STRAINED INDIUM ARSENIDE/GALLIUM ARSENIDE LAYERS FOR QUANTUM CASCADE LASER DESIGN USING GENETIC ALGORITHM

A Dissertation presented to
the Faculty of the Graduate School
at the University of Missouri

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by

DAVID W. MUELLER
Dr. Gregory Triplett, Dissertation Supervisor
MAY 2015
The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

STRAINED INDIUM ARSENIDE/GALLIUM ARSENIDE LAYERS FOR QUANTUM CASCADE LASER DESIGN USING GENETIC ALGORITHM

presented by David W. Mueller,
a candidate for the degree of Doctor of Philosophy, Doctor of Electrical and Computer Engineering and hereby certify that, in their opinion, it is worthy of acceptance.

________________________________________
Dr. Gregory Triplett

________________________________________
Dr. Scott Kovaleski

________________________________________
Dr. James Keller

________________________________________
Dr. Paul Miceli
ACKNOWLEDGMENTS

This dissertation is the written culmination of my doctorate research. The course of my dissertation cannot solely be expressed in words. It is however, held in the memories of the many people with whom I consulted and otherwise interacted with during my graduate studies. There are countless individuals who, along the way, impacted my path. Life is counted by the memories we share and I cherish them all.

To my advisor and mentor, Gregory Triplett, you stuck with me all along the way. I thank you for taking me in to your group and giving me many great years of learning, expanding, and most importantly consistent faith in my ability. To the members of my committee, Jim Keller, Scott Kovaleski, and Paul Miceli. Each of you played important roles intellectually in the contributions of this work, but more importantly each of you are colleagues to be admired. You each take great care in the journey of your students and I am fortunate to have learned from each of you.

There have been many laboratory colleagues with whom I have collaborated. To Denzil Roberts, Stanely Ipke, Reggie Jeff, and Hisham Abbas: we rocked the lab. To the many others and all the undergraduate researchers, thank you for the many contributions and best wishes on your journey.

To the many individuals across campus who’s efforts often go unnoticed, thank you. Specifically, I’d like to thank Shirley Holdmeier and Tami Beatty of the Department of Electrical and Computer Engineering. We shared many (probably too many) years. The Electron Microscopy Core, including Cheryl Jensen and Louis Ross, you taught me so much and gave me valuable opportunities.
The journey of my dissertation is only part of the story of my graduate experience. Many thanks to Noah Manring for giving me the opportunity to teach young minds as a part of the teaching fellows program. My teaching mentor, Robert O’Connell, has been an invaluable source of teaching wisdom and a most patient and caring person. To all my students, each of you have helped forge this path with me and I wish you all the best in your careers.

Finally, my graduate journey included joining the storied St. Pat’s Board to help continue the many traditions of Engineers’ Week. To all the fellows and friends gained. Specific thanks to the club’s advisors, Roger Fales and Steven Borgelt.

There are many more unnamed individuals who contributed in some way to my life as a graduate student. The many memories will forever be a reminder of the path walked.
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS** ......................................................... ii

**LIST OF TABLES** .......................................................... viii

**LIST OF FIGURES** ......................................................... ix

**ABSTRACT** ................................................................. xv

**CHAPTER**

1 Introduction ............................................................... 1

1.1 Motivation ............................................................... 2

1.1.1 Mid-Infrared (3-5 μm) Devices ................................. 2

1.1.2 InAs and GaInAs growth on GaAs (111)B .................... 5

1.2 Thesis Outline and Scope ............................................ 6

1.2.1 Background ......................................................... 6

1.2.2 Growth Study of InAs on GaAs (100) and (111)B .......... 7

1.2.3 Device Modeling and Optimization ........................... 7

1.2.4 Conclusion and Design Candidate ............................ 9

2 Mid-Infrared Sources: Quantum Cascade Lasers .................. 10

2.1 QCL Theory and Operation ......................................... 10

2.1.1 Principles of Quantum Mechanics ............................ 11

2.1.2 Schrödinger’s Wave Equation ................................. 13

2.1.3 The Infinite Quantum Well .................................... 17
2.1.4 Electronic States in Multiple Finite Semiconductor Quantum Wells ................................. 20
2.1.5 Multiple Quantum Well Active Region ................................................................. 22
2.2 Mid-Infrared QCLs in literature ................................................................. 25

3 Active Region Design ................................................................. 28
3.1 Extended Fundamentals ................................................................. 30
3.2 Rate Equation ................................................................. 32
3.3 The "Intersubband Toolbox" ................................................................. 34
3.3.1 Tunneling ................................................................. 34
3.3.2 Number of Periods ................................................................. 36
3.3.3 Upper-level confinement ................................................................. 37
3.4 Traditional QCL Design and Variations ................................................................. 38
3.4.1 Active Region Design Bounds ................................................................. 40
3.5 Strain Layer incorporation in the Active Region ................................................................. 42

4 Growth Study using GaAs (100) and (111)B ................................................................. 46
4.1 Molecular Beam Epitaxy ................................................................. 46
4.1.1 RHEED Characterization ................................................................. 50
4.2 Experimental Setup ................................................................. 51
4.3 Strained and unstrained layers on GaAs (111)B ................................................................. 52
4.3.1 Experimental Procedures ................................................................. 52
4.4 Results ................................................................. 54
4.4.1 GaAs Results ................................................................. 54
4.4.2 AlGaAs Results ................................................................. 56
4.4.3 InAs Results ........................................... 58

5 Genetic Algorithms and Use for Design ......................... 63

5.1 Multi-objective Genetic Algorithms .......................... 63

5.1.1 Genetic Algorithm Terminology ........................... 64

5.2 Genetic Algorithms and QCLs ................................. 68

5.3 Genetic Algorithm as Implemented ......................... 69

5.3.1 Initial Population ..................................... 72

5.3.2 Candidate Generation: Crossover and Mutation ....... 73

5.3.3 QCL Input Files ...................................... 74

5.3.4 nextnano$^3$ ........................................... 75

5.3.5 Data Collection ....................................... 75

5.3.6 Fitness Evaluation .................................... 77

5.3.7 Parent Selection ...................................... 78

6 Genetic Algorithm Simulation Results ....................... 80

6.1 Top Candidates by Individual Fitness Rank on GaAs .... 80

6.2 Average Fitness Values ................................. 88

6.3 Top Candidates for Modified Objectives .................. 92

7 Summary and Concluding Remarks ............................ 97

APPENDIX

A Project Code ............................................. 100

A.1 Main Code Routines .................................... 100

A.1.1 main.m ............................................. 100
A.1.2 createnn3in.m ........................................... 102
A.1.3 getnn3template.m .................................... 103
A.1.4 runnn3.m .............................................. 104
A.1.5 getqcldata.m .......................................... 104
A.1.6 findfitness.m ......................................... 106
A.1.7 getnextgen.m .......................................... 108

A.2 Subroutine Code ........................................ 109
A.2.1 createcstruct.m .................................... 110
A.2.2 calcgenes.m ......................................... 110
A.2.3 convgenes.m ........................................ 111
A.2.4 genes2ar.m .......................................... 112
A.2.5 binary_crossover.m ................................. 113
A.2.6 wmutate.m .......................................... 114
A.2.7 genrandgene.m ...................................... 114
A.2.8 proxcheck.m ...................................... 115
A.2.9 findlevels.m ....................................... 115
A.2.10 obj1fit.m ......................................... 117
A.2.11 signf.m ........................................... 117

A.3 Template File ........................................... 117

BIBLIOGRAPHY .............................................. 127

VITA ......................................................... 134

vii
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Growth Procedures</td>
<td>53</td>
</tr>
<tr>
<td>4.2 Growth Experiments Process Conditions</td>
<td>53</td>
</tr>
<tr>
<td>5.1 General Simulation Parameters for the GA</td>
<td>71</td>
</tr>
<tr>
<td>5.2 Application Specific Simulation Parameters for the GA</td>
<td>72</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Atmospheric transmission spectrum showing chemical absorption points. Source: US Navy</td>
<td>2</td>
</tr>
<tr>
<td>2.1 Single Quantum Well - AlGaAs,GaAs,AlGaAs - showing the separation of energy levels.</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Potential function representation of the infinite quantum well.</td>
<td>17</td>
</tr>
<tr>
<td>2.3 Simulation of a single quantum well with widths of 5.0nm (a) and 2.5nm (b). Shows the increase in energy level separation by variation of the well width.</td>
<td>19</td>
</tr>
<tr>
<td>2.4 Three well laser structure showing the energy levels and the electron probability functions.</td>
<td>22</td>
</tr>
<tr>
<td>2.5 Simulation of a diagonal transition QCL.</td>
<td>25</td>
</tr>
<tr>
<td>3.1 Band structure of the theoretical device proposed in [53]. Superlattice structure under applied electric field with alignment of the ground state of the upstream well just below the second excited state of the downstream well.</td>
<td>29</td>
</tr>
</tbody>
</table>
3.2 Schematic conduction band diagram of the first proposed quantum cascade laser. ................................................................. 31

3.3 Graphic representation of the three state lifetimes pertinent to the efficient operation of the quantum cascade active region. ................. 32

3.4 Two states coupled by a tunnel barrier. The lifetime of the upper state is controlled by the thickness of the coupling barrier. ................. 35

3.5 Electron tunneling into the continuum (left): $\tau_{esc} \approx e^{-2\kappa b L_b}$ and $\tau_2 < \tau_1$. Electrons are not allowed to tunnel into a minigap (right): $\tau_2 > \tau_1$. 38

3.6 Computed transmission of an injector designed to provide maximum reflection at the energy of the excited state. Credit: [4] .................. 39

3.7 Schematic conduction band profile and design of the three-well active region. a) Two-quantum separate quantum wells. b) Coupled wells with applied electric field. Resonance of the lower wells causes energy level splitting. c) Third well added upstream with the energy level aligned with the upper level of the second well. Produces diagonal transitions. d) energy level is above the upper level of the second well producing vertical transitions. Reproduced from [4] ......................... 40

3.8 Simulation of a strained layer active region QCL on GaAs(111). Shows reduction of emission wavelength and reduced operating voltage (electric field) over [57]. Reproduced from [56] ................................. 43

3.9 Three states coupled by a tunnel barrier. ................................. 44

3.10 Frequency mixing by introduction of a fourth state into the active region. 45

4.1 Substrate side of EPI 930 MBE system during operation. .......... 47

4.2 Source shroud side of the EPI 930 MBE system during operation. . 47
4.3 Control racks containing 2408 temperature controllers and 2kW Sorensen dc power supplies for each material cell of the MBE. Other equipment includes a RHEED gun controller, vacuum controllers, and vacuum meters.

4.4 Representative drawing of a typical MBE system with control and processing equipment. Copied from [4].

4.5 Illustration of RHEED oscillations during monolayer growth. [62].

4.6 50 nm layer of GaAs on (100) surface at $T_s = 580$ °C. SEM image nothing remarkable and confirms surface is flat.

4.7 SEM images taken at 30,000X magnification* ($\sim 10$ µm$^2$) of GaAs (111)B samples with varying $T_s$ and BEP ratio. All samples had some form of surface defect, but the sample grown at $T_s = 570$ °C and BEP ratio 15:1 would be acceptable for high-quality electronic and photonic devices. *The sample with $T_s = 620$ °C is shown at 60,000X magnification in order to reveal the surface structure more clearly.

4.8 10 µm$^2$ AFM images of Al$_{0.45}$Ga$_{0.55}$As structures with varying $T_s$ and a BEP ratio of 20:1. The smoothest surface was obtained at $T_s = 600$ °C, where the surface qualities are very similar to the optimal GaAs surface obtained in the preceding sample set. A transition in defect structure occurs for samples grown above and below this optimum temperature similar to the transition seen in GaAs samples.
4.9 5 μm² AFM images for 17 ML InAs samples grown with T_s = 480 °C and BEP ratio from 10-50:1. Note how plateaus of InAs are formed at the defect points on the underlying GaAs layer. As BEP ratio increases plateau size decreases while density increases. This is due to the suppression of adatom migration by increasing As4 overpressure.

4.10 5 μm² SEM images for InAs samples grown with a BEP ratio of 30:1 for two different thickness, 4 ML and 17 ML. Note the increase in plateau size as the thickness increases.

4.11 (a) Roughness and (b) Root-mean-square surface model fit for MBE-grown InAs/GaAs samples on (111)B 2° → [2̅1̅1] surfaces. Adapted from [20].

4.12 Simulation of a 3 quantum well structure with wells at 2.8 nm and barriers at 2.2 nm. The outer wells are GaAs while the middle well is binary InAs. The barriers are Al_{0.45}Ga_{0.55}As.

5.1 Example of per-bit binary crossover for an 8-bit gene sequence that produces two children from the set of parents.

5.2 Example of per-bit binary crossover for an 8-bit gene sequence that produces a mutated gene from the original and the random mutation gene.

5.3 Diagram of the flow of the genetic algorithm.

5.4 Schematic flow for the initial population.

5.5 Generic gene sequence with value ranges.
6.1 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs a) (100) and b) (111) with top fitness rank for the emission objective. 81

6.2 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs a) (100) and b) (111) with top fitness rank for the LO phonon resonance objective. 82

6.3 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs a) (100) and b) (111) with top fitness rank for the injector level alignment objective. 84

6.4 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs a) (100) and b) (111) with top fitness rank for the gain metric objective. 85

6.5 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs a) (100) and b) (111) with top aggregate fitness rank. 87

6.6 Average fitness value over the generations for Objective 1 on GaAs a) (100) and b) (111) 89

6.7 Average fitness value over the generations for Objective 2 on GaAs a) (100) and b) (111) 89

6.8 Average fitness value over the generations for Objective 3 on GaAs a) (100) and b) (111) 90

6.9 Average fitness value over the generations for Objective 4 on GaAs a) (100) and b) (111) 91

6.10 Total aggregate average fitness value over the generations on GaAs a) (100) and b) (111) 91

6.11 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top fitness rank for the modified emission objective. 92

6.12 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top fitness rank for the modified LO phonon objective. 93
6.13 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top fitness rank for the injector alignment objective. ........................................... 94

6.14 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top fitness rank for the modified gain metric objective. ............................... 95

6.15 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top aggregate fitness rank. .......................................................... 96
ABSTRACT

Achieving high power, continuous wave, room temperature operation of mid-infrared (3-5 µm) lasers is difficult due to the effects of auger recombination in band-to-band (type I) designs. Intersubband laser designs (type II) such as quantum cascade lasers reduce the effects of recombination, increasing efficiency and have advantages in large tunability of wavelength ranges (3 µm - 25 µm and into the THz spectrum). Highly efficient quantum cascade laser designs are typically used in lasers designed for >5 µm wavelength operation due to the small offset of conduction band energy $\Delta E_c$ in lattice matched materials. Some promising material systems have been used to achieve high-power output in the first atmospheric window (3-5 µm) but still suffer from low efficiency due to the lack of electron confinement. Larger $\Delta E_c$ is attainable through the use of strained (lattice mis-matched) materials such as InGaAs/AlGaAs on GaAs. However, this material system has limitations on the traditional (100) crystal orientation due to the large strain and low critical thickness, $h_c$. The necessity for controlled two-dimensional, optical quality layer growth limits the amount of strain incorporation due to defect formation in highly lattice mis-matched layers. The material systems used in this study are GaAs (100) and (111)B, AlGaAs, and (Ga)InAs. In the initial stage of research, I found that pseudomorphic growth of highly strained InAs layers on GaAs (111)B is possible. However, the growth window is very narrow and necessitates precise control over growth temperature and anion overpressure to achieve optical quality layers. As a result, a second stage of research explores the design space made available by this finding by using genetic algorithm based design and simulation of devices with a Schrödinger-Poisson solver - nextnano.
This genetic algorithm is designed to rank candidate solutions on four important objectives for achieving novel QCL designs: operation between 3-5 µm, $E_{\text{inj}}$ alignment just below $E_3$, $E_2 - E_1 \approx E_{LO}$ for optimal scattering for depopulation of carriers at the $E_2$ energy level, and a gain metric $\tau_3(1 - \tau_2/\tau_3)$. This approach was generally successful at finding unique designs, which have promise to work. However, since the algorithm ranked candidates based on only these four objectives, several interesting designs trends emerged. Many of the designs are far from the classic QCL structure and may not emit, however the trends help reveal the important design criteria. Discussion of this approach is compared to the traditional design philosophy and suggestions for improvement are centered around incorporating more objectives to guide the algorithm’s search. While improvements to the search mechanism are certainly necessary, several candidate solutions emerged and show great promise toward the goal of using this novel surface index and material system for efficient quantum cascade lasers that operate in the first atmospheric window.
CHAPTER 1

Introduction

The aim of this dissertation work is to explore the design possibilities of quantum cascade laser (QCL) devices [1] based on strained and unstrained (Ga)InAs/(Al)GaAs structures operating in the 3-5 µm wavelength range. This is accomplished by a three step approach starting with the experimental optimization and understanding of the growth of strained InAs on GaAs (100) and (111)B surfaces, followed by a simulation-based optimized design study using a genetic algorithm, and finally analysis of optimal designs.

Devices that operate in the mid-infrared region of the electromagnetic spectrum are of great interest in many application areas. Notable application areas include molecular identification through chemical spectroscopy, atmospheric composition detection, infrared countermeasures, thermal imaging, and free-space optics [2], [3]. Important molecules, such as CO$_2$, H$_2$O, and O$_3$ absorb specific energies of mid-infrared radiation allowing for precise identification. Two atmospheric transparency regions at 3-5 µm and 8-12 µm wavelengths, with minimal attenuation, are of interest for low-loss free space communication. The sub-range between 3.2-4 µm offers the highest transmission coefficient at >90%. Figure 1.1 shows the transmittance of infrared light in the atmosphere with labels for the absorption by gaseous chemical compounds.
These applications require efficient and high-power optical devices with spectral tunability. However, this is a significant hurdle for current generation devices. For laser applications, QCL designs offer the best efficiency and tunability and offer the most promise for overcoming these hurdles for next generation devices [4], [5].

1.1 Motivation

1.1.1 Mid-Infrared (3-5 µm) Devices

The 3-5 µm range corresponds to photon energies in the 0.25-0.41 eV range. This energy requirement has forced many designers to focus on laser designs using band-to-band transitions (Type I). Using these designs, it is difficult to achieve high-efficiency allowing for room temperature continuous wave operation due to inherent losses asso-
associated with Auger recombination [6]. Quantum cascade lasers [1] utilize intersubband transitions in the conduction band (making them unipolar) and show enhancement in Auger lifetime by greater than an order of magnitude [7], [8], resulting in increased efficiency over Type I designs.

The history of QCLs starts with ideas first expressed by the use of intersubband transitions in superlattice structures [9]. However, it wasn’t until the development of controlled epitaxy techniques such as Molecular Beam Epitaxy (MBE) [10] and metal-organic vapor phase epitaxy (MOCVD) that such devices could be produced. These techniques allowed for precise atomic layer control and enable the realization of sharp discontinuities in semiconductor heterostructures [11]. These developments allowed for research into a new class of devices that took advantage of the quantum size effects on charge carriers at the nanometer scale. For many devices this provided a large range of tunability by simply changing the thickness of quantum wells.

Traditional QCL designs focus on the high-mid to far-infrared regions due to the small intersubband spacing possible with lattice matched or strain-compensated systems – GaAs/AlGaAs, GaSb/AlSb, GaInAs/GaInSb/InP, etc. Of these systems, InP based devices are dominant for the shorter mid-infrared wavelengths due to the larger conduction band offset $\Delta E_c$ when matched with AlGaAs/InGaAs. However, InP based-devices have strict limitations on the value of $x$ in $\text{In}_x\text{Ga}_{1-x}\text{As}$ for a lattice match without introducing large amounts of strain. This limits the amount of design tuning available in the system. However, for all values of $x$, the GaAs/Al$_x$Ga$_{1-x}$As
system is lattice-matched providing one specific theoretical advantage of the system allowing for more design level tunability. Additionally, the GaAs system is technologically more mature and has achieved 3% higher wall-plug efficiencies for room temperature (RT), continuous-wave (CW) operation [12].

