Compound Semiconductor Quantum Wells and Dots for Light-Emitting Diode Applications: Insights on Structural and Optoelectronic Properties

by

Christina M. Jones

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Applied Physics) in the University of Michigan 2017

Doctoral Committee:
Professor Roy Clarke, Co-chair
Assistant Professor Emmanouil Kioupakis, Co-chair
Professor Rachel Goldman
Professor Çagliyan Kurdak
ACKNOWLEDGEMENTS

Personal acknowledgements

My graduate school experience here at the University of Michigan has not just been academic, but has been a truly wonderful life experience. There are many people who have contributed to the learning, growth, challenge, support, sense of accomplishment, and joy that I have felt throughout these years that I would like to acknowledge.

First and foremost, I would like to thank my research advisors, Roy Clarke and Manos Kioupakis. I could never have asked for more skilled, brilliant, kind, supportive, and patient people to advise me through these graduate years. Roy took a chance on me four years ago when I pestered him persistently asking to join his group, and he offered a wonderful project and collaboration for me to be a part of, which has not only provided challenging and fun technical skills to learn, but also motivated me and excited me with its meaningful cause; I have become a strong and enthusiastic advocate of nitrides and their solid-state lighting capabilities, telling everyone how great LED light bulbs are wherever I go. Roy has shown great insight into my work and true care for not just my technical progress but also my well-being. Furthermore, Manos graciously accepted me into his group as part of this collaboration. He has spent countless hours with me teaching and explaining concepts, brainstorming, troubleshooting, and guiding. His dedication as a mentor and advisor led to one of the most rewarding experiences of my graduate career, when I attended a nitrides conference and won a poster award among world leaders in our field. Both Roy and Manos offer an incredible wealth of knowledge ranging from technical expertise to stories of interesting R&D successes and failures to language and culture and beyond. They have calmed my nerves, allayed my anxieties, and been endless sources of optimism. They have listened to and supported my goals and aspirations for my Ph.D., including travel, attending new conferences, and gaining specific skills. I can’t express how fortunate I feel to have had the great privilege of working with my advisors and
mentors, Roy and Manos.

I would also like to thank my committee members, Professors Rachel Goldman and Çagliyan Kurdak. They have donated much of their time to me these past few years to review my work and provide insight and suggestions. I appreciate the energy and dedication they have given to me and my projects.

Professor Emeritus Yizhak Yacoby was also an invaluable mentor to me throughout my Ph.D. experimental work, guiding and teaching through all of our experiments with his knowledge of X-ray rodscans and COBRA analysis. His enthusiasm, energy, and drive for science are inspiring, as is his patience and dedication for teaching. My time working at Argonne National Lab and learning COBRA in Israel was some of the most memorable time in my graduate studies. I feel so fortunate to have been able to work with Yizhak and am so grateful for his hospitality during my travels.

There have been many people who have facilitated my work throughout my graduate studies. Shira Yochelis and Wesley Hong both worked with me at the beamline taking X-ray data, and they were instrumental to the success of many of our experiments. They were also fun, wonderful company and I enjoyed learning from them about their expertise. Our beamline scientists at Sector 33-ID, Christian Schlepütz and Zhan Zhang, proved over and over again to not only be great people but also never failed to impress me with their knowledge and skill. I am also thankful for Kai Sun, Haiping Sun, and Bobby Kerns from the Michigan Center for Materials Characterization, who have always been happy to spend time helping and teaching me various characterization and analysis skills ever since my first REU here at Michigan seven years ago. In particular, Kai Sun’s expertise with STEM was integral for the understanding of my experimental project. Finally, I am grateful to Stefan Birner from nextnano, who has answered numerous emails and questions on software use and troubleshooting, no matter the day of week or the hour of day. His responses helped teach me not just about the software use but also about the $k \cdot p$ method in general.

My groupmates have been an endless source of support, insight, challenging questions, and interesting discussions throughout graduate school. They have also spent much of their time answering my questions and instructing me on new techniques or skills, which has been an invaluable part of my growth as a researcher. From Roy’s group I would like to thank Dr. Nancy Senabulya, James Mathis, Dr. Renee Harton, Dr. Abe Oraiqat, Dr. Yongsoo Yang, and Dr. Vladimir Stoica. From Manos’s group, I would like to thank Dr. Guangsha Shi, Dr. Dylan Bayerl, Andrew McAllister, Jihang Lee, Logan Williams, Kelsey Mengle, Nocona Sanders, Keyi Cai, Dr. Alex
Toulouse, Annika Fash-White, Dr. Suvadip Das, and Dr. Michael Waters. All my
groupmates have provided a welcoming, fun, happy, incredibly intelligent, and won-
derfully quirky atmosphere. Every day I am thankful that I love getting to see the
people in my office. Even more than workmates, they have become my friends.

Thank you to the Applied Physics Program for all your support: Çağliyan, Cyndi,
Lauren, and Chuck. I love our Applied Physics family and the cheerfulness, encour-
gagement, and girl talk the staff provides every time I walk into the AP office.

I would like to acknowledge several friends who had a particularly important
impact on my graduate experience. First, I would like to thank Mac, who was my
first graduate student friend and who proved to be a constant source of encouragement
and interesting discussions throughout my studies. It’s remarkable how much he has
been able and willing to help me despite us being in different fields, whether it be with
homework, Matlab, LaTeX, X-ray data sources, investing, car maintenance, podcasts,
or dishes. Second, I thank Heather, who has been my dear friend, workout buddy,
and whose natural generosity made her one of the people who was most essential in
helping me recover from my concussion; she gave me hours and hours of her time and
welcomed me into her home, and for this I am forever grateful. Third, I would like to
thank Nancy, a fabulous lady who has always been a genuine and fun partner in crime
in Ann Arbor and at the beamline. I am incredibly thankful that she was willing to
dedicate over a week straight of her time to travel to help with my experiments during
my concussion recovery. Lastly, I would like to thank Marika. Marika has listened to
many, many, many stories about the good and the less than good in my life, and has
always provided thoughtful, empathetic, and encouraging feedback. She has been one
of the most reliable and impactful sources of support during this graduate journey
and has enriched my life with her love of learning, natural positive attitude, and easy
joy. I can’t imagine my graduate experience without each of these friends.

I would like to thank my family for their endless love and support. My dad instilled
in me a love of science, a drive to excel, and a desire to question and understand the
world around me. My mom has been a genuine example of kindness and patience,
who has taught me to appreciate the beauty in the world. My sister, Andrea, the twin
I never had, has worked around my crazy schedule so that we could visit each other
and have much-needed nights of cheesy movies, take-out Thai food, and girl talk. My
family has demonstrated their unwavering faith in there being no limits to what I can
achieve, has always expressed so much pride in my efforts and accomplishments, and
has been a constant source of unconditional love.

Finally, my expression of thanks would never be complete without acknowledg-
ing my soon-to-be-husband, Justin. Justin has brainstormed ideas with me, worked through code with me, helped me with LaTeX, empathized with me, and has always been incredibly understanding of the demands of my work. He never fails to push me to be the best researcher and the best person I can be. His dedication to critically reading and re-reading my NSF fellowship essays was one of the largest factors in bringing them to the level that earned me the NSF fellowship award, which is one of my most meaningful academic accomplishments. His love, support, thoughtfulness, understanding, discipline, drive, and natural gift for humor during everyday situations inspire me every day and make my life whole.

Despite this section being so long, this list is certainly not exhaustive, and I regret that I could not include everyone and all the ways in which they provided support to me these past six years.

Funding acknowledgements

The projects throughout this work were funded through various sources. C. M. Jones gratefully acknowledges support from the National Science Foundation (NSF) Graduate Research Fellowship Program (Grant No. DGE-1256260), the Applied Physics Program, the University of Michigan’s Rackham Merit Fellowship, the Rackham Graduate Student Research Grant, and Rackham Conference Travel Grants. Additional sources of funding include the NSF program in Electronic and Photonic Materials through Grant No. DMR-0906909 (experimental work), the Office of Science at the Department of Energy through Grant No. DE-FG02-06ER46273 (experimental work), and the NSF CAREER Award through Grant No. DMR-1254314 (computational work). Use of the Advanced Photon Source was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. X-ray data was collected on the APS Surface Scattering & Microdiffraction beamline 33-ID-D at the Advanced Photon Source, Argonne National Laboratory. This research used resources of the National Energy Research Scientific Computing (NERSC) Center, a DOE Office of Science User Facility supported under Contract No. DE-AC02-05CH11231. Finally, we acknowledge the Michigan Center for Materials Characterization and the University of Michigan College of Engineering for the use of the Veeco Dimension Icon Atomic Force Microscope, the FEI Nova 200 Nanolab SEM/FIB (NSF Grant No. DMR-0320740), and the JEOL 3100R05 Double Cs Corrected TEM/STEM (NSF Grant No. DMR-0723032).
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................................. ii
LIST OF FIGURES ..................................................................... ix
LIST OF TABLES ..................................................................... xv
ABSTRACT .............................................................................. xvi

CHAPTER

I. Introduction ................................................................. 1

1.1 Nanostructures ........................................................... 1
1.2 InGaN and GaN for solid-state lighting applications ....... 2
   1.2.1 Polarity of nitride structures ................................. 3
   1.2.2 White LED technologies ...................................... 4
   1.2.3 Dominant recombination processes and the $ABC$ model 9
   1.2.4 Efficiency droop issue ........................................... 10
   1.2.5 Green gap issue .................................................. 12
1.3 GaAsSb/GaAs for optical communications .................. 14
1.4 Organization of thesis .................................................. 16

II. Experimental and computational methods ....................... 18

2.1 Experimental Methods ................................................ 19
   2.1.1 Introduction to techniques .................................... 19
   2.1.2 X-ray diffraction .................................................. 20
   2.1.3 Atomic force microscopy (AFM) ........................... 41
   2.1.4 Scanning transmission electron microscopy (STEM) 42
2.2 Introduction to computational techniques ...................... 45
   2.2.1 $k \cdot p$ theory ..................................................... 45
   2.2.2 Theory of recombination for localized, envelope wave functions 52
### III. Surface studies of GaN QDs grown using droplet epitaxy on bulk, native substrates

3.1 Background and motivation ........................................... 59
   3.1.1 Introduction ................................................. 59
   3.1.2 GaN substrate growth methods .............................. 59
   3.1.3 GaN quantum dots ........................................... 62
   3.1.4 Quantum dot growth methods ............................... 62
3.2 Project objectives .................................................. 63
3.3 Bulk GaN: bare substrate .......................................... 64
   3.3.1 Dow Corning substrates .................................... 64
   3.3.2 Kyma substrates ........................................... 69
3.4 GaN QDs on bulk GaN ............................................... 73
3.5 Rapid thermal anneal ............................................... 83
3.6 Discussion on bulk GaN substrate quality and availability .. 86
3.7 Conclusions .......................................................... 87

### IV. Impact of carrier localization on recombination in InGaN quantum wells and the efficiency of nitride light-emitting diodes: insights from theory and numerical simulations

4.1 Motivation .......................................................... 89
4.2 Project objectives .................................................. 90
4.3 Computational methods and analysis ................................ 91
4.4 Wave function and recombination rates .......................... 93
4.5 Comparison to fine grid and rough alloy composition ........ 97
4.6 Carrier density and temperature dependence .................... 100
4.7 Comparison to published work .................................... 102
4.8 Effects of carrier localization .................................... 104
4.9 Discussion .......................................................... 104
4.10 Conclusions and future work ..................................... 107

### V. Strain relaxation and band alignment of GaAsSb/GaAs quantum wells

5.1 Motivation .......................................................... 109
5.2 Project objectives .................................................. 110
5.3 Methods ............................................................ 111
5.4 Band edge and alignment results .................................. 111
5.5 Band alignment type transition .................................... 114
5.6 Comparison to published experimental work ..................... 115
5.7 Discussion and conclusions ....................................... 116

### VI. Summary, implications, and future work

119
6.1 Surface studies of GaN QDs on bulk GaN ................. 119
6.2 Effect of alloy fluctuations on recombination in InGaN QWs . 121
6.3 Strain and band alignment of GaAsSb/GaAs QWs ........ 123
6.4 Concluding words ............................................. 124

BIBLIOGRAPHY ................................................. 125
LIST OF FIGURES

FIGURE

1.1 The capabilities of InN, GaN, and AlN to produce alloys that span the visible spectrum. .......................... 3
1.2 Illustration of the wurtzite polar (0001), non-polar (1-100, -1-120), and semi-polar (11-2-2, 20-21, 20-2-1) planes. .......................... 4
1.3 Illustration of polarization in In$_{0.15}$Ga$_{0.85}$N quantum wells (top) and how the polarization separates carriers (bottom). (a) The polarity of the crystal structure creates spontaneous polarization charges at the QW interfaces, and the compressive strain exerted on the QW causes piezoelectric polarization. The net polarization across the well separates carriers in the growth direction. (b) With no net polarization, the carrier wave functions overlap to a high degree and in the center of the well. ........................... 5
1.4 (a) Schematics of two prominent white LED technologies. Nitride-based LEDs produce a white spectrum through partial down-conversion of blue or UV light by phosphors. (b) Comparing the spectra and color rendering index (CRI) of these two LED technologies to that of sunlight, which has a perfect CRI value of 100. ........................... 6
1.5 (a) Schematic of white, solid-state LED based on direct color mixing. (b) Illustration of the color mixing required to produce white light using the additive (RGB) color model used for light. ........................... 8
1.6 Three dominant recombination processes in nitride LEDs: (a) Shockley-Read-Hall (SRH), (b) radiative, and (c) Auger recombination. SRH and Auger are non-radiative recombination processes. ........................... 9
1.7 (a) Direct and (b) indirect Auger recombination. ......................... 10
1.8 Efficiency droop behavior of nitride-based LEDs, where efficiency peaks at low current density and droops with high current density. ......................... 11
1.9 Maximum achieved external quantum efficiency (EQE) of commercial devices, which use multiple quantum wells (circles) and single quantum well devices (stars). The lack of high efficiency devices across the 500~600 nm range is coined the “green gap.” ........................... 13
1.10 Illustration of conduction and valence band edges in type-I and type-II band alignments. ........................... 14
2.1 Schematic of the Bragg’s Law representation of X-ray diffraction.

2.2 Graphical representation of the convolution of two functions, \( f(x) \) and \( g(x) \) for two different cases. \(^9\)

2.3 Describing the crystal lattice, atomic basis of one atom, and measurable Bragg peak signals using Fourier transforms and the convolution theorem for an infinite, ideal bulk crystal with a single-atom basis. \(^9\)

2.4 The origin of Bragg rods, also called crystal truncation rods, understood using the convolution theorem. \(^9\) The top illustrates multiplying an infinite crystal by a step function in order to obtain a truncated crystal with an infinitely sharp surface. The bottom illustrates the same construct in reciprocal space, where the diffraction pattern of an infinite crystal is convoluted with the Fourier transform of a step function, \( 1/q_z \)-shape, in order to obtain crystal truncation rods.

2.5 Schematic illustrating the effect on Bragg rods of a rough crystal surface as compared to an ideal truncated surface. \(^9\)

2.6 Schematic illustrating the effect on Bragg rods of translating the top atomic layer of the crystal by \( \pm 5\% \) and \( \pm 10\% \) of the out of plane lattice constant. \(^9\)

2.7 The Advanced Photon Source systems map, illustrating the main components of the facility. \(^10\) Electrons are first accelerated in the linac before being injected into the booster synchrotron. They are then further accelerated and injected into the main storage ring, which contains a beam current of 100 mA. Experiments are performed at the 70+ beamlines in the experiment hall at research stations. The beam for these experiments is created by insertion devices or at bending magnets.

2.8 Example of COBRA results: electron density of Si-capped Ge/Si QDs as a function of sample height with beam energy 11.08 keV. \(^11\) Sharp peaks indicate a highly ordered crystal, and spatially extended peaks indicate atomic layer bowing. The nominal substrate surface occurs at \( z = 0 \) nm.

2.9 The algorithm used in COBRA to iteratively solve for the phase of the measured structure in order to reconstruct the sample electron density.

2.10 Graphical representation of COBRA’s method of determining the correction amplitudes and phases, \( U \). \(^9\) The method is essentially the completion of matrix addition, where the known values are the magnitudes of \( T_1 \) and \( T_2 \) and the magnitudes and directions of \( S_1 \) and \( S_2 \). (a) and (b) represent values for two adjacent points along the Bragg rod. The correct solution of \( U_a \) and \( U_b \) are the ones that are most similar between (a) and (b) due to the rod being slowly varying.
Illustration of the (a,b) incredibly fast convergence of the COBRA method, and (c,d) the increase in error that can result from employing too many iterations. Notably, the majority of the convergence occurs within the first 10 iterations, as illustrated by the inset of (a), which is the first 15 iterations of (a).

The major components of a scanning transmission electron microscope (STEM). An electron source produces an electron beam at 100-300 keV, which is directed down the column and focused to a spot size $\sim 1 \, \text{Å}$ onto the sample. As the lenses raster the beam across the sample, an annular dark field detector detects Rutherford-like scattered electrons throughout the scan, producing an image that is highly sensitive to atomic mass.

Illustration of the surprising accuracy of a simple effective mass approximation (equation 2.28; black line) in predicting effective mass $m^*$ of common, low-bandgap III-V semiconductors.

Auger matrix elements of GaN for (a) electron-electron- hole ($eeh$) and (b) hole-hole-electron ($hhe$) Auger recombination as a function of the momentum transfer involved in the Auger process. For short-range Auger scattering (i.e., large momentum transfer), which dominates Auger recombination in wide-band-gap nitrides, the dominant matrix elements reach values that are approximately independent of momentum transfer (indicated with dotted line).

AFM images of the Dow Corning bare, bulk GaN substrate. Several roscans (blue line) plotted with their corresponding COBRA model fits (red line) for the Dow Corning bare, bulk GaN substrate. The model fit is very good giving an FOM of 0.0872.

Electron density of Dow Corning bare, bulk GaN substrate, represented as two line profiles, one through each of the two gallium atoms in the GaN primitive cell (inset).

Electron density of Dow Corning bare, bulk GaN substrate. (a) Cross-section along $c$-plane, taken within the bulk of the substrate and demonstrating the 6-fold symmetry we expect for GaN. (b) Cross-section along $a$-plane, illustrating electron density drop-off and structure change above substrate surface.

AFM images of a Kyma bare, bulk GaN substrate, Kyma4a. Several roscans (blue line) plotted with their corresponding COBRA model fits (red line) for the Kyma bare, bulk GaN substrate (Kyma4a). The model fit is good, giving an FOM of 0.1434.

Electron density line profiles for the bare substrate Kyma 4a along the $c$-axis (a) and cross-section along $a$-plane (b), illustrating a very rapid intensity drop-off above the substrate surface (nominally $z = 0$ Å).
3.8 AFM images of as-deposited GaN QDs grown on Kyma bulk GaN substrates. All images were taken of a 500 nm × 500 nm surface area. Sample nos. 1-5 correspond to growth temperatures 550°C, 250°C, 185°C, 185°C, and 185°C, respectively. Sample nos. 1-4 growths were performed with medium brightness nitrogen plasma conditions, and sample no. 5 was grown with a high-brightness nitrogen plasma condition.

3.9 Several rodscans (blue line) plotted with their corresponding COBRA model fits (red line) for sample no. 1 with GaN QDs on a Kyma bulk GaN substrate (RMBE1145, Kyma4b). The model fit is good, giving an FOM of 0.1225.

3.10 Electron density line profiles for the five QD samples on bulk GaN. Only sample no. 3 (RMBE1197) demonstrates significant coherence.

3.11 STEM micrographs of sample no. 1 (RMBE1145, Kyma4b) viewed from the m-plane, demonstrating limited coherence between QDs and substrate.

3.12 STEM micrographs of sample no. 3 (RMBE1197) viewed from the a-plane, demonstrating a large degree of coherence between QDs and substrate.

3.13 AFM images of as-deposited (left) and annealed (right) GaN QDs grown on Kyma bulk GaN substrates. All images were taken of a 500 nm × 500 nm surface area.

3.14 Electron density line profiles for QD samples as-deposited (left) and annealed (right). No significant difference in coherence is observed before and after annealing.

3.15 Two representative photo images of our X-ray detector signal during sample alignments at Argonne National Lab. We observed textured rings instead of Bragg peaks, indicating polycrystalline instead of single-crystal substrates. As a result, we were not able to measure rodscans. (a) location where we would expect the (102) peak for QD sample RMBE1266, and (b) location where we would expect the (113) peak for QD sample RMBE1267.

4.1 (a,b) A randomly generated, local InGaN distribution in a nominal In$_{0.15}$Ga$_{0.85}$N polar quantum well (z || c) at (a) the xy-plane (c-plane) at the QW center (z = 1.5 nm) and (b) a cross-section in the xz-plane (y = 0 nm). Electron (c; z = 2.3 nm) and hole (d; z = 0 nm) wave functions for the alloy distribution in (a,b).
4.2 Individual electron-hole wave-function overlaps between the first 40 electron and 40 hole states, illustrating the symmetry breaking introduced by alloy fluctuations. (a,b) Wave-function overlaps for a virtual-crystal alloy in the (a) polar and (b) non-polar orientation. (c,d) Wave-function overlaps for a fluctuating alloy in the (c) polar and (d) non-polar orientation. The weighted-averaged matrix element is larger for the fluctuating alloy because the reduction of individual matrix elements by spatial separation is overcompensated by the larger number of allowed electron-hole combinations due to translational symmetry breaking.

4.3 Radiative ($B$) and Auger ($C_{eeh}$, $C_{hhe}$) coefficients of InGaN calculated as a function of indium concentration for polar and non-polar orientations ($10^{19}$ cm$^{-3}$ carrier density, 300 K carrier temperature). In all cases, alloy composition fluctuations increase the coefficients compared to virtual-crystal alloys.

4.4 (a,b) A random alloy distribution for a 3 nm polar In$_{0.15}$Ga$_{0.85}$N QW, representing a more atomistic approach (0.3 nm grid and no spatial averaging) than the main study (0.6 nm grid and spatial averaging). (a) shows the $xy$-plane at the QW center, $z = 1.5$ nm, and (b) shows a cross-section at $y = 0$ nm. (c,d) The square modulus electron and hole ground-state wave functions, respectively, for the random alloy distribution shown in (a,b). Each wave function is shown at their most intense, respective $xy$-planes ($z = 2.7$ nm for electron, $z = 0.3$ nm for hole). The electron is less confined than our studies with a coarser (0.6 nm) grid and spatial averaging. Calculations of $B$, $C_{eeh}$, $C_{hhe}$ uphold our main results of Auger processes increasing by a larger factor than radiative rates, resulting in alloy fluctuations decreasing internal quantum efficiency as compared to virtual-crystal alloys.

4.5 Radiative ($B$) and Auger coefficients ($C_{eeh}$, $C_{hhe}$) of In$_{0.15}$Ga$_{0.85}$N as a function of carrier density (300 K carrier temperature). In all cases, composition fluctuations increase Auger coefficients more than the corresponding radiative coefficients.

4.6 Radiative ($B$) and Auger coefficients ($C_{eeh}$, $C_{hhe}$) of In$_{0.15}$Ga$_{0.85}$N as a function of carrier temperature ($10^{19}$ cm$^{-3}$ carrier density). Except for very low temperatures, Auger coefficients of fluctuating alloys increase over those of virtual crystals by larger factors than radiative coefficients.

4.7 (a,b) Internal quantum efficiency (IQE) of In$_{0.15}$Ga$_{0.85}$N quantum wells with fluctuating alloys (dashed line) and virtual crystals (solid line). Fluctuating alloys exhibit lower IQE than virtual crystals. (c,d) Ratio of radiative ($B$) to Auger ($C$) coefficients as a function of indium concentration (300 K carrier temperature, $10^{19}$ cm$^{-3}$ carrier density). The ratio decreases with indium concentration faster for fluctuating alloys than for virtual crystals, consequently alloy fluctuations exacerbate the green-gap problem.
5.1 Band-alignment diagrams of (a) strain-relaxed (type-I) and (b) pseudomorphic (type-II) GaAs0.65Sb0.35/GaAs QWs. Solid black lines denote the band edges and dashed blue lines show the energy levels of confined states. Energy levels are defined with respect to the unstrained GaAs valence band maximum.  

