Investigation of the Impact of Impurities on the Properties of Nitride Semiconductors Grown by RPECVD

Shawn Skerget

In partial completion of the requirements for
M.Sc. in Electrical and Computer Engineering

Electrical and Computer Engineering
Faculty of Engineering Lakehead
University, Ontario, Canada
September 2014

Table of Contents

Abstract

Progress toward the improvement of optical emission from InGaN optical active region devices is made through a combination of the tailoring of p-type gallium nitride growth recipes and bandstructure calculations on a graded InN-GaN bandstructure that provides insight into the causes of anomalous electroluminescence observations.

Using a 7nm thick amorphous aluminum nitride buffer layer, the quality of certain thin film, ~150nm thick GaN growths performed at 550C at 1360mTorr by nitrogen plasma RPECVD is increased fourfold. Growths of buffered GaN show an XRD intensity increase by four times with smaller FWHM, half the droplet roughness on the surface as identical recipe unbuffered GaN growths, and are believed to show superior electrical properties. An increase to lateral growth, crystallite nucleation and nitrided metal coalescence from the buffer layer are believed to account for the improvements.

The same buffer layers also improve the quality of p-type GaN fourfold, when grown by similar recipes to the GaN. A low carrier concentration is identified by Hall effect, which seems to indicate that the solubility limit of magnesium in gallium droplets is being reached; since at this limit and room temperature the hole concentration after this doping should be of the magnitude of the background electron concentration of GaN grown by the system. AFM smoothness figures also show improvement, with optical transmission characteristics indicating an empty conduction band edge.

Finally, calculation is performed using the linear combination of atomic orbitals method (LCAO) of the electron band structure of intermittent stoichiometry values between various types of primitive cells occurring in InGaN growths contaminated with background oxygen. The intermittent stoichiometry bandstructure between InN and GaN is important because it represents the bandstructure of the gradient of diffusion of indium into GaN at phase separation boundaries

between isolated phases of InN and GaN in InGaN growths. Calculations of oxygen on intermittent InN stoichiometries identifies it as a viable culprit for the yellow electroluminescence defect seen in InGaN light emitting devices. Calculations of the real part of dielectric susceptibility are performed on InGaN contaminated with oxygen; as well as in indium rich conditions of InGaN with indium substituted for nitrogen in the cell.

Declaration and Acknowledgements

I would like to thank Dr. Dimiter Alexandrov for his support of the research performed for this thesis. His model for bandstructure calculations laid the theoretical foundation that served as the underpinning for calculation of the bandstructures done in this work. Also, his technical help in the lab for growth and characterization of grown films was invaluable.

I must also thank all personnel who worked with me in the Lakehead University semiconductor laboratory. They selflessly answered my many questions and showed me what I needed to know in order to operate the lab equipment, identify my thesis topic, and fit in seamlessly to the rhythm of the research being done. My thanks goes out also to all personnel who have helped me to operate the equipment located outside of the lab.

Finally, my thanks goes out to my family, who have supported my education all throughout, encouraged the continuing of my education, and helped smooth over the ripples of life in an utterly unmentionable number of ways.

1 Chapter 1: Introduction

1.1 Purpose of this Work

The central purpose of this work was to increase the optical emission properties of nitride thin films grown using the Lakehead RPECVD growth system. Two major roadblocks exist along this path. The first is toward the development of p-type gallium nitride. While such a material has been achieved using different thermodynamic growth conditions, it has yet to be reported using a low temperature and low pressure RPECVD process like the one used in this work. Obtaining such a doped material is essential for the injection of holes into an optically active region; so that light generation by electron-hole recombination can be improved. The second roadblock is understanding the origins of a yellow-light defect that plagues devices using the alloyed compound semiconductor InGaN as the optically active region. A theoretical model based on the LCAO method of electron orbital calculation is used in the context of a graded InN-GaN stoichiometry to calculate the intermittent bandstructures of such graded transitions. From these calculations, we may understand the defect type and atomic species responsible for the optical emission of InGaN devices in the yellow region of the spectrum only and not others.

The purpose of the lab research performed is to improve the quality of p-type gallium nitride growable by low temperature MOCVD processes. One reason for the demand for high quality pGaN grown at low temperature is for incorporation of a hole injecting layer into device structures utilizing InN and nitride semiconductor alloys containing Indium – particularly InGaN.

InGaN exhibits phase separation above about 550C into separate regions of GaN and InN. Hence, to preserve the device structure we must find a material that acts as good source of holes whilst providing a compatible work function, bandgap, and radiative and transport properties seen

in InGaN. P-type conductivity in InN has not yet been realized, and hence p-type GaN is the natural candidate for research and development.

Devices utilizing InGaN can be used both as light emitting diodes (LED's) and as solar cell photovoltaics (PV). Pure GaN has a bandgap energy that corresponds to photon emission in the ultraviolet (UV) region and hence emits light as wavelengths invisible to the human eye. By alloying GaN with indium to create InGaN, the bandgap of the material can be tuned down to wavelengths visible by the human eye or within the emission spectrum of the sun by changing the mole fraction of indium to gallium.

In order for photon emission or absorption to continue to occur at a consistent rate, a consistent rate of electron and hole injection is required. Since crystal defects decrease the concentration of electrons and holes in n-type and p-type materials (respectively) and increase the resistance, it is desirable to grow these charge-source materials to a high crystal quality so that the diffusion forces driving injection/collection of charge carriers for LED's/solar cells can be maximized and the resistivity of the device minimized. Crystal quality also has bearing on energy lost by charge carriers as vibrational (thermal) energy. Hence, to achieve high performing LED's and solar cells, high quality doped nitrides are required.

The transport properties – namely charge mobility - are also degraded by crystal defects, and so for photovoltaics it is important to have high quality charge collectors in order for the electron-hole pairs generated by incoming solar radiation to migrate to these collector quickly before recombining (in which case the energy is lost again). These created charge pairs are separated by forces induced from potential gradients provided by the doped charge collector regions. The strength of these gradients are proportional to the charge carrier concentration of the charge collectors – which is in turn proportional to the crystal quality. Hence, a higher crystal quality will increase the efficiency of both LED's and PV's.

1.2 Overview of Nitrides

1.2.1 Background for Gallium Nitride

Wurtzite Gallium nitride's advantages over traditional silicon as a semiconductor are manifold. Firstly, gallium nitride has a large band gap and is very resistant to radiation – far more so than gallium arsenide which is also used for similar applications. It also has high temperature tolerance and conductivity, and can withstand electric fields over ten times greater than silicon (greater breakdown field). This means that both the voltage and current handling capability of GaN will be far greater than silicon before crystal damage occurs.

GaN has a breakdown field of about $5X10^6$ V/cm on account of its large bandgap; allowing greater reverse blocking capabilities. Reverse breakdown can occur either as the Zener effect or the avalanche effect. In the zener effect the device electric field becomes great enough to allow electrons to tunnel from the valence band to the conduction band. In the avalanche effect, charges are accelerated through the field to velocities where the kinetic energy is sufficient to free an atomically bound electron on collision. This leads to a cascading effect through the rest of the field. The high bandgap of GaN makes tunneling for the zener effect more difficult; thus the avalanche effect tends to be the dominant mode of breakdown. GaN avalanche photodiodes are useful for detecting faint ultraviolet or blue-region radiation without the noise from the rest of the visible spectrum.

The band gap of gallium nitride, as measured in monocrystalline films grown on buffer layers by molecular beam epitaxy is 3.44eV. Measured values tend to be lower than this on account of electron energy states created by crystal lattice defects. For typical MOCVD grown GaN it may measure as low as 3eV. This is the amount of energy to move a valence electron in an atom's highest energy occupied electron state to the lowest energy state where the electron

is no longer bound to the atom and free to move around the crystal and contribute to conduction current. Bound “valence” electrons can also move between atoms, and in the presence of an electric field will do so in the direction of the field. Since electrons in the valence band have slower conduction velocities than electrons in the conduction band, and since conduction band electrons can only bind to an atom in an unoccupied valence band site (due to the Pauli exclusion principle), the movement of these vacancies in the valence band can be thought of as particles with mass higher than conduction band electrons that move like positively charged particles in an electric field. Called holes, the recombination of an electron with a hole releases energy in the form of a photon equal to the energy to liberate an electron from a crystal atom into the conduction band – the bandgap. The bandgap energy can be measured from the energy of this photon; which can be obtained from the wavelength via the Planck relation.

The electron transport properties of GaN are conducive to research in the high frequency, high power spectrum of application. Electron mobility is a parameter that directly relates to the conductivity of a material. Knowing the density of electrons eligible to move within a material, mobility and hence conductivity can be determined from electron velocity in a weak electric field. The electron mobility of GaN is about 1500$cm^2/V*s$, which is low compared to GaAs. However, GaN boasts an electron saturation velocity of 2.5 times that of GaAs. This is the velocity of charges under a high electric field where atomic collisions serve as a limiting parasitic for increasing electron momentum.

1.2.2 Contacts

Good contacts to gallium nitride with can be made utilizing gold, aluminum, or titanium/aluminum; but low resistance contacts can only be made to n-type GaN. Al/Ti bilayers yield the lowest resistance contact to undoped and n-type gallium nitride, while gold/nickel

contacts and others with a high metal work function [1] tend to be better to p-type gallium nitride and typically form low Schottky barriers. These gold/nickel contacts are utilized for contacts to pGaN for the Hall measurements done in this work.

The work function is the average energy required to bring an electron from deep free space into the metal for conduction, and hence represents the effective energy of the conduction electrons within. The work function is especially important because the GaN crystal has a high ionic character - unlike silicon which is covalent in character [2]. Also, in this work the films always have the negatively charged gallium metals on the exposed face of the film. P doping of the crystal will change the nature of accumulation of charge at the surface where the metal is places. Thus, extra care must be taken in the engineering of metal contacts to minimize contact potential.

High quality contacts to nGaN can be deposited from a boule of tin as metal vapor via electron-beam evaporation, with subsequent annealing at 500-900C in N_2 ambient for 30 seconds using rapid thermal annealing [3]. Only 500-550C is viable for the purpose of indium preservation. Annealing at lower temperatures of 60C has showed barrier height increase from 200meV to 250meV, while 230C has yielded upwards of 450meV [4]. These effects are explained by chemical interaction with the substrate through the Schottky-Mott model of contacts.

The indium-tin contact is useful to n-type GaN and InGaN as well since annealing of the contact creates a graded $In_xGa_{(1-x)}N$ structure at the interface through a combination of solid state diffusion and reaction. This creates a quasi-ohmic contact with low resistance at the metal-InGaN interface as electrons are able to traverse the graded junction under a light electric field with significant contribution from thermal energies. This method is preferred in this work because it works well even when the contacts are simply soldered on – allowing quick and easy contact placement.

P contacts need to be fairly transparent in order to be viable on p-type epilayers for LEDs; and hence must be made thin. Contacts less than 0.03um have been used for over 50% transmission of light in the upper part of the visible spectrum [2]. N contacts are easily achieved with Schottky or ohmic electrical characteristics and with good mechanical robustness (low interdiffusion, corrosion resistance, thermal resistance, low film stress ect.); however P contacts have proved particularly difficult to achieve to meet the slew of electrical and mechanical requirements of commercially viable systems. The best p-type contacts reported were an alloy of Pd/Ag/Au/Ti/Au on the order of $1X10^{-6}\Omega cm$ [5], but these are unusable with growths containing indium since they require annealing at 800C under nitrogen. They are also rather complicated to prepare. It was reported that the anneal was supposed to create a p-type alloy between the pGaN and the silver and gold atoms under solid state diffusion.

For purposes of indium preservation the best contact scheme proposed is a multi-layered metallization scheme of platinum, nickel, and gold in different thickness proportions; proposed by Ja-Soon et al [6]. This process only requires an anneal of 350C for 1 minute in N_2 atmosphere and yields an ohmic contact with $5.1X10^{-4}\Omega cm$ under dopant concentrations on the order of $10^{17}cm^{-3}$.

1.2.3 Compound Semiconductors vs. Silicon

Perhaps the greatest use for nitrides is due to the fact that they are efficient radiation emitters and absorbers. This is on account of their direct band gaps. As discussed earlier, they also have a slew of features that make them physically tougher that silicon in many applications, and boast higher threshold electrical properties so that they are useful for high power and frequency applications.

Silicon has the main advantage that it can be grown quite easily and cheaply as a single crystal with minimal purities and defects using the Czochralski pulling process. Since silicon

crystal is monatomic, stoichiometry is not an issue. The low defect and impurity density allows very small features at high density to be implanted/printed on the substrate while maintaining a high yield for these devices and keeping them commercially viable. Full processors have been printed with gallium arsenide (a compound semiconductor) using specially designed CMOS logic gates that account for the discrepancy between electron and hole mobilities; however these never came to fruition for the mainstream because of materials costs and poor scaling capabilities due to limitations by defects and thermal conductivity. These devices are supposed to have a considerably higher switching speed and lower power loss than silicon based devices.

Today, III-V compound semiconductors continue to be of interest for their superior optical properties and switching speeds. Nitride devices dominate the upper range of the visible spectrum for LEDs and lasers, while the arsenides and phosphides and their alloys fill out the lower to middle of the visible spectrum, respectively.

Material	Bandgap (eV)	Electron Mobility (cm2/Vs)	Hole Mobility (cm2/Vs)	Critical Field *Ec* (V/cm)	Thermal Conductivity σT (W/m•K)	Coefficient of Thermal Expansion (ppm/K)
InSb	0.17, D	77,000	850	1,000	18	5.37
InAs	0.354, D	44,000	500	40,000	27	4.52
GaSb	0.726, D	3,000	1,000	50,000	32	7.75
InP	1.344, D	5,400	200	500,000	68	4.6
GaAs	1.424, D	8500	400	400,000	55	5.73
GaN	3.44, D	900	10	3,000,000	110 (200 Film)	5.4-7.2
Ge	0.661, I	3,900	1,900	100,000	58	5.9
Si	1.12, I	1,400	450	300,000	130	2.6
GaP	2.26, I	250	150	1,000,000	110	4.65
SiC (3C, b)	2.36, I	300-900	10-30	1,300,000	700	2.77
SiC (6H, a)	2.86, I	330 - 400	75	2,400,000	700	5.12
SiC (4H, a)	3.25, I	700		3,180,000	700	5.12
C (diamond)	5.46-5.6, I	2,200	1,800	6,000,000	1,300	0.8

Figure 1 A table of basic electrical and physical characteristics of many commonly used semiconductors and SiC substrates. For a table featuring more electronic information, see the appendix.

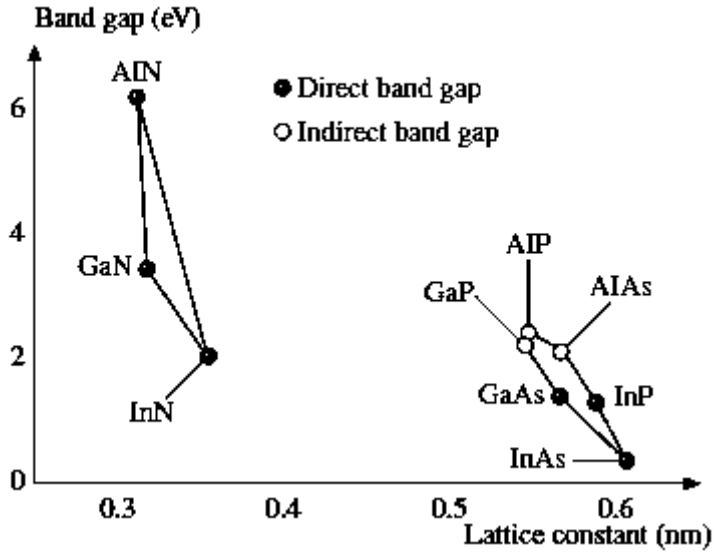

Figure 2 Bandgap vs. Lattice constant for many common binary compound semiconductors from the nitride, arsenide, and phosphide families.

1.2.4 Substrates for Gallium Nitride

Gallium Nitride is usually deposited on one of three substrates: sapphire, silicon <111> or silicon carbide (usually 4H or 6H). Sapphire is the most common substrate for deposition of devices because it provides the best balance between crystal quality and price. Typical applications of sapphire substrates with nitrides are for UV and blue/violet LEDS, lasers, and high frequency transistors. Widely available are commercial substrates with doped/undoped GaN or AlN buffer layers.

Sapphire is used mainly in high power and high frequency – typically RF – applications; but can also be viably used for LED's, being far less expensive than germanium substrates. It can also be used with a silicon substrate on top of one part so that dense logic can be fabricated on the silicon with high power handling devices on the sapphire for a monolithic solution to many communications demands. Porcelanium is a separating substrate that when used with a sapphire base has the ability to allow many-layered structures of different epitaxy families on a single base substrate. The sapphire serves as an insulating base with good thermal conductivity for high power handling and dissipation.

Sapphire has a fairly large lattice mismatch with GaN. In p-type gallium nitride doped with magnesium, the magnesium atoms tend to exert a compressive force to the GaN film which may serve to mitigate mismatch strain by matching the film more closely to the sapphire substrate. In Si <111> substrates where the inherent mismatch creates tensile forces on GaN, magnesium dopants compound that force. Hence, p-type doping of GaN grown on silicon has been very problematic and usually a GaN buffer layer is required. It is worth noting that sapphire also has the advantage of being insulating – negating the need for the passivation techniques seen with silicon. Sapphire also has the advantage of being viable for use with optical transmission measurements or electroluminescent measurements utilizing light collection from the bottom of the substrate.

Lattice mismatch for GaN on <111> silicon is around 17% tensile on GaN, on sapphire is around 16.1% compressive on GaN, and on SiC is around 3.5%.

Silicon oriented <111> is gaining momentum as a viable substrate for nitride LED development, and may yet replace sapphire as the preferred commercial substrate for these structures if improvements in crystal quality continue the current trend. They are cheaper than sapphire, but roadblocks in nitrides film quality keep silicon from overtaking sapphire. The silicon may be doped n-type for passivation, but an AlN buffer layer of around 50nm is almost always deposited on silicon before nitrides deposition. This buffer layer serves to block the diffusion of silicon into the nitride films, where it would otherwise act as an n-type dopant and contribute to a large defect density from relaxed tensile stresses [7]. Narrowing electroluminescence spectra and energy-increasing Stokes shifts have been observed for moderately silicon-doped GaN; with nonradiative recombination dominating the band-edge recombination in lightly Si-doped GaN [8]. Since diffusion of silicon into nitrides tends to be light to moderate (~10^{16} cm^{-3} – 10^{17} cm^{-3}), this will decrease the efficiency of LEDs grown without a diffusion blocking layer considerably. An overall linear coefficient characterizing the relationship between biaxial stress of GaN films (like from silicon dopants) and shift of the luminescent bandgap is given by [7] as 21.1±3.2meV/GPa.

When GaN grown on an intrinsic GaN buffer layer is doped with silicon, it becomes a viable n-type donor with excellent surface morphologies up to doping concentrations of ~2X10^{19} cm^{-3} grown by ammonia-based MOCVD [9]; suggesting a better incorporation of silicon into lattice gallium vacancies driven by the force of the underlying GaN buffer layer.

Silicon carbide substrates have the advantage of providing a very near match to gallium nitride's lattice constant; whilst still providing the temperature and radiation hardness that sapphire provides. Unfortunately, the price of 6H SiC is likely to restrain its use to industrial applications where environmental factors need it, as well as for photonic, thin film, and sensing/instrumentation

devices where film quality is paramount and/or the dielectric properties of sapphire create unacceptable phase interference.

Substrates are prepared for growth by heating in an air-ambient oven at 1050C for 6 hours. This is to remove the oxygen contamination in sapphire and SiC substrates. Treatment with nitrogen plasma generated at 600W for 10 minutes can be used to improve substrate surface conditions in a similar manner as hydrogen plasma treatment improves RHEED patterns in MBE [10]. This is done in the growth chamber before a nitridation of the surface at 300W deposits lower energy nitrogen to smooth the surface in preparation for metal pulsing.

1.2.5 Other Growth Techniques for Nitrides

MBE is the most common growth technique alternative to MOCVD. MBE growth operates at magnitudes of pressure lower than low pressure MOCVD (usually $<10^{-3}$ torr), as a result the metal species do not need to be delivered to the substrate in the form of inert precursors and dust formation from precursor interaction in the gas phase is not a problem. Instead of precursors, metals are delivered in elemental form by effusion cells that operate by sending a single species of metal to the substrate surface through a controllable shutter. Since the dissociation of precursors isn't a factor, stoichiometry can be controlled very accurately and very high quality films with low contaminant concentrations can be grown. Metal modulated epitaxy using pulsed effusion cell shuttering with nitrogen plasma is capable of growing almost single crystalline GaN. InN suffers from problems of desorption at the low pressures, but has also been grown to a high quality nonetheless. MBE systems have developed good growth rates, but still suffer from problems with the scalability of the process. Manufacturing yield needs to be improved if MBE is to compete with MOCVD or HVPE in the mass-consumer LED market.

Hydride vapour phase epitaxy (HVPE) utilizes metal-halides to deliver metal species from the liquid form to the substrate in a gaseous form. These then react with ammonia to create the crystal on the substrate. The process boasts a high growth rate, and has popular in industry for this reason; as well as for the scalability of this process. Hydrogen halides are the usual byproduct of this process, and must be disposed of properly.

Liquid Phase epitaxy has been used to grow very high quality and purity arsenides, which are the dominant LED in the red-infrared part of the spectrum. The process is actually similar to MOCVD using pulsed growth regimes (metal modulated epitaxy), in that metal in the liquid form is deposited before subsequent treatment by a group V element in the gas phase. Phosphides, which account for much of the middle of the visible spectrum of emission, are being grown by MOCVD.

1.3 Growth System Used in this Work – Low Temperature RPECVD

1.3.1 Overview

The Lakehead University Semiconductor Lab is home to a unique growth system used for growth of III-V nitride semiconductors. It is a plasma-based migration-enhanced metal-organic chemical vapor deposition system developed for compound semiconductor growth at low temperatures. Lower growth temperatures give lower operating costs (6kW-11kW) and less maintenance. It Uses nitrogen carrier gas and compressed metal-organics only – there are no toxic substances like ammonia used and emissions are safe without post-processing.

It utilizes fully computerized growth regimes for precise control of the growth conditions at the monolayer resolution; however films are usually grown at least a few monolayers at a time for the sake of faster output. For pure films, faster than monolayer growth is usually without significant degradation of the quality; however for alloy nitrides (like InGaN or AlGaN) a monolayer growth

regime is generally preferred in order to reduce defect density from phase separation caused by inhomogenous species delivery across the film surface.

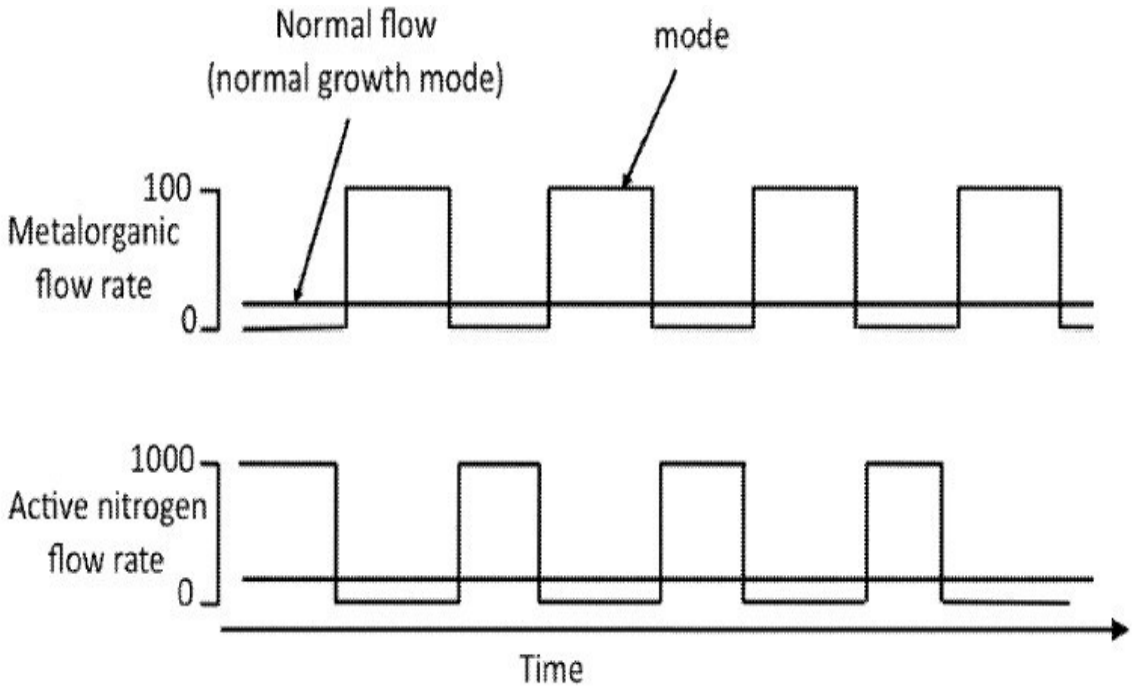

Figure 3 Sccm flows of metalorganic and active nitrogen plasma species in time for a fully pulsed growth regime. Also available are pulsing regimes utilizing constant flow of nitrogen species (common) or of both (less common)

Growth of nitrides is usually performed at temperatures below 600C instead of the 1000C range that is usually used for nitride growth by MOVPE. This is to make the system compatible with high indium content growths in order to realize the prospects of indium compound semiconductors. It also has the advantage of allowing temperature sensitive substrates like glass to be explored for growth. These unusually low growth temperatures mean that ammonia cannot be thermally decomposed in the chamber and so a nitrogen plasma source is used to provide active nitrogen species. This also means that growth pressure must be lower than usual in order to maintain the plasma species' kinetic energy for when it reaches the film surface. While GaN is

usually grown in the range of hundreds of torr for standard MOCVD growth under ammonia, the Lakehead growth system works at around 1-3 torr.

The Lakehead growth system is equipped with a separately pumped, gated load lock and a separately pumped residual gas analyzer (RGA) chamber for leak detection. A tapered valve allows sampling of the growth gases to the RGA for analysis of gas content during growth. Overall the system can be used for growth of III-Nitrides with use of nitrogen plasma, III-Oxides with use of oxygen plasma, and for modification of surfaces with Hydrogen plasmas like required for Graphene synthesis and patterning.

1.3.1.1 Substrate Bowing and Internal Stresses

The lower growth temperatures of the Lakehead MOCVD system mean less bowing of thin films from thermal expansion mismatch during cooling. Since the plasma source is also very scalable while maintaining uniformity of species delivery across the substrate, the growth system can be scaled up to large substrates (>12”) without many of the scaling problems that occur in standard MOCVD systems. The cost of operation is also reduced – operating at 6-11kW. Hence, the Lakehead growth system is potentially a very economic and environmentally friendly alternative to other methods of nitride growth.

Sample bowing can be tracked layer-by-layer in a build-up measurement process by utilizing a spatially sensitive receiving sensor and laser that scans the sample in a line to create a line-profile of bending after each layer is deposited. Utilizing the “effective” substrate (the lumped model of everything beneath the current layer), the Stoney equation can be used to deduce spatial film stress data from the line-profile. The true values usually depend strongly on the film’s vertical temperature gradients during growth, composition and composition-dependent lattice constants, and thermal expansion coefficient of the layers [11]. These all need to be separated. Hence, much information is required to discern precisely the origin of substrate bowing

and resulting crystal stresses. The Lakehead MOCVD system reduces defect-causing stresses in the sample from bowing by growing at a lower temperature (with nitrogen plasma) than needed for typical systems that generally use ammonia for active nitrogen delivery.

Compound nitride semiconductors grown with the Lakehead University MOCVD system include InN, GaN, amorphous AlN, and tertiary semiconductors thereof like InGaN and AlGaN. InN is a candidate material showing promise for use in future devices – it has extremely high electron saturation velocity and mobility, with applications in bandgap tunability when alloying GaN. High speed (RF HEMT transistors, diodes), high efficiency (multi-quantum-well (MQW) heterostructure and single-junction solar cells), and radiation and temperature resistant devices operating at high frequencies are all candidate applications.

1.3.2 Pulsed Growth Regimes for Improved Film Quality and Defect Management

For unalloyed films (GaN, InN, AlN) the growth regime follows a simple alternating pulsing of metalorganic and plasma. The plasma flow may be stopped during the metalorganic pulses (see Figure 3), or may be held continuous throughout the entire cycle with only metalorganic being pulsed in. Pulsing of the metalorganic is done to accommodate the migration of dissociated metal species across the film surface to energetically favorable conditions before binding with active nitrogen plasma species to form the film. This is done in order to help counteract the lack of on-the-fly growth-time “annealing” provided by higher temperatures.

The concept of pulsing to accommodate species migration is a potentially powerful idea for the improvement of films grown at low temperatures. While the donor oxygen tends to incorporate into metal lattice positions without much induced strain, P-type conductivity in nitrides (notably gallium nitride and indium nitride) is rather difficult to achieve even in higher temperature (~1000C) growth conditions due to defects from generated mechanical stresses. These stresses

tend to be compressive in nature for magnesium and their methods of relief by crystal dislocation depends heavily upon the growth conditions and underlying film thickness.

While edge dislocations generally dominate the transport characteristics in MOCVD grown gallium nitride films, it has been shown by transmission electron microscope (TEM) images of defect-selective etching by HCl in the vapor-phase that screw and mixed screw dislocation density directly detracts from photoluminescence (PL) intensity [12]. Edge dislocations did not have an appreciable effect on PL for a given current. It can therefore be inferred that threading dislocations of the screw-type act as nonradiative recombination centers while edge dislocations act primarily as radiative recombination centers. Hence, dislocation density increase by mainly edge-type defects from the lowering of growth temperature and pressure would explain the large drop in conductivity/mobility with only small loss of electroluminescence of Lakehead's CVD-grown gallium nitride films as compared to structures grown by MOCVD at more typical conditions of higher temperature and pressure. Typical screw dislocation density for a high quality GaN film grown on an AlN buffered SiC substrate (best lattice match) is ~10^9 cm^{-3}, while for edge dislocations ~$1X10^8$ cm^{-3} [13].