The availability of any value of \( x \) in Al\(_x\)Ga\(_{1-x}\)As is why this material system is used in designs with emission from 7.4 µm [13] to 190 µm [14]. Moving beyond the second atmospheric window and into the first requires a larger conduction band offset, \( \Delta E_c \). However, values of \( x > 45\% \) leads to an indirect band-gap and degraded efficiency and performance [15].

One potential solution to this problem is to intentionally incorporate non-lattice matched (strained) materials into the laser’s active region design [16]. In the case of GaAs based QCLs, incorporation of indium into the well regions forms ternary In\(_x\)Ga\(_{1-x}\)As and increases the electron affinity (GaAs 4.07 eV and InAs 4.9 eV). This decrease in the bandgap increases the achievable barrier height allowing for larger intersubband transitions while maintaining a direct band-gap and carrier confinement.

Traditional designs using the (100) surface are limited to only small incorporation of indium, only a few percent, due to three-dimensional structure formation arising from the compressive strain a result of the lattice mismatch [17], [18]. In binary InAs on GaAs, the lattice mismatch is 7% on the (100) surface; the lattice constants of InAs and GaAs are 6.06 Å and 5.65 Å, respectively. Using a novel surface-index, such as (111), research, including this work [19]–[21], have revealed the possibility of pseudomorphic, layer-by-layer formation of binary InAs on GaAs (111). Overcoming these strain related issues allows for the building and design of efficient devices, specifically QCLs, which operate in the first atmospheric windows of the mid-infrared.
spectrum. In addition, growth on the (111) surface opens up a range of design possibilities which take advantage of useful phenomena such as a built-in field from the piezoelectric effect, reduced threshold current, and increased non-linear susceptibility [16].

1.1.2 InAs and GaInAs growth on GaAs (111)B

Interest in the high-index surfaces and especially the (111)B surface of GaAs began in 1970 with the pioneering work of A. Y. Cho [10] and revolves around using the piezoelectric properties induced by strained layer incorporation [22]–[25]. This study is interested in the application of (111)-oriented multi quantum-well structures. Specifically, QCL designs, which require controlled, sustained, and repeatable growth of near atomically flat layers with thickness of 10-15 monolayers (MLs) for the thinnest layers and several microns for the thickest layers [16].

QCL devices on (111) with uniquely incorporated piezoelectric effects offer wavelength extendability towards the near infrared region through the use of nonlinear optical processes. Specifically, the incorporation of frequency conversion through the use of (111) orientation-patterned, quasi-phase matched substrates also offer additional design flexibility [26], given that high-quality layer growth can be achieved. Management of strain and control of interface quality is critical in these structures. It is necessary to optimize the growth in each layer of these multi-layer structures.

When compared to growth on the (100) surface using traditional molecular beam epitaxy (MBE) techniques, the growth parameter window on (111) is significantly narrower making optical quality growth very challenging when incorporating vastly different material systems. Much of the previous characterization of the surface qual-
ity of GaAs, AlGaAs, InGaAs, and InAs layers grown on the (111)B surface was limited to reflection high-energy electron diffraction (RHEED) studies and determination via the mirror-like qualities of the resulting samples [27]–[33]. Some studies included Nomarski optical characterization, photo-luminescence studies, and X-ray diffraction [29], [33]. A few other studies using relatively simple p-i-n diode and single quantum laser diode structures revealed that device quality layers are achievable and confirmed the utility of the piezoelectric effect in such devices [24]. The collective work is strongly correlated and reveals high quality materials can be achieved within a 30 °C narrow temperature region that lies just below the RHEED reconstruction transition from \((\sqrt{19} \times \sqrt{19})R23.4° \rightarrow (1 \times 1)HT\). The temperature of this transition is dependent on the As\(_4\) flux and can be moved up and down by increasing or decreasing the flux, respectively [33].

1.2 Thesis Outline and Scope

The following describes the layout of this dissertation and sets the scope of work completed and discusses interesting findings as well as suggestions for future work.

1.2.1 Background

Background information on the theory and operation and a review of the state of QCLs is given in Chapters 2 and 3. Chapter 3 details the traditional methods for design of a QCL and presented in [4]. Included in Chapter 3 is a brief discussion of strain related effects on QCL active region design and second harmonic generation.
1.2.2 Growth Study of InAs on GaAs (100) and (111)B

This research extends previous growth studies by examining the sensitivity of process conditions and strain on the surface quality for GaAs, AlGaAs, and InAs layers grown on (111)B 2° → [211] Si-doped substrates using nano-scale microscopy techniques, SEM and AFM, that allow for a better determination of the optimal process conditions for multi-layer photonic devices [19]. The data corresponds with the previously mentioned narrow growth window for these structures and better reveals the type and severity of defects that occur in the non-ideal growth zones. Further statistical analysis of the samples revealed pseudomorphic growth with an average surface roughness of 0.08 nm over a 10 µm area for the smoothest samples and the correlated data was used to develop a model to better understand the effects of each growth parameter [20]. The group has found that growth rate, anion overpressure, and substrate temperature are the most important factors in producing quality pseudomorphic layers for use in optical devices [19]–[21]. Further discussion and an experimental overview is given in Chapter 4. Chapter 4 concludes with a detailed discussion of experimental results of the growth studies.

1.2.3 Device Modeling and Optimization

QCLs are semiconductor hetero-junction superlattice structures designed with two parts: the injector and the active region. The active region determines the emission properties of the device. Active region designs vary, but the underlying principals are the same: quantization of energy levels in the well structures, the corresponding wavefunctions, and tunneling of the carriers. Diagonal transition active regions
traditionally consist of three quantum wells and their respective barriers. In order to provide the active region with electrons at the correct energy level, an injector region is used. The injector region is a series of quantum wells, which progressively transition the electron to the higher energy level under the presence of an electric field. The injector is designed to provide a high probability density state that lies just below the upper state of the active region.

A self-consistent Schrödinger-Poisson solver (nextnano) is used to simulate the devices. The simulation data is analyzed and evaluated against a set of objectives needed for high-efficiency and lasing action. These objectives are discussed further in the Chapter 5 and follow from the QCL theory discussed in Chapters 2 and 3. The objective-specific fitness of each device design is used in a genetic algorithm to produce further designs that are simulated and evaluated. The purpose of the genetic algorithm is to search the enormous design space for optimal designs for a given criteria. The effort given to properly characterize and evaluate designs systematically pays dividends in the reduction of experimental and fabrication effort and the increased confidence in the optimality of a given design. Discussion of particular methods and the theory of genetic algorithms is also presented in Chapter 5.

An additional outcome of this work is the analysis and comparison of simulated designs with both (100) and (111) surface orientations. Significant effort has been given to the theoretical advantages of the (111) surface. The focus is on designs, which incorporate piezoelectric and non-linear susceptibility as a means to achieve higher efficiency and wavelength extend-ability towards the 3-5 ũm range. Chapter 6 includes the results and comparisons of the simulation work.
1.2.4 Conclusion and Design Candidate

Development of the algorithms for evaluation, sorting, and reproduction of new designs constitute a major portion of this work. Work continues to progress in the development of techniques necessary for the sustained growth of optical quality strained layers on (111)B GaAs. The contribution of this work is the delivery a set of laser designs to be implemented and tested as future work beyond this dissertation. In Chapter 7, final concluding thoughts and a discussion of future work in implementation of the candidate device designs as well as improvements to the genetic algorithm’s (or other search mechanism) effectiveness are given.
Mid-Infrared Sources: Quantum Cascade Lasers

In this chapter, the theory and operating principals of the quantum cascade laser (QCL) is presented. The QCL is a unipolar semiconductor laser that uses a single carrier (electrons), to produce optical transitions between electronic states defined by spatial confinement within multi-quantum wells. Quantum engineering of the electronic wavefunctions is performed at the nanometer scale, defining the operating characteristics of the device. In principle, the QCL may be realized using most any semiconductor material system. For the purposes of this study, the focus is on the Al$_x$Ga$_{1-x}$As/GaAs and Al$_x$Ga$_{1-x}$As/In$_x$Ga$_{1-x}$As material systems. This chapter is a review of quantum mechanics pertinent to the technical theory of QCLs. The final section of this chapter briefly overviews the current state of mid-IR QCL achievements. Active region design and discussion of strained layer incorporation are discussed in the next chapter.

2.1 QCL Theory and Operation

The quantum cascade laser is a series of potential barriers and wells formed by semiconductor heterojunctions (i.e., materials with different bandgaps). The fundamental phenomena for operation is the quantum confinement of carriers and carrier tunneling. Figure 2.1 is an example of a quantum well structure formed by the double heterojunction of AlGaAs/GaAs/AlGaAs. The narrowing of the well width, on the
order of the DeBroglie wavelength, leads to separation of energy states into quantized levels. These quantized energy levels are tuned for the desired energy transitions via band structure engineering. Band structure engineering is the process of varying the well width and barrier thicknesses, and band offset to tailor specific quantum mechanical properties such as: the placement of energy levels, carrier scattering rates, optical dipole matrix elements, and tunneling times. These properties determine the population densities and transition probabilities for the various energy states.

Figure 2.1: Single Quantum Well - AlGaAs,GaAs,AlGaAs - showing the separation of energy levels.

2.1.1 Principles of Quantum Mechanics

The behavior of the QCL can be described using 1D quantum mechanical principles through application of Schrödinger’s wave equation to multiple quantum wells. The three quantum mechanical principles: energy quanta, wave-particle duality, and the uncertainty principle set the stage for understanding the operation of the QCL. Here a brief history is described before delving into the mathematics.
In 1900, Planck postulated that thermal radiation is emitted from a heated surface as discrete energy packets called quanta. Knowing this postulation, in 1905, Einstein applied this postulate to interpret experiments demonstrating the photoelectric effect, suggesting that light also contains discrete packets of energy. These particles, called photons, have energy given as $E = hv$. When a photon of sufficient energy, greater than or equal to the material’s work function, strikes a surface it will remove an electron. Excess photon energy is transferred into the kinetic energy of the photoelectron. This phenomena demonstrates the discrete nature of the photon and it’s particle-like behavior. The energy of an electromagnetic wave is given as

$$E = hv = \frac{hc}{\lambda} \quad (2.1)$$

where $h$ is Planck’s constant ($h = 6.625 \times 10^{-34}$ J-s), $v$ is the frequency of radiation, $c$ is the speed of light, and $\lambda$ is the wavelength.

Experiments involving phenomena such as the photoelectric effect and the Compton effect demonstrate the particle-like behavior of electromagnetic waves. In 1924, de Broglie postulated that since waves exhibit particle-like behavior, particles should be expected to demonstrate wave-like behavior. This hypothesis is known as the wave-particle duality principle. The momentum of a photon is given by

$$p = \frac{h}{\lambda} \quad (2.2)$$

and, similarly, the wavelength of a particle is

$$\lambda = \frac{h}{p} \quad (2.3)$$
At times, electromagnetic waves behave as particles (photons) and sometimes particles behave as waves. As it turns out, this principle of quantum mechanics also applies to other small particles such as electrons, protons, and neutrons.

A third principle, the Heisenberg uncertainty principle, given in 1927, also applies to small particles. This principle relates conjugate variables, energy & time and position & momentum. The uncertainty principle states that it is impossible to simultaneously describe the conjugate variables. For instance, it is not possible to accurately measure a particle's energy and know with certainty the time at which it had that energy. Because of the uncertainty principle, descriptions of a particle's position, for instance, are given as probabilities.

2.1.2 Schrödinger’s Wave Equation

In 1926 Schrödinger introduced a formulation that combined the principles of energy quanta given by Planck and wave-particle duality introduced by de Broglie. This wave theory is used to describe the motion of electrons in a crystal. Accurate computation of the energy states in multiple quantum well structures is essential for design of working devices. However, for the purposes of illustrating important quantum mechanical principles, a simplified version of the wave equation will now be described, followed later by the 8-band $k \cdot p$ approximation as used in the simulations. The one-dimensional, non-relativistic Schrödinger’s wave equation is given by

$$\frac{-\hbar^2}{2m} \cdot \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x)\Psi(x, t) = j\hbar \frac{\partial \Psi(x, t)}{\partial t}$$  \hspace{1cm} (2.4)
where $\Psi(x,t)$ is the wave function, $V(x)$ is the time independent potential function, $m$ is the mass of the particle. Using the technique of separation of variables it is possible to determine the time-dependent and position-dependent portions of the wave function. Assume the wave function may be written as

$$\Psi(x,t) = \psi(x) \phi(t) \quad (2.5)$$

where $\psi(x)$ is a function of $x$ only and $\phi(t)$ is a function of $t$ only.

By substitution, the wave equation is

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x)\psi(x)\phi(t) = j\hbar \psi(x) \frac{\partial \phi(t)}{\partial t}. \quad (2.6)$$

Dividing both sides by $\psi(x)\phi(t)$ the wave function becomes

$$-\frac{\hbar^2}{2m} \frac{1}{\psi(x)} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x) = j\hbar \frac{1}{\phi(t)} \frac{\partial \phi(t)}{\partial t} \quad (2.7)$$

with the left side of the wave function a function of $x$ only and the right side a function of $t$ only. The solution to time-dependent portion of the wave equation takes the form

$$\phi(t) = e^{-j(\eta/\hbar)t} \quad (2.8)$$

where the parameter $\eta$ is a separation constant. This form is a classical exponential form of a sinusoidal wave. From Equation (2.1) we have $E = \hbar \omega/2\pi = \hbar \omega$. In the previous equation $\omega = \eta/\hbar$, therefore the separation constant, $\eta$, is the total energy $E$ of the particle. The time-independent portion of Schrödinger’s wave equation can
now be written as
\[
\frac{-\hbar^2}{2m} \frac{1}{\psi(x)} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x) = E. \tag{2.9}
\]

Rearranging the equation yields
\[
\frac{\partial^2 \psi(x)}{\partial x^2} + \frac{2m}{\hbar^2} (E - V(x)) \psi(x) = 0. \tag{2.10}
\]

Schrödinger’s wave equation, for the purposes of the this work, is meant to describe the behavior of an electron in a crystal solid. The total wave function is the product of the position-dependent and time-dependent portions. Expanding Equation (2.5) gives
\[
\Psi(x,t) = \psi(x) \phi(t) = \psi(x)e^{-j(\eta/\hbar)t} = \psi(x)e^{-j\omega t}. \tag{2.11}
\]

However, the total wave function \( \Psi(x,t) \) is a complex function and therefore cannot represent a real physical quantity.

Also in 1926, a German physicist named Max Born postulated that the function \( |\Psi(x,t)|^2 \, dx \) is the probability of finding a particle between \( x \) and \( x + dx \) at a given time. This leads to a probability density function that can be written as the product of the total wave function and its complex conjugate.
\[
|\Psi(x,t)|^2 = \Psi(x,t) \cdot \Psi^*(x,t) \tag{2.12}
\]

From Equation (2.11) the complex conjugate can be written as
\[
\Psi^*(x,t) = \psi^*(x)e^{j\omega t}. \tag{2.13}
\]
The time-dependent portions of the total wave function and its conjugate cancel, leaving

\[ |\Psi(x,t)|^2 = \psi(x)\psi^*(x) = |\psi(x)|^2. \tag{2.14} \]

This result highlights that, in quantum mechanics, the position of a particle cannot be determined precisely, but rather is found in terms of a probability.

Since Equation (2.14) is a probability density function there are a few bounding conditions for its use. First, the total probability of finding a single particle in space is certain.

\[ \int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1 \tag{2.15} \]

The probability density function \(|\psi(x)|^2\) and the wave function \(\psi(x)\), at a given \(x\), must be finite and single-valued so as not to violate the uncertainty principle. Third, in the physical world the total energy \(E\) and the potential \(V(x)\) are finite, implying that the second derivative of the wave function is also finite. This means the first derivative is continuous, but since it is related to the particle momentum, it is also finite and single-valued. This, in turn, implies that the wave function itself is continuous.
2.1.3 The Infinite Quantum Well

The classic example of quantum confinement or binding of a particle is shown in the hypothetical infinite potential well, illustrated in Figure 2.2. While all the boundary conditions described in the last subsection do not apply to the infinite barriers, they do apply to the well region. However, it is trivial to solve Equation (2.10) in Regions I and III. Given that $E$ is finite then $\psi(x) = 0$ in these regions. This means the particle is bound completely in region II.

![Potential function representation of the infinite quantum well.](image)

**Figure 2.2:** Potential function representation of the infinite quantum well.

In Region II, we define the potential to be $V = 0$ making the time-independent wave equation

$$\frac{\partial^2 \psi(x)}{\partial x^2} + \frac{2mE}{\hbar^2} \psi(x) = 0.$$  \hspace{1cm} (2.16)
This is a classic 2\textsuperscript{nd} order differential equation whose solution is of the form

\[
\psi(x) = A_1 \cos(kx) + A_2 \sin(kx) \tag{2.17}
\]

where

\[ k = \sqrt{\frac{2mE}{\hbar^2}}. \tag{2.18} \]

The wave function \( \psi(x) \) must be continuous between all regions therefore \( \psi(x = 0) = \psi(x = a) = 0 \). Applying this condition at \( x = 0 \) makes \( \sin(kx) = 0 \) necessitating that \( A_1 = 0 \). At \( x = a \), this makes

\[
\psi(x = a) = 0 = A_2 \sin(kx). \tag{2.19}
\]

This result has infinite solutions, but they are expressed as \( ka = n\pi \), where \( n = 1, 2, 3, \ldots \). The parameter \( n \) is a quantum number, and this solution illustrates the quantization into discrete energy levels. Substituting this result back into equation \( 2.18 \) gives

\[
k_n = \frac{n\pi}{a} = \sqrt{\frac{2mE}{\hbar^2}} \rightarrow k_n^2 = \frac{n^2\pi^2}{a^2} = \frac{2mE}{\hbar^2}. \tag{2.20}
\]

Rearranging gives an expression for the total energy as

\[
E = E_n = \frac{\hbar^2 n^2 \pi^2}{2ma^2}. \tag{2.21}
\]
The previous equation shows a given quantized energy level has a \( \frac{1}{a^2} \) dependence. This dependence in addition to the particle effective mass, show in equation as \( m \), are the foundational methods for engineering the energy level separation. Figure 2.3 illustrates the consequences of varying well width.

![Figure 2.3](image)

**Figure 2.3:** Simulation of a single quantum well with widths of 5.0nm (a) and 2.5nm (b). Shows the increase in energy level separation by variation of the well width.

To finish the solution to the wave equation, the value of the coefficient \( A_2 \), the normalization boundary condition \( \int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1 \) is used. Assuming the wave function solution \( \psi(x) \) is a real function, then \( \psi(x) = \psi^*(x) \). Substituting the wave function into the normalization boundary condition gives

\[
\int_{-\infty}^{\infty} A_2^2 \sin^2(kx) \, dx = 1.
\]  

(2.22)

Evaluating this integral gives

\[
A_2 = \sqrt{\frac{2}{a}} \tag{2.23}
\]
making the time-independent wave solution

\[ \psi(x) = \sqrt{\frac{2}{a}} \sin \left( \frac{n \pi x}{a} \right) = \sqrt{\frac{2}{a}} \sin(k_n a). \]  

Again, this result demonstrates the quantization of the energy of a particle into discrete energy levels. These energy levels are related to quantum states. This result is the primary motivation for engineering devices with nanoscale multiple quantum well structures. Again, of particular note is that the result shows the separation of energy levels is a function of the well width and the particle effective mass. This example serves to illustrate principles that are applicable to the much more complicated case of multiple finite quantum wells.

