEMT structures, this translates to gate-to-drain breakdown voltages of approximately 100 V/μm. While record breakdown voltage over 500 V have been demonstrated for large gate-to-drain spacing, the more important result is for a device layout consistent with good microwave performance (i.e. gate-to-drain spacing of
1-2 μm). In this latter case, a breakdown voltage of ~80 - 100 V can be maintained in a microwave transistor that also has the high channel current mentioned previously. Figure 2-15 clearly illustrates these advantages compared to conventional high-speed compound semiconductor technologies. Due to the high breakdown field, III-N HEMTs are able to maintain large breakdown voltages in smaller geometry (i.e. short channel) devices that retain excellent high-speed performance.

To make full use of the high voltage capability, a transistor must not be thermally limited. GaN also has an advantage over GaAs in this regard, with thermal conductivity up to 2.0 W/cm·K for GaN versus 0.46 W/cm·K for GaAs. Furthermore, GaN HEMTs can be grown on semi-insulating SiC with a thermal conductivity of 3.3 W/cm·K. The high breakdown field and good thermal conductivity allow these devices to be used in high efficiency (theoretical efficiency of 78%) class B push/pull amplifiers at full power rating. GaAs microwave devices, on the other hand, can only be implemented in a class B push/pull amplifier by backing off the voltage bias (and hence the power level) to accommodate the higher voltage swing in this configuration as compared to class A, single-ended operation.

Over the past few years, dramatic progress has been made in understanding the AlGaN HEMT device physics and in demonstrating record microwave power performance. The potential of any microwave device technology is first realized in small periphery devices with microwave power density being the relevant figure of merit. These results are summarized in Figure 2-18, where state-of-the-art output power densities at X-band are shown as a function of time. Note the steady progress in power
Figure 2-15. Breakdown voltage vs. unity-gain cutoff frequency for various compound semiconductor device technologies, illustrating high frequency, high voltage capability of the III-N material system. After [177].
performance since the mid-1990s. The best results to date are 9.8 W/mm with 47% power added efficiency (PAE) at 8 GHz [173]. This power density is close to seven times the best result obtained with any GaAs technology and over twice that achieved with SiC MESFETs at this frequency [174]. Compilations of reported power density and PAE from AlGaN/GaN HEMTs and GaAs pHEMTs are given in Figures 2-17 and 2-18, respectively. Power added efficiency values are comparable between the two technologies, but the power densities achievable with AlGaN/GaN HEMTs are over an order of magnitude higher. With continued improvements in material quality and device design, a power density of 12 W/mm or higher appears plausible for AlGaN/GaN HEMTs. Other impressive results include 6.6 W/mm with 35% PAE at 20 GHz [175]. The 20 GHz performance clearly illustrates the potential for high power III-N electronics at K-band and beyond.

It has also been predicted and shown that AlGaN/GaN HEMTs can achieve low microwave added noise figures (NF = 0.6 dB at 10 GHz) while maintaining a large breakdown voltage (>60 V) and hence a large dynamic range [176]. These results imply that these devices can be used to perform the active transmit and receive functions in more robust, higher dynamic range modules.
Figure 2-16. Output power density of AlGaN/GaN HEMTs at X-band (8-12 GHz) as a function of demonstration time. The data points represent values reported in the literature. Note the superiority of III-N HEMTs over GaAs and SiC. Adapted from [178].

Figure 2-17. Reported output power per unit gate periphery comparing GaN HEMT and GaAs pHEMT technologies. Note that GaN HEMT power densities are more than an order of magnitude higher for the frequency range 3 - 20 GHz.
Figure 2-18. Reported power added efficiency for frequencies from 0 - 50 GHz comparing GaN HEMT and GaAs pHEMT technologies. The GaN HEMT PAE values remain comparable, despite much larger output power density.

2.7.2 Bipolar Transistors

Group III-nitride bipolar junction transistors (BJTs) and heterostructure bipolar transistors (HBTs) are attractive due to their potential for high current handling, high breakdown voltage, and device linearity. Additional advantages for HBTs over FETs include increased threshold voltage uniformity, lower phase noise, and higher transconductance [179]. Applications include radar, satellite, and communications applications in the 1-5 GHz range, with the potential for operation up to ~30 GHz predicted as the device technology matures [180]. Despite the many attractive features, progress in III-N bipolar electronic devices has greatly lagged the tremendous successes enjoyed by AlGaN/GaN HEMTs due to difficulties associated with both growth and processing.
As discussed in Sections 2.3.2 and 2.5.1.2, the deep Mg acceptor leads to relatively low p-type doping concentrations in GaN. Combined with the large bandgap of 3.39 eV, low resistance p-type Ohmic contact formation is difficult. In n-p-n BJTs and HBTs, these difficulties are related with the p-base. In the base, it is desirable to have a relatively high doping concentration to increase hole injection. Decreasing the hole concentration leads directly to a decreased collector current, reducing gain. The poor base contact resistance and conductivity introduces a large parasitic element and is manifested by a common emitter offset voltage.

For n-p-n devices, the emitter mesa etch must stop on the buried p-base. In GaN, point defects due to dry etch-induced lattice damage tend to be donorlike in nature. Type conversion (i.e. p-type layers becoming n-type) of the exposed base surface is possible if sufficient damage is introduced during the etch. Such a requirement presents a difficulty specific to BJTs and HBTs, since LEDs and LDs are typically grown with buried n-layers and p-type surfaces. For LEDs and LDs, etch damage is not as critical, since increased donor doping actually aids n-ohmic contact formation. In contrast, low-damage dry etch conditions are critical for successful HBT processing. Increased recombination at the emitter-base junction may result from a damaging etch.

The first III-N HBT was demonstrated by Ren and co-workers in 1998 [181]. This device was grown by MBE on an MOCVD GaN buffer/sapphire substrate. The device epilayer structure is given in Figure 2-19. The emitter mesa etch was stopped ~100Å from the base to minimize damage. The remaining emitter material was oxidized and stripped in acid. The low hole concentration in the p-base (~low 10^{17} cm^{-3}) led to strong series resistance effects and small low bias current at room temperature. Device
performance improved when tested at 300°C, presumably due to increased thermionic emission and Mg ionization. At 300°C, a DC current gain of ~10 was demonstrated, as illustrated in the Gummel plot of Figure 2-20.

Extensive use of simulation and modeling has been employed to improve layer structure design and predict device performance [182-186]. More recent experimental efforts have focused on various approaches to reduce the undesirable effects of the p-base. An n-p-n AlGaN/GaN HBT with selectively regrown GaAs:C was investigated to reduce base contact resistance [187]. BJTs and HBTs with regrown emitter have been investigated due to the lack of a need for an emitter mesa etch [188,189]. Superlattice structures have been shown to increase the hole density of Mg-doped GaN by decreasing the effective ionization energy [190,191]. A GaN BJT employing a superlattice base and regrown emitter was demonstrated with a current gain of ~10 and operation at a power density of 150 kW·cm⁻² [189]. Some p-n-p devices have been investigated to circumvent the p-type base. However, this geometry suffers from the inferior transport properties of holes, leading to reduced switching speeds.
500 nm GaN Si-doped (8E18)
20 nm AlGaN graded to GaN Si-doped (8E18)
80 nm AlGaN Si-doped (8E18 and Al = 0.2)
20 nm GaN nm undoped
220 nm GaN nm Mg-doped (5E17)
400 nm GaN Si-doped (1E16)
500 nm GaN Si-doped (8E18)
GaN Buffer Layer

Figure 2-19. Epilayer structure of AlGaN/GaN npn HBT. The GaN buffer layer was grown by MOCVD and all subsequent layers were grown by MBE [192].

Figure 2-20. Gummel plot at 300°C for first AlGaN/GaN HBT, illustrating current gain of ~10 [181].
2.7.3 Rectifiers

GaN electronic device research has been largely dominated by FETs. MESFETs and HEMTs for high frequency, high power applications have been developed to exploit the attractive material properties of the III-nitrides, as described in Section 2.7.1. HBTs and BJTs are promising, but still suffer from the materials and processing-related issues discussed in Section 2.7.2. A III-N device which has received considerably less attention is the power rectifier. Applications for these devices are numerous and include radar, satellite systems, hybrid-electric vehicles, utilities transmission and distribution, and “smart” power modules.

One of the immediate applications for GaN power rectifiers stems from the electric power utility industry. A major problem in the current grid is momentary voltage sags, which affect motor drives, computers, and digital controls. Mechanical switches are presently used to control electric current flow across utilities transmission and distribution lines. Opening or closing these switches can lead to large voltage or inductance spikes delivered to the load. Such spikes may be detrimental, for instance, to major computing centers or other sensitive electronic equipment. An outage of less than one cycle, or a voltage sag of 25% for two cycles can cause a microprocessor to malfunction [9]. As a result of these potential fluctuations, the electric power grid must be operated at capacities well below its rated value, leading to reduced energy efficiency. A system for eliminating power sags and switching transients would dramatically improve power quality [193]. Solid state devices, if available, are expected to show “clean” switching and could potentially eliminate line transients and allow more efficient operation of the grid. Since typical power devices are required to operate at elevated temperatures due to the power dissipation associated with switching large currents and
voltages, wide bandgap materials are attractive due to their increased tolerance to temperatures above the limits of silicon. Reduction of bulky, expensive cooling equipment should be possible, leading to decreased system complexity and cost.

Motors consume up to 50% of the electricity produced in the United States each year [9]. Repair costs for these motors are estimated at $5 - 10 billion annually. These expenses are expected to be greatly reduced by high power electronic devices that permit smoother switching and control. In addition, control electronics could dramatically improve motor efficiency, reducing the amount of power consumption. Other end uses for III-N high power switches include lighting and HVAC systems [193,194].

Wide bandgap materials are attractive for use in power devices due to their ability to sustain extremely high electric fields, leading to large blocking voltages. The on-state resistance of a diode rectifier is related to the breakdown voltage and critical electric field of the semiconductor by:

\[ R_{on} \propto \frac{4V_B^2}{\mu_n e \varepsilon E_c^2} \]  

(2.4)

where \( V_B \) is the reverse breakdown voltage, \( \mu_n \) is the electron mobility, \( \varepsilon \) is the semiconductor permittivity, and \( E_c \) is the critical electric field for breakdown. From Equation 2.4, it can be seen that for a fixed value of breakdown voltage, materials with larger critical electric breakdown field have a smaller on-resistance (given comparable \( \mu_n \) and \( \varepsilon_s \)), leading to reduced power dissipation in the forward, conducting state.

Analogously, for a given value of \( R_{on} \), large \( E_c \) corresponds to large breakdown voltage.

There have been a number of reports of mesa and lateral geometry GaN and AlGaN Schottky and p-i-n rectifiers fabricated on heteroepitaxial layers on Al\(_2\)O\(_3\) substrates [195-202]. However, these results still fall well below the theoretical
predictions for the III-N material system [8,195,203]. Bandic et al. have predicted that a 50 kV blocking voltage should be sustained by 20 μm of Al$_{0.2}$Ga$_{0.8}$N doped to $1 \times 10^{16}$ cm$^{-3}$. Such performance has not been achieved in practice, most likely due to the highly defective nature of III-N material at present. It has been amply demonstrated in other materials systems such as SiC that the presence of defects (dislocations, nanopipes) leads to premature breakdown in diodes and rectifiers [204-206].
CHAPTER 3
AlGaN / GaN HIGH ELECTRON MOBILITY TRANSISTORS

3.1 Introduction

One of the most noteworthy advantages of the III-nitrides over other wide bandgap semiconductors, such as SiC, is the availability of AlGaN/GaN heterostructures. The Type I band alignment between AlGaN and GaN has been shown to form a potential well and a 2-dimensional electron gas (2DEG) at the heterointerface [165,207]. When these materials are brought into contact, thermal equilibrium requires alignment of their respective Fermi levels ($E_F$). This induces conduction ($E_c$) and valence ($E_v$) band bending in both the AlGaN and GaN layers and can cause the GaN conduction band at the interface to drop below $E_F$, as illustrated in Figure 3-1. Since (for n-type material) the Fermi level can be viewed as an electrochemical potential for electrons, majority electrons will accumulate in the narrow gap material just below the heterointerface to fill the quasi-triangular potential well between $E_c$ and $E_F$. These electrons are confined by a distance shorter than their deBroglie wavelength, causing quantization of the allowed energy levels in the potential well. Depending on the structure, there may be more than one allowed energy level below the Fermi level, although only the lowest allowed level will be substantially populated at room temperature. With the heterointerface on one side and a potential barrier on the other, electrons in the 2DEG are only free to move in along the plane of the interface. A thin ‘sheet’ of negative charge (electrons) results.
Modulation doped field effect transistors (MODFETs) are a class of heterostructure FET that use selective barrier doping to spatially separate ionized donors from the electrons in the 2DEG, leading to an increase in channel mobility. For this reason, these devices are also known as high electron mobility transistors, or HEMTs. A typical AlGaN/GaN HEMT structure is given in Figure 2-14. The AlGaN barrier layer is typically grown with the same Al content throughout, but with varying doping levels. A heavily doped donor layer supplies electrons to the 2DEG. The ionized donors remain in the AlGaN, while the electrons are transferred as mobile carriers into the GaN. An unintentionally doped (UID) AlGaN layer between the donor layer and the gate metal
allows formation of a Schottky contact to the wide bandgap barrier. A thin AlGaN spacer layer between the donor layer and channel (≈30 Å) is used to screen the Coulomb potential of the ionized impurities. Without this spacer, Coulomb scattering from positively charged impurities in the donor layer would lead to reduced electron mobility in the 2DEG.

AlN, GaN, and their alloys are polar crystals. Polarization occurs due to the lack of inversion symmetry for typical growth along the polar (0001) axis of the wurtzite crystal structure. The strong spontaneous and piezoelectric polarization effects in the III-nitrides can strongly influence the electron density and potential profile of heterostructure devices. [208,209]. The polarization field is such that sheet electron densities on the order of $10^{13}$ cm$^{-2}$ can be realized even in undoped HEMTs [210].

In this chapter, a thorough discussion of all aspects of AlGaN/GaN HEMT device fabrication, characterization, and small-signal modeling will be given. In Section 3.2, the details of device processing, as well as DC, small signal, and large signal device characterization will be presented for MOCVD-grown HEMTs on sapphire substrates. In addition, s-parameter modeling will be used to model the small signal performance and extract information on the intrinsic elements of the FET. In Section 3.3, a direct comparison of AlGaN/GaN HEMTs of identical structure grown on Al$_2$O$_3$ and SiC substrates will be given. In Section 3.4, the device performance of HEMTs grown on sapphire substrates by reactive molecular beam epitaxy will be discussed.
3.2 MOCVD-Grown HEMTs on Al₂O₃ Substrates

3.2.1 Device Processing

All HEMTs in this study were fabricated by a 4-step process sequence consisting of dry etch mesa isolation, Ohmic metallization, Ohmic anneal, and gate metallization. The mesa etch was performed in a load-locked Plasma-Therm SLR 770 with a 2 MHz, 3 turn coil ICP source. A conductive adhesive was used to mount the samples to an anodized Al carrier plate which was cooled from the backside by He gas. A 13.56 MHz rf bias was superimposed on the substrate. The clear field etch mask was photoresist AZP 4330, which was spun at 4000 rpm for 30 sec., baked at 90°C for 90 sec., exposed for 6.5 sec. at 20 mW/cm², and developed for ~125 sec. in MF319. Surface profilometry indicated a photoresist (PR) thickness of ~3.75 µm, which was highly reproducible across the surface of the sample. The ICP process gases and flowrates were 8.0 sccm BCl₃, 32.0 sccm Cl₂, 5.0 sccm Ar. The addition of BCl₃ to the Cl/Ar chemistry was important for maintaining smooth surface morphology of the etched surface [211]. The chamber temperature and pressure were maintained at 2 mTorr and 25°C, respectively. An ICP source power of 500 W and a substrate RF power of 40 W was used to minimize surface damage. By using high ICP and low RF powers, a high density plasma is created which maintains fast etch rates with reduced ion energy. The DC self bias was below -90 V under these conditions. Typical etch rates were ~1300 Å/min. Etch depths (mesa step heights) were 1300 - 1700 Å. The stability of the 4330 photoresist was excellent.

Ohmic lithography was performed with AZ 5214 PR in both positive tone (dark field mask) and image reversal (clear field mask). The linewidth resolution was approximately the same for both processes, while the thickness of the positive tone PR
was slightly larger (1.46 μm vs. 1.35 μm for 5000 rpm spin speed). A hexamethyl disilazane (HMDS) treatment and dehydration bake was used prior to PR application to improve PR adhesion. The PR procedure for positive tone was: 5000 rpm spin for 30 sec., bake 90°C for 90 sec. on hotplate, pattern expose ~6.5 sec. at 20 mW/cm², and develop ~110 sec. in 1:1.4 MF312:H₂O. Immediately before loading the patterned wafers for metallization, a 60 sec.1:10 HCl:H₂O dip was used for removal of the native oxide. Electron beam evaporation was initiated at chamber pressures of ~3 × 10⁻⁷ torr. An Ohmic metallization scheme of Ti/Al/Pt/Au (250/1000/450/1500Å) was deposited at rates from 2 - 10 Å/sec, depending on the metal. This quad-layer scheme was annealed in a rapid thermal annealing (RTA) furnace at 850°C for 30 sec. under a N₂ ambient. Four-probe transmission line method (TLM) analysis resulted in specific contact resistances from high 10⁻⁷ to low 10⁻⁶ Ω·cm² with transfer resistances of 0.3 - 0.4 Ω·mm. A typical TLM plot of resistance vs. contact spacing is given in Figure 3-2. From such a plot, sheet resistance, transfer resistance, and specific contact resistivity can be simultaneously determined. Figure 3-3 (bottom) illustrates the very rough morphology of the Ohmic metallization after the high temperature anneal. The top picture is unannealed Ti/Al/Pt/Au for comparison. Note, however, that the edge acuity of the contacts (which affects the channel length) appears to be quite good. This will be further illustrated by scanning electron micrographs in Section 3.2.3. Some reports claim a tradeoff between contact resistance and surface roughness [212], which can be tailored by varying layer thicknesses or annealing conditions. For this study, an Ohmic metallization experiment was performed in which Ti/Al/x/Au (250/1000/450/1500Å) was deposited on AlGaN/GaN HEMT structure, where x = Ni, Pt, or Ti. Each of the metallizations was
annealed at 800, 825, 850, 875, and 900°C for 30 sec. in N₂. Circular TLM patterns were used to construct Figs. 3-4 and 3-5 by averaging 3 sets of TLM data for each experimental data point shown. Note that the general trend is for decreasing contact resistance at higher annealing temperature. However, as temperatures were increased above ~850°C, the contact morphology significantly degraded. An 850°C-annealed Ti/Al/Pt/Au Ohmic metallization was chosen for all devices presented in this work, unless otherwise noted.

Optical (contact) gate lithography was carried out with AZ 5214 photoresist, in a process analogous to that used for the Ohmic level. Gate lengths for the optically patterned devices were 0.8 and 1.2 μm. Electron beam lithography was used to define submicron gates using a bilevel PR scheme of 6% PMMA (bottom) and 9% P(MMA-MAA) copolymer (top). The copolymer is more sensitive than the PMMA and therefore develops out wider than the PMMA resulting in a T-shaped cross section. Additionally, the copolymer may be selectively developed with respect to the PMMA by choosing an appropriately low dose that is sufficient to expose the copolymer but not the PMMA. Therefore, the ‘wing’ portion of the mushroom gate may be extended by placing lightly-dosed boundaries on either side of the main gate dose. The main gate boundaries were weighted in the CAD file and doses were optimized to achieve the desired gate lengths.

The 6% PMMA was spun at 6600 rpm for 30 sec. and baked a minimum of 15 min. on a hotplate at 170°C, which produced ~4000 Å resist. The 9% P(MMA-MAA), was spun at 4400 rpm for 30 sec. and baked a minimum of 15 min. on hotplate at 170°C for ~5500 Å resist. Patterns were exposed at 50 kV, ~4 nA beam current in 480 × 320 μm fields, with a 25 nm step distance. A typical large area dose was 225 μC/cm². Typical
gate doses ranged from 375 μC/cm² for 0.5 μm gates to 800μC/cm² for the 0.1 μm gates. For the Cascade Probe FET mask design used in this study, the write time was 27 minutes per unit cell (1 cm²). The majority of the write time was for the pads and device labels. The exposed PR was developed for 2 minutes in a 1:2 mixture of MIBK:IPA (methyl isobutyl ketone: isopropyl alcohol) and dried with N₂. Prior to metallization, a 1:10::HCl:H₂O oxide strip was performed. For optimum liftoff using this recipe, the thickness of the deposited metal should be between 4000 and 5500 Å. For all submicron devices presented in this study, gate metallization consisted of e-beam evaporated Ni/Au (200/4500 Å). A gentle acetone rinse peeled away the PR in one thin sheet. The gate yield was close to 100% (i.e. no gate metal adhesion problems).

Figure 3-2. Plot used to determine contact resistance by the 4-probe TLM method.
Figure 3-3. Ti/Al/Pt/Au Ohmic metallization before (top) and after (bottom) 850°C RTA.
Figure 3-4. Specific contact resistivity as a function of annealing temperature for Ti/Al/x/Au, where $x = \text{Ti, Pt, or Ni}$.