5.2 Conduction and valence band edges for pseudomorphic (solid lines) and strain-relaxed (dashed lines) GaAsSb/GaAs QWs as a function of Sb mole fraction and QW width. Band edges are defined by the first confined energy level, if present. The horizontal black line indicates the conduction band edge of GaAs; QW conduction band edges below (above) this line result in type-I (type-II) band alignment. Energy levels are defined with respect to the unstrained GaAs valence band maximum.  

5.3 Conduction and valence band edges for 7 nm GaAsSb/GaAs QWs as a function of Sb mole fraction. Band edges are shown for strain-relaxed and pseudomorphic wells, as well as for partially strain-relaxed wells with 20%, 40%, 60%, and 80% of pseudomorphic strain relaxed. The right panel zooms in on the region enclosed by the dashed box in the left panel.  

5.4 The percentage of pseudomorphic strain relaxation for which the type-II to type-I band-alignment transition occurs in GaAsSb/GaAs QWs with a well width of 3, 5, 7, and 10 nm.
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Substrate crystallographic miscut, DE growth parameters, QD morphology, and QD coherence results as determined by COBRA for samples of GaN QDs grown on bulk GaN.</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>QD morphology and coherence results as determined by COBRA for samples of GaN QDs grown on bulk GaN before and after annealing.</td>
<td>83</td>
</tr>
<tr>
<td>4.1</td>
<td>Weighted-averaged overlap results of original In\textsubscript{0.15}Ga\textsubscript{0.85}N QW study with 0.3 nm grid and spatially averaged alloy composition.</td>
<td>99</td>
</tr>
<tr>
<td>4.2</td>
<td>Weighted-averaged overlap results of original In\textsubscript{0.15}Ga\textsubscript{0.85}N QW study with 0.6 nm grid and spatially averaged alloy composition.</td>
<td>99</td>
</tr>
<tr>
<td>5.1</td>
<td>Comparison of our calculated results for the band alignment type and optical transitions in GaAsSb/GaAs QWs with published experimental work.</td>
<td>117</td>
</tr>
</tbody>
</table>
ABSTRACT

Semiconductor nanostructures such as quantum wells (QWs) and quantum dots (QDs) have revolutionized solid-state devices through their unique electronic, optical, and strain properties. To understand how best to engineer nanostructures and optimize devices, we must understand their structural and optoelectronic properties as well as the influence of structural properties on optoelectronic behavior. The combined insight of experimental and computational techniques offers a powerful basis for correlating structural and optoelectronic properties of semiconductor nanostructures. In this dissertation project, we have employed both experimental and computational techniques to study epitaxial III-V nanostructures for light-emitting diode (LED) applications. In particular, we have focused on the gallium nitride (GaN), indium gallium nitride (InGaN), and gallium arsenide antimonide (GaAsSb) material systems.

GaN and its applications in LEDs play an integral part in efficient, solid-state lighting, as evidenced by its recognition in the 2014 Nobel Prize in Physics. To push this technology towards higher efficiency, improved reliability, and lower cost, we must understand device growth on bulk GaN substrates, which have lower defect densities and strain than template GaN grown on foreign substrates. We present our experimental findings on the surface properties of GaN QDs grown on commercial, bulk crystal wafers of GaN. This is the first detailed structural study of homoepitaxial GaN QDs. The QDs are grown using the droplet epitaxy method and characterized
using surface X-ray diffraction and an analysis technique called Coherent Bragg Rod Analysis (COBRA), which uses phase retrieval to reconstruct atomic positions near the substrate surface with sub-Angstrom resolution. We find that while several QD growth conditions in our study produce dense QDs, COBRA reveals that the GaN QD sample grown with the lowest nitridation temperature demonstrates significant coherence with the substrates, while all other samples demonstrate limited coherence. Results are supported with atomic force microscopy and scanning transmission electron microscopy. This work allows us better understand surface properties of bulk GaN and how to produce coherent QDs for highly efficient LEDs in the visible spectrum.

We build upon our understanding of materials for III-nitride LEDs by using computational methods to examine the effect of carrier localization due to random alloy fluctuations on recombination rates in InGaN quantum wells. In particular, we focus on radiative and Auger recombination mechanisms. We explore these recombination rates as a function of alloy composition, crystal orientation, carrier density, and temperature. Allowed energy states and wave functions are determined using Schrödinger-Poisson calculations within the framework of the semi-empirical $k \cdot p$ method. Our results show that alloy fluctuations reduce individual transition matrix elements by the separate localization of electrons and holes, but this effect is overcompensated by the additional transitions enabled by translational symmetry breaking and the resulting lack of momentum conservation. Hence, we find that localization increases both radiative and Auger recombination rates, but that Auger recombination rates increase by one order of magnitude more than radiative rates. Furthermore, we demonstrate that localization has an overall detrimental effect on the efficiency-droop and green-gap problems of InGaN LEDs, which has significant technological implications for the performance and design of nitride-based commercial LEDs and lasers.

xvii
In addition to nitride nanostructures, we apply computational methods to GaAsSb quantum wells in GaAs, another III-V material system. GaAsSb/GaAs quantum well-based LEDs are of great interest for several applications, including optical communications; however, their band alignment properties are not fully understood, particularly at 35% Sb alloy concentration used for emission at 1.3 µm. We use numerical simulation methods based on the $8\times8 \mathbf{k} \cdot \mathbf{p}$ theory to explore the effects of GaAsSb/GaAs quantum well composition, width, and strain on the band alignment. Strain-relaxed wells demonstrate type-I alignment and pseudomorphic wells demonstrate type-II alignment, regardless of quantum-well composition or thickness for wells wider than 3 nm. For partially strain-relaxed wells, we determine the band alignment as a function of the interplay of composition, width, and strain. Our calculated results at various strain conditions agree well with published experimental data. This work provides insight on band alignment of GaAsSb/GaAs quantum wells, as well as that of embedded quantum dots with strong confinement along the out-of-plane direction.
CHAPTER I

Introduction

1.1 Nanostructures

Nanostructures have transformed the development of electronic and optoelectronic devices through fascinating and highly applicable material properties not observed in bulk materials. The physical confinement created by reducing dimensions from bulk to nanoscale results in quantized energy levels, which greatly affect the electronic band structure and optical properties.

Quantum wells, which confine electrons in a single direction, are easily fabricated and used in many current commercial technologies, including light-emitting diodes (LEDs) and lasers. Their optoelectronic properties largely depend on QW thickness, composition, and strain. In fact, in this work we systemically explore these properties for the GaAsSb material system to predict band edge alignment properties. Additionally, we explore the effect of statistical variations in composition and crystal orientation on allowed states and recombination rates for the InGaN material system.

Similar to QWs, the optoelectronic properties of quantum dots (QDs), which are zero-dimensional nanoscale structures, can be tuned using QD shape, size, composition, and strain. They can be grown in a variety of shapes, including lenticular, pyramidal, and conical, and they have already demonstrated effectiveness in applications such as LEDs, photodetectors, solar cells, and thermoelectrics.\textsuperscript{16–21}
An important feature of embedded nanostructures in semiconductor devices is the quality of the interface between the nanostructure and its surrounding material. Defects at the interfaces can lead to non-radiative recombination in the form of the Shockley-Read-Hall mechanism. In order to reduce defects and dislocations, it is imperative to form a continuous crystallographic interface between materials. We use “coherence” to describe this continuous crystal behavior across material boundaries. In a coherent structure, the atoms of the nanostructure are atomically registered with the substrate. This is also referred to as an epitaxial structure. We focus much of our experimental characterization efforts in this work on determining the crystallographic coherence of quantum dots with the substrate in order to gain understanding that will allow for optimizing QD-based device performance.

1.2 InGaN and GaN for solid-state lighting applications

Indium gallium nitride (InGaN) is a particularly important material for LED applications because its direct band gap can be tailored to span the entire visible range by varying indium concentration (Figure 1.1). Additionally, it is currently the only material that yields commercially viable LEDs in the blue end of the visible light spectrum. This makes it an excellent material for use in white light sources, where it has revolutionized the solid-state lighting industry.

Before the introduction of new LED light sources, incandescent bulbs, fluorescent tubes, and compact fluorescent lights dominated the lighting markets. However, their efficiencies (luminous efficacies) leave much to be desired at 2-5% (15 lm/W), 25% (60-100 lm/W), and 20% (35-80 lm/W), respectively. White LED bulbs, which offer efficiencies \( \sim 50\% \), are not only significantly more efficient than these conventional light sources, but also last about 25\times longer than traditional incandescent bulbs and dont require special recycling due to toxic mercury content like fluorescent bulbs. By incorporating solid-state lighting into the market, the US has saved more than $2.8
billion to date, as reported by the Department of Energy (DOE) in a 2017 report, and almost 95% of its potential has yet to be utilized. The continued adoption of solid-state lighting technologies could reduce US lighting electricity use by 75% by 2035, saving the equivalent energy needs of 45 million homes for an annual total savings of $50 billion (in today’s dollars). This would have tremendous beneficial impacts on carbon reductions and energy security.

1.2.1 Polarity of nitride structures

InGaN as an LED material offers incredible advantages, as described above, but also has some inherent issues. InGaN and its binary components, GaN and InN, have a wurtzite crystal structure, which exhibits both spontaneous and piezoelectric polarization. Figure 1.2 illustrates the polar, non-polar, and semi-polar orientations of this crystal structure. When nanostructures or interfaces are produced using these materials along a polar or semi-polar growth direction, polarization charges form at the interfaces, which result in strong electric fields across the material that separate carriers (Figure 1.3). This is termed the quantum-confined Stark effect (QCSE). The
decreased wave function overlap resulting from this separating electric field poses issues for efficient radiative recombination. Typically, individual quantum well thicknesses are therefore limited to approximately 3 nm so as to prevent the complete separation of the electron and hole wave functions across the well. One method of addressing the QCSE issue is to grow devices along a non-polar orientation to avoid polarization charges at interfaces that separate carriers. Despite the possible benefits, devices grown on a non-polar surface are unlikely to penetrate commercial markets due to the industry’s investment in c-plane devices with established manufacturing equipment and processing.

1.2.2 White LED technologies

Currently, commercial InGaN LEDs utilize phosphor conversion processes to generate efficient white light. UV, violet, or blue light is generated using InGaN LEDs,
Figure 1.3: Illustration of polarization in In$_{0.15}$Ga$_{0.85}$N quantum wells (top) and how the polarization separates carriers (bottom). (a) The polarity of the crystal structure creates spontaneous polarization charges at the QW interfaces, and the compressive strain exerted on the QW causes piezoelectric polarization. The net polarization across the well separates carriers in the growth direction. (b) With no net polarization, the carrier wave functions overlap to a high degree and in the center of the well.$^3$
Figure 1.4: (a) Schematics of two prominent white LED technologies. Nitride-based LEDs produce a white spectrum through partial down-conversion of blue or UV light by phosphors. (b) Comparing the spectra and color rendering index (CRI) of these two LED technologies to that of sunlight, which has a perfect CRI value of 100.4 after which the light passes through phosphor particles that down-convert a portion of the light to longer wavelengths. If the LED emission wavelength and the phosphor are chosen appropriately, the spectral combination of the original, LED-emitted light and the phosphor-converted light combine to form what the human eye perceives as white light. Two dominant phosphor-based LED technologies are shown in Figure 1.4a; a blue LED is combined with yellow phosphors, or an ultraviolet LED is combined with either blue and yellow phosphors or red, green, and blue phosphors.

In order to discuss the evaluation of LED technologies for white light applications, it is useful to first define several metrics for judging a white-light source. The first metric is called luminous efficacy. Luminous efficacy is a measure of how efficiently the power going into the device produces visible light. Therefore, it is the ratio of luminous flux that is detectable by the human eye to input power, where the luminous flux is obtained by convoluting the emitted spectrum with the spectral sensitivity of the human eye. Its SI units are lumens per watt. The second metric is the color rendering index (CRI), which represents the ability to reproduce an ideal white-light spectrum. The CRI incorporates the effect of balancing the spectral primary colors, red, blue, and green, in order to create white light. The maximum, “perfect” value
of CRI is 100, which represents the spectrum of the sun. A CRI value of 80 is a typical standard for a spectrum quality that is sufficient for indoor use. The third metric is perhaps the most well-known by consumers: the color temperature. The color temperature arises from considering black body emission and spectrum. Lower temperatures such as 2,000 K produce light that appears more red (or “warm”), and higher temperature such as 7,000 K produce light that appears more blue (or “cold”), where the intermediate range may appear neutral. The color temperature scale ranges from 1,000 to 10,000, but approximate ranges for indoor lighting applications are 2000-3000 K, 3100-4500 K, and 4600-6500 K for “warm,” “neutral” or “cool,” and “daylight” spectrums, respectively. Luminous efficacy, CRI, and color temperature are three primary metrics for evaluating the performance of white-light sources, including white LED light bulbs.

Each of the phosphor-based white LEDs shown in Figure 1.4a has strengths and weaknesses, which can be quantified using the metrics luminous efficacy, CRI, and color temperature. InGaN-based, blue LEDs with yellow phosphors (such as Y3Al5O12:Ce) have high efficacy, making them one of the more cost-effective white LED choices. However, their CRI values are very poor (Figure 1.4b), making them unsuitable for indoor lighting applications. Nitride-based, ultraviolet LEDs with various phosphor mixtures have lower efficacy than the blue LED schemes, but they offer higher CRI values (Figure 1.4b), making them better indoor light sources. Furthermore, these most efficient, commercially available white-light LED technologies unfortunately must rely on phosphor conversion and they therefore necessarily suffer a ~20% conversion loss.

In order to avoid the efficiency limitations and the spectrum quality issues present with phosphor-based white LEDs, it is ideal to utilize direct color-mixing techniques to generate white light. This not only avoids the 20% losses from phosphor conversion, but also allows great flexibility in tuning the color or shade of white light. With
direct color mixing, either two or three LEDs could directly be used to generate the white spectrum, as opposed to one LED with phosphor down conversion. Figure 1.5 provides a schematic of a solid-state LED utilizing color mixing and illustrates the visual primary colors and their mixing behavior. As seen in this diagram, white light can be created using a variety of primary and secondary color combinations. For example, all three primary colors red, blue, and green can create white light. Alternatively, in the correct ratios, a primary and secondary color or four colors (red, green, yellow, and blue) can be blended to generate a white spectrum.

It is important to note that the overall efficiency of a color-mixed device is limited by the lowest-efficiency component. Therefore, in order to create white light using direct color mixing, it is essential that LEDs can emit efficiently across a broad range of the visible spectrum to comprise a balance of white light components. In particular and highly relevant to this work, a component from the green portion of the spectrum is essential for high-quality, efficient white light. However, the efficiency of LED technologies in this spectral region is low. We discuss this concept further in this chapter and throughout this thesis as motivation for our work on the nitride material system and its LED applications.
There are three dominant recombination processes commonly used to describe nitride LED performance (Figure 1.6). The first is the radiative combination process, which involves a single electron and single hole recombining and emitting light. The other two are non-radiative forms of recombination, and therefore are highly undesirable for LED performance. Shockley-Read-Hall (SRH) recombination is recombination mediated by defects, and Auger recombination is a three-carrier process in which an electron recombining with a hole gives its energy to a third carrier instead of emitting light. Auger recombination in which the third carrier is an electron excited high into the conduction band is termed electron-electron-hole (eeh) Auger. Likewise, if the third carrier is a hole excited lower into the valence band, the process is termed hole-hole-electron (hhe) Auger. Auger recombination has been observed and studied in not just nitrides, but a variety of materials with various applications, including infrared devices and scintillators.24–27

Auger recombination can be broken down into several different forms depending on the process by which the electron and hole provide energy to the third carrier (Figure 1.7). Direct Auger occurs when the third carrier receives energy directly from
the recombining electron-hole pair, as the name suggests. Indirect Auger in nitrides can occur through phonon-mediation or alloy-scattering, which provide additional momentum and increase the number of available final states.

A common model used to calculate the internal quantum efficiency, $\eta$, of nitride devices is the $ABC$ model (Eqn 1.1). It calculates efficiency using a simple ratio of radiative recombination to total recombination, where total recombination is represented by the three dominant recombination processes described above: SRH, radiative, and Auger recombination. Each recombination process is modeled as a coefficient, $A-C$, multiplied by a carrier density power; $An$ is the SRH term, $Bn^2$ is the radiative term, and $Cn^3$ is the total Auger term, where the different forms of Auger are added to obtain the total Auger recombination term.

$$\eta = \frac{Bn^2}{An + Bn^2 + Cn^3}$$  \hspace{2cm} (1.1)

1.2.4 Efficiency droop issue

Despite their commercial success and increasing adoption, InGaN devices still suffer from significant issues that restrict their performance. First, InGaN exhibits an “efficiency droop,” where its efficiency is very high in the low power regime, but
efficiency peaks at low current density and droops with high current density.\(^7\)

drops when operated in the ideal, high power and therefore high brightness regime (Figure 1.8). Reducing efficiency droop would allow for brighter, more efficient devices so that fewer LEDs would be required for a given lighting application, resulting in lowered material and manufacturing cost and more compact packages.

Several causes of droop have been proposed and resulted in intense debate, including Shockley-Read-Hall recombination,\(^{28}\) carrier leakage,\(^{29}\) carrier overshoot,\(^{30}\) and Auger recombination.\(^6,7\) Auger recombination is an accepted cause of droop, particularly after pioneering work by Shen et al. (experimental) and Kioupakis et al. (theoretical).

Kioupakis et al. used \textit{ab initio} calculations to show that indirect Auger recombination is the primary mechanism causing droop. Their work highlights two indirect Auger processes: Auger mediated by phonons and Auger mediated by alloy scattering. They found that phonon-mediated Auger contributes more to the droop effect if indium content is low, and alloy scattering becomes the dominant Auger mechanism at higher indium compositions.\(^{31}\)
Importantly, Kioupakis et al. showed that the binary compound GaN itself demonstrates efficiency droop, indicating the effects of alloying are not the only mechanism behind efficiency droop. Once indium is introduced to form InGaN, alloy scattering begins to contribute to indirect Auger recombination. Therefore, the total indirect Auger recombination increases with indium concentration, and the efficiency droop effect is more important for longer-wavelength devices. Our work described in Chapter IV contributes to this by determining that alloy composition fluctuations further exacerbate the efficiency droop.

1.2.5 Green gap issue

Another issue with InGaN LEDs is that it is very difficult to fabricate efficient optoelectronic devices across the full visible spectrum. Efficient blue emission is easily accessible through InGaN-based devices, and efficient red light emission through phosphide-based devices. However, the devices that can emit across the green region of the visible spectrum have low efficiency (Figure 1.9). This problem has been coined the “green gap.” The green gap limits the availability of semiconductor materials for green and white spectrum applications.

There are several causes proposed for the green gap that are commonly discussed. The first is due to the degradation of crystal quality as higher concentrations of indium are incorporated. The lattice mismatch between InN and GaN is very large (11%), and incorporating high percentages of indium, such as the \( \sim 30\% \) indium required for green emission, becomes very difficult and creates a large number of dislocations and defects. This can partially be understood by the fact that typical QW widths are comparable to the critical thickness of green/yellow-emitting InGaN QWs on GaN.\textsuperscript{32} Additionally, the higher indium concentrations require lower growth temperatures, which results in higher point defect and parasitic impurities and further leads to structural deformities.\textsuperscript{33} Despite crystal quality seeming a likely cause of the green
Figure 1.9: Maximum achieved external quantum efficiency (EQE) of commercial devices, which use multiple quantum wells (circles) and single quantum well devices (stars).\(^8\) The lack of high efficiency devices across the 500\~600 nm range is coined the “green gap.”

It is not supported by experimental evidence. The \(A\) constant of the \(ABC\) model does not significantly change over the blue to green spectral range.\(^{34}\)

The other commonly discussed cause of the green gap is the polarization fields spatially separating the electrons and holes across the QW. As more indium is incorporated into the QW to reach the green portion of the visible spectrum, larger polarization charges build at the interfaces, producing a stronger carrier-separating force. The green gap, however, is observed to a similar degree in both polar and semi-polar oriented devices, which challenges polarization fields as the cause of the green gap.\(^{35}\)

In summary, despite the intense discussion and effort put into understanding the green gap, the community has not yet reached a consensus on the primary cause, and research is ongoing. Our work described in Chapter IV supports carrier localization by alloy fluctuations being a contributing cause to the green gap problem.
1.3 GaAsSb/GaAs for optical communications

Gallium arsenide antimonide (GaAsSb) nanostructures in gallium arsenide (GaAs) are interesting from both an applications and a fundamental point of view because they are theorized to demonstrate either type-I or type-II band type alignment (Figure 1.10). In the case of type-I band alignment, the conduction band edge of the QW is lower and the valence band edge is higher than those of the barrier material. Therefore, both the electron and the holes are spatially confined within the QW. In GaAsSb/GaAs type-II alignment, however, the conduction band edge is higher in the QW than the barrier. As a result, the hole is confined within the QW but the electron is spatially separated from the barrier. There is also the intermediate case where the QW conduction band edge is lower than the GaAs conduction band edge, but the conduction band offset is small enough that there is no allowed confined state, so electrons can only occupy extended states. In this work, we define this alignment as type-II because the spatial overlap of electrons and holes is small, similar to strong type-II overlaps.

The ability of GaAsSb nanostructures to exhibit two band alignment types makes them versatile in applications such as LEDs and laser diodes, where type-I is preferred, and in photodetectors and charge-based memory elements, where type-II is preferred. In particular, GaAsSb/GaAs structures are exciting materials for optical communications because at ≈35% antimonide concentration, devices emit at 1.3 µm,
which is at a minimum chromatic dispersion wavelength in optical fibers. This wave-
length dominates short- to mid-distance transmission, an indespensible capability in
our technology-driven society. Other typical semiconductor systems that emit at this
wavelength are GaInAsP and AlGaInAs on InP, and GaInNAs(Sb) on GaAs, with
GaInAsP/InP as the current dominant technology. However, GaAsSb offers advan-
tages in that the material is less complicated as a ternary as opposed to quaternary
alloy. Additionally, GaAsSb avoids indium, which has very low earth-abundance and
is being used extensively in many other technologies such as touch screens electrodes
and solid-state lighting. Furthermore, the type-II band alignment offers interesting
possibilities for band gap engineering not achievable with most device material sys-
tems.

Both Type-I and Type-II band alignment types have been reported in the GaAsSb
experimental literature. In particular, GaAsSb quantum wells (QWs) and unstrained
GaSb QDs have been measured as type-I, and strained GaSb QDs as type-II.\textsuperscript{20,36,37,39,40}
Computational studies on GaSb/GaAs band alignment type have been performed for
pseudomorphic QWs.\textsuperscript{36,41} However, the calculations have focused on either a single
case or a function of one parameter (chemical composition). One study by Pryor and
Pistol calculated band alignment as a function of composition for both pseudomorphic
and strain-relaxed cases; however, the calculation was performed at 0 K.\textsuperscript{42} In addi-
tion to performing calculations at LED operating temperatures, it would be useful
to study the effects of partially strain-relaxed conditions. This would be particularly
insightful for QDs where one direction provides significantly more confinement than
other directions since QDs experience complicated strain conditions.

Zech et al. recently demonstrated type-I band alignment in buried GaSb/GaAs
QDs, indicating the dots are essentially strain relaxed based upon the current pub-
lished work.\textsuperscript{37} However, there may be more subtleties not yet explored computa-
tionally with nanostructure size, composition and strain parameters and their effect on
band alignment type. A better understanding of the mechanisms that determine structural and electronic properties of this system is therefore essential for furthering its application in devices, particularly for QD-based devices.\textsuperscript{43}

1.4 Organization of thesis

This thesis combines experimental and computational studies of structural and optoelectronic properties of several III-V materials for nanostructure-based LED applications. It is organized into the following five chapters.

In Chapter II, we discuss both the experimental and computational methods used throughout this thesis and the relevant physics and background for the methods. The experimental methods focus on structural characterization techniques, and the computational methods focus on $k \cdot p$ theory and its application in calculating electronic band properties and wave functions. Additionally, we describe how to use the $k \cdot p$ results to calculate recombination rates.

In Chapter III, we present our experimental characterization of the structural coherence of GaN QDs on bulk GaN substrates. We discuss the results and analysis of our surface X-ray diffraction studies as well as complementary characterization results from atomic force microscopy and scanning transmission electron microscopy.

Chapter IV contains details on computational studies of the localization of carriers in InGaN QWs with statistically random composition fluctuations. We use allowed energies and wave functions from the $k \cdot p$ software package nextnano to calculate radiative and Auger recombination rates as a function of QW composition, crystal orientation, carrier density, and temperature.