2.1.4 Electronic States in Multiple Finite Semiconductor Quantum Wells

As mentioned at the beginning of the last sub-section, accurate modeling of the electronic states is necessary for design of working devices. The derivations above serve to illustrate the quantization of energy levels into discrete states using a simplified version of Schrödinger’s wave equation. The simulation software used in this work uses single-band with effective mass, 6-band, 8-band \( \mathbf{k} \cdot \mathbf{p} \) approximations within a finite differences grid. The basic 8-band model is now presented in brief as given in [4]. The Schrödinger equation for a crystal is

\[
\left( \frac{p^2}{2m_0} + V(r) + \frac{\hbar^2}{4m_0^2c^2} (\sigma \times \Delta V) \cdot \mathbf{p} \right) \psi(r) = E \psi(r) \tag{2.25}
\]
The wavefunction $\psi(r)$ written in terms of the Bloch wavefunctions is

$$
\psi_{n,k}(r) = e^{i k \cdot r} u_{n,k}(r)
$$

(2.26)

where $n$ is the band and $k$ is the wavevector. The action of $p$ on $\psi$ is

$$
p \cdot (e^{i k \cdot r} u_{n,k}(r)) = -i \hbar \nabla \cdot (e^{i k \cdot r} u_{n,k}(r))
= \hbar k e^{i k \cdot r} u_{n,k}(r) + e^{i k \cdot r} p \cdot u_{n,k}(r)
= e^{i k \cdot r} (p + \hbar k) u_{n,k}(r)
$$

(2.27)

Using this relation, Schrödinger’s wave equation is obtained for $u_{n,k}(r)$:

$$
\left( \frac{p^2}{2m_0} + \frac{\hbar k \cdot p}{m_0} + \frac{\hbar^2 k^2}{2m_0} + V(r) \right) u_{n,k}(r) = E_{nk} u_{n,k}(r)
$$

(2.28)

and the Hamiltonian $H = H_0 + W(k)$ can be split into $k$-independent

$$
H_0 = \frac{p^2}{2m_0} + V(r)
$$

(2.29)

and $k$-dependent

$$
W(k) = \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} k \cdot p
$$

(2.30)

parts. The fundamental idea of the $k \cdot p$ approximation is to use $u_{n,0}(r)$ as the basis for expansion of the wavefunction and energies at a finite $k$ value.
2.1.5 Multiple Quantum Well Active Region

In this section a description of the key objectives for lasing are discussed as is relevant to this work. Additional discussion of various active region design schemes is given in Chapter 3. This work focuses on diagonal transition designs in which the radiative transitions occur as the carriers tunnel from one quantum well of the active region to the other. Typically, these structures are designed for confinement of a single energy band in each well. Figure 2.4 shows a rendition of a traditional three well structure with the pertinent energy bands highlighted for each well.

![Figure 2.4: Three well laser structure showing the energy levels and the electron probability functions.](image)

Traditional diagonal transition designs strive to have 3 active region energy levels; listed from highest first, these energies are: $E_3$, $E_2$, and $E_1$. The radiative transition occurs between $E_3$ and $E_2$, with energy

\[
E_{ph} = h\nu = E_3 - E_2
\]
In order for transition to occur, the purpose of the injector region is to relax an electron such that under an applied bias the electrons energy is near the \(E_3\) energy level with a high injection efficiency, \(\eta_{\text{eff}}\). When effective, the high probability of electrons existing in the \(E_3\) level assures a sufficient population of electrons in the first active region quantum well. In order for these electrons to fall to the \(E_2\) energy level, a necessary condition for emission, the lower energy level must have unoccupied states. This condition is known as population inversion. Population inversion is assured when the electron relaxation time from states 3 to 2 is greater than the electrons lifetime in state 2. Therefore, we can quantify the relationship needed for population inversion as the ratio of the two time constants:

\[
\rho_{32} = \frac{\tau_2}{\tau_{32}} \tag{2.32}
\]

The subband separation from state 2 to 1 is designed to be on the order of an electron-longitudinal-optical (LO) phonon resonance energy, which for GaAs is 35 meV. This scattering process occurs in less than a pico-second and allows for the quick depopulation of electrons from state 2, assuring population inversion. Equation 2.33 shows this relationship.

\[
E_{21} \sim \text{LO Phonon} \tag{2.33}
\]
From a basic description of the QCL, four objectives have been identified for achieving photon emission and lasing: the three equations above and the level of the electron out of the injector region. The latter objective can be quantified as the energy difference between $E_{inj}$ and $E_3$,

$$\Delta E_a = E_3 - E_{inj} \quad (2.34)$$

As mentioned, one of the main requirements for lasing is population inversion, but population inversion is also the most influential parameter for modal gain as shown in equation 2.35.

$$g_m = \frac{\xi}{L_p} = \tau_3 \left(1 - \frac{\tau_2}{\tau_{32}}\right) \frac{4e\pi z_{32}^2}{\lambda_0 n_{eff} \varepsilon_0 L_p} \frac{1}{2\gamma_{32}} \quad (2.35)$$

Population inversion, then, is proportional to the lifetime of the lasing states (3 and 2) and the scattering rate of electrons between these states as shown in the equation 2.36.

$$\Delta N \propto \tau_3 \left(1 - \frac{\tau_2}{\tau_{32}}\right) \quad (2.36)$$

Equation 2.35 also reveals a few other parameters that affect gain: the overall length of the active region, $L_p$, the previously mentioned injection efficiency, and the dipole matrix element, $z_{32}$. Preference will be given to designs with shorter overall active regions, but this is not a part of the simulation objectives for this work.

Discussion of additional objectives for future genetic algorithm simulations is given in Chapters 6 and 7. Figure 2.5 shows a simulated AlGaAs/GaAs(111) based diagonal transition QCL.
2.2 Mid-Infrared QCLs in literature

The mid-infrared region is from 3 µm to 25 µm, corresponding to photon energies of 0.41-0.083 eV. These smaller photon energies lend themselves well to intersubband transition designs where one can take advantage of the large oscillator strengths possible with these devices, improving efficiency over interband designs. However, there are significant inherent challenges. Perhaps the most significant is the short upper state lifetimes of superlattice structures that limit the potential gain. This challenge however, is met with need for efficient devices operating in these regions because of the poor performance of alternative designs.

Figure 2.5: Simulation of a diagonal transition QCL.
At the time of this writing, perhaps the most comprehensive reviews of quantum cascade lasers is given by [5]. In addition, the technical operation and design of a quantum cascade laser was, until the book by Faist [4], confined to a series of journal articles exploring aspects of the topic. The number of articles with the term 'quantum cascade laser' has increased linearly from the time of the seminal paper [1] in 1994 to over 500 articles in 2014 according to ISI: Web of Science. Progress has been rapid in the development of these sources.

The first QCL [1] operated with limited power output and at cryogenic temperatures. But just 3 years later in 1997, the first single frequency QCL operating at room temperature was demonstrated [34] using distributed feedback mechanisms. Another 5 years in 2002, brought the first continuous wave operation of a mid-IR QCL [35]. This result was significant, in fact, as no other semiconductor laser source operated at room temperature in the mid-IR range. This result brought the start of commercial applications for mid-IR QCLs, which have been mostly confined to gas sensing and infrared counter measures. However, with continued improvement in the efficiency of the sources, these applications could include communications networks. It is worth noting that since 2002 [36], quantum cascade lasers have also been used in terahertz wavelength applications due to the broad tunability of the intersubband design. An additional type of cascade laser, known as the interband cascade laser (ICL), was developed in 1995 [37]. ICLs have radiative transitions between the conduction and valence bands and have been operated at or close to room temperature in the 3-4 µm range [38].
QCLs may be realized using a myriad of semiconductor material systems. Most of the best performing devices have been confined to four different material systems: GaInAs/AlInAs grown on InP substrates [39], GaAs/AlGaAs grown on GaAs substrates [40], AlSb/InAs grown on InAs substrates [41], [42], and InGaAs/AlInAsSb, InGaAs/GaAsSb or InGaAs/AlInGaAs grown on InP substrates [43]–[45]. Material selection is important for realizing the most efficient devices and is influenced most by the effective mass $m^*$ of the wells and barriers [46] and the conduction band offset $\Delta E_c$. The GaAs material system is perhaps the most technologically mature, but suffers from low achievable conduction band offset $\Delta E_c$, limiting the output range to above 8-9 µm.

Over the last 20 years of QCL development, operation in the Mid-IR spectrum has reached impressive performance levels. Multi-watt output power with continuous wave operation with efficiencies of up to 21% at room temperature over 50% at cryogenic temperatures have been achieved [47]–[49]. Spectral coverage of QCLs have covered the entire Mid-IR and much of the THz frequency ranges [50]. The most challenging range is perhaps the first atmospheric window (3-5 µm) as a large conduction band offset $\Delta E_c$ is required, which limits the material system and contributes to efficiency losses due to lack of confinement of the upper excited state where radiative transitions occur. While there are recent demonstrations of room temperature and higher operation of QCLs in this range [41], [45], [50]–[52], much exploration is still needed. The primary motivation of this work is to extend the use of one of the more mature material systems, AlGaAs/GaAs, into this range in effort to explore possibilities for improved efficiency and novel operation using strained layer active regions [16], [19], [20].
CHAPTER 3

Active Region Design

In this chapter, a deeper understanding of the elements necessary for working designs of a QCL are developed. These concepts are contextualized for use in automating the design of a QCL using a genetic algorithm in conjunction with a nanoscale simulation tool, nextnano\textsuperscript{3}. With that purpose in mind, initial discussion will be on the traditional elements used for optimization of the QCL’s active region (AR). Some historical perspective and variations of design are discussed afterwards. The chapter concludes with some discussion on more novel AR designs including, strained layer incorporation and second-harmonic generation.

The previous chapter introduced the concepts of quantum mechanics with some extended discussion on solutions to Schrödinger’s wave equation. In the early 1970’s, Kazarinov and Suris [9], [53] presented the possibility of amplification of light by use of a semiconductor superlattice structure. In the latter paper, they examined the optical response of a semiconductor superlattice under the influence of a strong electromagnetic field. The superlattice wells were designed such that the ground state of one well was situated just below the second excited state of the neighboring well when under the influence of an electric field, which tilts the band structure so that this well is downstream - with respect to the flow of electrons. This investigation made clear the mechanism for population inversion, as is represented in Figure 3.1. The strength of this interaction between the ground state and the second excited state,
which is similar to injection efficiency in a QCL, is arbitrarily increased or reduced by adjusting the thickness of the tunnel barrier. Population inversion in this structure is a natural consequence of a particle’s desire to be at the ground state. Since in this design the ground state of one well is also the upper-state of the downstream well population inversion is a robust feature independent of temperature.

![Figure 3.1: Band structure of the theoretical device proposed in [53]. Superlattice structure under applied electric field with alignment of the ground state of the upstream well just below the second excited state of the downstream well.](image)

This structure and its theoretical elements have yet to be demonstrated experimentally. The electric field required to align the ground state and the upper state of adjacent levels puts it into an unstable region of negative differential resistance, as demonstrated in [54]. A more realistic structure would have the ground state aligned with the first excited state in the downstream well. This structure is the fundamental concept employed in design of an injector region. To extend this concept further toward the traditional design of the QCL, a second stage of the structure would then align the ground state to the second excited state of an adjacent well, as is the case for the last well of a QCL’s injector region.
This fundamental theoretical work by Kasarinov and Suris and the experimental demonstrations by K.K. Choi, who acknowledged Capasso, led Faist and Capasso toward the first successful demonstration of a QCL in 1994 [1].

3.1 Extended Fundamentals

According to Faist, the inventor of the QCL, as stated in [4] the design of a successful intersubband laser has to overcome a number of problems. It needs to operate in a stable electrical range. It needs to achieve population inversion. Finally, the gain should overcome waveguide and other active region losses. A quick note: this work does not discuss waveguide design, nor does it explicitly include concessions for active region losses (rather they are inherently considered). The first intersubband laser design to simultaneously achieve all these requirements was the QCL [1].

As has already been mentioned, the QCL design is far from a simple periodic superlattice structure. Rather, it is a periodic arrangement of cells, with an active (gain) region and an injector region, each consisting of a complex heterostructure potential with a number of well and barriers. Figure 3.2 show the generalized conduction band diagram of the quantum cascade laser as first proposed.

The gain region is a structure designed to create at least 3 electron energy states, such that electrons are injected into the \( n = 3 \) state and population inversion is maintained between states \( n = 3 \) and \( n = 2 \). Ideally, electrons would be injected exclusively into the \( n = 3 \) state, which mathematically is \( \eta_3 = 1 \). Population inversion is related to the following in terms of electron state lifetimes: \( \tau_{32} > \tau_2 \). Figure 3.3 indicates a simplified picture of the pertinent lifetimes and radiative processes in the active region.
Figure 3.2: Schematic conduction band diagram of the first proposed quantum cascade laser.

The gain region is followed by an injection/relaxation region. The purpose of this region is to provide electrons, via resonant tunneling, into the next active region period. To achieve resonant tunneling, the electron energy must be raised to align with the upper excited state of the active region. This is achieved by a sequence of alternating quantum wells and barriers with a changing duty cycle. Additionally, often times a portion of the injector is doped to provide an "electron reservoir" for the next period and to prevent the formation of a space-charge region. In contrast to the intersubband design proposed by Kazarinov, this graded injector region does not require a large electric field and promotes a cooling of the electron distribution towards the lattice temperature, hence the alternate term relaxation region.
3.2 Rate Equation

The requirements for design of a QCL may be achieved by an atomic-level analysis of the active region. For simplicity, the injector region will consist of only one level with a constant population $n_g$, aligned with the upper level of the next period, as was shown in Figure 3.3. Electrons in the ground state of the injector tunnel to the upper state ($n = 3$) of the active region at a rate equal to $J/e$. Electrons can scatter from this state to the $n = 1$ and $n = 2$ states. The scattering times for these transitions are defined as $\tau_{32}$ and $\tau_{31}$. The total lifetime of an electron in the $n = 3$ state is given by the sum of scattering rates: $\tau_{3}^{-1} = \tau_{32}^{-1} + \tau_{31}^{-1} + \tau_{esc}^{-1}$, where $\tau_{esc}^{-1}$ is the escape rate into the continuum. At large temperatures, $\tau_{esc}$ as well as thermally activated electrons entering state $n = 2$ are an important factor. The thermal equilibrium concentration of carriers (electrons) in state $n = 2$ is expressed by the equation $n_{2}^{therm} = n_{g}exp(-\Delta/kT)$, where $\Delta$ is the energy difference between
the Fermi level of the injector and the $n = 2$ state of the laser transition. With that background, the rate of change of carrier population of the three active region levels may be expressed as

$$\frac{dn_3}{dt} = \frac{J}{e} - \frac{n_3}{\tau_3} - S g_c (n_3 - n_2) \tag{3.1}$$

$$\frac{dn_2}{dt} = \frac{n_3}{\tau_{32}} + S g_c (n_3 - n_2) - \frac{n_2 - n_{2\text{therm}}}{\tau_2} \tag{3.2}$$

$$\frac{dS}{dt} = \frac{c}{n_{\text{refr}}} \left[ (g_c (n_3 - n_2) - \alpha_{\text{tot}}) S + \beta \frac{n_3}{\tau_{sp}} \right] \tag{3.3}$$

where $\alpha_{\text{tot}}$ is the sum of the waveguide and mirror losses (total modal loss), $S$ is the photon density, defined as the photon flux per period and unit active region width, $\beta$ is the fraction of the spontaneous emission coupled in the laser mode, and $\tau_{sp}$ is the spontaneous emission lifetime.

While the rate equations give a general understanding of the mechanisms and rates of electrons in a given active region, these equations are defines for only one period. Consequently when considering a cascaded laser with many periods one usually defines the gain cross-section $g_c$ as one of all the periods in the active region put together with a total overlap factor $\Gamma$:

$$g_c = \frac{4\pi e^2}{\varepsilon_0 n_{\text{refr}} \lambda 2\gamma_{32} L_p} \frac{z_{32}^2}{\Gamma} \tag{3.4}$$

where $z_{32}$ is the dipole matrix element, $2\gamma_{32}$ the FWHM broadening of the transition, $\lambda$ the wavelength, and $L_p$ the length of one period. As mentioned in the previous chapter, these factors are important when considering the design of the device.
3.3 The "Intersubband Toolbox"

Faist [4] identifies four conditions for obtaining a low-threshold, high-efficiency laser:

- large ratio of upper-state to lower-state lifetimes $\tau_{32}/\tau_2$.
- low waveguide loss $\alpha_w$.
- narrow transition linewidth $\gamma_{32}$
- long upper-state lifetime $\tau_3$

These conditions contain complicated trade-offs and have resulted in many different QCL designs. It is also the motivation for developing automated methods, for example the genetic algorithm developed in this work, for achieving optimal designs. Despite the availability of such automated methods, researchers have developed what Faist refers to as the "intersubband toolbox", which will now be described in brief. A full description of these methods is found in [4].

3.3.1 Tunneling

The first fundamental consideration in designing a QCL is the injection of electrons from the lower state of the last injector well to the upper state of the first active region well. In general, the coupling of these energy states is achieved by bringing the two levels together near resonance, i.e. aligning the levels. Because the final injector state 1 is also the ground (or initial) state of the active region (injector), one must
consider the coupling energy $\Omega_{1'1}$ between the initial state $1'$ and the final state $1$, illustrated in Figure 3.4. Mathematically these states are localized in their respective wells when the coupling energy is much smaller than the energy difference between the states: $\Omega_{1'1} << E_{1'} - E_1$.

![Figure 3.4](image)

**Figure 3.4:** Two states coupled by a tunnel barrier. The lifetime of the upper state is controlled by the thickness of the coupling barrier.

Scattering between these states is related to the coupling energy $\Omega_{1'1}$ which is proportional to the tunneling probability:

$$\exp(-\kappa_b L_b) \text{ where } \kappa_b = \sqrt{2m^*(V_b - E)}/\hbar.$$  \hspace{1cm} (3.5)

$L_b$ is the thickness of the tunnel barrier and scattering times will be lowered by simply increasing the tunnel barrier thickness, regardless of the scattering mechanism. However, the coupling energy $\Omega_{1'2}$ is proportional to the intensity transmission of the tunnel barrier, i.e. to $\exp(-2\kappa_b L_b)$, which also contains decreases as the tunnel barrier increases. However, careful examination shows that while the scattering rates and oscillator strength decrease with $L_b$ the initial rate of decrease of the scattering
rate is faster than that of the oscillator strength \( f \). The oscillator strength is

\[
f = \frac{2m_0}{\hbar^2} z^2_{ij} E_{ij}
\]  

(3.6)

where \( E_{ij} \) is the energy difference between the i and j states. One important note is that necessity to have a non-zero value for the energy difference. In an optimal design we want to bring the injector’s final state into resonance with the upper state of the active region, but they must not align perfectly. This is an important trade-off as lowering of this state increases the coupling energy to the lower states of the active region. A merit factor \( f \tau \), the product of the oscillator strength and the lifetime, is given to optimize the design of the tunneling barrier. This merit factor increases slightly with barrier thickness and is the justification for investigation of quantum cascade lasers based on photon-assisted tunneling transitions [55].

### 3.3.2 Number of Periods

In contrast to the single quantum well, superlattice active regions exhibit an advantage in oscillator strength because transitions occur between excited states. Because of this one must consider the consequences of increasing the number of periods in the QCL cell. The same merit factor, \( f \tau \), increases as the number of periods \( N \) increases. Every well/barrier pair introduces new states decreasing the spacing of adjacent states. Additional states contribute to the broadening of the transition and reduce gain. Also, upper states separated by less than \( kT \) will decrease the injection.
efficiency as the population of the electrons in the injector will be spread over many states. In practice, these factors limit the number of periods to around 8-10. For the purpose of this work, the number of periods is 8 (3 for active region and 5 for the injector).

