Thermal and Waveguide Optimization of Broad Area Quantum Cascade Laser Performance

Matthew Suttinger
University of Central Florida

Part of the Electromagnetics and Photonics Commons, and the Optics Commons

Find similar works at: https://stars.library.ucf.edu/etd

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
THERMAL AND WAVEGUIDE OPTIMIZATION OF BROAD AREA QUANTUM CASCADE LASER PERFORMANCE

by
MATTHEW MICHAEL SUTTINGER
B.S. University of Central Florida, 2013

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the College of Optics & Photonics at the University of Central Florida Orlando, Florida

Fall Term
2017

Major Professor: Arkadiy Lyakh
Quantum Cascade Lasers are a novel source of coherent infrared light, unique in their tunability over the mid-infrared and terahertz range of frequencies. Advances in bandgap engineering and semiconductor processing techniques in recent years have led to the development of highly efficient quantum cascade lasers capable of room temperature operation. Recent work has demonstrated power scaling with broad area quantum cascade lasers by increasing active region width beyond the standard ~10 μm. Taking into account thermal effects caused by driving a device with electrical power, an experimentally fitted model is developed to predict the optical power output in both pulsed and continuous operation with varying device geometry and minor changes to quantum cascade laser active region design. The effects of the characteristic temperatures of threshold current density and slope efficiency, active region geometry, and doping, on output power are studied in the model. The model is then used to refine the active region design for increased power out in continuous operation for a broad area design. Upon testing the new design, new thermal effects on rollover current density are observed. The model is then refined to reflect the new findings and more accurately predict output power characteristics.
This work is dedicated to my wife, Kehan Yu, who is likely the only person outside of my committee who would ever bother to read it in its entirety.
ACKNOWLEDGMENTS

This work was accomplished with the help of Arkadiy Lyakh, the principal investigator of the projects that this work covers, the designer of the quantum well superlattices that are used, and a provider of guidance and advise throughout the process of this work. Ankesh Todi and his work in his master’s thesis provided initial data on power scaling with quantum cascade laser active region ridge width. Processing efforts by Pedro Figueiredo, Rowel Go, and Jason Leshin provided the devices that this thesis discusses. Justin Cook provided formatting help and guidance with the writing of this thesis.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ viii

LIST OF ACRONYMS/ABBREVIATIONS .................................................................... xiv

CHAPTER 1: INTRODUCTION ......................................................................................... 1

CHAPTER 2: OVERVIEW OF QUANTUM CASCADE LASER THEORY ................... 3

  Section 1: Quantum Confinement ........................................................................... 3
  Section 2: Lasers ..................................................................................................... 6
  Section 3: Superlattice Design and Characteristics ................................................ 8
  Section 4: Waveguides ............................................................................................ 10
  Section 5: Pulsed Characteristics .......................................................................... 14
  Section 6: I/L Measurements .................................................................................. 17
  Section 7: Thermal Effects ..................................................................................... 18
  Section 8: Continuous Wave Operation .................................................................. 19

CHAPTER 3: INITIAL DESIGN MEASUREMENTS .................................................. 20

  Section 1: Quantum Cascade Laser Growth and Processing ............................... 20
  Section 2: Device Measurements ......................................................................... 21

CHAPTER 4: THE SEMIEMPIRICAL MODEL .............................................................. 25

  Section 1: Waveguide Simulations ....................................................................... 25
  Section 2: Pulsed Power ....................................................................................... 26
  Section 3: Continuous Wave Threshold .................................................................. 27
  Section 4: Continuous Wave Power ..................................................................... 28
LIST OF FIGURES

Figure 1: A single, finite, 1D quantum well with two confined states. The vertical axis represents energy level (potential energy in black, eigenenergy for the state as a dashed line.) The horizontal axis is position along the growth direction, z. Wavefunctions for each state are plotted as normalized curves anchored with 0 probability density at their energy levels (dashed lines.) The first (red) and second (green) order modes are the two confined states allowed by this quantum well.......................................................... 4

Figure 2: A double quantum well potential. First and second order states for the coupled wells displayed as wavefunctions anchored at state energy (shown as dashed lines) for 0 probability density. The lower modes in red and green show the coupled first order modes for the two interacting wells, with energy levels separated by a small amount. The upper modes in red and blue are coupled second order modes, slightly more separated in energy............................................. 5