In Chapter V we discuss our $k \cdot p$ results exploring quantum well thickness, composition, and strain of GaAsSb QWs in GaAs in order to determine the conditions for type-I and type-II band alignment and for the transition between band alignment types.
Finally, in Chapter VI, we summarize the work described in this thesis, highlight the implications of the work for the field, and discuss suggestions for future studies.
CHAPTER II

Experimental and computational methods

The combination of experimental characterization methods and predictive computational methods offer powerful insight into the structure and optoelectronic properties of semiconductor systems. Surface X-ray diffraction, atomic force microscopy, and scanning transmission electron microscopy all provide data regarding QD morphology and QD coherence with sub-nanometer resolution. Each method’s strengths and weaknesses prohibit each method from providing all necessary information individually, but together they form a complete picture of structure and coherence over a large range of length scales ($10^{-10}$ to $10^{-3}$ m). We use these methods to study GaN nanostructures, details for which will be presented in Chapter III. Furthermore, computational studies add to our knowledge of materials for an incredible variety of applications. Methods such as density functional theory (DFT) and $k \cdot p$ theory allow us to better understand material crystal structure, stability and thermodynamics, elastic properties, electronic properties, and optical properties. Several of many possible examples include UV detection and emission,\textsuperscript{44,45} electrochemical properties of solid electrolytes for batteries and supercapacitors,\textsuperscript{46} high-k dielectrics for devices such as memristors,\textsuperscript{47–49} novel materials for visual light absorption\textsuperscript{50–52} and thermoelectrics,\textsuperscript{53,54} and ultra-thin nanowire\textsuperscript{55} and quantum well configurations.\textsuperscript{44} Computational methods, in particular, can allow us to explore systems or conditions which
may be difficult to control or impossible to achieve experimentally, allowing for understanding of the physical mechanisms behind the resulting behavior and motivating future experimental work. They also build on our knowledge of nanostructures by allowing us to inexpensively explore a wide range of compositions, sizes, and conditions. Of particular interest for this thesis, efficient computational methods are excellent for studying the effects of material conditions such as alloy inhomogeneities and strain on nanostructures. We employ the $k \cdot p$ method to study InGaN and GaAsSb quantum wells, which are described in detail in Chapters IV and V, respectively.

In this chapter, we present the theory, methods, and analysis for experimental and computational techniques used in this thesis. We begin with experimental methods, where we discuss surface X-ray diffraction (SXRD), atomic force microscopy (AFM), and scanning transmission electron microscopy (STEM). In the experimental section, we also discuss our SXRD analysis technique, Coherent Bragg Rod Analysis (COBRA). In the following section, we discuss the semi-empirical computational method $k \cdot p$, for which we focus on $8 \times 8 k \cdot p$ and effective mass techniques. We finish by describing the formalism behind the recombination rate analysis we perform using envelope wave functions obtained from $k \cdot p$ calculations.

\section{Experimental Methods}
\subsection{Introduction to techniques}

In this chapter, we describe the relevant background and primary experimental and analysis techniques used for our experiments evaluating the coherence of GaN QDs on bulk, native substrates. The experimental methods are SXRD with COBRA, AFM, and STEM.
2.1.2 X-ray diffraction

2.1.2.1 Background

X-ray diffraction (XRD) is a powerful, non-destructive technique which utilizes X-ray scattering to determine material properties. It is capable of providing information on crystal structure, lattice constants, crystal orientation, mosaicity, grain size, and coherence, among other properties. Additionally, it requires little to no sample preparation and can probe sub-surface features.

Despite XRD having been first demonstrated more than 100 years ago, XRD is still one of the primary methods used to determine crystal structure properties of new materials. For example, in the past few years it was used to study an emerging new compound semiconductor material, ZnSnN$_2$, an earth-abundant alternative material for solar cells that was synthesized for the first time in 2013.$^{56}$ Senabulya et al. used XRD for the first demonstration of single-crystal, epitaxial ZnSnN$_2$ while identifying the crystal structure and measuring lattice constants with high accuracy.$^{52}$

There are two primary formalisms used to understand XRD. The first is Bragg’s law (Figure 2.1), which is conceptually simple and largely geometrical in nature. Bragg’s law states that when X-rays are elastically scattered off of a crystal lattice plane at specific angles, the scattered beam will interfere constructively and produce a bright signal called a Bragg peak. Under the scattering condition for maximum intensity, also called the constructive Bragg’s law condition, the difference in path lengths that scatter from nearby lattice planes is equal to integer number of wavelengths, allowing us to calculate the lattice spacing. Furthermore, the lattice spacing and Bragg peak index can be used to characterize the lattice point group symmetry. We note that this method mathematically highlights the necessity of using X-rays to measure lattice spacings as opposed to other forms of electromagnetic radiation; the incident wavelength must be comparable to, or smaller than, the lattice spacings of
The second description of XRD is the Laue formulation. Max von Laue described XRD using incident and outgoing plane waves interacting with individual atoms in the crystal, where all outgoing wave directions for a given ingoing wave are considered. The difference in wavevector between the incident and diffracted beam is the scattering vector. If the scattering vector is equal to reciprocal lattice vectors of the crystal, then the conditions for diffraction are satisfied, and we can describe the conditions by the Laue equations:

\begin{align*}
\mathbf{a} \cdot \Delta \mathbf{k} &= 2\pi h \\
\mathbf{b} \cdot \Delta \mathbf{k} &= 2\pi k \\
\mathbf{c} \cdot \Delta \mathbf{k} &= 2\pi l.
\end{align*}

(2.1)

where \(\mathbf{a}, \mathbf{b},\) and \(\mathbf{c}\) are lattice vectors, \(\mathbf{k}\) is the wavevector, and \(h, k,\) and \(l\) are Miller indices. Diffraction peaks are observed when these Laue equations are satisfied, or equivalently, when waves scattered from all lattice points interfere constructively. The Laue diffraction conditions can be reduced to Bragg’s law; therefore, even though the Laue formulation does not require the concept of atoms forming reflection planes, Bragg’s law and the Laue equations are equivalent representations of XRD in a periodic crystal.

The Bragg and Laue equations identify diffraction conditions, but they do not
give the magnitude of the peaks. The atomic form factor describes the intensity of
diffraction from an individual atom, or the scattering cross-section of the atom. It can
be calculated using the Fourier transform of the sample’s electron density distribution,

\[ f_j(q) = \int \rho_j(r)e^{iq \cdot r} dr, \]  

(2.2)

where \( q \) is the wavevector transfer, \( j \) denotes the \( j \)th atom, and \( \rho(r) \) is the electron
density distribution function of the atom. Notably, the diffraction scattering intensi-
ties increases with the atomic mass. This means that the diffraction intensity is very
low for light atoms making them very difficult to detect, and there is little contrast
between atoms with similar atomic number.

When the information obtained by the Laue equations and the form factor are
combined, we can define the structure factor \( F_{hkl} \). For \( j \) atoms in a unit cell for a
single crystal,

\[ F_{hkl} = \sum_j f_j e^{2\pi i(hx_j+ky_j+lz_j)}. \]  

(2.3)

This can similarly be described by an integral when all atoms are considered,

\[ F_{\text{struct}}(q) = \int \rho(r)e^{-2\pi i q \cdot r} dr, \]  

(2.4)

where \( \rho(r) \) is the total electron density for all atoms in the system. Equation 2.4
highlights that the structure factor is the Fourier transform of all the atoms and their
electron distributions in the unit cell. It is an expression that describes diffraction
based on the position of each atom and the scattering form factor of each atom.

Of particular importance to this work is the concept that this complex structure
factor can be decomposed into amplitude and phase components:

\[ F_{\text{struct}}(q) = |F_q|e^{-\phi_q}, \]  

(2.5)
where $|F_q|$ is the amplitude and $\phi_q$ is the phase of the structure factor. The measured diffraction intensity is proportional to the square modulus of the structure factor,

$$I \propto |F_{\text{struct}}|^2 \propto |F_q| e^{-\phi_q} F_q e^{\phi_q}$$

and the electron density of the system can be determined using an inverse Fourier transform of the structure factor. Of course, to perform this inversion directly requires knowledge of the phase factors $e^{-\phi_q}$, and therein lies the challenge of X-ray structure determination. One approach to solve this problem (i.e., not knowing the phase of the scattered X-ray beam), is the COBRA phase-retrieval algorithm specific to the case of epitaxial thin films. COBRA is described in detail in section 2.1.2.5.

### 2.1.2.2 Bragg rods and surface X-ray diffraction

Standard XRD experiments study bulk crystal properties, which follow the XRD theory described above; however, many experimental questions involve studying surfaces, interfaces, and other two-dimensional structures. Understanding these types of structures requires more sophisticated theoretical models. SXRD theory is dependent on the knowledge presented above, that the diffracted intensity is proportional to the square modulus of the Fourier transform of the electron density. Additionally, we present the convolution theorem, which is integral to understanding SXRD theory.

The convolution theorem of two generic functions, $f(x)$ and $g(x)$ is shown below:

$$f \otimes g = \int_{-\infty}^{\infty} f(x') \cdot g(x - x') dx'.$$

The concept of convolution is illustrated in Figure 2.2. It can be thought of the overlap of function $g(x')$ as it is passed along $x$ and across $f(x')$, which is stationary. The result is the integrated area of the product of both functions for each value of
Figure 2.2: Graphical representation of the convolution of two functions, \( f(x') \) and \( g(x') \) for two different cases.\(^9\)

For the specific case of one of the functions being a delta function, as is found in XRD, the convolution has the shape of the other function but at the position of the delta function (Figure 2.2b).

Particularly important for XRD is how the convolution function works with Fourier transforms:

\[
\mathcal{F}(f \cdot g) = \mathcal{F}(f) \ast \mathcal{F}(g).
\]  

(2.8)

The Fourier transform of two functions multiplied by one another is the convolution of the Fourier transform of each function individually.

Before focusing on how these mathematical tools help us understand SXRD, we first describe how they influence traditional XRD and improve our understanding of the diffracted signals from bulk crystals.

Bulk crystals can be described by a unit cell, which serves as a building block that can reproduce the full bulk structure. The unit cell is composed of a lattice, which is a mathematical construct describing the locations of atoms and the symmetry of the crystal, and a basis, which describes the atom(s) associated with each lattice point. The Fourier transform of the lattice is called the reciprocal lattice.

As described by Bragg and von Laue, if waves that scatter off of the crystal interfere constructively, they produce a highly localized, intense signal called a Bragg peak. All Bragg peaks combined, i.e. the diffracted signal, can be mathematically described with a convolution of the crystal lattice and individual atom’s electron
density, and then taking the Fourier transform and squaring the result (Figure 2.3, top). Equivalently, because of the convolution theorem, the diffracted signal can be described as the Fourier transform of the crystal lattice and the Fourier transform of the individual atom’s electron density multiplied and squared (Figure 2.3, bottom). This construct allows us to understand why, even for the case of an ideal, bulk, infinite crystal, we observe the intensity of the Bragg peaks drop off from the experimental origin, causing the signal to be weaker with higher-index peaks.

Using the same concepts applied to a bulk crystal, we can now use the convolution theorem to understand the diffraction properties for the purpose of SXRD. When we measure a sample, X-rays are incident not upon an infinite crystal, but the surface of a truncated crystal. This means that in the z-direction, which we define to be perpendicular to the sample surface, the sample’s periodicity is broken. As a result, the symmetry is broken and the diffraction conditions in the z-direction are relaxed, producing a continuous diffraction signal. Figure 2.4 illustrates this concept of Bragg
Figure 2.4: The origin of Bragg rods, also called crystal truncation rods, understood using the convolution theorem. The top illustrates multiplying an infinite crystal by a step function in order to obtain a truncated crystal with an infinitely sharp surface. The bottom illustrates the same construct in reciprocal space, where the diffraction pattern of an infinite crystal is convoluted with the Fourier transform of a step function, $1/q_z$-shape, in order to obtain crystal truncation rods.

“rods.” Using the convolution theorem, we can understand the crystal surface’s effect on the diffracted signal. In real space, we can produce a truncated crystal by multiplying an infinite crystal by a step function (Figure 2.4, top). In reciprocal space, the truncated crystal’s diffraction pattern can be obtained by convoluting the infinite crystal’s diffraction pattern with the Fourier transform of a step function, a $1/q_z$ shape function. This results in slowly varying diffraction intensity centered around Bragg peaks and producing a continuous signal in the $z$-direction. These extended signals are called Bragg rods, or crystal truncation rods (CTRs).

Because Bragg rods are a direct result of the crystal truncation, the rods are
highly sensitive to properties of the crystal surface. For example, Bragg rods can also provide valuable information on heterostructure properties, such as thin films. Since this work studies homoepitaxy, we refer further reading on those topics to an excellent resource, Christian Schlepütz’s Ph.D. dissertation, which was written with the quality and thoroughness of a text book. More relevant to homoepitaxy are the effects of surface roughness and planar displacement. In the case of surface roughness, we consider the case of a broadened crystal truncation as opposed to an ideal, abrupt surface (Figure 2.5a). The effect of this surface roughness is that the shoulders around the Bragg peaks fall off faster and the valleys between Bragg peaks dip lower than for a perfectly flat surface (Figure 2.5b). We can also consider the effect of shifting the top atomic layer ±10% of the unit cell spacing, which is consistent with atomic plane relaxations. This small shift can produce dramatic effects, as seen between the Bragg peaks in Figure 2.6; the signal between the peaks becomes highly asymmetric. These effects are so dramatic because the bulk signal between Bragg peaks is negligible as a result of destructive interference, so deviations from the bulk periodicity don’t completely destructively interfere and they dominate the signal. These two examples illustrate the information Bragg rods can provide on sample surface structure.

SXRD is an excellent technique for investigating Bragg rods arising from two-
Figure 2.6: Schematic illustrating the effect on Bragg rods of translating the top atomic layer of the crystal by ±5% and ±10% of the out of plane lattice constant.\(^9\)

dimensional crystal features such as surfaces, interfaces, thin films, and heterostructures. However, SXRD is difficult to perform because the signals produced by those from two-dimensional features between the Bragg peaks are significantly smaller than the bulk crystal signals, necessitating very bright X-ray sources and detectors that have very good signal-to-noise ratios. Fortunately, modern synchrotron facilities provide an ideal tool for such measurements.

2.1.2.3 Advanced Photon Source at Argonne National Lab

Very bright X-ray sources in the form of synchrotrons are excellent for SXRD studies. In this work, we use the Advanced Photon Source (APS) for data collection, which is a third-generation X-ray radiation source at Argonne National Lab. For these sources, X-rays are produced using electrons accelerated to relativistic speeds. At APS, electrons are first produced using thermionic emission from a cathode heated to \(~1,000^{\circ}\)C. The electrons are then accelerated to \(>99.999\%\) of the speed of light \((\gamma > 224)\) and with an energy of 450 MeV using a linear accelerator, also called the
Next, electrons are injected into the booster synchrotron ring, in which the electrons are further accelerated to \( \gamma > 7,071 \) and with an energy of 7 GeV within a half-second angular divergence using radiofrequency (RF) cavities. Bending and focusing electromagnets are used along with the RF fields to direct the electrons along the circular path of the booster synchrotron.\textsuperscript{58} The electrons are then injected into the main storage ring, which contains more than 1,000 electromagnets. In order to protect staff and users in the experiment hall outside the ring from stray radiation, the ring is surrounded by a concrete, radiation-proof wall. The electrons accelerate around the nominally circular ring in an aluminum-alloy chamber held under vacuum. The ring is composed of 40 straight sections, or sectors, each of which is 5 meters long and contains two dipoles, ten quadrupoles, and seven sextupoles to direct the beam.\textsuperscript{59} Of the 40 sectors at APS, one is used for injection into the ring, four have RF accelerators to replenish the energy lost due to X-ray emission, and 35 are equipped with insertion devices and used to provide highly collimated X-ray radiation for experiments. Each of these 35 sectors includes two beamlines, one at a bending magnet (BM) and another at an insertion device (ID), which extend into the experiment hall around the ring. This means that at full operation, at least 70 different experiments can be running concurrently. Each beamline includes equipment that is specifically designed for different, specialized applications and types of experiments. The BM lines at each beamline use the stray radiation emitted as the electrons accelerate in an arc between straight sections in the ring. The ID lines at each beamline include an insertion device containing an undulator that accelerates the electrons at different oscillation period lengths (1~6 cm) and numbers of oscillations (40-140 periods) to form a high-quality beam with specific properties for the types of experiments performed at each beamline.\textsuperscript{60} Many of the beamlines operate at a single wavelength, requiring monochromators to select one-millionth of the X-ray energy bandwidth for the experiment at that beamline.
These X-rays are passed down to the experiment hutch, which is lined with lead for radiation protection and contains the sample or system under investigation and detectors to collect diffracted signals after the X-ray-sample interaction.

APS is an impressive facility. We show a diagram of the facility in Figure 2.7, including the components relevant to the operations described above. Operating at 7 GeV, the APS is the highest-energy X-ray source in the Western Hemisphere. The only X-ray producing synchrotron that operates at higher electron energies is the 8-GeV Super Photon Ring (SPring-8) in Japan. Because the electron energy is so high, the APS is capable of providing hard, highly-penetrating X-rays. This high energy is also what determines the large size of the ring; with a circumference of 1,104 m (approximately 0.68 mi), the storage ring is large enough to encircle Chicago’s Wrigley Field.

An interesting fact about the experiment hall at APS is that the floor is one large slab of 1-ft thick, poured concrete. Usually when large areas of concrete are filled, sections are created so as to allow the slabs to move independently and avoid the creation of cracks. The single slab of concrete at APS provides a very stable surface for experiments, allowing for incredibly high alignment accuracy. Regular surveys are performed on the concrete floor and, based on shrinkage, settlement and transitory effects, the floor moves ~0.2 mm per year on average with parts moving up to 6 mm.

The main parameters used to evaluate X-ray source quality are flux and brilliance. Flux is the number of photons per second per unit area, and is used as the metric for experiments that use the entire, unfocused beam. Brilliance is a similar metric, but includes a measure of divergence of the beam, describing the smallest spot size achievable with the focused beam. The beam brilliance produced by third-generation sources like APS are more than $10^{15}$ times the brilliances of conventional lab-source XRD equipment. Maximization of both flux and brilliance are essential factors in
Electrons are first accelerated in the linac before being injected into the booster synchrotron. They are then further accelerated and injected into the main storage ring, which contains a beam current of 100 mA. Experiments are performed at the 70+ beamlines in the experiment hall at research stations. The beam for these experiments is created by insertion devices or at bending magnets.

third-generation synchrotrons and make APS one of the premier synchrotron X-ray facilities in the world.

2.1.2.4 PILATUS detector

Another technology essential to SXRD studies is a high-performing detector. The detector used in this work is a PILATUS, a two-dimensional X-ray detector with 100k pixels. It is a single photon counting detector which uses a two-dimensional array of silicon-based p-n diodes to convert individual photons into an electrical signal. The
use of silicon p-n diodes allows for a high counting rate of $>10^6$ counts/pixel/s, which is ideal for the high fluxes produced by synchrotron sources. An advantageous design to the Pilatus is that it produces no dark current or read-out noise, and the detector has a high dynamic range of $10^6$. An important factor in the performance of the PILATUS is the geometry of the pixels. In a single layer of pixels there are gaps between pixels. To avoid this dead area between adjacent chips, the device includes larger pixels that span the gaps between pixels. This leads to the detector using a “virtual pixel” procedure, and the signal is redistributed to resolve the position of the detected X-ray. The development of two-dimensional detectors revolutionized XRD experiments, allowing for significantly faster and increased data acquisition as compared to one-dimensional, point detectors, which required significantly more scanning to obtain a similar amount of information.

The goal of XRD is to provide information about the crystal structure of our material. To this end, we use the measured diffraction intensities to probe the structure factor for the crystal by taking the square of the intensity:

$$|F_{hkl}| \propto \sqrt{I_{hkl}}.$$  \hspace{1cm} (2.9)

Unfortunately, XRD does not measure the phase information and we can only experimentally determine the magnitude of the structure factor. In order to reconstruct the atomic positions and electron density of the crystal, the real and imaginary components of the structure factor are both required. This problem is coined the “phase problem” in crystallography.

Many methods have been demonstrated to solve the phase problem. One of these methods is simply to “guess and check.” This was most notably used to determine the double-helix structure of DNA, which led to the 1962 Nobel Prize in Physiology or Medicine. This method, however, is rarely suitable, especially in the case of complex systems or structures. Another notable method is the Patterson method, which is
useful for inorganic structures if there is only one heavy atom in the unit cell. It involves taking the Fourier transform of the intensity of the diffraction pattern, called a Patterson map. The heavy atom’s diffraction signal dominates the response, so this information would be used to determine the complex structure factor of the heavy atom alone. The lighter atoms would then appear in the electron density following analysis. The Patterson method was used to determine the structure of vitamin B_{12} and contributed to Dorothy Hodgkin winning the Nobel Prize in Chemistry in 1964.

As the capabilities of computation increased, more so-called direct methods became available. Direct methods simply employ the measured diffracted intensities and a priori knowledge of the system. The phases are then reconstructed using iterative fits, allowing us to reconstruct the electron density of the system. In this work, we use a direct method called COBRA to solve for the phase and reconstruct electron densities.

2.1.2.5 COBRA phase-retrieval analysis

COBRA is a powerful phase-retrieval X-ray technique used to reconstruct the averaged structure and morphology of crystalline thin films and QDs. It was developed by Yizhak Yacoby of Hebrew University in collaboration with our group and first demonstrated electron density reconstructions in 2003.\textsuperscript{65} It has demonstrated ultra-high resolution structure and morphology reconstruction of Ge/Si, InAs/GaAs and InSb/GaAs QDs, among others.\textsuperscript{66,67} As an example, Ge/Si QD electron density results generated using the COBRA technique are shown in Figure 2.8.\textsuperscript{11} In this work, we use COBRA to reconstruct electron densities of GaN QDs grown on bulk, native substrates (Chapter III).

Before using COBRA, we must first perform several corrections to the rodscan data. First, we normalize the two-dimensional detector rodscan images using the detector flat-field signal to correct for inhomogeneities in the detector response. Sec-
Figure 2.8: Example of COBRA results: electron density of Si-capped Ge/Si QDs as a function of sample height with beam energy 11.08 keV.\textsuperscript{11} Sharp peaks indicate a highly ordered crystal, and spatially extended peaks indicate atomic layer bowing. The nominal substrate surface occurs at $z = 0$ nm.

ond, we define a region of interest (ROI) and subtract the background signal from the scans. We take the series of two-dimensional images obtained from the PILATUS detector along the rod, and for each rodscan define the largest ROI that avoids any stray signals on the detector images. Then we integrate along either the horizontal or vertical direction, depending on the specific rod geometry and other features or stray signals in the image. Once we have a one-dimensional profile of the averaged signal for each image, we can fit a line to the background signal and subtract this from the detector image in order to extra the data from the detector images. Next, we apply geometrical corrections to the rodsca, which are dependent on our experimental setup. Our first geometrical correction is for polarization, which is dependent on the beam’s angle of incidence on the sample and has a cosine dependence. Another correction is the rod interception and Lorentz factor, which has a sine dependence because of the angle at which the surface normal and Ewald sphere intersect. Finally, we correct for the area of the X-ray beam incident on the sample surface, or the beam footprint. This correction is substantial since our angle of incidence is very small,
Figure 2.9: The algorithm used in COBRA to iteratively solve for the phase of the measured structure in order to reconstruct the sample electron density.

usually 5°, resulting in a much larger beam footprint than beam cross-section. The correction factor is based on simple geometry and has a 1/sine dependence. Following these corrections, the rodscan data is ready for COBRA.

The basic concepts behind COBRA are the following. COBRA utilizes an initial reference, which is compared to the observed diffraction signal and adjusted using magnitude projection in order to agree better with the observed signal. This adjustment and comparison to experimental data is repeated iteratively to minimize least squares error. It also uses the principle that the complex structure factor is a continuously varying function along the Bragg rod, thus implying that the phase must also be continuous. We now describe the principles behind COBRA in more detail. The algorithm steps are illustrated in Figure 2.9.