Figure 3-5. Transfer resistance as a function of annealing temperature for Ti/Al/x/Au, where $x = \text{Ti, Pt, or Ni}$.
3.2.2 Optical Gate Devices

The device structures were grown with conventional precursors at 1040°C in an impinging jet MOCVD system with rotating substrate. The growth sequence consisted of a thin AlN nucleation/strain relief layer followed by ~1.6 μm of unintentionally doped GaN, deposited at a growth rate of ~300 Å/min. The modulation doped cap layer consisted of 30Å UID Al0.2Ga0.8N, 200 Å Al0.2Ga0.8N:Si (n~1 x 10¹⁸ cm⁻³), and 100 Å UID Al0.2Ga0.8N. A schematic of the device structure is given in Figure 3-6 and an optical micrograph of a completed HEMT is shown in Figure 3-7.

The DC characteristics were measured in common source mode with an HP 4145A Parameter Analyzer. Maximum drain current densities of 0.5 A/mm were obtained for 0.8 x 100 μm² gate dimension devices, as shown in Figure 3-8. For the same devices, the extrinsic transconductance was ~135 mS/mm at V_G = -1.5 V and V_DS = 3 - 5 V (Figure 3-9). Slight self-heating effects were evident for the higher power levels associated with wider gate devices (i.e. 200 μm devices) biased to V_DS =10 V. The Schottky gate contacts exhibited leakage current on the order of 0.3 - 3mA.

Scattering parameters were measured with an HP 8510 Vector Network Analyzer calibrated via the short-open-load-thru (SOLT) method to ~40 GHz. The small signal current gain (h_{21}) and Unilateral gain (|U|) are shown in Figure 3-10. From this plot, a cutoff frequency (f_T) of 9.7 GHz and maximum frequency of oscillation (f_{max}) of 23.5 GHz were determined using -20 dB/decade extrapolation to the frequency axis.

Figure 3-11 is a summary of the microwave performance of 0.8 and 1.2 μm gate length devices with gate widths of 100, 150, and 200 μm. These data were taken from immediately adjacent devices in an effort to minimize the effects of material non-
uniformity. In general, measurements at various positions on the wafer provided similar results. Both \( f_T \) and \( f_{\text{max}} \) were larger for the 0.8 \( \mu \)m gate length devices than for 1.2 \( \mu \)m devices, although the magnitude of this difference was smaller for \( f_{\text{max}} \). However, the cutoff frequency-gate length product (\( f_T \cdot L_G \)) was \(-20\%\) lower for the 0.8 mm gate length devices. The \( f_T \cdot L_G \) product of \(-10\) GHz\( \cdot \mu \)m is comparable to results from similar devices in the literature [213,214], although values as high as 18.3 GHz\( \cdot \mu \)m have been reported [215]. From simple charge control theory:

\[
\nu_{\text{eff}} = 2\pi f_T L_G \quad (3.1)
\]

where \( \nu_{\text{eff}} \) is the effective electron velocity in the 2DEG and \( L_G \) is the gate length.

Equation 3.1 provides a convenient estimate of electron velocity if both \( f_T \) and \( L_G \) are known. Using the 1.2 \( \mu \)m gate length device data from Figure 3-11, \( \nu_{\text{eff}} \) is calculated to be \((6.2 \pm 0.1) \times 10^6\) cm\( \cdot \)s\(^{-1}\). Still, this is roughly a factor of 2 below the theoretical value predicted for this material system, clearly pointing toward the importance of reducing scattering mechanisms (such as interfacial roughness) during crystal growth. A possible explanation for the discrepancy between the \( f_T \cdot L_G \) product of the 0.8 and 1.2 \( \mu \)m devices is the value of \( L_G \) used in the calculation. Dimensions of 0.8 \( \mu \)m are very near the resolution limit for contact lithography using a standard Karl Suss MJB-3 aligner. It is possible the actual gate lengths were slightly larger than the nominal values, which would lead to both a smaller percentage difference between the 2 gate lengths and a larger calculated effective electron velocity (closer to the theoretical value).
Figure 3-6. Layer structure of MOCVD-grown AlGaN/GaN HEMT.

<table>
<thead>
<tr>
<th>Layer Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 Å UID Al_{0.2}Ga_{0.8}N</td>
</tr>
<tr>
<td>100 Å Al_{0.2}Ga_{0.8}N:Si, -1E18 cm^{-3}</td>
</tr>
<tr>
<td>30 Å UID Al_{0.2}Ga_{0.8}N</td>
</tr>
<tr>
<td>~200 Å 'transition' GaN</td>
</tr>
<tr>
<td>~1.7 μm UID GaN</td>
</tr>
<tr>
<td>~200 Å AlN</td>
</tr>
<tr>
<td>(0001) Sapphire</td>
</tr>
</tbody>
</table>

Figure 3-7. CCD image of optical HEMT with gate dimension given in device label (0.8 × 100 μm²). Source, drain, and gate contact pads are indicated.
Figure 3-8. Low voltage DC output characteristics of $0.8 \times 100 \ \mu m^2$ gate dimension HEMT

Figure 3-9. DC output characteristics of HEMT at 25°C.
Figure 3-10. Small signal characteristics of 0.8 × 100 μm² device illustrating 23.5 GHz \( f_{\text{max}} \).

Figure 3-11. Gate length and gate width dependence of \( f_T \), \( f_{\text{max}} \), and \( f_T \cdot L_G \) product.
3.2.3 Submicron Gate Devices

Submicron gate devices were fabricated using the processing sequence detailed in Section 3.2.1 on samples with 230 Å (sample NF1214D, Figure 3-6) and 430 Å (sample NF1110A, Figure 3-12) AlGaN layers (spacer + donor + cap). Scanning electron microscopy (SEM) was used to examine the quality of the lithographic processing steps and estimate the gate lengths of the submicron devices. Figure 3-13 shows an SEM micrograph of a 0.1 μm gate crossing the mesa. Note the use of an anchoring technique at the end of the gate to improve adhesion. The anchors were necessary for the 0.1 μm devices, but not for the longer gate lengths. A cross-sectional view of the same gate finger (0.1 × 100 μm²) is shown in Figure 3-14. Note the excellent edge definition on both the gate and the source/drain contacts. The surface morphology of the MOCVD-grown epilayers is also excellent, with no evidence of defects or tilted growth boundaries. Figure 3-15 shows an SEM micrograph of a 0.25 μm gate length device with the mesa step in the background. Once again the edge acuity of the metal is excellent. This micrograph gives a good view of the T-gate and shows that the lithography and lift-off processes produced well-defined gate contacts. From detailed SEM analysis, gate lengths of 0.12, 0.23, and 0.52 μm were measured for the nominal 0.1, 0.25, and 0.5 μm values, respectively.
Figure 3-12. Layer structure of MOCVD-grown HEMT.

<table>
<thead>
<tr>
<th>Layer Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 Å UID Al_{0.2}Ga_{0.8}N cap</td>
</tr>
<tr>
<td>200 Å Al_{0.2}Ga_{0.8}N:Si, ~1E18 cm³</td>
</tr>
<tr>
<td>30 Å UID Al_{0.2}Ga_{0.8}N spacer</td>
</tr>
<tr>
<td>~200 Å 'transition' GaN</td>
</tr>
<tr>
<td>~1.0 (\mu)m UID GaN buffer</td>
</tr>
<tr>
<td>~200 Å LT AlN</td>
</tr>
<tr>
<td>(0001) Sapphire</td>
</tr>
</tbody>
</table>

The maximum drain current density for the NF1110A devices approached 1 A/mm before self-heating effects caused a decrease in the output current. Characteristic drain I-V curves are given in Figures 3-16 (top and bottom) and 3-17 for 0.1, 0.25, and 0.5 \(\mu\)m gate length devices, respectively. Note the excellent pinch-off of the 0.25 and 0.5 \(\mu\)m devices. The 0.1 \(\mu\)m devices do not maintain complete pinch-off and exhibit significant output conductance due to short-channel effects. These are caused by the 2-dimensional nature of the electric field for short gate lengths and large drain bias, and are common in GaAs- and InP-based HEMTs. For the shortest gate length devices, the extrinsic transconductance exceeded 200 mS/mm. The transfer characteristics for each gate length are given in Figures 3-18, 3-19, and 3-20. The peak of the extrinsic transconductance curve shifted toward more positive voltage with increasing gate length.
and the magnitude of \( g_m \) scaled with inverse gate length (Figures 3-21 and 3-22). For these devices, high temperature operation with excellent pinch-off characteristics has been verified to 400°C, the temperature limit of the probe station heater. Drain current density decreases monotonically with increasing temperature over the range 25° - 400°C, as shown in Figure 3-23.

Maximum gate-source breakdown voltages were \( \sim 25 \) V for NF1110A and \( > 60 \) V for NF1214 devices with \( 0.25 \times 150 \mu m^2 \) gates measured in 3-terminal mode with source and drain shorted to ground (Figure 3-24). The gate leakage current for 1214D was \( \sim 200 \mu A \) at a reverse bias of -60 V. The leakage from 1110A devices was much higher (\( \sim 1 \) mA prior to hard breakdown). The reverse current density of the 0.25 and 0.5 \( \mu m \) gate length devices was similar, but leakage from the 0.1 \( \mu m \) device was significantly higher, as shown in Figure 3.25. This may be due to the higher dose (\( \sim 800 \mu C/cm^2 \)) during e-beam writing of the shorter gate lengths. This effect should be investigated in additional detail to optimize the submicron lithography process.

Gate current-voltage-temperature characteristics from 1214D were measured on a heated chuck from 25°C to 395°C. The gate leakage current decreased by nearly 3 orders of magnitude upon heating the device (in ambient air) from room temperature to 395°C (Figure 3-26). The turn-on voltage (\( V_{on} \)) also increased with temperature, as shown in Figure 3-27. Note that from 100°C - 150°C, \( V_{on} \) and the on-resistance (\( R_{on} \)) increases slightly. From 150°C - 200°C, the on-resistance is restored to approximately its initial value. From 200°C -300°C, there is very little change in the forward I-V characteristics. Above 300°, \( V_{on} \) begins to increase with no change in \( R_{on} \). This behavior is likely caused by the hypothesized e-beam-induced damage to the AlGaN surface. The damaged
regions under the gate contact are partially repaired upon heating. It is possible that the anomalous forward I-V characteristics are caused by a characteristic activation energy required for damage repair. This phenomenon emphasizes the sensitivity of the (Al)GaN surface to disruption by energetic particle-enhanced processes (others include dry etching or plasma enhanced chemical vapor deposition). In the early days of GaN device processing it was thought that the material was relatively impervious to these problems. Part of the reason for the dramatic improvement in device performance over the past few years has been the development of processing conditions that minimize surface damage. Continued progress in this area is paramount to the realization of the full potential of these devices.

Table 3-1. Summary of experimental DC and RF characteristics of MOCVD-grown AlGaN/GaN HEMTs.

<table>
<thead>
<tr>
<th>Nominal Gate Length, Lg (μm)</th>
<th>Gate Width, Wg (μm)</th>
<th>Maximum Current Density, IDS (A/mm)</th>
<th>Extrinsic Transconductance, gm (mS/mm)</th>
<th>Cutoff Frequency, fT (GHz)</th>
<th>Maximum Oscillation Frequency, fmax (GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>100</td>
<td>0.89</td>
<td>207</td>
<td>58</td>
<td>90</td>
</tr>
<tr>
<td>0.25</td>
<td>150</td>
<td>0.79</td>
<td>186</td>
<td>29</td>
<td>63</td>
</tr>
<tr>
<td>0.5</td>
<td>200</td>
<td>0.73</td>
<td>172</td>
<td>19</td>
<td>39</td>
</tr>
</tbody>
</table>

The cutoff frequency and maximum frequency of oscillation were extracted from measured s-parameters and are given in Figures 3-28, 3-29, and 3-30 for gate lengths of 0.1, 0.25, and 0.5 μm devices, respectively. These values are also listed in the Table 3-1, which summarizes the key DC and RF parameters of the 1110A devices. Maximum fT and fmax of 58 GHz and 90 GHz, respectively, were achieved for the 0.1 x 100 μm² devices, clearly illustrating the potential of AlGaN/GaN HEMTs for high frequency...
operation. The 0.25 μm devices exhibited an $f_T$ of 29 GHz and $f_{\text{max}}$ of 63 GHz, while the same parameters for the 0.5 μm device were measured to be 19 and 39 GHz, respectively. The high $f_{\text{max}}$ values are encouraging, since this is one of the most important metrics of device performance for applications such as satellites where overall gain is critical.

Figures 3-31 and 3-32 give a graphical representation of the gate length dependence of $f_T$ and $f_{\text{max}}$. A linear $f_T$ vs. $L_{\text{G}}^{-1}$ relationship is observed, as predicted from Equation 3.1. Calculation of the effective electron velocity from Figure 3-32 and Equation 3.1 gives $v_{\text{eff}} = 6.5 \times 10^6$ cm/s, consistent with the value estimated in Section 3.2.2, but below the theoretical value of $\sim 2 \times 10^7$ cm/s derived from Monte Carlo simulations.

Microwave power characteristics were measured with a load pull system at frequencies up to 10 GHz. The output power of a 0.25 × 150 μm device at 3 GHz is shown in Figure 3-33 as a function of DC input power. The output power increased monotonically with input power in the range 100 - 800 mW. At 1 dB compression, this device produced 22.5 dBm output power with a gain of 19.2 dB. The associated power-added efficiency (PAE) was 20.6%. At the 3 dB compression point, the same device produced 26.2 dBm output power, 17.2 dB gain, and 56.3% PAE, corresponding to an output power density of 2.75 W/mm (Figure 3-34). These power values at 3 GHz are approximately a factor of 5 higher than can be obtained with GaAs pHEMTs. At 10 GHz, an output power density of 1.7 W/mm was measured. This is roughly a factor of 3 higher than state-of-the-art GaAs pHEMTs at this frequency. As has been well documented in the AlGaN/GaN HEMT literature, the output power of devices grown on sapphire is severely limited by the poor thermal conductivity of the substrate. Flip-chip
bonding or other means of heat sinking would allow higher power densities to be realized.

Figure 3-13. SEM micrograph of 0.1 μm gate length HEMT showing Ohmic contacts, mesa step, and gate profile.
Figure 3-14. SEM micrograph of 0.1 μm gate length device showing gate and channel definition and gate anchoring technique to improve adhesion. Note smoothness of ICP-etched GaN surface and mesa sidewalls.
Figure 3-15. SEM cross-section of 0.25 μm Ni/Au gate. The scale bar at lower right is 100 nm.
Figure 3-16. Room temperature drain I-V characteristics for (a) $0.1 \times 100 \, \mu m^2$, (b) $0.25 \times 150 \, \mu m^2$ HEMTs.
Figure 3-17. Room temperature drain I-V characteristics for $0.5 \times 150 \, \mu m^2$ HEMT.

Figure 3-18. Transfer characteristics of $0.1 \times 100 \, \mu m^2$ gate dimension MOCVD-grown HEMT.
Figure 3-19. Transfer characteristics of $0.25 \times 100 \, \mu m^2$ gate dimension MOCVD-grown HEMT.

Figure 3-20. Transfer characteristics of $0.5 \times 100 \, \mu m^2$ gate dimension MOCVD-grown HEMT.
Figure 3-21. Comparison of extrinsic transconductance for HEMTs of 3 different gate dimensions. Note the positive voltage shift of the $g_m$ peak with increasing gate length and width.

Figure 3-22. Transconductance vs. $L_g^{-1}$ illustrating inverse gate length dependence.
Figure 3-23. Drain current density (measured at $V_G = +1$) for $0.25 \times 100 \mu m^2$ HEMT.

Figure 3-24. Gate reverse current for MOCVD-grown HEMTs. Note that gate characteristics from multiple devices are shown for each wafer. Device 1110A exhibits hard breakdown at $-26$ V (current compliance set to $-2.5$ mA).
Figure 3-25. Reverse current density as a function of gate dimension for 0.1, 0.25, and 0.5 μm gate length HEMTs. Note that the reverse current density for the longer gate length devices is very similar, but leakage from the short gate length device is much higher.

Figure 3-26. Reverse gate I-V-T characteristics illustrating significant decrease in leakage at elevated temperature.
Figure 3-27. Forward Gate I-V-T illustrating increase in turn-on voltage at elevated temperature.

Figure 3-28. Experimental $h_{21}$ and $|U|$ for $0.1 \times 100 \mu m^2$ device, illustrating $f_t = 58$ GHz and $f_{max} = 90$ GHz.
Figure 3-29. Experimental $h_{21}$ and $|U|$ for $0.25 \times 150 \ \mu m^2$ device, illustrating $f_T = 29$ GHz and $f_{max} = 63$ GHz from -20 dB/decade extrapolation.

Figure 3-30. Experimental $h_{21}$ and $|U|$ for $0.5 \times 200 \ \mu m^2$ device, illustrating $f_T = 20$ GHz and $f_{max} = 39$ GHz from -20 dB/decade extrapolation.
Figure 3-31. Gate length dependence of $f_T$ and $f_{\text{max}}$ for MOCVD-grown HEMTs.

Figure 3-32. Cutoff frequency, $f_T$, vs. inverse gate length for 0.1, 0.25, and 0.5 $\mu$m gate length HEMTs.
Figure 3-33. Output power of $0.25 \times 150 \, \mu m^2$ HEMT as a function of DC input power.

Figure 3-34. 3 GHz power performance of $0.25 \times 150 \, \mu m^2$ HEMT.
3.2.4 Small Signal Modeling

The small signal FET model is extremely important for active microwave circuit work. Models provide a vital link between measured scattering parameters and electrical processes in the intrinsic device. In the small signal model, each circuit element provides a lumped approximation to some aspect of the device physics. Properly extracted models allow prediction of device performance beyond measurement capability. In addition, equivalent circuit models can be scaled with gate width for a given foundry. This is important for scaling AlGaN/GaN HEMTs from small discrete devices (~100 μm gate width) to the large periphery devices (>1 mm) necessary for power amplification circuitry.

At microwave frequencies, it is difficult to accurately measure the absolute current and/or voltage at the input and output ports of the device under test (DUT). Scattering parameters measure traveling waves, not total voltage or current, and are therefore convenient for measurement and analysis. The s-parameters are algebraically related to the impedance (z) parameters and admittance (y) parameters. For a 2-port microwave network (1 input and 1 output, such as for an FET), the scattering matrix is 2 × 2 and can be written:

\[ B = SA \quad (3.2) \]
\[
S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \quad (3.3)
\]

where A and B represent the complex amplitudes of the incoming and outgoing traveling wave, respectively, and S is the frequency-dependent 2-port scattering matrix containing the 4 s-parameters. The s-parameters are complex quantities and contain both real and imaginary parts. The first subscript number is the output port and the second is the input.
port (e.g. $s_{21}$ relates the output power at port 2 to the input power at port 1). $S$-parameters also give information about the stability of a microwave network or amplifier.

\[
\begin{bmatrix}
  s_{11} & s_{12} \\
  s_{21} & s_{22}
\end{bmatrix}
\]

Figure 3-35. 2-port network illustrating input and output traveling waves and scattering matrix $S$.

Small signal modeling of AlGaN/GaN HEMT NF1110A (shown schematically in Figure 3-12) was performed with the commercially available ICCAP device simulation package. Devices of $0.1 \times 100$, $0.25 \times 150$, $0.5 \times 200$, and $0.8 \times 100$ $\mu m^2$ gate dimensions were modeled. The simulated scattering parameters were fitted to experimental data collected on an HP 8510 Vector Network Analyzer. The equivalent circuit model used for the simulation is given in Figure 3-36 and contains source, drain, and gate resistances and inductances; gate-source resistance and capacitance; intrinsic channel resistance; drain-source resistance and capacitance; a voltage-controlled current source ($g_m$); and gate-drain resistance and capacitance. A physical representation of each of the equivalent circuit elements is given in Figure 3-37. The simulation was performed by optimization of equivalent circuit parameter values, restricting values within a physically reasonable user-defined range. Complete simulations typically required several hours for convergence.
Figures 3-38 and 3-39 give the Smith Chart representation of experimental and simulated input ($s_{11}$) and output ($s_{22}$) reflection coefficients, respectively for a 0.25 μm gate length device. Note the excellent agreement for both curves at frequencies up to 40 GHz. Figure 3-40 shows the experimental and simulated $s_{21}$ forward transmission coefficient (essentially the gain) of the DUT and excellent agreement is again obtained. From the modeled s-parameters, the extrinsic cutoff frequency and the de-embedded intrinsic cutoff frequency were calculated (Figures 3-41 and 3-42). For the $0.25 \times 150 \ \mu m^2$ device, an extrinsic $f_T$ of 30 GHz was obtained. From experimental extrapolation of $h_{21}$ at -20dB/decade for the same device, the extrinsic $f_T$ was 29 GHz (see Figure 3-29), illustrating the excellent fit of the model. The intrinsic $f_T$ from the simulated data was 33 GHz, suggesting minimized parasitics that do not significantly decrease the speed of the intrinsic device.