Figure 3: Illustration of the simplified two level model for a quantum cascade. The potential, shown in black, is a series of quantum wells, coupled by proximity. The slope is introduced as a voltage bias contributing to potential energy for the electron experiencing the potential. Representations of wavefunction locations (horizontal position) and energy levels (vertical position) are shown as the red and blue upper and lower energy levels. Lasing is modelled as stimulated emission from an upper to lower laser level within a well. Electrons tunnel through barriers from lower levels in one well to the upper level in the next well with lower potential energy due to the voltage bias.................................................................................................. 7

Figure 4: The active region design used for the first iteration of devices tested in this work. The potential energy structure making up the quantum well superlattice is shown as the thin black
lined background pattern. The superlattice is a series of Aluminum Indium Arsenide (barriers) and Indium Gallium Arsenide (wells) layers biased by a voltage to introduce the slope in the potential profile. Curves represent wavefunctions generated by this potential. Wavefunctions are anchored at their state energy, with the flat regions representing areas of zero probability density, while high amplitude portions of the wavefunctions show locations of high probability density for electron location. Key states are shown as more thickly drawn curves. Upper laser level, 4, and lower laser level, 3, in red. State g is the injector state leading from the previous period. State 2 is a state of particularly long lifetime, affecting lower level population in state 3. State 5 is a parasitic state where electrons may transition without the probability of lasing, generating heat and reducing efficiency.

Figure 5: Schematic of the waveguide used for the devices in this work. The Active Region is composed of a superlattice of Indium Gallium Arsenide and Aluminum Indium Arsenide. Cladding layers are composed of Indium Phosphide doped with Silicon to carry charge and modify index. Pure Indium Phosphide sidewalls confine current and optical mode to a fixed width.

Figure 6: Fundamental TM mode profile of a representative device, taken from COMSOL calculation. Electric field magnitude is plotted, with red representing high field magnitude, and blue signifying little or no field magnitude. The single lobe structure indicates that it is a fundamental mode. The majority of the intensity exists evanescently in the cladding, with roughly 20% existing the active region.

Figure 7: Pulsed L-I-V for a representative narrow ridge device use in this work. The blue curve is voltage as a function of current. Voltage demonstrates a diode-like turn on voltage, followed by a linear regime over the operational range of the device, and finally a runaway effect at rollover current. The red curve is pulsed peak power as a function of current. Power
demonstrates linear behavior after threshold current, developing sublinearity and rolling over at rollover current.

Figure 8: 1/L measurement for first iteration of active region design. Data in red shows the linear relationship between the reciprocal of slope efficiency and cavity length. Data in blue shows the linear relationship between inverse cavity length and threshold current density.

Figure 9: Characteristic temperature measurement for first iteration device. Data in red shows the exponential growth of threshold current density with increasing active region temperature. Data in blue shows the exponential decay of slope efficiency with increasing active region temperature.

Figure 10: Pulsed L-I characteristics for three devices of varying ridge widths. Peak pulsed power is plotted against driving pulsed current. The 10.4 μm (blue,) 21.2 μm (purple,) and 31.0 μm (red) devices show standard linear threshold and sublinear rollover behavior. Rollover current density is fixed, so rollover current travels as a function of device width. Rollover current for the 31.0 μm device is predicted to be ~11 A, beyond the maximum capability of the driver used to gather the data.

Figure 11: CW L-I characteristics for three devices of varying ridge widths. CW power is plotted against driving current. The 10.4 μm (blue,) 21.2 μm (purple,) and 31.0 μm (red) devices show a general curvature throughout their range, and experience a thermal rollover before the rollover current shown in pulsed operation.

Figure 12: Threshold current and temperature as a function of CW driving current for a hypothetical narrow ridge device with threshold current density characteristic temperature $T_0 = 75 \, K$. Temperature (red) increases as a smooth curve with driving current. Threshold current (blue) is calculated as a function of temperature, itself a function of driving current. Driving current is
represented as a dashed line of slope 1 to indicate where threshold current becomes equal to its corresponding driving current. A device can operate in CW in the driving current range between the two interception points of the driving current line and the threshold current curve, provided it does not first reach rollover current.

Figure 13: Comparison of projected and measured peak pulsed power as a function of pulsed driving current for first iteration devices. 10.4 μm (blue,) 21.2 μm (purple,) and 31.0 μm (red) device peak pulsed power are shown as data points for measured results and solid and dashed lines for projected results in the model.