We begin COBRA with experimental rodscans and an initial reference structure, which is our model for the anticipated structure. The electron density of the reference and of the real structure differ by an unknown amount. In other words, the total of
the reference electron density and the unknown difference in electron density adds to the electron density of the real structure. Since Fourier transforms are a linear operation, we find a similar expression for the structure factors of the reference \((S)\), unknown difference \((U)\), and real structure \((T; \text{ where } T \text{ denotes “total”})\):

\[
T(q) = S(q) + U(q)
\]  

(2.10)

If we consider a small change in the scattering vector, \(\Delta q\), from either direction of \(q\), we can write the following two formulas for the two structure factors along the Bragg rod:

\[
T(q - \frac{\Delta q}{2}) = S(q - \frac{\Delta q}{2}) + U(q - \frac{\Delta q}{2}),
\]

\[
T(q + \frac{\Delta q}{2}) = S(q + \frac{\Delta q}{2}) + U(q + \frac{\Delta q}{2}).
\]

(2.11)

As stated above, in COBRA we assume that the amplitudes and phases varying continuously and slowly along the Bragg rod. We can therefore assume

\[
U(q - \frac{\Delta q}{2}) \simeq U(q + \frac{\Delta q}{2}) \equiv U_a(q).
\]

(2.12)

We note that this approximation is only valid if the rate of change of \(U_a\) is much smaller than that of \(S\). This is very reasonable in the case of a thin epitaxial film (<10 nm thick). If we plug \(U_a(q)\) into Equation 2.11, we obtain the following equations:

\[
|T_1| \equiv |T(q - \frac{\Delta q}{2})| = |S(q - \frac{\Delta q}{2}) + U_a(q)| \equiv |S_1 + U_a|,
\]

\[
|T_2| \equiv |T(q + \frac{\Delta q}{2})| = |S(q + \frac{\Delta q}{2}) + U_a(q)| \equiv |S_2 + U_a|.
\]

(2.13)

\(|T_1|\) and \(|T_2|\) are amplitudes, which we have stated above are proportional to the square root of the measured diffraction intensities from experiment. The next step, where we approximate a value for \(U_a\), is best illustrated graphically (Figure 2.10a), where we can see there are two solutions to Equation 2.13. The lengths of the vectors
represent the amplitude and the direction represents the phase of the structure factors. We only know the magnitude and therefore vector length of the measured structure factor, $|T_1|$ and $|T_2|$, but we know the amplitude and phase for the reference structure, $S_1$ and $S_2$. Graphically (Figure 2.10a), we first position $S_1$ and $S_2$, where the vectors begin at the same point. We then place $|T_1|$ and $|T_2|$ at the end of the respective $S_1$ and $S_2$ vectors, where we rotate the vectors to find where they intersect; there will be two locations. We then use these two intersection points to determine two possible solutions for $U_1$. In order to determine which solution is more suitable, we must do the same analysis for the next point along the Bragg rod, which is denoted by points 2 and 3 instead of 1 and 2 (Figure 2.10b). The appropriate $U$ values are then determined using the concept discussed at the introduction of COBRA: the amplitude and phase of adjacent points along the Bragg rod are continuous and slowly varying. We can assume this because the epitaxial films are very thin, resulting in the Fourier transform and diffracted signal is very broad and slowly varying. We therefore determine the $U_1$ and $U_2$ values to be the pair that have the smallest difference between Figure 2.10a and Figure 2.10b.

This calculation gives a first guess at the unknown structure factor, which we add to the structure factor of the reference structure. We then take the inverse Fourier transform of the updated reference structure in order to produce an updated total electron density.

Essential to direct methods is applying real-space constraints. Namely, we require the electron density to be positive, and generally localized around atoms in the system without infinite extent. After applying these real-space constraints, we have a new reference structure. We perform a Fourier transform in order to obtain updated Bragg rod for the reference structure, which we compare to the experimental Bragg rod scans for the real structure, for which we calculate a figure of merit (FOM), $r_1$, which is the least squares difference between the reference and experimental rodscans.
Figure 2.10: Graphical representation of COBRA’s method of determining the correction amplitudes and phases, $U$.\textsuperscript{9} The method is essentially the completion of matrix addition, where the known values are the magnitudes of $T_1$ and $T_2$ and the magnitudes and directions of $S_1$ and $S_2$. (a) and (b) represent values for two adjacent points along the Bragg rod. The correct solution of $U_a$ and $U_b$ are the ones that are most similar between (a) and (b) due to the rod being slowly varying.
We can then restart the calculations described above using the updated reference structure. The calculation continues iterating for \(\sim 10\) iterations. After a certain point, noise contributions start to dominate the result and \(r_1\) values start to increase; the calculation must be manually interrupted before this point in order to obtain the best fit.

One disadvantage to COBRA is that because of the iterative scheme, the algorithm may relax into a relative minimum as opposed to a global minimum, leading to an inaccurate result. In order to minimize the possibility of this, the initial reference structure must be chosen very carefully in order to best describe the real structure. In fact, it is prudent to test several structures and evaluate the reliability of each one before determining one as a solution.

Understanding the form of the electron density is essential to interpreting the results. COBRA’s electron density results are in the form a folded structure. The electron density of each atom is folded laterally so that all of the in-plane lattice information is contained within a single unit cell. This means that the results show averaged information in-plane, and therefore provide the most useful information in the out-of-plane direction. If all atoms are highly periodic and commensurate with the substrate, then the electron densities appear highly localized and sharp at their designated lattice points. However, if there is some relaxation or deviation from bulk periodicity, the electron density will appear to spread, while localized at the average lattice site. This could indicate lattice plane bowing, if the smearing is out-of-plane, or structural relaxation/surface reconstruction if the smearing is in-plane.

A benefit of COBRA over other direct methods is its incredible speed because of the direct phasing step; it reaches convergence in \(3\sim 12\) iterations with a total calculation time of approximately 30 minutes, as demonstrated in the inset of Figure 2.11a. In fact, using more iterations can result in noise beginning to dominate the results and give worse agreement between model and experiment.
Figure 2.11: Illustration of the (a,b) incredibly fast convergence of the COBRA method, and (c,d) the increase in error that can result from employing too many iterations. Notably, the majority of the convergence occurs within the first 10 iterations, as illustrated by the inset of (a), which is the first 15 iterations of (a).
The detailed insight that COBRA gives on 2D properties of semiconductor materials makes it an excellent method for studying the coherence of GaN QDs on bulk GaN substrates. Its sensitivity to structural coherence can provide unparalleled insight on epitaxial QD coherence over a large surface area relative to other techniques.

2.1.3 Atomic force microscopy (AFM)

Atomic force microscopy (AFM) is a scanning probe technique that maps the surface topography of a sample with resolution capabilities of less than a nanometer. It was developed in 1982 in order to extend scanning tunneling microscopy (STM), which earned its developers the 1986 Nobel Prize in Physics, to non-conductive materials. AFM is now one of the most important and prevalent surface imaging tools. Its operation is based on rastering a cantilever with a tip radius $\ll 10$ nm across the sample using piezoelectric scanners. A laser reflected off the tip of the cantilever onto a detector is used to determine the tip deflection and therefore the force between the cantilever and the sample surface during the scan. The variation in force throughout the scan is what is used to construct the surface image.

Advantages of AFM over other methods include that it produces a 3-D profile of the sample surface across $\sim 100$ nm-$100$ µm, it can provide incredibly high resolutions comparable to STM and TEM, that the sample does not require any special preparation or treatments for imaging, and that the measurement can be performed in ambient pressure conditions. While it has a larger scanning range than TEM, it has a smaller range than scanning electron microscopy. The disadvantages of standard AFM operation include slow scan speeds, which can lead to thermal drift, and an inability to probe structures below the surface. For these reasons, AFM is an excellent tool for verifying QD formation and measuring QD morphology, but other techniques are necessary to study the internal structural coherence of the QDs.

Several AFM operational modes are available, including contact, non-contact, and
tapping mode. For this work, we use tapping mode, which provides higher resolution than non-contact mode and minimal sample damage, unlike contact mode. Tapping mode is when the AFM tip intermittently makes contact with the sample, which relies on the tip oscillating at a well-defined resonant frequency. The force of each tip contact, or tap, with the sample is higher than in contact mode, but the overall damage is reduced compared to contact mode because the tip is less likely to drag or scrape at features of or contamination on the surface.

The AFM images shown in this work were taken by either the author of this work or one of our collaborators, Dr. Sunyeol Jeon. AFM image analysis was performed in order to calculate QD density, diameter, and height using commercial software Scanning Probe Image Processor (SPIP) and open-source software ImageJ.

2.1.4 Scanning transmission electron microscopy (STEM)

Scanning transmission electron microscopy (STEM) is a powerful technique allowing us to image the structure directly with sub-Angstrom resolution. Under this level of resolution, individual atoms, crystal symmetries, and defects can be observed, making it an excellent tool for evaluating structural coherence. STEM uses magnetic lenses to raster a highly converged beam of electrons across the sample and generate an image (i.e. micrograph), giving it higher resolution capabilities than traditional transmission electron microscopy (TEM). Figure 2.12 illustrates the major components of an STEM. Most STEM techniques detect electrons after transmission through the sample, requiring the samples to be very thin <100 nm. Due to the transmissive nature of this technique, although the samples are three-dimensional, the resulting image is a two-dimensional projection of the sample.

For this work, we use the high-angle annular dark field (HAADF) STEM imaging technique. For this method, an annular detector is used with a large center radius, allowing electrons that are directly transmitted through the sample to pass, and
Figure 2.12: The major components of a scanning transmission electron microscope (STEM). An electron source produces an electron beam at 100-300 keV, which is directed down the column and focused to a spot size \( \sim 1 \) Å onto the sample. As the lenses raster the beam across the sample, an annular dark field detector detects Rutherford-like scattered electrons throughout the scan, producing an image that is highly sensitive to atomic mass.
detecting the electrons that are incoherently and elastically scattered at large angles by interactions with atomic nuclei in the sample. Since the scattering, and therefore atomic contrast, is directly related to the atomic number, the HAADF method is excellent at providing easily interpretable Z-contrast of a sample and avoiding phase-contrast effects, which can be difficult to analyze.

STEM can probe ranges ∼200-500 nm of the sample, making it a very useful technique for highly localized information, but can make it difficult to verify with STEM alone that the micrographs are representative of the whole sample. Therefore, it is an especially powerful technique in conjunction with other techniques. For this work, we use cross-sectional STEM to provide insight on structure and coherence on a small scale, and COBRA to provide coherence information on a much larger scale (∼0.5 mm).

Preparing samples for cross-sectional STEM is highly involved, and the quality of the sample preparation can strongly influence the quality of the STEM micrographs. For this work, samples were prepared by first depositing a thin layer of carbon on the surface of the sample to protect the sensitive nanostructures from being affected by the processes in the following step: depositing a protective layer of platinum and performing focused ion beam (FIB) lift-out in a scanning electron microscope (SEM). The platinum is heavy enough to withstand gallium-ion milling used for the lift-out procedure. Using the FIB, a cross-sectional portion of the sample is cut and lifted out of the substrate that is 50-100 nm thick, and then it is attached to a grid for STEM imaging. All STEM imaging was performed on a JEOL 3100R05 Double Cs Corrected TEM/STEM by Kai Sun, Associate Research Scientist at the University of Michigan’s Michigan Center for Materials Characterization; however, all analysis was performed separately by the author of this work using open-source software ImageJ.
2.2 Introduction to computational techniques

The $k \cdot p$ method is a well-established, semi-empirical method used to calculate the electronic properties of semiconductor materials and nanostructures, including the band structure, the effect of strain on bands, and optoelectronic properties. It uses a basis of Bloch states, which puts the Schrödinger equation in a form that allows us to easily use perturbation theory around band extrema. This inherently makes it an excellent method for calculating properties of nanostructures at low computational cost. We employ $8 \times 8$ $k \cdot p$ and effective mass methods for optoelectronic calculations of InGaN and GaAsSb quantum wells, as described in Chapters IV and V.

In this section, we present relevant background useful for understanding electronic properties of semiconductors, and then we derive for the $k \cdot p$ and effective mass equations and matrices. We discuss how this method is implemented using commercial software nextnano. Finally, we present formalism for calculating radiative and non-radiative recombination rates using wave function overlaps of the envelope wave functions produced by $k \cdot p$ and effective mass calculations.

2.2.1 $k \cdot p$ theory

The majority of semiconductors that are used for optical devices are direct-gap materials, meaning the conduction band minimum and the valence band maximum occur at the same point in k-space, namely the $\Gamma$ point ($k_x = 0, k_y = 0, k_z = 0$). In fact, the materials studied in this work GaN, InGaN, and GaAsSb, are all direct-gap semiconductors. Since the $\Gamma$ point is the lowest-energy point for both electrons and holes, this is the region of the band structure that determines much of the electronic and optical behavior. Therefore, most of the physical phenomena of these direct gap materials can be understood by focusing on the band structure very near the $\Gamma$ point. The $k \cdot p$ method uses perturbation theory around a given k-point (such as the $\Gamma$ point) to describe band structure, making it an ideal method for inexpensive computation.
of electronic and optical properties. Additionally, it is very well suited for alloys and heterostructures because it can quickly map the \( \Gamma \) point band structure properties as a function of positional variation in material properties such as composition, strain, and polarization charges. This makes it especially powerful in studying embedded quantum wells and quantum dots with either a homogeneous or fluctuating-alloy composition.

### 2.2.1.1 Derivation of the \( k \cdot p \) Schrödinger equation

The band structure of a material is determined by calculating the energy-eigenvalue solution for the Schrödinger equation,

\[
H_k \psi_{n,k} = \left[ \frac{p^2}{2m_0} + V(r) \right] \psi_{n,k} = E_{n,k} \psi_{n,k},
\]

(2.14)

where \( p \) is the momentum operator, \( n \) is the band index, \( k \) is the wave vector, \( V(r) \) is the potential, and \( m_0 \) is the mass of an electron. This energy solution is given as a function of \( k \); therefore, the full band structure is depicted as an \( E \) vs. \( k \) diagram.

The \( k \cdot p \) method relies upon the highly useful assumption that we can solve the Schrödinger equation with wave functions that are of the form of Bloch waves. We can describe Bloch wave functions in the following way. We consider a periodic lattice with a total volume of \( V(r) \) and period \( R(r) \). The Bloch theorem states that the solution of the Schrödinger equation for a periodic lattice is of the following form:

\[
\psi_{n,k}(r) = e^{ik \cdot r} u_{n,k}(r),
\]

(2.15)

where \( \psi_{n,k}(r) \) is the Bloch wave function and \( u_{n,k}(r) \) is the Bloch lattice function, which has the same periodicity as the crystal lattice. These wave functions satisfy the completeness and the orthogonality relations. We note that the periodic functions \( u_{n,k}(r) \) are orthogonal for the same \( k \) vector, but not for different \( k \) vectors.
Using the momentum operator’s derivative form, \( p = -i\hbar \nabla \), substituting the Bloch wave function into the Schrödinger equation, and using the following math to expand and simplify the left-hand side of the equation,

\[
(-i\hbar \nabla)[e^{ik \cdot r} u_{n,k}(r)] = 2\hbar k e^{ik \cdot r} u_{n,k}(r) - 2i\hbar e^{ik \cdot r} \nabla u_{n,k}(r),
\]  
(2.16)

we find that \( e^{ik \cdot r} \) cancels from every term, and we obtain the following expression:

\[
\frac{p^2}{2m_0} + \frac{\hbar}{m_0} (k \cdot p) + \frac{\hbar^2 k^2}{2m_0} + V(r) \psi_{n,k} = E_{n,k} \psi_{n,k}.
\]  
(2.17)

Notably, this is incredibly similar to the original Schrödinger equation but with an additional term in the Hamiltonian.

\[
\left[ H_0 + \frac{\hbar}{m_0} (k \cdot p) + \frac{\hbar^2 k^2}{2m_0} \right] \psi_{n,k} = E_{n,k} \psi_{n,k}
\]  
(2.18)

This additional term acts as a perturbation,

\[
H' = \frac{\hbar}{m_0} (k \cdot p) + \frac{\hbar^2 k^2}{2m_0}.
\]  
(2.19)

There are several common methods to solve equation 2.17, of which we will discuss two: the effective mass method and the \( 8 \times 8 \) \( k \cdot p \) method.

### 2.2.1.2 Effective mass theory

If the bandgap is sufficiently large or we can otherwise assume non-interacting bands, we can use standard perturbation theory to solve for the energies and wave functions. For the energies, we use perturbation theory’s energy solution to second order,

\[
E_{n,k} = E_{n,0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} k \cdot p_{nn'} + \frac{\hbar^2}{m_0^2} \sum_{n \neq n'} \frac{|k \cdot p_{nn'}|^2}{E_n(0) - E_{n'}(0)},
\]  
(2.20)
and the wave function solution to first order,
\[ u_{n,k}(r) = u_{n,0}(r) + \sum_{n \neq n'} \left[ \frac{\hbar}{m_0} \frac{k \cdot p_{nn'}}{E_n(0) - E_{n'}(0)} \right] u_{n',0}(r) \] (2.21)

\[ \psi_{n,k}(r) = e^{ik \cdot r} u_{n,k}(r) \] (2.22)

where \( p_{nn'} \) is the momentum matrix element

\[ p_{nn'} = \langle u_n | p | u_{n'} \rangle . \] (2.23)

The momentum matrix element is often determined experimentally. We note that the reason we use second order energy perturbations but first order wave function perturbations is that when we observe an extremum \( k_0, E_{n,k_0} \) must depend quadratically on \( k \), and \( p_{nn'} = 0 \). Therefore, the first nonzero energy perturbation term at an extremum is the second-order term.

We can rewrite the expression for energy dispersion, \( E_{n,k} \), using the effective mass approximation, which has the following, simple form:

\[ E_{n,k} = E_n(k) = E_n(0) + \frac{\hbar^2 k^2}{2m^*_n}, \] (2.24)

where we define the effective mass for the \( n \)th band, \( m^*_n \), using the relation

\[ \frac{1}{m^*_n} = \frac{1}{m_0} \left[ 1 + \frac{2}{m_0 k^2} \sum_{m \neq n} \left| \langle u_{n,0}(r) | k \cdot p | u_{m,0}(r) \rangle \right|^2 \right]. \] (2.25)

For example, in the case of the conduction band,

\[ m^*_c \approx \frac{m_0}{1 + \frac{2p_{cv}^2}{m_0 E_g}}. \] (2.26)

The value of \( 2p_{cv}^2/m_0 \) is remarkably constant for most semiconductors, typically \( \sim 20 \)
Figure 2.13: Illustration of the surprising accuracy of a simple effective mass approximation (equation 2.28; black line) in predicting effective mass \(m^*\) of common, low-bandgap III-V semiconductors.\(^{13}\)

eV. This allows us to approximate the conduction band effective mass as

\[
m_c^* \approx \frac{m_0}{1 + \frac{20eV}{E_g}}.
\]

(2.27)

For semiconductors where \(E_g \ll 20\) eV, we can further approximate equation Equation 2.27 as

\[
m_c^* \approx \frac{m_0 E_g}{20eV}.
\]

(2.28)

Applying equation 2.28 to a variety of common semiconductors, we obtain the linear curve shown in Figure 2.13, which is compared to the accepted values. It is easy to observe that the simple relationship shown in Equation 2.28 predicts III-V effective masses remarkably well. Despite the simplicity of the expression, the effective mass approximation gives a powerful result.
We note that it is conceptually understandable that the effective mass increases with band gap, as observed in Figure 2.13. As the conduction and valence bands become farther in energy from each other, they interact less; therefore, their “pushing” effect on one another decreases, flattening out the band and resulting in a larger effective mass.

2.2.1.3 $8 \times 8 k \cdot p$ theory

For the case of interacting bands, we solve the Schrödinger equation with perturbation terms (Equation 2.17), we form a matrix to solve for the energy eigenvalues. We first assume we know the Bloch functions and energies at $k = 0$, $u_{n,0}(r)$ and $E_{n,0}$. We can then expand the Bloch functions for $k \neq 0$ in terms of the functions for $k = 0$:

$$u_{n,k}(r) = \sum_n C_n(k)u_{n,0}(r).$$

(2.29)

This allows us to rewrite equation 2.17:

$$\sum_n C_n(k) \left[ E_{n,0} + \frac{\hbar}{m_0}(k \cdot p) + \frac{\hbar^2 k^2}{2m_0} \right] u_{n,0}(r) = E_{n,k} \sum_n C_n(k)u_{n,0}(r)$$

(2.30)

If we utilize the orthogonality of $u_{n,0}(r)$ by multiplying $u^*_{n',0}(r)$ from the left and integrating over the crystal’s primitize cell, we obtain a system of linear equations that allow us to solve for the unknown energies $E(k)$ and coefficients $C_n(k)$,

$$\sum_n \left\{ \left[ E_{n,0} - E_{n,k} + \frac{\hbar^2 k^2}{2m_0} \right] \delta_{nn'} + \frac{\hbar}{m_0} k \cdot p_{nn'} \right\} C_n(k) = 0,$$

(2.31)

where $p_{nn'}$ is again the momentum matrix element. The solutions $E(k)$ and $C_n(k)$ can be found by diagonalizing the matrix.
\[
\begin{bmatrix}
E_{n,0} - E_{n,k} + \frac{\hbar^2 k^2}{2m_0} \delta_{nn'} + \frac{\hbar}{m_0} k \cdot p_{nn'}
\end{bmatrix} = 0.
\]

(2.32)

For example, the matrix for a 2-band case is

\[
\begin{bmatrix}
E_{n,0} + \frac{\hbar^2 k^2}{2m_0} - E_{n,k} & \frac{\hbar}{m_0} k \cdot p_{nn'} \\
\frac{\hbar}{m_0} k \cdot p_{n'n} & E_{n',0} + \frac{\hbar^2 k^2}{2m_0} - E_{n,k}
\end{bmatrix} = 0
\]

(2.33)

However, it is typical to consider four bands [conduction, heavy hole (HH), light hole (LH), and split-off (SO) bands] with spin degeneracy, producing an $8 \times 8$ matrix. This is referred to as the $8 \times 8$ $k \cdot p$ method. In order to consider spin degeneracy and produce this $8 \times 8$ matrix, we add a spin-orbit term to the Hamiltonian:

\[
H = H_0 + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} k \cdot p + \frac{\hbar}{4m_0^2 c^2} (\sigma \times \nabla V) \cdot (\hbar \mathbf{k} + \mathbf{p})
\]

(2.34)

where $\sigma$ is the set of Pauli spin matrices. Similarly, the effects of strain and crystal field splitting (relevant for wurtzite structures) are incorporated with a perturbative term to the Hamiltonian.

### 2.2.1.4 Implementation of $k \cdot p$ methods

We employ the $8 \times 8$ $k \cdot p$ and the effective mass methods through a commercially available software package called nextnano.\(^{69}\) Nextnano can self-consistently solve the Schrödinger and Poisson equations to calculate band structure while incorporating strain and polarization charges. Many studies have utilized nextnano for QD, quantum well, and nano-device applications with a variety of material systems.\(^{66,69–73}\)

As described in the computational background information, the $k \cdot p$ and effective mass methods are semi-empirical, so they rely upon material parameters obtained experimentally or through \textit{ab initio} computational methods. nextnano uses a database of material properties which it modifies accordingly based upon the conditions of the
material at each grid point in the simulation box. Some of the values important to this work include band gaps, band offsets, deformation potentials, lattice constants, and expansion coefficients. For our work on InGaN and GaAsSb, we used material parameter sources by Vurgaftman, Ambacher, Zunger, and Yan.\textsuperscript{1,74–76}

### 2.2.1.5 nextnano program work flow

The calculation begins with an initial calculation set up. The program reads the input file, creates the simulation grid, and maps the requested structure geometries onto the grid. The grid is discretized using the finite difference method with periodic boundary conditions. The program also pulls the necessary material values from the database, and allocates and writes the required arrays for the calculation.

The next step of the program calculates the strain in the material using the continuum elasticity model and by minimizing the elastic energy of the system. This strain is then used to modify the band edge properties and induce appropriate band splitting using the deformation potentials. If doping is included in the calculation, ionization of dopants is calculated, followed by the calculations of piezoelectric and pyroelectric polarization charges. As the final preparatory step, the built-in potential of the system is calculated based upon the defined potential at the contacts. The Schrödinger and Poisson equations are then solved iteratively until the calculation reaches the defined convergence energy tolerances.