3.3.3 Upper-level confinement

Under an applied electric field, the states of a quantum cascade laser are resonances with a limited lifetime caused by a finite tunneling probability. In other words, these states are occupied because of electrons having a probability of tunnelling through the barriers due to the autoionization energy. This energy also translates to an escape possibility from the upper-state to the continuum. This necessitates a large barrier potential, so that the states are bound. However, a large barrier potential reduces the tunneling probability and generates reflections. The transmission coefficient of a single electron with energy $E$ incident upon a finite barrier with energy $V_0$ for the case when $E << V_0$ is approximated by

$$ t \approx 16 \left( \frac{E}{V_0} \right) \left( 1 - \frac{E}{V_0} \right) \exp(-2\kappa_2 a). \quad (3.7) $$

One method for limiting the escape possibility is to design the injector region such that the upper-state of the laser transition faces a minigap in the energy levels, promoting the scattering to the lower excited state of the active region as opposed to the continuum. This again necessitates the reduction of periods in the injector region, so as not to introduce multiple states. This is shown in Figure 3.5.
Figure 3.5: Electron tunneling into the continuum (left): $\tau_{esc} \approx e^{-2\kappa_b L_b}$ and $\tau_2 < \tau_1$. Electrons are not allowed to tunnel into a minigap (right): $\tau_2 > \tau_1$.

Escape time can be computed by invoking Heisenburg’s relation. The escape time is given as

$$\tau_{esc} = \frac{\hbar}{E_3 t}$$  \hspace{1cm} (3.8)

where $t$ is the transmission coefficient. This relationship shows that a reduction of the transmission coefficient has a positive effect on the escape time. In fact, Figure 3.6 shows that optimal placement of the upper-state energy is near the middle of the minigap, which will provide maximum reflection.

### 3.4 Traditional QCL Design and Variations

The development of the first QCL [1] was preceded by study of single, double, and superlattice designs for optical transition. Traditionally the QCL has a three-quantum-well active region, although other design possibilities do exist. The three-quantum-well active region can be decomposed as shown in Figure 3.7. This design starts with a pair of quantum wells (a) a thinner well and a thicker well. When these wells are brought together (b) the can be resonantly aligned by adjusting the electric field and
the barrier thickness. Generally speaking the wells need to be of differing thickness so that the excited states are not in resonance. A third quantum well is then added upstream (c) and for diagonal transition designs the lower state of this well is brought in resonance with the upper state of the second well.

One variation of the three-quantum-well active region is the active region which introduces another ground state by addition of a fourth downstream well. This ground state is situated such that the energy gap between the preceding lower state is that of the LO-phonon energy. This type of design [35] utilizes double phonon resonance and can aid in the depopulation of the lower excited state and increase population inversion.
Figure 3.7: Schematic conduction band profile and design of the three-well active region. a) Two-quantum separate quantum wells. b) Coupled wells with applied electric field. Resonance of the lower wells causes energy level splitting. c) Third well added upstream with the energy level aligned with the upper level of the second well. Produces diagonal transitions. d) energy level is above the upper level of the second well producing vertical transitions. Reproduced from [4]

3.4.1 Active Region Design Bounds

For this subsection, consider only the thickness of each layer in the active region of a QCL including the last well/barrier from the injector to the active region. From a computational standpoint the design possibilities are quite large. With four potential wells and five barriers this simple design has a total of seven (8) layers. Specification of a range of thicknesses for each layer is necessary. Luckily, at the atomic level the thicknesses are discrete and defined in terms of monolayers (ML). A reasonable restriction of the range of thicknesses for each layer is 2ML – 100ML. This gives a
result of 99 potential thicknesses for each layer. Given a 8 layer design like this one this means there is 99 to the power of 8 total different combinations. This is \(9.2 \times 10^{15}\) design possibilities an enormous number. Certainly, there are several more constraints that can be defined to restrict the solution space further, but in reality there are several other variables that can be changed in a design such as the material composition of the layers and the surface. Furthermore, the actual design of a quantum laser is dependent upon the injector design.

Given the enormity of the design challenge it is actually quite astonishing that scientist have achieved designs that work reasonably well. In designing a quantum laser structure several important objectives must be met and balanced in order to have lasing. Most laser designs utilize only 2 different materials which determine the barrier height (well depth). The separation and positioning of energy levels are dependent on the well widths and barrier thicknesses. It does not take long when changing just a few of the parameters to discover that the energy levels and probability functions change dramatically. Additionally a change, for example, of one well width affects the energy levels and probabilities in another well. This means that the variables are dependent upon each other. It is this dependence that complicates the design and makes it an excellent candidate for multi-objective evolutionary computation for searching potential solutions.

In this work, a multi-objective genetic algorithm was developed to optimize the design of a quantum cascade laser designed to operate in the 3-5µm atmospheric window of the mid-infrared region. The genetic algorithm is discussed in detail in chapter 5. As is discussed in the following sections and the next chapter, one motivation for automating the design is to provide context for developing a novel QCL
with strain intentionally incorporated into the active region. This type of design has shown the possibility for reduction in threshold voltage [16]. Additionally, a second type of design is possible that utilizes the enhanced non-linear susceptibility in the (111) crystal orientation for second harmonic generation.

3.5 Strain Layer incorporation in the Active Region

One of the unique aspects of the work completed in this thesis is the incorporation of strained layers in the active region. This approach is relatively new and has been a part of the focus of our research group for some time [16], [19], [20]. In this way, the concept of strained layer active regions is an extension of the work described in Denzil Roberts’ Ph.D. thesis [56]. For extended discussion on the physics of strained material, the reader if referred to his work. In this work, a reference QCL design [57] is simulated, using nextnano3, and compared to the same design simulated on the (111) substrate. He found that the $E_3$ level is situated further into the well, a result of a Stark-shift influenced by the piezoelectric properties of this orientation, and that the $E_{32}$ energy increased, lowering the emission wavelength. To further extend the work, strained In$_x$Ga$_{1-x}$As layers were used in the active region at values of $x$ between 0-6%. Figure 3.8 shows a simulated strained layer QCL with 6% indium in the active region.

This work showed the promise of improved active region performance by inclusion of strained layers on the mis-oriented substrate. This thesis intends, in part, to explore this concept further but without the use of a reference design, rather by generation of a design through genetic algorithm search.
One additional concept proposed by our group [20], [56] is second harmonic generation by means of improved non-linear susceptibility and the presence of a fourth excited state in the upper bounds of the first active region well. Close examination of Figure 3.8 shows this fourth state. Since the energy levels are lowered for a given well width on GaAs(111) vs GaAs(100) and with the deepening of the well by indium incorporation, a fourth state becomes bound and introduces a coupling energy from the injector this this new level. Figure 3.9 shows this coupling.
Electron populations in both the injector’s excited state as well as electrons that exist in the $E_3$ level of the active region could jump up to this $E_4$ level if given energy. If $E_4 - E_3 = E_3 - E_2 = E_{\text{photon}}$ the electron will have a probability of absorbing an incident photon. In doing so, the energetic electron may be lost into the continuum or emit a photon through transition form $E_4$ to $E_2$. Figure 3.10 illustrates this concept of frequency mixing. The design given in [56] and shown in Figure 3.8 was grown, by MBE, in our lab and was recently tested for presence of second harmonic generation [20]. Raman spectroscopy results showed a second peak at the expected energy for second harmonic generation from the $E_4$ to $E_2$ transition.

Each of the conceptual ideas made possible by incorporation of strained layers into the active region of a QCL represent new avenues for exploration in the quest to improve the efficiency of devices that operate in the first atmospheric window. However, investigations into these types of devices are relatively new and optimal designs have not been engineered. This thesis explores traditional three quantum well designs on both (100) and (111) surfaces with indium incorporation into the active region. The objectives laid out for the genetic algorithm do not specifically address
Figure 3.10: Frequency mixing by introduction of a fourth state into the active region.

the concept of second harmonic generation, however some discussion is given in the simulation results in Chapter 6. The genetic algorithm described later in this work is, on the other hand, designed to produce designs with direct $E_3$ to $E_2$ transitions in the 3-5 µm range.
CHAPTER 4

Growth Study using GaAs (100) and (111)B

4.1 Molecular Beam Epitaxy

Superlattice structures like those theoretically analyzed by Kazarinov [9] and the more complicated superlattice based structures like the QCL [1] are not possible without a highly controlled growth processes. The advent of molecular beam epitaxy (MBE) [10], [58], [59] enabled single monolayer control of growth with high quality surface morphologies. The near atomic level flatness of each layer allowed for development of structures with abrupt interface transitions. MBE is preferred deposition technique for superlattice structures thanks to it’s low impurity levels, growth control, and reproducibility of composition and doping [10], [58]–[61].

At the time of these growth experiments, the University of Missouri, under the direction of Dr. Gregory Triplett maintained a EPI Systems 930 Molecular Beam Epitaxy System. Figures 4.1, 4.2, and 4.3 show pictures of the MBE and its control equipment.
Figure 4.1: Substrate side of EPI 930 MBE system during operation.

Figure 4.2: Source shroud side of the EPI 930 MBE system during operation.
Figure 4.3: Control racks containing 2408 temperature controllers and 2kW Sorensen dc power supplies for each material cell of the MBE. Other equipment includes a RHEED gun controller, vacuum controllers, and vacuum meters.
Figure 4.4: Representative drawing of a typical MBE system with control and processing equipment. Copied from [4].

Figure 4.4 shows an drawing of the working internals of a typical MBE. The EPI 930 MBE at MU is outfitted with a 15 keV electron gun and RHEED screen with a 12-bit digital camera. This tool is used for in-situ monitoring of the growth process. The electron gun is positioned at a low (<1°) grazing angle and causes a diffraction pattern of scattered electrons on the phosphor-coated screen. The pattern is used to determine the surface morphology and the quality and shape of the pattern clues in on the crystallinity and phase of layer growth.
4.1.1 RHEED Characterization

RHEED characterization is also used to determine, along with pyrometry or thermocouple, the substrate surface temperature. Bare substrates will contain a thin oxide layer that must be removed before growth. To accomplish this, the substrate is slowly heated under a small arsenic overpressure. As the temperature increases the oxide layer is desorbed. For a given substrate material, such as GaAs at 580 °C, this happens at a precise temperature. This allows the user to calibrate the system using optical pyrometry or a thermocouple (as is the case for MU’s system) so that the substrate temperature $T_s$ is precisely known. During desorption of the oxide layer the RHEED pattern transitions from an diffuse glow (amorphous surface) to a horizontal streaks, indicative of a crystalline surface emerging.

The diffraction of electrons on a crystalline surface produces a pattern that is material dependent (crystal structure). Using a digital camera the intensity of the specular spots of the pattern can be detected. As layers form on the surface the adatoms (Ga and As) nucleate on the surface. This causes additional scattering of the electron beam and dims the specular spots. As the surface become more and more complete, the intensity of the specular spot will recover. The oscillation of intensity is used to determine the growth rate of layers on the surface. Figure 4.5 taken from [62] illustrates this mechanism.
Figure 4.5: Illustration of RHEED oscillations during monolayer growth. [62].

4.2 Experimental Setup

The results discussed below are from work completed in the last few years. Many of the results of each of these studies are now published [19]–[21]. Collectively the work from our lab has significantly enhanced the understanding of the effects of process conditions on the quality of crystal growth on GaAs (111)B substrates. This work affirms the possibility of using strained layers in the design of optical devices and is the motivation for the effort to design and optimize QCLs via simulation and genetic algorithm based search.
4.3 Strained and unstrained layers on GaAs (111)B

To further understand the conditions necessary for pseudomorphic growth of strained and unstrained layers on GaAs(111)B, this work expands on previous studies by examining the sensitivity of process conditions and strain on the surface quality for GaAs, AlGaAs, and InAs layers grown on (111)B $2^\circ \rightarrow [2\overline{1}\overline{1}]$ Si-doped substrates using nano-scale microscopy techniques, scanning electron microscopy (SEM) and atomic force microscopy (AFM). These techniques allow for a better determination of the optimal process conditions for multi-layer photonic devices. This data corresponds with the previously mentioned narrow growth window for these structures and better reveals the type and severity of defects that occur in the non-ideal growth zones.

4.3.1 Experimental Procedures

Table 4.1 outlines the steps taken for sample preparation and growth in the MBE reactor. For purposes of direct comparison, both GaAs (100) and (111)B $2^\circ \rightarrow [2\overline{1}\overline{1}]$ wafer pieces were indium mounted onto a molybdenum sample block, where (in most cases) the (100) oriented substrate was placed in the block center for in-situ RHEED analysis.

Three sets of experiments were performed to determine the optimal substrate temperature (Ts) and V-III ratio parameter ranges for layers of GaAs, Al$_{0.45}$Ga$_{0.55}$As, and InAs on GaAs (111)B. Table 4.2 outlines the process conditions used for the experiments. Since the InAs and AlGaAs layers rely on optimized conditions for GaAs, the former preceded others.
Table 4.1: Growth Procedures

<table>
<thead>
<tr>
<th>Step</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mount (111) and (100) sample on moly block</td>
</tr>
<tr>
<td>2</td>
<td>Pre-stage sample in load lock</td>
</tr>
<tr>
<td>3</td>
<td>Pump down load lock and equalize pressures</td>
</tr>
<tr>
<td>4</td>
<td>Load Sample in Growth Chamber</td>
</tr>
<tr>
<td>5</td>
<td>Pump Growth Chamber</td>
</tr>
<tr>
<td>6</td>
<td>Set As flux to 10⁻⁷ Torr</td>
</tr>
<tr>
<td>7</td>
<td>Initially raise $T_s$ at 20 °C/min</td>
</tr>
<tr>
<td>8</td>
<td>at ~50 °C below deox lower ramp rate to 5 °C/min</td>
</tr>
<tr>
<td>9</td>
<td>Watch for 2x4 RHEED reconstruction</td>
</tr>
<tr>
<td>10</td>
<td>Reach deox point, raise $T_s$ to 610 °C</td>
</tr>
<tr>
<td>11</td>
<td>Hold for 5 mins (assures complete deox)</td>
</tr>
<tr>
<td>12</td>
<td>Set $T_s$ at desired point for growth</td>
</tr>
<tr>
<td>13</td>
<td>Deposit whetting layer of GaAs</td>
</tr>
<tr>
<td>14</td>
<td>Determine RHEED surface reconstruction</td>
</tr>
<tr>
<td>15</td>
<td>if good proceed with growth</td>
</tr>
</tbody>
</table>

Table 4.2: Growth Experiments Process Conditions

<table>
<thead>
<tr>
<th>Material</th>
<th>Thicknesses (nm)</th>
<th>BEP Ratio</th>
<th>$T_s$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs</td>
<td>50</td>
<td>10:1-25:1</td>
<td>550-620</td>
</tr>
<tr>
<td>Al₀.₄₅Ga₀.₅₅As</td>
<td>10</td>
<td>20:1</td>
<td>550-680</td>
</tr>
<tr>
<td>InAs</td>
<td>2.2-5</td>
<td>10:1-50:1</td>
<td>480</td>
</tr>
</tbody>
</table>

For GaAs experiments, a single 50 nm layer of GaAs was deposited where $T_s$ was varied within the 550-620 °C range, and the BEP ratio was varied from 10:1-25:1. The substrate was then removed from the growth chamber and AFM and SEM images were obtained to determine surface quality. For subsequent AlGaAs and InAs experimental sets, the range of process conditions which provided the best results for GaAs were employed for the GaAs layers that lay within sample structures in this case the optimal $T_s$ was 580 °C and the BEP ratio was 15:1. In this way, the quality of the AlGaAs and InAs layers is determined via isolation.
For the AlGaAs experimental set, a tri-layer HEMT like structure was studied. A 100 nm thick GaAs layer was followed by a 35 MLs (∼10 nm) of AlGaAs, which was capped by an 18 ML (∼5 nm) GaAs passivation layer. $T_s$ was varied for the AlGaAs layer from 550-700 °C with a BEP ratio of 20:1. For the InAs experimental set, an initial 100 nm thick GaAs layer was deposited at the same optimal conditions followed by 4 MLs or 17 MLs (∼5nm) of InAs. $T_s$ was held constant at 480 °C, and the BEP ratio was varied from 10:1-50:1.

SEM images were obtained using a FEI Quanta 600F under high vacuum with an acceleration voltage of 5 kV. AFM images were obtained using an Agilent Technologies 5500 AFM operated using AC tapping mode and a standard cantilever type tip with a diameter of <10nm.

4.4 Results

4.4.1 GaAs Results

In this section, the focus is primarily on results which show the sensitivity of the growth conditions for GaAs (111)B oriented substrates. Previous work has shown that at optimal growth conditions for a (111) oriented sample, a (100) oriented sample would also have a high quality surface, refer to Figure 4.6. With this in mind, the design of experiment centered around process conditions previously established to produce high quality mirror-like surfaces on GaAs (100) [$T_s = 570-580 °C$ and a BEP ratio of 15-25:1]. For GaAs (100) samples produced within the experimental range, all of the samples were found to be mirror-like with no distinguishable features revealed under an optical microscope at 80X magnification, with the exception of samples
grown under the BEP ratio 10:1, which were arsenic starved and resulted in Ga dots on the surface as confirmed with Energy-dispersive X-ray spectroscopy (EDX). SEM and AFM revealed that the best sample was obtained at $T_s = 570 \, ^\circ\text{C}$ and a BEP ratio of 15:1, where the surface is presumably atomically flat. Consistent with the myriad of literature on the subject, we found that a wide range of conditions can be used to produce high quality layers on GaAs (100).

In comparison, GaAs (111)B samples showed a great variety in surface quality over the experimental range, as shown in Figure 4.7. In fact, of the samples studied in this experimental set only one, $T_s = 570 \, ^\circ\text{C}$ and BEP ratio 15:1, could be considered near atomically flat with only a few minor surface defects per 5 $\mu\text{m}^2$. While the two samples [$T_s = 570 \, ^\circ\text{C}$ with BEP ratio 10:1 and $T_s = 590 \, ^\circ\text{C}$ with a BEP ratio 25:1] have the poorest surface quality. They reveal the importance of adatom migration length on the resulting surface quality for the (111)B orientation, as the samples are very

Figure 4.6: 50 nm layer of GaAs on (100) surface at $T_s = 580 \, ^\circ\text{C}$. SEM image nothing remarkable and confirms surface is flat.
similar in the type of defect observed. It can be assumed, in this case, that the adatom migration length is similar for both sample conditions, but the limiting mechanism is different: the lower BEP ratio compensates for the reduced substrate temperature and the higher BEP ratio compensates for the increased temperature. This concept is also revealed, although not so obviously, for the two samples \( T_s = 590 \, ^\circ\text{C} \) with a BEP ratio of 10:1 and \( T_s = 590 \, ^\circ\text{C} \) with a BEP ratio of 15:1 by the similarities in their triangular pit defects. Although not shown in this experimental set, it is conceivable that a high quality surface is achievable at other combinations of \( T_s \) and BEP ratio as long as the optimum adatom migration length is satisfied. As adatom migration length is increased from the optimum point, either through a decrease in BEP ratio or an increase in \( T_s \), the surface quality suffers through the formation of triangular pit defects which increase in density and depth as the migration length is increased.