Scaling analysis can be performed using small signal models to predict performance of devices with larger gate periphery. This is useful in scale-up from the discrete device level to multiple gate finger devices suitable for use in amplifiers. A scale factor can be defined that is equal to the ratio of the old and new gate widths. This analysis leads to simple predictive expressions for all equivalent circuit elements. However, such an analysis is only valid for scale-up at constant gate length. The e-beam gates of wafer NF1110A used for the small signal study were written with 3 lengths, each with a different gate width. Therefore, scaling analysis was not possible. Simply for illustration, Figure 3-43 gives a plot of capacitance values extracted from the model as a function of gate area, and the expected linear dependence is observed in all cases. Future
devices were written with varying gate widths for each gate length, in order to allow proper gate width scaling of all equivalent circuit parameters.

Figure 3-36. Equivalent circuit representation of FET used for small signal modeling.

Figure 3-37. Physical representation of equivalent circuit elements.
Figure 3-38. Measured and modeled input reflection coefficient ($s_{11}$) of $0.25 \times 150 \, \mu m^2$ gate dimension HEMT.

Figure 3-39. Measured and modeled output reflection coefficient ($s_{22}$) of $0.25 \times 150 \, \mu m^2$ gate dimension HEMT.
Figure 3-40. Measured and modeled $S_{21}$ gain of $0.25 \times 150 \ \mu m^2$ gate dimension HEMT.

Figure 3-41. Simulated extrinsic cutoff frequency of 30 GHz from small signal model. The experimental value was 29 GHz, as shown in Figure 3-29.
Figure 3-42. Simulated intrinsic cutoff frequency from small signal model. The intrinsic \( f_T \) is 33 GHz, only 3 GHz higher than the extrinsic value.

Figure 3-43. Gate-source, gate-drain, and drain-source capacitance extracted from small-signal linear model.
3.3 Direct Comparison of HEMTs Grown on Al₂O₃ or AlN/SiC

Currently, most III-N devices are fabricated on heteroepitaxial material grown on sapphire because of the relatively low cost and general availability of these substrates. However, for power applications it is desirable to have a substrate with higher thermal conductivity, such as SiC [216,217]. In this section, results are presented from a detailed investigation of the dc performance of AlGaN/GaN HEMTs grown side-by-side on either sapphire substrates or AlN/SiC templates. The gate length dependence of the dc parameters as well as the temperature dependence of transconductance, drain current and forward and reverse gate current characteristics were measured for both types of devices.

The defect density in the epilayers grown on AlN/SiC templates was significantly lower than those grown on sapphire, as measured by transmission electron microscopy (TEM). Reverse breakdown voltages above 40 V were obtained for 0.25 μm gate length devices on both types of substrate. Extrinsic transconductances of ~200 mS/mm for HEMTs on sapphire and ~125 mS/mm for devices on AlN/SiC were achieved, with the latter devices showing significantly lower self-heating effects. Both types of HEMTs showed similar trends of drain current and transconductance with increasing temperature. There was a clear signature of longitudinal optical phonon scattering in the HEMTs grown on both substrates.

The AlN/SiC template was fabricated by hydride vapor phase epitaxy (HVPE) of ~1000 Å AlN on the (0001)Si face of on-axis 6H-SiC substrates (conducting, n~10¹⁷ cm⁻³). Such AlN layers are single crystal and displayed electrical resistivity of about 10⁹ Ω·cm at room temperature and about 10⁶ Ω·cm at 700 K. Properties of these AlN layers
Table 3-2: Layer Structure of AlGaN/GaN HEMTs grown on Sapphire or AlN/SiC templates.

<table>
<thead>
<tr>
<th>Layer Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 Å UID GaN</td>
</tr>
<tr>
<td>100 Å UID Al(<em>{0.2})Ga(</em>{0.8})N</td>
</tr>
<tr>
<td>100 Å Al(<em>{0.2})Ga(</em>{0.8})N:Si, ~1E18 cm(^{-3})</td>
</tr>
<tr>
<td>30 Å UID Al(<em>{0.2})Ga(</em>{0.8})N</td>
</tr>
<tr>
<td>~1 µm UID GaN</td>
</tr>
<tr>
<td>~200 Å AlN buffer</td>
</tr>
<tr>
<td>(0001) Sapphire (NF1205B) or AlN/n-6H SiC (MS 397)</td>
</tr>
</tbody>
</table>

Figure 3-44. Layer structure of AlGaN/GaN HEMTs grown on sapphire substrates or AlN/SiC templates.

grown by HVPE on SiC substrates were reported by Melink et al. [218]. The Al\(_2\)O\(_3\) and AlN/SiC templates were first deposited with a thin AlN buffer, and then the layer structure shown in Figure 3-44 was grown by MOCVD at ~1040°C using conventional precursors. Hall measurements at room temperature showed a sheet carrier density of 8.5 \(\times 10^{12}\) cm\(^{-2}\) with a mobility of 1045 cm\(^2\)/V·sec for the structure grown on sapphire and a corresponding sheet electron density of 1.38 \(\times 10^{13}\) cm\(^{-2}\) and mobility of 731 cm\(^2\)/V·sec for that on the AlN/SiC template, as summarized in Table 3-2. The higher mobility in the sapphire structure may have been related to the slightly smoother surface, which may imply a smoother AlGaN/GaN interface. Figures 3-45 and 3-46 show atomic force...
microscopy (AFM) scans over $1 \times 1 \, \mu m^2$ areas for epilayers grown on sapphire and AlN/SiC, respectively. The structure grown on AlN/SiC showed a root-mean-square (RMS) roughness of 0.33 nm, while that on sapphire had a value 0.17 nm over the same area. However, the defect density visible by cross-sectional TEM was significantly lower in the structures grown on AlN/SiC (Figures 3-47 and 3-48) due to the closer lattice match (see Table 2-2). The AlN buffer is clearly effective in improving the crystal quality of the overlying GaN, as is particularly evident in Figure 3-48 (bottom).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Mobility (cm$^2$.V$^{-1}$.s$^{-1}$)</th>
<th>Electron Sheet Density (cm$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al$_2$O$_3$</td>
<td>1045</td>
<td>$8.54 \times 10^{12}$</td>
</tr>
<tr>
<td>AlN/SiC</td>
<td>731</td>
<td>$1.38 \times 10^{13}$</td>
</tr>
</tbody>
</table>

HEMT processing proceeded as described in Section 3.2.1. Gate lengths of 0.1, 0.25, 0.5, and 0.75 μm were written by e-beam lithography. For these devices, gate widths of 100, 150, and 200 μm were fabricated for each gate length. The DC characteristics were measured on a temperature-controlled chuck with an HP4145B parameter analyzer.

Figure 3-49 shows $I_D$-$V_{DS}$ characteristics for 0.75 μm gate length HEMTs on both types of substrates. Higher drain currents were obtained with the devices on sapphire but for both types of HEMTs there was no current collapse observed due to hot electron
Figure 3-45. $1 \times 1 \, \mu m^2$ AFM image of HEMT grown on sapphire substrate.
Figure 3-46. $1 \times 1 \, \mu m^2$ AFM image of HEMT grown on AlN/SiC template.
Figure 3-47. Cross-sectional TEM images of HEMT structure grown on sapphire.
Figure 3-48. Cross-sectional TEM images of HEMT structure grown on AlN/SiC. Note the greatly reduced density of threading dislocations.
Figure 3-49. $I_D$-$V_{DS}$ characteristics of 0.75 μm gate length devices on sapphire (top) and AlN/SiC (bottom).
Figure 3-50. $I_D-V_{DS}$ characteristics of 0.25 μm gate length devices on sapphire (top) and AlN/SiC (bottom).
injection and trapping in the buffer layer. This attests to the quality of the MOCVD GaN buffer layers, since the hot electron trapping is typically associated with defects or dislocations below the FET channel. Similar behavior was obtained with shorter gate length devices, as shown in Figure 3-50 for 0.25 \( \times \) 100 \( \mu \text{m}^2 \) devices on both types of substrates. There is excellent pinch-off for both types of devices and slightly higher peak drain currents for those on sapphire. Figure 3-51 illustrates operation at higher drain bias. Note that 40 mA peak current was obtained at \( > 40 \text{V}_{DS} \) for the devices on AlN/SiC with no self-heating (i.e. \( \geq 1.6 \text{W} \) of power in devices not optimized for power performance).

Typical HEMT transfer characteristics for 0.75 \( \mu \text{m} \) gate length devices are shown in Figure 3-52. Note that \( g_m \geq 200 \text{mS/mm} \) for devices on sapphire, with maximum values of \( \sim 135 \text{mS/mm} \) for those on AlN/SiC. The difference likely results from the fact that the growth conditions are optimized for sapphire, but not for the AlN/SiC template. It is fairly clear that with additional experience of growth of the AlGaN/GaN structures on AlN/SiC, equivalent or superior results could be achieved due to the lower defect density. The transfer characteristics for 0.25 \( \mu \text{m} \) gate length devices are shown in Figure 3-53. The \( g_m \) values are nearly equal to those from the larger gate length HEMTs, with maximum values of \( \sim 190 \text{mS/mm} \) for devices on sapphire and \( \sim 120 \text{mS/mm} \) for those on AlN/SiC. Similar relative changes in drain current were observed for the shorter gate lengths.

Normalized transconductances for 0.25 \( \mu \text{m} \) gate length devices on both types of substrates are shown in Figure 3-54 for drain-source voltages in the range 5-40 V. Note the sharper decreases in transconductance for devices on sapphire at high drain voltages.
Figure 3-51. $I_D-V_{DS}$ characteristics at high bias for 0.25 $\mu$m gate length devices on sapphire (top) and AlN/SiC (bottom).
Figure 3-52. Transfer characteristics for 0.75 μm gate length devices on sapphire (top) and AlN/SiC (bottom).
Figure 3-53. Transfer characteristics for 0.25 µm gate length devices on sapphire (top) and AlN/SiC (bottom).
Figure 3-54. Normalized transconductance as a function of drain voltage for 0.25 µm gate length devices.
due to negative differential output conductance (self-heating effects). The data were normalized to the $+5 \ V_{DS} \ g_m$ value for each device. This representation clearly emphasizes the need for the high thermal conductivity SiC substrates if high power conditions are needed in order to minimize the self-heating of the HEMTs.

The forward gate current density characteristics for both types of devices are shown in Figure 3-55 as a function of gate length. In both types of devices there is low-bias current which is probably a result of the high defect density in the material. The barrier heights were $\sim 0.96 \ eV$ for e-beam gate devices on sapphire and $1.1 \ eV$ for optical lithography devices fabricated from the same wafer.

The reverse gate current density characteristics are shown in Figure 3-56. The current densities are comparable for both types of HEMTs for a given gate length. Forward gate I-V comparison plots are given in Figure 3-57 for $0.25 \ \mu m$ gate lengths, illustrating the higher turn-on voltage for the HEMTs on AlN/SiC. The entire comparison curve is shown on a semilog scale in Figure 3-58 At a given bias, the sapphire devices exhibit larger forward current and smaller reverse leakage current.

Figure 3-59 shows drain I-V characteristics for $0.25 \ \mu m$ gate length devices as a function of measurement temperature from $25 - 300^\circ C$. The drain currents decrease with temperature for devices fabricated on both types of substrates. Note that for the HEMTs on sapphire, the overshoot in the I-V characteristics decreases as the temperature increases. This is expected, given that self-heating becomes less significant if the device is already at elevated temperature.
Figure 3-55. Forward gate characteristics as a function of gate length at 25°C for devices on sapphire (top) and AlN/SiC (bottom).
Figure 3-56. Reverse gate characteristics as a function of gate length at 25°C for devices on sapphire (top) and AlN/SiC (bottom).
Figure 3-57. Comparison of 25°C gate I-V characteristics of HEMTs grown on sapphire substrates and AlN/SiC templates. The same data are plotted on linear (top) and semilog (bottom) scales to illustrate the differences in forward turn-on.
Figure 3-58. Comparison of 25°C gate I-V characteristics for 0.25 × 100 μm² devices grown on sapphire substrates and AlN/SiC templates.
Figure 3-59. $I_D$-$V_{DS}$ characteristics as a function of temperature for 0.25 μm gate length devices grown on sapphire substrates (top) and AlN/SiC templates (bottom).
Figure 3-60 shows normalized I-V-T comparison for 0.25 μm gate length HEMTs, where the drain current was normalized to a value of unity for the maximum current measured at room temperature. This representation of the data clearly emphasizes the absence of self-heating effects in the HEMTs on AlN/SiC. The transfer characteristics as a function of measurement temperature are shown in Figure 3-61 for both types of devices. Note that the HEMTs on sapphire do not maintain pinch-off at temperatures > 25°C, whereas the devices on AlN/SiC still pinch-off at 300°C. The transconductances for the same devices are shown in Figure 3-62 as a function of measurement temperature. At 300°C both types of devices retain approximately 40-50% of their original gm values.

A summary of the temperature dependence of gm and Ip is shown in Figure 3-63 for the 0.25 μm gate length devices. The data was normalized to the 25°C values and Ip, max was taken from the +1V_G trace at each temperature. The same basic trends for both parameters are observed in the devices on the two different types of substrates, with slightly higher absolute values for the devices on sapphire. Once again, it is likely that this is due to the MOCVD growth recipe which was optimized for sapphire.

Figure 3-64 gives the forward gate I-V characteristics from 0.25 μm gate length HEMTs as a function of temperature. In general, the gate current increases with temperature, as expected since the conduction mechanisms are thermally activated. Note that there is not a vast difference in the characteristics for both types of devices over this broad measurement temperature range, emphasizing the ability of AlGaN/GaN HEMTs to accommodate high temperature operating requirements. The forward and reverse gate
Figure 3-60. Normalized $I_D$-$V_{DS}$ characteristics as a function of temperature for 0.25 μm gate length devices, clearly illustrating the benefits of the high thermal conductivity SiC substrate.
Figure 3-61. Transfer characteristics for 0.25 μm gate length devices grown on sapphire (top) and AlN/SiC (bottom) as a function of temperature.
Figure 3-62. Transconductance of 0.25 \mu m gate length devices grown on sapphire (top) and AlN/SiC (bottom) as a function of temperature.
Figure 3-63. Normalized $I_D$ and $g_m$ as a function of temperature for $0.25 \times 100 \, \mu m^2$ gate dimension devices.
Figure 3-64. Forward gate characteristics as a function of temperature for 0.25 μm gate length devices grown on sapphire (top) and AlN/SiC (bottom).
I-V characteristics from 0.25 µm gate length devices are shown in Figure 3-65 as a function of measurement temperature. The reverse leakage currents are slightly lower in the devices on sapphire, although there are not any major differences in the overall behavior of the devices on the two types of substrate.

Figure 3-66 shows low drain-source voltage output I-V characteristics \((V_\text{DS} = 0-0.4 \text{ V})\) as a function of measurement temperature for the 0.25 µm gate length devices. The overall magnitude of the drain current is larger in the devices on sapphire, but the relative decrease with temperature is similar in both types of devices. In an effort to determine any differences in the primary electron scattering mechanism for the two types of devices, the temperature dependence of the effective mobility was extracted from the linear region of the I-V-T curves of Figure 3-66. In the linear regime of the output I-V curve for HEMTs, the drain current is proportional to the electric field in the semiconductor (i.e. drain-source voltage), with a proportionality constant including the effective mobility. Hence, measuring the slope of the I-V-T curve as a function of temperature at low drain bias gives an estimate of the temperature dependence of the 2DEG mobility. The resulting data are shown in Figure 3-67 for chuck temperatures from 25 - 300°C. The clear inverse exponential dependence observed for the devices grown on sapphire indicates that longitudinal optical phonon scattering is dominant. This is consistent with both theoretical predictions [219] and experimental results from \(\text{Al}_{0.15}\text{Ga}_{0.85}\text{N}/\text{GaN}\) heterostructures [220]. The results from the device grown on AlN/SiC are linear for \(T \geq 100°C\), although the datum taken at room temperature does not fit this trend. It is unclear if this is value is anomalous or if a secondary scattering mechanism dominates at lower temperatures for the heterostructure grown on AlN/SiC.
Figure 3-65. Gate I-V characteristics as a function of temperature for 0.25 μm gate length devices grown on sapphire (top) and AlN/SiC (bottom).
Figure 3-66. Low bias $I_D$-$V_{DS}$ characteristics as a function of temperature for 0.25 $\mu$m gate length devices.
Figure 3-67. Temperature dependence of electron effective mobility for 0.25 μm gate length devices grown on sapphire (top) and AlN/SiC (bottom).
The subthreshold characteristics for 0.25 \( \mu \text{m} \) gate length devices are shown in Figure 3-68 as a function of temperature. As discussed earlier, the HEMTs on sapphire did not pinch-off for temperatures > 25\(^\circ\)C, whereas those on AlN/SiC had good pinch-off characteristics even at 300\(^\circ\)C. This is clearly depicted in the semilog representation of the transfer characteristics.

In summary, a direct comparison of the dc performance of AlGaN/GaN HEMTs grown on sapphire substrates or AlN/SiC templates showed the following main points:

1. Devices on AlN/SiC do not show any self-heating effects even for relatively high voltage (40 V) operation. By sharp contrast, the HEMTs on sapphire show pronounced self-heating effects on the I-V characteristics for measurement temperatures up to \( \sim 200\degree\)C.

2. Structures grown on AlN/SiC show significantly lower defect densities measured by TEM than those grown on sapphire.

3. The \( g_m \) and \( I_D \) values were larger in the HEMTs grown on sapphire, but this is most likely a result of the non-optimized growth on the AlN/SiC templates.

4. HEMTs on AlN/SiC show much less fall-off in \( g_m \) with drain voltage due to the absence of self-heating effects. In addition, devices on sapphire did not maintain pinch-off for temperatures above 25\(^\circ\)C.

5. Devices on both types of substrate retain 40-50% of their room temperature \( g_m \) and \( I_D \) values at 300\(^\circ\)C.

6. Longitudinal optical phonon scattering was the dominant electron scattering mechanism in devices fabricated on sapphire substrates and AlN/SiC templates.
Figure 3-68. Semilog plot of $I_D-V_G$ characteristics as a function of temperatures for 0.25 \( \mu \text{m} \) gate length devices grown on sapphire (top) and AlN/SiC (bottom). Note lack of pinch-off for sapphire devices.
3.4 AlGaN/GaN HEMTs on RF-assisted MBE-grown Epilayers

A common feature of reported HEMT work to date is the fact that molecular beam epitaxy (MBE) material appears quite competitive with that grown by MOCVD [175]. This has not been the case with photonic devices, in which the generally better optical properties obtained by MOCVD growth have led its dominance in LED and LD manufacturing. In this section, the fabrication and DC performance of MBE-grown AlGaN/GaN HEMTs is investigated.

MBE growth was proceeded using a standard Ga effusion cell and RF atomic nitrogen plasma source. Growth details have been given previously [221-223], but in brief consist of a low temperature AlN (300Å) buffer on sapphire (0001), 2 μm of undoped GaN grown at 750°C under Ga-rich conditions, 250Å of undoped Al0.2Ga0.8N, and a 30Å undoped GaN cap. A growth rate of 0.5 - 1 μm·hr⁻¹ was used for all depositions. The layer structure is given in Figure 3-69, which also includes schematic representation of the source, gate, and drain contacts.

Device processing proceeded as described in Section 3.2.1. E-beam gate lengths of 0.1, 0.25, 0.5, and 0.75 μm were written. In addition, 1.2 μm gates were fabricated by contact lithography on a separate sample from the same parent wafer. The $I_D-V_{DS}$ characteristics for devices of each gate length are shown in Figures 3-70, 3-71, and 3-72. In all cases the gate width was 100 μm, and the devices were biased with +1V gate bias ($V_G$), with a step of -1V. The drain characteristics are free of kinks in these unpassivated devices, with no sign of the current collapse phenomenon attributed to hot electron injection and trapping in the buffer layer [224]. In this mechanism, as a high drain voltage is reached, electrons are injected into the GaN buffer layer where they are trapped...
Figure 3-69. Schematic diagram of AlGaN/GaN HEMT with MBE-grown epilayer structure.

... in deep states. This trapped charge acts to deplete the two-dimensional electron gas from the buffer side of the channel and thereby reduces the drain current for the subsequent traces [224]. The shortest gate length devices show a crowding of the I-V traces at high grain source voltage and do not pinch-off. Figure 3-73 shows the $I_D-V_{DS}$ characteristics from a $0.25 \times 100 \, \mu m^2$ date device operated at high-bias conditions. Operation at $>40 \, V_{DS}$ with $>50 \, mA$ peak $I_D$ was obtained. It is clear that GaN HEMTs on sapphire substrates will not be as well-suited for power operation as those on SiC substrates. Flip-chip bonding is an attractive option to avoid the poor thermal characteristics of the sapphire.
Figure 3-70. Drain I-V characteristics of $0.1 \times 100 \, \mu m^2$ (top) and $0.25 \times 100 \, \mu m^2$ MBE-grown HEMTs.
Figure 3-71. Drain I-V characteristics of $0.5 \times 100 \, \mu m^2$ (top) and $0.75 \times 100 \, \mu m^2$ MBE-grown HEMTs.
Figure 3-72. Drain I-V characteristics of $1.2 \times 100$ MBE-grown HEMT.