Threading dislocation density as a function of film thickness in GaN grown by MOCVD on a log-log plot has been shown to have a more or less constant negative slope for film thicknesses over 1um [13]. If the marginal absorption coefficient resulting from product of film-absoptivity and (marginal) epilayer thickness is properly factored out, a clear correlation between marginal TD density and marginal loss in photoluminescence intensity at the band edge as a function of film thickness will be observed.

The prevention of the formation of stress-related defects via film migration-enhancement may be possible at low temperatures. This is due to the fact that the added time can allow for reduction of residual entropy in the film caused by energetically unfavorable dopant inclusion.

Interstitial inclusions of magnesium introduce harmful compressive stresses in the film that can be mitigated by altering film growth conditions to allow these atoms to assume the position of nitrogen vacancies instead. It is potentially this vacancy-filling mechanism that accounts for the somewhat better conductivity and film quality observed in nitrogen-lean growth conditions of pGaN. Likewise, the complex that magnesium forms with hydrogen impurity atoms from metalorganic precursors can possibly be mitigated in films by allowing time between metalorganic/dopant delivery and plasma delivery for the hydrogen to be removed by the chamber by the vacuum source. Magnesium tends to dissolve into the liquid gallium droplets that form on the film surface more strongly than hydrogen, and hence the added time between metal pulsing and plasma delivery can ensure better distribution of magnesium throughout the droplet and the better removal of impurities like hydrogen from the droplets by diffusion before the plasma is introduced to form the film.

1.3.3 Metalorganics and species delivery

The metal is delivered to the chamber for deposition in the form of metalorganic precursor molecules. This is done so that the metal species do not react with contaminants like oxygen before reaching the substrate surface, and the precursors are chosen so that they easily dissociate under the conditions there. These molecules are, for the most part, a simple complex of methyl (methane) ligands bonded covalently to the central metal atom. They exist primarily in liquid form under pressure in canisters residing in a separate cabinet from the chamber equipped with a safety vent. The configuration of the vent canisters is shown in Figure 4. They operate on pneumatic valves and have manual shutoff valves just after the canister and to the purge line at the bottom of the diagram to accommodate flushing of the lines for changing the canisters. The canisters for gallium and aluminum precursors don't require heat or backing gas pressure for delivery since growth pressures are low enough for them to flow on their own. All precursors flow

into a mixing manifold where they are mixed with 99.99999% pure N_2 that has been ran through an in-line oxygen scrubber to increase purity. The mix flows consistently during the growth and is pulsed between two valves seen to the left of the chamber in Figure 4. One valve allows entry to the chamber and the other is a purge valve that shunts the flow to the turbomolecular pump for evacuation.

The N_2 flowing into the mixing manifold provides the background chamber pressure for growth, while the plasma source can accept either the (essentially) pure N_2 or oxygen contaminated N_2 from a separate bottle if n-type doping is desired. All flows are controlled through dedicated mass flow controllers.

The magnesium doesn't exist as a simple methyl ligand complex like the gallium and aluminum, but rather as bis-cyclopentadienyl magnesium. The vapor pressure of this species is low and when exposed to chamber growth pressures and requires a backing gas of N_2 in order to achieve proper flow. Figure 4 shows the bypass lines (both pneumatic and manual) for magnesium – these are used to maintain continuous and proper flow of the backing gas during the plasma-only part of the growth pulsing regime and also to purge the line of lingering metalorganics before changing the canister. Magnesium and Indium precursors both require a temperature higher than room temperature for flow; they are both fed electrical heating and wrapped in aluminum foil to seal the temperature in. In the case of indium, a specialized electrical heating configuration and mass flow controller are required to keep the delivery lines before the flow controller at a proper temperature to prevent condensation of indium metal to the pipe walls. The Indium precursor species readily dissociates and burns in atmospheric conditions, and is also known to be toxic to humans via respiratory delivery; so extra care must be taken when handling the setup. The Indium metalorganics don't require backing gas pressure for flow.

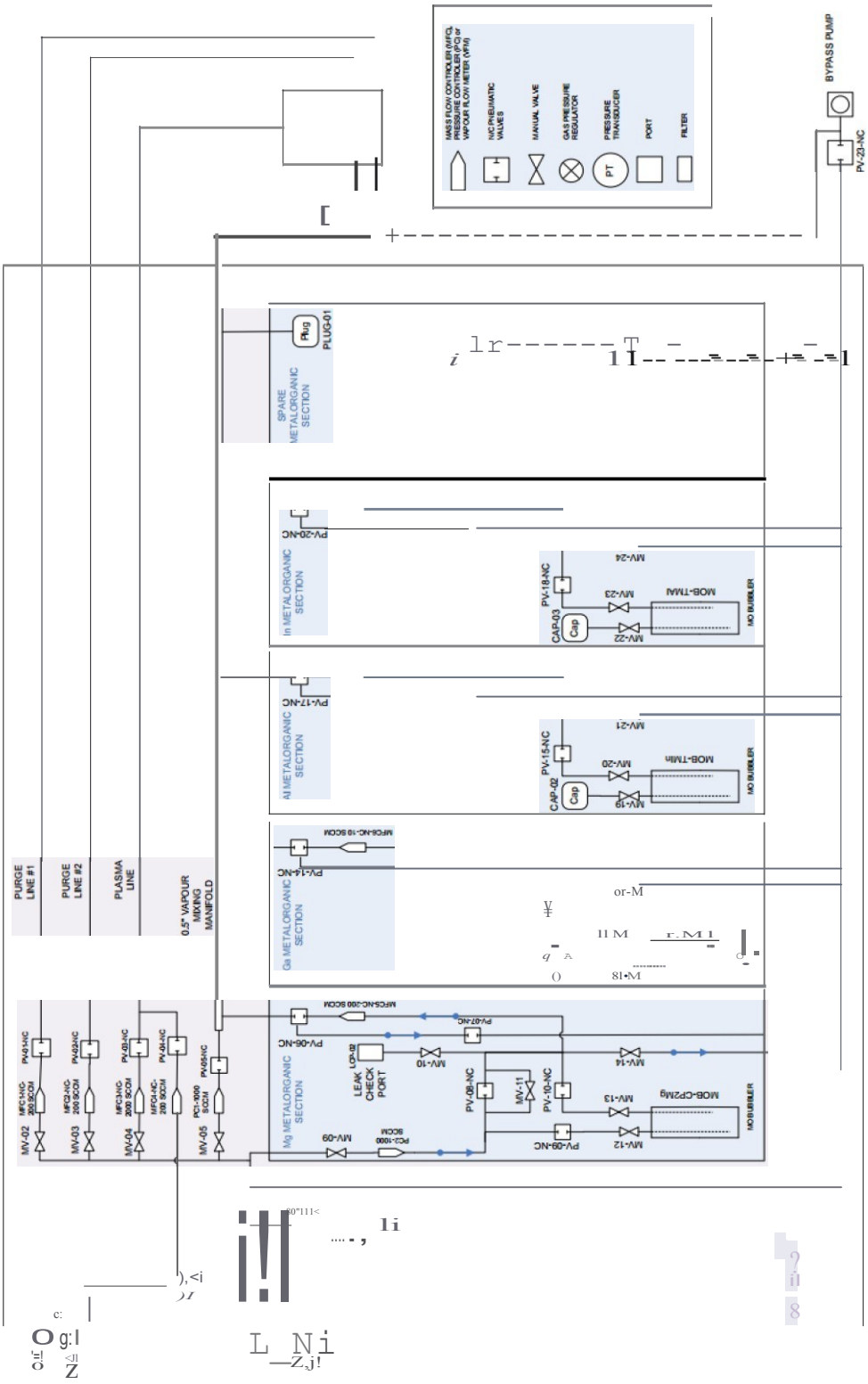

Figure 4 Process and instrumentation diagram for metalorganics. An N2 supply and oxygen-impure N2 supply tanks serve as inputs, outputs are to the turbo-molecular pump (the bypass pump is the base of this pump)

The Lakehead growth system utilizes a dry turbomolecular pump with a rotary backing pump to maintain static pressures as low as 10E-8mtorr and supports up to 2500sccm growth flow. The turbomolecular pump is equipped with a computerized throttle valve that allows for precise control of chamber growth pressure under a specific flow.

The purging during growth is not done directly to the rotary pump, but rather in shunt to the turbomolecular pump. This means that there is no need to maintain constant pressure in the purge line to prevent oil vapor backflow into the system since the turbomolecular pump always provides isolation to the main chamber. Hence, the purge line avoids utilizing a nitrogen background pressure and does not need the one-way valves and foreline traps seen in other systems.

1.3.4 Hollow Cathode Plasma Source

The plasma source in the Lakehead University RPECVD growth system is of the hollow cathode design and provides mainly positive molecular nitrogen plasma species to the substrate. Inside the plasma source, molecular nitrogen gas passes between two concentric metal plates that are fed a fluctuating electric field via coaxial cable from an RF generator/matching box. The field has a frequency of 13.56MHz and typically runs 600W; although 300W power is used to create less excited species for the initial nitridation of the substrate surface before growth begins.

The source is actually very similar to a capacitively coupled plasma (CCP) source in that it uses an RF electric field between two plates in a hollow cavity. The field is greater than the critical voltage for N_2 to initiate a glow discharge and strike the plasma. However, unlike CCP sources which tend to give a fairly low charge density, the hollow cathode plasma source offers electron densities of up to $9X10^{11}cm^{-3}$ – 50X the density of a typical CCP source used for MOCVD.

This allows for a higher proportion of active nitrogen species to inactive and hence more flexibility to the minimum growth pressure, higher growth rates, and even purer films.

An uncompensated Langmuir probe was used by Butcher et al. to measure the density of excited species. It was measured just below the holes of the hollow cathode to be about $9X10^{11}cm^{-3}$ by measuring the current and voltage through the probe. Calculating the charge density from the Langmuir probe equation (given in [14]) is not accurate because electron temperature is not known exactly, but electron concentration can still be found from the ion saturation.

The plasma is introduced into the growth chamber through a series of holes (see Figure 5). The growth can be scaled to larger wafer and chamber sizes simply by increasing the number of holes. The same source can also be used for high surface area MBE as it offers preservation of low growth pressures.

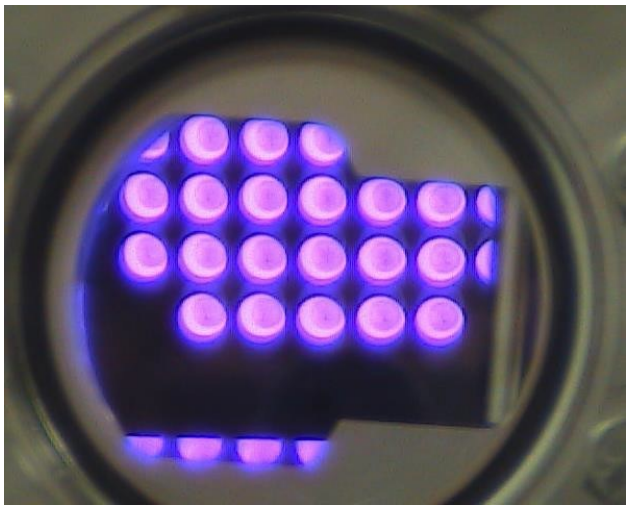

Figure 5 The custom built hollow cathode plasma source delivers nitrogen plasma for growth in a scalable manner through the holes seen here into the chamber. Consists mainly of excited molecular species.

1.3.5 Background Electron Concentration of Nitrides and Doping/Codoping in P-GaN

GaN and InN both exhibit intrinsic n-type carrier concentrations. For GaN, the concentration is $n \sim 10^{16} cm^{-3}$, with $\sim 10^{17}$ cm^{-3} not uncommon [15]. The nitrogen vacancy has been proposed as a viable source for this electron background by Pankove et al [16], and is one of the most prominent among many theories. Given binding energies of that donor ranges from 17meV to 42meV in the literature [17] [18].

Other theories for GaN's intrinsic donor have been reviewed in [14], with the most popular model being an oxygen donor on substituting on gallium sites. Proof that this is in fact a viable donor is indirectly demonstrated in the band structure calculations for single-atom defects in GaN elementary cells performed in chapter 5. Also reviewed in [14] is the afforementioned theory involving unfilled nitrogen vacancies which allow tunneling to the conduction band by electrons of the surrounding gallium atoms. However, this case wasn't covered by the calculations done in this work.

When p-type Gallium Nitride is codoped with oxygen, there are two effects that come into play. Firstly, as noted by Korotkov [19], the hole concentration increases. They measured an increase by an order of magnitude (to $2X10^{18}$ cm^{-3}) and suggested via results from temperature-sweep Hall effect that in the presence of a lattice with oxygen in gallium anti-sites, the activation energy of magnesium decreases from 170(±5) meV to 135(±5) meV. This goes a long way to activating a larger fraction of the magnesium since the average of the thermal Boltzmann energy at room temperature is only about 26meV. It is known that oxygen tends to migrate to grain boundaries and droplet surfaces since this is where the balance of lattice stress energy and energy of nitrogen bonds is most minimized. Perhaps the binding of the more nimble oxygen atoms to these sites before magnesium serves to alter the way magnesium atoms incorporate into the lattice during growth.

Secondly, as noted by [19] and explained by G. Togtema [14], there is drastic decrease to the resistivity of Mg:GaN films codoped with oxygen. For example, it was observed in [19] by Hall effect derivation that the resistivity of a certain p-GaN film changed from 8 to 0.2 Ωcm by regrowth under slight oxygen in the ambient. This may be explained by the weighting provided by carrier mobilities. Since the mobility of holes in p-GaN is orders of magnitude inferior to electron mobility (10 for holes vs. 400 for electrons [15]), the equations

$$\rho = \frac{1}{q\left(\mu_n n + \mu_p p\right)} \tag{1}$$

$$\rho_H = \frac{\left(p\mu_h^2 - n\mu_e^2\right)}{q\left(p\mu_h + n\mu_e\right)^2} \tag{2}$$

clearly demonstrates how the mobility weighting of a change in electron concentration can lead to a notable decrease in resistivity despite being far smaller than the change in hole concentration from p-type doping. It is important to realize that due to this, net hole concentration can exist in these samples despite Hall measurements of net electron (n-type) conductivity. This explains why electroluminescence is still observed for p-GaN films grown on n-GaN with both the magnesium and oxygen doped films measuring n-type by Hall effect.

As observed by G. Togtema [14], commercial samples of p-type GaN also demonstrated variable resistivity and an inconsistent Hall effect measurements of carrier concentration – often measuring n-type (-1.58X10^{18}cm^{-3} to +1.12X10^{17}cm^{-3} and resistivities of 0.038Ω to 4.66Ω). This is clear evidence of codoping effects in some commercially grown p-GaN films by MOCVD, but not others. Nonetheless, consistent electroluminescence was still observed universally for n-GaN growths upon these samples.

A possible use of this codoping is for the deposition of LED structures using a p-GaN base layer. Usually the resistivity of p-GaN is far too high to place on the bottom of a structure where an extensive distance of lateral conduction takes place; and must be deposited atop another layer so that only vertical conduction through the (thin) film contributes to device resistance. Being able to place n-GaN on the top face of the structure may be useful for increasing the luminous output from non flip-chip LED structures, since n-GaN layers tend to exhibit less light absorption due to a smaller internal defect density and thinner top oxide layers.

2 Chapter 2: Nitride Semiconductor Growth Regimes

2.1 Substrate Preparation

C-plane sapphire and <111> silicon substrates for MOCVD GaN growth may be prepared by ambient air anneal at 1050C for 6 hours in the case of sapphire or an etch of KOH solution for silicon. The annealing is done to remove hydrocarbon-related impurities like hydroxides; as well as to help relieve damages from polishing. Since the annealing is done in ambient air, a combination of oxygen and nitrogen incorporation can be expected. However, chamber preheating under vacuum will remove much of the diffused-in oxygen and the subsequent nitridation will smooth the substrate over anyways. If a more thorough etch is desired, substrates of either type can be prepared using a cleaning process of degreasing and chemical etching as per Christou et al [20]. This process is described below.

For degreasing, the substrate is usually immersed in an ultrasonic bath containing either xylene or trichloroethylene for a few minutes and then they are rinsed off with deionized water or methanol of a certain resistivity. Subsequent wet chemical etching removes contaminants from the surface and improves stoichiometry of metal-rich grows; and also creates a thin oxide layer for passivation. The precise process involves exposure in a water-diluted HF solution (1:10) for 5 minutes followed by rinsing with deionized water, 20 seconds in a NH3OH:H20 (10:1) solution to remove carbon and halogen ion contamination from the surface, 15 seconds in a HCl:H20 (1:10) to create the thin surface oxide, and a short exposure to the diluted HF solution from earlier with subsequent rinse to trim down the oxide layer to a precise thickness.

For GaN films grown by the Lakehead MOCVD system the electrical characteristics tend to gravitate toward semi-insulating with mobilities of less than 10. Hence, the passivation layer is not usually required. It was found that the efficacy of contaminant bakeout from commercial sapphire by ambient air anneal is about on par with HF:H20 and NH3OH:H20 treatment when

utilizing the films for our purposes. Hence, this entire chemical etching process has been replaced for our purposes by a simple annealing process in a small annealing oven with air as the background gas. Annealing in the chamber under nitrogen background instead of ambient air has not shown any noticeable improvement to film qualities.

2.2 Growth Kinetics and Ambient Conditions

2.2.1 Typical Growth Conditions and Implications of Growth Parameters

For both GaN and doped GaN, typical growth temperatures used are between 500C and 600C. It is mentioned elsewhere in this chapter that the dissociation fraction of the trimethylgallium precursor species varies drastically in this range of temperatures. Hence changes in growth temperature tend to play a large role in the stoichiometry. Film quality is also severely affected by temperature in the case of p-type gallium nitride as the undissociated TMG species are the source of the hydrogen atoms that form the Mg-H inactivation complex. This incorporation of impurities is also seen in pure GaN films as well. Thermal annealing to deactivate the Mg-H complex is successfully implemented at 600-650C throughout the literature, and hence growths near 600C should be expected to have different dopant compensation character throughout the film as compared to growths done at 500C.

The maximum solubility of magnesium species in the gallium metal droplets is of course temperature dependent; while the resonance of absorption/desorption of nitrogen on the droplet surface as theoretically predicted [21] will also vary with temperature. Hence, the stoichiometry of the latent magnesium in the metal droplets and on the surface of the droplet will both be affected differently by the temperature dependencies of the separate gallium and nitrogen species. Thus it can be expected that the doping will be inconsistent throughout the film and will contain such a memory effect. Perhaps annealing also helps to resolve this inhomogeneity.

Typical pressure used of growth for GaN is about 1.4 torr. The ratio of incorporated nitrogen species to impurities of course depends on the partial pressure of nitrogen. It may then seem like a good solution to reduce impurity concentrations and increase growth rate by increasing nitrogen partial pressure. However this has implications on the desorption of already-incorporated and incorporating nitrogen. There is a background desorption rate for both gallium and nitrogen species during growth, and the partial pressures of each delivered species will affect this desorption. Hence, to achieve good film stoichiometry the rates of desorption need to be compensated for. As desorption is mainly a problem with nitrogen and not gallium (as discussed in 2.2.1), the stoichiometry of the film will vary strongly with applied nitrogen background pressure.

2.2.2 Pulsing and Powder Formation

The hollow cathode plasma source creates both species of molecular and atomic nitrogen plasma. However, the distance of the source to the substrate (~12.5cm) combined with the chamber pressure of growth (~1 torr) means that the atomic nitrogen has mostly been recombined into molecular form by intermittent gas collisions (see Figure 10). This is particularly advantageous for films grown by migration enhanced epitaxy (MEE), and the higher growth pressure is a reason for the viability of MOCVD as a low temperature growth regime for nitrides over MBE where pressures are too low for energetic plasma species to be quelled. Gallium nitride and indium nitride films grown by remote-plasma-assisted MBE in the 400-800C range utilizing MEE have consistently reported film damage. This is caused by high energy excited atomic nitrogen plasma species and ionic/radical species that are not de-energized by collision in transit to the substrate on account of the low growth pressures typical of MBE (typically <E-4 torr) [22] [23].

The pulsing of metal and plasma separately should allow for a higher quality film because the metal adatom species on the film surface have more time to migrate to energetically favorable surface positions before the nitrogen plasma is introduced. When the metal and nitrogen species are delivered to the substrate simultaneously, they react in the gas phase before reaching the substrate. This results in the buildup of dust and powder on the film surface. Pulsing should be able to reduce the amount of powder and dust formation for obvious reasons. Film consistency and surface roughness measures will benefit from this.

Despite these supposed advantanges, during long film growths using the Lakehead MOCVD system, it was found that lightly nitrided metal powders form either directly on the film surface or in the gaseous plasma space immediately above the film. Either way, the metal powder accumulates on the surface. It was found that this buildup can be stopped by periodically lowering film temperature intermittently during growth [24]. In agreement with reports from MBE growths, indium metal nitrides more easily than gallium metal. When the powder is subsequently fully nitrided, the crystal orientations of grown nucleations as seen by XRD tend to be very random in InN and slightly less so with GaN. This seems to suggest that metal powder itself shields the powder-surface nucleation sites from the electromagnetic forces from the underlying film that dictate polarity for unpowdered regions of the film. As electromagnetic shielding is characteristic of metals and less so in dielectric materials, this effect should be expected to be stronger in more lightly nitrided and hence more conductive powders.

2.2.1 Nitrogen Plasma Species, Low Pressure Desorption, and Growth Rate

Optical emission spectroscopy measurements on the nitrogen plasma above the substrate have shown it to be composed mainly of an excited molecular nitrogen species. This is verified by varying power supplied to the nitrogen plasma in the hollow cathode plasma source and

observing the change in the spectral emission. These measurements were performed in the lab by K.S. Butcher et al. The wavelengths corresponding to energies of relaxation for electrons in excited molecular nitrogen states show a great surge in intensity as the plasma source power is moved from a value where the plasma is obviously weak to a value where a strong supply of plasma can be seen (see Figure 6).

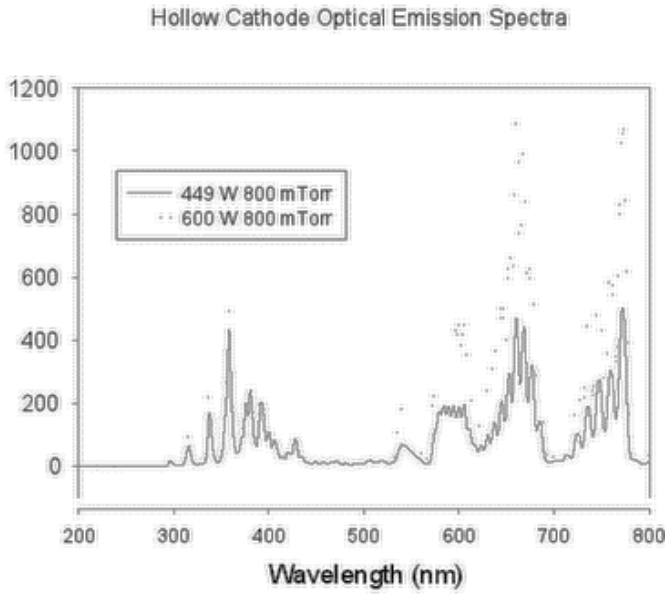

Figure 6 Optical emission spectra for plasma above the sample holder from the Lakehead MOCVD system hollow cathode plasma source. The wavelengths that show great increase with RF plasma power correspond to the relaxation energies of electrons in excited molecular nitrogen states

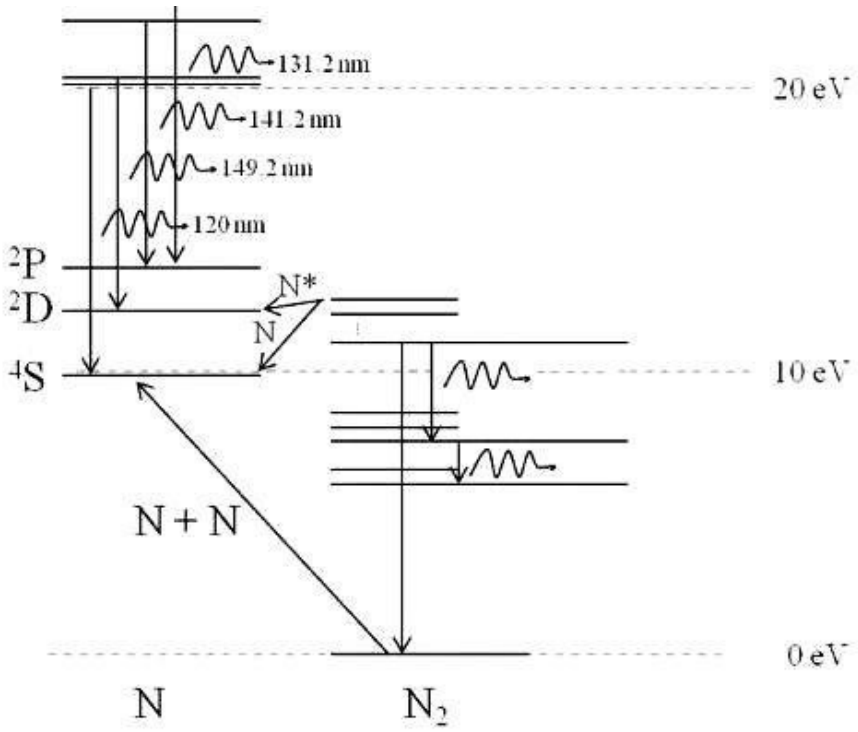

Figure 7 Energy levels of molecular nitrogen placed next to energy levels for orbitals of atomic nitrogen. The two main relaxation energy gaps of molecular nitrogen are represented in the optical emission spectra, as is the decomposition of excited molecular nitrogen above 10eV into atomic nitrogen with 4s orbital filled.

In high temperature deposition of gallium nitride by MBE, lattice decomposition and desorption of gallium is a contentious problem [25]. This is due to the low pressure of growth. Nitrogen plasma with nitrogen background is known to be a good alternative to ammonia based growth for managing this issue due to the fact that the growth temperature may be lowered. This is because the higher temperatures (>1000C) are not needed for the thermal decomposition of ammonia. In the Lakehead University MOCVD system, desorption during low temperature nitride growth is not a problem because the growth pressure is much higher than MBE.

Nitrogen plasma species may be delivered to the substrate for growth as either atomic or molecular species, depending on the plasma source. They contain electrons in excited (non-ground) states rather than being ions. It seems that molecular nitrogen is the better choice for

growth of gallium nitride, since atomic nitrogen seems to cause gallium atom desorption from the lattice at temperatures over 650C [25]. This slows the growth rate and effects film stoichiometry.

The molecular plasma species must first decompose before reacting with gallium adatoms on the film surface. This presents a thermodynamic barrier that may complicate and inhomogenise film stoichiometry. While atomic nitrogen should theoretically allow for a more consistent film due to the more trivial reaction pathway with gallium, it may be that it removes gallium from the film by forming N_2 gas with the nitrogen in the film.

In the discussion presented by [14], a study of GaN decomposition and Ga desorption in plasma assisted MBE was presented to explain the advantages of molecular over atomic nitrogen plasma species. In the referenced study, two plasma sources were used - one that creates excited atomic nitrogen and one that creates excited molecular nitrogen. The excited atomic nitrogen showed a decrease in growth rate with temperature corresponding to the predicted desorption of gallium atoms from the lattice starting at 650C. The molecular nitrogen showed a slowing growth rate with temperature that seems to correspond with theoretical predictions of GaN lattice decomposition. The molecular nitrogen plasma source was tested at two growth rates, with the higher rate starting to show what may be Ga desorption under excited molecular nitrogen plasma. However, it was not noted that growth temperatures and mainly pressures of the Lakehead MOCVD system are well below the thresholds presented by the cited research. Hence, in our case of experiments with gallium nitride, thermally-based desorption could be safely ignored. However, for the case of InN, the contribution was not yet clear.

By observing temperature-series growths of InN at pressures greater than the threshold of plasma damage (so that plasma damage was not an issue), it was determined that the growth rates of these films were not being affected noticeably by temperature. Plasma damage effects

are covered in detail in 2.2.2. In fact, growth rate increased slightly positively with temperature; seemingly in accordance with the temperature-exponential Arrhenius relation that loosely governs most simple spontaneous chemical reactions in typical thermodynamic conditions.

2.2.1 3D vs. Planar growth

Gallium Nitride grown with the Lakehead MOCVD system can achieve consistently Ga-face Gallium Nitride with RMS surface roughness less than 0.5nm for ~200nm thick films. Ga-face GaN is usually much smoother on the surface than N-face GaN, which tends to have a pyramidal surface character. Hence, Ga-face is more suitable for device growth. Films are grown on c-plane sapphire annealed at 1050C in an air-ambient oven for 6 hours.

ω-2θ XRD scans of InN samples at the (0002) peak give values as low as 290 arcseconds full-width at half maximum (FWHM) [24]. When the film grows out in a lateral fashion from the deposited metal droplets, the film shows a surface morphology of terracing from these droplet sites. This is a good sign that growth has proceeded mainly in a lateral fashion per deposition cycle as opposed to three dimensional growth caused by upward metal migration to the plasma. Three dimensional growth results in rougher film surfaces and defect-ridden inconsistencies in stoichiometry throughout the film caused by differences in the nitridation between the droplet surface to the center. These droplet structures are clearly visible in AFM; however droplets may simply be from incomplete film nitridation and do not necessarily signal poor lateral growth conditions. Thorough examination of metal droplet shape as seen by AFM or (preferably) SEM can serve as indication of stoichiomtery and growth character. In two dimensional growth with

good nitridation, the droplets tends to flatten out laterally across the film and good nitridation will tend to cause coalescence between the droplets to occur.

In order to assist with 2D growth, an electrically charged metal grid emanating an electrostatic field was placed just above the substrate holder to cancel the field from the ambient plasma. This positive field was causing negative metal species to migrate upward from the film during growth.