### 2.2.2 Theory of recombination for localized, envelope wave functions

Here we derive the equations that relate the radiative and Auger coefficients of fluctuating alloys to the corresponding coefficients of virtual-crystal alloys and weighted-averaged overlaps squared of envelope functions. We start by expressing the localized wave functions of fluctuating alloys as a linear combination of virtual-crystal bulk Bloch functions:
\[ \psi_J(r) = \frac{1}{\sqrt{V}} \sum_k c_J(k)e^{ik \cdot r} u_{nk}(r) \cong \varphi_J(r) u_n(r) \quad (2.35) \]

where the periodic component of the Bloch wave function can approximately be considered independent of the wave vector,

\[ u_{nk}(r) \cong u_n(r) \quad (2.36) \]

and the envelope function of the localized carriers is given by,

\[ \varphi_J(r) = \frac{1}{\sqrt{V}} \sum_k c_j(k)e^{ik \cdot r}, \quad (2.37) \]

where the composite index,

\[ J \equiv (j, n), \quad (2.38) \]

is a combination of the index of the localized state \( j \) and the bulk band index \( n \). In this basis, the radiative recombination rate (i.e., the number of recombining carriers per unit time per unit volume) is determined by Fermi’s golden rule,

\[ R_{rad} = 2 \frac{2\pi}{\hbar} \sum_{1,2} f_1(1 - f_2) |\langle 1 | H_{el-phot} | 2 \rangle|^2 \delta(\epsilon_1 - \epsilon_2 - \hbar \omega). \quad (2.39) \]

The matrix elements of the electron-photon perturbation Hamiltonian between localized states can be expressed in the basis of delocalized Bloch orbitals as:

\[ \langle 1 | H_{el-phot} | 2 \rangle = \sum_{k_1, k_2} c^*_1(k_1)c_2(k_2) \langle n_1k_1 | H_{el-phot} | n_2k_2 \rangle. \quad (2.40) \]

In bulk virtual-crystal alloys, the translational crystalline symmetry conserves the crystal momentum in radiative transitions, and the matrix elements are given by:
\[
\langle n_1 k_1 | H_{el-phot} | n_2 k_2 \rangle \propto p_{n_1 n_2}(k_1) \delta_{k_1, k_2} \simeq p_{n_1 n_2} \delta_{k_1, k_2},
\]

where the interband matrix elements are further assumed independent of wave vector. In this case, the electron-photon matrix elements between localized electrons and holes are proportional to the electron and hole envelope function overlap according to:

\[
\langle 1 | H_{el-phot} | 2 \rangle \propto p_{n_1 n_2} \sum_{k_1} c_1^*(k_1) c_2(k_1) \propto p_{n_1 n_2} \int \varphi_1^*(r) \varphi_2(r) d\mathbf{r}
\]

Therefore, the radiative recombination rate is given by:

\[
R_{rad} = B np \propto \sum_{1, 2} f_1(1 - f_2) \left| \int \varphi_1^*(r) \varphi_2(r) d\mathbf{r} \right|^2;
\]

i.e., it is proportional to the weighted sum of electron-hole envelope function overlap squared, the weights being the electron and hole Fermi-Dirac occupation factors. The radiative coefficient is subsequently given by

\[
B \propto \frac{\sum_{1, 2} f_1(1 - f_2) \left| \int \varphi_1^*(r) \varphi_2(r) d\mathbf{r} \right|^2}{\sum_{1, 2} f_1(1 - f_2)} = \frac{F_{eh}^2}{F_{eh}^2},
\]

i.e., it is proportional to the weighted-averaged overlap squared of electron and hole wave functions. The last equation allows the comparison of radiative recombination coefficients of virtual-crystal and fluctuating-alloy quantum wells to bulk virtual crystals according to:

\[
\frac{B_{fluct/VCA,QW}}{B_{VC,A,Bulk}} = \frac{F_{eh}^{fluct/VCA,QW}}{F_{eh}^{VC,A,Bulk}}.
\]

We extend the previous analysis to encompass Auger recombination. We examine electron-electron-hole recombination (the hole-hole-electron case is analogous). The
recombination rate is given by Fermi’s golden rule according to:

\[
R_{\text{Auger, eeh}} = \frac{2\pi}{\hbar} \sum_{1,2,3,4} f_1 f_2 (1 - f_3) (1 - f_4) |\langle 12 | H_{\text{Auger}} | 34 \rangle|^2 \delta(\epsilon_1 + \epsilon_2 - \epsilon_3 - \epsilon_4),
\]

(2.46)

where the Auger recombination matrix elements are the screened Coulomb interaction matrix elements between the initial and final pairs of states. Expressed in the Bloch basis of virtual crystals, the matrix elements are given by:

\[
\langle 12 | H_{\text{Auger}} | 34 \rangle = \int \int \psi_1^*(r) \psi_2^*(r') W(r, r') \psi_3(r) \psi_4(r') dr dr'
\]

\[
= \sum_{k_1, k_2, k_3, k_4} c_1^* (k_1) c_2^* (k_2) c_3 (k_3) c_4 (k_4) \langle n_1 k_1 n_2 k_2 | W | n_3 k_3 n_4 k_4 \rangle.
\]

(2.47)

In general, the Coulomb matrix elements are nontrivial functions of the wave vectors. They involve the wave-vector dependence of the screened Coulomb interaction, as well as the wave-vector dependence of the overlaps between the periodic parts of Bloch functions across bands.\textsuperscript{77} However, the matrix elements can be simplified in the wide-band-gap group-III nitrides. Since the energy released by the recombining electron-hole pair is large, on the order of 2.4–3.0 eV for InGaN used for green to violet LEDs, the resulting Auger electron is excited to high conduction-band states at energies higher than the conduction-band minimum by approximately the value of the gap. The momentum transfer involved for these high-energy Auger transitions is large, which indicates that Coulomb scattering in wide-gap-nitrides is short-ranged.\textsuperscript{31} As demonstrated in Figure 2.14, the dominant Auger matrix elements in group-III nitrides are approximately constant and independent of wave vector at short range (large momentum transfer). We can therefore approximately express the Coulomb matrix elements between virtual-crystal states as a constant times a Kronecker delta
for momentum conservation:

$$\langle n_1 k_1 n_2 k_2 | W | n_3 k_3 n_4 k_4 \rangle \cong W \delta_{k_1+k_2-k_3-k_4}.$$  \hspace{1cm} (2.48)

Under the above approximation, valid for other wide-band-gap materials as well, we can write the Auger matrix elements in terms of the envelope functions as:

$$\langle 12 | H_{\text{Auger}} | 34 \rangle \cong W \sum_{k_1,k_2,k_3} c_1^*(k_1) c_2^*(k_2) c_3(k_3) c_4(k_1+k_2-k_3)$$

$$= W \sum_{k_1,k_2,k_3} \int \varphi_1^*(r_1) e^{ik_1 \cdot r_1} \varphi_2^*(r_2) e^{ik_2 \cdot r_2} \varphi_3(r_3) e^{-ik_3 \cdot r_3} \varphi_4(r_4) \times$$

$$e^{-i(k_1+k_2-k_3) \cdot r_4} dr_1 dr_2 dr_3 dr_4.$$  \hspace{1cm} (2.49)

Now, we make use of the identity:

$$\sum_{k_1} e^{ik_1 \cdot (r_1-r_4)} = V \delta(r_1-r_4)$$  \hspace{1cm} (2.50)

to reduce the integral to a simple overlap integral of the four envelope wave functions of the states involved in the Auger transition:

$$\langle 12 | H_{\text{Auger}} | 34 \rangle \propto \int \varphi_1^*(r) \varphi_2^*(r) \varphi_3(r) \varphi_4(r) dr$$  \hspace{1cm} (2.51)

Now, state 4 is highly excited and its energy significantly exceeds the localization energy. It is therefore highly delocalized over the entire volume of the material, and its envelope function can be approximated to be constant. The Auger matrix elements can therefore be reduced to triple overlap integrals according to:

$$\langle 12 | H_{\text{Auger}} | 34 \rangle \propto \int \varphi_1^*(r) \varphi_2^*(r) \varphi_3(r) dr.$$  \hspace{1cm} (2.52)

Correspondingly, the Auger recombination rate is given by a weighted sum of the
Figure 2.14: Auger matrix elements of GaN for (a) electron-electron-hole (eeh) and (b) hole-hole-electron (hhe) Auger recombination as a function of the momentum transfer involved in the Auger process. For short-range Auger scattering (i.e., large momentum transfer), which dominates Auger recombination in wide-band-gap nitrides, the dominant matrix elements reach values that are approximately independent of momentum transfer (indicated with dotted line).
squares of triple overlaps, the weights given by the occupation numbers of the recombining electrons and holes according to:

\[ R_{\text{Auger,eeh}} = C_{\text{eh}} n^2 p \propto \sum_{1,2,3} f_1 f_2 (1 - f_3) V_{\text{QW}} \left| \varphi_1^*(r) \varphi_2^*(r) \varphi_3(r) dr \right|^2, \quad (2.53) \]

where \( V_{\text{QW}} \) is the volume of the quantum well, and thus the Auger coefficients are proportional to weighted-averaged triple overlaps squared according to:

\[ C_{\text{eh}} \propto \frac{\sum_{1,2,3} f_1 f_2 (1 - f_3) V_{\text{QW}} \left| \varphi_1^*(r) \varphi_2^*(r) \varphi_3(r) dr \right|^2}{\sum_{1,2,3} f_1 f_2 (1 - f_3)} \equiv \frac{F_{\text{eh}}^2}{F_{\text{eh}}} . \quad (2.54) \]

By evaluating the weighted-averaged triple overlaps squared of fluctuating alloys and comparing to virtual crystals we can determine the impact of carrier localization on the Auger coefficients according to:

\[
\frac{C_{\text{eh}}^{\text{fluct/VCA,QW}}}{C_{\text{eh}}^{\text{VCA,Bulk}}} = \frac{F_{\text{eh}}^{\text{fluct/VCA,QW}}}{F_{\text{eh}}^{\text{VCA,Bulk}}} .
\quad (2.55)
\]

and similarly for the hole-hole-electron process:

\[
\frac{C_{\text{hhe}}^{\text{fluct/VCA,QW}}}{C_{\text{hhe}}^{\text{VCA,Bulk}}} = \frac{F_{\text{hhe}}^{\text{fluct/VCA,QW}}}{F_{\text{hhe}}^{\text{VCA,Bulk}}} .
\quad (2.56)
\]
CHAPTER III

Surface studies of GaN QDs grown using droplet epitaxy on bulk, native substrates

3.1 Background and motivation

3.1.1 Introduction

GaN and InGaN devices thus far have predominantly been fabricated on heteroepitaxial (often referred to as “foreign” or “template”) substrates. For template substrates, several hundred microns of GaN was grown on a sapphire, silicon carbide, or silicon substrate as the base for a device. Because of the 13.9% lattice mismatch between sapphire and GaN, for example, the GaN templates have high dislocation densities of $10^8 \sim 10^{10}$ cm$^{-2}$. Reducing the dislocation density of the substrates is an essential step in obtaining the highest possible commercial device performance.

3.1.2 GaN substrate growth methods

Improvements in substrate dislocation densities first began with the development of the A-DEEP (advanced dislocation elimination by the epitaxial-growth with inverse-pyramidal pits) method. This method starts with a heteroepitaxial substrate on which a thick GaN film is grown simultaneously while employing a technique during growth to consolidate defects into pits and reduce the net defect density.
result is a thick GaN layer where the substrate portions with high defect densities are separated from portions with low defect densities of \( \sim 2 \times 10^{5} \text{ cm}^{-2} \). The GaN substrate is then released from the heteroepitaxial substrate using lapping and polishing. The use of these freestanding substrates in laser diodes was shown to increase lifetime by over an order of magnitude, but due to the fact that defects were consolidated and not eliminated, devices have to be positioned on low-defect regions of the substrate. Other methods of producing freestanding GaN substrates with reduced defect densities have subsequently been developed, including the void-assisted separation (VAS) method and the random-islands facet-initiated epitaxial lateral overgrowth (r-FIELO) technique, each producing substrates with defect densities of \( \sim 2 \times 10^{6} \text{ cm}^{-2} \). The use of GaN bulk substrates for InGaN devices is seen as the future of highly efficient, commercial white LEDs. In fact, InGaN LEDs fabricated on bulk, native substrates have been called “second-generation LEDs” by LED leading expert, Michael Krames, and Nobel laureate Shuji Nakamura. Great effort is being put into bulk GaN growth methods in order to reduce the dislocation density and increase the wafer size of the bulk substrates, thereby increasing their commercial impact.

Several primary bulk fabrication methods are ammonothermal growth in autoclaves, ammonothermal growth in high temperature and high pressure apparatuses, high-pressure nitrogen solution growth, sodium-flux-based growth, and hydride vapor phase epitaxy (HVPE). We briefly describe each growth method in the following paragraphs.

The ammonothermal method is based on the well-established hydrothermal method used for the commercial growth of quartz. It involves dissolving polycrystalline GaN into supercritical ammonia under high pressure, often using mineralizers to increase the solubility. The dissolved GaN is then transferred to a lower solubility, growth zone using a temperature gradient, and the GaN vapor crystallizes onto GaN seed crystals, growing into bulk GaN boules. The ammonothermal method produces crystals with
very low defect densities \( \sim 10^3 \text{ cm}^{-2} \).\textsuperscript{81}

On the other hand, the high-pressure nitrogen solution method uses a direct reaction between gallium and nitrogen. Because GaN has such strong bonding, this growth method must be performed at very high pressures (\( \sim 1 \text{ GPa} \)) and temperatures (\( \sim 1700^\circ \text{C} \)), which increases the nitrogen solubility in gallium.\textsuperscript{81} This direct method is very slow and results in small crystallites (\( \sim \text{mm} \)), however the crystals demonstrate the lowest bulk GaN dislocation densities (\( \sim 10^2 \text{ cm}^{-2} \)).\textsuperscript{81}

In contrast, in the sodium-flux-based approach, nitrogen solubility is increased using sodium so the growth can be performed at much lower pressures (\( \sim 4 \text{ MPa} \)) and temperatures (750-900\(^\circ\text{C}\)).\textsuperscript{81} Six-inch wafers have been created using this method; however, its main creator, Yusuke Mori, has announced in 2015 that it is likely a better method for seed growth than large bulk crystal growth.\textsuperscript{82}

The final method is also the most common commercial method: the HVPE method. This method utilizes two zones: a source zone and a deposition zone. In the source zone, gallium reacts with HCl to produce a reactive, chloride gas of gallium, after which it is carried to the deposition zone by H\(_2\) and inert gases. In the deposition zone, the GaCl and GaCl\(_3\) gases react with NH\(_3\) to form GaN. This method is performed at low pressures (1 atm) but very high temperatures (\( > 1000^\circ \text{C} \)). It provides growth times an order of magnitude faster than the other methods, but the resulting strain is higher and the dislocation densities are at least an order of magnitude higher than in crystals grown using other methods. Additionally, this method requires a seed crystal which is usually a foreign substrate (\( \text{Al}_2\text{O}_3, \text{GaAs, SiC, Si} \)).\textsuperscript{82} Combining the HVPE and ammonothermal methods by using a bulk GaN seed crystal grown using the ammonothermal method for fast HVPE growth may facilitate increased commercial production and higher quality substrates.\textsuperscript{82}

Overall, the largest bulk wafer that has been realized is 6 inches in diameter; however, the largest wafers for large-scale production have mostly been grown using
HVPE and have diameters of 2 inches.\textsuperscript{81} These bulk crystals can be used to produce wafers in polar, semi-polar, and non-polar orientations.

Impressive performance enhancements have been realized on these bulk, native substrates: a 4\% efficiency droop compared to 18–27\% droop on foreign substrates. Additionally, devices on native substrates produce $10 \times$ high power density.\textsuperscript{78} Improvements in material quality through these improved substrate methods not only increase device performance, but will also allow for larger area substrates compatible with higher throughput manufacturing techniques and result in overall lower device cost. The device performance results demonstrate the promise of bulk native substrates’ contribution towards nitride applications, and they motivate further understanding of these new substrates in order to facilitate high efficiency and low cost in commercial devices.

### 3.1.3 GaN quantum dots

Using QDs instead of thin films is a promising way to address some of the performance issues in InGaN devices. QDs can accommodate strain more easily than thin films, allowing for higher quality crystals, particularly in the green regime, and lower defect-assisted recombination. Additionally, less strain results in weakened QCSE and less polarization-induced carrier separation. QDs can also allow for easier control of emission wavelength through confinement. For example, embedded QDs of different compositions and dimensions can individually emit at blue, green, and red, effectively producing white light emission.\textsuperscript{22} Finally, confinement induced by the QDs could result in a larger electron-hole overlap and higher carrier recombination.

### 3.1.4 Quantum dot growth methods

A common growth technique used to grow QDs is the Stranski-Krastanov (SK) method, which relies on a large lattice mismatch between the substrate and growth
material in order to induce self-assembly. Because this method is driven by strain, QD size, density, and shape can be limited or difficult to control. Additionally, the strain-driven mechanism makes the SK method not suitable for homoepitaxial growth.\(^{83}\)

An exciting alternative to the SK growth method is the droplet epitaxy (DE) method. The first step of DE is depositing liquid-phase nano-droplets of metal (gallium in this case) onto the substrate surface. In the case of III-V semiconductors, the group-III element is used for this droplet deposition. The second step is to expose the liquid droplets to a group-V-element vapor. Under appropriate conditions including temperature, vapor pressure, and plasma power, the droplet crystallizes into a III-V material that is structurally coherent, or atomically registered, with the substrate.

Because the DE method does not rely on a lattice mismatch between the substrate and growth material, it offers greater flexibility in tuning nanostructure size and density and also offers more substrate flexibility than the SK method. It can also produce QD growth without a wetting layer and allow for precise QD position control through pre-patterned substrates. The DE method has already been demonstrated in growing QDs using several material systems, including InSb,\(^{67}\) InAs,\(^{67}\) GaSb,\(^{39}\) GaN,\(^{84}\) InN,\(^{85}\) AlN,\(^{84,86,87}\) and AlGaN.\(^{88-90}\) For the case of GaN, the resulting QD size and density can be controlled using various growth parameters: gallium flux (for droplet formation), nitridation temperature, and plasma power. Smaller, higher density QDs tend to be produced using lower gallium fluxes and lower nitridation temperatures.\(^{84,91-93}\)

### 3.2 Project objectives

This work is motivated by several current questions and issues of InGaN LEDs. Bulk GaN is certainly the future substrate of highly efficient devices, but its commercial availability is still limited and its properties for device growth not fully understood. We would like to explore the physics and surface properties of \textit{bulk} GaN for all
of our studies. In addition, in order to address several of the crystal quality and polarization issues observed in conventional LEDs, we would like to explore nitride QDs instead of quantum wells in order to take advantage of the many benefits described in Section 3.1.3. Finally, in order to grow nitride materials on native substrates, we are interested in exploring the use of the DE growth method for producing structurally coherent, or atomically registered, QDs.

In order to understand the effects of homoepitaxial QD growth on bulk GaN substrates using droplet epitaxy, our objectives are threefold. First, we seek to understand the surface properties of commercial, bulk GaN for fundamental physical reasons and also as a basis for future QD and device-growth studies. Second, we want to identify droplet epitaxy growth conditions that produce structurally coherent QDs. Third, we seek to study the degree of crystallographic coherence within the QDs. We will accomplish these goals through the combined insights of surface X-ray diffraction, COBRA analysis, AFM, and STEM. Together, the knowledge we gain from these techniques provides the basis for potential future work on embedded InGaN QDs, which would be essential towards integration into LED devices. Accomplishing these goals will help us better understand surface properties of bulk GaN and how to produce coherent QDs for highly efficient light-emitting devices.

3.3 Bulk GaN: bare substrate

3.3.1 Dow Corning substrates

We first study a bulk GaN substrate obtained from Dow Corning. The substrate is single-crystal, c-plane GaN with a Ga-terminated surface. These substrates are still under commercial development by Dow Corning and are not yet available for purchase. For this reason, details on the growth and many of the material properties were not disclosed.
We performed AFM to verify the quality of the substrate surface (Figure 3.1). We observe the surface to be clean with very low surface roughness (0.4-0.6 nm), which is approximately the height of one unit cell. In larger area scans (5 \(\mu\)m \(\times\) 5 \(\mu\)m), we observe some “waviness” in the atomic steps, indicating the surface is not completely flat on the micron scale. Of particular interest are the atomic steps visible in the image, which are present due to a small miscut of the surface with respect to the crystallographic plane; each step is the termination of an atomic plane at the sample surface. These atomic steps visible in AFM are particularly useful for our analysis in that they can be used to calculate the substrate miscut (also called off-cut or misorientation). For wurtzite structures grown along the c direction, each atomic step is a monolayer and therefore has the height of half the wurtzite c lattice constant.\(^{94-96}\) We therefore anticipate the step height for GaN to be 2.593 Å, which agrees well with 0.22 nm measured using the height scale in the AFM images. In this work, we use ideal step heights to calculate miscut as opposed to those measured from AFM due to the large relative error in the AFM technique on these scales. Using simple geometry, the step height, and the step widths observed in AFM, we calculate a 0.25° miscut for the Dow Corning bulk GaN substrate.

We performed Bragg rod scans of the Dow Corning bare substrate at Sector 33-
Figure 3.2: Several rodscans (blue line) plotted with their corresponding COBRA model fits (red line) for the Dow Corning bare, bulk GaN substrate. The model fit is very good giving an FOM of 0.0872.

ID at the Advanced Photon Source of Argonne National Lab and reconstructed the electron density using COBRA. We measured six irreducible Bragg rods to construct reciprocal space data for the sample. Our beam energy was 10.30 keV, which is just under the gallium K-absorption edge of 10.367 keV, allowing us to avoid background intensity from K-shell fluorescence.

Figure 3.2 shows several of the rodscans we measure for the Dow Corning bulk substrate and the fit of our COBRA model for this sample. The data exhibits smooth dips between Bragg peaks, which is the behavior we expect for a bare substrate with low surface roughness. We note that for the (00\ell) rod (where \ell is the reciprocal lattice constant for the out-of-plane direction), selection rules do not allow a Bragg peak at odd-numbered \ell values, so we do not observe a Bragg peak at (001) or (003). Overall, the fit is very good and gives a figure of merit (FOM) value (r_1) of 0.0872. The FOM
is the least squares difference between the reference and experimental rodscans, and strong FOM values are below 0.10, while good values, particularly for the GaN system with 10 Bragg rods included in the fit, are below 0.15.

The electron density results are shown in Figure 3.3 as a line profile along the growth direction, (0001), and the nominal surface is shifted to zero nanometers. In these results, we observe behavior consistent with a bare substrate with variation of surface height. The first \(\sim 11\) monolayers occur within the bulk, after which the electron density layers drop off in intensity over a range of \(\sim 6-9\) monolayers before the signal approaches the level of noise. The two profile lines correspond to a line through each of the two gallium atoms in the primitive unit cell (visible in the Figure 3.3 inset). Therefore, each profile line provides a different view of the same model results. Higher peaks correspond to gallium atoms, and lower peaks correspond to nitrogen atoms; since gallium has 31 electrons and nitrogen has 7 electrons, it is easy to distinguish between their electron density peaks in the bulk. This electron density drop-off occurs over a larger length scale than might be anticipated for a bare substrate with a surface roughness of \(\sim 0.5\) nm. We believe this drop-off occurs over more monolayers than expected because the variation in surface height observed in AFM is on the order of microns; with an incident X-ray beam footprint of approximately 0.1 mm × 0.5 mm, we cover many of these surface height variations.

In Figure 3.4, we illustrate 2D electron density results for the Dow Corning bare substrate for the (0001) plane (\(c\)-plane; left) and (11-20) plane (\(a\)-plane; right). The insets illustrate these cross-sectional planes with respect to a schematic of the wurtzite conventional cell. The electron density \(c\)-plane cross-section was taken at a plane within the bulk of the substrate, and it illustrates the 6-fold symmetry and in-plane lattice constants we expect for a wurtzite GaN structure. The \(a\)-plane cross-section also shows expected symmetry and \(c\)-lattice constant. Above the nominal substrate surface though, we observe that the bulk structure does not persist at the surface;
atoms appear in the interstitial positions, and the spacing between the gallium and nitrogen atoms become more evenly spaced along each line of atoms in the $z$-direction. This altered crystal structure at the surface does not agree with the likely possibilities of $\beta$-$\text{Ga}_2\text{O}_3$ or GaN in the zinc blende structure. It is very likely that the interstitial signals are artifacts of the fit. These can arise from small errors during the experiment that result in COBRA producing additional Fourier components for the structure factor. It is particularly likely that the surface features are artifacts because the altered crystal structure occurs most strongly as the electron density result approaches the noise level. Therefore, these weak features are likely not trustworthy. COBRA error analysis, developed by Yizhak Yacoby and coworkers, \cite{Zhou:2012} may provide additional insight into whether these weak occupations in the interstitial sites are trustworthy; however, this analysis is highly involved and would only be potentially advantageous if we were to continue work with these substrates. This is the only sample with which we observed features in the interstitial sites and slight spacing changes between atoms at the surface. In order to help reduce the likelihood of spurious solutions and artifacts going forward, we measured 10 instead of 6 Bragg
Figure 3.4: Electron density of Dow Corning bare, bulk GaN substrate. (a) Cross-section along c-plane, taken within the bulk of the substrate and demonstrating the 6-fold symmetry we expect for GaN. (b) Cross-section along a-plane, illustrating electron density drop-off and structure change above substrate surface.

rods for all other samples.