A summary of the peak $I_D$ values obtained as a function of both gate length and gate width (100, 150, or 200 $\mu$m) is shown in Figure 3-74. A maximum current density of $\sim 0.44$ A·mm$^{-1}$ was obtained in 1.2 $\mu$m gate length devices for the shortest gate width. The variations in $I_{D,max}$ at 0.5 $\mu$m gate length are likely random and do not represent a trend. There seems to be little change in peak $I_D$ with decreasing gate length for the e-beam written gates, despite the expected $I_D$ increase at shorter lengths.

Typical HEMT transfer characteristics for $0.25 \times 100$ $\mu$m$^2$ devices are shown in Figure 3-75. There is little fall-off in $g_m$ over a range of drain-source biases, with a maximum transconductance of $\sim 80$ mS·mm$^{-1}$. A summary of the transconductance data as a function of gate length is shown at the bottom of Figure 3-76. Maximum $g_m$ values of $\sim 90$ mS·mm$^{-1}$ were obtained with $1.2 \times 100$ $\mu$m$^2$ devices.
From these results, RF plasma-assisted MBE growth of AlGaN/GaN HEMTs appears well-suited for producing high-performance devices. The epilayers surface morphologies and uniformity of device DC performance are excellent. Reasonable current densities and excellent scaling properties were achieved over a range of gate lengths and gate widths. As described in Chapter 2, the first high quality films for III-N light emitting devices were all grown by CVD. However, with the development of activated nitrogen species, MBE has since become more widely used for GaN-based devices. With additional optimization, MBE may even surpass MOCVD for devices where precise control of doping profiles or minimization of interfacial roughness is critical.
Figure 3-73. 40V, 2W operation of $0.25 \times 100$ MBE-grown HEMT.

Figure 3-74. Drain current density as a function of gate length for devices with gate widths ranging from 100 - 200 $\mu$m.
Figure 3-75. Transfer characteristics of 0.25 μm gate length MBE-grown HEMT.

Figure 3-76. Extrinsic transconductance as a function of gate length for devices with gate widths ranging from 100 - 200 μm.
CHAPTER 4
GROWTH AND ELECTRICAL CHARACTERIZATION OF NOVEL GATE OXIDES
FOR GaN MOS DEVICES

4.1 Introduction

Advances in material quality and device processing have led to promising results for III-nitride electronic devices for high temperature and high power applications. GaN MESFETs and AlGaN/GaN HEMTs exhibiting excellent device characteristics have been reported by numerous groups. However, leakage from the Schottky gate of such devices is a major concern for operation at elevated temperatures. An attractive alternative to GaN MESFETs and HFETs is the use of an insulated gate structure, reducing leakage and power consumption and increasing gate bias swing. Unfortunately, the availability of a thermally stable gate dielectric with sufficiently low density of interface states has been a historically troublesome issue for III-V materials. The native oxide of GaN and GaAs is Ga$_2$O$_3$, which has been shown to have very poor interface characteristics on both materials [225]. Heteroepitaxial growth techniques are therefore necessary to realize MIS structures on GaN or AlGaN layers. Of these, molecular beam epitaxy (MBE) is attractive due to its atomic layer control of film thickness and in-situ heating, cleaning, and film analysis. This chapter investigates MBE growth and electrical characterization of novel oxide films Gd$_2$O$_3$ and ScO for potential use as gate dielectrics in GaN-based MOS devices.
4.2 Molecular Beam Epitaxial Growth

Surface preparation is crucial to the achievement of a high quality semiconductor/insulator interface. Any residual impurities at the interface prior to oxide growth can contribute to increased interface state density, defect density, or surface roughness. GaN surface preparation has been studied in detail by Lee et al. [226]. The ex-situ surface preparation used for all oxide growths in this work consisted of 3 min. HCl:H2O (1:1), DI rinse, 25 min. UV-O3 exposure, 5 min. HF:H2O (1:3) or buffered oxide etch (BOE), and final DI rinse. The HCl etch removes any native oxides on the GaN surface. The UV-O3 treatment oxidizes residual surface carbon contamination, which is removed during the subsequent HF etch. After indium mounting the samples on molybdenum blocks and loading into the MBE chamber, the samples were heated to 700°C for 10 min. to desorb any residual impurities, then cooled to the growth temperature before initiating oxide growth.

All oxides were grown in a Riber 2300 gas-source MBE (GSMBE) system equipped with a Wavemat MPDR 610 ECR plasma source operating at 2.54 GHz. A schematic of the MBE system is given in Figure 4-1. Standard effusion cells were used for evaporation of the high purity (4 to 7N) elemental metallic precursors, with maximum heated cell temperatures near 1300°C. Two molecular sources are also available for electron beam deposition. The substrate temperature was measured using a backside thermocouple calibrated from the melting points of InSb and GaSb. Reflection high energy electron diffraction (RHEED) allows in-situ determination of substrate surface reconstruction prior to growth, as well as crystallinity of the surface of the growing film. After heating to 700°C prior to growth, the “clean” GaN usually exhibited a sharp (1 × 3)
surface reconstruction, illustrated by the RHEED pattern in Figure 4-2. This RHEED pattern is typical of the starting GaN surfaces for all oxide films in this study. The typical chamber pressure and forward power in the ECR plasma head during growth were $1 - 5 \times 10^{-4}$ torr and 150 - 200W, respectively.

Figure 4-1. Schematic of Riber GSMBE system used for oxide growths.
Figure 4-2. RHEED pattern of GaN after *ex-situ* and *in-situ* surface preparations, prior to initiation of oxide growth. The pattern was recorded along the (11-20) crystal plane and illustrates the (1x3) (0001) GaN surface reconstruction.

4.3. Quantification of Oxide Quality

Before the various dielectrics can be compared, it is first necessary to understand the important characteristics of a "good" gate insulator. The important quantities are related to inherent physical properties of the material as well as the atomistic structure of the film and interface. Fixed oxide charge, interface state density, border trap density, mobile ion density, breakdown electric field, thermal stability, and reliability are important quantities that relate directly to the quality of the resultant MOS device. Many of these are summarized schematically in Figure 4-3.

Of the various charges and/or traps present in the oxide, fixed oxide charge is perhaps the easiest to describe and understand. As the name implies, these charges are spatially fixed within the oxide, and cannot migrate throughout the oxide during device operation. Fixed oxide charges may be positive, negative, or a combination of both
leading to an effective (measured) fixed charge. The amount of fixed oxide charge is
easily determined from the C-V measurements described in Section 4.4.

Interface states, or interface traps, are unsatisfied dangling bonds at the
oxide/semiconductor interface which introduce energy levels within the forbidden energy
gap of the semiconductor. These energy levels can capture (fill) or emit (empty) charge
carriers as the MOS device is operated. Interface states are manifested as a bias-
dependent non-ideality that may result in poor modulation of the charge in the underlying
semiconductor. In the limit of extremely high interface state density ($D_{it}$), field-effect
devices are not possible.

Border traps are similar to interface traps, but are spatially located a short distance
into the oxide from the interface. These traps can communicate with the semiconductor
via tunneling, and therefore behave very much like interface traps. However, due to their
location in the oxide, the characteristic time constants of these traps are much longer than
those for interface traps. For this reason, border traps are sometimes referred to as "slow
states" while interface traps are termed "fast states".

Mobile impurities are particularly troublesome in MOS devices, since they can
lead to erratic, unstable device behavior. Mobile ions are free to move within the oxide
as a function of applied bias. Therefore, their spatial location varies depending on the
magnitude and sign of the gate bias. Furthermore, the magnitude of the voltage shift
caused by the mobile ions is dependent on their position in the oxide, with charges nearer
the oxide/semiconductor interface causing much larger nonidealities. Bias-temperature
stress measurements can easily confirm or deny the presence of mobile ionic charge.
Early Si MOS was plagued by instabilities that were later ascribed to ionic sodium contamination.

The breakdown electric field ($E_B$), also known as dielectric strength, is the maximum electric field that can be sustained within the oxide without electrical breakdown. Due to the large breakdown field of GaN (~3 MV/cm), a large oxide $E_B$ is critical in order to maximize power handling capability for high-field electronic devices, such as power MOSFETs. For a given gate bias, the transverse voltage drop is shared by both the oxide and semiconductor. The dielectric constants of both materials determine the ratio of the voltage drop or field experienced by each material. The ratio is calculated from:

$$E_s \varepsilon_s = E_B \varepsilon_{ox}$$

(4.1)

where $E_s$ is the electric field in the semiconductor, $\varepsilon_s$ is the static dielectric constant of the semiconductor, and $\varepsilon_{ox}$ is the static dielectric constant of the oxide. From this, it is shown that for a given $E_s\varepsilon_s$ product in the semiconductor, gate oxides with a small dielectric constant will experience a larger electric field. This especially troublesome, for example, in the SiO$_2$/SiC material system. SiC is a wide bandgap semiconductor that can be thermally oxidized to form SiO$_2$. However, the dielectric constant of SiO$_2$ is rather small (3.9) compared to that of SiC (~10). This leads to an electric field in the SiO$_2$ that is (~10 / 3.9) = 2.5 times larger than the field present in the SiC. Since the blocking voltage of a power device scales quadratically with maximum semiconductor electric field, it is often found that high voltage SiO$_2$/SiC MOS devices are limited by breakdown in the oxide, not in SiC. All oxides investigated in this study have dielectric constants comparable to
or greater than that of GaN, and are therefore attractive for use in high power applications if a sufficiently large oxide $E_B$ is obtained.

Thermal stability of the oxide is imperative for high temperature device operation. In addition, typical GaN processing steps (e.g. Ohmic and/or implant anneals) can expose the oxide layer to temperatures of $\sim1000^\circ$C. If the electrical properties as-grown oxide degrade during device processing or elevated temperature operation, the material is of little use as a gate insulator. Thermally-enhanced diffusion or migration of the gate metal through the oxide is one potential source of thermal instability. The temperature stability and reliability of candidate gate dielectrics can be evaluated with bias-temperature or constant current stress measurements.

![Diagram of MOS capacitor](image)

Figure 4-3. Pictoral representation of quantities important to MOS devices. Adapted from [227].

4.4. The MOS Capacitor

The most basic MOS device is the MOS capacitor (MOSC), also known as the MOS diode (Figure 4-4). The MOSC is a unipolar device, with an ohmic contact formed
on the semiconducting layer or substrate and a gate contact deposited on the surface of the gate oxide. The MOSC is a useful tool for evaluating many of the material and electrical properties of both the oxide and the semiconductor. For such purposes, fabrication of MOSCs generally does not require a great deal of complex processing. Ohmic contact can either be formed from the backside of the substrate or by etching the oxide and depositing metal onto the exposed underlying semiconductor (as is typically the case for GaN MOS devices fabricated on insulating sapphire substrates). The gate contacts are usually circular ‘dots’ of known diameter deposited onto the oxide surface lithographically or through a shadow mask.

![Schematic of MOS capacitor and external voltage source](image)

Figure 4-4. Schematic of MOS capacitor and external voltage source for electrical characterization. Both frontside and backside Ohmic contacts are shown although in practice, both would not be necessary on the same device. For MOS capacitors on conducting substrates, backside metallization is typical. For GaN MOS structures grown on insulating sapphire substrates, contact to the semiconductor must be made from the front side.

The MOSC is operated by applying a potential difference across the gate and ohmic contacts. In Figure 4-4, the semiconductor is held at ground potential and the
voltage source applies a positive or negative potential \((V_G)\) to the gate contact. For n-type semiconductors, application of a positive potential to the gate results in a forward biased condition and negative gate potential results in reverse bias. If the oxide is assumed to be a perfect insulator, no current flows as a result of the applied bias in the normal operation range of the device. Instead, the gate contact serves to create a transverse electric field in the direction normal to the surface. This electric field modulates the charge in the underlying semiconductor by attracting or repelling majority and/or minority carriers. Since there is no current flow, the MOSC can be described completely by electrostatics, greatly simplifying the mathematical treatment. A understanding of ideal MOSC behavior is facilitated through the use of energy band diagrams and ideal capacitance-voltage (C-V) curves.

4.4.1 Collection of Capacitance-Voltage Data

Capacitance-voltage (C-V) is one of the most useful electrical techniques for MOS characterization. From C-V data, information on fixed oxide charge, interface state density, border trap density, and mobile ion density can be quantified. With modern equipment, implementation of a C-V measurement is relatively straightforward. All C-V data in this work were collected using an automated HP4284A LCR connected via GPIB to a LabVIEW™-based PC interface. During the C-V measurement, the LCR meter generates a voltage signal of superimposed AC and DC components, as shown in Figure 4-5. The voltage input to the device at any time can be expressed as:

\[
V = V_{DC} + \delta V
\]  

(4.2)

where \(V\) is the input voltage signal to the MOSC, \(V_{DC}\) is the direct current, and \(\delta V\) is the contribution from the time varying, sinusoidal AC signal. The adjustable DC parameters
are voltage limits (operation window) and sweep rate. AC input parameters are frequency and amplitude of the sinusoidal waveform. The 4284 measures the complex impedance \((Z)\) of the MOSC, given by:

\[
Z = R + \frac{1}{j\omega C}
\]  

(4.3)

where \(R\) is resistance, \(j = \sqrt{-1}\), \(\omega\) is angular frequency, and \(C\) is capacitance. Series or parallel resistance and capacitance are determined from the real and imaginary portions, respectively, of the complex impedance. C-V plots are constructed by plotting the differential capacitance \([= \delta Q(t) / \delta V(t)]\) as a function of the applied DC bias, where \(Q\) is charge and \(V\) is voltage.

![Figure 4-5. Schematic representation of superimposed voltage signal seen by MOS capacitor during C-V measurement.](image)

4.4.2 Ideal MOS Capacitor

An ideal MOSC curve is given in Figure 4-6 for an n-type semiconductor. For p-type, the curve would be simply reflected about the ordinate (capacitance axis). From Figure 4-6, 3 distinct bias regimes are noted: accumulation, depletion and inversion. The
C-V curves are frequency-independent in accumulation and depletion, but become strongly frequency-dependent in inversion. In accumulation and depletion, *majority* carriers are responsible for the observed capacitance, while in inversion *minority* carrier response determines the shape of the C-V curve. Also shown in Figure 4-6 is the region of "deep depletion", to be discussed in detail in Section 4.4.2.5.

Figure 4-6. Ideal capacitance-voltage curves for a MOS capacitor at high and low frequency.

4.4.2.1 Accumulation

For an n-type semiconductor, accumulation refers to application of a positive gate potential, which serves to attract majority carriers (electrons) to the surface. This is illustrated by the energy band diagram of Figure 4-7(a), where the conduction band bending causes energetically-favored accumulation of electrons at the
oxide/semiconductor interface. Also defined in Figure 4-7 are the band bending ($\psi_s$), surface potential ($\phi_s$), and bulk potential ($\phi_B$). For accumulation bias, these three quantities are all positive, following the convention of increasing positive potential in the direction of decreasing electron energy (down on a standard energy band diagram). The block charge diagram of Figure 4-7(b) illustrates that the positive charge on the gate electrode is exactly balanced by the negative surface charge of majority electrons in the semiconductor, resulting in an effective parallel plate capacitor with separation equal to the oxide thickness, $d_{ox}$. The capacitance is given by the familiar parallel plate equation:

$$C_{ox} = \frac{\varepsilon_o \varepsilon_{ox} A}{d_{ox}}$$  \hspace{1cm} (4.4)

where $C_{ox}$ is the accumulation (oxide) capacitance, $\varepsilon_o$ is the permittivity of free space ($8.854 \times 10^{-14}$ F·cm$^{-2}$), $\varepsilon_{ox}$ is the oxide dielectric constant, $A$ is the capacitor area, and $d_{ox}$ is the oxide thickness. This capacitance is independent of frequency, since majority carrier response is typically several orders of magnitude faster than the frequency of the AC signal. This can easily be shown by a simple calculation of the characteristic response time of majority carriers to a point charge fluctuation. The characteristic screening length in the semiconductor is given by the Debye length:

$$\lambda = \sqrt{\frac{\varepsilon_s \varepsilon kT}{q^2 n}}$$  \hspace{1cm} (4.5)

where $\varepsilon_s$ is the semiconductor dielectric constant, $k$ is Boltzmann’s constant, $T$ is absolute temperature, $q$ is the electron charge, and $n$ is electron concentration. Majority carriers within a local Debye length will rearrange, or respond, to compensate the point charge. The field experienced by these electrons is approximated by $(kT/q) / \lambda$. The majority carrier response time is approximated by [228]:
Thus, even for the maximum practical measurement frequency of 1 MHz, majority carriers will always follow the AC signal and contribute to the accumulation capacitance. In GaN, for example, assuming a donor concentration of $10^{17}$ cm$^{-3}$, the majority carrier response time is on the order of 0.1 ps. This is approximately 7 orders of magnitude faster than the period of a 1 MHz AC signal.

\[
\tau_{maj} = \frac{E_s}{q\mu_n n} \quad (4.6)
\]

Figure 4-7. (a) Energy band diagram of MOSC biased in accumulation; (b) Block charge diagram illustrating equal and opposite charges residing across dielectric thickness, $d_{ox}$.
4.4.2.2 Depletion

A MOSC under moderate reverse bias is said to be in depletion (Figure 4-8). In depletion, \( \psi_s \) becomes negative, repelling majority carriers and creating a depletion region under the gate electrode. The surface potential may be either positive or negative, depending on the magnitude of the depletion bias, while the bulk potential remains positive, since it is not a function of the gate bias. In the depletion region, the positively-charged ionized donors exactly balance the negative charge on the gate contact. At a given temperature, the width of the depletion region depends on the magnitude of the reverse bias and the doping density of the semiconductor. For heavily doped material, the band bending will be more abrupt and the depletion region will be shallower than for an identical device with lighter doping density at the same gate bias. Since the measured capacitance consists of oxide and depletion region capacitances in series, the equivalent capacitance of the network is less than the accumulation capacitance. Correspondingly, the C-V curve decreases into depletion from the accumulation plateau as shown in Figure 4-6. As the magnitude of the DC reverse bias is increased, the depletion width (\( W_D \)) widens to compensate the charge on the gate contact.

Majority carrier response still governs the C-V behavior, since the depletion region responds to the AC signal by majority carrier movement into and out of the depletion region boundary (i.e. the point closest to the surface at which \( \phi(x) = \phi_B \), where \( \phi(x) \) represents the potential at any depth \( x \) measured from the surface (\( x = 0 \)) into the semiconductor). The depletion region therefore constantly changes its width by a small distance which uncovers sufficient ionized donor charge to balance the magnitude of the AC signal (typically \( \sim 10\text{mV} \)). Thus, on the positive half of the AC cycle \( W_D \) decreases.
and on the negative half-cycle it increases. Again, since majority carriers are solely responsible for the depletion capacitance, there is no frequency dependence.

Figure 4-8. (a) Energy band diagram of MOSC biased in depletion; (b) Block charge diagram illustrating positive ionized donor charge compensating negative potential on gate contact.
4.4.2.3 Inversion

In depletion, the surface charge due to ionized donors increases approximately with the square root of semiconductor band bending. After the onset of inversion \( \psi_s \approx \phi_B \), minority carrier surface charge begins to increase exponentially with band bending and quickly overtakes that of the ionized impurities. Since the large minority carrier surface charge effectively screens the depletion region from the DC signal, the depletion width reaches a fixed maximum (threshold) value \( W_T \). Additional gate charge is compensated by the accumulation of minority carriers (holes for n-type) at the semiconductor surface. This condition is known as inversion, since the conductivity of the surface is inverted (e.g. from n-type to p-type). Since the minority holes are present in much greater concentrations than majority electrons, the response of holes to the AC signal determines the C-V behavior of the MOSC in inversion.

In the absence of an external source of minority carriers (such as a diffused p\(^+\) region), bulk traps are responsible for the generation of minority carriers in the MOSC. For GaN at moderate temperatures (such that diffusion rates are relatively slow), hole generation stems from bulk traps in the depletion region of the semiconductor. For these bulk traps to be efficient minority carrier generation/recombination (G/R) centers, their energy levels must be simultaneously near \( E_i \) and \( E_F \) [228]. For narrow bandgap materials such as Si and Ge at room temperature and low measurement frequencies, minority carriers can respond completely since generation and recombination rates are faster than the period of the AC signal. In this case, holes (for an n-type semiconductor) are provided to the surface exactly as needed and the degenerate surface acts exactly as in accumulation, only with reversed polarity. The capacitance rises to the parallel plate
value given in Equation 4.1, as illustrated in Figure 4-6. A “U-shaped” low frequency C-V curve results.