AFM roughness measurements have shown that GaN films grown by a similar low-temperature PECVD processes have smoother films when grown at lower temperatures (450-500C) as compared to higher temperatures (550-650C) [26]. This has been attributed to the fact that higher temperature growths incorporate less oxygen at the surface of the growing film [27], causing the resulting crystallites to be larger. Hence it can be inferred that oxygen plays a pivotal role in the growth of the crystal; with lower temperatures resulting in smaller crystallites and hence tending to yield lower quality films. This is important because it shows that AFM roughness figures alone cannot gauge exactly how planar a growth was. The droplet size, growth temperature, and oxygen incorporation must all be considered in concert.

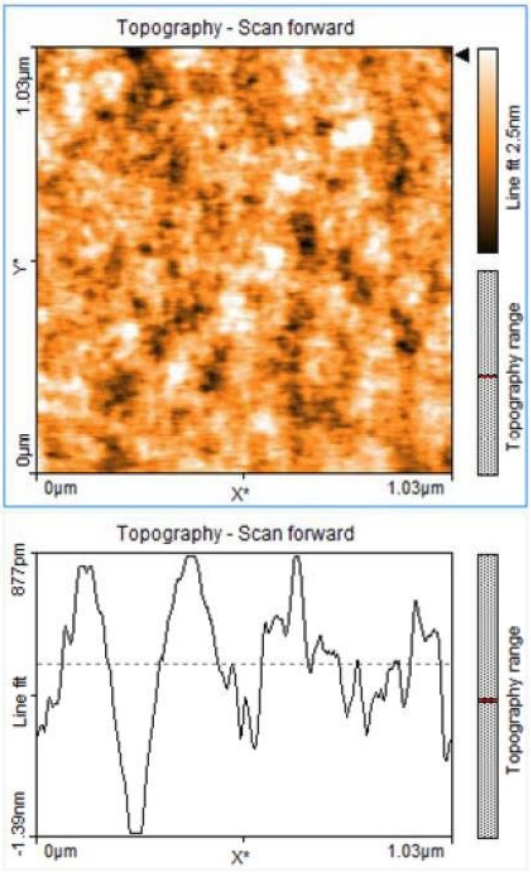

Figure 8 (Right) AFM image of GaN sample with (RMS) surface roughness of 0.46nm. Growth temperature was 665C at 1 torr

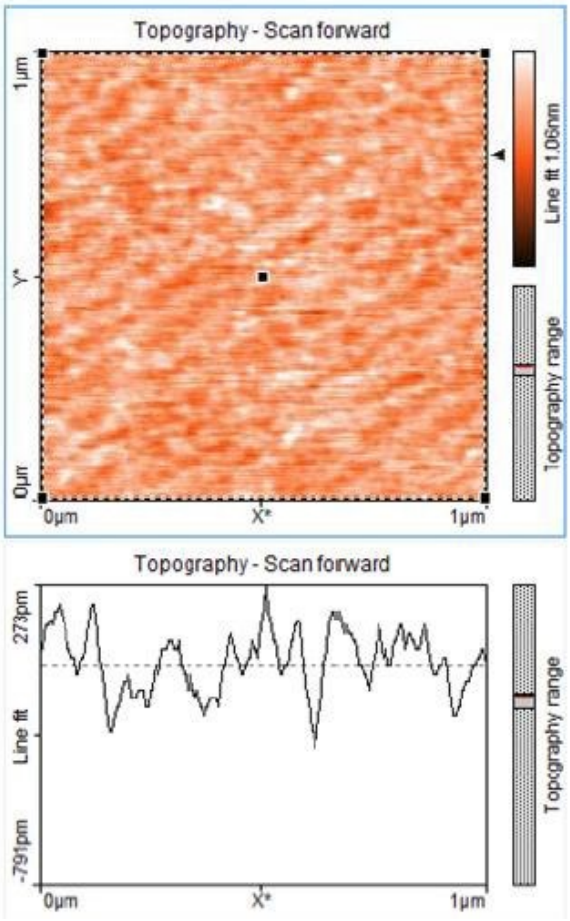

Figure 9 (Left) AFM of InN sample clearly showing molecular terracing which is strong evidence of lateral (2D) growth. Film thickness is 22nm and RMS surface roughness is only 0.10nm. Growth temperature was only 450C. Growth was performed at around 2.3 torr as per requirements of growth pressure to mitigate plasma species induced film damage (see Figure 10)

2.2.2 Thermodynamics of Stoichiometry and Precursor Dissociation

Since the reactions are taking place at a fairly consistent temperature and pressure, the Gibbs free energy and entropy change of the reaction can be used as criteria for spontaneity. Reactions with negative ΔG and positive ΔS are taken to be spontaneous. ΔG for GaN crystal formation is indeed negative, but since ΔS for a crystal is usually negative, at room temperature and pressure the reaction is not necessarily spontaneous. ΔS can be made negative in the

aggregate (amongst all reactants and products involved) and hence the reaction that forms the crystal can be made spontaneous by overcoming enthalpy barriers to dissociation. This is what occurs at growth conditions, and the conditions are adjusted within this requirement to give the highest quality crystal possible with the best growth rate possible.

While changes in desorption rates may not be particularly sensitive to the range parameters used in this work, there is still a present rate that effects the required stoichiometry of the growth. A kinetic model was published by Koleske et al [21] using results of film quality published in the literature that linked the growth temperature and III/V ratio to the quality of films. The primary result was that the delivered ratio of group V to group III elements needed to be slightly greater than the corresponding ratio of desorption rates. Also of particular interest for this discussion is the fact that at high temperatures (>1000C), nitrogen was found to have a desorption rate so high as to be semi-volatile. It may be logical to conclude that the above two facts are related. Nevertheless, desorption of the volatile nitrogen species may still be significant even at lower growth temperatures in the presence of nitrogen plasma, since the plasma can have a pulling effect on the nitrogen in attempt to form molecular nitrogen. Also noted is the square relation between diffusion length and diffusion time ($L \alpha t^{1/2}$); which clearly shows the diminishing returns of metal migration time to achieving 2D, lateral growth.

It was noted in [14], using values for densities of nitrogen species and metalorganic precursors from thermodynamic tables, that the ideal-gas-like flow for these species meant that the delivery of nitrogen species in the Lakehead growth system was over thirty times that of gallium species. It was also noted that under the results of ion density obtained from Langmuir probe measurements that only some of the nitrogen species had been activated by the plasma source; and only a fraction of those would retain sufficient energy on the way to the film surface. However, based on the analysis in [26] of GaN growths performed between 450-640C, it can be found that the initial thermal dissociation of the trimethylgallium precursor in nitrogen plasma does

not occur in its entirety until 580C at ~1.5 torr growth pressure. This discovery was based off of publications that cited these temperatures for 4 torr partial pressure of TMG, but was also found to be valid at ~1.5 torr of total growth pressure by Chen [26]. Based off the results cited, the majority of the TMG species are still not dissociated even at 525C when the concentration curve takes a sharp downward dive. This is significant information for the Lakehead growth system, where GaN is routinely grown between 500-550C in the range of 1.5-2.5 torr. Thus, for GaN growths around this temperature the stoichiometry is further complicated and the temperature dependence of growth will be deviant from the common models.

For a given pressure, nitrogen source, and growth regime, a plot of stoichiometry character vs. growth temperature will yield a pattern reminiscent of solid/liquid/vapour/gas phase diagrams commonly used with pure substances. Such a diagram can be found in [14] for typical higher pressure and temperature MOCVD. Below a certain threshold for III/V ratio, the crystal will fail to grow due to insufficient gallium pressure to prevent desorption. This ratio is around 1; and increases slightly with temperature on account of the widening difference between nitrogen and gallium incorporation/desorption rates. Above this threshold, increasing temperature will move the growth outcome from a film consisting almost entirely of gallium droplets through a phase of gallium rich conditions and eventually into a mixed condition. The mixed condition is a transition region between gallium rich conditions and nitrogen rich conditions. Mixed conditions with lightly nitrogen rich composition is usually the desired region of operation as it results in the lowest defect density on account of complete nitridation of gallium droplets. Slightly nitrogen rich conditions are especially desired for p-type growth since this trivially suppresses the electron background concentration of gallium nitride. If temperature is increased even further, the solubility and activity of nitrogen will become such that it actually displaces gallium and again results in a nitrogen rich film; despite a high III/V ratio.

2.2.3 Plasma-induced Film Damage and Electrostatic Field Management

Despite the fact that the plasma species utilized for growth is mainly molecular nitrogen, the pressures used for growth are on the edge of safety. For gallium nitride the consequences of growing at low pressure (<1 torr) are not adverse because the lattice it fairly tightly bound. However, in the case of indium nitride films grown at around 420-500C must usually be grown at pressures over 2.3 torr to avoid problems with decomposition during growth. The effects of damage on InN can be seen in the X-ray diffraction measurements of Figure 10.

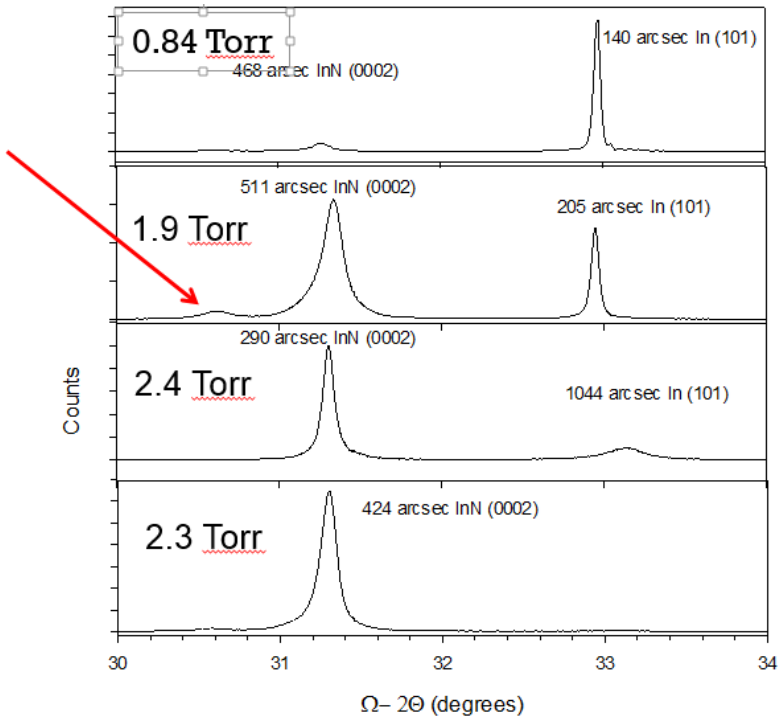

Figure 10 High indium content is seen at low pressure due to high energy plasma-induced dissociation of InN. More molecular collisions between the plasma species and background N_2 occur at higher pressures – less film damage and hence better stoichiometry control is achieved. The red arrow indicates a peak of indium oxide due to a substrate that wasn't heated long enough to achieve complete bake-out of oxygen impurities. Picture taken from [28].

During film growth, the cationic plasma species in the 15cm gap between the plasma source and substrate act as a cloud of positive charge that creates a corresponding electric field. As a result, negatively charged gallium species being incorporated into the lattice experience an upward force. Left unmitigated, this forces causes an upward, columnar migration of metal

droplets that ruins stoichiometry in the area and inhibits good lateral growth. Hence, a metal grid was placed above the sample holder which can receive biases from -80V to +80V to counteract this field.

Incidently, it has been shown that the grid can also help manage plasma species induced damage. A negative grid bias will decrease velocity of ionic and excited atomic plasma species traveling between the grid and the sample holder and reduce the kinetic energy of these species at the film. P-GaN growths with negative grid bias measured by AFM have showed reduced droplet size and count; as well as better nitridation of the surface. Also, when used with a positive bias and a delay between introduction of the plasma and charging of the grid, the grid field can actually accelerate nitrogen plasma species toward the substrate. 1 second should be used in the Lakehead system since it is the smallest interval allowed by the software. This might prove to be acceptable as an etch mechanism in for silicon substrates; in lieu of hydrogen plasma etching or external KOH / HF acid etching.

Capacitively coupled argon plasma sources with two frequency sources have been reported, used for successful SiO_2 etching. One source at the upper capacitor plate is ran at high frequency in the hundreds of megahertz while the other serving as a substrate bias is run at less than 1 MHz. This way, a periodically alternating injection of positive and negative ions to the surface is achieved [29] [30]. It would seem from these experiments that a high frequency applied to the grid may be able to replicate this rapidly varying surface bombardment and achieve in situ, high quality etching of silicon substrates. The lack of a DC field would mean that liberated secondary electrons from the plasma would not contribute appreciably to the dissociation or ionization of molecular nitrogen, and hence the molecular weight of the species on arrival would be similar to argon. Further research is needed in this area.

3 Analysis Techniques for Films and Devices

3.1 Atomic Force Microscopy

3.1.1 Theory

AFM utilizes a scanning tip on a cantilever to scan a sample surface in lines. As the tip moves along the surface, a collection of evenly spaced points in space of the bending of the cantilever is taken. The points from each line can then be interpolated between and a 3D map of the approximate film topography is obtained. A laser is shone onto the back of the cantilever tip, and based on the deflection of the tip on the sample surface, the laser will reflect onto a different portion of a grid of receiving opto-detectors tuned to the laser frequency. Hence, the output bit-pattern from the grid indicates the particular attitude of the tip at that moment. The most common arrangement of detectors is in a quadrature configuration.

There are many modes of AFM scanning, and many different mechanical properties can be derived from different AFM scan regimes using different probe types. Even chemical properties may be derived in some specialized cases.

The three modes of operation for topographic imaging are contact mode, non-contact mode, and tapping mode. In contact mode, the tip is attracted to the surface through Van der Waals forces caused by attraction of spontaneous dipole pairs between the tip and surface (London forces), as well as from forces between permanent dipoles in the surface and corresponding induced dipoles in the tip (Debye forces). The tip moves toward the surface under these forces until the restoring force of the cantilever and the electrostatic repulsion between electrons balances out the Van der Waals forces. The tip is usually a tough material like alloyed silicon and it delivers high (atomic) resolution as it follows the surface of the material very closely. However, this mode of scan is inevitably the slowest and can damage fragile or soft surfaces. The tips used are also expensive and breakage is a prohibitive problem on rough surfaces.

In tapping mode, the tip is fed an oscillating voltage so that the electrostatic repulsion with electrons can be felt out as the tip taps the surface while scanning across a line. This mode is an intermittent between contact and noncontact; offering good scan rates with a good representation of the topography on account of the close contacts made with the surface. It does not damage surfaces like contact mode, so it is useful for soft materials like surface oxides [26]. Variation in the amplitude of vibration allows local variation in the sample hardness and composition to be determined.

Noncontact mode is the mode preferred in this work as it gives the fastest scan rates. Also, with phase contrast imaging it is possible to obtain high resolution images comparable to the other scan regimes. It has been reported that phase contrast in the tapping mode can reveal fine information on grain boundaries and other crystal discontinuities invisible to standard AFM scans [31]. In noncontact mode the tip is fed an oscillating voltage source so that the tip oscillates above the sample. By measuring the phase shift between the applied tip voltage and the mechanical oscillation of the probe, electromagnetic and hardness properties of the surface can be seen. Precise information on the cantilever's mechanical properties are required for precise determination of the contribution from the surface; however most modern AFM systems are capable of handling these calculations and require only the tip model and vendor as input.

The most commonly used and referenced figure of merit for AFM of films in this work is the root mean square roughness; usually taken across a 4X4um square scanning area. However, the roughness average may provide a more mathematically relevant insight and the line roughness can be insightful for inconsistent growths.

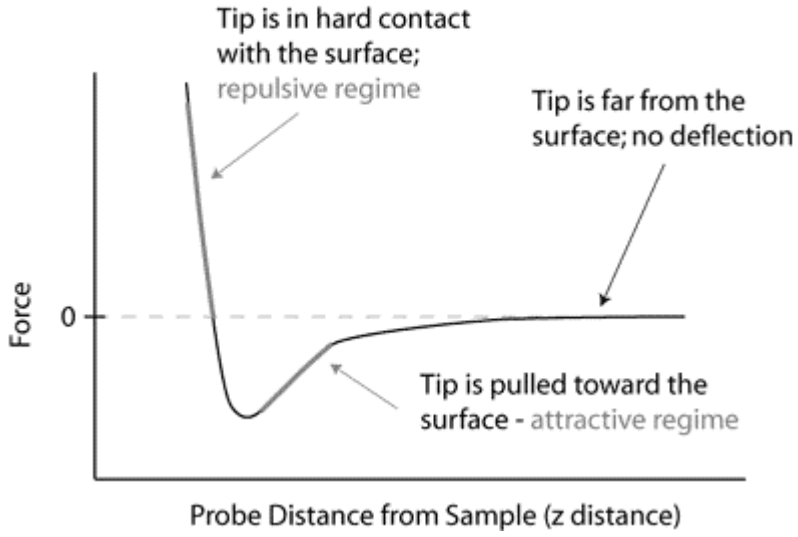

Figure 11 Simple illustration showing the competition between Van der Waals attractive forces and repulsive electrostatic forces as the tip is brought towards the sample surface. The minimum is where the marginal electrostatic repulsion with distance becomes greater than the marginal attractive forces with distance. By using the slope of the linear region of the repulsive regime and subtracting the stiffness contribution from the cantilever, the mechanical stiffness of the surface can be measured. For a tapping regime this graph will in fact show hysteresis, wherein the area enclosed by the oscillation circuit represents the energy lost per cycle.

3.1.2 Measurement Technique

Films are measured using an AppNano system equipped with ACL-A silicon probes. The probes feature a long silicon cantilever with a 30nm aluminum coating on the reflex side to increase the quality and precision of the deflected laser signal. Scanning is done using the phase contrast mode with 100mV supplied to the tip for oscillation. Tip height is controlled by a PI controller, with proportional gain typically set to 10,000 and integral gain set to 1,000. Tip approach to the sample and tip distance are controlled automatically. Scans are typically done in sizes of 4X4um and 2X2um for higher resolution. One to two seconds are typically taken per scan line, with longer times used for publication quality measurements. Typical scans consist of 256 to 512 lines. The typical figure of merit used for film smoothness is the root mean square roughness,

although line roughnesses can provide useful insight on inconsistent samples or films with droplets.

3.2 Hall Effect

Hall effect measurements are used mainly to yield measurements of majority carrier concentration, resistivity/conductance, and majority carrier mobility.

3.2.1 Theory

At the heart of the measurement theory is the Lorentz force equation. This equation

$$F = q(\vec{E} + \vec{v} \times \vec{B}) \tag{3}$$

gives the force resultant of a charge q moving with velocity of magnitude v through a magnetic field of strength B. When a current flows lengthwise through a rectangular conducting member with a normally oriented, widthwise magnetic field of known value, the current will initially be deflected according to the equation – in a direction normal to both the flow and the magnetic field with direction depending on whether the charges are negative (electrons) or positive (holes). Henceforth, a charge accumulation will occur on the current-deflected side of the material. Thus, steady state is reached when the force of the charge accumulation on the flowing current equals the force of the magnetic field on the flowing current. At this point the current flows more or less like usual with a charge accumulation off to the side of the material.

A powerful characteristic of this phenomenon is that it differentiates between the movement of electrons and holes. Given a voltage supplying the current applied lengthwise across the member, it can be seen from the Lorentz force equation that an electron current flowing one way will cause a charge accumulation on the opposite side of the member as a hole current

flowing the opposite way. Hence, a measurement of voltage widthwise across a semiconductor to measure the potential created by the charge accumulation will change in polarity depending on whether electron conduction is dominant or hole conduction is dominant. The sign indicates if the material is n-type or p-type, while the magnitude of the potential (knowing the applied current-inducing voltage and applied magnetic field) indicates how much charge is accumulated. Charge accumulation is proportional to force exerted on the flowing current by the magnetic field, and hence is proportional to the product of current density and carrier velocity.

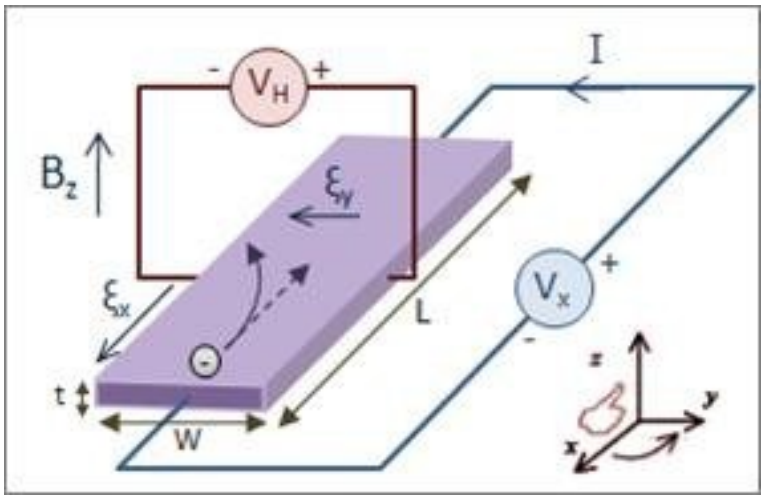

Figure 12 A diagram of the hall effect for electron flow through a rectangular member in the -x direction. The voltage V_H is measured for applied values of B and current. Also shown is the left hand rule of the Lorentz force equation for negative charges, which demonstrates that curling the left hand fingers from the direction of current flow towards the magnetic field will leave the thumb pointing in the direction of electron accumulation. Picture was taken from the internet as a public domain work.

For conduction with a dominant charge carrier species (no comparable codoping), Hall field (between widthwise sides of the member) observed is given by the relation

$$E = \frac{JB}{(p - \square)\square} = -\frac{V}{\square} \qquad (4)$$

where n is the concentration of the dominant charge carrier, V is the hall voltage, and W is the width (widthwise length) of the sample. Also, for electron dominated conduction,

$$J = nev = \frac{ne^2\tau}{m^*}E = \sigma E \tag{5}$$

Where

$$\sigma = \frac{1}{\rho} = e(\mu^n n + \mu^p p) \tag{6}$$

Hence current density of the main current J, as well as E widthwise and B can be measured to find the other useful electrical characteristics σ and μ. Tau is given as the mean time between conduction electron collisions with lattice atoms, and is essential in the derivation of conservation of momentum in time. In this model, momentum given to the electron by the external electric field per unit time must equal the momentum lost per unit time on average by electron collision. Momentum imparted by the field is the simple electric field force equation multiplied by time, while the collision side of the conservation is given by a simple differential equation for the concentration of electrons which have not participated in a collision with exponential scaling factor τ. Dividing the resulting momentum conservation equation out by effective mass m^*, the charge drift velocity v is obtained. Henceforth, using the product of charge density and drift velocity, current density can be linked to the drift velocity in the field as per the above equation. This relation is given by conductivity σ, which is the reciprocal of resistivity. Hence, we have established a full relationship between measurables (current, Hall voltage, magnetic field) and unmeasurables (drift velocity, current density, mobility, carrier concentration).

An often used value to simplify Hall expressions is

$$R_H = \frac{1}{(p - \square\square}} \tag{7}$$

and is the Hall resistance. To express these equations in terms of easily measureable values like currents and voltages, a bit of manipulation is still needed. Using the above equation for Hall field, we can easily see that

$$\frac{1}{q_s} = \frac{IB}{qd} = p - n \tag{8}$$

And this is the predominant equation to determine carrier density. Note that J was replaced by current 'I' by multiplication by material dimensions. We can also see the fact that $J=Ee(\mu^n n-\mu^p p)$ for mixed electron/hole conduction where $v=\mu E$ for a small field. In the case when strong codoping is present and thus one of the Hall resistance denominator concentrations cannot be ignored, the relation

$$R_H = \frac{n_h^2\mu_h^2 - n_e^2\mu_e^2}{q(n_h\mu_h - n_e\mu_e)^2} \tag{9}$$

may be required to explain the outcome of Hall measurements. In this equation, the contribution of a carrier species to the resistivity (and hence hall resistivity) is weighted by the mobility of that species. For example, in p-type gallium nitride with typical magnesium dopant concentration (~$10^{18}cm^{-3}$ – $10^{19}cm^{-3}$), the ratio of codoping (oxygen) species to activated magnesium donor species is actually greater than the ratio of hole mobility ($10cm^2/Vs$) to electron mobility ($400cm^2/Vs$). The electron minority carrier actually dominates the determination of the Hall resistivity, and hence the carrier concentration will falsely give an n-type measurement despite prominent p-type doping. Electroluminescence studies as first performed by G. Togtema [14] on magnesium doped gallium nitride demonstrate the existence of a large hole population orders of magnitude larger than the background electron concentration; despite Hall measurements yielding net electron concentrations.

3.2.2 Measurement Technique

The Van der Pauw measurement technique is the most convenient technique for Hall effect because it allows the above data to be obtained without knowing any problematic specific values

of the material. Full measurements can be made only with knowledge of the film thickness and

resistivity – contact placement and distance is not a factor as the properties dependent on these values are factored out in the derivation [32]. Thickness must be measured separately, but most Hall effect systems include an I-V characterization module to collect resistivities between contacts.

There are only a few demands of the sample so that the assumptions and simplifications of the derivations hold valid. Firstly, the contacts to the material need to be small compared to the sample. Also, when using currents of a certain magnitude it is important to ensure that the contacts are essentially ohmic and with insignificant Schottky behavior so that contact potential does not skew measurements of weak charge accumulations. This is very important in the measurement of semi-insulating nitride semiconductors since for the maximum current output of most Hall measurement systems, a magnet of insufficient intensity cannot usually be found to cause a charge accumulation strong enough to entirely trump Schottky contact potentials. Selection and application of good contacts to gallium nitride is discussed in 1.2.2.

The second demand presented is the contacts be at the edge of the sample, with the sample being preferably disc-like or square. Lastly, the sample needs to be of consistent thickness throughout. This can create error with measurement in certain MBE systems where the effusion cells tend to create thicker films toward the middle of the substrate.

The measurement process involves first finding the resistivities between four diametrically placed contacts. Then the main current is rotated through the permutations of contact pairs; measuring the change in resistivity between the widthwise (charge-accumulation) contacts each time the current is changed. The equation

$$\mu = \frac{\Delta R^{bd\ ac}}{\rho_{abcd} + \rho_{bcda}} \frac{2 * \ln 2}{\pi B} \qquad (10)$$

Gives the mobility in terms of the change in resistance measured between contacts b and d (lettered clockwise) when a current is applied flowing from contact a to c. Note that this technique only gives a single value for mobility and cannot isolate hole and electron mobilities. Thus, for codoped materials any combination of temperature and pressure carrier freezeout techniques must be used to separate each mobility. This was attempted in [14] on magnesium doped gallium nitride but failed to return consistent p-type measurements.

The denominator resistances of the above equation are the no-current resistances with superscripts defined likewise. To find the resistivity of the material, the equation

$$\rho = \frac{R^{ab,\,cd} + R^{bc,\,da}}{2} \frac{t\,\pi}{\ln 2} \tag{11}$$

Is used. The symbol t is the sample thickness Note that this equation is only valid if the type of transverse hall measurement implied by the resistances in the equation is isotropic (rotationally identical). In other words, the terms in the numerator must be approximately equal. Thus, this equation does not work well for orientable materials with substantial anisotropic behavior. With mobility and resistivity determined, the carrier concentration can be easily calculated.

An Ecopia HMS3000 DC Hall Measurement System is used to perform the measurements. After the sample is mounted on a pre-routed board and the leads connected to the contacts, an I-V curve is taken to determine static resistivites and to check for good contacts. Then, with a permanent dipole magnet of 1T and the input of film thickness and desired current (usually around 1nA to tens of mA, depending on film resistivity), the first half of measurements are automatically performed. Then the magnet is flipped in polarity and the second half of measurements are taken. Measurements are averaged out, and final calculation results are displayed.

Some highly resistive films cannot be measured due to the limitations either from current output from the device or from measurements below the noise floor. A stronger electromagnet and possibly AC Hall measurements are a viable solution to fighting the noise floor, while a more powerful Hall system is required for highly resistive films.

3.3 Scanning Electron Microscopy

3.3.1 Theory

Functioning of the SEM starts with an electron gun. A high voltage supply is applied between an anode and a cathode which is either an electrical filament or has been fitted with a heater. The cathode emits electrons due to the effect of thermal liberation of electrons from an atom – thermionic emission. A series of anodes, a Wehnelt cylinder, or some other setup acting as a control grid guides the electrons out of the electron gun. The electron beam is focused further to a very fine focus through a column equipped with a series of electromagnetic lenses - three for the device used in this work. The first lens from the electron gun is the condenser, which feeds the electron beam to a point so as to be rarefied by a narrow slit called the beam blanker. The second is the objective lens, which refocuses the beam after the blanker; and finally are the scanning coils. The scanning coils function as an Einzel lens to guide the electrons to a spot shape on the sample surface in the chamber.

Because the electron beam is so narrow, it is able to penetrate rather deeply into the surface of the specimen and yields a high imaging depth of field. The scanning spot is moved around the surface, and the resulting images can be used to make inferences about a range beneath the sample surface; and not necessarily the surface alone. An SEM chamber usually operates in high vacuum conditions – the unit used for measurements in this work uses a vacuum on the range of 10^{-4}torr. This is necessary to preserve the momenta of the electron beam to the

sample and the secondary electrons from the sample to the detector against collision with ambient species. Typical resolutions for SEMs mounted on separated, floating concrete pads are less than 1nm; however the resolution of the device used in this work is around 5nm.

When high energy electrons or ions (usually protons, as usually used for PIXE measurements) are directed toward the sample surface, they may liberate higher energy electrons close to the nucleus. Subsequently, an electron from an outer orbital will fill this vacancy and and X-ray will be emitted. Collection and measurement of the energies of these X-rays by an energy-dispersive spectrometer allows mapping of the discrete energies to the known orbital gap energies of elements. This allows the identification of the atomic species present in the sample; as well as the relative concentrations of them. This is known as energy-dispersive x-ray spectroscopy (EDX). This measurement technique is often a useful alternative to secondary ion mass spectroscopy (SIMS) measurements, and was used in this work to gauge impurity concentrations in films examined transversely.

Secondary electron detectors are the most commonly used to image specimen. However, also commonly used with the SEM are collectors for backscattered electrons (BSE) and characteristic X-rays. BSEs have undergone a process with the specimen called elastic scattering, whereby an electron from the beam simply deflects with little loss of kinetic energy. Because this process is mechanical in nature, it will depend heavily on the atomic mass of the specimen's constituent atoms. The collection and counting of these electrons along with the characteristic X-rays forms the basis for mass spectroscopy measurements like EDX.