4.4.2 AlGaAs Results

Results shown in Figure 4.8 for samples containing 35 ML of AlGaAs follow the same trends observed for the GaAs samples: a narrow range for an optimal surface at 600 \(^\circ\text{C}\) with dense and deep pit defects occurring for larger \( T_s \) and undulating saw-like defects occurring for lower \( T_s \). The increase in optimal \( T_s \) from 570 \(^\circ\text{C}\) to 600 \(^\circ\text{C}\) is attributed to an increase in BEP (20:1) ratio for the AlGaAs layer. These trends indicate once again that adatom migration length is the primary factor in determining the optimal growth conditions. Careful examination of the sample at \( T_s = 600 \) reveal a striking similarity to the GaAs layer, which lies below [assumed from the GaAs trends shown in Figure 4.7]. Since the AlGaAs layer will assume the surface quality
Figure 4.7: SEM images taken at 30,000X magnification* (\(\sim\)10 \(\mu\)m\(^2\)) of GaAs (111)B samples with varying \(T_s\) and BEP ratio. All samples had some form of surface defect, but the sample grown at \(T_s = 570 \, ^\circ\text{C}\) and BEP ratio 15:1 would be acceptable for high-quality electronic and photonic devices. *The sample with \(T_s = 620 \, ^\circ\text{C}\) is shown at 60,000X magnification in order to reveal the surface structure more clearly.
of the layer below, it raises the question: are higher quality layers of GaAs achievable with an increase BEP ratio to 20:1. The surface quality changes dramatically with a change in $T_s$ of just 10 °C below and above this optimal point indicating that control of the temperature is critical for high quality growth on (111) orientations.

4.4.3 InAs Results

Up to this point, only a few published attempts have shown binary InAs on GaAs (111)B [17], [18]. It is believed that pseudomorphic growth is possible, but it remains unclear what conditions are necessary to do so. Using the experimental conditions outlined above, none of the samples showed full coverage of pseudomorphic layers, but the growth was also not 3D, see Figure 4.9. Instead the growth follows a nucle-
Figure 4.9: 5 µm² AFM images for 17 ML InAs samples grown with $T_s = 480$ °C and BEP ratio from 10-50:1. Note how plateaus of InAs are formed at the defect points on the underlying GaAs layer. As BEP ratio increases plateau size decreases while density increases. This is due to the suppression of adatom migration by increasing As⁴ overpressure.

ated Frank-van der Merwe (FM) type mode where the adatoms find an opportune nucleation site and proceed to grow in layer by layer fashion from that nucleation point. EDX was used to confirm the composition of materials. The surface around the islands does not contain In and does not suggest the formation of a wetting layer.
Figure 4.10: 5 μm² SEM images for InAs samples grown with a BEP ratio of 30:1 for two different thickness, 4 ML and 17 ML. Note the increase in plateau size as the thickness increases.

Figure 4.10 details this trend, showing an increase in island/plateau size as the growth proceeds from 4 MLs to 17 MLs. While this result is not the desired pseudomorphic layer growth, it does indicate the possibility that under the right conditions InAs deposition on GaAs (111)B is achievable. Consider the results from Wen [17], where they showed 2D growth of InAs on small stepped layers of GaAs. The atomic level stepped layers provided a myriad of opportune nucleation sites for InAs, which allowed for quick realization of pseudomorphic layers.

Further analysis was conducted on samples with InAs on GaAs substrates of both the (100) and (111)B orientations [20]. Regression models were utilized to determine the most important parameters influencing the smooth growth of InAs on GaAs (111)B surfaces. These models depicted in Figure 4.11 indicate that substrate temperature, beam equivalent pressure, and deposition rate are all significant factors in smooth film growth. The highest t-ratios exist for substrate temperature and deposition rate, which both significantly influence adatom diffusion length.
Figure 4.11: (a) Roughness and (b) Root-mean-square surface model fit for MBE-grown InAs/GaAs samples on (111)B \(2^\circ\rightarrow [2\bar{1}\bar{1}]\) surfaces. Adapted from [20].

It should be noted at this point that although pseudomorphic growth of binary InAs layers may be possible on GaAs (111)B \(2^\circ\rightarrow [2\bar{1}\bar{1}]\) surfaces the conduction band diagram of such a device if used only in the active region of a QCL with AlGaAs/GaAs injector regions would be impractical, see Figure 4.12. The purpose in confining the strained layers to the active region would be to mitigate the compounding of strain effects through a multilayer cascaded device. The large well conduction band offset \(\Delta E_w = 1\text{eV}\) make use of this form of device impractical for QCL as electrons would not be able to tunnel out of the InAs wells if the \(n = 1\) state were low in the well. Therefore, a more realistic conduction band profile would be obtained by a lower indium concentration. This has the effect of reducing the overall strain in the active region. This would likely relate to better quality pseudomorphic growth of the InGaAs layers as was demonstrated in [56]
Figure 4.12: Simulation of a 3 quantum well structure with wells at 2.8 nm and barriers at 2.2 nm. The outer wells are GaAs while the middle well is binary InAs. The barriers are Al$_{0.45}$Ga$_{0.55}$As.
CHAPTER 5

Genetic Algorithms and Use for Design

List of Symbols

\( \mu \)  Mean - as used in a distribution
\( \psi^2 \)  Square of the wavefunction
\( \sigma \)  Standard deviation
\( \tau_2 \)  Lifetime of electrons in the \( E_2 \) level
\( \tau_3 \)  Lifetime of electrons in the \( E_3 \) level
\( \tau_{32} \)  Transition lifetime of electrons from the \( E_3 \) to \( E_2 \) energy levels
\( n \)  Quantum number
\( E_1 \)  Energy level of the third active region well
\( E_2 \)  Energy level of the second active region well
\( E_3 \)  Energy level of the first active region well
\( E_{inj} \)  Energy level of the last injector well

5.1 Multi-objective Genetic Algorithms

Evolutionary computation is a set of optimization algorithms that use the basic processes of Darwinian nature: selection, breeding, mutation, inheritance, etc. to make decisions about a population based on objectives. This process is iterative and the intended outcome is progress, based on the objectives, in the population. In this
way, these techniques can be used as a global search tool. These algorithms also lend themselves to parallel computation in several ways, but primarily by simulating and evaluating the candidates in each population simultaneously. In this work, the techniques of evolutionary computation, specifically genetic algorithms (GA), are used to search the massive design space of the diagonal transition QCL for both the (100) and (111) crystal surface orientations.

5.1.1 Genetic Algorithm Terminology

In this section, a brief introduction of the terminology and techniques used in the implemented GA are given. This research is not intended to explore novel GA implementations, but rather to implement standard GA procedures in a novel application area. By far the biggest challenge of the GA technique is application specific implementation of the fitness evaluation. Additionally, a second challenge is determining a proper technique for representing a design that is amenable to traditional Ga techniques.

The first step of a GA is the creation and/or input of an initial population. The population is a set of candidate solutions. The candidate solutions are defined by a set of properties, called chromosomes, that represent the candidate. In the application of QCLs, these properties could consist of the number of wells and barriers, widths of each, and material compositions. The individual parts of the chromosome are called genes.
The chromosomes of a population can be expressed in several forms, but one of the most common is binary representation. Each gene, which is a design element such as the width of a well, is represented by a binary number. This plays a role in the methodologies used to form a child candidate from a parent or a set of parents. Binary crossover is a popular technique in which a binary mask of the same size as the chromosome is randomly generated. The ones in the mask will constitute selection of a property from the first parent and the zeroes in the mask will select the corresponding property from the second parent. Figure 5.1 illustrates the process of binary representation and crossover.

<table>
<thead>
<tr>
<th></th>
<th>Binary Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent 1</td>
<td>1 0 1 1 0 0 1</td>
</tr>
<tr>
<td>Parent 2</td>
<td>0 1 1 0 1 1 0</td>
</tr>
<tr>
<td>Mask</td>
<td>1 0 0 1 1 1 0</td>
</tr>
<tr>
<td>Child 1</td>
<td>1 1 1 1 1 0 0 1 if mask is 1, from Parent 1, else 2</td>
</tr>
<tr>
<td>Child 2</td>
<td>0 0 1 0 1 1 0 1 if mask is 1, from Parent 2, else 1</td>
</tr>
</tbody>
</table>

**Figure 5.1:** Example of per-bit binary crossover for an 8-bit gene sequence that produces two children from the set of parents.

Binary crossover can be implemented in many forms. Sometimes a crossover point, a value between 0 and the number of bits in a gene, is randomly generated. In this scheme, bits to the left of crossover point are selected from the first parent and bits to the right are selected from the second parent, and vice-versa for the second child. Children may also be derived from the mutation of genes from a parent or even the mutation of the genes derived after crossover. One method for mutation is shown in Figure 5.2. In this method, a random gene sequence is derived as well as a random
mask modified by the given mutation probability. The original gene is the subjected to binary crossover with the random gene sequence using the mutation mask. Another simple alternative to use the mutation mask as a flip-bit selector. If the mutation mask is 1 at a given bit then flip.

<table>
<thead>
<tr>
<th></th>
<th>Binary Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>1 0 1 1 0 0 1</td>
</tr>
<tr>
<td>Mutation</td>
<td>0 1 1 0 1 1 0 1</td>
</tr>
<tr>
<td>Mask</td>
<td>0 0 0 1 0 0 0 0</td>
</tr>
<tr>
<td>Child 1</td>
<td>1 0 1 0 1 0 0 1</td>
</tr>
</tbody>
</table>

*Figure 5.2:* Example of per-bit binary crossover for an 8-bit gene sequence that produces a mutated gene from the original and the random mutation gene.

Each candidate is simulated and the simulation output is evaluated. In the case of multi-objective algorithms, evaluation of each objective is considered and a fitness value is calculated. The fitness function can be an aggregate of the individual evaluation of each objective or it can be a fitness function matrix. In either case, this fitness function is utilized in the selection and generation of the next population. This process continues for either a set amount of iterations or until a stopping criterion is satisfied.

The key to genetic algorithms is in how each successive population is derived and the quality of the evaluation of each candidate solution. Several techniques are often used to derive the next population. One technique, designed to improve the search space, is fitness sharing. Fitness sharing reduces the fitness of candidate solutions
that are close to other candidate solutions. This increases the search space. With elitism, unimodal domains are searched with ease, but multimodal domains require the identification of more than one optimum region. In short, it pushes candidate solutions away from each other. Fitness sharing occurs before the selection phase. The algorithm takes the fitness of each of the candidate solutions and transforms them into a new fitness based on their old fitness and a distance measure. The distance measure can be the distance between the fitness, or it can be a distance between each chromosome to the other.

In genetic algorithms, the new population is derived from the old population through a selection of parents, carry-over candidates, and newly generated candidates. The parents are used to derive a new set of chromosomes which will define a child candidate. The selection method for parents will dictate the behavior of the search.

The roulette wheel is a fitness proportionate selection method. The sum of the fitness of each chromosome is normalized to 1, and then converted to a probability of selection, pi. A random number is then picked, and compared to the cumulative sum of p. The closest value is the winning parent. In some cases the roulette wheel eliminates solutions with high fitness. This allows for more search, but may result in a suboptimal solution. Other methods look to overcome that limitation.

Rank selection is a method of selection that is not based on how much better one fitness is than the other, but just if its higher or not. The fitness are ranked from highest to lowest. In some implementations the top candidates are selected as parents, a form of elitism. Alternatively, to help expand the search, randomization of the selection may be used. In this case, the highest fitness does have a higher chance of being selected, although it is not proportionate to the fitness. Just like in the
roulette selection, a random number is then used to pick the winning parent. Weaker parents are more likely to be selected in this method, but the selective pressure is more equal for all chromosomes. This is better for multimodal searches because it prevents one suboptimal from dominating the selections. Several other techniques exist which influence the behavior of the search process and will ultimately determine its success, but exhaustive description is beyond the scope of this thesis.

5.2 Genetic Algorithms and QCLs

Only a few examples are known to exist which apply the techniques of evolutionary programming to the design of QCLs [63]–[65]. The earliest known application [63], is by a group whose interests lie in the comparison of evolutionary algorithm techniques rather than the quality of the resulting QCL. This is evidenced by the selection of objectives. They used a multi-objective approach, but only two aggregate objectives were identified. The first aggregate consisted of the following: aligning of the injector energy level with the $E_3$ level and depopulation of the $E_2$ level to insure population inversion. The second aggregate consisted of a measure of the overlap ratio of each probability function, which they claim to be a determination of the ability of the electron to transition. They do not take into consideration factors such as LO-phonon scattering, electron lifetimes, or the length of the active region factors which play a major role in the lasing potential of the device.

Another group [64], defined their approach as inverse-quantum-engineering. In their methods they used a reference design as a comparison point to evaluate two case-studies. The first case-study explored the ability to change the emission output. The fitness evaluation is vague, but is defined to be the deviation of the wave function
shapes and the transition energies to that of the reference design. In the second case-
study, the focus was on the shifting of the confinement energy by means of changing
the material composition. The fitness evaluation was the same, as the goal was to
replicate the reference design within a different material system.

The third group [65], focused solely on the optimization of the active region under
the influence of a magnetic field and made no concession for the role of the injector
region. Also, they used an aggregate to form a single fitness function which primarily
was a measure of the optical gain. This group did, however, have the most parameter
variance. The barrier height was varied as well as the thickness of the each layer.
Both the second and third groups focused on the AlGaAs/GaAs material system. It
is unclear what material system the first group used. No genetic algorithm based
studies have been found which incorporate the full extent of design principals from
a multi-objective approach. Nor are there any studies known to exist which look at
strained material systems such as In(Ga)As on GaAs (100) or (111).

5.3 Genetic Algorithm as Implemented

In this section, the genetic algorithm implemented in this research is described. The
flow sequence of the algorithm is shown in Figure 5.3. Each stage is described in
more detail in the following subsections. The actual code implemented is found in
the Appendix. MATLAB was used to code the algorithm due to its built in library
of mathematical functions and the ease in which it can operate in parallel, utilizing
the power of today’s multicore processors.
Figure 5.3: Diagram of the flow of the genetic algorithm.

The general parameters of this genetic algorithm are shown in Table 5.1. The application specific parameters are shown in Table 5.2. As previously discussed in Chapter 3, the ideal injector region consists of 5-7 well/barrier pairs. Because of limitations in the input file of the software, this value was chosen to be fixed at 5 well/barrier pairs. Additionally, as will be seen, the genetic algorithm makes no effort to test the effectiveness of the injector region other than to assure the last stage of the injector has an energy level just below that of the upper excited state, E_3. Although each injector well and barrier may be varied by the search they would likely not be optimized. In general, this algorithm is designed (in it’s current state) to optimize the active region and the last injector well. Further discussion on this is given in the
next chapter. The design of interest is the traditional three well diagonal transition active region. Five (5) injector pairs and 3 active region pairs makes for a total of 16 layers which may be varied in the device. The other 2 variables are the concentrations of each well and barrier.

Table 5.1: General Simulation Parameters for the GA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runs</td>
<td>6 (3 pairs)</td>
</tr>
<tr>
<td>Initial Population Size</td>
<td>8</td>
</tr>
<tr>
<td>Population</td>
<td>64</td>
</tr>
<tr>
<td>Generations</td>
<td>40</td>
</tr>
<tr>
<td>Variables</td>
<td>18</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
</tr>
</tbody>
</table>

At the time of this writing, a total of 6 runs (3 pairs - (100) and (111) orientations) of the GA have completed. These runs consist of a population of 64 candidates in each generation and 40 generations. This makes for a potential for 2,560 individual candidates to be tested in each run. However, this algorithm includes elitism via carryover, therefore the number of unique candidates is lower. Some changes to the algorithm were made between successive pairs of runs. These changes and their effects are analyzed in more detail in the next chapter.
Table 5.2: Application Specific Simulation Parameters for the GA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value(s)/Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Indices</td>
<td>(100) &amp; (111)</td>
</tr>
<tr>
<td>Well Widths</td>
<td>1/2 ML - 64 ML</td>
</tr>
<tr>
<td>Barrier Widths</td>
<td>1/2 ML - 64 ML</td>
</tr>
<tr>
<td>Well Alloy %</td>
<td>0 - 100</td>
</tr>
<tr>
<td>Barrier Alloy %</td>
<td>10 - 45</td>
</tr>
<tr>
<td>Active Region Wells</td>
<td>3</td>
</tr>
<tr>
<td>Active Region Barriers</td>
<td>3</td>
</tr>
<tr>
<td>Injector Well</td>
<td>5</td>
</tr>
<tr>
<td>Injector Barriers</td>
<td>5</td>
</tr>
</tbody>
</table>

5.3.1 Initial Population

In this genetic algorithm, an initial population of QCL candidates is generated from a set of 8 initial parents. These parents are known working QCLs or parts of working QCLs found in literature [56], [57], [66]–[69]. The initial stage of the algorithm converts the initial population into a gene sequence. Each of the parent genes is then crossbred with every other parent to generate full population of 64 candidates. These resulting candidates are also randomly mutated by a mutation rate (2%). Figure 5.4 gives a flow sequence of the conversion of the initial population.

![Figure 5.4: Schematic flow for the initial population.](image)

The initial population is given in values of percentage for the alloy concentrations and in nanometers for the well and barrier heights. The gene sequence in this algorithm is a real valued sequence of integers. The alloy percentages are carried over, but the nanometer lengths are converted into 1/2 monolayer values. The conversion
of the lengths takes into account the alloy percentages of the well and barrier and converts this into a lattice constant \( a \) using Vegard’s Law. A diagram of the gene sequence is given in Figure 5.5. Once converted the gene sequences are then subjected to crossover and mutation operations.

<table>
<thead>
<tr>
<th>Alloy %: AlGaAs: 10-45%</th>
<th>InGaAs: 0-100%</th>
<th>...</th>
</tr>
</thead>
</table>

**Figure 5.5:** Generic gene sequence with value ranges.

### 5.3.2 Candidate Generation: Crossover and Mutation

The initial population of 8 candidates, once converted to a gene sequence, is crossbred with every other candidate. The crossover operation feeds into a mutation operation. These operations are real valued. Therefore, crossover occurs as described in the previous section but an entire gene is taken from one parent or the other based on a binary mask. Mutation occurs post real-valued crossover operation on the two resulting children. A random gene sequence is generated with uniform distribution over the given ranges. Then a mutation mask is randomly generated with a 2% probability over a uniform distribution. If the mask value for a given gene is true, then the new gene is inserted into the sequence for the child candidate.
The real-valued crossover and mutation operations are the only operations performed on the selected parents from generation to generation. However after the initial generation, the population is no longer derived from the starting population, but rather through the selection of parents from the previous generation, which is described below. One aspect of this process is that random genes are generated and parts of these genes are inserted into the population via crossover and mutation with some of the selected parents.

5.3.3 QCL Input Files

After the generation of the new population has occurred, the program then converts each gene sequence into an input file that can be read into the nextnano software. This is made possible through the use of a template file which uses variables to determine the geometry and material parameters of the device and generate the grid which is used to solve for the wavefunctions and other parameters. The entire template file is imported into a MATLAB cell array. Rows of this cell array are manipulated at the string level and then this cell array is written to a new file, which becomes the input file for a given candidate. Note that this operation, as well as the running of the nextnano software, the data collection and fitness evaluation all occur in parallel (multiple candidates at a time) via a parfor loop in MATLAB.

Because the values in the template file are given in floating point nanometers and the gene sequence is an array of integers, the values are converted using Vegard’s Law (taking into account the alloy percentages) into decimal string values. This conversion occurs only once for a given gene sequence so as not to introduce conversion errors that could cause problems when collecting the data necessary for fitness evaluation.
5.3.4 \textit{nextnano}^3

As was discussed previously at a theoretical level in Chapter 3, \textit{nextnano}^3 is a Schödinger-Poisson solver. While it is capable of solving both the 8x8 and 6x6 kp approximations, this level of accuracy would take lots of computational time and is not necessary for the extracting pertinent data of the QCL device. Instead, the software is configured to solve the single-band approximation while taking into account the effective mass. It does this self-consistently with Poisson’s equation, which is initially solved. The software is configured, via the input file, to generate the first 20 eigenvalue solutions to the wave equation. The software allows specification for the data to be output in text files. These text files are used to collect data from the simulation.