3.3.2 Kyma substrates

The second bare substrate we study is a bulk GaN substrate from Kyma Technologies (based in Raleigh, NC USA), oriented in the c-plane direction with a Ga-terminated surface. The substrates are specified to have a dislocation density \( \leq 5 \times 10^6 \) cm\(^{-2}\), surface orientation within 1° of the c-plane, surface polished to a roughness <0.5 nm, and they arrive epi-ready. Kyma achieves bulk GaN growth using the HVPE method.

AFM images of the Kyma substrate surface (Figure 3.5) show the substrate to arrive clean with a low surface roughness \( \sim 0.5 \) nm, as advertised by the company. Atomic steps are easily visible in this sample, as observed with the Dow Corning substrate. The miscut was quoted by the company to be <0.61°, and using the atomic steps visible in the AFM image we calculate the miscut in this region of the substrate to be 0.30°.
We performed Bragg rod scans of the Kyma bare substrate and reconstructed the electron density using COBRA. Our beam energy was 15 keV for this sample, which allowed us to reach 10 irreducible Bragg rods to construct reciprocal space data for the sample. While this beam energy is above the gallium K edge and therefore we measure some additional background signal from fluorescence, we benefit by reaching more Bragg rods and being able to scan farther along each rod, which allows us to produce higher resolution results and higher accuracy by including more data into the model fit.

Similar to the Dow Corning results, we observe smooth dips between Bragg peaks in our rod scans (Figure 3.6). Our best model fit has a FOM value of 0.1434, which is not as low our value for the Dow Corning substrate, but is standard for our analysis of Kyma GaN substrates. One possible reason for a higher FOM in the Kyma samples is the increased amount of data (more rods, longer scans along each rod), which is inherently more difficult to fit. The fit is particularly good along “shoulders” of the peaks, which are the regions of highest interest.

Electron density results in the form of two line profiles along the out of plane direction for a bare Kyma bulk substrate are shown in Figure 3.7a. As with the Dow Corning electron density results, each line corresponds to the profile through each of the two gallium atoms in the primitive unit cell, and the nominal surface has been
Figure 3.6: Several rodscans (blue line) plotted with their corresponding COBRA model fits (red line) for the Kyma bare, bulk GaN substrate (Kyma4a). The model fit is good, giving an FOM of 0.1434.
shifted to 0 nm. Below the surface we observe expected electron density behavior that agrees with our observations in the Dow Corning substrate; peaks for gallium atoms are significantly higher than those of nitrogen atoms, and the results are uniform throughout the bulk. The cross-sectional view of the electron density through the $a$-plane (3.7b) corroborate this behavior. In particular, the electron density appears to drop off quickly above the substrate surface.

Notably, with the Kyma bulk substrate, we observe a much faster electron density drop-off at the surface than with the Dow Corning substrate. In the Kyma substrate, the drop-off occurs over only $\sim$4 monolayers. This indicates that the substrate is flatter on the 100-micron scale (size of the beam footprint) than the Dow Corning sample, which we can verify for the micron scale using the AFM images since we observe straight atomic steps. The smoother surface properties of the Kyma bare substrate make it a better substrate than the Dow Corning substrate for our COBRA studies because it is easier to determine whether atomic layers above the nominal substrate surface can be attributed to structurally coherent QDs as opposed to extended substrate surface layers. We use the bulk GaN substrates from Kyma for all of our subsequent studies.

Figure 3.7: Electron density line profiles for the bare substrate Kyma 4a along the $c$-axis (a) and cross-section along $a$-plane (b), illustrating a very rapid intensity drop-off above the substrate surface (nominally $z = 0$ Å).
Table 3.1: Substrate crystallographic miscut, DE growth parameters, QD morphology, and QD coherence results as determined by COBRA for samples of GaN QDs grown on bulk GaN.

<table>
<thead>
<tr>
<th>QD sample no.</th>
<th>Sample name</th>
<th>Substrate miscut (°)</th>
<th>Nitridation temperature (°C)</th>
<th>Nitrogen plasma power (W)</th>
<th>Nitrogen flow rate (sccm)</th>
<th>QD density (cm⁻²)</th>
<th>QD diameter (nm)</th>
<th>QD height (nm)</th>
<th>Coherent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RMBE1145, Kyma4b</td>
<td>0.61 (Kyma), 0.30 (AFM, this work)</td>
<td>550</td>
<td>350</td>
<td>1</td>
<td>2.20×10¹¹</td>
<td>23 ± 6</td>
<td>1.3</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>RMBE1194A</td>
<td>0.38 (Kyma)</td>
<td>250</td>
<td>350</td>
<td>1</td>
<td>1.40×10¹¹</td>
<td>20</td>
<td>3.2</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>RMBE1197A</td>
<td>0.50 (Kyma)</td>
<td>185</td>
<td>350</td>
<td>1</td>
<td>2.1×10¹¹</td>
<td>22</td>
<td>4</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>RMBE1263</td>
<td>0.18 (Kyma), 0.06 (AFM, this work)</td>
<td>185</td>
<td>350</td>
<td>1</td>
<td>8.3×10¹⁰</td>
<td>19 ± 7</td>
<td>2.7 ± 0.6</td>
<td>N</td>
</tr>
<tr>
<td>5</td>
<td>RMBE1265</td>
<td>0.37 (Kyma), 0.39 (AFM, this work)</td>
<td>185</td>
<td>450</td>
<td>0.55</td>
<td>1.5×10¹¹</td>
<td>15 ± 4</td>
<td>4 ± 1</td>
<td>N</td>
</tr>
</tbody>
</table>

3.4 GaN QDs on bulk GaN

GaN QDs were grown on Kyma GaN bulk substrates by our collaborators Dr. Sunyeol Jeon and Professor Rachel Goldman at the University of Michigan. The QDs were grown using the DE method in a molecular beam epitaxy (MBE) system, which is an ultra-high-vacuum system (<3×10⁻¹⁰ Torr) that allows for atomic-layer level control over growth of high quality films and nanostructures. The substrate is heated to a specific temperature, and molecular gas beams are passed over the surface, resulting in epitaxial growth. For the case of GaN QDs, the bulk GaN substrate first undergoes a thermal cleaning in the MBE chamber, in which it is heated to 900°C for 10 min. 5.5 ML of dense gallium liquid droplets are then deposited at a rate of 0.1 ML/s for 55 s on a substrate held at 185°C. The nitrogen valve is then opened to allow nitrogen gas, created using a nitrogen plasma, to flow over the substrate for 30 min and perform nitridation, and therefore crystallization, of the QDs. Further information about DE growth methods can be found in Dr. Jeon’s Ph.D. dissertation.83 For this study, the substrate temperature during nitridation and the nitrogen plasma and flow conditions were varied (Table 3.1).

We note that the samples reported in Table 3.1 do not represent all the samples.
grown; they represent the samples that produced dense QDs for which we were able to obtain rodscans and perform COBRA analysis. Therefore, this list does not represent all growth parameters explored nor does it represent all growth parameters that produced QDs or other similar nanostructures (i.e. rings).

Following growth of the GaN QDs on native, bulk substrates, AFM images were taken to confirm QD growth and measure QD statistics (Figure 3.8). The images all show beautiful QD shapes with high density. We note that the QDs are very flat in a shape one might call a “pancake.” For one sample, sample no. 1 (RMBE 1145) in which nitridation was performed at 550°C, there is question as to whether the QD features visible with AFM are QDs or whether they are gallium droplets which did not successfully undergo nitridation due to an experimental problem during growth. The composition and/or crystallinity of the nanoscale features can be confirmed using TEM. Additionally, a selective chemical etch could be used as a possible destructive technique to determine their composition. The density for all samples reported is higher than $10^{10} - 10^{11}$ cm$^{-2}$, making them suitable for COBRA analysis; in order to detect the electron densities above the substrate surface, high QD density is required.

Next we took rodscan data and performed COBRA analysis on the GaN QD samples, allowing us to produce electron density results of each sample. We used a 15 keV beam energy for all QD samples, which allowed us to reach 10 irreducible Bragg rods. Only one of our samples, sample no. 3 (RMBE1197A), demonstrated structural coherence between the GaN QDs on bulk GaN. QDs for this sample were grown using a nitridation temperature of 185°C.

Rodscan data and the COBRA model fit for the first four Bragg rods of sample no. 1 are shown in Figure 3.9 as a representative example. We observe that the rodscans look qualitatively very similar to those of the bare GaN substrate, including very similar minimum intensities between Bragg peaks. A very subtle difference is that the substrates with minimally coherent QDs have slightly higher sloped “shoulders”
Figure 3.8: AFM images of as-deposited GaN QDs grown on Kyma bulk GaN substrates. All images were taken of a 500 nm × 500 nm surface area. Sample nos. 1-5 correspond to growth temperatures 550°C, 250°C, 185°C, 185°C, and 185°C, respectively. Sample nos. 1-4 growths were performed with medium brightness nitrogen plasma conditions, and sample no. 5 was growth with a high-brightness nitrogen plasma condition.
Figure 3.9: Several rodscans (blue line) plotted with their corresponding COBRA model fits (red line) for sample no. 1 with GaN QDs on a Kyma bulk GaN substrate (RMBE1145, Kyma4b). The model fit is good, giving an FOM of 0.1225.

on the sides of each Bragg peak, resulting in more flat behavior in the middle between Bragg peaks than in the bare substrate. For all samples, the COBRA models produce very good fits for these samples; the FOMs were all lower than 0.14.

Using COBRA, we reconstructed the electron densities for each of the five bulk substrates with GaN QDs. In Figure 3.10, we show the two line profiles along the out of plane direction. We observe that sample nos. 1, 2, 4, and 5 exhibit very similar behavior to the bare substrate in that the electron density intensities above the nominal surface drop off very quickly (within several atomic layers). Since these samples with QDs appear the same as a bare substrate, this means that COBRA was not able to construct a structure factor that includes the QD atomic layers. This is directly indicative that the QDs have different crystallographic orientations and therefore phases than the substrate, and we cannot form a structure factor and
Figure 3.10: Electron density line profiles for the five QD samples on bulk GaN. Only sample no. 3 (RMBE1197) demonstrates significant coherence.

reconstruct electron densities for those layers. In short, because the electron densities for sample nos. 1, 2, 4, and 5 look qualitatively indistinguishable from the bare substrate electron densities, the QDs demonstrate very limited coherence with the substrate.

We emphasize that when we don’t observe QDs in the COBRA electron density results, this simply means the QD atomic layers are not structurally coherent with the substrate; this does not indicate that the QD layers are not present. The AFM images indicate that QDs are present, but are not able to determine coherence of
the QDs with the substrate. COBRA is incredibly sensitive to any phase changes from the substrate phase, and requires a consistent phase and therefore QDs that are atomically registered with the substrate for QD electron density reconstruction.

Although sample nos. 1, 2, 4, and 5 did not demonstrate any significant QD coherence, sample no. 3 (RMBE1197A) demonstrated an excitingly large degree of QD coherence. In Figure 3.10, we observe for sample no. 3 that coherent atomic layers persist above the nominal substrate surface; we see a long, ~10-12 layer drop-off of electron density above the substrate surface. In the wurtzite GaN crystal structure, there are two monolayers of gallium and nitrogen in each unit cell. Since two gallium (or equivalently, nitrogen) peaks occur within the height of one unit cell, we can conclude that approximately five QD unit cell layers are coherent with the substrate, equating to ~2.6 nm of QD layers being coherent. For this sample, AFM analysis returns a QD height value of $4.0 \pm 0.6$ nm. This discrepancy could indicate that not all of the crystalline atomic layers in the QD are coherent with the substrate. It is also likely that the first and possibly second atomic layer of the QDs formed an oxide and therefore would not be reproducible by COBRA, although those layers would be measured by AFM, resulting in a 0.5–1.0 nm larger QD height from the AFM data than the COBRA data. Regardless, sample no. 3 was the only sample we measured in which we found significant QD coherence with the substrate.

We observe that the electron density peak heights between the atomic planes of gallium and the atomic planes of nitrogen do not maintain the same relative ratio between samples. Specifically, gallium to nitrogen peak height ratios for the bare substrates and sample no. 1 are relatively small, which they are larger in sample nos. 2-5. We highlight two possible explanations for this. First, atomic vibrations may be higher in some samples than in others, potentially affecting the height and width of the electron density peaks. Rodscans for all samples were taken under the same temperature conditions; however, some samples were under X-ray irradiation.
longer than others due to alignment and experimental issues, which may increase sample temperature slightly during measurement. Perhaps more likely is that the issue of manually selecting the location in the electron density results to generate a line profile. There is a possibility that the location chosen to visualize a line profile was not chosen exactly down the center of a line of atoms. It is possible then to see inaccurate peak heights and widths. In either case, we could verify the ratio of gallium to nitrogen by integrating the electron density in 3D space to determine the number of electrons associated with each atom and compare to the theoretical values of 31 electrons in gallium and 7 electrons in nitrogen. In short, we see no reason to believe that the gallium to nitrogen ratio is significantly different than the ideal ratio for GaN for all our samples; rather, we believe the apparent discrepancy in gallium to nitrogen peak height ratios are a result of small experimental variations or variations in visualizing the electron density results.

Additionally, we observe that the homoepitaxial QDs in the GaN system using droplet epitaxy appear to form dots with a “wetting” characteristic shape on the substrate surface. This is compared to a non-wetting QD characteristic shape observable in other, heteroepitaxial studies. The wetting behavior is understandable since there is no step in surface energy between the substrate and QDs so the surface tension would be very low. Because there is a tail of atomic layer occupations on the surface as opposed to a QD geometric shape, COBRA alone cannot be used to determine both the crystallinity and the coherence of the QDs. For this reason, STEM techniques are especially valuable. While STEM is not able to provide information on nearly the same length scales as COBRA, it is very powerful in providing snapshots of insight to crystallinity and coherence of the sample.

In order to gain more insight on the QD coherence of these samples, we perform STEM on two samples: one that demonstrated limited coherence (sample no. 1) and
one that demonstrated significant coherence (sample no. 3) according to COBRA.

STEM micrographs for sample no. 1 (550°C nitridation temperature) are shown in Figure 3.11. In all micrographs, the surface of the substrate is easily visible, and the 2D projection of the QDs appearing as clouds within ~5 nm of the substrate surface. All micrographs were taken such that we view the \(m\)-plane \(\{10-10\}\) (illustrated in 3.11b), making individual atoms in the substrate and coherent portions of the QDs visually distinguishable and allowing us to determine crystallinity and orientation. We observe in all micrographs very little structural coherence of the QDs. Not only do we see QD regions with very little or undetermined crystallinity (3.11a), but we also see QD regions with crystallinity rotated with respect to the substrate (3.11c-d). This is supported by looking at fast Fourier transforms (FFTs) of various portions of the projection through the QDs. Visually, in 3.11c, the atomic planes of the QD appear as horizontal streaks as opposed to individual points, indicating the QD has a crystalline orientation rotated about the \(c\)-axis. The FFT confirms that we see periodicity in the out of plane direction but negligible periodicity in plane, confirming that the crystal is rotated around the \(c\)-axis with respect to the substrate crystal orientation. Furthermore, we observe various different orientations in one micrograph in 3.11d. FFT confirms that there are crystal orientations that demonstrate either 2-fold or 6-fold symmetry that is very obviously rotated and demonstrates a different symmetry compared to the GaN substrate. These STEM micrographs provide great insight into the limited structural coherence of GaN QDs in sample no. 1 by illuminating that the QDs are mostly either not strongly crystalline or have crystallinity rotated with respect to the substrate.

In contrast to sample no. 1, we observe significant QD coherence in the STEM micrographs for sample no. 3 (Figure 3.12). In this set of micrographs, we observe the sample from the \(a\)-plane \(\{11-20\}\) (illustrated in 3.12b). Notably, in this sample, we are not able to clearly observe the original substrate surface like we are with sample
Figure 3.11: STEM micrographs of sample no. 1 (RMBE1145, Kyma4b) viewed from the m-plane, demonstrating limited coherence between QDs and substrate.
Figure 3.12: STEM micrographs of sample no. 3 (RMBE1197) viewed from the a-plane, demonstrating a large degree of coherence between QDs and substrate.

no. 1. Instead the sample appears to have a rolling surface with height variations up to about 3 nm. These height variations are slightly less than but similar to the QD height observed using AFM, further indicating the GaN QDs are highly coherent with the GaN substrate. We use FFT on regions of the samples to confirm that the QDs demonstrate the same crystallographic symmetry as the substrate.

These STEM and FFT studies for both sample nos. 1 and 3 confirm the conclusions from our COBRA electron density results: QDs in sample no. 1 exhibit a low degree of coherence, while QDs in sample no. 3 exhibit a very high degree of
Table 3.2: QD morphology and coherence results as determined by COBRA for samples of GaN QDs grown on bulk GaN before and after annealing.

<table>
<thead>
<tr>
<th>QD sample no.</th>
<th>Sample name</th>
<th>QD density (cm⁻²)</th>
<th>QD diameter (nm)</th>
<th>QD height (nm)</th>
<th>Coherent QDs?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RMBE1145, Kyma4b</td>
<td>2.20×10¹¹</td>
<td>23 ± 6</td>
<td>1.3</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>RMBE1194A</td>
<td>1.40×10¹¹</td>
<td>20</td>
<td>3.2</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>RMBE1197A</td>
<td>2.1×10¹¹</td>
<td>22</td>
<td>4</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>RMBE1145, Kyma4b-RTA</td>
<td>1.1×10¹¹</td>
<td>21 ± 10</td>
<td>1.4</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>RMBE1194B</td>
<td>4.8×10⁹</td>
<td>57 ± 25</td>
<td>19 ± 6</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>RMBE1197B</td>
<td>9.9×10¹⁰</td>
<td>18 ± 5</td>
<td>4 ± 1</td>
<td>Y</td>
</tr>
</tbody>
</table>

Coherence. They also provide additional insight for sample no. 1 in particular in that the QDs are at least partially crystalline, but largely rotated with respect to the substrate.

3.5 Rapid thermal anneal

We employ a rapid thermal anneal (RTA) on several samples (sample nos. 1-3) in order to investigate the effects of annealing on structural coherence. The proposed mechanism is that the thermal, activation energy provided by the anneal can promote atomic mobility in the QD structures and allow them to shift to a lower energy, more coherent configuration. The thermal anneal was computer-controlled for accuracy and reproducibility, and it entailed a 20 s ramp up to 850°C for a 30 s anneal in a nitrogen-rich (1,000 sccm) environment. The annealing temperature 850°C was chosen as a compromise between high temperature annealing, which demonstrated severe coarsening or destruction of the QDs, and low temperatures, which have negligible effect on the morphology or the crystallinity.

Following the RTA of samples nos. 1-3, we observed through AFM a range of QD coarsening; the density reduced and the QD size changed by varying amounts (Table 3.2). This coarsening is visible in the AFM images (Figure 3.13, right), which we compare to AFMs before the RTA (Figure 3.13, left). The amount of coarsening for
Figure 3.13: AFM images of as-deposited (left) and annealed (right) GaN QDs grown on Kyma bulk GaN substrates. All images were taken of a 500 nm × 500 nm surface area.
Figure 3.14: Electron density line profiles for QD samples as-deposited (left) and annealed (right). No significant difference in coherence is observed before and after annealing.

Sample nos. 1 and 3 does not change the density or dimensions enough to affect our ability to perform COBRA; therefore, these samples are suitable for our X-ray studies. The QD density of sample no. 2 is questionable with respect to sufficient detection for our X-ray studies, and may perhaps affect our ability to determine whether the RTA increased the coherence of the QDs. In fact, it is possible that the RTA induced decomposition of the QDs in sample no. 2, allowing the nitrogen to evaporate and leaving gallium droplets with an order of magnitude larger diameters and heights.

We compare the electron density results for the annealed QD samples (Figure
to the electron density profiles for the sample QD samples before the RTA (Figure 3.14, left). No difference in electron density drop-off above the surface is found between the as-deposited and the annealed samples, indicating no difference in QD coherence. This could be because the RTA is not at a high enough temperature to allow for atomic mobility that affects crystal structure, even though the QD dimensions are affected. Because we do see some change in QD morphology, if there is an RTA temperature that improves coherence without destroying the QDs, it is likely not much higher than our current RTA temperature of 850°C. Because we did not observe any difference in QD coherence after applying a RTA on our early samples (nos. 1-3), we did not perform an anneal on any of the later samples (nos. 4-5).

3.6 Discussion on bulk GaN substrate quality and availability

It is important to note for this project the limitations due to bulk GaN substrate quality and availability. On more than one occasion, our substrate supplier had an incredibly limited stock of c-plane, single-crystal, bulk GaN substrates available for purchase. In these cases, we had little choice in substrate quality (i.e. number of macrodefects or cracks, miscut). In the case of the RMBE1263 and RMBE1265 growth run in particular, the only substrates available were of the lowest quality “rider”, or grade D). According to AFM, these substrates appeared appropriate for QD growth, but once we began measurements at Argonne National Lab, we found many of the substrates to in fact be polycrystalline, despite being advertised as single-crystal, to the degree where we could not obtain an orientation matrix to align the sample (Figure 3.15). In short, the substrate quality impeded our ability to obtain rodscans and perform COBRA analysis. For future work, it would be prudent to find a more reliable substrate supplier. Ammono (out of Poland) is another commercial c-
Figure 3.15: Two representative photo images of our X-ray detector signal during sample alignments at Argonne National Lab. We observed textured rings instead of Bragg peaks, indicating polycrystalline instead of single-crystal substrates. As a result, we were not able to measure rodscans. (a) location where we would expect the (102) peak for QD sample RMBE1266, and (b) location where we would expect the (113) peak for QD sample RMBE1267.

We look forward to the improvement of bulk GaN production so that high-quality, commercial substrates will be more readily available to us, other researchers, and the nitride device industry.

3.7 Conclusions

In this work, we studied bare, bulk GaN substrates and GaN QDs grown on bulk GaN in order to study the structural coherence of homoepitaxial nitride QDs grown using droplet epitaxy. Between the two bare, bulk GaN substrates we studied, we found the substrate from Kyma to be better suited for these studies because of smoother surface properties and faster electron density drop-off behavior in the COBRA results. We measured rodscans at Argonne National Lab for five QD samples and reconstructed their electron densities using COBRA, allowing us to analyze QD coherence with the substrate. We found that the lowest nitridation temperature (185°C) with a medium nitrogen plasma condition was the only QD sample to demonstrate QD coherence. These results were confirmed using STEM. Additionally,
we explored the effect of rapid thermal annealing on QD coherence, and found that the anneal has little to no effect. Proposed next steps and future studies are discussed in Chapter VI.