At high frequencies, minority carrier response is slower than the frequency of the AC signal. Thus, minority carriers contribute nothing to the measured capacitance. For slow changes to the DC bias, minority carriers are still able to maintain thermal equilibrium and are generated to compensate the gate charge. However, since the minority surface charge is not able to respond to the AC signal, the depletion region must compensate the AC charge fluctuations. This is accomplished exactly as in accumulation, but with a fixed $W_D$ (perturbed only by the AC response). Since the depletion width is at $W_T$ (its maximum value), the high frequency capacitance saturates at a minimum value and remains constant with additional reverse bias, as shown in Figure 4-6.
4.4.2.4 Flatband

Flatband refers to a MOSC biased such that $\psi_s = 0$, which by definition implies $\phi_B = \phi_s$. The voltage at which this condition is achieved is known as the flatband voltage ($V_{FB}$). In an ideal MOSC, the semiconductor flatband capacitance is given by:

$$C_{FBS} = \frac{\varepsilon_o \varepsilon_r A}{\lambda_{ex}}$$

(4.7)

where $\lambda_{ex}$ is the extrinsic Debye length given as in 4.5 using $N_D$ for the electron concentration. The overall flatband capacitance ($C_{FB}$) of the device is given as the series network of $C_{ox}$ and $C_{FBS}$. As the oxide thickness increases, $C_{ox} << C_{FBS}$ (Figure 4-10) and $C_{FB} \approx C_{ox}$. This indicates that increasing oxide thickness causes the flat band
capacitance to lie nearer the accumulation capacitance plateau (at a given semiconductor doping density). For an ideal MOSC with no fixed oxide charge (Qox) or metal-semiconductor work function difference (ΦMS), CFB occurs at Vg = 0V. That is, at thermal equilibrium in the absence of external bias, Qox, or ΦMS, there is no difference in the potential at any point in the semiconductor. Differences in the work functions of the semiconductor and gate metal causes a built-in potential (Vbi) which shifts the flat band voltage by an amount in V equal to ΦMS in eV. Any additional shift in VFB is attributed to charges in the oxide, as will be discussed in Section 4.4.3.1.

Figure 4-10. Relative contribution of semiconductor flatband capacitance calculated with doping density as a parameter for ND from $1 \times 10^{16}$ to $5 \times 10^{17}$ cm$^{-3}$. The curves were generated for a Gd$_2$O$_3$/GaN MOSC at room temperature, assuming $\varepsilon_{\text{GaN}} = 9.5$ and $\varepsilon = 11.4$. 
4.4.2.5 Deep Depletion

In Section 4.4.2.3, the inversion capacitance was shown to depend on the AC frequency, resulting in the high and low frequency C-V curves illustrated Figure 4-6. For each of these, the MOSC is in thermal equilibrium with the DC signal at each bias point. A special, non-equilibrium, condition known as deep depletion can result if the DC sweep rate is too high. As the MOSC is biased from depletion into inversion, the Fermi level moves toward the valence band and thermal equilibrium requires generation of minority carriers to compensate the DC bias. In deep depletion, minority carrier generation cannot respond fully to the DC sweep rate, and the depletion region width must extend beyond $W_T$ in order to maintain charge neutrality. Thus, the deep depletion capacitance falls below the high frequency inversion limit (since $W_D > W_T$) and continues to decrease as inversion bias is increased. The deep depletion limit is known as total deep depletion, in which no minority carriers are present and the entire negative gate charge is balanced by ionized donors in the depletion region. In deep depletion, $W_D$ will increase until the generation of minority carriers in the widened depletion region is sufficiently enhanced to restore equilibrium or until avalanche breakdown provides holes via impact ionization. Since deep depletion is a non-equilibrium condition, generation of minority holes will eventually restore the capacitance to a high frequency inversion value if biased in inversion for a sufficient time. However, for GaN this time is well beyond the limits of a practical experiment.

To understand or predict the minority carrier response, the time constants of the bulk traps responsible for minority carrier generation and recombination must be considered. Shockley-Read-Hall (SRH) statistics have been shown to accurately model the dynamics of deep trap levels. From SRH theory:
where $t_r$ is the time required for generation of minority carriers from bulk traps, $n_i$ is the intrinsic carrier concentration, and $\tau_T$ is an effective lifetime term related to both majority and minority carriers. For an extrinsic doping density of $10^{17}$ at $25^\circ$C, $t_r$ from Equation 4.8 is on the order of seconds for Si, on the order of minutes for GaAs, and approximately $10^{11}$ years for GaN (comparable to the age of the universe). The extremely slow minority carrier generation rate in GaN is due to its large energy gap. Thus, thermal minority carrier generation in GaN at room temperature is impossible on practical timescales, and external sources of minority carriers are necessary to form inversion layers at the surface. Deep depletion C-V curves will always be obtained from GaN MOSCs with uniformly doped epilayers measured in the dark at room temperature.

4.4.3. Non-Ideal MOS

Capacitance-voltage techniques are sensitive to many of the non-idealities discussed in Section 4.3. Fixed oxide charge density, mobile ion density, border trap and interface trap density are all calculable from proper analysis of C-V data. Of these, only fixed oxide and interface trapped charges will be considered in this work. No experimental evidence of mobile charge in Gd$_2$O$_3$ or ScO films deposited on GaN has ever been observed. Also, for wide bandgap MOS, time constants for border traps are long enough that, in most cases, they can be effectively grouped with fixed oxide traps and collectively treated as an effective oxide charge density, $N_{\text{eff}}$.

4.4.3.1 Effective Oxide Charge

The effective oxide charge cannot communicate, or exchange charge, with the semiconductor or gate electrode. The charge is “fixed” in both spatial position and
charge magnitude. In the presence of oxide charge, the C-V curve is shifted along the voltage axis by an amount proportional to the magnitude of the effective charge. The slope of the curve is unchanged. The voltage shift may be positive or negative, depending on the polarity of the oxide charge. For positive effective oxide charge, the C-V curve shifts toward negative voltage from the ideal. For negative oxide charge, the C-V curves shifts in the opposite direction, toward more positive voltage. This is illustrated in Figure 4-11.

As an illustrative example, assume a positive effective oxide charge with an arbitrary charge distribution. Further assume a MOSC without metal-semiconductor workfunction difference held at $V_G = 0V$ (flatband). The positive oxide charge creates a mirror charge in the semiconductor in order to maintain charge neutrality (i.e. $\psi_s$ becomes slightly positive to balance the positive oxide charge with an accumulated electron charge in the semiconductor). Although the gate bias is still 0V, the MOSC is not longer at flatband conditions. A slightly negative gate potential is necessary to repel the mirror charge and restore flatband conditions in the semiconductor. Thus, the net effect of the positive oxide charge is a negative shift of the flatband voltage, as illustrated in Figure 4-11. A similar argument can be made for each point along the C-V curve, illustrating parallel shift without change in slope. The magnitude of the effective charge is usually calculated by measuring the flatband voltage shift ($\Delta V_{FB}$) between the real (including oxide charge) and the ideal (considering work function difference) C-V curves. The effective oxide charge is then given by:

$$N_{eff} = \frac{\Delta V_{FB} C_{ox}}{qA}$$ (4.9)
where \( A \) is the area of the MOSC. Typical values of \( N_{\text{eff}} \) are \( 10^{11} - 10^{12} \, \text{cm}^{-2} \) for the oxides investigated in this study.

![Normalized C-V curve for MOSC with (dashed) and without (solid) positive oxide charge. The flatband voltage shift is indicated.](image)

Figure 4-11. Normalized C-V curve for MOSC with (dashed) and without (solid) positive oxide charge. The flatband voltage shift is indicated.

4.4.3.2 Interface Trapped Charge

Interface trap density \( (D_{it}) \) is manifested as a bias-dependent "stretchout" of the capacitance-voltage curve. Interface traps are energy levels in the forbidden energy gap that are free to exchange charge with the semiconductor. In the presence of interface traps, a portion of the gate charge serves to fill or empty these traps, and the resultant band bending in the semiconductor is reduced from the ideal:

\[
\delta Q_{\text{gate}} = \delta Q_s \quad \text{(ideal)}
\]

\[
\delta Q_{\text{gate}} = \delta Q_s + \delta Q_{it} \quad \text{(with interface traps)}
\]
where \( Q_{it} \) represents the interface trapped charge and is a function of applied bias. Minimization of \( D_{it} \) is perhaps the single most important aspect of MOS materials (i.e. dielectric) development. If the interface state density is too high, or if it is not reproducible, the resultant MOS devices lose all practical utility. Various schemes have been developed to reduce \( D_{it} \), both during and after oxide growth. Proper surface preparation, growth conditions, and post-growth annealing can influence \( D_{it} \). It should also be noted that \( D_{it} \) depends on semiconductor surface orientation. For this reason, virtually all modern silicon MOS devices are grown on the (100) surface, since it has a smaller surface atom density than (111). This effect was not a factor in this study, since all insulators were grown on the (0001) hexagonal face of GaN.

Various methods have been developed for quantification of interface state density from capacitance-voltage data. A partial list includes the Gray-Brown method [229,230], Terman method [231], and high-low method [232]. An excellent chapter concerning \( D_{it} \) extraction from C-V data is given in Ref. [228]. Terman’s method is a simple technique used to give approximate \( D_{it} \) values from high frequency C-V data. For GaN MOS devices, high frequency techniques are attractive since low frequency inversion behavior is difficult to achieve in practice. The Terman method is based upon the principle that at high frequencies (where interface traps contribute nothing to the measured capacitance) the capacitance of an ideal MOSC and a MOSC with interface traps is exactly the same if and only if their band bending is the same. Mathematically:

\[
C (\psi_s)_{\text{ideal}} = C (\psi_s)_{\text{real}, \omega \to \infty}
\]

(4.12)

Generation of an ideal C-V curve is necessary for comparison with the experimental curve. Such C-V curves can be calculated by the following procedure:
1. Use semiconductor doping density and dielectric constant to calculate semiconductor flat band capacitance ($C_{FBS}$).

2. Solve Poisson’s equation to obtain potential as a function of position in the semiconductor.

3. Calculate the electric field in the semiconductor with band bending as the independent variable.

4. From the surface charge and the electric field, calculate the semiconductor surface capacitance as a function of $\psi_s$.

5. Calculate total capacitance (oxide and semiconductor in series) as a function of $\psi_s$.

6. Relate $\psi_s$ to applied bias (gate voltage, $V_G$) to obtain theoretical $C$ vs. $V_G$.

Comparison of the calculated and experimental curves gives the required relationship between $\psi_s$ (ideal) and $V_G$ (real). In practice, the two plots ($C_{\text{ideal}}$ vs. $\psi_s$ and $C_{\text{real}}$ vs. $V_G$) are typically plotted adjacently on the same y-axis scale. A horizontal line drawn at any constant capacitance intersects the 2 curves at values of $\psi_s$ and $V_G$. Scanning this line throughout the measured range of capacitance gives a locus of data used to plot $\psi_s$ vs. $V_G$. Graphical or numerical differentiation of this curve gives the differential change in band bending with gate bias. The interface state density is inferred directly from this value of $(d\psi_s/dV_G)$. This method gives one data point (one $D_{it}$ value) for each pair of $(\psi_s, V_G)$ data. The interface state density can be correctly positioned in the semiconductor energy gap by determining the energy immediately opposite the surface Fermi level at each bias point. This is done with the help of the energy band diagram, where it can be shown that this energy level is given by [228]:

$$\frac{E_c - E_f}{q} = E_i - (\psi_s + \phi_B)$$  \hspace{1cm} (4.13)
where \((E_c - E_T)\) is the discrete trap energy level with respect to the semiconductor conduction band. A plot of \(D_{it}\) vs. \((E_c - E_T)\) normally increases near the band edge and decreases toward midgap. It should be noted that there can be interface states with energy levels actually above the conduction band. However, these are almost always dwarfed by the conduction band density of states.

Special care must be exercised when inferring interface state densities from wide bandgap MOS C-V data. The emission rates of electrons from interface traps decreases exponentially with increasing separation from the conduction band edge. Shallow traps can effectively emit to the CB on practical timescales. However, emission from deeper lying states is too sluggish, and the charge in these traps is frozen during the C-V measurement. Such traps are qualitatively analogous to a fixed charge located at the interface. Hence, they contribute a flatband shift and not a distortion of the experimental C-V curve. For this reason, information about interface state density near midgap cannot be probed at room temperature using conventional techniques. High temperatures are necessary to sufficiently reduce deep level emission times. Capacitance-voltage characterization for GaN MOS devices should be collected at temperatures \(> 400^\circ\text{C}\) in order to “see” states near midgap. Both an n-type and a p-type sample are necessary in order to evaluate \(D_{it}\) throughout the bandgap: the n-type sample gives data from \(E_c\) to \(E_i\) and the p-type sample gives information from \(E_v\) to \(E_i\).

4.5. Gadolinium Oxide

Gallium gadolinium garnet, \((\text{Ga}_2\text{O}_3)\text{Gd}_2\text{O}_3\), evaporated from a single crystal \(\text{Ga}_5\text{Gd}_3\text{O}_{12}\) has been shown to effectively passivate GaAs and InGaAs, and both n-and p-type enhancement mode MOSFETs have demonstrated [233,234]. A similar approach
was used to fabricate Ga$_2$O$_3$(Gd$_2$O$_3$)/GaN MOS diodes with low $D_{it}$ [235]. However, compositional control of the (Ga$_2$O$_3$)$_x$(Gd$_2$O$_3$)$_y$ is difficult due to its strong dependence on substrate temperature during oxide growth. SIMS profiles suggest that the initial monolayers of Ga$_2$O$_3$(Gd$_2$O$_3$) are strongly Gd-rich, and possibly consist entirely of the binary Gd$_2$O$_3$. This naturally leads to the investigation of Gd$_2$O$_3$ as a candidate gate dielectric. Growth of stoichiometric Gd$_2$O$_3$ has been demonstrated on GaAs by electron beam deposition from a compound source [236,237]. The material properties of Gd$_2$O$_3$ make it attractive for use as a gate insulator on GaN, as summarized in Table 4-1 [238-240]. In addition to the large bandgap and dielectric constant and high melting temperature, Gd$_2$O$_3$ is expected to show increased compositional uniformity in the growth direction relative to Ga$_2$O$_3$(Gd$_2$O$_3$) [241]. Gd$_2$O$_3$ exists in the Bixbyite (Mn$_2$O$_3$) crystal structure, which exhibits similar atomic symmetry in the (111) plane as the GaN (0001) basal plane, suggesting the potential for growth of an epitaxial dielectric. Single crystal growth of Gd$_2$O$_3$ on GaAs has, for example, been demonstrated [242,243]. However, in the case of GaAs, the expected orientation between the Gd$_2$O$_3$ and GaAs is (100) rather than (111). Further, the bond length mismatch between the Gd$_2$O$_3$ and GaAs (100) planes is smaller than between the Gd$_2$O$_3$ (111) and GaN (0001), 4.2% vs. 20%. Thus it is not obvious from the GaAs results that single crystal Gd$_2$O$_3$ can be grown on GaN. Based upon results with other dielectrics [225], it is expected that polycrystalline Gd$_2$O$_3$ will not provide the low interface and oxide trap densities needed for acceptable device performance, necessitating the growth of single crystal films.
Table 4-1. Properties of candidate gate dielectrics for GaN MOS device applications

<table>
<thead>
<tr>
<th></th>
<th>SiO₂</th>
<th>Si₃N₄</th>
<th>AlN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
<td>Amorph.</td>
<td>Amorph.</td>
<td>Hex.</td>
</tr>
<tr>
<td>Lattice Mismatch to GaN (%)</td>
<td>--</td>
<td>--</td>
<td>2.5</td>
</tr>
<tr>
<td>Bandgap (eV)</td>
<td>9.0</td>
<td>5.0</td>
<td>6.2</td>
</tr>
<tr>
<td>Dielectric Constant</td>
<td>3.9</td>
<td>5.0</td>
<td>8.5</td>
</tr>
<tr>
<td>Melting Point (°C)</td>
<td>1993</td>
<td>2173</td>
<td>3273</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Ga₂O₃</th>
<th>Gd₂O₃</th>
<th>ScO</th>
<th>MgO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
<td>Hex./Mono.</td>
<td>Bixbyite</td>
<td>Bixbyite</td>
<td>NaCl</td>
</tr>
<tr>
<td>Lattice Mismatch to GaN (%)</td>
<td>56</td>
<td>~20</td>
<td>9.2</td>
<td>4.216*</td>
</tr>
<tr>
<td>Bandgap (eV)</td>
<td>4.4</td>
<td>5.3</td>
<td>6.32</td>
<td>8</td>
</tr>
<tr>
<td>Dielectric Constant</td>
<td>10</td>
<td>11.4</td>
<td>14.5</td>
<td>8.1-9.6</td>
</tr>
<tr>
<td>Melting Point (°C)</td>
<td>2013</td>
<td>2670</td>
<td>2300</td>
<td>2800</td>
</tr>
</tbody>
</table>

* lattice constant (Å)
4.5.1. High Temperature Gd$_2$O$_3$

Single crystal Gd$_2$O$_3$ films were grown on n-GaN and characterized by RHEED, AFM, XRD, TEM, and SIMS. Typical Gd-cell and substrate temperatures during growth were 1230° and 650°C, respectively. A high-resolution cross-sectional transmission electron micrograph of the Gd$_2$O$_3$/GaN interface is given in Figure 4-12. From the image, the single crystalline nature of the oxide is obvious. Clear registry of the Gd$_2$O$_3$ lattice planes is also observed. In fact, additional TEM and AFM investigation indicated that the oxide served to planarize the rather rough GaN starting surface. The excellent registry of the interface is encouraging for effective surface passivation with small interface density. Still, a high dislocation density is expected in these films due to the large lattice mismatch with GaN. From Figure 4-12, at least 2 defects are noted in the oxide, which appear to be edge dislocations.

The thermal stability of the single crystal Gd$_2$O$_3$ films were investigated with x-ray diffraction (XRD). A Philips high resolution XRG 3100 5-crystal diffractometer was used to collect omega scans of a characteristic single crystal film both before (as-grown) and after rapid thermal annealing to 1000°C for 30 seconds. Figure 4-13 presents the data both before and after annealing. The as-grown sample exhibited a small peak shoulder, which actually decreased after 1000°C exposure. Thus, from x-ray data, the crystallinity of the Gd$_2$O$_3$ improved at the high temperatures required for device processing. This result is also encouraging for the long-term thermal stability of the device.
Figure 4-12. High resolution cross-sectional TEM of Gd₂O₃/GaN interface. Note single crystalline nature of oxide and excellent registry of lattice planes.
MOS capacitors were fabricated from single crystal Gd$_2$O$_3$ films by etching the oxide in concentrated HCl (5 min.) and 1:1 HCl:H$_2$O (5 min.) to open Ohmic windows, e-beam evaporation of Ti/Al frontside Ohmic metallization, and evaporation of circular Pt/Ti/Au (100/200/500 Å) Schottky contacts of 50 - 400 μm through a shadow mask. The Gd$_2$O$_3$/GaN MOSCs were characterized by I-V with an HP 4145B, and by C-V with an HP 4284A. Current-voltage characterization showed the single crystal diodes to be leaky, and this leakage is likely attributable to dislocations in the oxide acting as current leakage paths. The forward I-V characteristics from a typical single crystal Gd$_2$O$_3$/GaN MOSC are given in Figure 4-14, illustrating that the oxide does little to enhance the Schottky barrier height of ~0.8 V. Certain MOSCs were able to withstand higher forward voltages (Figure 4-15), but these results were not reproducible. Repeat measurements on the same device usually led to catastrophic failure after 2 to 3 measurements, even with a low current compliance. From TEM investigation, it was determined that rough GaN substrate surfaces led to increased defect density in the single crystal oxide. The discrepancies between diodes of equal diameter fabricated on the same oxide epilayers are likely caused by variations in surface quality of the nitride substrate prior to oxide growth. After failure, the forward I-V characteristics typically resembled those of a small barrier height Schottky contact (0.3 - 0.6 V). Reverse breakdown voltages were generally <10 V, and the previously described irreversible failure was observed. The significant oxide leakage precluded C-V characterization and modeling, since the equations developed for C-V analysis assume electrostatic behavior. Furthermore, an accumulation capacitance plateau could not be observed in the presence of forward current conduction, making it impossible to determine the oxide thickness self-consistently. Instead of
Figure 4-13. XRD omega scans of as-grown Gd$_2$O$_3$ film and same film after RTA at 1000°C. Note the slight decrease in peak width, corresponding to an increase in crystal quality upon annealing.

Figure 4-14. Forward I-V characteristics of typical single crystal Gd$_2$O$_3$/GaN MOSC.
reaching a constant accumulation capacitance, the single crystal Gd$_2$O$_3$ MOSCs displayed C-V curves as shown in Figure 4-16, where the capacitance decreases with increasing forward voltage. This is easily understood, since the forward current flow serves to dissipate the stored charge across the oxide layer, decreasing the capacitance.

The Gd$_2$O$_3$/GaN interface appeared promising, but the oxide could not withstand sufficient bias for a high voltage FET. To circumvent this problem, a thin SiO$_2$ film was evaporated onto the surface of the Gd$_2$O$_3$ to terminate the leakage paths associated with oxide dislocations. Metal oxide semiconductor field effect transistors (MOSFETs) fabricated from such a stacked gate dielectric will be described in detail in Chapter 5.