Another techniques viable with SEM are wavelength dispersive X-ray spectroscopy (WDS), which is similar to XRD but more precise in that the ambient vacuum allows the detection of weaker photons from lighter elements. Smaller phases that may or may not contain light

elements can also be mapped and identified more precisely; since the X-ray beam used is much more focused than in XRD.

SEM images don't necessarily represent the visual appearance of the specimen since it is measuring secondary electron generation from the higher energy orbitals (valence orbitals mostly) due to the bombardment of focused electrons; and not the optical reflectance and phosphorescence that determines visual appearance to the unaided eye.

3.3.2 Measurement Technique

To avoid buildup of electrons within the sample from the continuous stream of electrons - and to avoid surface effects due to the buildup of electrons on surface states caused by background fields – a thin layer of carbon paint is applied to the films before inserting them into the chamber. Samples must be broken so that they can be mounted vertically; if the characteristics of the film layers (as opposed to just the surface) are to be measured. Mounting is done in a simple slotted metal disc, fastened tight by a set screw. The secondary electron detector (detecting valence electrons mostly) is used for imaging, and another is used for EDX measurements of sample composition.

Typically measured is the sample thickness. Working distance for the material from the lens is chosen to be the minimum acceptable value (4mm) as it gives the smallest spot and hence the highest resolution. Also, the since the sample is never perfectly oriented vertically, oblique images of the surface may also be taken. These images tend to reveal much information about the effects of droplets and inclusions on nearby surface morphology, and they are spectacular in appearance due to a high depth of field that SEM measurements offer. Also studied are film cracks developed from thermal stresses, fan-out of defects near the substrate to the upper layers, substrate integrity (to gauge the effectiveness of thermal annealing and effect of defects on the

film), split droplet cross sections, and comparative studies between single films and identically grown films with additional layers grown on top.

3.4 X Ray Diffraction

3.4.1 Theory and Measurement Technique

The basic model used to introduced the nature of this measurement are the Bragg conditions. The Bragg conditions state that the X-rays directed toward the crystal lattice are deflected specularly – ie. the angle of departure equals the angle of incidence as measured relative to the plane of the specimen. This leads to the formulation of a simple equation for positive interference

$$2d\sin\theta = n\lambda \tag{12}$$

Where d is the spacing between interatomic planes looking downward into the film, and λ is the wavelength of the X-rays. This equation simply states that when the path length difference between successive atomic layers of the crystal is equal to some integer multiple n of λ, there will be constructive interference of the reflected waves that can be measured. Thus, if we vary the angle of the X-ray emitter and collector such that their angles made with the plane of the film stays the same, the collector will detect peaks of constructive interference for angles that satisfy the Bragg equation for the z-axis lattice parameter of the crystal under investigation. These positively interfering waves induce charge flows in the receiver which are counted by a collector.

Note that this condition assumes that the crystal is oriented perfectly along the z-axis. If the surface is polycrystalline, then identical crystals of certain different orientations will provide peaks at other angles as well. Also, when the lattice develops strain the profile of the diffraction peaks will widen. The width of the peaks is often given as a figure of merit for films, and is known as the full-width at half maximum (FWHM). The measuring procedure for this figure is self

explanatory. If a variable skew is placed between the collector angle and the emitter angle, or if the angle of the sample on the sample holder is varied, information about crystallite orientation on the surface caused by edge defects and dislocations can be obtained.

If it is considered that X-rays are simply the carriers of electromagnetic radiation, we can see that when an X-ray gets close to an atom in the crystal it will distort the electron orbital clouds of the atom. This distortion has the same periodicity as the incident radiation. Since this distortion tends to be isotropic, we can in fact treat the atoms of the lattice as isotropic re-emitters of electromagnetic waves and we have a Huygens-like wave model of Rayleigh scattering. Hence, the reflection in XRD is not specular, but it is a diffraction at scattering centers given by the atoms. Using this model we can infer that there will be positive interference of the diffraction patterns anywhere in space where the path difference between scattering centers is such that the wavevectors of the diffracted waves (which are equal to the wavevector of the incident photon) give a multiple of 2π when multiplied by this difference. With some simple manipulation of the equation, we realize that the condition for positive interference is that the difference in wavevectors should be reciprocal space vectors. This is the Von Laue condition. The case of positive interference is often displayed by placing the two wavevectors head-to-head, drawing the resultant vector between the tails, and separating it by a normal plane – the Bragg plane of the first Brillouin zone.

Because the intensity (height) of the peaks depends on many system factors as well as the scan speeds and regimes used, peak intensity is usually only comparable between samples scanned by the same system under the same scan procedure.

The scanning regime used in this work is the common ω-2θ scan. In this scan the angles made with the specimen plane between both the emitter and collector are the same, and the angle is varied at a constant rate (ω) rather than moved to discrete angles. Scans are usually done from

2θ = (30-45) because this range usually captures all of the important features of InN, GaN, AlN, sapphire, and alloys. Peak 2θ positions can be referenced against databases of peak positions, like that of the ICDD.

The X-ray emitter is a copper coil or bar, bombarded by electrons from an electron gun and emitting based off the principle of secondary electrons and characteristic X-rays of orbital leaps. There are three predominant X-ray energies created by the copper – labeled as the k-α_1, k-α_2, and k-β. The k-β is usually too weak to be detected, but the k-α 1 and 2 are both strong enough. The k-α_1 is chosen as the basis of measurement because it gives the strongest peaks; but since the k-α_2 is close in energy and intensity it must be filtered for an accurate diffractogram to be obtained. This filtering is done by software, and it also filters the k-β for the case of strong peaks.

3.5 Transmission Measurements

3.5.1 Theory

An electron can absorb a photon to increase its energy in accordance to the relationship E=hf where f is the photon frequency and h is Planck's constant. In a direct bandgap semiconductor crystal, the absorption of a photon of light of sufficient energy is the only requirement for the promotion of an electron from the valence band to the conduction band.

When a beam of monochromatic light is shone through a semiconductor film, it will show an absorption in accordance to the available energy levels in its band structure to which electrons can move which are separated by the photon energy from the current energy of the electron. In other words, by sweeping a monochromatic light source through a range of frequencies, we obtain a picture of the average distribution of energy levels relative to the typical population of electrons under the film conditions (thermodynamic, mechanical, chemical) during the scan. Like most types

of measurements, transmission scans are done at room temperature, but they may be done at typical device operating conditions for a more representative scan; or to reveal more information about the distribution and density of electron states.

A common phenomena that occurs during measurements involves the internal reflection of the light at the film interfaces. When the wavelength λ of the monochromatic light has the relation

$$N\lambda = 2nt \tag{13}$$

then the light is able to form a standing wave in the film and the positive interference allows for the greatest possible transmission of light through the film. Likewise, when wavelength is such that a light wave crest is necessarily incident upon at least one of the film boundaries, the light will tend to reflect away at the interface and be lost. Hence, the semiconductor acts like a Fabry-Perot cavity and oscillations will occur on an optical transmission plot of transmission vs. frequency. Note that n is the refractive index of the film and is needed to adjust for the speed of light in the medium (where frequency f is fixed identical in all media) so that the optical path length is used rather than the physical length.

Since the value of N is not necessarily known at any given transmission spectrum oscillation maxima (due to incomplete information about film thickness), error can be avoided by using the difference between two peaks instead of only one. This way, using the wavelengths of two peaks λ_1 and λ_2, it is easy to show that

$$t = \frac{A}{2\left(\frac{1}{\lambda 1} - \frac{1}{\lambda 2}\right)} \tag{14}$$

Where A represents the number of peaks in between the two measured peaks (A=1 for no peaks in between, A=2 for 1 peak in between, ect.). Note that if SEM measurements of film thickness are taken accurately, a refractive index can be obtained from this equation and used as a closer

approximate than values typically give in the literature, which are usually for very high quality GaN.

Another prominent feature of the transmission plot of a semiconductor is the quick rolloff that occurs with increasing wavelength. This is due to the bandgap energy of the semiconductor. A rough way to find the bandgap from this dropoff is to follow it down linearly from the upper linear portion to the x axis. The x intercept will be roughly the wavelength corresponding to a photon of energy equal to the bandgap energy E_g.

For a more accurate estimation of band gap energy, take first the equation for film induced light attenuation for photon absorption near the band edge transition

$$I(l) = Io\, e^{a(hf-Eg)^{\gamma} l} \tag{15}$$

Where 'a' is a constant of proportionality, t is film thickness, Io is the light intensity before entering the film, and I(l) is light intensity at a length of travel l into the film. γ is a derating constant pertaining to the probability of photon absorption. The value is ½ for direct transitions such as the dominant band transition in GaN, and gives the strongest attenuation. The value is 2 for dominant indirect band transitions like in silicon. It has also be used in the characterization of forbidden transitions with values of 3/2 for direct and 3 for indirect [33].

If the logarithm of both sides is taken, it can be seen that

$$\ln \frac{Io}{I(l)} \alpha \sqrt{h\nu - E_g} \tag{16}$$

Where $(\ln \frac{Io}{I(l)})^2$ is called the optical density squared (ODS). If ODS is plotted with photon energy on the x-axis (energy of the photons from the spectrometer), then an interpolation from the exponential region of the plot down to the x-axis (where ODS=0) will give the energy at which $hf=E_g$ since all other factors in the original equation are necessarily nonzero. Thus, this energy is the bandgap energy.

Locating the exponential region of the ODS curve can be difficult because it is not linear. It often helps to identify the point where the region ends by taking the logarithm of the y-axis and noting the energy where the linear region just above the x-intercept ends. This should be made a point on the original non-log graph, with another point taken at an energy just above it. Then, the linear interpolation through these two points down to the x-axis gives the energy of the bandgap. The theory behind this technique is that the exponential region of the ODS plot is in fact a linear region corresponding to the logarithm of the (exponential) density of states (DOS) in the conduction band. The ending of this linear region represents the transition into photon energies which are smaller than the bandgap and thus tend to correspond to absorption by electron populations which either do not have exponential DOS above them (valence band electrons) or have a small DOS above them (conduction band electrons). Taking the logarithm of the logarithm helps to remove misleading behavior caused by defect and forbidden states near the conduction band minimum; as well as conduction band-edge electrons with defective DOS above them, and errors in the transmission measurement caused by light scattering, ect.

It was noted by [26] in his own analysis of transmission measurement techniques that the slope of the ODS plot gives a rough qualitative measure of film thickness. Steeper slopes would correspond to thicker films.

Another feature of transmission % vs wavelength measurements is the softening of the slope of the bandgap-induced rolloff at low values. Called an Urbach-Martienssen tail, it is caused by the filling of conduction band edge states by localized impurity donor states [34]. The film obtains a steeper rolloff in the first part of the tail (non-Urbach region) due to the empty defect states, but a more gradually sloped tail for energies that the donor carriers have partly filled. The knee represents the unlikely transition from degenerate strain states just above the valence band to degenerate states just above the conduction band. While the impurity states are localized, the filling of the conduction band states by these donors is not. The absorption in the Urbach tail can

be expressed as an exponential function of the photon energy minus the photon energy at the tail's inception, divided by a scaling factor called the Urbach energy, with the out front coefficient being absorption coefficient at the energy of its inception. Hence, it is simply an exponential interpolation from the tail starting point with scaling factor the Urbach energy [35] [26].

If the bottom of the conduction band is for some reason partly or completely filled, the optical transmission measurement will show a shift of the bandgap rolloff to a shorter wavelength compared to the same material with a conduction band edge populated in the regular way. This is prevalent in GaN samples where the background electron concentration can vary considerably from sample to sample. Called the Berstien-Moss effect, it is due to the fact that valence band edge electrons must traverse an energy gap that spans over both the band gap and the filled conduction band states. Lastly, a dropoff in transmission intensity is often seen for samples as wavelength is increased from the band edge. This is due to the product of electron promotion from the conduction band to higher conduction band states with the density of states. Therefore, it will tend to be more prominent in samples which also display a Berstein-Moss effect due to conduction band edge filling. The promoted electrons will soon re-emit the photon, but because there is no stimulated emission the photon will depart in an arbitrary direction and the probability of entering the narrow detector aperture is low.

3.5.2 Measurement Technique

Measurements are done using a Cary 5e spectrophotometer. The sample is mounted on a sample holder such that the beam strikes at normal incidence to the top of the film. A baseline of 0% is done with a beam blocking apparatus placed in front of the beam emitter so that the receiver can record the background ambient light in the measurement chamber. A baseline transmission measurement is then done on a blank piece of sapphire if that is the substrate; or if an AlN buffer layer is used for growth a suitable sapphire substrate with an AlN buffer layer only is used.

Scanning is usually done from the near infrared (1200nm) to the near ultraviolet (300nm), with a detector switch being done once during scanning. The switching usually creates a noticeable glitch artifact on the transmission graph.

4 Results: Film Growths

While the utilization of aluminum nitride buffer layers was indeed pioneered elsewhere, the technological achievements in this work regarding GaN and p-GaN on amorphous AlN and the corresponding recipes for this particular prototype growth system are unique and original; being presented originally in this work. Furthermore, the recipes and many of the growth techniques utilized with the different film types – namely plasma pulsing - with this prototype growth system are unique and original to this work.

4.1 Aluminum Nitride Buffer Layers

The use of AlN buffer layers atop sapphire was a pivotal step in the development of high quality GaN. Thin amorphous AlN layers were first used by Amano et al. [36] to create smooth, crack free surfaces using MOCVD at around atmospheric pressure. It was also found in non-migration-enhanced MBE studies that monocrystalline AlN buffer layers increase the quality of InN to the point of being comparable to migration-enhanced growth [37]. This means that high quality films can be grown faster. Mobility increased with buffer layer thickness, and background electron concentration decreased.

Proper nitridation of the sapphire substrate to form AlN_xO_{1-x} is beneficial to GaN growth using amorphous layers. Annealing studies of nitridation layers showed that flat nitridated layers of sapphire assisted in the favoring of two-dimensional growth by GaN films grown above. Annealing the nitridation layer resulted in densely spaced protrusions that gave favor to 3-D GaN growth [38].

ZnO buffer layers grown by MBE with atomic species oxygen plasma have also been used with success for high quality growth of GaN sandwich structures [39]. They use an AlGaN active region with n and p type AlGaN collectors in the middle, with continuously graded stoichiometry

of Al/Ga to the outer GaN collectors. This creates an ohmic contact device with excellent carrier injection.

4.1.1 Growth Recipe and Results

Before the aluminum nitride buffer layer is deposited, the sapphire substrate is nitrided by a 100W RF nitrogen plasma for one minute. The details are available in the growth recipe for PGaN discussed later in this work (section 4.2.1).

The AlN buffer layer used for this work was deposited at 575C with 1700sccm nitrogen plasma at 600W, 12.56 MHz RF source power and 0.88sccm flow of trimethyl-aluminum at 1370mTorr. 50 cycles of an 8 second plasma-only pulsing time and a 4 second metal-only pulse time creates a 7nm thick layer of amorphous AlN atop sapphire. The stoichiometries for the buffer layer were taken from the recipe used by Amano et al. [36] that gave high quality GaN films, and were suitably adapted for the species used by the Lakehead RPECVD system.

Three different buffer layer thicknesses were tried, with thicknesses assumed to be 5nm, 7nm, and 11nm by interpolation from the growth rate calculated for the 7nm layer, whose thickness was measured by SEM. Thickness was controlled simply by changing the number of deposition cycles. However, only two of them (7nm, 5nm) were studied by AFM, and only one (7nm) was studied by XRD, SEM, and transmission measurements. This is because the independent buffer layer growths were done after growths of GaN on similarly grown buffer layers, and the 7nm layer gave the best quality GaN.

The first task of measurement was to observe the film thickness. Since the layer is expected to be only 8.9nm thick (as per [36]), SEM measurements did not show much detail. However, enough clarity was achieved to gauge the thickness to the nearest nanometer; and the film thickness looked to be about 7nm.

Measurements of the film surface by AFM showed a generally very smooth and continuous surface with some rougher regions consisting of pits and deep ridges that would cause breaking of the scanning tip. While the depths of the (typically round) pits could be seen to be on the scale of a nanometer or two, the exact depth of the ridges could not be measured because of consistent tip breakage; even when slow speeds were attempted. Only the surfaces corresponding to buffer layer thicknesses of 7nm and 5nm were investigated, and both had a very similar morphology.

XRD measurement of the 7nm AlN layer showed a very large peak corresponding to the sapphire substrate (not shown in Figure 13). However, accompanying this peak was a peak thousands of times smaller corresponding to the (0002) orientation of aluminum nitride. This clarifies that the layer is indeed amorphous, however it indicates the presence of some highly oriented AlN crystallites in the metal. It is unknown whether these crystallites seeded atop of amorphous metal or if they seeded from defects either inherent on the sapphire or created by the nitridation process. A possible future work is to see if higher power plasma treatment of the sapphire surface - and for a longer duration – leads to a large increase in the nucleation of AlN crystals; and a corresponding increase in the XRD peak. It may also be useful to use mixed crystalline / amorphous buffer layers as seeds for nitride films, if the crystallites are highly oriented.

Transmission measurements revealed a primarily metallic behavior of electron states. While the characteristic of the curve is not perfect of a metal or of a semiconductor, the gradual rolloff with photon frequency shows no prominent bandgap other than starting from below 800nm. The material itself can be understood to be a scarce mixture of crystallites in a matrix of aluminum metal with covalently held nitrogen atoms – perhaps a semimetal.

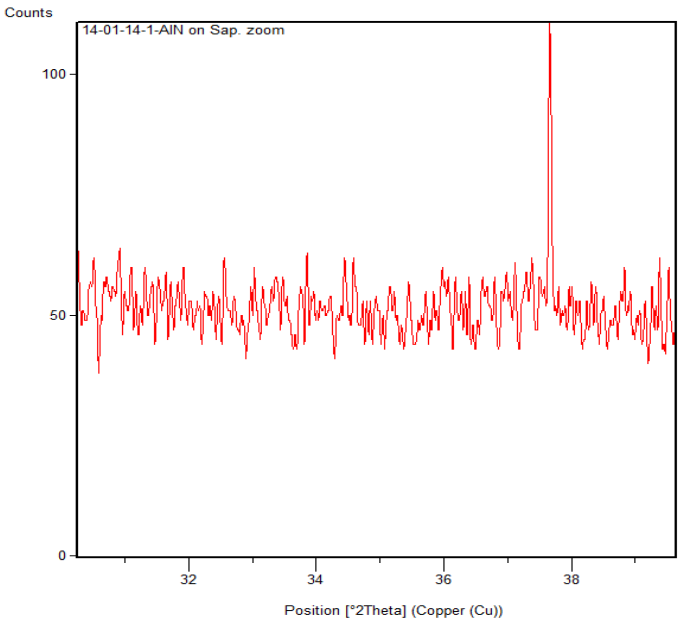

Figure 13 XRD scan of a 7nm thick aluminum nitride buffer layer. While the layer itself is taken to be mainly amorphous, the narrow peak in this diffractogram indicates the existence of some highly oriented crystallites.

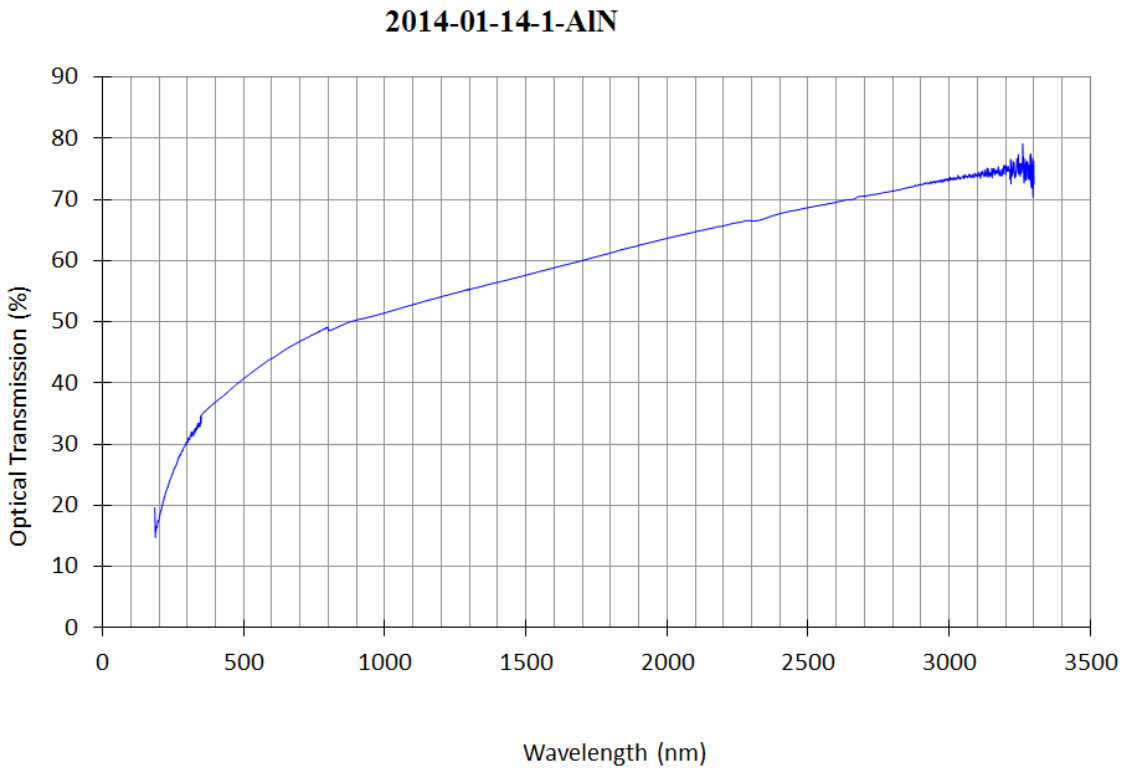

Figure 14 Transmission measurement of 7nm thick AlN buffer layer. The gradual decrease of transmission seen here through increasing energy is characteristic of the electron states in a metal, semimetal. The material itself is a mixture of scarce crystallites in a matrix of Al with covalently held nitrogen atoms.

4.2 P-type Gallium Nitride on Sapphire

There are four sources of problems for p-type doping of GaN by magnesium. Collectively, they make achieving high hole concentrations at room temperature difficult. They are the Mg-H inactivation complex, the somewhat low limit on Mg solubility in gallium droplets during growth, self-compensation of magnesium acceptors in high doping concentrations, and low activation fraction of magnesium donors at room temperature. Unfortunately, the only two problems that can be addressed by changes in the growth regime are the Mg-H complex and self compensation.

In this work, it will investigated if the problem of the magnesium-hydrogen complex can be reduced by using species pulsing – ie. pulsing of the nitrogen plasma and precursor metalorganics separately. This way, the hydrogen impurity atoms originating from the metalorganic precursors are given time to thermally diffuse out of the gallium droplets and off of the surface and be evacuated from the chamber before the nitrogen plasma species is pulsed in to form the film. Before this study, only continuous plasma flow growth regimes have been tried with this RPECVD growth system, and p-type conductivity in GaN has not been achieved. Also, the magnesium flow rate is varied to check for effects of Mg-Mg self-compensation.

Increase in the hole concentration by dissociation of the Mg-H complex by low energy electron beam irradiation (LEEBI) was first reported by Amano et al. [40] Despite reports that LEEBI only served to do this by heating the lattice, depth-resolved cathodoluminescence (CL) studies of the spatial distribution of injected carriers vs. depth of LEEBI treatment strongly suggest that the complex dissociates via electron-hole recombination [41] [42]. In this theory hydrogen doesn't leave the lattice, but takes the place of nitrogen vacancies and creates additional recombination pathways.

It has been shown that high quality p-type GaN can be grown at atmospheric pressure and temperatures of ~1000C by MOCVD by using a thin AlN buffer layer with a thick GaN overlayer. In fact, this discovery paved the way for the creation of the first blue LEDs by using a GaN p-n junction. In this work, the feasibility of these techniques for low temperature and pressure RPECVD techniques is investigated.

4.2.1 Growth Recipe and Results

Recipe

The growth recipe used for the film result showed in this work is similar to the growths of which it is compared. The comparison growth is a growth that was performed for [14] as part of a pGaN temperature series (many consecutive growths done with step increments in temperature between each one); where the selected growth represents the best results in the series for XRD crystal quality, AFM surface smoothness and apparent stoichiometry, and Hall effect electrical characteristics (film resistivity, carrier concentration/mobility). All the growths in that series gave wildly inconsistent Hall effect readings for carrier concentration, but offered variable resistivities. They were all measured with gold contacts which were placed on with a soldering iron and annealed in the growth chamber; and were essentially ohmic for the Hall currents used. No etching on was needed on the films; as they were cooled in a nitrogen background pressure and removed from the chamber at room temperature.

The growths performed for this work were done in the same way, They were performed on a quarter wafer of 2" diameter, 0.33±0.03mm thick top-polished sapphire heated to approximately 900C for 5 hours, on a sample holder with a rotation of 60rpm; since the metalorganic shower head that pulses in the metal species is off on one side of the sample holder. The blank spots of the sample holder were covered with a combination of an alumina disc and other "dummy growth" sapphire substrates to even out the surface thermal conditions and to prevent deposited metal on the sample holder from evaporating and re-depositing on the film at inappropriate times during the growth cycle.

The substrate was first heated to 550C under 1000sccm of N_2 flow (not plasma) under a chamber pressure of 1370mtorr. After 1 hour of heating, the sample was nitrided for 1 minute at 750mTorr under 100W of RF plasma power; still at 1000sccm flow and 550C. Then, the AlN buffer layer was deposited at 575C with 1700sccm nitrogen plasma at 600W source power and

0.88sccm flow of trimethyl-aluminum at 1370mTorr. This AlN buffer layer used 8 seconds of plasma pulsing time and 4 seconds of metal pulsing time. It takes 50 cycles of deposition to deposit the 7nm thick buffer layer.

The main growth was performed by the same regime as the buffer layer; however the temperature is lowered 25C to 550C. Flow of trimethyl-gallium was 0.47sccm during metal pulse-in, and Cp_2Mg aggregate flow with the nitrogen carrier was 100sccm at 350torr and 40C bubbler conditions. It was roughly calculated in [14] using the thermodynamic conditions of the bottle that the flow of magnesium at 100sccm aggregate flow is about 0.00453sccm. This was done using an ideal gas assumption to calculate the partial pressure for nitrogen and an empirical partial pressure equation for Cp_2Mg. Finally, the flow rate of Cp_2Mg was found by using an ideal gas assumption on the aggregate flow, ie. the 100sccm aggregate flow rate was multiplied by the ratio of Cp_2Mg partial pressure to the sum of N_2 and Cp_2Mg partial pressures. Note that the aggregate pressure was used as an approximation to the N_2 partial pressure to calculate Cp_2Mg flow; and that there is a diminishing return on Cp_2Mg flow with increasing N_2 flow.

After 1500 film growth cycles (for a film thickness ~140nm), the growth is stopped and a 1 hour annealing is performed at 600C (the maximum temperature of the heater) in attempt to activate the magnesium donors. Once this is done, cooling begins.

Because cracks develop in GaN films when the heater power is simply shut off, the power to the heater is reduced periodically in increments to allow for a slower cooling. The time period at the beginning of cooling is the most important because – due to the decaying exponential nature of cooling - it tends to give the fastest cooling rate and the largest change in film-substrate mismatch per degree Celsius of cooling. Hence, the power to the heater is decreased slowly at first, with the temperature increments increasing as the cooling progresses. N_2 (non plasma) is passed over the sample during the cooling process in order to maintain a sufficient chamber

pressure so that thermal desorption of the film species does not occur. Oxide etching is usually not required on samples because they are cooled under nitrogen pressure and removed at room temperature.

Optimization of the Recipe

The preceding section describes the growth recipe for the best result obtained in a batch of 10 p-type GaN growths performed on sapphire. In this series, first the aggregate flow of Cp_2Mg (in N_2 carrier gas) was varied between growths from 100sccm to 150sccm to 200sccm to check for the point at which magnesium self-compensation began. While Hall measurements generally did not show a decreasing electron concentration with increasing Cp_2Mg flow, the resistivity increased due to magnesium-related stresses, and films quickly became difficult to measure. Thus, 100sccm was deemed the best flow for growths to mitigate resistivity effects on measurements. While there was no way to pinpoint the reason why Hall carrier concentrations didn't change with increased Mg doping, two theories are probable. The first is that the solubility limit of magnesium in the gallium droplets had been reached at 100sccm; ~$5X10^{19}cm^{-3}$ in GaN grown at atmospheric pressure and ~1000C (this would result in a hole concentration of ~$10^{17}cm^3$ at room temperature – 1 in 30 donors - due to thermal limits on donor activity) [14]. The second is that the magnesium atoms are self-compensating in this span of dopant concentrations. Hence, any thermally activated magnesium atoms added would simply be donated an electron by other magnesium atoms added. Due to the increasing resistivity of the sample with dopant flow rate, the second theory seems more plausible. This is still to be verified by mass spectroscopy measurements.

The second step in the growth sequence was to determine the optimal buffer layer thickness. Three identical undoped GaN growths were performed on top of 35, 50, and 80 pulse cycles of (amorphous) AlN; corresponding roughly to thicknesses of 5nm, 7nm, and 11nm. The 7nm thickness was chosen first as per the 8.9nm thickness given by Amano et al. [36], and was

the final thickness chosen for p-type GaN growths because it gave the best qualitative XRD measurements for GaN.

The final step was to optimize the stoichiometry of the gallium and nitrogen in the film growth. Growth rate of the best pGaN film was calculated to be 89.3 nm/hour (~0.13nm/cycle); and was calculated from the number of cycles multiplied by the time per cycle. Hence, this means the growth progresses at one monatomic layer per cycle. Note that this growth rate did not include the 1 hour heating or the 3 hour activation and cooling period.

Since with this recipe the growth rate progresses at about a monolayer each cycle, the growth per cycle was not tampered with. Instead, the plasma flows, metal flows, and pulse-in times were varied while attempting to keep the growth per cycle constant. Using combinational variations that represented only a slight deviation away from the film growth rates (growth per unit time) used for the best growths of pure GaN films with the system, a four-fold increase in crystal quality was achieved over previously grown pGaN films without AlN buffer layers. Since SEM pictures of film thickness could not be taken on many films, the reliance of growth rate on these parameters could not be studied in detail.