In conclusion, we would like to discuss the issue of error and noise in phase-retrieval results from X-ray rodscans. Phase-retrieval analysis methods are still fairly new, and errors in this analysis can not be treated in the same way as traditional XRD. It remains an ongoing challenge to analyze errors in the phase-retrieval context. Zhou et al. describe in detail a method of analyzing uncertainties of COBRA electron density results, which is essential towards identifying the level of noise in these experiments and the confidence in results at low values. However, this calculation is complex and cumbersome to perform on many samples. For the purposes of this experiment, we have confidence in distinguishing little QD coherence from significant QD coherence using COBRA. A useful way of approximating the noise and supporting our qualitative results is to assume the electron density values in the interstitial sites are at the approximate level of noise. We can then compare the peaks to this approximate noise value. In order to quantitatively determine the number of QD atomic layers which are coherent with the substrate and more rigorously describe the degree of QD coherence between samples, we would perform the COBRA error analysis in order to identify the electron density peaks that are above the level of noise.
CHAPTER IV

Impact of carrier localization on recombination in InGaN quantum wells and the efficiency of nitride light-emitting diodes: insights from theory and numerical simulations

4.1 Motivation

Indium gallium nitride (In$_x$Ga$_{1-x}$N) has transformed solid-state lighting by enabling efficient light-emitting diodes (LEDs)\textsuperscript{98} and lasers\textsuperscript{99} in the short-wavelength part of the visible spectrum. Despite their commercial success, InGaN devices suffer from several issues that may be further exacerbated by the localization of carriers. The internal quantum efficiency (IQE) of InGaN LEDs reduces at high power (efficiency droop),\textsuperscript{100} which is further exacerbated for LEDs operating at longer wavelengths (green-gap problem).\textsuperscript{22} Auger recombination, a three-carrier scattering process that consumes the energy of a recombining electron-hole pair to excite another electron higher in the conduction band (electron-electron-hole Auger, or \textit{eeh}) or another hole lower in the valence band (hole-hole-electron Auger, or \textit{hhe}), has been demonstrated to be an important non-radiative recombination mechanism in nitride LEDs and a plausible cause of the efficiency-droop and green-gap problems.\textsuperscript{6,7,101–103}
Despite extensive studies on carrier localization in InGaN, its impact on radiative and Auger recombination remains unclear. Optoelectronic device simulations frequently use the virtual-crystal approximation (VCA), which models alloys as a crystalline solid with interpolated properties between the two end compounds. However, virtual crystals miss several physical features of disorder, such as translational symmetry breaking and the localization of carriers. InGaN alloys in particular contain statistically random composition fluctuations at the nanometer scale\(^ {104-106}\) that spatially localize carriers\(^ {107-110}\). Several studies have reported the effect of InGaN composition fluctuations on radiative and Auger recombination. Simulations by Yang et al. found radiative recombination rates in devices with fluctuating-alloy quantum wells (QWs) to be much higher than virtual-crystal alloys.\(^ {109}\) In contrast, Auf der Maur et al. used atomistic tight-binding and reported optical matrix elements and radiative recombination coefficients in fluctuating alloys as smaller than those of virtual crystals.\(^ {8}\) Experimentally, part of the reduction of the quantum efficiency of LEDs has also been attributed to hole localization.\(^ {111}\) With regards to Auger recombination, there are also conflicting results. Work using a semi-empirical model suggests that hole localization reduces Auger recombination,\(^ {112}\) but optical experiments indicate that localization increases Auger recombination.\(^ {113}\) Further studies are therefore necessary to understand the effects of alloy fluctuations on InGaN devices.

### 4.2 Project objectives

In this work, we use Schrödinger-Poisson simulations in the effective-mass approximation to investigate the effect of carrier localization on the radiative and Auger recombination rates in polar and non-polar InGaN QWs. This computational method has been validated for nanostructures and local fluctuations in previous work.\(^ {114-116}\) We compare results for simulations incorporating alloy composition fluctuations to virtual-crystal InGaN QWs and to virtual-crystal bulk InGaN to assess the modifica-
tion of recombination rates by localization. We find that carrier localization increases Auger recombination rates by one order of magnitude more than radiative rates, and is therefore detrimental to the IQE of InGaN LEDs and lasers.

4.3 Computational methods and analysis

We first demonstrate that the radiative and Auger recombination rates of localized carriers in InGaN are approximately proportional to the weighted-averaged overlaps squared of the localized wave functions of the recombining carriers. The detailed derivation (provided in the supplementary material) is similar to the analysis for one-dimensional confinement in QWs. The ratio of radiative coefficients of fluctuating or virtual-crystal alloys compared to bulk virtual crystals is determined by

\[
\frac{B_{\text{fluct/VCA,QW}}}{B_{\text{VCA,bulk}}} = \frac{F_{\text{eh}}^{\text{fluct/VCA,QW}}}{F_{\text{eh}}^{\text{VCA,bulk}}}, \tag{4.1}
\]

where

\[
F_{\text{eh}}^{2} = \frac{\sum_{1,2} f_{1}(1 - f_{2}) \int \psi_{1}^{*}(r) \psi_{2}(r) dr |^{2}}{\sum_{1,2} f_{1}(1 - f_{2})} \tag{4.2}
\]

are weighted-averaged overlaps squared of localized electron and hole envelope functions \(\psi\), weighted by the Fermi-Dirac occupation numbers \(f\). We subsequently demonstrate that the ratio of Auger coefficients of fluctuating (virtual crystal) alloys versus bulk virtual crystals is given by

\[
\frac{C_{\text{eeh}}^{\text{fluct/VCA,QW}}}{C_{\text{eeh}}^{\text{VCA,bulk}}} = \frac{F_{\text{eeh}}^{\text{fluct/VCA,QW}}}{F_{\text{eeh}}^{\text{VCA,bulk}}}, \tag{4.3}
\]

for the case of \(eeh\) Auger recombination, where

\[
F_{\text{eeh}}^{2} = \frac{\sum_{1,2,3} f_{1}f_{2}(1 - f_{3}) V_{QW} | \int \psi_{1}^{*}(r) \psi_{2}^{*}(r) \psi_{3}(r) dr |^{2}}{\sum_{1,2,3} f_{1}f_{2}(1 - f_{3})}, \tag{4.4}
\]
are weighted-averaged triple overlaps squared of localized electron and hole envelope functions (and similarly for hhe Auger). Polkovnikov and Zegrya derived a similar expression for Auger recombination in QWs.\textsuperscript{117} Shockley-Read-Hall rates are also approximately proportional to by similar arguments as discussed in Ref. 103 for localization along the QW confinement direction. Parameters for the carrier-density and temperature-dependent bulk and coefficients are obtained from density functional theory (DFT) results.\textsuperscript{31,118} For a carrier density of $10^{19} \text{ cm}^3$ and a temperature of 300 K (typical LED operating conditions) the bulk coefficients are $B_0 = 2 \times 10^{11} \text{ cm}^3/\text{s}$, $C_{eeh,0} = 2 \times 10^{31} \text{ cm}^6/\text{s}$, and $C_{hhe,0} = 4 \times 10^{31} \text{ cm}^6/\text{s}$.\textsuperscript{7,31} DFT calculations include alloy disorder but the simulation cells are too small (32 atoms) to exhibit localization of carriers, and thus are approximately closer to VCA results. For the bulk Shockley-Read-Hall coefficient we used a typical experimental value ($6 \times 10^7 \text{ s}^{-1}$).\textsuperscript{7}

We subsequently numerically evaluate wave functions and energies for realistic 3 nm InGaN QWs with a commercial Schrödinger-Poisson solver (nextnano\textsuperscript{69}). We consider the effect of the occupied states (for 99% of the carriers) in the effective-mass approximation for all heavy hole, light hole, and crystal-field split-off bands, which are all occupied by carriers and have been found to be important in, e.g., Ref. 119. For $\text{In}_x\text{Ga}_{1-x}\text{N}$ compositions spanning the blue-green spectrum (10-35%), we simulated ten random 3D alloy distributions using the method described in Ref. 109. Since several studies have reported the distribution of indium and gallium atoms in InGaN to be statistically random,\textsuperscript{105,116,120,121} we use a random-number generator to assign the atomic occupancies (indium or gallium) on a real-space point grid according to a binary (Bernoulli) distribution based on the nominal indium composition. We subsequently average the composition in the volume ($\pm 1.2 \text{ nm}$ along the $x$, $y$, and $z$ directions) surrounding each grid point to better approximate atom probe tomography (APT) results. The resulting alloy distributions reproduce the statistical distributions of APT data and follow binomial composition distributions.\textsuperscript{105,122} Indium
clusters form solely due to random composition fluctuations. We also incorporated a Gaussian out-of-plane distribution profile. Fermi levels and state occupancies are evaluated in the rigid-band approximation. The rigid-band approximation ignores carrier interactions and assumes that carrier energies and wave functions are not affected by carrier occupations, i.e., the conduction/valence structure of doped or excited semiconductors is the same as the un-doped ground-state one. The validity of the approximation has been demonstrated in, e.g., carrier transport in thermoelectric materials. Interactions between localized holes may affect their electronic structure, but we expect their localization properties to remain similar upon the inclusion of interactions. Our main results are evaluated with a spatial averaging of the alloy composition and a simulation grid spacing of 0.6 nm, which is smaller than the localization length of 1-3 nm (5-10 nm) for holes (electrons) reported for InGaN. We validated the conclusions of our study with calculations on a finer grid (details below and in supplementary materials). We examine two QW polarization extremes: the polar, c-plane and the non-polar, a-plane (1120) orientation. These two polarizations also approximate no screening and complete screening of polarization by carriers, respectively. Realistic devices with externally applied voltages and screening by free carriers lie between these two polar and non-polar extremes. Our study therefore enables us to separate the effect of polarization and localization on carrier overlaps and recombination rates.

4.4 Wave function and recombination rates

We then evaluate the weighted-averaged double and triple wave-function overlaps to quantify the impact of localization on the recombination rates. Our results show that both electrons and holes are localized by composition fluctuations (Figure 4.1), similar to published work. Overlaps for virtual crystals obey selection rules, and only electron and holes states of the same wave vector demonstrate nonzero
Figure 4.1: (a,b) A randomly generated, local InGaN distribution in a nominal In$_{0.15}$Ga$_{0.85}$N polar quantum well ($z \parallel c$) at (a) the $xy$-plane ($c$-plane) at the QW center ($z = 1.5$ nm) and (b) a cross-section in the $xz$-plane ($y = 0$ nm). Electron ($c; z = 2.3$ nm) and hole ($d; z = 0$ nm) wave functions for the alloy distribution in (a,b).

---

94
Figure 4.2: Individual electron-hole wave-function overlaps between the first 40 electron and 40 hole states, illustrating the symmetry breaking introduced by alloy fluctuations. (a,b) Wave-function overlaps for a virtual-crystal alloy in the (a) polar and (b) non-polar orientation. (c,d) Wave-function overlaps for a fluctuating alloy in the (c) polar and (d) non-polar orientation. The weighted-averaged matrix element is larger for the fluctuating alloy because the reduction of individual matrix elements by spatial separation is overcompensated by the larger number of allowed electron-hole combinations due to translational symmetry breaking.\(^\text{14}\)

overlap. However, random alloy fluctuations and carrier localization break the translational symmetry and wave-vector conservation, allowing nonzero overlaps between any hole and electron states. These nonzero overlaps are typically much smaller than 1 (Figure 4.2) since carriers localize in different spatial regions. Overall, however, by including the contribution from all occupied states, localization increases the electron and hole weighted overlaps compared to virtual crystals (Figure 4.2). For In\(_{0.15}\)Ga\(_{0.85}\)N QWs, and for a carrier density of \(10^{19}\) cm\(^3\) and a temperature of 300 K, alloy fluctuations increase the electron-hole weighted-averaged overlap by 64% (31%) over the virtual crystal in the polar (non-polar) orientation.

Our calculations demonstrate that both the radiative and the Auger coefficients are larger in fluctuating alloys than in virtual crystals; however, Auger coefficients in-
Figure 4.3: Radiative ($B$) and Auger ($C_{eeh}, C_{hhe}$) coefficients of InGaN calculated as a function of indium concentration for polar and non-polar orientations ($10^{19}$ cm$^{-3}$ carrier density, 300 K carrier temperature). In all cases, alloy composition fluctuations increase the coefficients compared to virtual-crystal alloys.\textsuperscript{14} 

Carriers increase significantly more than radiative coefficients, indicating an overall detrimental effect of carrier localization on the LED IQE (Figure 4.3). All ten generated random alloy distributions were used for calculating $B$, but only two of these distributions were used for calculating $C$ because the triple wave-function overlaps are computationally intensive. The two distributions were those that gave $B$ values closest to the ten-distribution average and thus best represent the randomness of the alloy, which likely explains why the $C$ values for a given indium concentration have a smaller relative spread than the $B$ values. For a carrier density of $10^{19}$ cm$^{3}$ and a temperature of 300 K, we observe a 30-50% increase in $B$ and 390-540% increase in $C$ in fluctuating alloys compared to virtual crystals.
4.5 Comparison to fine grid and rough alloy composition

To facilitate comparison with previous work that uses atomistic, tight-binding calculations,\textsuperscript{8,110} we performed a set of calculations for a fluctuating alloy on a finer simulation mesh. We use a fluctuating In\textsubscript{0.15}Ga\textsubscript{0.85}N alloy with a smaller grid size than the main part of our study (0.3 nm instead of 0.6 nm), which is similar to the in-plane lattice constants of InGaN. Additionally, the fine-grid calculations do not spatially average the alloy distribution in order to represent potential fluctuations at the atomic scale (Figure 4.4).

All wave functions obtained with the fine-grid calculations exhibit similar localization behavior to the coarse-grid study, except for some of the lowest-energy states. The first few electron states of the coarse-grid study, which cover the lowest 15–20 meV of electron energy states, demonstrate localization by compositions fluctuations. Above this energy, electron states are still affected by alloy fluctuations, but they become extended and cover the span of the simulation domain, while their amplitude decreases as expected for more extended states. In comparison, electron wave functions for the fine-grid study are also influenced by the alloy distribution, but they are extended even for the lowest-energy states and do not exhibit confinement. However, only the first three electron states out of the >45 states exhibit this different electron localization behavior between the coarse- and fine-grid studies. Therefore, the large majority of electrons are not affected by localization at low energies for the coarse-grid, spatially averaged alloy study, similar to the fine-grid results.

We compare weighted-averaged overlaps for fluctuating alloys in the fine-grid study (Table 4.1) to the coarse-grid results (Table 4.2), and find that most overlaps increase by factors of 1.2–2.7 when using finer grids. The exception is the non-polar electron-electron-hole overlap, for which the coarse-grid study predicts a larger overlap by a factor of 1.4. The physical origin of the exception is the reduced localization of electrons in the finer-grid study, which becomes particularly important
Figure 4.4: (a,b) A random alloy distribution for a 3 nm polar $\text{In}_{0.15}\text{Ga}_{0.85}\text{N}$ QW, representing a more atomistic approach (0.3 nm grid and no spatial averaging) than the main study (0.6 nm grid and spatial averaging). (a) shows the $xy$-plane at the QW center, $z = 1.5$ nm, and (b) shows a cross-section at $y = 0$ nm. (c,d) The square modulus electron and hole ground-state wave functions, respectively, for the random alloy distribution shown in (a,b). Each wave function is shown at their most intense, respective $xy$-planes ($z = 2.7$ nm for electron, $z = 0.3$ nm for hole). The electron is less confined than our studies with a coarser (0.6 nm) grid and spatial averaging. Calculations of $B$, $C_{eeh}$, $C_{hhe}$ uphold our main results of Auger processes increasing by a larger factor than radiative rates, resulting in alloy fluctuations decreasing internal quantum efficiency as compared to virtual-crystal alloys.\(^\text{14}\)
Table 4.1: Weighted-averaged overlap results of original In$_{0.15}$Ga$_{0.85}$N QW study with 0.3 nm grid and spatially averaged alloy composition.

<table>
<thead>
<tr>
<th></th>
<th>Polar</th>
<th>Non-polar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F^2_{eh}$</td>
<td>$F^2_{eeh}$</td>
</tr>
<tr>
<td>Fluctuating alloy</td>
<td>0.0350</td>
<td>0.0084</td>
</tr>
<tr>
<td>Virtual crystal</td>
<td>0.0079</td>
<td>0.0007</td>
</tr>
<tr>
<td>Fluctuating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>virtual crystal</td>
<td>4.4</td>
<td>12.5</td>
</tr>
</tbody>
</table>

Table 4.2: Weighted-averaged overlap results of original In$_{0.15}$Ga$_{0.85}$N QW study with 0.6 nm grid and spatially averaged alloy composition.

<table>
<thead>
<tr>
<th></th>
<th>Polar</th>
<th>Non-polar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F^2_{eh}$</td>
<td>$F^2_{eeh}$</td>
</tr>
<tr>
<td>Fluctuating alloy (averaged for 10 distributions)</td>
<td>0.0175</td>
<td>0.0068</td>
</tr>
<tr>
<td>Virtual crystal</td>
<td>0.0114</td>
<td>0.0012</td>
</tr>
<tr>
<td>Fluctuating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>virtual crystal</td>
<td>1.5</td>
<td>5.4</td>
</tr>
</tbody>
</table>

for $eeh$ Auger recombination (since it involves two less-localized electron states) and for the non-polar direction (for which the carrier localization by fluctuations is not overshadowed by the carrier separation by polarization fields).

We also observe that the non-polar virtual-crystal weighted-averaged overlaps agree very well between the two studies, while the polar virtual-crystal fine-grid overlaps are smaller than those of the original study. This is because the increased resolution in the out-of-plane direction for the polar case allows the spatial separation of electrons and holes from the polarization field to increase slightly (approximately 2.4 nm instead of 2.3 nm), which reduces the net overlap.

By examining the net effect of fluctuations in the fine-grid study compared to the original, coarse-grid results, we find that the coarser grids represent the physics and qualitative behavior very well. In all cases examined with the fine-grid calculations, we observe that fluctuations significantly increase overlaps as compared to virtual-
crystal alloys, as also observed with coarser grids and spatial averaging. Because the coarser-grid calculations predict smaller overlaps for fluctuating alloys and larger overlaps for virtual-crystal alloys, they slightly underestimate the net effect of alloy fluctuations and carrier localization on the recombination behavior of InGaN quantum wells compared to calculations using finer meshes. Importantly, the fine-grid study predicts an even greater efficiency reduction of the internal quantum efficiency due to alloy fluctuations: 9% (12%) for the polar (non-polar) orientations for the fine-grid study as compared to 4% (9%) for the same carrier density, $10^{19} \text{ cm}^{-3}$, in the coarse-grid study.

4.6 Carrier density and temperature dependence

We also evaluate the radiative and Auger coefficients over a realistic range ($10^{17}$–$10^{20} \text{ cm}^{-3}$) of carrier densities for LED operation for the alloy distribution producing the $B$ value closest to the average $B$ value for the ten initial distributions. Compared to virtual crystals, we find that composition fluctuations consistently increase Auger coefficients more than the radiative ones for all examined carrier densities (Figure 4.5). At low carrier concentrations, the enhancement of fluctuating-alloy coefficients over virtual crystals is particularly prominent because localization effects are strongest for the lowest-energy states that dominate state occupancies. At high carrier densities, state occupancies are dominated by less confined states and fluctuating-alloy coefficients approach those of virtual crystals. The lower smoothness of the Auger curves (Figure 4.3(c)-(f)) compared to the radiative ones (Figure 4.3(a) and (b)) is likely due to statistical limitations and could be improved with additional calculations for more random alloy distributions and larger simulation cells.

We note that $C_{hhe}$ is larger than $C_{eeh}$ for polar QWs, in agreement with bulk DFT calculations.\textsuperscript{31} Interestingly, for non-polar QWs (and, approximately, screened polar QWs), we find $C_{eeh}$ to be larger than $C_{hhe}$ by a factor of 1–14, depending
Figure 4.5: Radiative ($B$) and Auger coefficients ($C_{eeh}$, $C_{hhe}$) of In$_{0.15}$Ga$_{0.85}$N as a function of carrier density (300 K carrier temperature). In all cases, composition fluctuations increase Auger coefficients more than the corresponding radiative coefficients. $^{14}$
on the carrier density, which agrees with a recent experimental report that $C_{eeh}$ is larger than $C_{hhe}$ by a factor of 1–12.\textsuperscript{124} Our results therefore reveal that the interplay of QW composition, carrier localization, polarization fields, quantum confinement, carrier density, and screening plays an important role in determining the dominant Auger recombination type ($eeh$ or $hhe$) in InGaN QWs.

We further determine the temperature dependence of the radiative and Auger coefficients in fluctuating alloys and compare to virtual crystals. As with the carrier density study, for the temperature study we use the alloy distribution that produced the $B$ value closest to the average of the ten initial distributions. Auger coefficients increase more than the radiative ones due to alloy fluctuations consistently across the 100–600 K temperature range (Figure 4.6). Auger coefficients for fluctuating alloys increase with increasing temperature up to 100 K, at which point they start to decrease. This peak with temperature was observed experimentally in GaInP alloys, which also exhibit carrier localization.\textsuperscript{125} At low temperatures (below 100 K), carriers are strongly localized and wave-function overlaps are small, therefore the spatial separation of carriers dominates and reduces the radiative coefficients compared to virtual crystals. Above 100 K, the occupations of higher-energy extended states increases with increasing temperature and localized states contribute a smaller fraction to the weighted-averaged overlaps, hence the coefficients decrease with temperature and approach VCA results.

4.7 Comparison to published work

Our analysis also resolves the apparent disagreement between published studies on the impact of carrier localization on the radiative and Auger rates. Our results agree with other reported simulations in that for the ground state and other individual state overlaps, fluctuations reduce the overlaps compared to virtual crystals.\textsuperscript{8,112} However, by considering the occupations of all electron and hole states for realistic
Figure 4.6: Radiative ($B$) and Auger coefficients ($C_{eeh}$, $C_{hhe}$) of In$_{0.15}$Ga$_{0.85}$N as a function of carrier temperature ($10^{19}$ cm$^{-3}$ carrier density). Except for very low temperatures, Auger coefficients of fluctuating alloys increase over those of virtual crystals by larger factors than radiative coefficients.\textsuperscript{14}
carrier densities and temperatures under device-operating conditions, we predict an overall net increase of the radiative recombination coefficients. Our results therefore disagree with Auf der Maur et al. and Karpov et al., and they are consistent with the work of Yang et al. and Shahmohammadi et al., who found localization to increase recombination rates.\textsuperscript{109,113}

### 4.8 Effects of carrier localization

Our results further demonstrate that carrier localization exacerbates the efficiency-droop and green-gap problems of both polar and non-polar LEDs. The difference between the IQE of fluctuating alloys and virtual crystals is up to 9\% (12\%) in the polar (non-polar) orientation (Figure 4.7). Additionally, the ratio decreases more with increasing indium concentration for fluctuating alloys than for virtual crystals, indicating that alloy fluctuations aggravate the green-gap problem (Figure 4.7). Random composition fluctuations are an inevitable result of alloying, therefore the detrimental effects of carrier localization on the efficiency droop of InGaN LEDs cannot be eliminated. One avenue to avoid localization is to use binary nitride compounds and tune the gap by engineering the polarization fields or the quantum-confinement dimensions.\textsuperscript{44,55} Another promising avenue to mitigate Auger recombination is to reduce the steady-state carrier density with fewer, thicker InGaN QWs (e.g., a double heterostructure design up to 12 nm thick).\textsuperscript{101,126}

### 4.9 Discussion

While atom probe tomography is a very powerful characterization technique, the atomic detection efficiencies are lower than 100\%, which may affect the spatial behavior of fluctuations.\textsuperscript{127} Hence, results obtained with APT for InGaN\textsuperscript{105,121} may depend on experimental and reconstruction parameters.\textsuperscript{128,129} Transmission electron mi-
Figure 4.7: (a,b) Internal quantum efficiency (IQE) of In$_{0.15}$Ga$_{0.85}$N quantum wells with fluctuating alloys (dashed line) and virtual crystals (solid line). Fluctuating alloys exhibit lower IQE than virtual crystals. (c,d) Ratio of radiative ($B$) to Auger ($C$) coefficients as a function of indium concentration (300 K carrier temperature, $10^{19}$ cm$^{-3}$ carrier density). The ratio decreases with indium concentration faster for fluctuating alloys than for virtual crystals, consequently alloy fluctuations exacerbate the green-gap problem.\textsuperscript{14}
croscopy is another useful technique for providing local composition information but may also produce experimental artifacts since radiation damage from overexposure produces localized strain contrast and the artificial appearance of clusters.\textsuperscript{105,122,130} Due to the limitations of each technique, it is important to use complementary characterization techniques to verify distribution results. Rigutti et al. used effective mass calculations to correlate APT and micro-photoluminescence results for AlGaN, allowing them to statistically correct the composition measured using APT.\textsuperscript{127,131} Additionally, for InGaN QWs, Mehrtens et al. found scanning transmission electron microscopy (STEM) and APT composition profiles to be in agreement.\textsuperscript{132} We note that our calculations did not rely on APT data, but we instead theoretically generated statistically random atomic distributions that agree with experimental distributions. Deviations from a completely random distribution (e.g., indium clustering) cause further localization of carriers and exacerbate the droop and green-gap problems. We thus expect that our qualitative results also hold for other disordered atomic distributions that deviate from complete randomness.