4.5.2 Low Temperature Gd$_2$O$_3$

To eliminate leakage from crystalline defects, low temperature Gd$_2$O$_3$ growth was investigated. By reducing the substrate temperature to 100°C, quasi-amorphous Gd$_2$O$_3$ films could be grown. The 100°C substrate temperature was effectively a lower limit due to the geometry of the growth system, since radiative heating from the Gd effusion cell maintained this temperature even with zero power to the substrate heater. RHEED patterns suggested that the films were amorphous, but XRD data indicated the presence of crystallites which oriented along the (111) and (321) Gd$_2$O$_3$ planes. The low temperature Gd$_2$O$_3$ films were stoichiometric, as determined by Auger depth profiling. From the same measurements, the interface was found to be abrupt and carbon-free within detection limits. The surface roughness was comparable to the high temperature, single crystal Gd$_2$O$_3$ (0.1 - 0.3 nm). This was augmented by SEM investigation, where no distinguishable features were noticed on the surface (Figure 4-17). Energy dispersive x-ray spectroscopy (EDX) measurements indicated the presence of only Ga, Gd, N, and
O species in the film, as shown in Figure 4-18. MOSCs from low temperature Gd$_2$O$_3$ films were fabricated as described previously, with only a difference in the oxide etch step. It was found that the etch rate of amorphous Gd$_2$O$_3$ was orders of magnitude faster than that of single crystal material. In general, the etch rate depended on the conditions for growth. An etch standard was necessary for each sample in order to accurately determine the etch time necessary for appropriate pattern transfer. This is not a crucial step for MOSC fabrication, but becomes much more important for the smaller, more critical dimensions of FETs. The fastest etch rates obtained from amorphous Gd$_2$O$_3$ films were $>2000$ Å/minute in 1:8 HCl:H$_2$O.

Current-voltage measurements from these films showed much better insulating properties. Typical forward and reverse characteristics are given in Figure 4-19 for a 50 μm diameter MOSC with oxide thickness of $\sim400$ Å. The reverse current was less than 1 nA at a bias of -15 V. In addition, there was very little current flow (< 10 nA) at forward bias to +10 V. The breakdown field for the low temperature Gd$_2$O$_3$ was estimated to be $\geq 3$ MV/cm, a significant improvement over the single crystal films. Capacitance voltage data were analyzed by the Terman Method to provide an estimate of the interface quality. The experimental data are given in Figure 4-20, overlayed with the calculated (theoretical) C-V curve. As expected, the real curve is stretched along the voltage axis due to the presence of interface traps. Following the analysis described in Section 4.4.3.1, the calculated relationship between gate voltage and band bending is given in
Figure 4-15. Forward I-V characteristics of ‘best’ single crystal Gd$_2$O$_3$/GaN MOSC.

Figure 4-16. C-V in accumulation region for single crystal Gd$_2$O$_3$/GaN MOSC, illustrating capacitance decrease at $V_G > 1.2$V due to oxide leakage.
Figure 4-17. SEM plan view micrograph of amorphous Gd$_2$O$_3$ film grown on n-GaN. The large feature at top right is a dust particle.

Figure 4-18. EDX scan of amorphous Gd$_2$O$_3$ on n-GaN.
Figure 4-19. Forward and reverse I-V characteristics of 50 μm diameter amorphous Gd$_2$O$_3$/GaN MOSC.

Figure 4-20. Capacitance-voltage from amorphous Gd$_2$O$_3$ MOSC.
Figure 4-21. From this relationship, the interface state density was appropriately positioned in the GaN bandgap according to Equation 4.13. The resultant plot of $D_{it}$ vs. $(E_c - E_t)$ is given in Figure 4-22, where $E_t$ represents the energy level of the trap state. The calculated interface state density of $6 \times 10^{11}$ cm$^{-2}$eV$^{-1}$ at $\sim 0.5$ eV below $E_c$ is higher than desired for GaN MOS device applications. Interface state densities below $10^{11}$ cm$^{-2}$eV$^{-1}$ are strongly preferred, and for silicon (100), values below $10^{10}$ cm$^{-2}$eV$^{-1}$ are easily obtained via post-oxidation annealing. In addition, room temperature Terman method data merely provide an estimate of $D_{it}$ and may be subject to gross errors (probably underestimates) depending on the conditions used in data collection (e.g. DC sweep rate). To appropriately determine $D_{it}$ for GaN MOSCs, it is necessary to collect capacitance-voltage or conductance-voltage data at elevated temperatures. This not only allows investigation closer to midgap, but also ensures that states near the band edges have sufficient time to respond to DC bias changes during the voltage sweep. Before attempting to measure elevated temperature device characteristics, verification of the thermal stability of the amorphous Gd$_2$O$_3$ films was attempted. Such an experiment is also necessary before any optimization of post-growth annealing conditions is undertaken.

A quasi-amorphous Gd$_2$O$_3$ film was annealed to 1000°C for 30 seconds in flowing N$_2$, similar to the experiment previously described for the single crystal Gd$_2$O$_3$ film. XRD data collected after the anneal strongly indicated the presence of recrystallization, with the appearance of diffraction peaks corresponding to a second Gd$_2$O$_3$ phase. Post-anneal electrical measurements indicated that the electrical properties of the film had severely degraded, and characteristics very similar to those from single
crystal films were observed. Combining the XRD and electrical data, it is clear that the increased leakage is attributable to grain boundaries formed after recrystallization. The polycrystalline films were shown to be electrically inferior to the amorphous films in all respects. Due to the relatively high interface state density and thermal instability, amorphous Gd$_2$O$_3$ was not investigated in additional detail.

A stacked dielectric consisting of a thin single crystal layer, followed by a thick low temperature layer was investigated to attempt to reduce $D_i$ while maintaining the insulating properties associated with the amorphous material. The thin (~100 Å) single crystal Gd$_2$O$_3$ was grown at high temperature, then the sample was cooled under an oxygen plasma to ~100°C to attempt amorphous growth. However, the single crystal template did not allow amorphous Gd$_2$O$_3$ to form at this growth temperature, and the resultant RHEED pattern indicated polycrystallinity throughout the low temperature film.

The electrical properties of this film were not tested, since it was previously demonstrated that both single crystal and polycrystalline films show excessive leakage. Even if the cap layer had grown amorphously, it is almost certain that the same recrystallization issues would have limited its use upon annealing to typical processing temperatures.
Figure 4-21. Gate voltage-band bending relationship used to determine $D_A$ for amorphous Gd$_3$O$_3$/GaN MOSC.

Figure 4-22. Interface state density of low temperature, amorphous Gd$_2$O$_3$ grown on (0001) GaN.
4.6. Scandium Oxide

Scandium oxide (ScO) was grown by GSMBE and characterized as a candidate dielectric for GaN MOS applications. ScO exists in the Bixbyite crystal structure, but the lattice mismatch between ScO and GaN is less than half that of Gd$_2$O$_3$ (9.2% vs. 20%), as given in Table 4-1. The ScO bandgap of 6.3 eV suggests larger band offsets than Gd$_2$O$_3$/GaN. The exact band alignment is unknown, since ScO/GaN heterostructures have never before been investigated.

Single crystal ScO films with ScO (111) // GaN (0001) were confirmed by high resolution XRD. This is the same orientation observed for Gd$_2$O$_3$ and is expected due to similar atomic symmetry along these lattice planes. Surface cleaning prior to GSMBE oxide growth was accomplished using the procedure detailed in Section 4.2. The substrate and Sc effusion cell temperatures were 575°C and 1130°C, respectively. Growth proceeded at $1 \times 10^{-4}$ in the ECR plasma head at 200 W forward power. The growth rate was ~8 Å/min. under these conditions. Surface roughnesses of ~0.6 nm for ScO grown on n-GaN were measured by AFM.

MOS capacitors were fabricated as described previously, with the exception of an ICP dry etch to open Ohmic windows. A 10 Cl$_2$ / 5 Ar chemistry with 100 W rf and 300 W ICP power etched the ScO at ~120 Å/min. Dry etching was employed due to the current lack of a suitable room temperature wet etch chemistry for this material. Wet etching in concentrated acids at 80°C has been demonstrated, but standard photoresists used for pattern transfer are quickly consumed in strongly acidic solutions at this temperature. After Ohmic deposition and annealing, gate contacts were defined photolithographically, instead of with a shadow mask.
The 25°C I-V characteristics of a 50 μm ScO MOSC are given in Figure 4-23 on both linear and logarithmic axes. The oxide thickness calculated from the growth rate was ~500 Å. Note the very small current scale of the Figure. The reverse current is ~20 pA at -4V, while the forward current is ~100 nA at +4 V. These data are far superior to any previous results from single crystal Gd2O3. Such results may be attributed to reductions in crystalline defects and dislocations in the oxide due to the smaller lattice mismatch. Alternatively, the microstructure of the films may be such that the dislocations are not vertical in nature, and thus may not act as current leakage paths from semiconductor to gate metal. TEM investigation would help to address many of these issues. The electrical properties of the oxide did not degrade with multiple voltage sweeps in a ±4 V window. Capacitance-voltage data were collected from the MOSC of Figure 4-23. Clear modulation from accumulation to depletion was observed, with capacitance values of the correct magnitude. However, the capacitance from near-flatband to accumulation increased monotonically with gate voltage to +4 V, and no accumulation plateau was observed in this range. Furthermore, the dielectric constant of ScO deposited by GSMBE on GaN is still unknown. An exact knowledge of the dielectric constant is necessary for calculation of any parameters from ideal C-V theory. Additional work to characterize both ScO and the ScO/GaN interface is warranted.
Figure 4-23. Linear (top) and semilog (bottom) I-V characteristics of ScO/GaN MOSC at room temperature.
CHAPTER 5
GaN METAL OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTORS

5.1 Introduction

Metal oxide semiconductor field effect transistors (MOSFETs) were demonstrated by Kahng and Atalla of Bell Labs in 1960 [244]. MOSFETs consume less power than bipolar devices, and are therefore much more suitable for modern ULSI circuitry. Since Kahng and Atalla’s first device, MOSFET technology has been completely dominated by silicon, due in part to its stable native oxide. Certain III-V compounds, such as GaAs, have been heavily researched for their potential for high speed MOS devices. However, GaAs does not have a stable native and until recently, surface state densities at the oxide/semiconductor interface prohibited the useful operation of such devices. In 1997, Ga$_2$O$_3$(Gd$_2$O$_3$) evaporated from a single crystal Ga$_5$Gd$_3$O$_{12}$ source was shown to effectively passivate GaAs and InGaAs, and both n-and p-type enhancement mode MOSFETs were demonstrated [233,234]. The gate dielectric for these devices was a mixed Ga$_2$O$_3$(Gd$_2$O$_3$), or GGG (gallium gadolinium garnet).

As described in Chapters 1 and 2, GaN FETs have shown impressive DC performance and record output power at high frequencies. However, leakage from the Schottky gate of such devices is a major concern for operation at elevated temperatures. In addition, a Schottky gate recess etch is difficult due damage induced in the surface region. The MOS structure reduces leakage current relative to GaN MESFETs, as well as exhibiting lower parasitic resistances, reduced power consumption, and improved linearity [245-247].
Like GaAs, GaN does not possess a stable native oxide with low density of interface states. Results from GGG on GaAs led to the demonstration of the first GaN MOSFET in 1998. Low leakage currents and good charge modulation were demonstrated, with breakdown field of 12 MV cm\(^{-1}\) demonstrated for the GGG insulator. This depletion mode device showed an extrinsic transconductance of 15 mS/mm and improved performance at elevated temperature relative to a conventional metal semiconductor FET (MESFET), but suffered from high contact resistance and parasitic capacitance. Later work by Ren et al. [248] investigated the effect of temperature on the gate leakage of depletion mode GGG/GaN MOSFETs. Results showed significantly reduced gate leakage at elevated temperatures.

Other amorphous oxides, namely SiO\(_2\) and Si\(_3\)N\(_4\), have been investigated [249-253] and MOS capacitors fabricated using this dielectric have shown promising results, with interface state densities below 10\(^{11}\) cm\(^{-2}\) reported for an SiO\(_2\)/Si\(_3\)N\(_4\)/SiO\(_2\) stacked dielectric [252]. Khan and co-workers have demonstrated impressive results from insulated gate HEMTs using PECVD-deposited SiO\(_2\) [254-256]. However, an insulator with dielectric constant comparable to or greater than that of GaN (\(\varepsilon_r \sim 9.8\)) is highly desirable. The small dielectric constant of SiO\(_2\) (\(\varepsilon_r \sim 3.9\)) may limit the high voltage performance of SiO\(_2\)/GaN devices. An alternative is the use of single crystal Gd\(_2\)O\(_3\), which has been described in detail in Chapter 4. In this Chapter, the fabrication and DC performance of depletion mode, n-channel Gd\(_2\)O\(_3\)/GaN MOSFETs will be described. The device results presented herein represent a significant improvement over previously reported GGG/GaN devices.
5.2 MOSFET Processing

The nitride epilayer structure for all devices in this chapter consisted of 2 μm unintentionally doped GaN and a 700Å Si-doped GaN channel grown by MOCVD on a c-Al₂O₃ substrate. Hall measurement indicated a channel doping level in the low-10¹⁸ cm⁻³. The wafer was prepared for oxide growth by a 3 min. dilute HCl dip, 25 min. UV-ozone exposure, and 5 min. dilute HF dip. The Gd₂O₃ was grown in a Riber 2300 gas-source molecular beam epitaxy (GSMBE) system equipped with a Wavemat MPDR 610 electron cyclotron resonance (ECR) plasma source. After loading, the wafer was heated to 700°C to desorb the native oxides, then cooled to the growth temperature of 650°C. Growth proceeded at 1×10⁻⁴ torr oxygen pressure in the ECR plasma head at 200W forward power, with a Gd cell temperature of 1230°C. X-ray diffraction showed a single crystalline Gd₂O₃ structure for the 700Å film.

Oxygen ion implantation for device isolation was simulated using the Transport-of-Ions-in-Matter (TRIM) code. The TRIM code uses statistical and quantum mechanical algorithms to simulate the interactions of ions with solids. The O⁺ implantation was performed at an angle of 7° off the surface normal through the gate insulator. The density of GaN was taken from tabulated values and that of Gd₂O₃ was calculated from unit cell dimensions and occupancy. The implant was designed to compensate the ~3 × 10¹⁸ cm⁻³ carriers in the epilayer between devices, reducing crosstalk. Due to the Gaussian shape of the ion stopping profile at a given implant energy, the use of multiple energies was necessary to provide a uniform vacancy concentration throughout the FET channel. Implant energies of 60, 145, 160, and 200 keV were employed, each at an implant dose of 4 × 10¹² cm⁻². The resultant vacancy profile of each implant and the overall profile are
given in Figure 5-1. Note the vacancy concentration of \( >5 \times 10^{20} \text{ cm}^{-3} \) in the channel and
\( >1 \times 10^{20} \text{ cm}^{-3} \) to a depth of \( \sim3000\text{Å} \). From previous reports concerning \( \text{O}^+ \) implantation
in GaN, this implant recipe should provide a sheet resistance on the order of \( 10^{12} \)</p>
\( \Omega/\text{square} \) [257].

![Vacancy profiles simulated by TRIM code for 4 discrete \( \text{O}^+ \) implant energies. Note the discontinuities at the \( \text{Gd}_2\text{O}_3/\text{GaN} \) interface caused by the abrupt change in density.](image)

After implant isolation, device fabrication began with a wet etch of \( \text{Gd}_2\text{O}_3 \) to open
source and drain ohmic contact windows using HCl as the etchant and photoresist \( \text{AZ-1818} \) as the mask. A 15 second Ar dry etch at 50 W rf and 200 W ICP was used to
remove any interfacial layer/contaminants remaining on the surface after wet etching.
Any damage induced from such an etch would produce donorlike defects on the surface
of the contact region, and have been shown to actually improved contact resistance.
Ti(200Å)/Al(600Å)/Au(1000Å) ohmic contacts were defined by electron-beam evaporation and liftoff. Gate metallization was e-beam deposited Pt/Ti/Au. A schematic of the processing sequence and final device structure of the planar n-channel depletion mode GaN MOSFET is given in Figure 5-2.

On some of the devices, a 300Å layer of SiO₂ was used to enhance the gate breakdown voltage. The thin SiO₂ layer was electron-beam evaporated onto the Gd₂O₃ surface in the photolithographically-defined gate contact region. This was followed under vacuum by evaporation of Ti(100Å)/Au(1000Å) gate contacts to the SiO₂ surface. The final structure of an SiO₂/Gd₂O₃/GaN MOSFET is given in Figure 5-3. Devices without this SiO₂ cap layer will be discussed in Section 5.3.1, while those with the stacked dielectric will be discussed in Section 5.3.2.
Figure 5-2. Processing sequence of Gd$_2$O$_3$/GaN MOSFET.
Figure 5-3. Schematic of SiO$_2$/Gd$_2$O$_3$/GaN depletion mode, n-channel MOSFET.

5.3 Results and Discussion

5.3.1 Gd$_2$O$_3$/GaN Devices

When deposited under optimum conditions, Gd$_2$O$_3$ was grown epitaxially on GaN, as verified by reflection high energy electron diffraction (RHEED) and cross-sectional transmission electron microscopy (XTEM). As discussed in Chapter 4, the Gd$_2$O$_3$ interface was quite smooth despite a somewhat rough GaN wafer surface. However, there was still a relatively large difference in lattice spacing between Gd$_2$O$_3$ and GaN, which contributed to a high number of dislocations in the oxide film. These dislocations acted as current leakage paths and caused the breakdown voltage of the as-grown Gd$_2$O$_3$ to be rather low. Functional MOSFETs were fabricated using the as-grown single crystal Gd$_2$O$_3$, but the forward gate voltage swing was enhanced little from that of a conventional Schottky contact ($\sim +1$ V). Typical drain I-V characteristics are given in Figure 5-4 for a $6 \times 100 \, \mu$m$^2$ gate dimension Gd$_2$O$_3$/GaN device. Figure 5-5 gives the gate current measured during the collection of the drain I-V data of Figure 5-4. The
figure clearly shows significant leakage at $V_G = +2 \text{ V}$. The device could not be modulated at more positive gate voltages due excessive gate current. Still, these devices represented the first demonstration of functional GaN MOSFETs using Gd$_2$O$_3$ as the gate insulator. Devices remained operational at 150°C, with very little change in output characteristics, as shown in Figure 5-6.

Figure 5-4. Drain I-V characteristics of $6 \times 100 \ \mu\text{m}^2$ gate dimension Gd$_2$O$_3$/GaN MOSFET.
Figure 5-5. Gate current traces taken during drain voltage sweeps of Figure 5-4.

Figure 5-6. Drain I-V characteristics of Gd$_2$O$_3$/GaN MOSFET at 150°C.
5.3.2 SiO$_2$ / Gd$_2$O$_3$ / GaN Stacked Dielectric Devices

In order to improve the breakdown and increase voltage swing, amorphous SiO$_2$ was deposited on the Gd$_2$O$_3$ surface under the gate contact. The formation of a stacked gate dielectric of SiO$_2$/Gd$_2$O$_3$ allowed the interfacial properties of Gd$_2$O$_3$/GaN to be maintained, while reducing leakage by terminating the dislocations in the crystalline oxide. Figures 5-7 and 5-8 illustrate the significant reduction in gate leakage after SiO$_2$ deposition. The breakdown field of the insulator improved from 0.3 to 0.8 MV/cm and, from Figure 5-6, the reverse gate current was only $\sim$10 pA at $V_G = -10$ V. From Figure 5-7, gate reverse leakage current was $\sim$10 nA at $V_{GS} = -70$V for the $1 \times 200$ $\mu$m$^2$ SiO$_2$/Gd$_2$O$_3$/GaN device, clearly demonstrating the benefit of the insulated gate MOS structure. Comparison with the gate characteristics before SiO$_2$ evaporation clearly indicate the improvement in insulating behavior.

Capacitance-voltage data were measured at room temperature for large gate area devices using a HP4284 LCR meter. Figure 5-9 illustrates C-V characteristics of a $36 \times 100$ $\mu$m$^2$ device measured in the dark at a sweep rate of 100 mV/second and frequencies from 1 kHz to 1 MHz. Clear charge modulation from accumulation to depletion is observed. The absence of surface inversion is due to the wide band gap and correspondingly low room temperature minority carrier generation rate of GaN, as explained in Chapter 4. Both the forward and reverse sweep data are included in Figure 5-9 for the SiO$_2$/Gd$_2$O$_3$/GaN structures, and a slight capacitance hysteresis is observed. However, additional data collected at 100 kHz from a 400 $\mu$m diameter Gd$_2$O$_3$/GaN diode illustrated negligible hysteresis, suggesting that the hysteresis in Figure 5-9 may be related to the SiO$_2$, not Gd$_2$O$_3$. More work is needed to characterize the interfacial and
bulk electrical properties of these Gd$_2$O$_3$/GaN structures. From a plot of $1/C^2$ vs. V for the long-gate MOSFET, the doping profile of the GaN layer structure was calculated and is plotted in Figure 5-10 as a function of applied gate bias. The doping concentration in the channel was $\sim 3 \times 10^{18}$ cm$^{-3}$, consistent with Hall measurements, and the carrier concentration in the bulk UID GaN was $\sim 8 \times 10^{15}$ cm$^{-3}$.