Longer post-growth anneal times (up to 3 hours) and longer pre-growth heating times (up to 4 hours) were tried as well, but didn't provide any noticeable change from films grown without such additions.

XRD

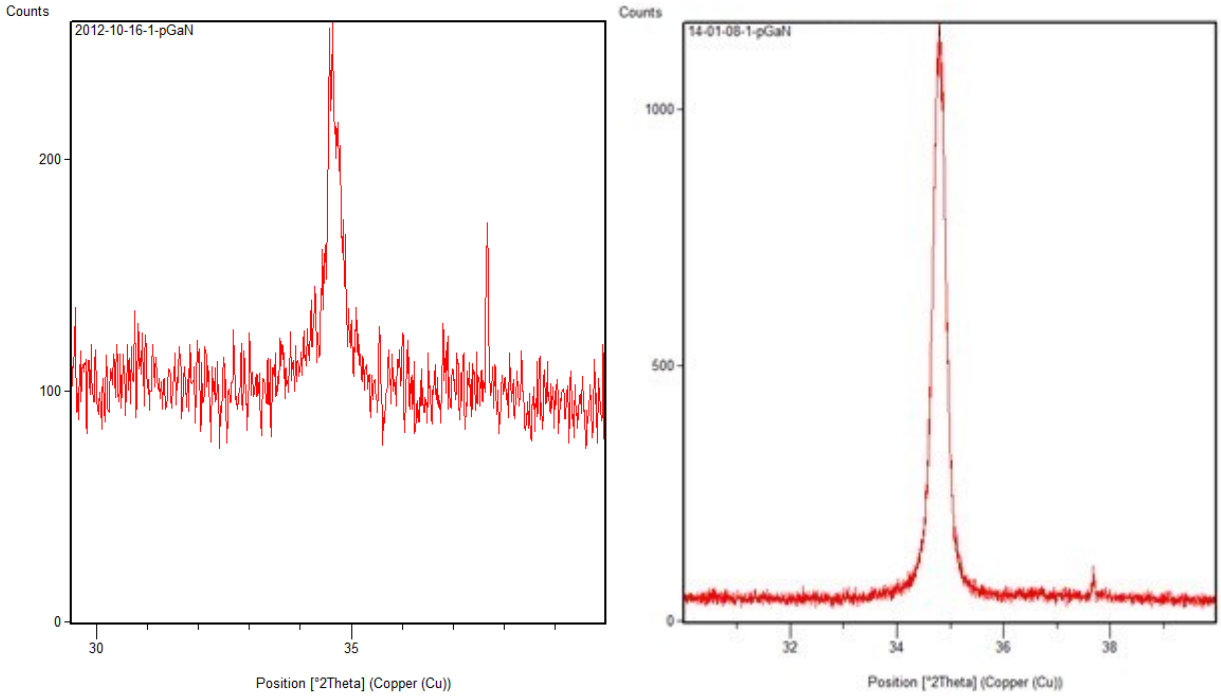

Figure 15 pGaN grown with the Lakehead RPECVD system. To the left is the best pGaN grown before using a buffer layer and pulsed growth regime, and to the left is the pGaN grown using these techniques. The difference in peak intensity differs by a fact of about 4.

As can be seen in Figure 15, the use of a buffer layer with pulsed growth techniques improved p-type GaN XRD measurements considerably. FWHM comparison studies are not particularly insightful when the character of the tail skirts (the bottom of the peak) are so different, but it should be mentioned that from observation the FWHM of the right-hand picture is considerably smaller than the left. The most prominent difference between the peaks is in the bottoms. While the left-hand figure shows clear signs of adverse lattice strain, the right-hand figure rounds out very sharply near the bottom. This may be indication of the reduction in crystal stresses by the aluminum nitride layer, or perhaps by reduced incorporation of contaminant hydrogen and carbon species from the MO precursors. It should be noted that the chamber was thoroughly flushed of contaminants like indium using nitrogen plasma before growth, so interspecies mixing to create InGaN, InN, or InO phases is not a problem. The effectiveness of such plasma flushes have been verified for other growths by EDX.

Hall Effect

Measurements by Hall effect were unsuccessful because the instrument, while operating at the maximum of its I*V product (maximum power output) simply could not make readings above the noise floor. This indicates a high material resistivity with a low electron concentration (the two values are closely related, see 1.3.5). This is actually a promising sign for successful p-type doping; since the expected mobility of holes is very low. For example, while the electron mobility of GaN for a good crystal quality is expected to be a little over 400cm^2/Vs (reported widely), typical readings for moderately oxygen-doped n-type growths done by this system in a similar way to the p-type GaN shown here are ~10-50 cm^2/Vs with 10 cm^2/Vs being more typical. Hole mobility has been reported at around 10 cm^2/Vs for the high quality pGaN growths under the conditions described above in 4.2. Thus, the expected hole mobility for growths done in this work is expected to be considerably less than 1 cm^2/Vs (on account of Mg induced strain), which is the threshold of measurement for the Hall system used. The fact that the measurements did not yield a measurement of net electron concentration like other pGaN growths may indicate that the acceptor concentration is sufficient to quell any background electron species present from oxygen contamination or the natural mechanisms that create the electron background seen in pure GaN [16].

Since XRD scans do not properly represent defect density, it is possible for the above intense and narrow peaks to be obtained while having an oxygen-codoped film. Since oxygen codoping has been reported to both increase the Hall-measured hole concentration and decrease sample resistivity [19], we must account for the high resistivity Hall measurements either by asserting that the hole mobility is very low, the defect density is very high, or both. Thus, it may be safe to argue that the films have a high hole concentration due to a combination of codoping, reduced film stress from the AlN buffer layer, and reduced incorporation of hydrogen due to

species pulsing; and a high resistivity due to the low hole mobility and defects from the growth process and the magnesium doping.

Transmission and Bandgap

The transmission measurements taken of this low electron concentration, high crystal quality pGaN growth is shown in Figure 16. A very gradual rolloff at the band edge may be accounted for by stress states created by the magnesium or a combination of magnesium and oxygen. However, it may just be due to effects from the AlN buffer layer, ie. interface and scattering effects. On the other hand, this anomalous gradual rolloff may be an indication that the conduction band edge states (including the defect-created states) have been thoroughly depleted by magnesium acceptors. This would be the first time that this effect has been observed in GaN growths by this technique; since until now only growths have been done with an electron background populating these states. Due to the film's surface morphology and higher crystallinity, it seems unlikely that this is merely the result of metal-rich conditions.

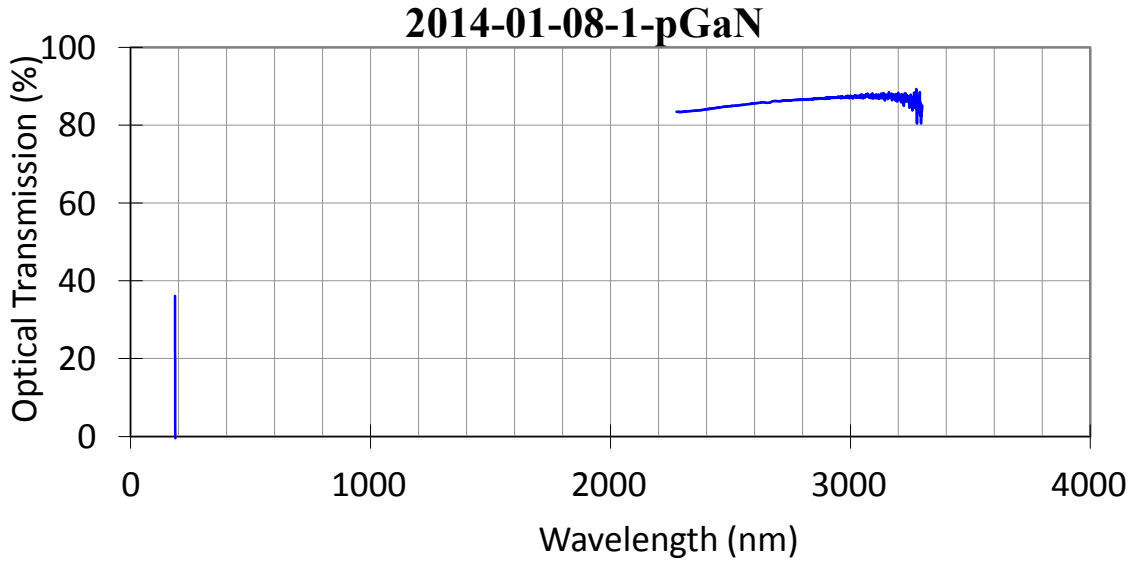

Figure 16 Optical transmission measurement of low codoping, high resistivity pGaN film with high crystal quality grown on an AlN buffer layer. A sapphire baseline with a similarly grown AlN layer was used.

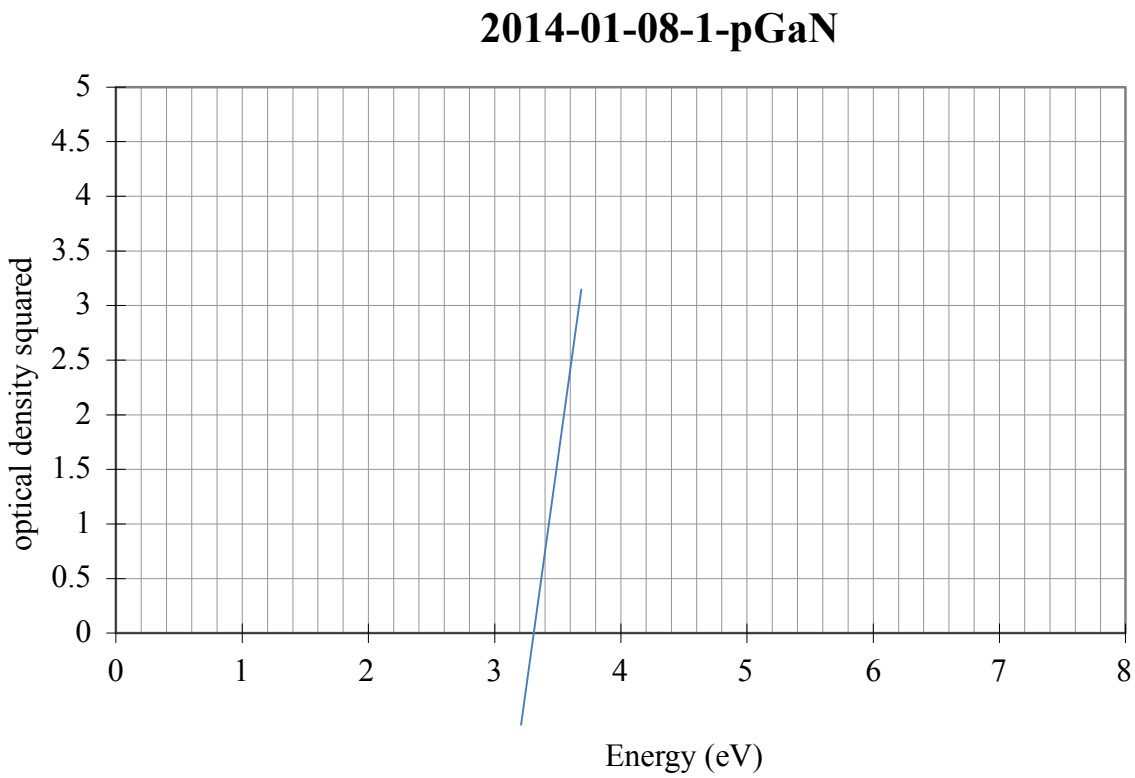

Figure 17 Plot of optical density squared (ODS) vs. photon energy for the high crystal quality, low electron concentration pGaN growth. Using the method described in the optical transmission theory, a line is traced back from the exponential section to the x axis to find a bandgap of 3.3eV.

For the Figure 17, the proper region of the optical density squared (ODS) is identified by taking the threshold of the linear region as the starting data point when the logarithm of the y-axis is taken. Linear interpolation through the linear region is then performed to the x-axis and this value is taken to be the material bandgap. The value obtained in this case was around 3.0eV; comparable to the 3.4eV value reported for high quality GaN and consistent with the lower values that are usually obtained due to sub-conduction band defect states. However, this value is still slightly lower than typical measured values for growths done with this RPECVD system (which are usually ~3.2eV); and the vacancies near and below the band edge created by magnesium acceptors could account for this.

The nonlinear regions correspond to where the exponent of the absorption coefficient of the light is different from the coefficient for a pure direct optical transition (defect states, nonradiative recombination). The other, straight line on the figure is to be ignored as it is simply an artifact of the calculation and graphing process.

AFM

AFM imaging of the film (Figure 18) reveals the formation of droplets on the surface. The topography image shows an accurate representation of the vertical sizes of the features, while the phase contrast scan doesn't. However, the phase contrast scan shows the contours of the film with more clarity; giving a more thorough overall impression of the features present on the film surface. Streaking can be seen on the either side of the droplets in the phase contrast image as the vibrations from the impulse to the probe from these droplets takes time to settle. Slowing the scan speed can mitigate this problem. In general, the severity of the streak is correlated to the size of the droplet and to the imperfection of the droplet shape.

The next two images display line scans in a line across the film according to the position of the arrow on the topography and phase scan images. Besides the droplets, the results on the surface of the film are indeed impressive – the right hand line scan clearly shows that the smoothness between the droplets is on the scale of a single GaN wurtzite unit cell. The real surface is expected to be even smoother due to artifacts from probe vibration and clinging atomic surface contaminants from the atmosphere. The smaller (non-droplet) bumps seen across the surface are in fact droplets in the previous layer of film growth. It can be seen from the phase contrast image that these droplets have in fact coalesced together very nicely. This is a testament to the efficacy of the metal-migration technique of growth and to molecular nitrogen plasma as a viable growth species.

The reason for the droplets on the surface is actually because the growth pattern loops from plasma pulsing to metal pulsing – in that order. In the last cycle, there was no subsequent

plasma delivery to the surface after pulsing. Hence, the droplets seen on the surface are the droplets that were deposited in the last cycle, but managed to accommodate into the surface using some of the residual heat and nitrogen on the film, and hence did not get evacuated by the turbomolecular pump. A few minutes of dilute (1:10) HCl Etching can be used to rid of these droplets and smooth out the droplets of the layer beneath.

A future prospect for film growths to increase the hole density in pGaN growths comes from [43], and involves utilizing magnesium flows during nitrogen pulsing to increase final Mg incorporation into the lattice. This way, the increased solubility of magnesium in nitrogen (acting as fillers for gallium vacancies) allows a large amount of magnesium to be held on the surface of a gallium droplet. This process works better for incorporating Mg into regions of good stoichiometry, rather than incorporating it into metal rich regions of the film through a solution of magnesium in gallium; which is what tends to happen when simultaneous gallium-magnesium pulses are used. Magnesium atoms would tend to be in nitrogen rich regions of the film instead - not as easily compensated by gallium anion atoms in metal rich regions.

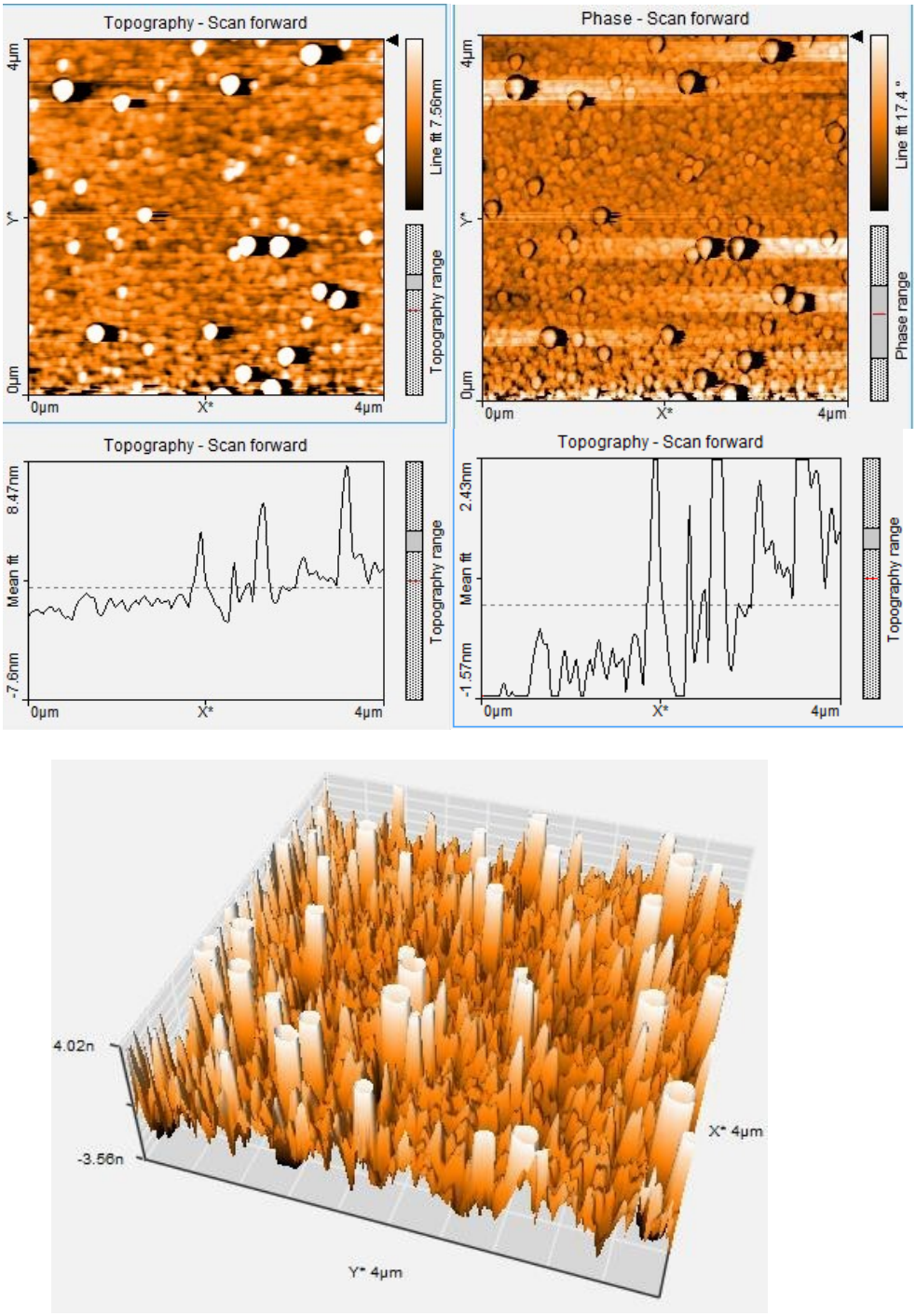

Figure 18 AFM images of sample 14-01-08-1-pGaN showing good crystal quality and low electron concentration. The top two images are topographical and tip-phase images, while the middle two are the same line scan showing droplet heights and smoothnesses between the droplets. Finally, a 3D topography image is shown of the surface.

4.3 GaN On Sapphire

Two GaN film growths are compared in this section. One was the best growth in a series of growths of varying stoichiometry, grown upon sapphire. The second growth was the best in series that used the aforementioned growth as a starting point, for growth upon aluminum nitride buffer layers. Three buffer layer thicknesses (5nm, 7nm, 11nm) were tried, but the 7nm gave the best results. Hence, GaN stoichiometry was varied on 7nm AlN growths, and the best result is compared in this section to the original GaN recipe as it was grown on bare (nitride) sapphire.

4.3.1 Results

Despite obvious pitting in AFM scans, the GaN film grown with a buffer layer clearly shows superiority to the growth without - a nearly fourfold increase in the crystal quality as measured by XRD. The FWHM value for the unbuffered GaN was 0.39, while for the buffered GaN it was 0.36. It is expected that these pits can be reduced by further experimentation with buffer layer parameters; as it was reported [36] that buffer layers too thick would result in deep cracks and pits in the GaN film. This is perhaps due to the excessive buildup of stress in the film; analogous to how stress buildup in tectonic plates causes a large quake on relief. Buildup of stress with a buffer layer would occur because relieving factors for stresses (like thermal gradients) would be quelled and evened out more quickly by the buffer layer; and a sort of strain absorbance would be expected. The layer would act as a spring to build up strain-related energy. A thinner buffer layer may allow a less damaging route of stress relief by allowing relief to occur more often. Note that the reference cited used GaN growths on the micrometer scale, while in this work the film thicknesses are only around 150nm. Buffer layer thickness used was ~7nm. Thus, some appropriate scaling of the buffer layer thickness may be necessary in future work.

Also, it is unclear whether the origin of the pitting in the GaN films on AlN is the result of stress buildup in the film on a proper buffer layer, if it is the result of a defective buffer layer. As the AFM scans of the buffer layer alone revealed pitting, the second scenario is the most probable; however it is expected that the nitridation of the buffer layer prior to growth should mitigate the transference of defects to the seeding GaN layer.

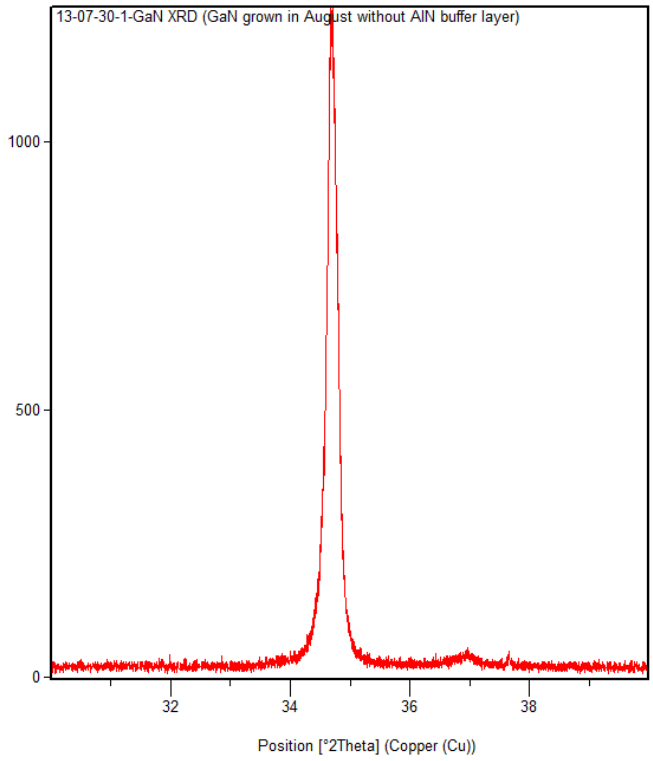

Figure 19 GaN grown without an AlN buffer layer

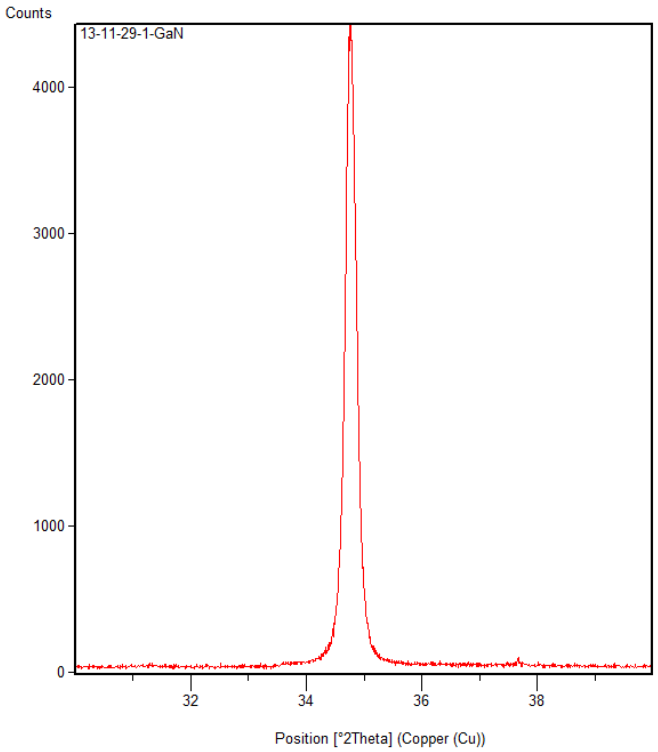

Figure 20 XRD scan of GaN grown with an AlN buffer layer. Other than the addition of a buffer layer, growth was the same as 13-07-30-1-GaN. FWHM was around 2θ=0.31

RMS surface roughness in the non-pitted regions of the GaN growth with buffer layer was around 1.0nm; while roughness for the growth without the buffer layer was 3.09nm. A notable difference between the surface morphology of the two regions is observed. The buffer layer GaN surface seems to consist of a fine collection nitrided droplets, while the non-buffer GaN seems consist of larger, better coalesced droplets. However, inspection of the topographical line-scan indicates that the comparative width of peaks is approximately the same. In fact, the buffer-layer GaN shows some droplets of almost identical dimension to the non-buffered GaN, but the buffered GaN consists mainly of droplets about half the height. Also, the buffered GaN seems much less likely to develop the larger droplets of ~9nm height seen in the unbuffered GaN. Hence, it can be observed that the buffered GaN allows better droplet

formation and movement across the surface, but seemed to hamper the coalescence of these droplets as compared to the unbuffered GaN.

When measured by Hall effect, both samples showed similar mobilities of ~5cm^2/Vs, with measurement sensitivity giving about ±3 cm^2/Vs from this value in this typical case. Good accuracies could not be obtained because the Hall effect measurement system was operating near its output limit in order to fight the noise floor. The unbuffered sample tended toward a mobility of 2 cm^2/Vs, while the buffered sample tended toward 5 cm^2/Vs on average. Over five measurements were averaged out for each film. However, the resistivity of the buffered sample was 5.1 Ωcm, while the resistivity of the unbuffered sample was an order of magnitude lower, averaging 6.1X10^{-1}Ωcm. Resistivities give very little variance between measurements. It is likely that oxygen contamination is the cause of these measurement differences, as the chamber pressure before growth for the unbuffered sample was slightly higher than the chamber pressure for the buffered sample. Measured background concentration was typically in the range of $10^{17}cm^{-3}$ for both samples, but the unbuffered sample sometimes gave measurements into the $10^{18}cm^{-3}$ range. Values of $10^{17}cm^{-3}$ are typical of undoped GaN background concentrations; however measurements in the range of $10^{18}cm^{-3}$ should require a background of oxygen to achieve. Hence, the unbuffered sample gave lower values of resistivity and higher carrier concentrations due to increased oxygen donor contamination, while the buffered sample gave better overall carrier mobility due to the increase in crystal quality.

Transmission measurements showed a prominent Berstein-Moss effect both in the shift in the bandgap tail and in the rolloff of transmission in higher wavelengths in the case of the unbuffered GaN. This reflects the fact that the oxygen contamination increased the conduction band electron concentration. The Urbach tails seem to be similar, which seems an unlikely outcome given the difference in oxygen doping. This may be accounted for by the difference in

crystal quality. Both films show a peak with exponential rolloff after the bandgap rolloff as per the nature of inter-conduction-band absorption. No inter-cavity reflection is observed.

Hence, the use of a thin amorphous layer of AlN for growth of nitride thin films is shown to be a viable technique for improving the crystal quality. The buffer layer seemed to improve surface smoothness of the film, indicating a more planar growth than unbuffered films. It is probable that with tuning of the growth speed and stoichiometry, the coalescence of droplets can be improved for these films, and resultant electrical properties should be notably superior to unbuffered films.

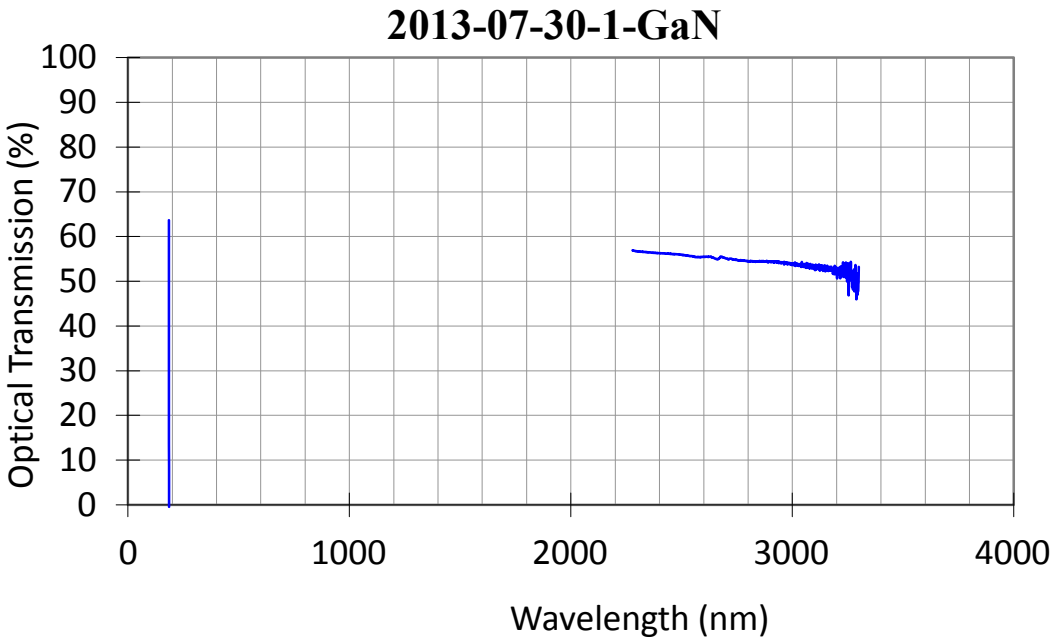

Figure 21 Optical transmission of 2013-07-30-1-GaN (unbuffered), showing a prominent Urbach characteristic tail, conduction band absorption peak and rolloff, and Berstein-Moss effect.