5.3.5 \textit{Data Collection}

Perhaps the biggest challenge of this genetic algorithm is importing the text based data that is output from the \textit{nextnano} simulation. Fortunately the data output is generally consistent in format for a given run, but the data is mixed with both text and numerical data in a variety of formats. Therefore, most of the data has to be read into MATLAB using primitive string commands and then converted appropriately. Comments given in the code contained in the appendix help the reader to understand different approaches to data collection and conversion.
The data that is collected for each candidate is designed around the needs to evaluate the fitness of a given candidate. This includes the energy level differences for each well of the active region and the last injector well. To determine the pertinent levels for each well of the active region, both the shift offsets and the $\psi^2$ data are imported. After testing several methods to sort the wavefunctions and determine the energy levels belonging to a given well, perhaps the most simple method worked best. Each of the wavefunctions were summed over the length of a given well. Since there are four pertinent wells, this resulted in a (4 x 20) matrix. This matrix was then sorted in descending order and the index of each was saved. This index was then used to determine the energy level from the shift offset matrix.

Once the $E_{inj}$, $E_3$, $E_2$, and $E_1$ levels were determined then the relevant electron lifetimes, $\tau_3$, $\tau_2$ and $\tau_{32}$ could be imported from the output data. One particular caveat to this is the possibility that levels determined by the previous sorting method were not in the correct order for proper QCL design: $E_3 > E_{inj} > E_2 > E_1$. If this or any level is out of order then the $\tau_{32}$ level would likely not exist in the data and could cause errors in the calculation of the gain metric. To avoid this problem, an error check was performed. In the special case that $E_2 < E_1$ the algorithm would sum the wavefunctions over the range from the start of well 2 to the end of well 3 and determine the top 2 levels. It would then order these levels so that it was assured that $E_2 > E_1$. However, this does not guarantee that all the levels are in order, so the full condition must still be tested. If this test fails then the $\tau_{32}$ lifetime data is
arbitrarily set to a value that will produce a low fitness for the gain metric. One additional piece of data that is being extracted from the nextnano output is the value of the LO-phonon for the particular well material. Since the material composition can change, this value is not always a constant.

5.3.6 Fitness Evaluation

This algorithm was designed around optimization of 4 objectives: emission in the 3-5 µm range, LO-phonon scattering between the \( n = 2 \) and \( n = 1 \) states, injector level alignment just below the \( E_3 \) level, and a gain metric based on the lifetimes of the \( n = 3 \) and \( n = 2 \) states. These objectives are mathematically shown at the end of Chapter 2. Each of the fitness values are designed to be maximized.

The difference in the \( E_3 \) and \( E_2 \) energy levels is normalized to a value between 0 and 1 using a piecewise function. If the value lies between 0.25 and 0.41 then the fitness value for this objective is 1, if the value lies outside this range then is given a fitness value of the normalized normal distribution with \( \mu = 0.25 \) and \( \sigma = 0.03 \) for values that lie below 0.25 and \( \mu = 0.41 \) for values that lie above 0.41.

The fitness value for the LO-phonon metric is calculated by taking the normal distribution, normalized for a value between 0 and 1, of the difference of \( E_2 \) and \( E_1 \) with \( \mu = \text{LO-phonon} \) and \( \sigma = 0.03 \). These normal distributions fall off in value pretty quickly and act to encourage tight selection of this objective in highly ranked candidates.

The injector level fitness value is taken by calculation of the extreme value distribution, normalized between 0 and 1, of the \( E_{\text{inj}} \) level where \( \mu = E_3 - 0.01 \) and \( \sigma = 0.03 \).
Since the gain metric does not have a target value, but rather needs to be maximized the fitness value of this metric is calculated by calculating the sigmoid function (was originally set to the hyperbolic tangent for the first 4 runs)

\[ y = \frac{1}{1 + e^{-ax+c}} \]  

where \( a = 0.5 \) and \( c = 1 \). This function produces values between 0 and 1 and is situated to have variance in the range where the calculated lifetime should lie.

While fitness functions are evaluated for each objective, aggregate functions are calculated for each pair of objectives by taking the mean. Finally the mean of all fitness values is also taken for a complete aggregate fitness.

5.3.7 Parent Selection

At this time, parents are selected via simple rank selection. Each candidate is ranked based on the fitness of each objective and the top 2 of each are selected for carry over into the next generation. Additionally, the top candidates from the aggregate ranks, mean of the two fitness values, of pairs of objectives are selected to carry over. For 4 pairs this equates to 6 selections. Finally, the top 2 from the complete aggregate fitness values are selected as the final carry over candidates. This means that 1/4 of the next generation are carry over’s from the previous generation.
This is very elitist, which has the potential to find solutions very quickly, but at the risk of being stuck in a locally optimized regime. To mitigate this, random genes are inserted into the population by mating (crossover) the top candidates from each of the fitness functions with these genes. Since each crossover operation produces 2 children, this means that 8 candidates in each generation are highly likely to contain new designs. The rest of the population is generated by breeding (crossover and mutation) the carryover population.
CHAPTER 6

Genetic Algorithm Simulation Results

The genetic algorithm results are classified into categories that align with the fitness evaluation objectives outlined in the previous chapters. Graphs of the QCL candidate bandstructures with corresponding wavefunctions are given for the top candidates in each objective ranking, as well as aggregate rankings. These graphs serve as representations of what is considered the 'top' output of the algorithm, however the practicality of these candidates is discussed. Additionally, some commentary is given on the performance of a given ranking scheme and how additional ranking metrics could improve the outcome of the algorithms. As a result, some additional experiments were performed with additional added fitness ranking metrics. These results are given at the end of the chapter.

6.1 Top Candidates by Individual Fitness Rank on GaAs

Figure 6.1 shows the top candidates by fitness ranking of objective 1 for devices simulated in the (100) and (111) orientations. The first objective is emission of the QCL in the 3-5 µm range and the fitness value is calculated by a piecewise function as follows:
Figure 6.1: \( \text{Al}_x\text{Ga}_{1-x}\text{As/In}_y\text{Ga}_{1-y}\text{As QC laser candidates on GaAs a) (100) and b) (111) with top fitness rank for the emission objective.} \)

Looking at these candidate solutions, there have been some major changes from the initial generation QCL designs. For this objective the well alloy percentage is high, especially for the (100) candidate at 25%, which is an unsurprising consequence. However, the barrier alloys are not at the max value of 45% as one might expect. Instead, they are 30 and 35 %, respectively. Alternatively, the larger energy separation is achieved by significant narrowing of the first active region well. For the (100) candidate, the significant narrowing of the first active region well has pushed the energy level near the continuum. This would likely lead to a device with low efficiency due to electron escape into the continuum. To combat this issue, it is proposed that the fitness function could be modified to have a penalty for energy levels the approach the upper value of the conduction band at the beginning of the active region.
The (100) candidate also has $E_1$ and $E_2$ levels that are reversed in their maximum probability in their respective wells. This appears to be a consequence of them not being in proper resonance with a near LO phonon energy separation. This kind of result, could potentially be avoided by using a better scheme of fitness sharing or a more sequential approach to optimization. For instance, because it is clear that objective 1 seems to modify the alloy percentages and the first active region well width, then these changes could be confined to a progressively more narrow range and then the candidates would be passed into a objective 2 minimization scheme. Further testing of both types of proposals would be warranted as future work.

Figure 6.2 shows the top candidates by fitness ranking of objective 2 for devices simulated in the (100) and (111) orientations. The second objective is $E_2 - E_1$ proximity to LO phonon and the fitness value is calculated by

$$f_2(x) = normdist(x, E_{lo}, 0.03)$$

Figure 6.2: Al$_x$Ga$_{1-x}$As/In$_y$Ga$_{1-y}$As QC laser candidates on GaAs a) (100) and b) (111) with top fitness rank for the LO phonon resonance objective.
The top candidates for the second objective do not have a particular trend with regard to alloy percentage and as can be seen in the (100) sample, do not have optimized emission. The optimized emission in the (111) sample is likely a result of the fitness sharing that is present in the parent selection and new generation breeding. Objective 2 has a lot more diversity in the well and barrier widths for all other layers except active region wells 2 and 3. For these wells, the objective narrows down on the range of widths that produce good LO phonon resonance. This kind of performance leads to further investigations using a sequential optimization of objectives. Alternatively, more objective sharing in the selection and breeding processes or increased complexity in the fitness evaluation may also produce better results. It is important to understand that this is a look at a single objective’s results and it would not be expected that these candidates would be working devices overall.

One interesting feature is the large barrier between the first and second active region wells. Since, this large barrier also exists in the top solutions for objectives 3 and 4 this is likely a result of the mutations that occurred during breeding. Optimization of those objectives does not account for a high transition probability for $E_{32}$, which would necessitate a much more narrow barrier.

One issue that is not present in these candidates, but did occur in other candidates in these runs and other runs was the possibility that the $E_2$ or $E_1$ levels fall below the first injector well. In this case, the tunneling probability from the end of the active region into the injector region would suffer. A suggested modification is to penalize objective 2 for this possibility.
Figure 6.3 shows the top candidates by fitness ranking of objective 3 for devices simulated in the (100) and (111) orientations. The third objective is $E_{\text{inj}}$ alignment just below the $E_3$ energy level and the fitness value is calculated by

$$f_3(x) = \text{evdist}(E_{\text{inj}}, E_b, .03)$$

(6.3)

The top candidates for objective 3 again show a lot of diversity in well and barrier widths overall and likely make them poor candidates for lasing, however they effectively align the $E_{\text{inj}}$ level, shown in blue, just below the $E_3$ level. This appears to be accomplished by changing the width of the last injector barrier and varying the width of the last injector well. One thing that would likely improve the performance of this objective is including functions that will direct the entire injector region to progressively increase the energy of the electron up to the desired level. In its current state, this would not necessarily be a consequence of the ranking mechanism and is evidenced by the large injector wells present in both candidate solutions. These large
wells lower the minimum energy level in a given well and introduce higher confined
states. The effect of this is unknown, but would likely decrease the injection efficiency
as well as other undesired consequences. Another resulting issue is the reduction or
elimination of the minigap feature that has previously been shown to improve tran-
sition efficiency.

Figure 6.4 shows the top candidates by fitness ranking of objective 4 for devices
simulated in the (100) and (111) orientations. The fourth objective is the gain metric
and the fitness value is calculated by

\[ f_4(x) = \text{sigmf}(x, [.5, 1]) \] (6.4)

Figure 6.4: Al\textsubscript{x}Ga\textsubscript{1-x}As/In\textsubscript{y}Ga\textsubscript{1-y}As QC laser candidates on GaAs a)
(100) and b) (111) with top fitness rank for the gain metric objective.
In these results, the injector regions have similar anomalies as was previously discussed but if these candidates do often also satisfy the LO phonon resonance objective. In fact, the actual (111) top candidate for this objective is the same as for objective 2. The pictured candidate is the 2nd top candidate for this objective, which happens to also be the 2nd top candidate for objective 2. This result is actually not surprising and follows from the population inversion criteria for high gain. It affirms that population inversion is improved when LO phonon resonance is achieved.

The gain objective is however not entirely redundant, as it also takes into account the overall $E_3$ lifetime. It appears from the results, that in the current form, this results in a propensity for a large barrier between active region wells 1 and 2. This decreases the scattering of electrons in the 1st well and results in a large value for $\tau_3$. However, this negatively affects the coupling between $E_3$ and $E_2$ and results in a low oscillator strength. A better gain metric needs to include the oscillator strength. It is clear that these values are competing and add to the $f \cdot \tau$ figure of merit methods Faust [4] used to determine the best barrier widths.

Another factor that might be a good addition to this objective or objective 3 is the addition of the $E_3$ level facing a minigap. This, in many ways, is mostly influenced by the design of the injector region as it supplies a vast majority of the energy levels that could potentially line up between the $E_3$ and $E_2$ levels.
Figure 6.5 shows the top candidates by aggregate fitness ranking (mean value) for devices simulated in the (100) and (111) orientations. The aggregate fitness is the sum of all the fitness function values in this case. If each of the fitness functions collectively do a good job of optimizing, then the ideal situation would be that the aggregate fitness ranking would produce the most viable candidate solution. Given the results of the individual functions, it is already clear that the injector region is not optimized.

![Figure 6.5: Al\textsubscript{x}Ga\textsubscript{1−x}As/In\textsubscript{y}Ga\textsubscript{1−y}As QC laser candidates on GaAs a) (100) and b) (111) with top aggregate fitness rank.](image)

The active regions of the top candidates from the total aggregate fitness are clearly unconventional in a traditional sense, but they do produce candidates that meet all the objectives quite well as they are calculated. Both candidates achieve a perfect score in all the objectives. Forgiving the lack of injector optimization, it is unclear if these devices would work from a traditional perspective.
As was seen with the first objective, these solutions both have very narrow first active region wells. Perhaps as a surprise, the alloy percentages are relatively low at 0 and 5% ((100) and (111)). The second objective is clearly achieved, but the large barrier in the (100) candidate is likely to cause low $E_{21}$ scattering rates. This is the same for the large barrier that is present between active region wells 1 and 2. As discussed, this is counter to the need for a large oscillator strength, which is another factor in the total gain equation.

### 6.2 Average Fitness Values

In this section, we explore the average fitness values vs generation for the GaAs (100) and (111) substrate orientations. Figure 6.6 shows the trend of the fitness value mean for each generation in achieving objective 1. Between the (100) and (111) runs, the average fitness value for objective 1 shows a different trend overall. The (100) sample initially increases, but then drops and has a more steady value. This is likely a sign of diversity introduced by the mutation operator and random gene insertion. However, this general trend doesn’t seem to hold for the (111) orientation run. It has a large initial increase and then, on average, shows a steady increase for the life of the run. The slight dips in some generations show the presence of diversity from one generation to another. Overall, the diversity is good for assuring adequate coverage of the search space; however, one would expect a steady average rising trend given a long enough run. Given the difficulty of achieving the wavelength in real life applications, the low steady value for the (100) run may be an indication of the difficulty of the algorithm to find multiple good and diverse candidates. Further studies are warranted.
Figure 6.6: Average fitness value over the generations for Objective 1 on GaAs a) (100) and b) (111)

Figure 6.7 shows the trend of the objective 2 fitness value mean for each generation. The average fitness values for both sets of runs show diversity but overall do show an increasing trend. It is notable that the average values are much higher for the (111) runs even though they both start in similar circumstances in Generation 1. This may be an indicator that the given objectives are easier to achieve using the (111) orientations. Further testing using multiple short runs is needed.

Figure 6.7: Average fitness value over the generations for Objective 2 on GaAs a) (100) and b) (111)
Figure 6.8 shows the trend of the objective 3 fitness value mean for each generation. For both sets of runs the objective shows a slow steady rise through the first 30 generations then a sudden increase. Since the runs are independent and random, this is most likely a coincidence. Again we see average values for the (111) run higher than the (100) run.

![Figure 6.8: Average fitness value over the generations for Objective 3 on GaAs a) (100) and b) (111)](image)

Figure 6.9 shows the trend of the objective 4 fitness value mean for each generation. Both runs show increasing trends, but the (100) run started to show some leveling off while the (111) run showed an relatively large rate of increase from one generation to the next. Values for the (111) run are higher than the (100) but a little less of a gap as in other objectives. It is clear that for each individual objective, the (111) run is on average trending toward higher fitness evaluation. The finesses are not biased in anyway, but this is an initial strong indicator that the (111) surface is preferable or at least has more potential candidate solutions.
Finally, figure 6.10 shows the trend of the aggregate fitness value mean over each generation. Both runs increase in average aggregate fitness over the course of the run. Nothing too notable that isn’t already present in the individual objectives as this is just a sum.

Figure 6.10: Total aggregate average fitness value over the generations on GaAs a) (100) and b) (111)
6.3 Top Candidates for Modified Objectives

In this section, a short 15 generation run is presented in which some modifications to original fitness calculations are used in hopes to improve the results that were given in the previous sections of this chapter. The run was only conducted on the (111) surface orientation. Figure 6.11 shows the top QCL design candidate for objective 1. The old fitness function is modified as shown in the following equation. The resulting solution shows improvement in this area, but the modification may be too strong given the large alloy percentage and the low position of the level in the first active region well.

\[ f_{1,\text{mod}}(x) = f_1(x) - \text{sign}(E_3, [10, CB_{\text{max}}]) \] (6.5)

![QCL Design Candidate](image)

**Figure 6.11:** Al_{x}Ga_{1-x}As/In_{y}Ga_{1-y}As QC laser candidates on GaAs (111) with top fitness rank for the modified emission objective.
Figure 6.12 shows the top QCL design candidate for objective 2. This objective was modified to penalize $E_1$ values significantly below the conduction band minimum. Mathematically, this is shown in the following equation. This seems to have the effect of pushing the algorithm toward results with very little or no alloy percentage. This result is not surprising, but it does make the candidate solution even less likely to work overall. In the given example, an $E_3$ level was not found, but this is more likely the result of the multiple wide wells in the injector region and the 20 eigenvalue calculation setting given for the simulation.

$$f_{2,\text{mod}}(x) = f_2(x) - (1 - \text{sgn}(E_1, [10(CB_{\text{min}} - .25)])$$

(6.6)

Figure 6.12: Al$_x$Ga$_{1-x}$As/In$_y$Ga$_{1-y}$As QC laser candidates on GaAs (111) with top fitness rank for the modified LO phonon objective.
Figure 6.13 shows the top QCL design candidate for objective 3. The given candidate solution is perhaps one of the more viable overall solutions produced as a top candidate for a given objective. This objective was not modified from the original. It is likely a result of the carryover from the initial population, given the likelihood of these initial solutions to have good injector alignment.

Figure 6.13: $\text{Al}_{x}\text{Ga}_{1-x}\text{As/In}_{y}\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top fitness rank for the injector alignment objective.

Figure 6.14 shows the top QCL design candidate for objective 4. The gain metric top candidate produced an active region that is also fairly viable, although it fails to meet the main emission design criteria. Unlike the previous modifications, modifying this objective by adding in the oscillator strength to the gain metric produced better
a result without the obvious compromise. It seems this approach, along with some modifications to assure a viable injector region and proper emission, is the key to producing good resulting solutions. However, a good test of this statement would be to start with all random solutions rather than an initial population.

Figure 6.14: $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top fitness rank for the modified gain metric objective.

Figure 6.15 shows the top aggregate fitness QCL design candidate. This candidates solution shows aspects of each modified fitness function solutions previously discussed. However, it seems the very large alloy percentage of 22% is unrealistic as it necessitates a very wide last injector well to align the level. This large alloy percentage also leads the large escape rate of electrons into the continuum as is evidenced by the $E_3$ probability curve.
Figure 6.15: $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ QC laser candidates on GaAs (111) with top aggregate fitness rank.

While experiments, further modifications to the objectives, as well as additional objectives are certainly warranted, several other approaches may also produce good results. For example, depth first search methods and/or Monte Carlo methods could also be effective. Any of these strategies still rely on appropriate and careful fitness evaluation of candidate designs. The evolutionary algorithm/genetic algorithm approach does help identify the key traits that are necessary for producing working devices. It is clear that an evolutionary strategy approach, which narrows the design variable ranges in a systematic way would also produce useful data and may prove to be just as effective in producing viable results.
CHAPTER 7

Summary and Concluding Remarks

This dissertation set out to explore the possibility of quantum cascade laser designs based on the non-traditional GaAs(111)B surface. This initial work on this project is a set of growth experiments that show the possibility of pseudomorphic growth of GaAs, AlGaAs, and most importantly highly strained InAs layers on the GaAs(111)B substrate. While the temperature range and As\textsubscript{4} overpressure requirements for quality growth of each of these materials is narrow when compared to the GaAs(100) surface counterpart, the results were conclusive that pseudomorphic growth is possible.