Drain I-V characteristics of a $1 \times 200$ $\mu$m$^2$ gate dimension SiO$_2$/Gd$_2$O$_3$/GaN MOSFET were measured with an HP4145A and are shown in Figure 5-11. The drain-source breakdown exceeded 80V. For these high breakdown devices, charge modulation could be demonstrated for gate voltages from +2 to -5V. Other depletion mode devices from the same wafer could be modulated to an accumulation bias of +7V, as shown in Figure 5-12. This large forward voltage clearly verifies the high quality of the Gd$_2$O$_3$/GaN interface. A maximum intrinsic transconductance of 61 mS/mm was measured at $V_{GS} = -0.5$ V and $V_{DS} = 20$ V. The high output conductance of the drain IV in Figure 6 may have been caused by short channel effects due to the high-$10^{15}$ cm$^{-3}$ doping in the buffer layer. The device had high knee voltage and parasitic resistance because the ohmic contact was annealed only to $\sim 300^\circ$C. Although the O$^+$ isolation implant provided a good sheet resistivity for the as-implanted sample, the thermal stability of the implanted samples was quite poor [257,258]. After annealing the sample at 400$^\circ$C for 1 min., the sheet resistivity has been shown to decrease by more than an order of magnitude. To optimize the ohmic contact annealing and device isolation processes, Cr$^+$ and Fe$^+$ implantation may be needed. Both of them showed good thermal stability to 600$^\circ$C [130].
Figure 5-7. Gate I-V characteristics of $1 \times 200 \, \mu m^2$ gate dimension MOSFET. Note the reverse current of only $\sim10 \, pA$ at $V_G = -10 \, V$.

Figure 5-8. Comparison of gate leakage before and after SiO$_2$ evaporation.
Figure 5-9. Capacitance-voltage of SiO$_2$/Gd$_2$O$_3$ dielectric stack MOSFET.

Figure 5-10. Doping profile obtained from C-V measurements. The channel doping density agrees well with Hall effect data.
Figure 5-11. High voltage output characteristics of GaN MOSFET.

Figure 5-12. Output characteristics of GaN MOSFET illustrating gate modulation to +7 V.
CHAPTER 6
GaN SCHOTTKY RECTIFIERS ON BULK GaN SUBSTRATES

6.1 Introduction

GaN electronic device research has been largely dominated by FETs. MESFETs and HEMTs for high frequency, high power applications have been developed to exploit the attractive material properties of the III-nitrides, as described in Chapter 3. A device which has received considerably less attention is the power rectifier. Wide bandgaps allow III-nitrides to sustain extremely high critical electric fields, leading to large blocking voltages. Applications for these devices are numerous, and their potential technological and commercial importance is beginning to take shape.

Mechanical and Si-based switches are presently used to control electric current flow across utilities transmission and distribution lines. Opening or closing these switches can lead to large power sags and switching transients delivered to the load. Such transients may be detrimental, for instance, to major computing centers, motor drives, computers, digital controls, or other sensitive electronic equipment. An outage of less than one cycle, or a voltage sag of 25% for two cycles can cause a microprocessor to malfunction. As a result of these potential fluctuations, the electric power grid must be operated at capacities well below its rated value, leading to reduced energy efficiency. A system for eliminating power sags and switching transients would dramatically improve power quality [193,259]. Solid state devices, if available, are expected to show “clean” switching and could potentially eliminate line transients and allow more efficient operation of the grid. In addition, typical power devices are required to operate at
Elevated temperatures due to the power dissipation associated with switching large currents and voltages. In this respect, wide bandgap switches are attractive due to their increased tolerance to temperatures above the limits of silicon. Reduction of bulky, expensive cooling equipment should be possible, leading to decreased system complexity and cost. Other end-uses include electronic motor controls, lighting, heating, and air-conditioning.

The GaN material system has a high critical field, good saturation electron velocity and reasonable thermal conductivity if bulk wafers are available. A key component of the inverter modules required for many of the previously mentioned applications is the simple rectifier. There have been a number of reports of mesa and lateral geometry GaN and AlGaN Schottky and p-i-n rectifiers fabricated on heteroepitaxial layers on Al₂O₃ substrates [10,198,261-264]. The geometry of such a device is illustrated in Figure 6-1. A major disadvantage of this approach is the poor thermal conductivity of sapphire (κ=0.5 W/cm·K) and the limited epilayer thicknesses employed. In this regard, better substrate choices would be either SiC or GaN itself,
since the latter has approximately the same thermal conductivity as Si, as shown in Figure 6-2. It should be noted that the typically cited value of thermal conductivity for GaN (1.3 W/cm·K) is effectively a lower limit, as suggested by recent studies. It has been shown that both defect density and carrier concentration can significantly affect the thermal conductivity [265]. Values near 2 W/cm·K have been experimentally demonstrated for GaN. The availability of bulk GaN substrates would allow fabrication of vertical geometry rectifiers capable of much higher current conduction than lateral rectifiers fabricated on insulating substrates [266].

Figure 6-2. Thermal conductivities of sapphire, silicon, GaN, and SiC. For GaN and SiC, the given data points represent lower limits. The actual values may fall along any point of the vertical error bars, depending on crystal quality, impurity concentration, polytype, or other factors.
6.2 Material Characterization and Diode Fabrication

Recently, high quality free-standing GaN templates have been grown by hydride vapor phase epitaxy (HVPE) [55,267]. Using this technique, thick GaN layers (~300 μm) are grown at a high deposition rate on conventional sapphire substrates and subsequently removed by laser heating using a photon energy greater than the GaN bandgap. Chemical mechanical polishing (CMP) is used to smooth the N-face (nearest substrate), while both CMP and dry etching are used to smooth the Ga-face (epilayer surface) after lift-off [58]. Wet chemical etching removed additional heavily damaged, highly conducting material from the N-face. The resultant ~200 μm free-standing GaN templates used in this study were specular and transparent, with very little cloudiness. The substrate dimensions were typically ~ 8 × 10 mm², as shown in Figure 6-3. Defect densities were ~10⁵ cm⁻² for the Ga-face and ~10⁷ cm⁻² for the N-face [59]. Atomic force micrographs of both faces of a typical GaN template are given in Figure 6-4, where the difference in the surface morphology is evident. Root-mean-square (RMS) roughnesses for the Ga- and N-faces were < 1 nm and ~10 nm, respectively. From Figure 6-4, scratch marks from the CMP process are easily discernable. The crystal quality of the as-grown templates was investigated by high resolution x-ray diffraction (HRXRD). A Philips high resolution XRG 3100 5-crystal diffractometer in triple axis mode was used to collect ω and ω/2θ scans from the symmetric (0002) reflection of both the Ga- and N-faces. Full widths at half maximum (FWHMs) from these data are summarized in Table 6-1 and the x-ray spectra are shown in Figure 6-5. Capacitance-voltage measurements showed an n-type doping level of ~8 × 10¹⁶ cm⁻³.
Table 6-1. FWHM values from (0002) reflection of bulk GaN template

<table>
<thead>
<tr>
<th></th>
<th>$\Omega/2\theta$ scan</th>
<th>$\omega$ scan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ga-face</td>
<td>6</td>
<td>808</td>
</tr>
<tr>
<td>N-face</td>
<td>19</td>
<td>818</td>
</tr>
</tbody>
</table>

On some of the substrates, 6.8 $\mu$m of undoped GaN was grown on the Ga-face by metalorganic chemical vapor deposition (MOCVD) at 1040°C using conventional precursors. After epilayer growth, photoluminescence (PL) data were collected from bare substrate samples as well as samples with MOCVD epilayers. Figure 6-6(a) shows PL spectra taken at 25K and 298 K from the Ga-face of the bare substrate. Note that band-edge luminescence is only observed at low temperatures. After growth of the 6.8 $\mu$m thick epilayer on top of the substrate the room temperature PL spectrum is dominated by the yellow band centered at around 660 nm.

Vertical geometry Schottky diodes were fabricated by electron beam deposition of full-area backside (N-face) Ti/Al/Pt/Au ohmic contacts, heated to 750°C for 30 secs. After annealing the Ohmic contacts, a 5000Å layer of SiO$_2$ was deposited on the Ga-face by RF (13.56 MHz) plasma-enhanced chemical vapor deposition (PECVD) using SiH$_4$ and N$_2$O as precursors. A dielectric window was patterned and the SiO$_2$ was etched with CF$_4$/O$_2$ reactive ion etching (RIE). After RIE, the sample was briefly immersed in diluted (10:1) BOE to ensure complete removal of the oxide in the Schottky contact regions. The Schottky metal was patterned with AZ1808 photoresist and front-side Pt/Au rectifying contacts with diameters 44-148 $\mu$m were defined by lift-off. Schematics of the processing geometry and final vertical rectifier structure are given in Figure 6-7. The Schottky window was slightly larger than the dielectric window so the metal would overlap the dielectric. Since the Schottky contact is terminated on the oxide, edge effects
leading to electric field crowding at the surface are reduced. Current-voltage (I-V) characteristics were recorded using a HP 4145B Parameter Analyzer with samples held at 25-300°C by a heated stage.

Figure 6-3. Photograph of typical free-standing GaN substrate, illustrating size and color clarity. Note the absence of cloudiness.
Figure 6-4. AFM scans of N-face (top) and Ga-face (bottom) of GaN substrate. Note the difference in the vertical scale. Scratch marks from the CMP process are visible.
Figure 6-5. Triple axis high resolution x-ray diffraction scans from Ga-face (top) and N-face (bottom) of free-standing GaN template.
Figure 6-6. Photoluminescence spectra from GaN substrate at 25 and 298 K (top) and 298 K PL from the 6.8μm thick MOCVD epilayer grown on the GaN substrate (bottom).
6.3 Current-Voltage Results and Discussion: Schottky Diodes

Figure 6-8 shows I-V characteristics from 75 μm Schottky rectifiers fabricated either directly on the GaN substrate, or on the epilayer/substrate structure. The reverse breakdown voltage is not yet observed for the voltages shown in Figure 6-8, but the bare substrate device conducts a discernable amount of current at reverse biases above -250 V. The breakdown voltages for the samples shown in Figure 6-8 were ~300 V for the bare substrate device and ~450 V for the epilayer/substrate device. In the case of the epilayer/substrate structure, larger values of $V_B$ were obtainable due to the lower doping ($\sim 5 \times 10^{16} \text{ cm}^{-3}$) in the MOCVD-grown epilayer relative to the HVPE substrate ($>10^{17} \text{ cm}^{-3}$). Improvements in HVPE growth, particularly decreasing the background impurity concentration in these thick layers, will be central toward the realization of higher blocking voltages in future devices.
Figure 6-8. I-V characteristics of 75 μm diameter GaN Schottky diodes on bare substrate and epilayer/substrate structure.

Figure 6-9 shows the reverse I-V characteristics as a function of Schottky contact diameter for rectifiers fabricated on the bare substrate and the epilayer/substrate structure, as a function of the contact diameter. The general trend is for a decrease in $V_B$ with increasing contact diameter. This reflects the higher probability for having defects within the active area that lead to premature breakdown, as also observed for SiC devices [205,268,269]. The same general trend is observed for both devices, but the overall magnitude of the $V_B$ values is higher for the devices with the MOCVD epilayer, as described above. These values are fairly typical of reported numbers for vertically-depleting GaN diode rectifiers [198,264-266]. Note that the data is too scattered to extract a quantitative relationship.
Figure 6-9. Reverse I-V characteristics for diodes of various Schottky diameters fabricated on bare substrate (top) and epilayer/substrate samples (bottom).
Figure 6-10 shows the variation of reverse current at -300 V reverse voltage for rectifiers fabricated on the epi/substrate structure, as a function of contact diameter. At this bias, it is difficult to determine if the current scales with contact diameter (suggesting surface-related leakage) or with contact area (suggesting bulk leakage). Surface-related leakage is typical of most of the GaN rectifiers reported to date and emphasizes the need for edge termination methods to reduce the influence of defects around the periphery of the Schottky contact. Applications for these types of rectifiers in power switching or power conditioning equipment will require very large area devices (many cm²), emphasizing the need to reduce the density of active defects in the GaN material in order to keep the reverse current to a manageable level.

Figure 6-11 shows the temperature dependence of the reverse I-V characteristics of epilayer/substrate rectifiers with 54 μm Schottky contact diameter. The $V_B$ values show a negative temperature coefficient $\beta$, where

$$V_B = V_{B0} \left[ 1 + \beta(T-T_0) \right] \quad (6.2)$$

where $V_{B0}$ is the breakdown voltage at room temperature ($T_0$) and $T$ is the absolute sample temperature. Previous results on both lateral and vertical rectifiers have indicated that $\beta$ is strongly dependent on device geometry and the defect density of the material. From the data of Figure 6-11, an estimate of $\beta = -2.5 \pm 0.6$ V·K⁻¹ is obtained for these rectifiers. In principle, the temperature coefficient should be positive for GaN, i.e. the breakdown voltage should increase with increasing temperature [198,263]. However the presence of a high defect density leads to a negative value [204,205]. It may be possible that only certain types of defects contribute to this effect. Recently, the inhomogeneous spatial distribution of leakage current in GaN Schottky diodes was attributed to the
presence of dislocations containing a screw component [270]. Certainly, more work is needed to understand the negative temperature coefficient of reverse breakdown and the underlying physical mechanisms responsible for it.

Forward current density-voltage-temperature (J-V-T) characteristics for 54 μm rectifiers fabricated on the epi/substrate structure are shown in 6-12. The forward turn-on voltage (V\text{on} -- defined as the forward voltage at which the current density is 100 A·cm\(^{-2}\)) was as small as 3 V at 25°C. The values of V\text{on} increased with temperature. This is somewhat surprising, given that more efficient electron emission over the Schottky barrier might be expected at higher temperatures. However, the positive correlation between V\text{on} and T is consistent with previously reported devices, where the reduction in barrier height was attributed to the temperature-dependent capture cross sections of surface states [271]. Figure 6-13 shows that the ideality factor of these rectifiers is ~1.25, as extracted from a the low voltage slope of the J-V plot. This value is not ideal, but is comparable to previously reported GaN Schottky devices and does indicate a reasonably good contact. The best on-state resistance was 20.5 mΩ·cm\(^{-2}\) for epilayer/substrate rectifiers with V\text{B} = 450 V, producing a figure-of-merit (V\text{B}\(^2\) / R\text{ON}) of ~10 MW·cm\(^{-2}\).
Figure 6-10. Reverse breakdown voltage as a function of Schottky contact diameter. The decrease in $V_B$ with increasing diameter is attributed to the enhanced probability of electrically active defects existing within the device area.

Figure 6-11. Reverse current at constant -300 V reverse bias for diodes of different Schottky diameters, illustrating the correlation between device size and amount of leakage current.
Figure 6-12. Reverse I-V-T characteristics of 54 μm diameter devices fabricated on the epilayer/substrate structure. The leakage current is greatly enhanced at elevated temperatures, a characteristic which will need to be controlled for successful implementation of these devices for high temperature applications.

Figure 6-13. Forward J-V-T characteristic up to 300°C. The turn-on voltage is 3 V at 25°C and increases to ~5 V at 300°C.
Figure 6-14. Low bias J-V curve illustrating ideality factor of 1.25 at 100°C for a 75 μm diameter device.

6.4 Large- and Small-Area Bulk GaN Rectifiers with Implanted Guard Rings

The simple Schottky diodes described in Section 6.3 used a dielectric overlap passivation technique to reduce field crowding along the Schottky contact periphery. Implantation of Mg⁺ guard rings has also been successfully employed as an edge termination technique in III-N rectifiers [261,272]. This section describes Schottky rectifiers with implanted p⁺ guard ring edge termination fabricated on free-standing GaN substrates. These bulk GaN rectifiers show significant improvement in forward current density and on-state resistance over previous heteroepitaxial devices.

The free-standing substrates were described in Section 6.2. P-type guard rings (30 μm diameter) were formed by selective area Mg⁺ implantation at 50 keV, $5 \times 10^{14}$ cm⁻². Then implant was followed by an 1100°C, 30 second anneal to remove residual lattice damage. Ohmic contacts were deposited as described in Section 6.2. On some of the samples, Ohmic contacts were also deposited on the front surface so comparisons
could be made between vertical and lateral device geometries. Schottky contacts of e-beam evaporated Pt/Ti/Au with diameters of either 54-75 μm for small-area devices or ~5 mm for large-area devices were placed on the front (Ga-face) surface. For the large area devices the contact was extended over a PECVD SiO₂ passivation layer. CCD images and schematic cross-sections of both the small and large area GaN rectifiers are shown in Figures 6-15 and 6-16.

The large-area diode had a reverse breakdown voltage, \( V_B \), of only 6 V. By sharp contrast, the small-area diodes show \( V_B \) values of ~120 V in the lateral mode and ~160 V in the vertical mode. The I-V characteristics from the small-area vertical diodes are shown in Figure 6-17. These results have several implications. Firstly, they demonstrate that the presence of defects in the depletion region or around the contact periphery have a major effect on the electrical characteristics. Since the defect density in the quasi-substrates was ~10⁵ cm⁻² as measured by combined photo-chemical etching and atomic force microscopy [58,59], the large area rectifiers are highly likely to include one or more defects. There will obviously be a higher concentration of total defects (by roughly a factor of 10⁴) in the region under the contact of the large area devices. In defect-free material, \( V_B \) is related to the maximum electric field strength at breakdown, \( E_M \), by the equation

\[
V_B = \frac{E_M W_M}{2}
\]

(6.3)

where \( W_M \) is the depletion depth at breakdown. However it has been amply demonstrated in other materials systems such as SiC that the presence of defects
Figure 6-15. CCD image of small-area implanted Schottky rectifiers (top) and schematic of device cross-section (bottom).
Figure 6-16. Photograph of large-area implanted Schottky rectifiers (top) and schematic of device cross-section (bottom).
surface defects around the periphery of the devices will also degrade the breakdown and increase reverse leakage. The second key implication from Figure 6-17 is that the small area diodes show a larger $V_B$ in the vertical geometry. This indicates that the surface plays a major role in determining $V_B$, because even though the vertical GaN thickness is ~200 μm and the Ohmic-gate spacing is ~30 μm for the lateral diodes, both types of devices are in the punch-through region.

![Graph](image)

**Figure 6-17.** Current-voltage characteristics from implanted vertical geometry rectifiers at 25°C.

An expanded view of the forward characteristics is given in Figure 6-18. Figure 6-19 shows the temperature dependence of reverse current in a typical 75 μm diameter vertical geometry rectifier (top), along with the temperature dependence of the forward current. These bulk rectifiers still show poor thermal characteristics, with both forward and reverse leakage increasing rapidly with operating temperature. The reverse breakdown still showed a negative temperature coefficient, similar to previous work,
suggesting that more work needs to be done to reduce the surface and bulk defect density. The reverse current was thermally activated with an activation energy of $0.11 \pm 0.04$ eV and this may represent the most prominent surface state giving rise to the current. Over the small range of diameters, reverse current was proportional to contact diameter. This indicates that the reverse current originates from surface periphery leakage.

![Semilog plot of implanted GaN diode forward I-V characteristics at 25°C. Note the change in slope due to series resistance beginning at ~2V.](image)

Figure 6-18. Semilog plot of implanted GaN diode forward I-V characteristics at 25°C. Note the change in slope due to series resistance beginning at ~2V.

Figure 6-21 shows a plot of avalanche and punch-through breakdown of GaN Schottky rectifiers calculated as a function of doping concentration and active layer thickness. For example, a 200 μm thick sample with a doping of $10^{15}$ cm$^{-3}$ (which is quite feasible by reducing the background Si and O content, or by approximate compensation) would have a predicted $V_B$ of $\sim 1.5 \times 10^4$ V [7]. These would have application for power control systems in the 13.8 kV class [193,275,276]. One can expect major improvements in $V_B$ in quasi-bulk GaN rectifiers as the background doping
is decreased. The data points included in Figure 6-21 represent $V_B$ values from the present work. The biggest issue facing GaN rectifiers achieving very high breakdown voltages is now to reduce the background doping in the free-standing substrates. Previous work on lateral diodes fabricated on fairly resistive GaN and AlGaN showed $V_B$ values of several kV, even though the defect density was very high($>10^8$ cm$^{-2}$) [272].

![Constant $V_R = -100$ V](image)

**Figure 6-19.** Temperature dependence of reverse leakage current in 75 μm diameter vertical geometry implanted rectifiers

The specific on-state resistances ($R_{ON}$) for the three types of rectifiers measured were 3 mΩ·cm$^2$ for the small-area vertical diodes, 1.7 mΩ·cm$^2$ for the small-area lateral diodes, and 3.4 Ω·cm$^2$ for the large-area devices. The small-area on-state resistance is the lowest reported for any GaN rectifier and shows that continued improvements in surface cleaning and contact technologies for this materials system have led to a rapid maturation of device processing techniques. To place these results in context, Figure 6-
22 shows a plot of $R_{\text{ON}}$ versus $V_B$ for GaN rectifiers reported in the literature. The lines show theoretical results for Si, SiC, and GaN rectifiers from Khemka et al. [277]. There has been a steady improvement in the on-state resistances over the past few years as material and processing quality have improved.