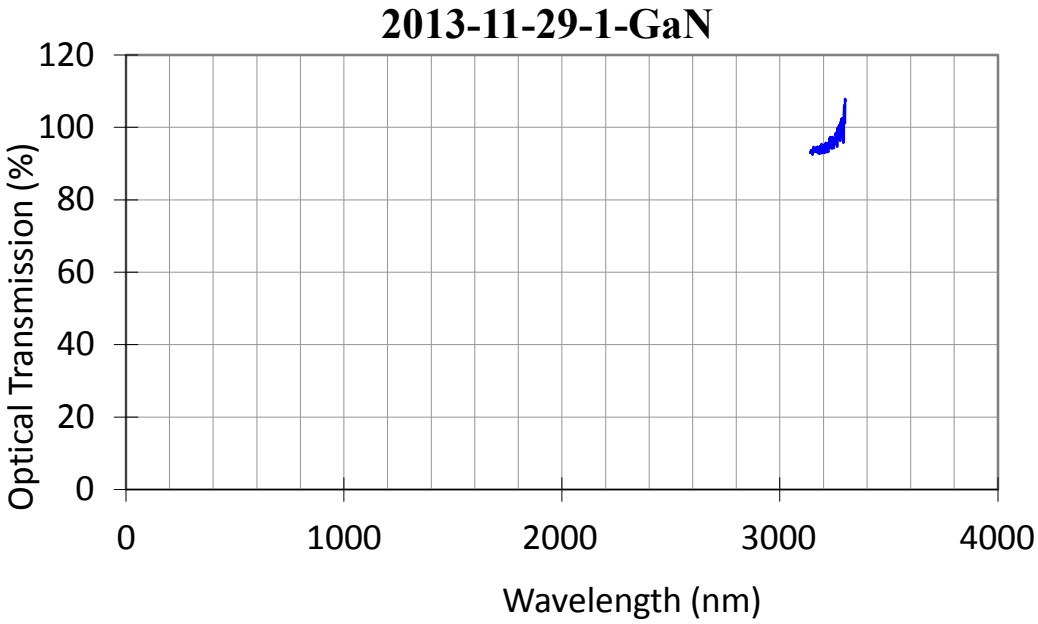

Figure 22 Optical transmission of 2013-07-30-1-GaN (buffered), showing an Urbach characteristic tail, cavity resonance peak, and an anomalous increase in transmission with wavelength – likely due to the buffer layer..

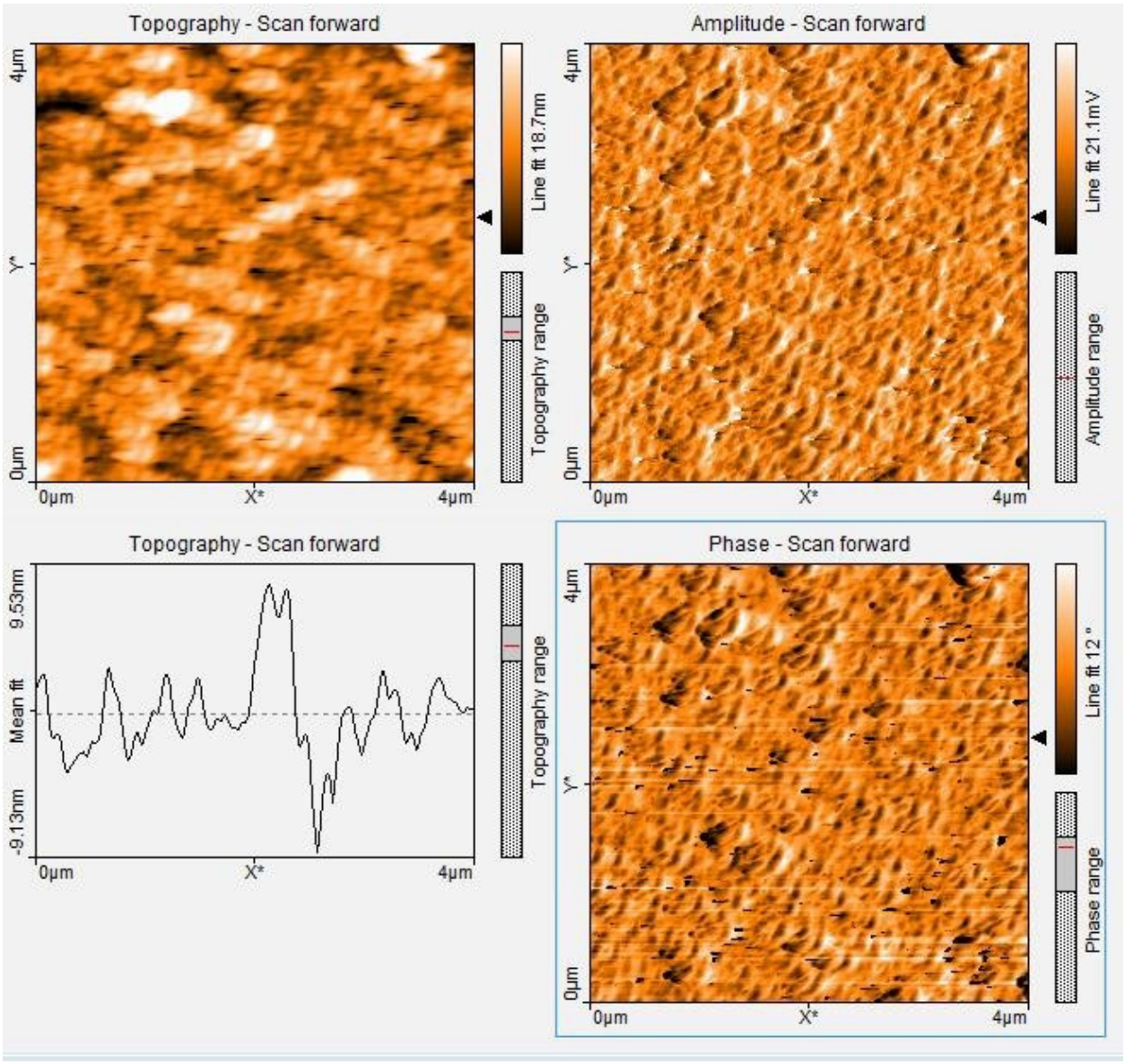

Figure 23 AFM of GaN grown without a buffer layer. RMS surface roughness was 3.09nm

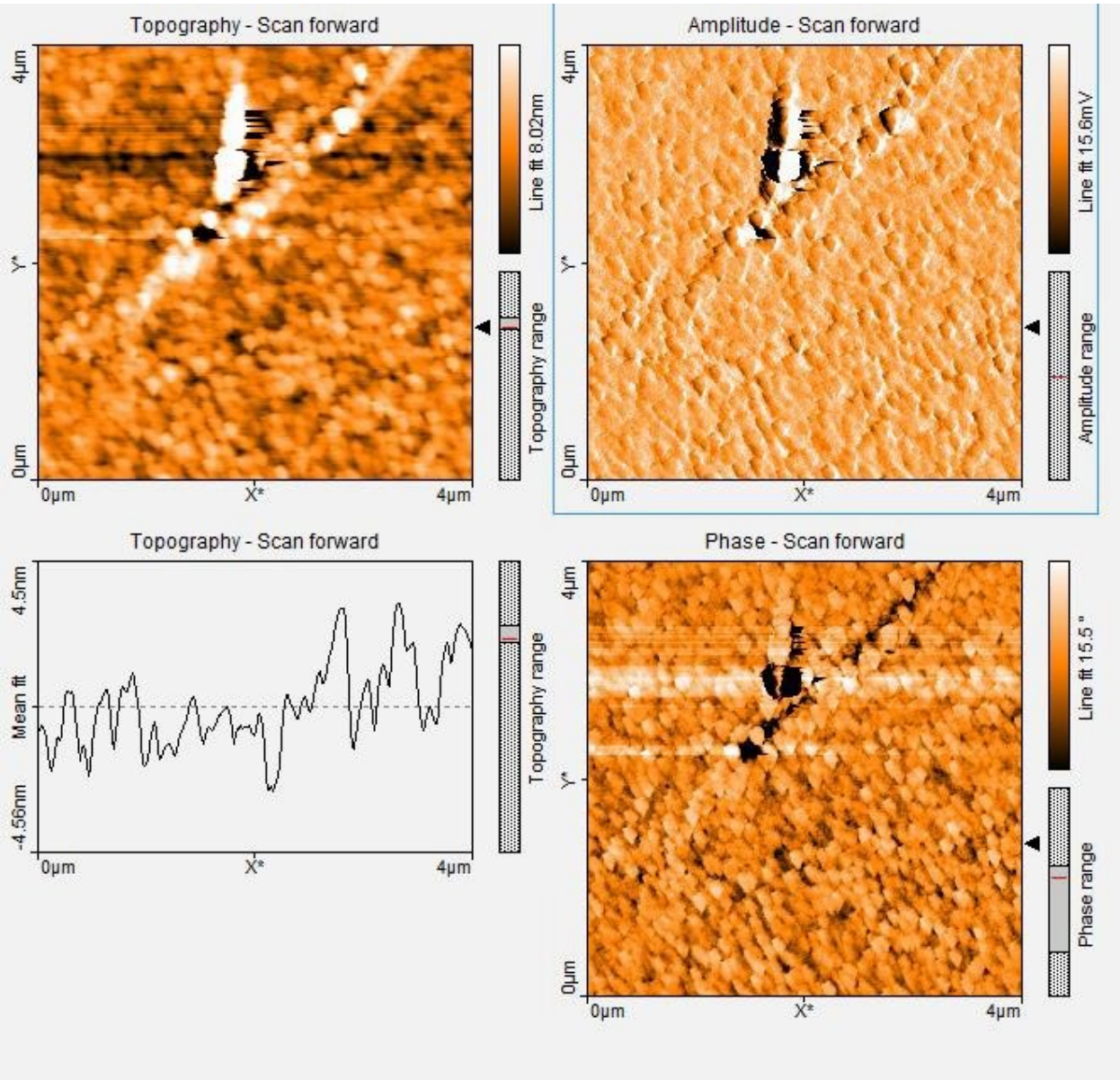

Figure 24 13-11-29-1-GaN, a repeat of 13-07-30-1-GaN grown on an AlN buffer layer. The AFM image was taken to show one of the pits. It is not a crack, but it fact is caused by the underlying buffer layer. It's unclear if it is seeded in the GaN from layer effects, or if it is the result of growth on a buffer layer defect. The pit was around 35nm deep in the worst region. The line scan shows an RMS roughness of about 1nm in the nicer regions.

5 Calculation of Electron Band Structure of Semiconductor Compound Alloys

5.1 Introduction

The semiconductor compound alloys based on binary compounds InN, GaN and AlN are very important in semiconductor physics and in design of optical semiconductor devices. Because of this, these three binary alloys have been the subject of much investigation [47 – 52] over the last few decades. The theoretical investigations are closely connected with calculations of the electronic band structures of InN, GaN, AlN [47] [48] [53 - 56]; and the obtained results have been connected successfully with the properties of these binary alloys found experimentally. The relatively recent observation of 0.7 eV photoluminescence for InN - and of absorption features near this energy - have been the basis of a number of papers [57] [58].

The exact band gap of indium nitride is still unclear. It has been proposed that these low energy features indeed indicate a 0.7 eV band-gap. However, the material had long been held to have a much higher band-gap of 1.9 eV. The phenomena accounting for the large differences between these values is not presently understood. The Burstein-Moss effect does not explain the variation in energies seen in materials of low carrier concentration, and Vegard's law[1] indicates that oxygen levels in the higher band-gap materials are insufficient to account for the difference.

Sample inhomogeneity offers a strong possible explanation. Typical nitrides growth conditions are known to form non-stoichiometric indium nitride. In light of this, one would now seek to investigate the possibility that these low energy features arise as growth artifacts due to

[1] A technique to find the approximate lattice constant of an alloy by linearly interpolating between the lattice constants of the individual unnalloyed elements based on the mole fractions of the alloying elements. Can also be used for bandgap energies of alloyed semiconductors.

an alloy formed by non-stoichiometric InN containing both single In substitutions on N sites and vice versa.

5.2 Electron Band Structure of $A_xB_{1-x}C$

(A and B are cationic atoms, and C is an anionic atom)

Mathematical formalism

The multinary crystal $A_xB_{1-x}C$ is considered to be a periodic crystal having as a basis large primitive super-cells which satisfy certain conditions of spatial periodicity. A partial case is the multinary crystal containing only one large primitive super-cell. It is considered that this primitive super-cell contains a finite number of the familiar brand of elementary cells. By definition, an elementary cell of the primitive super-cell is a cell having the same translational symmetry as of the primitive cells of the crystals' constituents AC and BC; if the symmetry is defined only on the basis of the positions of the cations and the anions without concerning oneself with the actual species of the atoms (in binary crystals the elementary cell is equivalent to the primitive cell).

The number of orbitals in an elementary cell is considered to be equal to *n*. For the sake of simplicity the number of elementary cells in the primitive super-cell in the direction of each crystal axis is considered to be an odd number and so $N_1 = N_2 = N_3 = 2N + 1$, where N is a positive integer. This means that, if one wants to decompose the bandstructure transitions between AC and BC by intermittent primitive super-cells of $A_xB_{1-x}C$ of intervallic stoichiometry, then these primitive super-cells may only have an odd number of constituent elementary cells. Only the interactions between orbitals belonging to a single primitive super-cell of the multinary crystal will be taken into consideration. In the one-electron approximation of the lattice, the electron wave function of the crystal can be expressed by LCAO terms:

$$\Psi_{\mathbf{p}}(\mathbf{r}) = \sum_{\upsilon} \sum_{\mathbf{q}} \sum_{g} C(\mathbf{p})_{\upsilon,\mathbf{q},g}\, \chi^{\upsilon\mathbf{q}}{}_{g} \qquad (48)$$

Where:

$\mathbf{p} = (p_1, p_2, p_3)$ and $\mathbf{q} = (q_1, q_2, q_3)$

Integers p_j and q_j (j = 1, 2, 3) run over $-N$, …..,0, ……,N

$\mathbf{r}$ is radius-vector of the electron,

$\sum_{\mathbf{q}}$ is a shorthand notation for $\sum^{N}_{q1=-N} \sum^{N}_{q2=-N} \sum^{N}_{q3=-N}$

$\sum_{U}$ is the sum over all primitive super-cells of the crystal

$\sum_{g}$ is the sum over all n orbitals belonging to the elementary cell

$\chi^{U\mathbf{q}}{}_{g} = \chi_g(\mathbf{r} - \mathbf{R_q} - \mathbf{r}_g)$ is the g^{th} atomic orbital (corresponding to an atom having position vector $\mathbf{r}_g$) in the elementary cell characterized by vector $\mathbf{R_q} = q_1\mathbf{a}_1 + q_2\mathbf{a}_2 + q_3\mathbf{a}_3$, where $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ are three basis vectors of the primitive super-cell defined by the elementary cell.

Because only the interactions between orbitals belonging to a single primitive super-cell will be considered, electron orbital $\varphi_{\mathbf{p}}(\mathbf{r})$ (belonging only to the primitive super-cell) will be used. In the one-electron approximation (applied strictly within the primitive super-cell) the orbital $\varphi_{\mathbf{p}}(\mathbf{r})$ becomes

$$\varphi_{\mathbf{p}}(\mathbf{r}) = \sum_{\mathbf{q}} \sum_{g} C(\mathbf{p})_{q,g}\, \chi^{\mathbf{q}}{}_{g} \tag{49}$$

And the expectation value of the electron energy is canonically

$$<\varepsilon>(\mathbf{p}) = < \varphi_{\mathbf{p}} \mid \hat{H} \mid \varphi_{\mathbf{p}} > / < \varphi_{\mathbf{p}} \mid \varphi_{\mathbf{p}} > \tag{50}$$

Where $\hat{H}$ is the Hamiltonian. The Ritz variational method can be performed over (50) for the coefficients in (49) and the result is the matrix equation for the primitive super-cell

$$\mathbf{H}\,\mathbf{C}(\mathbf{p}) = \varepsilon(\mathbf{p})\,\mathbf{S}\,\mathbf{C}(\mathbf{p}) \tag{51}$$

The hyper-matrix **H** of dimension $n(2N + 1)^3$ (n is the number of elementary cell orbitals) has sub-matrices $\mathbf{H}_{\mathbf{p},\mathbf{q}}$ of dimension n and the overlap hyper-matrix **S** has sub-matrices $\mathbf{S}_{\mathbf{p},\mathbf{q}}$. In conformity with the translational symmetry of the primitive super-cell (and the Born - von Karman boundary conditions), matrices **H** and **S** are cyclic hyper-matrices (each row is a cyclic permutation of the vector in row 1) and it is always possible to find a unitary transformation (a matrix that is a purely rotational isometry) which turns these hyper-matrices into diagonal hyper-matrices. The transformations of both matrices (**H** and **S)** are similar and only those concerning **H** will be given. The unitary matrix **U** of dimension $n(2N + 1)^3$ having $n \times n$ dimensional hyper-matrix blocks is defined as

$$\mathbf{U}_{\mathbf{p},\mathbf{q}} = (2N + 1)^{-3/2} \exp[(2\pi i\mathbf{p}.\mathbf{q})/(2N + 1)]\,\mathbf{1} \tag{52}$$

will transform **H** and **S** in block-diagonal forms (**1** is unit matrix having dimension $n \times n$):

$$\mathbf{H}' = \mathbf{U}^{+}\,\mathbf{H}\,\mathbf{U} \tag{53}$$

From the definition of (52) we can see that it can be shown that $\mathbf{U}^{+}\,\mathbf{U} = \mathbf{I}$, where the unit matrix **I** has dimension $n(2N + 1)^3$. The **p**th diagonal block of the block-diagonal matrix (53) has the form

$$\mathbf{H}'(\mathbf{p}) = \sum_{\mathbf{q}} \exp[(2\pi i\mathbf{p}.\mathbf{q})/(2N + 1)]\,\mathbf{H}(\mathbf{q}) \tag{54}$$

where matrix **H**(**q**) is sub-matrix of the matrix **H** before the unitary transformation. This sub-matrix corresponds to the elementary cell having radius-vector $\mathbf{R}_{\mathbf{q}}$. The equation (51) can be multiplied on the left by $\mathbf{U}^{+}$ and the unit matrix can be inserted in both sides of the same equation, yielding,

$$\mathbf{U}^{+}\,\mathbf{H}\,\mathbf{U}\,\mathbf{U}^{+}\,\mathbf{C}(\mathbf{p}) = \varepsilon(\mathbf{p})\,\mathbf{U}^{+}\,\mathbf{S}\,\mathbf{U}\,\mathbf{U}^{+}\,\mathbf{C}(\mathbf{p}) \tag{55}$$

Next, to express things more compactly, matrix **G**(**p**) can be defined as $\mathbf{U}^{+}\,\mathbf{C}(\mathbf{p}) = \mathbf{G}(\mathbf{p})$. Equation (55) can thus be expressed by **G**(**p**) and by equation (6) in following way

$$\mathbf{H}'\,\mathbf{G}(\mathbf{p}) = \varepsilon(\mathbf{p})\,\mathbf{S}'\,\mathbf{G}(\mathbf{p}) \tag{56}$$

However, **H**' and **S**' are block-diagonal (a diagonal matrix of blocks) and so equation (56) can be decomposed into simpler equations

$$\mathbf{H}'(\mathbf{p})\,\mathbf{g}(\mathbf{p}) = \varepsilon(\mathbf{p})\,\mathbf{S}'(\mathbf{p})\,\mathbf{g}(\mathbf{p}) \tag{57}$$

Where **H**'(**p**) was given by (54).

The primitive super-cell of the multinary crystal is very large and it can be seen that the number $(2N+1)^3$ of elementary cells belonging to a primitive super-cell is very big. The electron wave vector can be defined as

$$\mathbf{k} = k_1\,\mathbf{b}_1 + k_2\,\mathbf{b}_2 + k_3\,\mathbf{b}_3 \tag{58}$$

where $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are the basis vectors of the reciprocal elementary lattice and

$$k_j = (2\pi\, p_j) \,/\, [a_j\,(2N + 1)] \tag{59}$$

Again, (j = 1, 2, 3). The equation (57) can be rewritten in the following way

$$\mathbf{H}(\mathbf{k})\,\mathbf{g}(\mathbf{k}) = \varepsilon(\mathbf{k})\,\mathbf{S}(\mathbf{k})\,\mathbf{g}(\mathbf{k}) \tag{60}$$

Where

$$\mathbf{H}(\mathbf{k}) = \sum_{\mathbf{q}} \exp(i\mathbf{k}.\mathbf{R}_{\mathbf{q}})\,\mathbf{H}(\mathbf{q}) \tag{61}$$

Hence, the electron band structure of the multinary crystal can be found from (60) as eigenvalues of the matrix **H**(**k**). According to the transformation to (61), every energy eigenvalue of **H**(**k**) contains the corresponding eigenvalues of the matrices **H**(**q**). The corresponding electron energy is

$$E(\mathbf{k}) = \sum_{\mathbf{q}} \exp(i\mathbf{k}.\mathbf{R}_{\mathbf{q}})\, E(\mathbf{q}) \qquad (62)$$

The expression (62) shows that the electron energy in the multinary crystal $E(\mathbf{k})$ can be defined in terms of a set of energies $E(\mathbf{q})$. Each element $E(\mathbf{q})$ of this set is the electron energy in the elementary cell having radius-vector $\mathbf{R}_{\mathbf{q}}$.

Next, a Fourier transform will be applied on both sides of (62) and the dependence of the electron energy on the spatial coordinates will be found.

$$E(\mathbf{r}) = F\,[E(\mathbf{k})] = \sum_{\mathbf{q}} \delta\,(\mathbf{r} - \mathbf{R}_{\mathbf{q}})\, E(\mathbf{q}) \qquad (63)$$

where $\delta\,(\mathbf{r} - \mathbf{R}_{\mathbf{q}})$ is delta-function. The expression (63) shows that the electron energy depends on the electron radius vector, and it depends on the elementary cell (the elementary cells are all equivalent)

The electron band structure of the multinary crystal can be determined using interactions within the primitive super-cell, which determine the corresponding sub-bands. From here the quasi-elementary cell will be introduced. It has the same physical structure as the elementary cell, but the specific identities of the constituent atoms are considered (here they are In, Ga and N). According to expressions (62) and (63), the total electron band structure is the aggregate sum of electron band structures calculated for every elementary cell. This fact will be applied likewise for the quasi-elementary cells; because the orbital interactions within this cell have approximately 93% of the total contributions in electron energy, and these contributions will increase if the interactions between second-neighbor orbitals are taken into consideration.

The electron band structure of the multinary crystal contains the same sub-bands like those determined for the primitive super-cells of the multinary crystal without consideration for the localizations of the interactions. However, here the corresponding sub-bands are calculated

localized in the corresponding quasi-elementary cells and the electron band structure of the crystal depends on both **k** and $\mathbf{R_q}$. This is in conformity with (62) and (63) if both expressions are applied in approximation of the quasi-elementary cell. The electron band structure E(**k**, $\mathbf{R_q}$) of the multinary crystal determined in this way is a discrete sequence of electron band (sub)structures; and each of them is calculated with the corresponding quasi-elementary cell.

5.2.1 Electron Band Structure of $In_xGa_{1-x}N$ with Defects

The defects in $In_xGa_{1-x}N$ corresponding to the formation of an $In_xGa_{1-x}A_yN_{1-y}$ alloy - whose properties depend on A content – have been modeled. A is a species of atom different than both Ga and N that substitutes on N sites. Segregated A species within the interior of wurtzite GaN are not considered; only true alloy species (substitutional) with 'A' substituting on the nitrogen site are. For this arrangement, a new quaternary semiconductor $In_xGa_{1-x}A_yN_{1-y}$ with $0 < y < 1$ is presented. It has four binary constituents – InN, InA, GaN and GaA. Thus we have introduced new primitive cells of binary constituents GaA and InA; satisfying the following conditions:

a) the A atom substitutes for a N atom saving the tetrahedral symmetry of the crystal cell

b) *n* of the valence electrons of the isolated A atom occupy the valence *s*-orbital, and the other *m* valence electrons occupy the valence *p*-orbital. The *s*- and *p*- orbitals form sp^3 hybrid orbitals.

Some of the electrons of the above orbitals can come from other orbitals ϑ (where ϑ = *d*-, *f*-…) belonging to the same atom. The charge transfer ΔQ from the ϑ orbital of the A atom to the $4sp^3$ hybrid orbital of the Ga atom, and to the $5sp^3$ hybrid orbital of the In atom, is calculated by using the following expression, which comes from perturbation theory

$$\Delta Q = [2^{1/2} (-V_{s\vartheta\sigma} + 2*3^{1/2} V_{p\vartheta\sigma}) / 2]^2 / (\varepsilon_{\zeta h} - \varepsilon_{\vartheta})^2 \qquad (64)$$

Where $\varepsilon_{\zeta h}$ is the energy of the ζsp^3 (ζ = 4, 5 for Ga and for In atoms respectively) hybrid orbital, ε_{ϑ} is the energy of the ϑ orbital, and $V_{s\vartheta\sigma}$ and $V_{p\vartheta\sigma}$ are the matrix elements for interaction ζs-ϑ and σ-interaction ζp-ϑ respectively. It can be found that normally the matrix element $V_{p\vartheta\pi}$ of π-interaction ζp-ϑ has negligible value if the interaction ζsp^3-ϑ occurs in the tetrahedral cell. It has to be considered that ζs and ζp orbitals form the ζsp^3 hybrid orbital of either of the gallium or the indium atom.

The equation (64) states that the charge transfer is ΔQ electrons, which could be a significant value; determining, in fact, that one or more electrons can be moved from the ϑ orbital of the A atom to the neighboring ζsp^3 hybrid orbital of either the Ga atom or the In atom. Henceforth, two assumptions can be made:

i) the ϑ orbital of the A atom doesn't participate in either A-Ga or A-In ionic-covalent bonds although this orbital provides an electron for each of these bonds

ii) Both the A-Ga and A-In ion-covalent bonds are of the sp^3 type and it occurs via two hybrid orbitals – ζsp^3 of either the Ga atom or the In atom, and the corresponding sp^3 orbital of the A atom. For calculation purposes, the energy of the valence *p*-orbital can be determined for the A atom by the Hartree-Fock method [60] under consideration of the Koiller-Falicov assumptions [61].

It is necessary to use localized wave functions for calculations of the electron band sub-structure of a quasi-elementary cell. Linear combinations of atomic orbitals (LCAO) are used as the theoretical basis for understanding, and the electron band sub-structures are calculated using the LCAO method [62]. By definition, the primitive super-cell of the wurtzite $In_xGa_{1-x}A_yN_{1-y}$ contains elementary cells containing four atoms – two cations and two anions. The tetrahedral cell of

wurtzite $In_xGa_{1-x}A_yN_{1-y}$ is hence introduced to represent interactions between nearest neighbor orbitals.

The tetrahedral cell of wurtzite $In_xGa_{1-x}A_yN_{1-y}$ contains two atoms – the cation and anion. In this cell, an atom of one kind (either the anion or cation) is tetrahedrally surrounded by atoms of the other kind. In the quasi-elementary cells the cations are both metal (Ga and In) atoms, and the anions are the nitrogen atom and the atom 'A'.

The orbitals corresponding to both the *s*- and *p*-states of the valence electrons of atoms A, In, Ga and N are used. These are taken to be the only non-zero matrix elements; which represent interactions between nearest neighbor orbitals and between the second-neighbor orbitals (provided that the corresponding interaction will be significant). Due to the fact that the valence electron states consist of one *s* and three *p* states (p_x, p_y and p_z) per atom, the corresponding LCAO Hamiltonian matrix of a tetrahedral cell has 16 rows and 16 columns (4 neighbors X 4 interacting orbitals/neighbor).

There are five distinct types of matrix elements, $H_{\alpha\beta}$. Matrix elements $H_{\alpha\alpha}$ (for α = 1,2,3,……8) represent orbital energy terms that exist for the corresponding atom. Energy terms of the *s* and *p* states of the isolated (noninteracting) atoms A, In, Ga and N are designated by ε_s and ε_p respectively. The matrix elements $H_{\alpha\alpha}$ are connected to the widths of energy bands and they are defined in following way:

$$H_{\alpha\alpha} = C_\alpha \varepsilon_\alpha \tag{65}$$

Where α = *s*, *p* and C_α is necessarily a real number. The coefficients C_α are obtained by using the fitting method for each of wurtzite InN, GaN, InA and GaA. The energy terms ε_α are taken from Herman and Skillman [62] for atoms A, In, Ga and N.

Matrix elements $H_{\alpha\beta}$ for both (α = 9, 10,…….16 & β = 1,2,……8) and (α = 1,2,….8 & β = 8,9,…….16) represent the interactions between atomic orbitals belonging to a tetrahedral cell and atomic orbitals which are their second neighbors, respectively. They can be divided into two types: interactions between two orbitals of the same kind of atoms (anion-anion and cation-cation); and interaction between two orbitals of different kind of atoms (anion-cation).

The first type of interaction in the case of $In_xGa_{1-x}A_yN_{1-y}$ is the interaction between atomic orbitals belonging to a tetrahedral cell and the atomic orbitals which are their second neighbors A-A, In-In, In-Ga, Ga-Ga, A-N and N-N. The corresponding matrix elements are the off-diagonal $H_{\alpha\beta}$ for both sets of (α = 9, 10,…….16 & β = α - 8) and (α = 1,2,….8 & β = α + 8).

The second type of interaction in the case of $In_xGa_{1-x}A_yN_{1-y}$ is between the orbitals In-A, In-N, A-Ga and Ga-N. The corresponding matrix elements are $H_{\alpha\beta}$ for both (α = 9, 10,…….16 & β = 1,2,……8 & $\beta \neq \alpha$ - 8) and (α = 1,2,….8 & β = 8,9,…….16 & $\beta \neq \alpha$ + 8). The matrix elements of both types can be expressed in the following way:

$$H_{\alpha\beta} = E_{ll'} \Sigma_j \boldsymbol{e}^{i\,\mathbf{k}\,.\,\mathbf{d}_j} \qquad (66)$$

Where $\mathbf{d}_j$ is vector-distance between the orbitals *l* and *l'* that are engaged in the interaction, and $E_{ll'}$ is the magnitude of the matrix element. This sum is over all parts of interacting orbitals *l* (*l* = *s*, *p*) and *l'*(*l'* = *s*, *p*). Of course, the modulus $E_{ll'}$ is equal to zero if there is no interaction between two orbitals.