Excitons were not included in this work since they dissociate at the high carrier densities where droop occurs,\textsuperscript{44} or at temperatures near room temperature.\textsuperscript{113} However, our results at low temperatures and low carrier concentrations, for which excitons have been observed,\textsuperscript{113} need to be reexamined with higher-order calculations that account for excitonic effects. Since both radiative and Auger recombination involve one free electron-hole pair recombination (which needs to be updated with the corrected excitonic wave functions), we expect that, to lowest order of approximation, exciton formation increases both the radiative and Auger rates in a similar fashion. Hence, we do not anticipate that our qualitative conclusions regarding the localization effects on the rates and efficiency with be affected by the inclusion of excitonic corrections.

This work highlights the detrimental effects of carrier localization on radiative and
Auger recombination. Another interesting factor to consider is the effect of carrier localization on SRH recombination. Other work has suggested that carrier localization can benefit device performance from the perspective of non-radiative SRH recombination; localization can spatially separate carriers from interacting with defects, thereby allowing for functioning devices despite such high defect densities compared to many other material systems.\textsuperscript{3} We note that the studies mentioned in this chapter accept electrons to exhibit more extended than localized behavior and therefore would likely interact with a significant number of defects. Yet, if this is in fact true, we anticipate that localization may provide both beneficial and detrimental effects on device performance when its effects are studied for all recombination mechanisms. However, at this point, it is unclear whether there would be certain conditions or regimes where the beneficial effects would be particularly prominent. It would be interesting to study the combined effect of localization on SRH, radiative, and Auger recombination.

4.10 Conclusions and future work

In conclusion, we quantify the effect of carrier localization due to alloy fluctuations on the IQE of InGaN QWs. Localization increases the Auger rates more than radiative rates by one order of magnitude. Therefore, alloy fluctuations exacerbate the efficiency-droop and green-gap problems of InGaN LEDs and are detrimental to the IQE of InGaN LEDs and lasers.

In addition to our primary study, we simulated a In\textsubscript{0.15}Ga\textsubscript{0.85}N fluctuating alloy with a grid size similar to the in-plane lattice constant of InGaN (0.3 nm) and performed no spatial composition averaging. We found qualitatively identical results to our calculations with the 0.6 nm grid and spatial averaging. In fact, we find that the fine-grid results demonstrate an even larger reduction in efficiency due to alloy fluctuations than the coarser-grid study. Therefore, our fine-grid study supports the physics
and qualitative trends obtained with the coarser grid and spatial, alloy-composition averaging.

Future studies include considering Coulomb interactions between carriers. Holes in particular might experience repulsive Coulombic effects because of their highly localized behavior, which could reduce Auger recombination. However, the attractive effects of the electrons and holes could increase recombination. We anticipate incorporating Coulombic effects may change the recombination performance, however it is unclear what net effect we may observe on IQE, efficiency droop, and green gap.
CHAPTER V

Strain relaxation and band alignment of 
GaAsSb/GaAs quantum wells

5.1 Motivation

GaAsSb quantum wells (QWs) embedded in GaAs emit photons with wavelengths near 1.3 \( \mu \text{m} \) and are promising for optical communications applications.\textsuperscript{36,133–136} To achieve 1.3 \( \mu \text{m} \) emission, Sb mole fractions of approximately 35\% are often used. GaAsSb alloys with higher Sb concentrations are difficult to achieve due to phase separation and increased Sb-As exchange during growth, causing broadening of the heterointerface, as well as degradation in material quality and optical performance.\textsuperscript{135,137–141} Therefore, alloys with Sb percentages in the 15–40\% range are the most useful for device applications and are most commonly reported in the literature.\textsuperscript{134,135,137,142–144}

GaAsSb/GaAs QWs are particularly interesting from both an applications and a fundamental physics point of view because they have been shown to demonstrate both type-I and type-II band alignment depending on the QW composition and strain.\textsuperscript{36,42,135} Type-I band alignment occurs if the quantum well exhibits both electron and hole confined states within the well. Type-II band alignment in GaAsSb/GaAs occurs if the hole is confined within the quantum well, but the conduction band edge
of the quantum well is higher than that of the barrier material, causing a spatial separation of electrons and holes. Additionally, there exists an intermediate case where the conduction band edge of the quantum well is below that of the barrier material, but the energy difference between the conduction band edges is small enough that no confined electron state exists. We define this case to be type-II because the spatial overlap between electron and hole wave functions is small since the electron is not confined to the QW. Being able to predict band alignment type is imperative because it greatly affects device engineering.

Experimental and theoretical reports in the literature have demonstrated both type-I and type-II band alignment in GaAsSb/GaAs QWs without a clear consensus on how to predict the alignment type. Several publications have reported band diagrams for pseudomorphic and strain-relaxed QWs for a single composition. Additionally, Wang et al. reported conduction and valence band bowing results by fitting three points: theoretical band edges for GaSb/GaAs QWs, band edges of GaAs, and experimental band edges obtained for a single GaAsSb/GaAs composition. Pryor and Pistol used computational methods to report band edge properties of pseudomorphic and strain-relaxed GaAsSb/GaAs QWs at 0 K. However, these works still do not agree on the band alignment type of GaAs$_{0.65}$Sb$_{0.35}$/GaAs QWs, which is within the most important compositional range for applications. In order to properly engineer GaAsSb/GaAs QWs for device applications, we need to understand and predict the band alignment type under various growth conditions.

5.2 Project objectives

In this work, we use semiconductor device simulations to study the effects of composition, strain, and QW width on the band alignment type of GaAsSb/GaAs QWs. Our goal is to understand and predict the band alignment type as a function of structure and composition. Our results can be used to predict the band-edge
properties of GaAsSb/GaAs QWs for the full compositional range, including the effect of partial strain relaxation.

5.3 Methods

Our computational method is based on established device simulation techniques. We employ $8 \times 8 \mathbf{k} \cdot \mathbf{p}$ theory to calculate the band-edge properties of GaAsSb/GaAs QWs using the nextnano commercial software package with a self-consistent Schrödinger-Poisson solver. Strain was calculated in the continuum elasticity approach with elastic constants from Vurgaftman et al.\textsuperscript{146} Band offsets within the model-solid theory\textsuperscript{147} (values for which were obtained from Ref. 75) and deformation potentials (values for which are obtained from Ref. 148) were subsequently used to calculate the conduction and valence bands of the system. Band bowing was included with a conduction band bowing parameter of 0.37 eV and a valence band offset bowing of -1.06 eV (i.e., band gap bowing of 1.43 eV).\textsuperscript{146} Further information about the structure of the calculations can be found in published work.\textsuperscript{69,149,150} Grid spacing, barrier thickness between QWs, and the number of calculated electronic states were chosen to converge the desired band energies to within 10 meV. All calculations were performed at a temperature of 300 K.

5.4 Band edge and alignment results

Our results for 7 nm GaAs\textsubscript{0.65}Sb\textsubscript{0.35}/GaAs QWs demonstrate type-I band alignment for strain-relaxed systems and type-II alignment for systems with pseudomorphic strain (Figure 5.1). The strain-relaxed system exhibits one confined electron state and several confined hole states. We calculate the conduction-band offset between the QW and GaAs barrier to be 0.18 eV and the valence-band offset to be 0.40 eV. All confined energy levels, as well as the heavy-hole and light-hole band
Figure 5.1: Band-alignment diagrams of (a) strain-relaxed (type-I) and (b) pseudomorphic (type-II) GaAs\textsubscript{0.65}Sb\textsubscript{0.35}/GaAs QWs. Solid black lines denote the band edges and dashed blue lines show the energy levels of confined states. Energy levels are defined with respect to the unstrained GaAs valence band maximum.\textsuperscript{15}

edges, in the strain-relaxed system are degenerate, as expected under unstrained conditions. In contrast to the strain-relaxed system, the pseudomorphic system exhibits no confined electron states, and the QW conduction band edge is at a higher energy than the GaAs conduction band by 0.10 eV. The confined hole level degeneracy is split by strain, as is the heavy and light hole band edge within the QW, resulting in a valence band offset of 0.54 eV. These band offsets are consistent with published experimental\textsuperscript{144} and theoretical\textsuperscript{36,42,140} values.

We also find that strain-relaxed GaAsSb/GaAs QWs result in type-I band alignment, while pseudomorphic systems lead to type-II band alignment, regardless of QW composition or thickness (Figure 5.2). This is evident by the conduction band edges for the strain-relaxed QWs with any non-zero Sb mole fraction being lower than the GaAs conduction band edge and containing an electron confined state, and the conduction band edges for pseudomorphic QWs rising above that of GaAs. Thinner QWs exhibit larger band gaps than wider QWs due to quantum-confinement effects. Due to band bowing, a minimum QW conduction band energy occurs at Sb mole fraction of 60%. Therefore, for Sb mole fractions in the 0–60% range, the conduction-band offset can be increased by increasing the Sb mole fraction. Strain-relaxed (pseudo-
Figure 5.2: Conduction and valence band edges for pseudomorphic (solid lines) and strain-relaxed (dashed lines) GaAsSb/GaAs QWs as a function of Sb mole fraction and QW width. Band edges are defined by the first confined energy level, if present. The horizontal black line indicates the conduction band edge of GaAs; QW conduction band edges below (above) this line result in type-I (type-II) band alignment. Energy levels are defined with respect to the unstrained GaAs valence band maximum.\textsuperscript{15}

Pseudomorphic) QWs resulted in type-I (type-II) band alignment, even when we modified the GaAsSb band bowing parameter ±10%. These results indicate that while strain, QW thickness, and composition all affect band offsets, strain is the dominant factor that determines the band alignment type.

Since our results indicate that strain is the primary determinant of the band alignment type, we explore the effects of partial strain relaxation on the band alignment. Understanding the effect of partial strain release, which is an active area of research,\textsuperscript{136,151–154} on the band edges of GaAsSb/GaAs heterostructures and other materials is important for device design. Our results also apply to GaAsSb quantum dots embedded in GaAs, (which often experience strain relaxation\textsuperscript{37,139,155,156}) since the electronic properties of anisotropic quantum dots with only one strongly confined dimension can be approximated by QWs. To model partial strain relax-
ation, we artificially modify the lattice constant of the QW material to adjust the mismatch with the barrier, thereby altering the strain. For example, to achieve 50% strain relaxation, we set the QW lattice constant to a value halfway between the bulk QW and barrier materials. The calculation then employs deformation potentials to determine the effect of the resulting partial strain (which is lower than the full, pseudomorphic amount) on the band edges. No material parameters other than the QW lattice parameter were directly modified as they are dependent on composition and not strain.

For this strain study, we focus on Sb mole fractions in the 15–35% composition range, as this range is most useful for device applications. We note that our partial strain calculations result from a one-dimensional model of a partially strain-relaxed system. Experimentally, strain relaxation is achieved through dislocations. These dislocations may result in a QW with local variations in strain, while our model is a partial strain model that considers uniform, average strain throughout the QW.

5.5 Band alignment type transition

Our partial strain relaxation results indicate that a band-alignment-type transition from type-II to type-I occurs at different amounts of partial strain relaxation depending on the GaAsSb QW composition. Band edges for systems with partial-strain relaxation span the space between the strain-relaxed and pseudomorphic extremes (Figure 5.3). For the 7 nm GaAs$_{0.65}$Sb$_{0.35}$ QW with less than 45% strain relaxation, the alignment of the QW conduction-band edge with respect to the GaAs barrier does not lead to a confined electron state in the well. In this case, the system demonstrates type-II alignment. For more than 45% pseudomorphic strain relaxation, the energy of the QW conduction band edge dips below the GaAs conduction band edge by a sufficient amount to allow a confined electron state. Therefore, once 45% of the strain...
Figure 5.3: Conduction and valence band edges for 7 nm GaAsSb/GaAs QWs as a function of Sb mole fraction. Band edges are shown for strain-relaxed and pseudomorphic wells, as well as for partially strain-relaxed wells with 20%, 40%, 60%, and 80% of pseudomorphic strain relaxed. The right panel zooms in on the region enclosed by the dashed box in the left panel.

relaxes, the 7 nm GaAs$_{0.65}$Sb$_{0.35}$ QW switches from type-II to type-I. The type-II to type-I band alignment transition occurs at 37% and 33% strain relaxation for 7 nm QWs with 25% and 15% Sb mole fraction, respectively. Similarly to the 7 nm QW case, as the Sb mole fraction increases, QWs of thicknesses 3, 5, and 10 nm require more strain relaxation to achieve a type-II to type-I transition (Figure 5.4).

5.6 Comparison to published experimental work

Our calculated results of band alignment type and optical transition energies in GaAsSb/GaAs QWs under various strain conditions agree well with experimental measurements (Table 5.1). Sadofyev and Samal demonstrated 7 nm GaAs$_{0.64}$Sb$_{0.36}$ QWs to have type-II band alignment and an optical transition energy ranging from 0.946-0.972 eV, depending on growth conditions. These results agree well with our pseudomorphic well calculations, which have a dominant transition energy of 0.996 eV. A similar experimental study reported type-II alignment but with a smaller transition energy of 0.943 eV. This result more closely agrees with our calculations.
for a QW with 20% strain relaxation, which has a transition energy of 0.934 eV, indicating that strain may have partially relaxed within the QW. An additional work by Wang et al. used a series of QW barrier alloys, compositions, and thicknesses as well as a fitting model to determine that 7 nm GaAs$_{0.643}$Sb$_{0.357}$ QWs have type-I band alignment and a transition energy of 0.928 eV.\textsuperscript{36} Although this experiment reports a type-I band alignment for wells that are estimated to be pseudomorphic,\textsuperscript{158} the conduction-band offset is small with a large relative error (24 ± 24 meV).\textsuperscript{36} Therefore, no confined electron states exist in the QW, in agreement with our type-II definition. Furthermore, the AlGaAs cladding layers used in the experiment\textsuperscript{36} provide additional barriers that confine electron states within the combined GaAsSb/GaAs volume and raise the electron state energies.\textsuperscript{142} The lower of the two dominant transition energy results for a 7 nm pseudomorphic QW with type-II band alignment is 0.908 eV, which agrees reasonably well with the experimental transition of 0.928 eV reported by Wang et al.\textsuperscript{36} The difference may be attributed to the additional confinement energy of electrons in the GaAsSb/GaAs region due to the AlGaAs cladding layers. Although the work by Wang et al.\textsuperscript{36} attributes the band alignment type transition to the bowing of the conduction band of GaAsSb, our results suggest that strain relaxation (and not the Sb mole fraction) determines the band-alignment type. Other works report type-II band alignment at 77 K for GaAsSb QWs with Sb mole fractions of 0.30-0.36,\textsuperscript{137,142,144} which agrees with our results for highly strained QWs.

### 5.7 Discussion and conclusions

We note that our calculated results for QWs can also be applied to estimate the band alignment type in other quantum-confined GaAsSb structures, such as embedded GaAsSb/GaAs quantum dots. Our results are particularly applicable for dots with dimensions that are significantly smaller along the out-of-plane direction than in the plane.\textsuperscript{139,156,159,160} GaAsSb/GaAs quantum dots find applications in gas sens-
Figure 5.4: The percentage of pseudomorphic strain relaxation for which the type-II to type-I band-alignment transition occurs in GaAsSb/GaAs QWs with a well width of 3, 5, 7, and 10 nm.\textsuperscript{15}

Table 5.1: Comparison of our calculated results for the band alignment type and optical transitions in GaAsSb/GaAs QWs with published experimental work.

<table>
<thead>
<tr>
<th>QW Parameters</th>
<th>Band alignment type</th>
<th>Transition energy (eV)</th>
<th>% strain relaxed</th>
<th>Band alignment type</th>
<th>Transition energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 nm, $x \approx 0.36$\textsuperscript{143}</td>
<td>Type-II</td>
<td>0.946-0.972</td>
<td>0</td>
<td>Type-II</td>
<td>0.966</td>
</tr>
<tr>
<td>7 nm, $x \approx 0.35$\textsuperscript{143}</td>
<td>Type-II</td>
<td>0.943</td>
<td>20</td>
<td>Type-II</td>
<td>0.934</td>
</tr>
<tr>
<td>7 nm, $x \approx 0.357$\textsuperscript{36}</td>
<td>Type-I</td>
<td>0.928</td>
<td>0</td>
<td>Type-II</td>
<td>0.908</td>
</tr>
</tbody>
</table>

In summary, we performed semiconductor device simulations to investigate the
influence of composition, width, and partial strain relaxation on the band alignment type of GaAsSb/GaAs QWs. Our results indicate that pseudomorphic wells result in type-II band alignment and strain-relaxed well results in type-I band alignment. This is true for any Sb concentration and for QW thicknesses larger than 3 nm, making strain the primary determinant of band alignment type. The band alignment of QWs with partial strain relaxation depends on the interplay of composition, thickness, and strain. Band alignment and transition energy results from our calculations agree well with published experimental data. Our results not only provide insight on the band alignment of QWs, allowing easier device design and analysis, but are also a useful tool for understanding band alignment in thin quantum dots.
CHAPTER VI

Summary, implications, and future work

This dissertation describes three studies of compound semiconductor nanostructures for light-emitting diode applications. It combines insights of experiment, computation, and theory to illucidate the interplay of growth, structure, electronic, and optical properties of quantum wells and dots. Furthermore, we describe how knowledge of these properties aids in device development. To conclude this work, a brief summary of each project is included below along with the impact of the project on the field and discussions of possible future directions.

6.1 Surface studies of GaN QDs on bulk GaN

In Chapter III, we described surface characterization studies of GaN QDs grown on bulk GaN substrates using the droplet epitaxy method. We used surface X-ray diffraction and COBRA, an iterative phase-retrieval technique, to determine that the only growth recipe that produced QDs with a significant degree of coherence used the lowest nitridation temperature studied (185°C). This perhaps indicates that the growth mechanisms in this system are not dominated by kinetics, but perhaps by thermodynamic processes instead. Our surface X-ray diffraction results were supported through STEM and AFM measurements. Experiments were limited by the quality and availability of commerical, bulk GaN substrates.
This project speaks to the importance of implementing bulk substrates in semiconductor device development and commercialization in order to increase device performance and yield. It also speaks to the need to implement creative solutions towards addressing the efficiency droop and green gap issues in InGaN LEDs. Quantum dots are an interesting avenue towards reducing device issues and the droplet epitaxy method is an exciting possible growth method because of its ability to perform homoepitaxial growth.

Initial future work would include exploring a larger matrix of growth conditions to evaluate all reasonable parameters achievable using the droplet epitaxy method. This would allow us to evaluate the droplet epitaxy growth method for homoepitaxial growth in the InGaN material system. Ideally, these additional growth studies would be performed using substrates with higher and more reliable quality, possibly the Ammono substrates out of Poland mentioned in Section 3.6. Reliable bulk GaN substrates are not only imperative for the continuation of this project, but also for the commercial progress of InGaN devices in general.

Once successful growth parameters have been identified and characterized for GaN QDs on bulk GaN, QDs can be alloyed with indium to form InGaN QDs, and the QDs can be embedded for device fabrication. Structural characterization studies would continue to be important at this stage to ensure high crystal quality is maintained following alloying and capping. Optical characterization methods would also become highly important at this stage in the project. In order to construct and evaluate devices, doping and barriers layers could be included, followed by electrical characterization methods.

Finally, another interesting future study would be to look further at whether QD growth is dominated by kinetics or thermodynamics. This can be explored from both experimental and theoretical directions. Experimentally, the effect of different RTA parameters can be studied on QD morphology and structural coherence. In
particular, we could study slightly higher annealing temperatures (900-1,000°C). It is possible that higher temperatures may simultaneously increase the coherence while also inducing species evaporation, thereby destroying the QD; we anticipate that this is increasingly likely as we approach temperatures of 1,000°C as the MOCVD GaN growth temperatures are ~1,100°C. We could therefore study shorter annealing times (10-25 s) as we increase the annealing temperatures in order to maintain QD morphology. RTA studies could be performed before or after QD capping. Perhaps there are also growth variables that can be tuned to determine the effect of kinetics and thermodynamics during growth. Theoretically, simulations such as density function theory (DFT), molecular dynamics, or kinetic Monte Carlo and could be used to study whether QD formation in this system is a thermodynamically driven process.

6.2 Effect of alloy fluctuations on recombination in InGaN QWs

In Chapter IV, we described our work on the effect of carrier localization due to alloy fluctuations on InGaN QWs. In particular, we study the effect of radiative and Auger recombination within the QWs at typical LED operating conditions and as a function of alloy composition, crystal orientation, carrier density, and temperature. Our results show that alloy fluctuations reduce individual transition matrix elements by the separate localization of electrons and holes, but this effect is overcompensated by the additional transitions enabled by translational symmetry breaking and the resulting lack of momentum conservation. Hence, we find that localization increases both radiative and Auger recombination rates, but that Auger recombination rates increase by one order of magnitude more than radiative rates. Importantly, we demonstrate that localization has an overall detrimental effect on the efficiency-droop and green-gap problems of InGaN LEDs.
Our main study was performed with simulation grid sizes of 0.6 nm and employed a spatial average of the alloy distribution in order to reproduce APT data and statistics. We performed an additional study with a grid size of 0.3 nm, which is on the order of the GaN lattice constant, and used a binary composition (GaN or InN) at each lattice site with no spatial averaging. We found that the results from the fine grid study reproduced the physics and qualitative results of the coarse grid study. In fact, the fine grid study produces an even larger reduction in efficiency due to alloy fluctuations. This fine-grid study validates the alloy distributions used in the main study.

Our work provides important physical insight on the impact of carrier localization on recombination rates in InGaN LEDs. Additionally, we resolve an apparent discrepancy in the literature on the effect of composition fluctuations on both radiative and Auger recombination rates. Our methodological developments describe the fundamental physics of radiative and Auger carrier recombination in the presence of carrier localization in semiconductor alloys as a function of alloy composition, carrier density, and temperature. In addition, our results explain how InGaN alloy fluctuations reduce the internal quantum efficiency and aggravate the efficiency-droop and green-gap problems in nitride LEDs, which has significant technological implications for the performance and design of commercial LEDs and lasers.

There are several avenues of interesting future work related to this project. The first is to include excitonic effects in the simulations. We anticipate that this will affect the $hhe$ results more than the $eeh$, as the holes are more strongly localized. We are in the process of exploring Coulombic effects for the InGaN system in collaboration with Dr. Stefan Schulz and his coworkers. Another interesting future project is to apply these same techniques to the AlInN material system. Because of the very large bandgap of AlN and very small bandgap of InN, it might be possible to localize both holes and electrons, which may allow for more control over radiative and Auger
recombination processes. Finally, it would be interesting to explore wave function overlaps and recombination rates for ordered GaN/InN alloys.

6.3 Strain and band alignment of GaAsSb/GaAs QWs

Finally, in Chapter V, we used commercial software based on $k \cdot p$ theory to study the effects of confinement, composition, and strain (including partial strain) on the band-edge properties of GaAsSb/GaAs QWs. We found that regardless of QW width or composition, strain-relaxed wells result in type-I band alignment and pseudomorphic wells result in type-II band alignment. Since strain is the primary determinant of band alignment type, we further studied partial strain relaxation. The band edges of QWs with 20, 40, 60, and 80% of the pseudomorphic strain span the range between the strain-relaxed and pseudomorphic band edges. This allows us to determine a strain percentage at which the band alignment type of a GaAsSb/GaAs QW transitions from type-II and type-I, which depends non-trivially on QW thickness and composition.

Our work is important both from a fundamental and an applications point of view. On the fundamental side, our calculations uncover the relationship between QW thickness, composition, strain, and band alignment type. In terms of applications, our work can be used to design and predict optoelectronic properties of GaAsSb/GaAs QW devices, such as LEDs and laser diodes used for optical fiber communications. Additionally, our results apply towards devices utilizing QDs with small height-to-width ratios.

Interesting future work would be to explore the effect of shape, size, composition, and strain of GaAsSb/GaAs QDs as opposed to QWs. GaAsSb/GaAs QDs are of particular interest for intermediate band solar cells, charge-based memory devices, novel photonic devices, and fundamental physics studies. To the best of our knowledge, the effect of partial strain relaxation has not been reported for the
GaAsSb/GaAs QD system.

6.4 Concluding words

This concludes my Ph.D. dissertation. I hope the information herein has been interesting, informative, and thought-provoking. I look forward to the continued progress on the included topics by not just me, but also the research community as a whole.
BIBLIOGRAPHY