The most important aspect for controlled growth of AlGaAs and InAs layers on GaAs is precise control over adatom migration length by variation of temperature or As\textsubscript{4} overpressure. Al adatoms have inherently slower migration and therefore prefer higher substrate temperatures or lower BEP ratios. In contrast, In adatoms require lower temperature and higher BEP ratios.

At the end of Chapter 4, a conclusion was given that the usefulness of binary InAs on GaAs substrates for a device structure like the QCL is limited if the InAs wells are only situated in the active region. Therefore for practical reasons, the alloy content of In in the wells of the active region for such a device would be limited. This further enhances the promise that novel device structures based on strained layer incorporation in the active region are possible. Ternary InGaAs at reasonable alloy
content percentages would be less strained and will likely have growth qualities that are more like that of GaAs on GaAs (111). This is a perfect compromise, as devices with In incorporation can benefit from lower threshold voltages and higher efficiency while extending the operating range.

To explore this concept a multi-objective genetic algorithm was developed and used to search the design space made possible by the promise of psuedomorphic growth of InGaAs alloys on the GaAs(111)B surface. The algorithm was successful in searching the design space, but was ultimately limited in many of its conclusions because of the objectives. The algorithm found designs, which it ranked high in all or most of the objective categories, but would not work in real life. Further enhancement of the algorithm is possible by adding more objectives that will regulate the fitness of the candidate solutions more effectively. The algorithm was successful based on the objectives it was given, but these objectives by themselves do not conclusively produce working devices. Ultimately, there is hope that with further refinement, the algorithm would be successful in finding working devices and ranking non-working devices more effectively.

To that end, further development of the algorithm would begin with fitness evaluation or penalties for the following items:

- $E_3$ level in the continuum: tested in short run
- $E_1$ level below lower conduction band edge of the injector: tested in short run
- Wavefunction overlap, $z_{32}$, and/or oscillator strength $f_{32}$: tested in short run
- Working injector design based on graded transitions
- $E_3$ facing a minigap
To do this would require continued development and effort. Additionally, this algorithm uses simple mechanisms for its parental selection and enhancement of multimodal search range. Future algorithm development should expand the usage of more advanced parental selection methods to provide a much quicker and more effective search. More advanced evolutionary algorithms may be in order to help narrow the variable range. This could be accomplished using a sequential objective evaluation approach, where one objective is optimized and used to narrow the variable parameters, then another is objective, and so on. There are many potential approaches, but overall the results of this dissertation work show promising possibilities for the viability of an intelligent search approach to solving the problem of finding good device structures for a given set of objectives.
CHAPTER A

Project Code

A.1 Main Code Routines

Here, the main code routines the makeup the genetic algorithm are listed. All code
given is from the final run after the modified fitness functions.

A.1.1 main.m

This main function is responsible for the general sequence of the program. It mainly
serves to call the other functions that retrieve and calculate data and/or call the
simulation program. This routine also invokes many tasks in parallel using a parfor
loop. A matlab pool must be setup before running this routine in order to properly
execute in parallel.

```matlab
1 %main program for Dissertation work on QCL automated design using ...
2    genetic
3 %algorithm
4 clear pop gen ftemplate initpop candidates m n idx ipl retcode
5 tic
6 pop = 64; %population of candidates per generation, number is ...
7    set to 8X the initial population
8 gen = 15; %number of generations
9 fp = '.\output\';
10 ftemplate = getnn3template(5); %get the file template for making ...
    an input file for nn3 (note: 5 stands for 5 well injector)
11 %initial population data
12 initpop(1,:) = [45 0 17 18 20 26 46 34 30 28 30 30 11 11 28 19 54 ...
    48]; %300K operation of a GaAs-based QCL APL 78 pp3529-3531
13 % initpop(2,:) = [33 0 17 18 20 26 46 34 30 28 30 30 11 11 28 19 ...
    54 48]; %APL 73 pp3486; Electron Letters 35 1848
```
candidates = createcstruct(pop,1);

%generate children from initial population
parfor n = 1:ipl
    %convert initpop to genes with monolayer units
    candidates(n).genes = calcgens(initpop(n,:));
end

clear initpop;

for m = 1:(ipl-1)
    for n = 1:(ipl-m)
        %run binary crossover on initial population
        [candidates(idx).genes, candidates(idx+1).genes] = ...
            binary_crossover(candidates(m).genes, ...
                candidates(m+1).genes);
        idx = idx + 2;
    end
end

clear idx;

avgs = nan(4,gen);
totavgs = nan(1,gen);

% figure(1);
% plot(m,avgs(1,:)) xlabel('Generation'); ylabel('Average ...
    Fitness Value'); title('Plot of Average Fitness Value');
% axis([1 gen 0 1]); hold all; plot(m,avgs(2,:)); ...
    plot(m,avgs(3,:)); plot(m,avgs(4,:));

for m = 1:gen
    parfor n = 1:pop
        candidates(n).genestr = convgenes(candidates(n).genes);
    end
end
A.1.2 createnn3in.m

The createnn3in.m program uses an input file template formatted for nextnano$^3$ to generate a new input file based on the gene sequence of a given candidate. It does this by importing the template as a string array.
A.1.3 getnn3template.m

This function imports the nextnano\(^3\) template file.

```matlab
function [ ftemplate ] = getnn3template( numinj )
%getnn3template: import the nn3template file to a char array
% Detailed explanation goes here

filepath = '.\input\';
if numinj == 5
    fpfn = strcat(filepath,'1DQCL_vars.in');
elseif numinj == 6
    fpfn = strcat(filepath,'1DQCL_vars6.in');
elseif numinj == 7
    fpfn = strcat(filepath,'1DQCL_vars7.in');
elseif numinj == 8
    fpfn = strcat(filepath,'1DQCL_vars8.in');
end

fid = fopen(fpfn,'r');
umlines = 0;
tline = fgets(fid);
while ischar(tline)
    numlines = numlines+1;
end
fclose(fid);
```
A.1.4 runnn3.m

This program opens a command line and runs the nextnano\(^3\) program with certain input specifiers that dictate the input file. The input file is determined by the filename data in the QCL structure data.

```
function [ retcode ] = runnn3( QCL )
%runnn3: Run the DOS executable for the nextnano3 with the input ... 
  file and
%output directory information
basedir = 'D:\Dissertation Project Storage\nextnano\';
ifn = strcat(basedir,'input\',QCL.filen,'.in');
outdir = strcat(basedir,'output\',QCL.filen,'\');
codeloc = 'cd "D:\nextnano\nextnano3\Intel 64bit" & '
command = strcat(codeloc,'nn3_Intel_64bit.exe -inputfile ... 
            
            
retcode = dos(command);
movefile(ifn,outdir);
mfn = strcat(ifn,'.no_macro');
movefile(mfn,outdir);
end
```

A.1.5 getqcldata.m

The getqcldata.m program imports the useful output data from the simulation into the QCL structure. This data is used to determine the fitness values for each candidate.
function [ QCL ] = getqcldata( QCL )

%**************************************************************************
%getqcldata get data from output of nextnano3
% Detailed explanation goes here
%**************************************************************************
basepath = 'd:\Dissertation Project Storage\nextnano\output\';
filepath = strcat(basepath,QCL.filen, '\');

%get wavefunction data from psi_squared.dat
fpfn = strcat(filepath, ...
    'sg_1band1\cb001\qc001\sg001\deg001\dir.psi_squared_shift.dat');
P = importdata(fpfn);
QCL.spectrum(:,1) = P.data(1,2:21);
%QCL.spectrum = transpose(QCL.spectrum);
QCL.psi2sh = P.data(:, [1,22:41]);

%get wavefunction data from psi_squared.dat
fpfn = strcat(filepath, ...
    'sg_1band1\cb001\qc001\sg001\deg001\dir.psi_squared.dat');
P = importdata(fpfn);
QCL.psi2 = P.data;

%get conduction band profile data from cb1D_Gamma.dat
fpfn = strcat(filepath, 'band_struc1\cb1D_Gamma.dat');
CB = importdata(fpfn);
QCL.cb = CB.data;

%get total lifetime data from intraband_pz1D..._dir.dat
fpfn = strcat(filepath, ...
    'sg_1band1\intraband_pz1D\cb001\qc001\sg001\deg001\dir.dat');
QCL.IB = importdata(fpfn);
for n = 0:19
    LT(n+1) = textscan(QCL.IB{38+24*n,1},'%*s %*s %*s %f %*s');
end
LT = cell2mat(LT);
QCL.taui = LT;

%get dipole matrix element data from ...
    intraband_pz1D\cb001\qc001\sg001\deg001\dir_2Dplot.dat
fpfn = strcat(filepath, ...
    'sg_1band1\intraband_pz1D\cb001\qc001\sg001\deg001\dir_2Dplot.dat');
DME = importdata(fpfn);
QCL.dme = zeros(20,20);
for m = 1:20
    for l = 1:20
        QCL.dme(l,m) = DME((m-1)*20+l);
    end
end
A.1.6 findfitness.m

The findfitness.m file calculates the fitness of each objective and stores the data in the QCL structure.
%extract relevent values
E3 = QCL.spectrum(QCL.idx(1));
E2 = QCL.spectrum(QCL.idx(2));
E1 = QCL.spectrum(QCL.idx(3));
Einj = QCL.spectrum(QCL.idx(5));
i = QCL.idx(1);
j = QCL.idx(2);

ir1 = (QCL.ar(2)+9):1:(QCL.ar(2)+14);
ir2 = (QCL.ar(7)+15):1:(QCL.ar(7)+20);
CBmax = max(QCL.cb(ir1,2));
CBmin = min(QCL.cb(ir2,2));

%calculate fitness for target wavelength
QCL.fitness(1,1) = E3 - E2;
fit1adj = sigmf(E3, [10 CBmax]); %adjust target wavelength ...
    fitness value if E3 is in the continuum
QCL.fitness(1,2) = obj1fit(QCL.fitness(1,1)) - fit1adj; %combines ...
    uniform and normal distributions

%calculate the fitness for the E2 - E1 proximity to LO phonon
QCL.fitness(2,1) = E2 - E1;
fit2adj = 1 - sigmf(E1, [10 (CBmin-.25)]); %adjust LO phonon ...
    fitness if E1 level is below CBmin in the first well of injector
QCL.fitness(2,2) = ...
    normpdf(QCL.fitness(2,1),lo,.03)/normpdf(lo,lo,.03) - fit2adj; ...
    %normal distribution

%calculate the fitness for the injector alignment to the E3 level
QCL.fitness(3,1) = E3 - Einj;
if QCL.fitness(3,1) >= 0
    Eb = E3-.01; %move the max of the objective distribution just ...
                below the E3 level
    QCL.fitness(3,2) = evpdf(Einj,Eb,.03)/evpdf(Eb,Eb,.03); ...
                %extreme value distribution
else
    QCL.fitness(3,2) = 0.01;
end

%calculate gain/population inversion metric
QCL.fitness(4,1) = QCL.taui(i)*(1 - (QCL.taui(j)/QCL.tau32));
fit4adj = sigmf(QCL.osc32, [.5 1]);
QCL.fitness(4,2) = (sigmf(QCL.fitness(4,1), [.5 1]) + fit4adj) / 2;
end
A.1.7 getnextgen.m

This program generates the next generation by calling the binary crossover and mutation routines.

```matlab
function [ newgenes ] = getnextgen( genes, obj )

%getnextgen: Generates next generation population from the current
% generation.

%Elitism: Top 4 ranked from each objective carry over and mate
% as follows 1-2, 3-4 for a total of 16 offspring.
pop = length(genes);
newgenes = zeros(18,pop,'uint8');

%new for runs 7+ - calculates aggregate fitness values
[~, ix] = sort(obj,2,'descend');
objs(1,:) = mean(obj(1:2,:));
objs(2,:) = (obj(1,:) + obj(3,:))/2;
objs(3,:) = (obj(1,:) + obj(4,:))/2;
objs(4,:) = mean(obj(2:3,:));
objs(5,:) = (obj(2,:) + obj(4,:))/2;
objs(6,:) = mean(obj(3:4,:));
[~, idx] = sort(objs,2,'descend');
[~, ids] = sort(mean(obj), 'descend');

%****************
%Individual Objective Elitism
newgenes(:,1:2) = genes(:,ix(1,1:2));
newgenes(:,3:4) = genes(:,ix(2,1:2));
newgenes(:,5:6) = genes(:,ix(3,1:2));
newgenes(:,7:8) = genes(:,ix(4,1:2));

%Aggregate Elitism - added for runs 7+
newgenes(:,9) = genes(:,idx(1,1));
newgenes(:,10) = genes(:,idx(2,1));
newgenes(:,11) = genes(:,idx(3,1));
newgenes(:,12) = genes(:,idx(4,1));
newgenes(:,13) = genes(:,idx(5,1));
newgenes(:,14) = genes(:,idx(6,1));
newgenes(:,15:16) = genes(:,ids(1,1:2));

%Mate top candidates with mutation rate
for n = 1:2:16
    [newgenes(:,n+16), newgenes(:,n+17)] = ...
        binary_crossover(newgenes(:,n), newgenes(:,n+1));
```

108
%FITNESS SHARING: Mate top 4 candidates from a given objective ...
with top
%candidates from other objectives
for n = 1:6:24
    [newgenes(:,n+32), newgenes(:,n+33)] = ...
        binary_crossover(newgenes(:,n), newgenes(:,n+2));
    [newgenes(:,n+34), newgenes(:,n+35)] = ...
        binary_crossover(newgenes(:,n), newgenes(:,n+4));
    [newgenes(:,n+36), newgenes(:,n+37)] = ...
        binary_crossover(newgenes(:,n), newgenes(:,n+6));
end

%generate 4 random genes
tempgene = zeros(18,4,'uint8');
parfor n = 1:4
    tempgene(:,n) = genrandgene;
end

%RANDOM GENE INSERTION: mate each top candidate with random genes
[newgenes(:,57), newgenes(:,58)] = ...
    binary_crossover(newgenes(:,1), tempgene(:,1));
[newgenes(:,59), newgenes(:,60)] = ...
    binary_crossover(newgenes(:,3), tempgene(:,2));
[newgenes(:,61), newgenes(:,62)] = ...
    binary_crossover(newgenes(:,5), tempgene(:,3));
[newgenes(:,63), newgenes(:,64)] = ...
    binary_crossover(newgenes(:,7), tempgene(:,4));
newgenes = proxcheck(newgenes);
end

A.2 Subroutine Code

The supporting code routines are listed in this section. These codes are supporting
the main program flow and are often called by multiple main routine functions.
A.2.1 createcstruct.m

The createcstruct.m program is responsible for creating and initializing the candidate structure data for a given run.

```matlab
function [ candidates ] = createcstruct( pop, gen )
%UNTITLED2 Summary of this function goes here
% Detailed explanation goes here
candidates(1:pop,1)=struct;
dt = datestr(now,'mmmdd_HHMM');
gens = num2str(gen, '%03d');
for m = 1:pop
    spop = num2str(m, '%03d');
candidates(m).filen = strcat('1DQCL_vars.',gens,'.',spop,'.',dt);
candidates(m).genes = zeros(18,1,'uint8');
candidates(m).genestr = cell(18,1);
candidates(m).ar = zeros(7,1,'uint16');
candidates(m).inj = zeros(2,1,'uint16');
candidates(m).spectrum = zeros(20,1);
candidates(m).psi2sh = zeros(758,21);
candidates(m).psi2 = zeros(758,21);
candidates(m).cb = zeros(784,2);
candidates(m).IB = cell(495,1);
candidates(m).taui = zeros(20,1);
candidates(m).tau32 = 0;
candidates(m).osc32 = 0;
candidates(m).dme = zeros(20);
candidates(m).lo = 0;
candidates(m).fitness = zeros(4,2);
candidates(m).idx = zeros(5,1);
end
end
```

A.2.2 calcgenes.m

The calcgenes.m program converts the starting population whose lengths are given in angstroms to a gene sequence which is given in 1/2 monolayer units.

```matlab
function [ genes ] = calcgenes( initpop )
%calcgenes: Calculates the gene sequence from initial population
```
Detailed explanation goes here

```matlab
x = initpop(1)/100;
y = initpop(2)/100;
ag = 5.65325; % lattice constant of GaAs @300K in Angstroms
ai = 6.0583; % lattice constant of InAs @300K
aa = 5.6611; % lattice constant of AlAs @300K
ab = ag + (aa-ag)*x; % lattice constant of barriers - Vegards Law
aw = ag + (ai-ag)*y; % lattice constant of active region wells
```

```matlab
genes = zeros(18,1,'uint8');
genesis(1:2) = initpop(1:2);
for n = 3:18
    if n<8
        genes(n) = 4*initpop(n)/ab;
    elseif n<13
        genes(n) = 4*initpop(n)/ag;
    elseif n<16
        genes(n) = 4*initpop(n)/ab;
    else
        genes(n) = 4*initpop(n)/aw;
    end
end
```

A.2.3 convgenes.m

The convgenes.m program converts the gene sequence to a string array for use in creating the input file for the simulation program.

```matlab
function [ genestr ] = convgenes( genes )
% convgenes: Converts integer values of genes into array of strings
% Detailed explanation goes here

% get alloy constants
x = double(genes(1))/100; % barriers
y = double(genes(2))/100; % active region wells
ag = .565325; % lattice constant of GaAs @300K
ai = .60583; % lattice constant of InAs @300K
aa = .56611; % lattice constant of AlAs @300K
ab = ag + (aa-ag)*x; % lattice constant of barriers - Vegards Law
aw = ag + (ai-ag)*y; % lattice constant of active region wells
```

111
\% calculate the thickness of each layer in Angstroms
\% each monolayer constitutes 1/4 of lattice constant
\% strings need to be in units of [nm] where as genes are in ML
l = length(genes);
genestr = cell(l,1);
for n = 1:l
  if n < 3
    newv = double(genes(n))/100;
    genestr\{n\} = num2str(newv, '\%-4.2f');
  elseif n<8
    newv = ab*double(genes(n))/4;
    genestr\{n\} = num2str(newv, '\%-4.1f');
  elseif n<13
    newv = ag*double(genes(n))/4;
    genestr\{n\} = num2str(newv, '\%-4.1f');
  elseif n<16
    newv = ab*double(genes(n))/4;
    genestr\{n\} = num2str(newv, '\%-4.1f');
  else
    newv = aw*double(genes(n))/4;
    genestr\{n\} = num2str(newv, '\%-4.1f');
  end
end
end

A.2.4 genes2ar.m

The genes2ar.m program calculates the active region bounds from the gene string sequence produced by convgenes.m.

function [ ar, inj ] = genes2ar( genestr )
% genes2ar: Calculates the active region limits, including each ...
% well left
% and right values.
% Detailed explanation goes here
thicknesses = 10*str2double(genestr(3:18));
ar = zeros(7,1,'uint16');
inj = zeros(2,1,'uint16');
11 \texttt{ar(1) = sum(thicknesses(1:10))+thicknesses(13); \%ARW1 Left} \\
12 \texttt{ar(2) = ar(1)+thicknesses(14); \%ARW1 Right} \\
13 \texttt{ar(3) = ar(2)+thicknesses(11); \%ARW2 Left} \\
14 \texttt{ar(4) = ar(3)+thicknesses(15); \%ARW2 Right} \\
15 \texttt{ar(5) = ar(4)+thicknesses(12); \%ARW3 Left} \\
16 \texttt{ar(6) = ar(5)+thicknesses(16); \%ARW3 Right} \\
17 \texttt{ar(7) = ar(6)+thicknesses(13); \%ARB3 Right} \\
18 \texttt{inj(2) = ar(1)-thicknesses(5); \%INJW5 Right} \\
19 \texttt{inj(1) = inj(2)-thicknesses(10); \%INJW5 Left} \\
20 \texttt{end}