If $E_{ll'}$ is not zero, then there are few possible expressions depending on the types of interacting orbitals *l* and *l'*:

i) Two *s* orbitals [$E_{ss} = V_{ss\sigma}$]

ii) Between an *s* orbital and a *p* orbital, in which case

[$E_{sp} = V_{sp\sigma}\, f_{sp}(\gamma)$, where $f_{sp}(\gamma)$ is a numerical value depending on angle γ between the *p* orbital and the vector-distance $\mathbf{d}_j$]

iii) Two *p* orbitals, in which case

[$E_{xx} = V_{pp\sigma}\, f_{pp\sigma}(\gamma) + V_{pp\pi}\, f_{1pp\pi}(\gamma)$, $E_{xy} = V_{pp\sigma}\, f_{pp\sigma}(\gamma) + V_{pp\pi}\, f_{2pp\pi}(\gamma)$, where the meanings of values $f_{pp\sigma}(\gamma)$, $f_{1pp\pi}(\gamma)$ and $f_{2pp\pi}(\gamma)$ are the same as for the mixed orbital value $f_{sp}(\gamma)$]

The terms $V_{ss\sigma}$, $V_{sp\sigma}$, $V_{pp\sigma}$ and $V_{pp\pi}$ are matrix elements representing the interaction energies between correspondingly indexed orbitals. Again, these terms are designated by $V_{ll'\eta}$ where l, l' = *s*, *p* respectively and $\eta = \sigma, \pi$ represents the type of the corresponding (quasi-hybridized) *p* orbital.

The matrix elements $V_{ll'\eta}$ are given in [62] for interacting atomic orbitals in a tetrahedral cell. Unfortunately, the findings in [62] cannot be used for the matrix elements $H_{\alpha\beta}$ between orbitals belonging to second-neighbor atoms (or further). Instead, the muffin-tin orbital approach [63][64] is used to determine the $V_{ll'\eta}$.

The matrix elements $V_{sp\sigma}$ corresponding to anion-anion orbital interactions and cation-cation orbital interactions are taken to be equal to zero in accordance with the aforementioned orthogonality of LCAO Bloch orbitals.

The matrix element $V_{ss\sigma}$ for Ga-Ga and In-In has two different values. For Ga-Ga the first value corresponds to the case of Ga-Ga interaction in the AGa cell; and the second to the case of Ga-Ga interaction in GaN. For In-In the first value corresponds to the case of In-In interaction in InA; and the second to the case of In-In interaction in InN)

The matrix elements $H_{\alpha\beta}$ for α = 1, 2, …8 and β = 1, 2, …8 & $\alpha \neq \beta$ are found using [62] for interactions between orbitals of atoms which are nearest-neighbors. Here, $V_{ll'\eta} = \hbar^2 \varpi_{ll'\eta} / (m_0 d_f^2)$ where $\varpi_{ll'\eta}$ is a structural factor, and m_0 is mass of the electron. The interatomic distances are taken to be equal to the corresponding distances of the GaA cell if there is a A-Ga interaction; and likewise for GaN if the interaction is Ga-N; or of InN if the interaction is In-N and finally of InA if the interaction is I-A.

The matrix elements $H_{\alpha\beta}$ for α = 9, 10, …16 and β = 9, 10, …16 are taken to be zero because only the interactions between nearest-neighbor and second-neighbor orbitals are being considered here. Now the LCAO Hamiltonian matrix of a tetrahedral cell has been fully determined for all cases of matrix elements $H_{\alpha\beta}$ in the cases discussed above.

$$H = \left|\begin{array}{cccc:cccc} H_{1\,1} & H_{1\,2} & \ldots\ldots & H_{1\,8} & H_{1\,9} & H_{1\,10} & \ldots\ldots & H_{1\,16} \\ H_{2\,1} & H_{2\,2} & \ldots\ldots & H_{2\,8} & H_{2\,9} & H_{2\,10} & \ldots\ldots & H_{2\,16} \\ \multicolumn{8}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ H_{8\,1} & H_{8\,2} & \ldots\ldots & H_{8\,8} & H_{8\,9} & H_{8\,10} & \ldots\ldots & H_{8\,16} \\ H_{9\,1} & H_{9\,2} & \ldots\ldots & H_{9\,8} & 0 & 0 & \ldots\ldots & 0 \end{array}\right| \qquad (67)$$

The electron band sub-structure of quasi-elementary cells can be calculated by the Hamiltonian matrix (67) for a tetrahedral cell. There are actually five different Hamiltonian matrices (67) for wurtzite $In_xGa_{1-x}A_yN_{1-y}$ used in this work; corresponding to the five different tetrahedral cells.

Detailed investigation shows that both the optical absorption and the photoluminescence spectra for semiconductor compound alloys are the parts of the LCAO resultant electron band structures corresponding to configurations of the quasi-elementary cells giving the optimal

balance of deep energy pockets for the electrons in the conduction band, deep energy pockets for the holes in the valence band, and short distances between the two pockets.

For these three conditions to be realised, configurations of at least five different types of wurtzite quasi-elementary cells (each determining a corresponding sector of the LCAO electron band structure) - taken in the following order - must be used for $In_xGa_{1-x}A_yN_{1-y}$:

1) The pure quasi-elementary cell of either In-A or Ga-A surrounded by second-neighboring Ga and In cations

2) The pure quasi-elementary cell of either GaN or InN surrounded by second-neighboring Ga and In cations.

3) A mixed quasi-elementary cell of either GaA-N or InA-N containing 1.5 atoms of A and 0.5 atoms of N surrounded by second-neighboring Ga and In cations

4) A mixed quasi-elementary cell of either GaN-A and InN-A containing 1 atom of N and 1 atom of A surrounded by second-neighboring Ga and In cations

5) A mixed quasi-elementary cell of either GaN-A and InN-A containing 0.5 atoms of A and 1.5 atoms of N surrounded by second-neighboring Ga and In cations

5.2.2 Optical Properties of $In_xGa_{1-x}A_yN_{1-y}$

Many important optical properties of disordered semiconductors are investigated and reviewed in [59]. Given are two main types:

i) **Tunnel optical absorption**. At the core of this phenomenon is the overlap in space between the electron wave function $|i>$ of the initial electron state and the electron wave function $|f>$ of the final electron state. The second part of Fig.1 provides a spatial probability conceptualization of the electron wave functions of the initial state $|i>$ and of the final state $|f>$ in $In_xGa_{1-x}A_yN_{1-y}$.

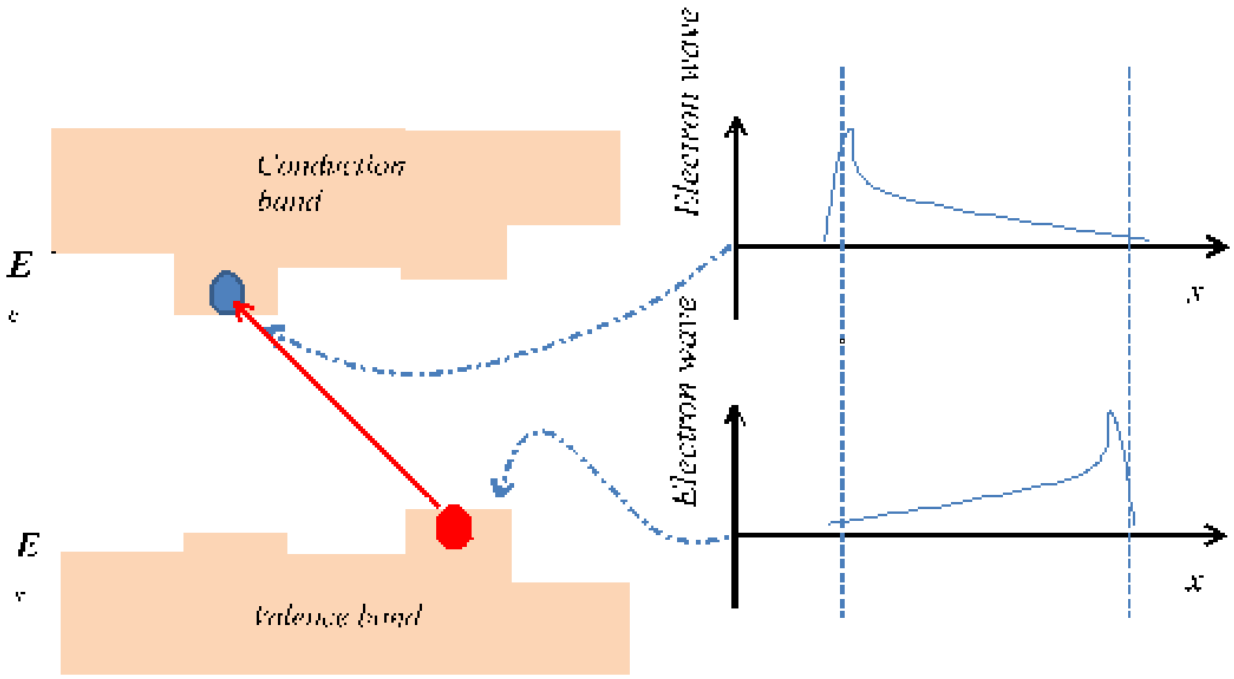

Figure 25 Tunnel optical absorption. The conduction band pocket corresponding to the band edge of super-cell 2 exhibits wavefunction overlap with the valence band edge of super-cell 4 pocket. Hence, charge tunneling between the two is possible.

ii) **Excitons of the Structure**. The basis of this phenomenon is the holding together of an electron-hole pair (separated in space) by the Coulomb interaction between an electron located in a pocket of the conduction band and a hole located in a pocket of the valence band. This quasi-particle can be characterized by two parameters: the binding energy and hydrogen-like energy levels.

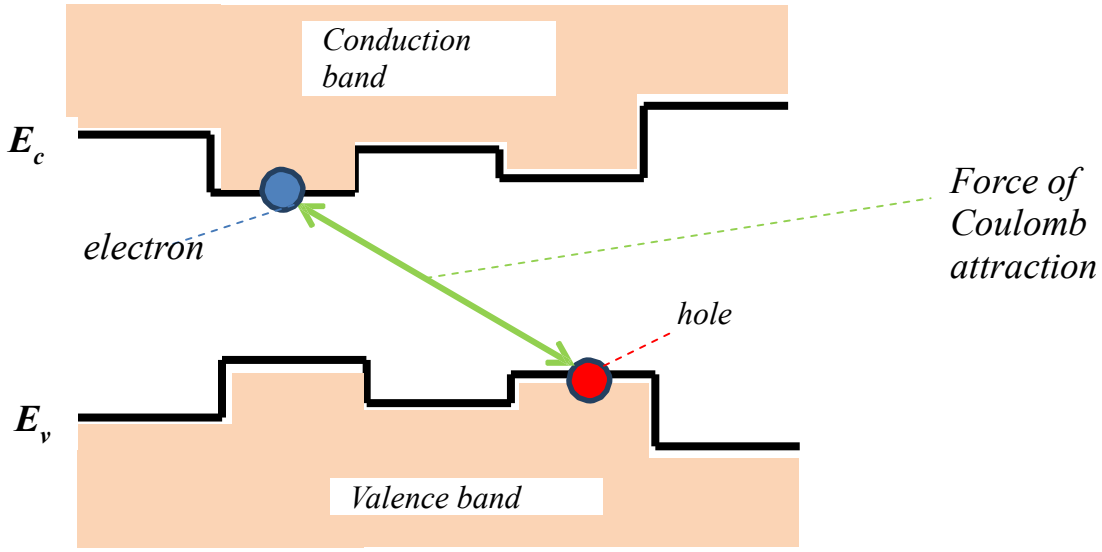

Figure 26 An exciton of the structure - a quasi-particle consisting of an electron-hole pair in which both entities are spatially confined by super-cell heterojunctions, but bound together by the Coloumb interaction

5.3 The Structure of Oxygen Impurity States in $In_xGa_{1-x}N$

Oxygen is a common contaminant in $In_xGa_{1-x}N$ – being introduced in abundance by any practical MOCVD process and still problematic even in MBE growths. There are two possible positions for oxygen atoms in the structure of $In_xGa_{1-x}N$: The oxygen atom can be alloyed to form a quaternary compound of $In_xGa_{1-x}O_yN_{1-y}$ where oxygen is substituting on the nitrogen site; or the oxygen atom can be an isolated interstitial atom in the $In_xGa_{1-x}N$ matrix. It can also be a non-isolated, interacting interstitial atom in the same structure if the concentration of the non-alloyed O impurities is high enough to allow such close spatial coupling of defects.

Real $In_xGa_{1-x}N$ typically exhibits phase segregation - containing clusters of GaN, InN, GaO_yN_{1-y}, InO_yN_{1-y}, non-stoichiometric $In_xGa_{1-x}N$/Ga, and clusters of non-stoichiometric $In_xGa_{1-x}N$/In (note that the "/" means "replaced by", and so GaO_yN_{1-y} = $(In_y/O_y)GaN_{1-y}$). The electron impurity levels of isolated interstitial O atoms in real $In_xGa_{1-x}N$ is investigated by investigating the

phases separately. The hydrogen-like (isolated orbitals) impurity atom analysis is applied for simplicity, and the behavior of interstitial O atoms is investigated for the following cases:

i) Incorporation of the O atom in a cluster of pure GaN

ii) Incorporation of the O atom in a cluster of GaO

iii) Incorporation of the O atom in a cluster of non-stoichiometric GaN:Ga (substitutions of Ga atoms on N sites)

iv) Incorporation of the O atom in a cluster of pure InN

v) Incorporation of the O atom in a cluster of pure InN/O

vi) Incorporation of the O atom in a cluster of non-stoichiometric InN/In (substitutions of In atoms on N sites)

The corresponding impurity levels are calculated and their positions in the electron band structure of real $In_xGa_{1-x}N$ are shown.

5.3.1 Electron Band Structures of Wurtzite $InxGa_{1-X}O_yN_{1-Y}$ and Non-Stoichiometric $In_xGa_{1-X}N$:In

LCAO electron band structures of both wurtzite $In_xGa_{1-x}O_yN_{1-y}$ and wurtzite non-stoichiometric $In_xGa_{1-x}N$/In for points Γ are calculated by the method given in [59]. In the case of $In_xGa_{1-x}O_yN_{1-y}$, oxygen is the atom 'A' that was used in the above discussions. The investigation shows that for the purpose of inter-band electron transitions for semiconductor compound alloys, the parts of the LCAO electron band structures can be taken (without considering the rest) corresponding to the configurations of the quasi-elementary cells giving the best combination of: deep energy pockets

for the electrons in the conduction band, tallest energy pockets for the holes in the valence band, and shortest distances between those pockets.

It turns out that two different cases should be considered:

A. GaN/O where the O atom substitutes for N in the GaN quasi-primitive cell in

$In_xGa_{1-x}O_yN_{1-y}$

In order for the above three conditions to be satisfied, configurations of five different types of wurtzite quasi-elementary cells taken in the following order must be used:

i) For GaN/O (energy separations shown in Figure 27):

1) GaO quasi-elementary cell surrounded by second-neighbor O anions
2) In-GaO-N quasi-elementary cell containing 0.5 In atoms, 1.5 Ga atoms, 0.5 N atoms and 1.5 O atoms having second-neighbor anions O and N;
3) A mixed In-GaO-N quasi-elementary cell containing 1 In atom, 1 Ga atom, 1 N atom and 1 O atom having second-neighbor anions O and N;
4) A mixed In-GaO-N quasi-elementary cell containing 1.5 In atoms, 0.5 Ga atoms, 1.5 N atoms and 0.5 O atoms having second-neighbor anions O and N;
5) InN quasi-elementary cell surrounded by second neighboring N anions.

B. InN/O,In where either O or In substitutes on the N site in the InN quasi-primitive cell in $In_xGa_{1-x}O_yN_{1-y}$.

In order for the above three conditions to be satisfied, configurations of five different types of wurtzite quasi-elementary cells taken in the following order must be used:

i) For InN/O (energy separations shown in Figure 28):

1) InO quasi-elementary cell surrounded by second-neighbor O anions;
2) In-GaO-N quasi-elementary cell containing 0.5 Ga atoms, 1.5 In atoms, 0.5 N atoms and 1.5 O atoms having second-neighbor anions O and N;
3) A mixed In-GaO-N quasi-elementary cell containing 1 In atom, 1 Ga atom, 1 atom O and 1 atom N having second-neighbor anions O and N;
4) A mixed In-GaO-N quasi-elementary cell containing 1.5 Ga atoms, 0.5 In atoms, 1.5 N atoms and 0.5 O atoms having second-neighbor anions O and N;
5) GaN quasi-elementary cell surrounded by second-neighbor N anions.

ii) For non-stoichiometric InN/In (energy separation shown in Figure 29):

1) (pure) InIn quasi-elementary cell surrounded by second-neighbor N anions;
2) A mixed InN-Ga quasi-elementary cell containing 3 In atoms, 0.5 Ga atoms and 0.5 N atoms, having second-neighbor N anions;
3) A mixed InN-Ga quasi-elementary cell containing 2 In atoms, 1 Ga atom and 1 N atom having second-neighbor N anions;
4) A mixed InN-Ga quasi-elementary cell containing 1 In atom, 1.5 Ga atoms and 1.5 N atoms, having second-neighbor N anions;
5) GaN quasi-elementary cell surrounded by second-neighbor N anions.

The LCAO electron band structure calculations for each quasi-elementary cell of the above cases was performed by method [59] described above. The electron energy terms ε_s and

ε_p for In, Ga, and Ga and N atoms, are taken from [65]. These same energy terms for the O atom substituting on the N site in the quasi-primitive GaN cell were determined to be ε_s = -22.58 eV and ε_p = -10.42 eV. It was also found that the charge transfer from the O atom to Ga atom is ΔQ=0.72; found with preliminarily determined matrix elements $V_{s\vartheta\sigma}$ = -0.43 eV and $V_{p\vartheta\sigma}$ = 3.81 eV. It should be understood that the value found for charge transfer means that the valence electron cloud of the O atom is partially displaced in the direction toward the Ga atom in the corresponding quasi-elementary cell.

The energy terms for an O atom substituting on the N site of a quasi-primitive InN cell are taken from [61]. The nearest-neighbor matrix elements were determined according to [62], and the required inter atomic distances were calculated using a perfect tetrahedral arrangement with the well-known radii of In, Ga, N and O atoms.

MATLAB code was written to perform all of the calculations. The calculated electron band structures are given in Figure 27 for GaN/O, Figure 28 for InN/O, and Figure 29 for non-stoichiometric InN/In. The energy levels Γ^{υ}_{c1} and Γ^{υ}_{v15} were determined by taking the energy of the vacuum level (free electron) as being equal to zero. Hence, all the charges are in a potential well. The energy difference $E^{\upsilon}{}_{g} = (\Gamma^{\upsilon}_{c1} - \Gamma^{\upsilon}_{v15})$ gives the energy band gap of sector υ. The shifts of the boundaries of the energy band gaps and the corresponding energy intervals are due to defects in the crystal lattice of InGaN – i.e. the existence of two sorts of atoms on both cationic and anionic sites.

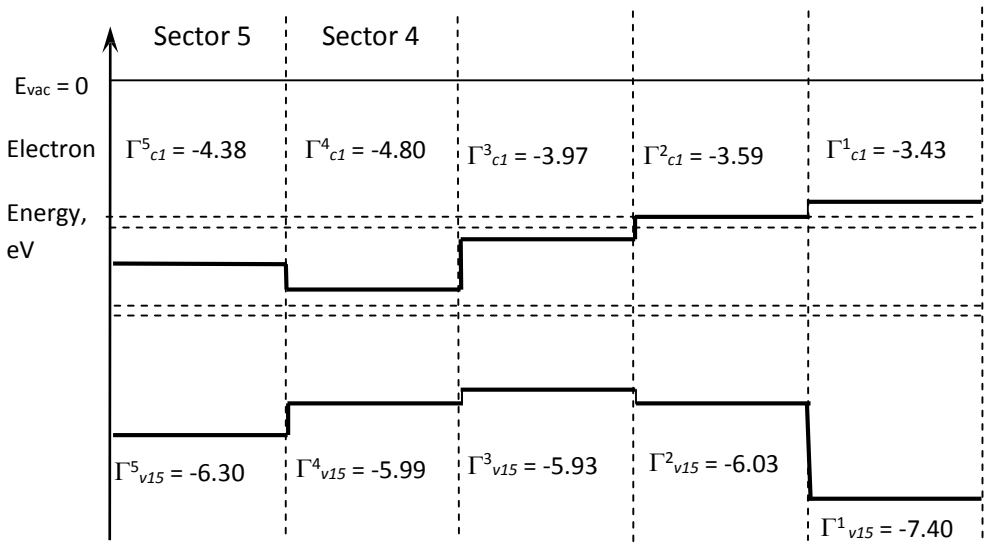

Figure 27 The electron band structure in the case of GaN/O for the Γ point only with O donor levels (not to scale)

5.3.2 Electronic Structures of Isolated Oxygen Impurity States

The electronic structure of isolated oxygen impurity states can be found by considering the case that the valence electrons of the incorporated interstitial O atom are not involved in any ion-covalent bonds in the quasi-primitive cells (InN, GaN, In-GaN, In-In, GaO and InO). In this case, a relatively long distance influence of the valence states is to be expected in the above clusters, and a hydrogen-like impurity atom analysis [66] can be applied to imagined isolated interstitial O atoms hosted in a real InGaN lattice. The Schrodinger equation (giving the electron wavefunction properties) for a hydrogen atom - but with effective electron mass m and dielectric constant κ values for the quasi-primitive cells (InN, GaN, In-GaN, In-In, GaO and InO) - can be used:

$$(- \hbar^2\Delta/(2m) - e^2/(\kappa r))\, F(\mathbf{r}) = E\, F(\mathbf{r}) \qquad (68)$$

Where Δ is the Laplacian and both involved charges in the potential were of value e. $F(\mathbf{r})$ is the ground-state wave function of the form

$$F(r) = (\pi a^3)^{-1/2} e^{-r/a} \quad (69)$$

Where a is the effective Bohr radius, which determines the localization radius of the wave function, and

$$a = \hbar^2 \kappa / (me^2) \quad (70)$$

From the basic Bohr model of the atom. Solving (68) with (69, 70) gives isolated electron state energies of the impurity atom

$$E_n = - e^4 m / (2n^2 \kappa^2 \hbar^2) \quad (71)$$

Where n is an integer number greater than zero. To determine the isolated electron states we consider the contribution from n = 1.

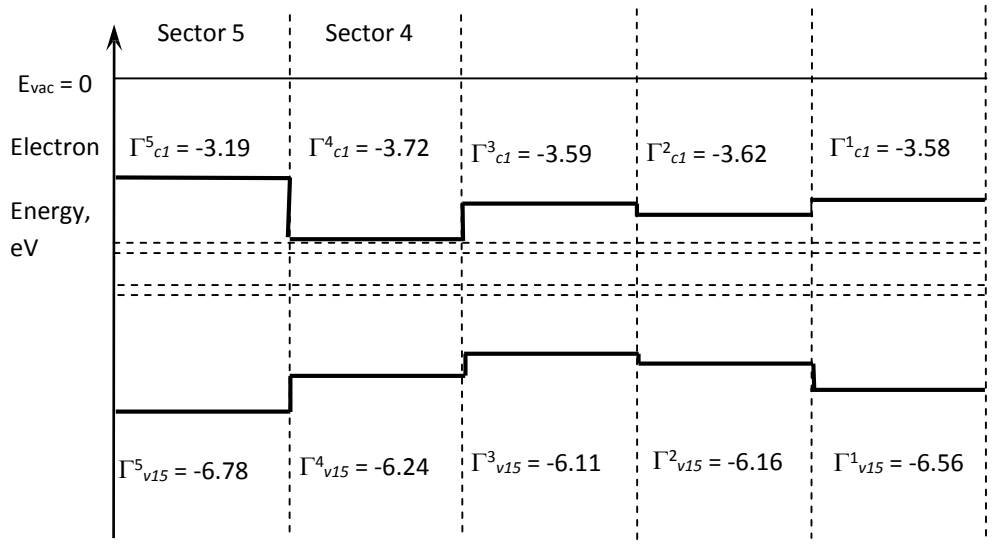

Figure 28 Electron band structure for InN/O for point Γ only with hydrogen-like O donor levels (not to scale)

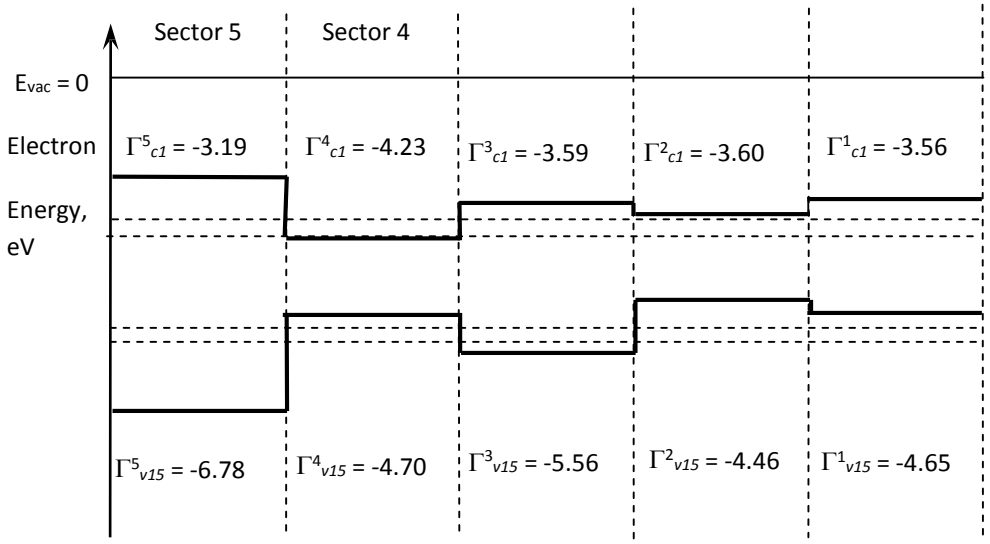

Figure 29 Electron band structure of non-stoichiometric InN/In for the Γ point only and hydrogen-like interstitial O levels (not to scale)

(70) and (71) have also been used to determine the localization radii and the electron energy states (respectively) of interstitially incorporated O atoms in quasi-primitive cells of GaN, In-GaN, In-In, GaO and InO (along with InN, which was shown). The values of κ were taken from [62], and the electron masses m were taken from [59]. To summarize, the following results have been obtained (declaring zero energy at the vacuum level):

i) If the O atom is interstitially incorporated in a GaN quasi-primitive cell, the impurity state has an energy of -3.79 eV (Figure 30), which acts as a donor level with ionization energy 0.09 eV. Also, in the case of InGaN this state is a donor level with an ionization energy of -1.42 eV.

ii) The impurity state has energy -3.86 eV (Figure 31), if the O atom is interstitially incorporated in a GaO quasi-primitive cell, which acts as a donor level with ionization energy 0.1 eV. Also, in the case of InGaN this state is a donor level with an ionization energy of -1.35 eV.

iii) The impurity state has energy -5.11 eV (Figure 33), if the O atom is interstitially incorporated in the InIn quasi-primitive cell, in which it acts as an impurity state (which is in the valence band of the cell). However, in this case this state still acts as a donor level for the InN quasi-primitive cell (0.33eV) and for the InO quasi-primitive cell (0.25eV). Lastly, in InGaN the level is also a donor level having an ionization energy of -0.1 eV.

iv) The impurity state has energy -5.15 eV (Figure 32), if the O atom is interstitially incorporated in the InO quasi-primitive cell, in which it acts as a donor level of 0.19. The quasi-primitive cell of InN for this case has ionization energy 0.19 eV. For the InGaN, this level is a donor level having ionization energy -0.06 eV.