The forward turn-on voltage, $V_F$, for a Schottky rectifier is given by

$$V_F = \frac{n k T}{q} \ln \left( \frac{J_F^*}{A^* T^2} \right) + n \phi_B + R_{\text{ON}} \cdot J_F$$

(6.4)

where $n$ is the ideality factor, $k$ is Boltzmann’s constant, $T$ the absolute temperature, $q$ the electronic charge, $A^*$ is Richardson’s constant, $J_F$ is the forward current density at $V_F$, and $\phi_B$ the Schottky barrier height (~1.1eV in this case for Pt on n-GaN). Defining $V_F$ as the voltage at which the forward current density is 100 A·cm$^{-2}$, we obtained values of ~1.8 V for the small-area rectifiers. This is roughly half of what has been reported previously for GaN Schottky rectifiers on heteroepitaxial layers. These are shown in Figure 6-23, along with previously reported results from GaN rectifiers. The solid lines are theoretical values assuming various Schottky barrier heights. The $V_F$ values in the present work are significantly lower than reported previously, indicating that surface cleaning and oxide removal steps have improved over time. The ideality factors were ~2 in the large area rectifiers, indicating that recombination was the dominant current transport mechanism. In the small-area rectifiers, $n$ values were ~1.5, which is consistent with dominant diffusion current.
Figure 6-20. Temperature dependence of forward current in 75 μm diameter vertical geometry implanted rectifiers.

Figure 6-21. Calculated breakdown voltage as a function of doping concentration and active layer thickness in GaN rectifiers.
Figure 6-22. Specific on-state resistance versus breakdown voltage for GaN rectifiers reported in the literature. The lines show theoretical results for Si, SiC, and GaN.
Figure 6-23. Forward turn-on voltage versus reverse breakdown voltage for GaN rectifiers. The curves show theoretical values expected for different barrier heights.

A thermal package was designed for the large area vertical rectifier shown in Figure 6-16. The diode was mounted on an FR-4 board with ½ oz copper on each side. The copper was overplated with 0.5 μm Ni and ~1 μm Au. The diode was adhered to the board with H2OE silver-loaded epoxy from EpoTek. The topside Schottky metal was connected to the pad with 1 × 5 mil gold ribbon, also mounted with EpoTek H2OE. A schematic of the package design is given in Figure 6-24. The packaged device is shown in Figure 6-25. Forward I-V characteristics of the packaged device were measured by applying a square wave voltage pulse (0 V - V_F) to the Schottky contact and monitoring the current using a wideband current probe connected to a 500 MHz Agilent Infinium
50662 oscilloscope. The reverse characteristics were taken from DC measurements using an HP4145B. Both the forward and reverse characteristics are shown together in Figure 6-26. As mentioned previously in this section, the reverse breakdown voltage of the large area device was small (~6V). However, pulsed forward current of 1.65 A was demonstrated at $V_F = 6$ V. This is the highest forward current ever obtained from a GaN rectifier. Despite the small breakdown voltage, clear rectification behavior is evident from Figure 6-26. For GaN-based rectifiers to become useful in the commercial power grid, they will not only be required to block large voltages, but also conduct significant forward currents. Previous small area GaN rectifiers have achieved impressive reverse characteristics, but forward characteristics have always been reported as current density. This device represents a large step toward the achievement of practical on-state current levels.
Figure 6-24. CAD design of thermal package for large area GaN rectifier. The package is necessary to avoid significant self-heating in the high current device.

Figure 6-25. Photograph of large area diode package. The approximate dimensions of the entire package are 1 cm × 3.5 cm.
Figure 6-26. Current-voltage characteristics of packaged GaN diode measured in pulsed voltage mode.

6.5 Switching Behavior

A key feature of wide band gap rectifiers is their expected ability to be switched from forward to reverse voltage without the large current overshoots and long switching times observed with Si diodes. Switching characteristics of the GaN rectifiers were measured with the reverse recovery apparatus shown schematically in Figure 6-27. A square wave voltage waveform from $+V_1$ to $-V_2$ was sent to the device under test from an Agilent 8110A pulse generator. During the positive bias portion of the waveform, the diode is forward biased, allowing current to flow. The square wave allows the polarity of the bias to be rapidly switched from forward to reverse, causing the device to turn “off”. The device output was monitored by a 500 MHz oscilloscope.
Figure 6-27. Experimental setup used to measure transient characteristics of GaN Schottky diodes [278].

The reverse recovery waveforms from a small-area vertical geometry GaN rectifier and from a standard “fast recovery” 1A, 600V, 200nsec Si diode are shown in Figure 6-28. The devices were switched from forward current densities of 400 A·cm\(^{-2}\) to a reverse bias of -25 V. The GaN data shows that nanosecond switching times are possible. Note also the significantly lower reverse current in the GaN device due to the smaller amount of stored charge. Following the analysis of Khemka et al [277], an estimate of ~15 nsec is obtained for the high injection level hole lifetime in these rectifiers. This is within the range of 1-20 nsec for previously reported minority carrier lifetimes in n-GaN [279].

In summary, there are some very promising features of the performance of these vertically depleting diodes fabricated on free-standing GaN templates, both with and without p-guard ring implantation. Excellent forward current densities were achieved compared to previously reported lateral rectifiers and very good reverse recovery characteristics were demonstrated. The temperature coefficient for \(V_B\) is still negative, indicating the need for further defect reduction in the material and continued optimization of epigrowth on these novel substrates.
Figure 6-28. GaN rectifier reverse recovery characteristics from 75 μm diameter device (top), compared to standard 1A, 600V Si diode (bottom). Note the absence of large current overshoot for the GaN device, indicating the potential for "cleaner" power switching.
CHAPTER 7
SUMMARY, CONCLUSIONS, AND FUTURE WORK

7.1. AlGaN / GaN HEMTs

Al$_{0.2}$Ga$_{0.8}$N/GaN HEMTs have been fabricated from MOCVD-grown epilayers on (0001) sapphire substrates. DC current densities from 0.5 to ~1A/mm were achieved for gate lengths ranging from 0.75 - 0.1 μm. A maximum transconductance of 207 mS/mm was measured for short gate length devices. From s-parameter measurements, $f_t$ of 59 GHz and $f_{max}$ of 90 GHz were extracted. Load pull results from $0.25 \times 150$ μm$^2$ gate dimension devices indicated an output power density of 2.75 W/mm at 3 GHz and 1.7 W/mm at 10 GHz. Small signal modeling of the measured s-parameters was also presented.

AlGaN/GaN HEMTs grown side-by-side on (0001) sapphire and 6H-SiC substrates were studied in terms of material quality and DC device performance. TEM micrographs clearly indicated that nitride epilayers grown on SiC contained a lower density of defects and dislocations due to the smaller lattice mismatch between epilayer and substrate. However, the DC performance of devices fabricated on sapphire substrates was superior in terms of peak drain current, transconductance, and 2DEG mobility. High temperature DC measurements confirmed the superior thermal characteristics of the SiC substrates, which demonstrated significantly lower self-heating effects. Longitudinal optical phonon scattering was the dominant electron scattering mechanism for devices fabricated on both types of substrates.
RF plasma-assisted MBE-grown HEMTs were fabricated and DC characteristics were reported. Operation at $>40 \ V_{DS}$ with $>50 \ mA$ peak $I_D$ was obtained (i.e. $>2 \ W$ of power in devices not optimized for power performance), suggesting competitiveness between MBE- and MOCVD-grown epilayers for GaN electronic device applications. Historically, MBE-grown material has proven inferior for LED and LD device fabrication. This does not seem to be the case for electronic devices. The surface morphology of the MBE-grown epilayers and uniformity of device performance were excellent. Reasonable current densities and excellent scaling properties were achieved over a range of gate lengths and gate widths.

Future efforts in this area should involve scaling the small periphery devices reported in this study to gate widths useful for microwave power amplification. Assuming adequate scalability of drain current, a 2 mm gate width device with performance similar to the 100 $\mu$m devices presented in Chapter 3 should produce $>2 \ A$ output current. Power levels of $>20 \ W$ seem feasible at X-band from a device of this size, if sufficient thermal management is provided. Additional improvements in material growth, device design, and processing techniques may extend the useful frequency range well beyond this, perhaps to Kα band or beyond.

All devices fabricated for this study employed Ni/Au Schottky gate metallization. Interesting phenomena were observed for many of the forward and reverse gate I-V characteristics. A slight decrease in the Schottky barrier height was reported for e-beam patterned gates, and this was attributed to surface damage during the gate writing process. Verification of this mechanism and development of lower damage e-beam recipes is warranted. In a potentially related phenomenon, the reverse gate leakage was shown to
decrease by nearly 3 orders of magnitude when measured at 400°C. This change may be associated with changes to the metallization itself or more probably with defect states on the surface or under the gate. Although unexplained, this result suggests the potential utility of a final low temperature “gate anneal” processing step to improve gate characteristics. It is not known if similar behavior would be observed with other Schottky metals, such as Pt or Pd.

Despite remarkable recent progress in AlGaN/GaN HEMTs, it has been frequently observed that the RF power output is much lower than expected from the DC characteristics. Although the mechanism for this effect is still unclear, it is very likely caused by electron capture and emission by “traps” in the device structure. These traps may be related to impurities, defects, dislocations, or polarization-induced bound charges. Understanding the origin and dynamics of these traps, and how to reduce or eliminate them through optimization of epilayer growth or device processing, is paramount to the realization of the full potential of these devices. Passivation of the surface sheet charge with Si₃N₄ has been shown to increase the output power of the device considerably. However, a clear mechanism for the improvement is yet to emerge. Although Si₃N₄ has shown promising results, other dielectrics such as SiO₂ have not been able to realize the same improvements in large signal performance. It is possible that the insulated gate approach with a low density of interface states will lead to devices with little DC-to-RF dispersion. This is an area of tremendous potential for additional contributions.

Initial reports of gas-sensing devices from III-nitrides have been given [280-282]. These is an interesting niche application for AlGaN/GaN HEMTs due to the chemical inertness of the material system. Sensors and detectors operating in harsh environments
and at elevated temperatures, such as in an exhaust stack, are envisioned. Using Pt as the Schottky gate may allow hydrogen detection due to the catalytic nature of the gate metal. In addition, the polarization-induced surface charge in unpassivated AlGaN/GaN heterostructures has may be an extremely sensitive monitor of various ambient conditions.

7.2. Gate Oxide Growth and Processing

Single crystal and amorphous Gd2O3 were grown by GSMBE and characterized by I-V and C-V. Initial results from single crystal ScO, a candidate dielectric with much smaller lattice mismatch to GaN, were presented. Ideal C-V theory and MOSC physics were each discussed to assist in data interpretation.

Capacitance-voltage characterization of wide bandgap MOSCs is complicated by the large energy bandgap of GaN. High measurement temperatures are necessary to decrease trap emission time constants to reasonable levels. Still, due to the exponential dependence of emission rates derived from Shockley-Read-Hall statistics, trap states near midgap are difficult to characterize at practical temperatures. High temperature measurements are not difficult in practice, assuming the stability of other aspects of the device structure. Due to the preliminary nature of this work, the temperature stability of MOSCs fabricated with these novel gate dielectrics devices has not been investigated. One of the primary benefits of the MOS structure is reduced gate leakage at elevated temperature (where the thermally-activated Schottky gate leakage becomes significant). Adequate temperature performance of the insulator is paramount to the utility of these devices for such applications. In addition to elevated temperature measurements, constant-current stress measurements are commonly used to assess the long-term stability
of the gate dielectrics. In this technique, a small current is forced through the oxide and the voltage drop is measured as a function of time.

It is standard practice in silicon MOS processing to perform a post-oxidation anneal to passivate or satisfy remaining dangling bonds that may contribute to the interface state density. Dramatic reductions in $D_{it}$ may be achieved by optimization of the annealing conditions (time, temperature, ambient). The effects of such anneals have also been described for the SiO$_2$/SiC system [283]. The achievement of large dielectric strength and a small interface state density in as-grown dielectrics may be only a portion of GaN MOS technology development. An equally important aspect may be determination of the appropriate conditions required to improve oxide and interfacial quality. Candidate gases for initial annealing experiments include forming gas, H$_2$, O$_2$, and air.

Various methods for extraction of interface state densities from capacitance-voltage measurements have been described (e.g. Terman, high-low, photoassisted, Gray-Brown). Each of these techniques has advantages and disadvantages, but none are considered as accurate and sensitive as AC conductance-voltage (G-V). G-V measurements are similar to C-V in data collection, but data interpretation is considerably more involved. The increased complexity of G-V is accompanied by the availability of both capture cross sections and trap time constants from proper data analysis. These parameters are only beginning to surface for the much more mature SiO$_2$/SiC interface. Very little, if any, of this data for GaN MOS devices has been reported. Such information is useful in the determination of the region of the forbidden energy gap from which interface state density can be determined at a given set of measurement conditions.
(i.e. temperature, frequency). Figure 7-1 is an illustration of this limit for an arbitrary MOS system measured at a DC sweep rate of 0.1 V/sec. The numerical values used for the calculation are given in the plot, and the capture cross sections were varied from $10^{-18}$ to $10^{-15}$ cm$^2$. The values plotted on the ordinate are limiting trap energy levels ($E_T$) below the conduction band ($E_c$) for an n-type GaN MOSC, calculated from Ref. [284].

![Graph](image)

Figure 7-1. Limiting interface trap energy level that can respond to capacitance-voltage experiment. An arbitrary interface state density, dielectric thickness, dielectric constant, and DC ramp rate were chosen. The capture cross section is used as a parameter, since these values are unknown for GaN.

As mentioned briefly in the previous section, passivation of AlGaN/GaN HEMTs with the novel, low $D_{it}$ oxides with may lead to improved passivation of surface effects responsible for large signal dispersion. Such passivation layers would be deposited after device fabrication, and appropriate non-damaging cleaning techniques need to be investigated to optimize the effectiveness of the passivation layer. These post-fabrication passivation steps will be limited by the maximum temperature tolerable by the Schottky contact. However, with the observed phenomenon of reduced Ni/Au gate leakage after
annealing, a 300° - 400° passivation deposition may simultaneously improve the gate characteristics and stabilize the surface charge of the HEMT.

7.3. MOSFETs and MOSHEMTs

Gate oxide growth and characterization for GaN MOS applications is still in its infancy. More than 20 years of research was necessary before the first GaAs MOSFETs were demonstrated using Ga$_2$O$_3$(Gd$_2$O$_3$) as the gate dielectric. The experience gained from GaAs-based devices was leveraged in the fabrication of GaN MOSFETs using Ga$_2$O$_3$(Gd$_2$O$_3$) [248] and binary Gd$_2$O$_3$ (this work). GaN based metal oxide semiconductor field effect transistors were demonstrated using a stacked gate oxide consisting of single crystal Gd$_2$O$_3$ and amorphous SiO$_2$. Gd$_2$O$_3$ provides a good oxide/semiconductor interface and SiO$_2$ reduces the gate leakage current and enhances oxide breakdown voltage. Charge modulation of the n-channel depletion mode MOSFET was achieved for gate voltage from +7 to -5 V. The source-drain breakdown voltage exceeded 80 V. An intrinsic transconductance of 61 mS/mm was obtained at a gate-source and drain-source bias of -0.5 V and 20 V, respectively. This is the first demonstration of epitaxial Gd$_2$O$_3$ growth on GaN and the first use of Gd$_2$O$_3$ as an insulating layer for nitride electronic device applications.

The MOSFET approach is attractive for reducing gate leakage current. However, the electron mobility of the n-channel MOSFET is limited by the doped channel. The bulk mobility in low dislocation density GaN is much less than that of AlGaN/GaN HEMT structures, such as the ones described in Chapter 3. The insulated gate approach can equally well apply to HEMTs, allowing fabrication of ‘MOS-HEMTs’. This would offer the advantages attributable to the heterostructure (i.e. increased mobility, large
electron sheet carrier concentration) and the insulated gate (i.e. reduced leakage, larger
gate voltage swing). Due to the buried channel in a MOS-HEMT, interface state density
may not be as critical as for the MOSFET, where interface states can communicate
directly with the channel electrons responsible for current flow.

GaN MOSFET and AlGaN/GaN HEMT fabrication differ by one key processing
step: the Ohmic window (oxide) etch. This etch is necessary to remove the oxide in the
Ohmic contact regions for subsequent metallization. Formation of intimate metal-
semiconductor contact is paramount to the minimization of contact resistance. The oxide
must be completely removed by this etch, or a thin interfacial layer will exist under the
Ohmic metal. This unetched oxide may act as a diffusion boundary for the various
diffusion steps necessary for formation of a low resistance alloyed Ohmic contact to the
nitride underlayers. The net result would be a device with high knee voltage and large
parasitic resistances, degrading the extrinsic properties of the FET. For Gd$_2$O$_3$, the oxide
etch was performed in HCl, and a final ‘clean-up’ was performed in an Ar plasma. The
wet etch was closely monitored, such that oxide undercut was minimized. When wet
etching the Ohmic window of short channel FETs, excessive undercut can entirely
remove the oxide from the channel region, producing a MESFET instead of the desired
MOSFET. In general, it was found that the etch rate of Gd$_2$O$_3$ in HCl was a strong
function of crystal quality and/or growth conditions. Hence, it was always necessary to
calibrate the etch rate prior to device processing using a ‘dummy’ sample. As oxide
development and optimization continues, investigations of the etching behavior of the
resultant films are essential to the successful processing of MOSFETs or MOS-HEMTs.
Particularly, development of dry etch chemistries is warranted, since any damage induced
in the source and drain contact regions may serve to increase the surface electron concentration, improving the Ohmic contact.

7.4. GaN Rectifiers on Bulk Substrates

Schottky rectifiers fabricated on free-standing GaN substrates in this work showed significant improvements in forward turn-on voltages (1.8 V), on-state resistance (1.7 mΩ·cm²), and reverse recovery characteristics relative to previously reported devices fabricated on GaN/sapphire. Reverse breakdown voltage decreased and reverse current increased with increasing temperature. The ideality factor of the Pt/Au Schottky contact extracted from the forward J-V characteristics was ~1.25 at 100°C. From reverse recovery measurements, the high injection level hole lifetime in these devices was estimated to be ~15 nsec. A large area (~20 mm²) rectifier was mounted to a heat sink package with conductive epoxy and wire bonded with 1 × 5 mil gold ribbon. This device, measured with 1 kHz voltage pulses, produced 1.65 Amperes of pulsed forward current at +6 V. This is the largest on-state current ever reported for a GaN rectifier.

The blocking voltage of the small-area rectifiers was not competitive with previous heteroepitaxial devices fabricated on insulating epilayers. The large area device also suffered from large reverse leakage currents. The low reverse voltages led to low figures-of-merit (= V_B² / R_ON) compared to devices able to withstand several hundreds or thousands of volts. Future efforts should focus on lowering the background doping level during HVPE growth of the GaN templates. Also, measurements of the temperature dependence of V_F, R_ON, and V_B for rectifiers fabricated on this optimized material are warranted. Previously reported heteroepitaxial rectifiers have shown a negative temperature coefficient for V_B, but this is expected to reverse sign in low defect material,
as was observed for SiC. The viability of GaN rectifiers in most applications depends on making very large area devices with high $V_B$, while retaining low $V_F$ and $R_{ON}$. One way to approach this may involve growth of p-i-n rectifiers on the bulk GaN. Previous results have shown the reverse breakdown voltage to be substantially larger for p-i-n diodes than for Schottky diodes fabricated on the same wafer [202]. Combining this advantage with the higher current levels available with a vertical geometry may lead to superior forward and reverse characteristics. For such devices to be realizable, optimization of epigrowth on the novel GaN substrates will be necessary.
REFERENCES


111. W. Schottky, Naturwissenschaften, 26 843 (1938).


BIOGRAPHICAL SKETCH

Jerry Wayne Johnson was born February 13, 1975 in Durham, NC. He is the son of Jerry Talmage Johnson and Shirley Eubanks Johnson of Chapel Hill, NC, grandson of Mr. and Mrs. John Talmage Johnson of Asbury, NC and Mr. and Mrs. Roland Percy Eubanks of Bynum, NC. After graduation from Northwood High School in Pittsboro, NC in June 1993, he enrolled at the Georgia Institute of Technology in Atlanta, GA, where he received the Bachelor of Chemical Engineering degree in June 1997 with a Certificate in Economics.

In August 1997, he began graduate studies in the Department of Chemical Engineering at the University of Florida. At UF, his early research involved various aspects of compound semiconductor growth and materials characterization. Later, his interests shifted to wide bandgap electronic devices. He is co-author of approximately 40 journal and conference papers dealing with compound semiconductor device technology.

During the summer of 1999, Wayne worked as an Engineering Intern at Deposition Sciences, Inc. in Santa Rosa, CA. At DSI, he was responsible for process optimization of LP-CVD deposited SiO₂, TiO₂, and Ta₂O₅ thin film optical interference filters. He spent January - May 2001 at Sandia National Laboratories in Albuquerque, NM in the RF Microsystems Technologies division. At Sandia, he conducted research on AlGaN/GaN HEMTs in collaboration with the University of Florida.

On October 10, 2001, Wayne and wife Traci welcomed the birth of their first son, Maxwell Scott Johnson.
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Fan Ren, Chairman
Professor of Chemical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Timothy J. Anderson, Co-Chair
Professor of Chemical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Stephen J. Pearson
Professor of Materials Science and Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Cammy R. Abornathy
Professor of Materials Science and Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2001

Pramod P. Khargonekar
Dean, College of Engineering

Winfred M. Phillips
Dean, Graduate School