\textit{A.2.5 \texttt{binary\textunderscore crossover.m}}

The \texttt{binary\textunderscore crossover.m} program performs real valued binary crossover on a set of parent and produces two child candidates.

\begin{verbatim}
function [ child1, child2 ] = binary\textunderscore crossover( parent1, parent2 )
%UNTITLED2 Summary of this function goes here
% Detailed explanation goes here

numbits = length(parent1);
child1 = zeros(numbits,1);
child2 = zeros(numbits,1);
xmask = rand(numbits,1)>.5;

parfor n = 1:numbits
    if xmask(n)
        child1(n) = parent1(n);
    else
        child1(n) = parent2(n);
    end
end

parfor n = 1:numbits
    if ~xmask(n)
        child2(n) = parent1(n);
    else
        child2(n) = parent2(n);
    end
end
\end{verbatim}
A.2.6 \textit{wmutate.m}

This code performs real valued mutation on a given candidate by using a random gene sequence and a mutation percentage.

\begin{verbatim}
function [ newgenes ] = wmutate( genes )
%wmutate: Mutates the alloy percentages or the width of a well or ...
    %barrier
%within the specified min/max using uniform mutation on an integer
mut = 0.02; %mutation probability
mutmask = rand(length(genes),1)<mut;
mutval = genrandgene;
newgenes = genes;
newgenes(mutmask) = mutval(mutmask);
end
\end{verbatim}

A.2.7 \textit{genrandgene.m}

Calculates a random gene sequence given the data format.

\begin{verbatim}
function [ gene ] = genrandgene( )
%genrandgene: generates a random gene based on the limits of each ...
    %gene
%value
gene = zeros(18,1,'uint8');
a = 10; %minimum barrier alloy percentage
\end{verbatim}
b = 45; %maximum barrier alloy percentage
gene(1) = a + (b-a)*rand(1);

a = 0; %minimum well alloy percentage
b = 25; %maximum well alloy percentage
gene(2) = a + (b-a)*rand(1);

a = 1; %minimum # of monolayers
b = 96; %maximum # of monolayers ~13nm
gene(3:18) = a + (b-a).*rand(16,1);

end

A.2.8 proxcheck.m

This routine checks to see if a candidate solution is identical any other candidate and
mutates the candidate until it is no longer identical.

function [ genes ] = proxcheck( genes )
%proxcheck: mutates genes that are identical until all genes are ... different.

for m = 1:63
  for n = (m+1):64
    while all(genes(:,m) == genes(:,n))
      genes(:,n) = wmutate(genes(:,n));
    end
  end
end

end

A.2.9 findlevels.m

This program finds the relevant energy levels in the injector and active regions. This
is accomplished by finding the peak value locations of the psi^2 data.
function [ QCL, retcode ] = findlevels( QCL )
%findlevels: Determines the 4 pertenent levels of the laser ... 
active region
%(if available)
% Detailed explanation goes here

%set index range for active region left (1) and right (2)
ir1 = (QCL.ar(1)+11):1:(QCL.ar(1)+14);
ir2 = (QCL.ar(7)+17):1:(QCL.ar(7)+20);

%determine conduction band maximums
cbmaxl = max(QCL.cb(ir1,2)); %just before ARW1L
cbmaxr = max(QCL.cb(ir2,2)); %just before ARB3R

%determine the conduction band minimas
cbminl = min(QCL.cb(ir1,2)); %left side of AR
cbminr = min(QCL.cb(ir2,2)); %right side of AR

 [~,I] = max(QCL.psi2(:,2:21)); %get indices of ... maxima
I = zeros(5,1);
[~,I(5)] = max(sum(QCL.psi2(QCL.inj(1):QCL.inj(2),2:21)));
[~,I(1)] = max(sum(QCL.psi2(QCL.ar(1):QCL.ar(2),2:21)));
[~,I(2)] = max(sum(QCL.psi2(QCL.ar(3):QCL.ar(4),2:21)));
[~,I(3)] = max(sum(QCL.psi2(QCL.ar(5):QCL.ar(6),2:21)));

if I(2) == I(3)
temp = sum(QCL.psi2(QCL.ar(3):QCL.ar(6),2:21));
[~,I(2)] = max(temp);
temp(I(2)) = 0;
[~,I(3)] = max(temp);
if I(3) > I(2)
Itmp = I(2);
I(2) = I(3);
I(3) = Itmp;
end
end
QCL.idx = I;

if ((QCL.idx(1) > QCL.idx(5)) && (QCL.idx(5) > QCL.idx(2)) && ...
(QCL.idx(2) > QCL.idx(3)))
    retcode = 0;
else
    retcode = -1;
end
end
A.2.10  \textit{obj1fit.m}

This program is called by the find fitness routine for objective 1. It calculates the value using the piecewise function described in Chapter 6.

```matlab
function [ fit1 ] = obj1fit( x )
%UNTITLED Summary of this function goes here
% Detailed explanation goes here

if (.248 < x) && (x <= .413)
    fit1 = 1;
elseif x < .248
    fit1 = normpdf(x,.248,.03)/normpdf(.248,.248,.03);
else
    fit1 = normpdf(x,.413,.03)/normpdf(.413,.413,.03);
end
end
```

A.2.11  \textit{sigmf.m}

Calculates the sigmoid function which produces a value from 0 to 1 for a given input.

```matlab
function [ y ] = sigmf( x, lims )
%UNTITLED Summary of this function goes here
% Detailed explanation goes here

a = lims(1);
c = lims(2);

y = 1 ./ (1 + exp(-a*(x-c)));
end
```

A.3  Template File

This section contain the template input file used for running the nextnano\textsuperscript{3} software. This file is modified by changing the variables at the beginning of the program listing.
%FunctionParser = yes

****** Start of design parameters ******!

%AlloyContent_barrier = 0.45
%AlloyContent_well = 0.06
%ThicknessInjBarrier_1 = 1.7
%ThicknessInjBarrier_2 = 1.8
%ThicknessInjBarrier_3 = 2.0
%ThicknessInjBarrier_4 = 2.6
%ThicknessInjBarrier_5 = 4.6
%ThicknessInjWell_1 = 3.4
%ThicknessInjWell_2 = 3.0
%ThicknessInjWell_3 = 2.8
%ThicknessInjWell_4 = 3.0
%ThicknessInjWell_5 = 3.0
%ThicknessARBarrier_1 = 1.1
%ThicknessARBarrier_2 = 1.1
%ThicknessARBarrier_3 = 2.8
%ThicknessARWell_1 = 1.9
%ThicknessARWell_2 = 5.4
%ThicknessARWell_3 = 4.8

****** End of design parameters ******!

--------------------------------------!
Start of calculated values
--------------------------------------!

****** Active Region *****************
%ARWell1_L = 0 + %ThicknessInjBarrier_5
%ARWell1_R = %ARWell1_L + %ThicknessARWell_1
%ARWell2_L = %ARWell1_R + %ThicknessARBarrier_1
%ARWell2_R = %ARWell2_L + %ThicknessARWell_2
%ARWell3_L = %ARWell2_R + %ThicknessARBarrier_2
%ARWell3_R = %ARWell3_L + %ThicknessARWell_3

****** Injector Right ****************
%InjRWell1_L = %ARWell3_R + %ThicknessARBarrier_3
%InjRWell1_R = %InjRWell1_L + %ThicknessInjWell_1
%InjRWell2_L = %InjRWell1_R + %ThicknessInjBarrier_1
%InjRWell2_R = %InjRWell2_L + %ThicknessInjWell_2
%InjRWell3_L = %InjRWell2_R + %ThicknessInjBarrier_2
%InjRWell3_R = %InjRWell3_L + %ThicknessInjWell_3
%InjRWell4_L = %InjRWell3_R + %ThicknessInjBarrier_3
%InjRWell4_R = %InjRWell4_L + %ThicknessInjWell_4
%InjRWell5_L = %InjRWell4_R + %ThicknessInjBarrier_4
%InjRWell5_R = %InjRWell5_L + %ThicknessInjWell_5

!****** Injector Left - REVERSE ORDER **!
%InjWell5_R = 0
%InjWell5_L = %InjWell5_R - %ThicknessInjWell_5
%InjWell4_R = %InjWell5_L - %ThicknessInjBarrier_4
%InjWell4_L = %InjWell4_R - %ThicknessInjWell_4
%InjWell3_R = %InjWell4_L - %ThicknessInjBarrier_3
%InjWell3_L = %InjWell3_R - %ThicknessInjWell_3
%InjWell2_R = %InjWell3_L - %ThicknessInjBarrier_2
%InjWell2_L = %InjWell2_R - %ThicknessInjWell_2
%InjWell1_R = %InjWell2_L - %ThicknessInjBarrier_1
%InjWell1_L = %InjWell1_R - %ThicknessInjWell_1

!****** Total Region Values ************!
%MinPos = %InjWell1_L - %ThicknessARBarrier_3
%MaxPos = %InjRWell5_R + %ThicknessARBarrier_5
%ARLength = %InjRWell1_L - %ThicknessInjBarrier_5
%InjLength = %InjRWell5_R - %ARLength
%DeviceLength = ( 2 * %InjLength ) + %ARLength + ...
%ThicknessARBarrier_3
%QuantumRegionMin = %MinPos
%QuantumRegionMax = %MaxPos
%GridSpacing = 0.1
%nodes = ( %DeviceLength / %GridSpacing ) - 1
%INT(nodes) = %nodes

!---------------------------------------!

! End of calculated values
!---------------------------------------!

!---------------------------------------!

$numeric-control
  simulation-dimension = 1
  zero-potential = no
  varshni-parameters-on = yes
$end_numeric-control

!---------------------------------------!

!****** OVERALL SIMULATION PARAMETERS **!
!---------------------------------------!
$simulation-dimension
dimension = 1
orientation = 0 0 1
$end_simulation-dimension

$global-parameters
lattice-temperature = 300d0
$end_global-parameters

$electric-field
electric-field-on = yes
electric-field-strength = 48d5
electric-field-direction = 0 0 1
$end_electric-field

$simulation-flow-control
flow-scheme = 21
raw-directory-in = raw_data1/
raw-potential-in = no
strain-calculation = homogeneous-strain
$end_simulation-flow-control

$domain-coordinates
domain-type = 0 0 1
z-coordinates = %MinPos %MaxPos
growth-coordinate-axis = 1 0 0
pseudomorphic-on = GaAs
hkl-x-direction-zb = 0 1 -1
hkl-z-direction-zb = 1 1 1
$end_domain-coordinates

****** END OVERALL SIMULATION PARAMETERS **!

****** REGIONS AND CLUSTERS ***************
$regions
!---------
! sequence 1
!---------
region-number = 1  base-geometry = line  region-priority = 1
z-coordinates = 0d0 %ARWell1_L
region-number = 2  base-geometry = line  region-priority = 1
z-coordinates = %ARWell1_L %ARWell1_R
region-number = 3  base-geometry = line  region-priority = 1
z-coordinates = %ARWell1_R %ARWell2_L
region-number = 4  base-geometry = line  region-priority = 1
z-coordinates = %ARWell2_L %ARWell2_R
region-number = 5  base-geometry = line  region-priority = 1
z-coordinates = %ARWell2_R %ARWell3_L
region-number = 6  base-geometry = line  region-priority = 1
z-coordinates = %ARWell3_L %ARWell3_R
region-number = 7  base-geometry = line  region-priority = 1
z-coordinates = %ARWell3_R %InjRWell1_L
region-number = 8  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell1_L %InjRWell1_R
region-number = 9  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell1_R %InjRWell2_L
region-number = 10  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell2_L %InjRWell2_R
region-number = 11  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell2_R %InjRWell3_L
region-number = 12  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell3_L %InjRWell3_R
region-number = 13  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell3_R %InjRWell4_L
region-number = 14  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell4_L %InjRWell4_R
region-number = 15  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell4_R %InjRWell5_L
region-number = 16  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell5_L %InjRWell5_R
region-number = 17  base-geometry = line  region-priority = 1
z-coordinates = %InjWell5_L %InjWell5_R
region-number = 18  base-geometry = line  region-priority = 1
z-coordinates = %InjWell4_L %InjWell4_R
region-number = 19  base-geometry = line  region-priority = 1
z-coordinates = %InjWell4_R %InjWell5_L
region-number = 20  base-geometry = line  region-priority = 1
z-coordinates = %InjWell5_L %InjWell5_R
region-number = 21  base-geometry = line  region-priority = 1
z-coordinates = %InjWell5_R %InjWell4_L
region-number = 22  base-geometry = line  region-priority = 1
z-coordinates = %InjWell4_L %InjWell5_L
region-number = 23  base-geometry = line  region-priority = 1
z-coordinates = %InjWell2_L %InjWell2_R
region-number = 24  base-geometry = line  region-priority = 1
z-coordinates = %InjWell1_R %InjWell2_L
region-number = 25  base-geometry = line  region-priority = 1
z-coordinates = %InjWell1_L %InjWell1_R
region-number = 26  base-geometry = line  region-priority = 1
z-coordinates = %MinPos %InjWell1_L

!------------
! sequence 2
!------------
region-number = 27  base-geometry = line  region-priority = 1
z-coordinates = %InjRWell5_R %MaxPos
$end

!-----------------------------------------------!
!-----------------------------------------------!
$grid-specification
grid-type       = 0 0 1
z-grid-lines    = %MinPos %MaxPos
z-nodes         = %INT(nodes)
z-grid-factors  = 1d0
$end_grid-specification

!-----------------------------------------------!
$region-cluster

!------------
! sequence 1
!------------
cluster-number = 1  region-numbers = 1
cluster-number = 2  region-numbers = 2
cluster-number = 3  region-numbers = 3
cluster-number = 4  region-numbers = 4
cluster-number = 5  region-numbers = 5
cluster-number = 6  region-numbers = 6
cluster-number = 7  region-numbers = 7
cluster-number = 8  region-numbers = 8
cluster-number = 9  region-numbers = 9
cluster-number = 10 region-numbers = 10
cluster-number = 11 region-numbers = 11
cluster-number = 12 region-numbers = 12
cluster-number = 13 region-numbers = 13
cluster-number = 14 region-numbers = 14
cluster-number = 15 region-numbers = 15
cluster-number = 16 region-numbers = 16
!------------------------------------------------
! sequence left of sequence 1 (in reverse order)
!------------------------------------------------
cluster-number = 17  region-numbers = 17
cluster-number = 18  region-numbers = 18
cluster-number = 19  region-numbers = 19
cluster-number = 20  region-numbers = 20
cluster-number = 21  region-numbers = 21
cluster-number = 22  region-numbers = 22
cluster-number = 23  region-numbers = 23
cluster-number = 24  region-numbers = 24
cluster-number = 25  region-numbers = 25
cluster-number = 26  region-numbers = 26

!------------
! sequence 2
!------------
cluster-number = 27  region-numbers = 27 28
$end_region-cluster

!****** END REGIONS AND CLUSTER ***************

!****** MATERIALS AND ALLOY PROFILES **********
!----------------------------------------------------------------!
$material
material-number = 1
material-name = Al(x)Ga(1-x)As
cluster-numbers = 1 3 5 7 9 11 13 15 18 20 22 24 26 27
alloy-function = constant

material-number = 2
material-name = GaAs
cluster-numbers = 8 10 12 14 16 17 19 21 23 25

material-number = 3
material-name = In(x)Ga(1-x)As
cluster-numbers = 2 4 6
alloy-function = constant
$end_material

!----------------------------------------------------------------!
$alloy-function
material-number = 1
function-name = constant
xalloy = 0.45d0
material-number = 3 !

function-name = constant

xalloy = %AlloyContent_well !

$end_alloy-function
!

!----------------------------------------------------------------!

293

!----------------------------------------------------------------!

$binary-zb-default

binary-type = GaAs-zb-default !

apply-to-material-numbers = 1 2

conduction-band-masses = 0.067d0 0.067d0 0.067d0 !

1.9d0 0.0754d0 0.0754d0 !

1.3d0 0.23d0 0.23d0 !

conduction-band-energies = 2.979d0 3.275d0 3.441d0 !

$end_binary-zb-default
!

!----------------------------------------------------------------!

$binary-zb-default

binary-type = AlAs-zb-default !

apply-to-material-numbers = 1

conduction-band-masses = 0.15d0 0.15d0 0.15d0 !

1.32d0 0.15d0 0.15d0 !

0.97d0 0.22d0 0.22d0 !

conduction-band-energies = 4.049d0 3.410d0 3.190d0 !

$end_binary-zb-default
!

!----------------------------------------------------------------!

$ternary-zb-default

ternary-type = Al(x)Ga(1-x)As-zb-default !

apply-to-material-numbers = 1 !

binary(x) = AlAs-zb-default !

binary(1-x) = GaAs-zb-default !

bow-conduction-band-energies = 0.4625d0 0d0 0.055d0 !

band-shift = 0.022719d0 !

$end_ternary-zb-default
!

!----------------------------------------------------------------!

!***** END MATERIALS AND ALLOY PROFILES ************************! 

!***** QUANTUM ****************************************************

$quantum-regions

region-number = 1 !

base-geometry = line !

region-priority = 1 !

z-coordinates = %QuantumRegionMin %QuantumRegionMax !

$end_quantum-regions
$quantum-cluster
  cluster-number = 1
  region-numbers = 1
  deactivate-cluster = no
$end_quantum-cluster

$quantum-model-electrons
  model-number = 1
  model-name = effective-mass
  cluster-numbers = 1
  conduction-band-numbers = 1
  number-of-eigenvalues-per-band = 20
  boundary-condition-001 = Dirichlet
$end_quantum-model-electrons

$output-1-band-schroedinger
  destination-directory = sg
  sg-structure = yes
  scale = 1d0
  conduction-band-numbers = 1
  cb-min-ev = 1
  cb-max-ev = 20
  intraband-matrix-elements = p
  intraband-lifetime = yes
$output-bandstructure
  destination-directory = band
  conduction-band-numbers = 1
  band-gap = no
  potential = no
  electric-field = no
$output-1-band-schroedinger
$end_output-bandstructure
$output-material

destination-directory = material_parameters/

LO-phonon-energy = LO_phonon_energy.dat

$end_output-material

!****** END OUTPUT **********************************************!
Bibliography


VITA

David Mueller was born on the 5th of November of 1981 to Gary and Laura Mueller in Columbia, Missouri. He is the oldest child and has a younger brother Daniel Mueller and a half-sister Brandi Mueller. David Mueller was the first in his immediate family to complete a college degree. Each of his academic achievements have been completed at the University of Missouri at Columbia. He received his Bachelor of Science degrees in Electrical & Computer Engineering in December of 2006. He then pursued and achieved his Master of Science degree in Electrical Engineering in August of 2008 under the advisement of Dr. Thomas Engel. He began doctoral research with his dissertation advisor, Dr. Gregory Triplett, in August of 2010. David successfully defended his dissertation on April 30, 2015. Along with his dissertation work, David pursued a graduate minor in College Teaching and completed the requirements in August 2014. He was inspired to complete this minor, in part, because of seven semesters in which he was awarded the department of Electrical and Computer Engineering’s teaching fellowship. This fellowship allowed him to gain valuable experience as instructor of record for several fundamental engineering courses.

David married the love of his life, Shelby (Connor) Mueller, on September 6, 2014 in Nashville, Tennessee. After graduation with his Ph.D., David will be working in academia as an Assistant Professor of Electrical Engineering at the University of the Pacific equipping young minds for their future careers in engineering.