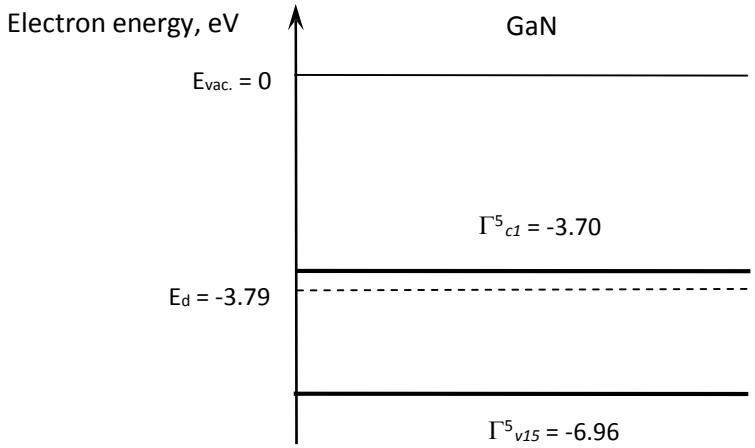

Figure 30 Quasi-primitive cell of GaN. Donor level Ed = -3.79 eV

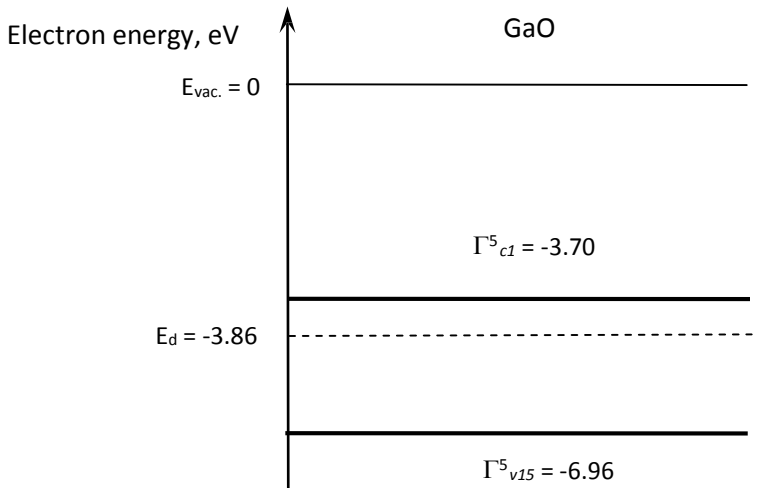

Figure 31 Quasi-primitive cell of GaO. donor level Ed = -3.86 eV

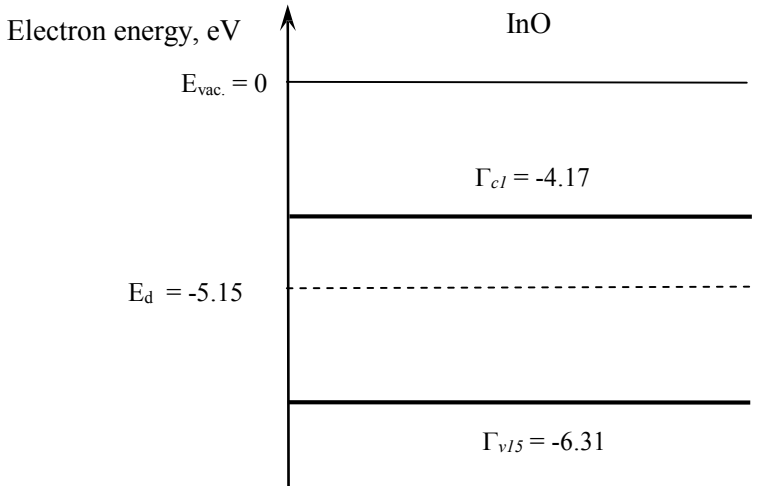

Figure 32 InO quasi-primitive cell with donor level E = -5.15 eV (not to scale). The results were previously calculated by D. Alexandrov

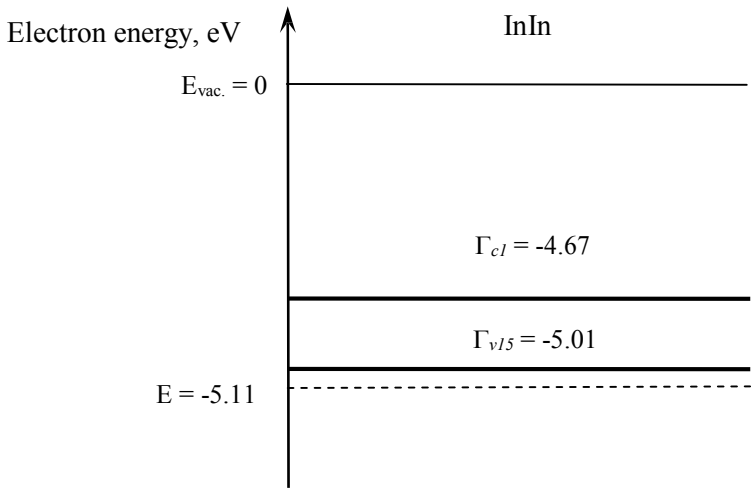

Figure 33 InIn quasi-primitive cell with level E = -5.11 eV (not to scale). The results were previously calculated by D. Alexandrov

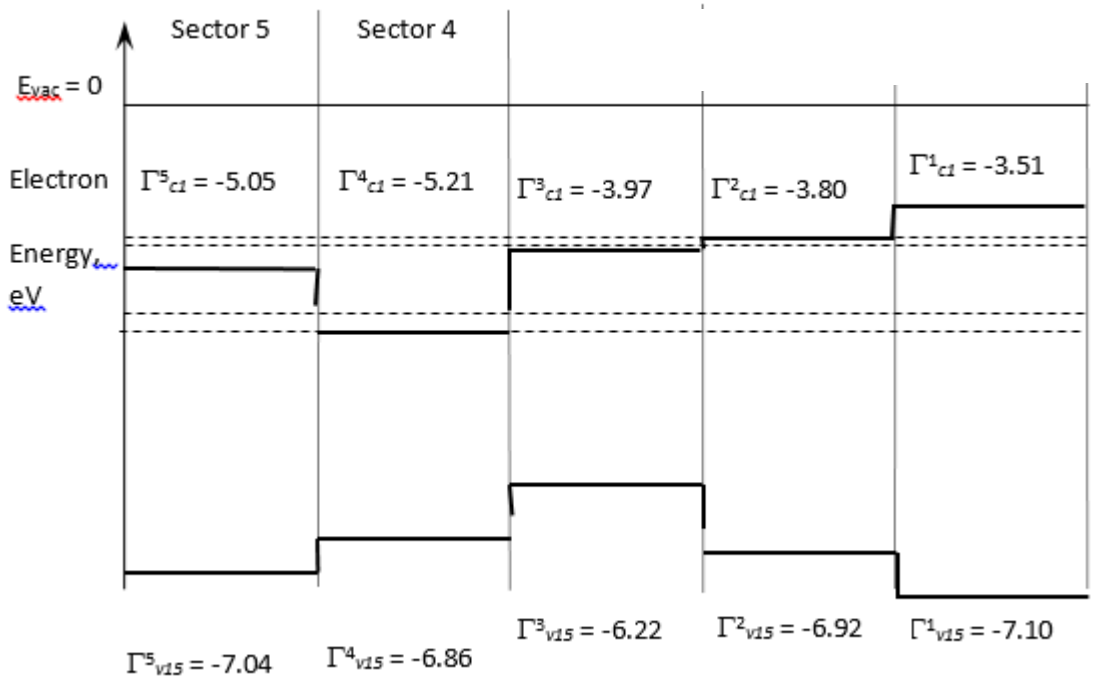

Figure 34 Energy band gap of InGaN with hydrogen-like O donor levels

The energy band gap of InGaN is provided in Figure 34 [59], which shows three intermittent quasi-primitive cells. The hydrogen-like O donor levels calculated above have been incorporated.

5.3.3 Optical Properties of InGaN Containing O Impurity Atoms

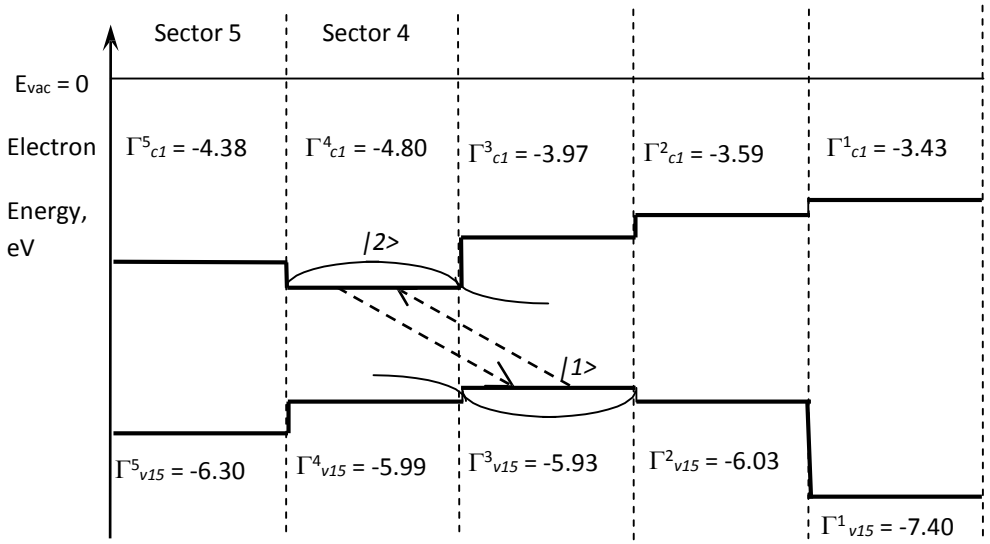

Figure 35 Electron band structure for GaN/O for the Γ point only. Tunnel optical transition wavefunctions between pockets of the super-cells in both the conduction and the valence bands are shown.

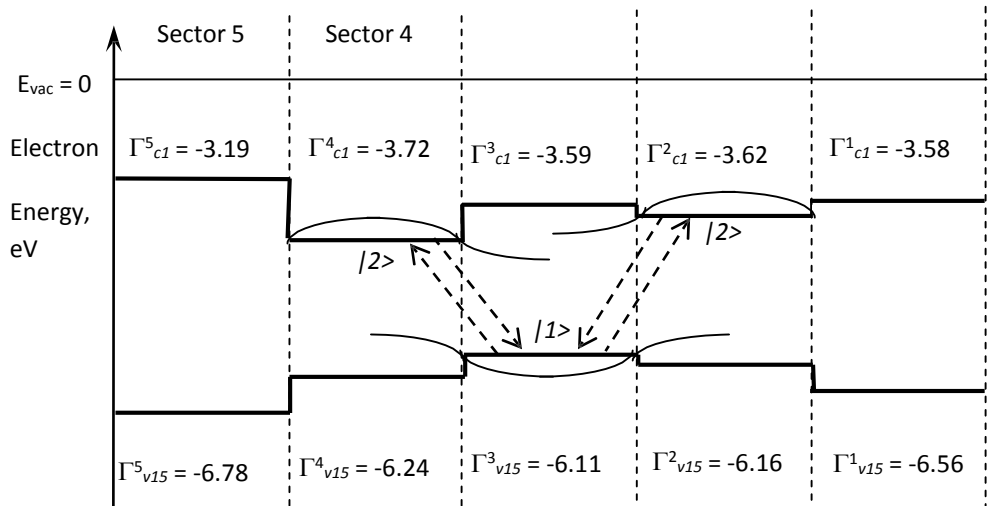

Figure 36 Electron band structure for InN/O for the Γ point only. Wavefunctions of super-cells with tails extending into other cells (creating the possibility for tunnel optical transition between pockets) are shown.

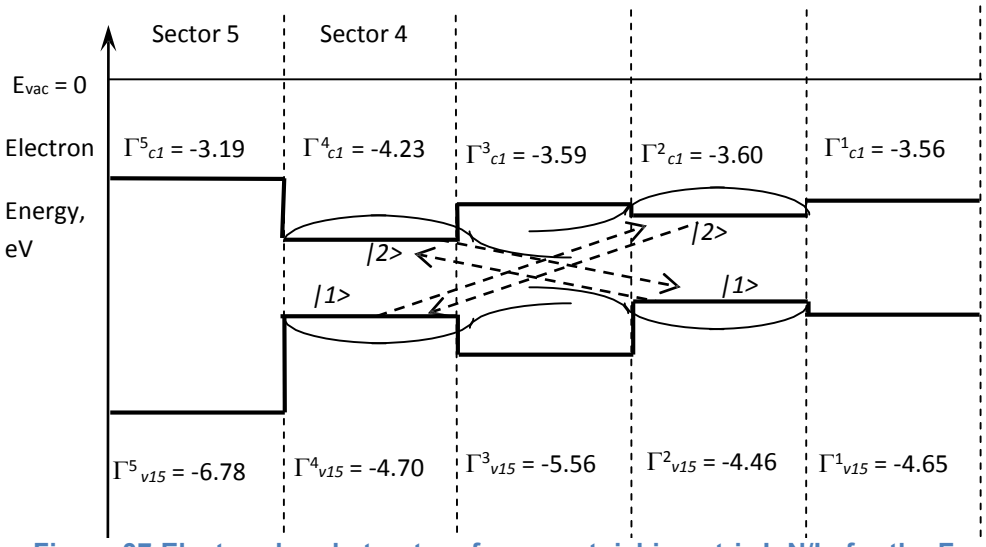

Figure 37 Electron band structure for non-stoichiometric InN/In for the Γ point only. Tunnel optical transitions from wavefunction overlap between super-cell pockets of the conduction and valence bands are shown.

Tunneling optical absorption and tunneling optical radiation occurs in both of: InGaN containing impurity O atoms and in non-stoichiometric InGaN/In. As can be seen above, the following optical transitions are allowed according to the phenomena in [59]:

i) In Figure 35 the electron wave function of state Γ^4_{c1} overlaps with the electron wave function of state Γ^3_{v15} and tunneling may occur. Thus, both tunneling optical absorption and tunneling optical radiation should be observed in the corresponding energy interval of 1.13 eV. The optical absorption spectra of InGaN samples show an optical absorption edge of ~1.20 eV that may be evidence of defect domination for the defect that is associated with the energy band diagram in Figure 35.

ii) There are two overlaps in Figure 36 – the electron wave function of state Γ^4_{c1} overlaps the wavefunction of state Γ^3_{v15} ; and the wavefunction of state Γ^2_{c1} overlaps the wavefunction of state Γ^3_{v15}. Thus, both tunneling optical absorption and tunneling optical radiation between both states should be observed at energy

intervals of 2.39 eV and 2.49 eV respectively. Incidently, a lot of InGaN samples show an optical absorption edge in the interval 2.40 – 2.50 eV, and this may be evidence of defect domination of the defect shown in the band diagram of Figure 36. Also, the observed optical emissions of some structures using InGaN as a light emitting layer in the yellow and yellow-green parts of the optical spectrum may be explained by the presence of the defects associated with the transitions seen in Figure 36.

iii) In Figure 37– the electron wave function of state Γ^4_{c1} overlaps the wavefunction of state Γ^2_{v15} and the wavefunction of state Γ^2_{c1} overlaps the wavefunction of state Γ^4_{v15} . Thus, both tunneling optical absorption and tunneling optical radiation between states should be observed in the energy ranges of 0.23 eV and 1.10 eV respectively. The optical absorption spectra of InGaN samples show an optical absorption edge of ~1.20 eV that may be evidence of defect domination by the defect that is associated with the energy band diagram seen in Figure 37.

5.3.4 Dielectric Susceptibility of InGaN

According to [62], both the real and imaginary parts of the dielectric susceptibility $\varepsilon = \chi_1(0) + i\chi_2(\omega)$ can be expressed as:

$$\chi_1(0)=(2e^2/\Omega)\,\Sigma_{\boldsymbol{k}(\text{full}),\,\boldsymbol{k}'(\text{empty})}\,|<\psi_{\boldsymbol{k}'}|\,x\,|\psi_{\boldsymbol{k}}>|^2\,/\,(E_{\boldsymbol{k}'}-E_{\boldsymbol{k}}) \tag{72}$$

and

$$\chi_2(\omega)=(\pi\, e\,\hbar/(m\,\omega\,\Omega))\,\Sigma_{\boldsymbol{k}(\text{full}),\,\boldsymbol{k}'(\text{empty})}\,|<\psi_{\boldsymbol{k}'}|\,\partial/\partial x\,|\psi_{\boldsymbol{k}}>|^2\,*\delta(E_{\boldsymbol{k}'}-E_{\boldsymbol{k}}-\hbar\omega) \tag{73}$$

Where *e* is the electron charge, *m* is electron mass, ω is the frequency of the electromagnetic (EM) field, Ω is volume, $\psi_{k'}$ is the electron wave function corresponding to an empty state having electron wave vector ***k'*** and electron energy $E_{k'}$, and ψ_k is an electron wave function corresponding to a filled state (having electron vector ***k*** and electron energy E_k). The sum is over all pairs of dissimilarly filled states belonging to the corresponding sub-bands; and the matrix elements are along the axis Ox. Finding $\chi_1(0)$ and $\chi_2(\omega)$ requires calculation of the corresponding electron band structures and subsequent determination of the dominating optical inter-band transitions (giving the most significant terms) for the sums in both (72) and (73).

In this work, $\chi_1(0)$ has been determined for above cases of InGaN defects. The energy differences in the terms ($E_{k'} - E_k$) in (72) were taken to be the narrowest energy intervals ΔE between the lowest sub-band of the conduction band and the highest sub-band of the valence band which also corresponded to approximately parallel sub-bands (optimal energy and spatial locality). The results of these calculations are summarized in Table 1.

Defect case	**ΔE, eV**	**$\chi_1(0)$**
$In(Ga/O_y)N_{1-y}$	6.63	4.0768
$Ga(In/O_y)N_{1-y}$	8.73	3.8437
InGaN/In	2.57	7.4752

Table 1 Calculated values of the real component of electrical susceptibility for the smallest energy gap of the band structures calculated for the quasi-elementary cells listed

6 References

[1] T. Mori et al, "Schottky barriers and contact resistances on p-type GaN," *Applied Physics Letters*, vol. 69, no. 23, p. 3537, 1996.

[2] D. Steigerwald et al, "III-V Nitride Semiconductors for High-Performance Blue and Green Light-Emitting Devices," [Online]. Available: *http://www.tms.org/pubs/journals/jom/9709/steigerwald-9709.html*. [Accessed 4 9 2014].

[3] E. Lin et al, "Low resistance ohmic contacts on wide band-gap GaN," *Applied Physics Letters*, vol. 64, no. 8, p. 1003, 1994.

[4] T. Hirscha et al, "Evolution of Ti Schottky Barrier Heights on n-Type GaN with Annealing," in *MRS Proceedings*, V. 449, 1996.

[5] V. Adivarahan et al, "Very-low-specific-resistance Pd/Ag/Au/Ti/Au alloyed ohmic contact to p GaN for high-current devices," *Applied Physics Letters*, vol. 78, p. 2781, 2001.

[6] Y. Seong et al, "Low-resistance Pt/Ni/Au ohmic contacts to p-type GaN," *Applied Physics Letters*, vol. 74, no. 1, pp. 70-72, 1999.

[7] G. Zhao et al, "Stress and its effect on optical properties of GaN epilayers grown on Si(111), 6H-SiC(0001), and c-plane sapphire," *Applied Physics Letters*, no. 83, p. 677, 2003.

[8] F. Schubert et al, "Optical properties of Si-doped GaN," *Applied Physics Letters*, vol. 71, no. 7, p. 921, 1997.

[9] S. Nakamura et al, "Si- and Ge-Doped GaN Films Grown with GaN Buffer Layers," *Japanese Journal of Applied Physics*, vol. 31, no. 1, p. 2883, 1992.

[10] E. Lin et al, "A comparative study of GaN epilayers grown on sapphire and SiC substrates by plasma-assisted molecular-beam epitaxy" *Appl. Phys. Lett.*, vol. 62, no. 26, p. 3479, 1993.

[11] F. Brunnera et al, "Quantitative analysis of in situ wafer bowing measurements for III-nitride growth on sapphire," *Journal of Crystal Growth*, vol. 310, no. 10, pp. 2432-2438, 2008.

[12] T. Hino et al, "Characterization of threading dislocations in GaN epitaxial layers," *Applied Physics Letters*, vol. 76, no. 23, p. 3421, 2000.

[13] A. Ponce et al, "Microstructure of GaN epitaxy on SiC using AlN buffer layers," *Applied Physics Letters*, vol. 67, no. 3, p. 410, 1995.

[14] G. Togtema, *p-type Gallium Nitride Semiconductor Development and Characterization of LEDs and other Devices*. M.Sc. Thesis, Lakehead University, Canada, May, 2013.

[15] "NSM archive of Physical Properties of Semiconductors," 08 2014. [Online]. Available: *http://www.ioffe.rssi.ru/SVA/NSM/Semicond/*. [Accessed 08 2014].

[16] J. I. Pankove, "Properties of Gallium Nitride," in *MRS Proceedings*, Volume 97, 1987, SERI, Golden, Colorado and University of Colorado, Boulder, Colorado, 1987.

[17] Landolt-Börnstein, "Gallium nitride (GaN), shallow impurities," in *Impurities and Defects in Group IV Elements, IV-IV and III-V Compounds. Part b: Group IV-IV and III-V Compounds. Group III Condensed Matter Volume 41A2b*, Springer Berlin Heidelberg, 2003, pp. 1-9.

[18] R. Dingle et al, "Absorption, Reflectance, and Luminescence of GaN Epitaxial Layers," *Phys. Rev.* B, vol. 4, p. 1211, 1971.

[19] R. Korotkov et al, "Electrical properties of p-type GaN:Mg codoped with oxygen," *Applied Physics Letters*, vol. 78, no. 2, pp. 222-224, 2001.

[20] A.Christou et al, *Applied Physics Letters*, vol. 44, pp. 796-798, 1984.

[21] D. Koleske et al, "Growth model for GaN with comparison to structural, optical, and electrical properties," *Journal of Applied Physics*, vol. 84, no. 4, 1998.

[22] T. Yamaguchi et al, "Novel InN growth method under In-rich condition on GaN/Al2O3 (0001) templates," *Phys. Status Solidi*, vol. C 6, no. S360, 2009.

[23] E. Trybus et al, "Extremely high hole concentrations in c-plane GaN," *Phys. Status Solidi,* vol. C 6, no. S788, 2009.

[24] K. S. A. B. et al., "Initial experiments in the migration enhanced afterglow growth of GaN and InN," *Phys. Status Solidi*, vol. C 9, no. 3-4, pp. 1070-1073, 2012.

[25] J. Pankove. et al, "The relation of active nitrogen species to high-temperature limitations for (000) GaN grown by radio-frequency-plasma-assisted molecular beam epitaxy," *Applied Physics Letters*, vol. 74, pp. 3836-3838.

[26] P. Chen, "Ch. 2, Preparation and Growth of GaN Thin Films," in *Remote Plasma Enhanced Chemical Vapour Deposition Growth and Characterisation of Gallium Nitride Films*, M.Sc Thesis, Macquarie University, Australia, 2002, pp. 2-6.

[27] K. S. A. Butcher et al, "Crystal size and oxygen segregation for polycrystalline GaN," *Journal of Applied Physics*, vol. 92, no. 6, 2002.

[28] K. S. A. Butcher et al, "InN grown by migration enhanced afterglow," *Physica Status Solidi* A, vol. 209, no. 1, pp. 41-44, 2012.

[29] K. Maeshige et al, "Functional design of a pulsed two-frequency capacitively coupled plasma in CF4/Ar for SiO2 etching," *Journal of Applied Physics*, vol. 91, no. 12, pp. 9494-9501, 2002.

[30] T. Kitajima et al, "Functional separation of biasing and sustaining voltages in two-frequency capacitively coupled plasma," *Applied Physics Letters*, vol. 77, no. 4, p. 489, 2000.

[31] Geoffrey K.H Panga et al, "Topographic and phase-contrast imaging in atomic force microscopy," *Ultramicroscopy*, vol. 81, no. 2, pp. 35-40, 2000.

[32] L. v. d. Pauw, "A method of measuring specific resistivity and Hall effect of discs of arbitrary shape," *Phillips Research Reports*, pp. 1-9, 1958.

[33] B. Gurbulak, "The optical investigation of TlGa0.999Pr0.001Se2 and TlGaSe2 single crystals," *Physica B: Condensed Matter*, vol. 293, no. 3-1, pp. 289-296, 2001.

[34] P. T. Chen et al, "High Energy Urbach Characteristics Observed for Gallium Nitride Amorphous Surface Oxide," *Elsevier, Thin Solid Films*, no. 496, pp. 342-345, 2006.

[35] M. Kranjčec et al, "Fundamental optical absorption edge and compositional disorder in γ1-(GaxIn1−x)2Se3 single crystals," *Physica Status Solidi*, vol. 144, no. 1, pp. 223-233, 2006.

[36] H. Amano et al, "Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer," *Applied Physics Letters*, vol. 48, no. 5, p. 353, 1986.

[37] H. Lu et al, "Effect of an AlN buffer layer on the epitaxial growth of InN by molecular-beam epitaxy," *Applied Physics Letters*, vol. 79, no. 10, p. 1489, 2001.

[38] K. Uchida et al, "Nitridation process of sapphire substrate surface and its effect on the growth of GaN," *Journal of Applied Physics*, vol. 79, no. 7, p. 3487, 1996.

[39] J. F. Schetzina, "Integrated heterostructures of Group III-V nitride semiconductor materials including epitaxial ohmic contact non-nitride buffer layer and methods of fabricating". United States Patent *US5670798 A*, 29 03 1995.

[40] H. A. et al, "P-Type Conduction in Mg-Doped GaN Treated with Low-Energy Electron Beam Irradiation (LEEBI)," *Japanese Journal of Applied Physics*, vol. 28, no. pt.2, 12, 1989.

[41] O. Gelhausen et al, "Influence of low-energy electron beam irradiation on defects in activated Mg-doped GaN," *Applied Physics Letters*, vol. 81, no. 20, p. 3747, 2002.

[42] J. J. Coleman et al, "Time-dependent study of low energy electron beam irradiation of Mg-doped GaN grown by metalorganic chemical vapor deposition," *Applied Physics Letters*, vol. 69, no. 11, 1996.

[43] G. Namkoong et al, "Metal modulation epitaxy growth for extremely high hole concentrations above E19cm-3 in GaN," *Applied Physics Letters*, vol. 93, no. 17, 2008.

[44] M. Sze, *Physics of Semiconductor Devices*, 3rd ed., New Jersey: Wiley, 2007.

[45] L. V. d. Pauw, "A Method of Measuring the Specific Resistivity and Hall Effect of Discs of Arbitrary Shape," *Phillips Research Reports*, vol. 13, pp. 1-9, 1958.

[46] M. Shur et al, "Electron Mobility in Two-Dimensional Electron Gas in AlGaN/GaN Heterostructures and in Bulk GaN," *Journal of Electronic Materials*, vol. 25, no. 5, 1996.

[47] Y. C. Yeo et al, *J. Appl. Phys*., 83, 1429, 1998

[48] K. Karch et al, *Phys. Rev*., B 57, 7043, 1998

[49] C. Noguez, *Phys. Rev*., B 58, 12641, 1998

[50] K. Amimer et al, *Appl. Phys. Lett*., 76 2580, 2000

[51] E. Stach et al, *Materials Research Society Symposium*, 617, J3.5.1, 2000

[52] S. Pantelides, W Harrison, *Phys. Rev*., B 11, 3006, 1975

[53] D. J. Dugdale et al, *Phys. Rev.*, B 61, 12933, 2000

[54] Z. Yang, Z. Xu, *Phys. Rev.*, B 54, 17577, 1996

[55] D. Vogel et al, *Phys. Rev.*, B 55, 12836, 1997

[56] D. E. Boucher et al, *Phys. Rev.*, B 59, 10064, 1999

[57] J. Wu et al, *Applied Physics Letters* 80, 3967, 2002

[58] V. Yu. Davydov et al, *Physica Status Solidi.* (b) 234, 787, 2002

[59] D. Alexandrov, *Journal of Crystal Growth*, 246, 325, 2002

[60] F. London et al, *Phys. 42*, 375, 1927

[61] L. O'Raifeartaigh, *The Dawning of Gauge Theory, part I*, 1997

[62] W.A. Harrison, *Electronic Structure and the Properties of Solids*, Dover Publ. Inc, 1989

[63] O. K. Andersen, *Sol. State Commun.*, 13, 133, 1973

[64] O. K. Andersen et al, *Phys. Rev.*, B 17, 1209, 1978

[65] F. Herman, B. Skillman, "*Atomic Structure Calculations*", Prentice Hall, 1963

[66] L. D. Landau, E. M. Lifshitz, "*Quantum Mechanics (Non-Relativistic Theory)*", 2nd revised edition, Pergamon Press, Oxford, 1965

[67] D. Alexandrov, S. Butcher, T. Tansley, *Journal of Crystal Growth*, 288, 261, 2006

Semiconductor			Crystal Struct.	Lattice Canst. at300 K(A)	Bandgap (eV)		Band	Mobility at 300 K (cm²/V-s)		Effective Mass		E_0
					300 K	OK		μ_n	μ_p	m_l/m_0	m_t/m_0	
	C	Carbon (djarnond)	D	3.56683	5.47	5.48	I	1,800	1,200	0.2	0.25	5.7
	Ge	Germanium	D	5.64613	0.66	0.74	I	3,900	1,900	1.64/0.082	0.04lh,0.28hh	16.0
	Si	Silicon	D	5.43102	1.12	1.17	I	1,450	500	0.98/0.19	0.16lh,0.49hh	11.9
IV-IV	SiC	Silicon carbide	W	a=3.086,c=15.117	2.996	3.03	I	400	50	0.60	1.00	9.66
III-V	AlAs	Aluminum arsenide	Z	5.6605	2.36	2.23	I	180		0.11	0.22	10.1
	AlP	Aluminum phosphide	Z	5.4635	2.42	2.51	I	60	450	0.212	0.145	9.8
	AlSb	Aluminum antimonide	Z	6.1355	1.58	1.68	I	200	420	0.12	0.98	14.4
	BN	Boron rutride	Z	3.6157	6.4		I	200	500	0.26	0.36	7.1
	"	"	W	a=2.55,c=4.17	5.8		D			0.24	0.88	6.85
	BP	Boron phosphide	Z	4.5383	2.0		I	40	500	0.67	0.042	II
	GaAs	Gallium arsenide	Z	5.6533	1.42	1.52	D	8,000	400	0.063	0.076/0.Shh	12.9
	GaN	Gallium nitride	W	a=3.189,c=5.182	3.44	3.50	D	400	10	0.27	0.8	10.4
	GaP	Gallium phosphide	Z	5.4512	2.26	2.34	I	110	75	0.82	0.60	11.1
	GaSb	Gallium antimonide	Z	6.0959	0.72	0.81	D	5,000	850	0.042	0.40	15.7
	lnAs	Indium arsenide	Z	6.0584	0.36	0.42	D	33,000	460	0.023	0.40	15.1
	InP	Indium phosphide	Z	5.8686	1.35	1.42	D	4,600	150	0.077	0.64	12.6
	InSb	Indium antimonide	Z	6.4794	0.17	0.23	D	80,000	1,250	0.0145	0.40	16.8
II-Vl	CdS	Cadmium sulfide	Z	5.825	2.5		D			0.14	0.51	5.4
	"	"	W	a=4.136,c=6.714	2.49		D	350	40	0.20	0.7	9.1
	CdSe	Cadmium selenide	Z	6.050	1.70	1.85	D	800		0.13	0.45	10.0
	CdTe	Cadmium telluride	Z	6.482	1.56		D	1,050	100			10.2
	ZnO	Zinc oxide	R	4.580	3.35	3.42	D	200	180	0.27		9.0
	ZnS	Zinc sulfide	Z	5.410	3.66	3.84	D	600		0.39	0.23	8.4
	"	"	W	a=3.822,c=6.26	3.78		D	280	800	0.287	0.49	9.6
IV-VI	PbS	Lead sulfide	R	5.9362	0.41	0.286	I	600	700	0.25	0.25	17.0
	PbTe	Lead telluride	R	6.4620	0.31	0.19	I	6,000	4,000	0.17	0.20	30.0

D = Diamond, W = Wurtzite, Z = Zincblende, R = Rock salt. I, D = Indirect, direct bandgap. *l,t,lh,hh* = Longitudinal, transverse, light-hole, heavy-hole effective mass.