Figure 14: Comparison of projected and measured CW power for first iteration devices. 10.4 μm (blue,) 21.2 μm (purple,) and 31.0 μm (red) device CW power are shown as data points for measured results and solid and dashed lines for projected results in the model. Spikes in power are due to multi-mode contributions (mode-hopping) beyond the scope of this work.

Figure 15: Projected pulsed and CW power of 21.2 μm device (blue,) and the optimized device (red.) The optimized device is a 21 stage device with an active region 25 μm wide with half the active region doping of the 21.2 μm design. The lower active region doping results in a compressed dynamic range (earlier threshold and rollover currents) with higher slope efficiency for a net improvement in CW power.

Figure 16: Surface of maximum CW power produced by given characteristic temperatures in the 31.0 μm device. Characteristic temperatures for threshold current density (T0) and slope efficiency (T1) are varied from 100 K to 1000 K. Purple indicates no power, with dark red indicating highest power. Particular sensitivity to T0 is shown as the steep gradient in the direction of increasing threshold current characteristic temperature.
Figure 17: Temperature distribution at thermal rollover for 31 μm design without (left) and with (right) additional 300 nm indium gallium arsenide spacers added on both the substrate and epitaxy side of the active region. Average temperature is shown to increase in the case of additional indium gallium arsenide spacers as a darker red throughout the active region. However, the difference is small, relative to the increased confinement, leading to increased CW power. ...... 35

Figure 18: 1D cross section of predicted maximum power in the 6D sweep. Shown are the power distributions with variable relative doping (top left,) stage thickness (top right,) number of stages (bottom left,) and ridge width (bottom right) around the optimal configuration. All plots encompass the optimal design of a 40 μm wide, half doped active region with 16 stages. A trend of increased CW power with decreasing stage thickness/height is observed, showing that a shorter injector, other parameters being equal, is a preferred design. .............................................. 36

Figure 19: Simulated and projected power for a 8.5 μm x 3 mm device with second iteration active region design. Peak pulsed power is blue while CW power is red. Measured data is shown as points while projections from the model are shown as solid lines. .............................................. 39

Figure 20: Rollover current density as a function of temperature for devices of variable ridge width. Three devices of different widths were tested; 8.5 μm (black,) 21.2 μm (red,) and 31.0 μm (blue.) The 8.5 μm device has anomalously high rollover current. The 21.2 μm and 31.0 μm devices both show a similar linear trend, with decreasing rollover current density with increasing temperature. .................................................................................................................. 41

Figure 21: Envelope (measured peak pulsed power divided by linear projected power) as calculated from four representative devices, and a best fitting curve. The value x is the parametrization from threshold (x = 0) and rollover (x = 1.0) current. An envelope value of 1.0
indicates direct correspondence to linear projected power. The envelope decreases near rollover current and rapidly decays at and after rollover current. .......................................................... 42

Figure 22: Projected power for two 21.2 µm wide devices of second iteration design. Peak pulsed power is blue while CW power is red. Measured data is shown as points while projections from the model are shown as solid lines. Due to the inclusion of new temperature effects, increased sensitivity to temperature projections and width measurements cause CW projections to properly approximated overall power with a shifted dynamic range. .......................................................... 44
LIST OF ACRONYMS/ABBREVIATIONS

1D – One Dimensional
2D – Two Dimensional
BH – Buried Heterostructure
CW – Continuous Wave
DC – Direct Current
HR – Highly Reflective
LASER – Light Amplification by Stimulated Emission of Radiation
L-I – Light-Current
L-I-V – Light-Current-Voltage
MBE – Molecular Beam Epitaxy
MCT – Mercury Cadmium Telluride
MOCVD – Metal Organic Chemical Vapor Deposition
QCL – Quantum Cascade Laser
$q_e$ – Electron Charge
$T_0$ – The Characteristic Temperature of Pulsed Slope Efficiency
$T_1$ – The Characteristic Temperature of Threshold Current Density
TE – Transverse Electric
TM – Transverse Magnetic
CHAPTER 1: INTRODUCTION

Since their first demonstration in 1994 [1], Quantum Cascade Lasers (QCLs) have rapidly developed into a technology of interest for scientific research, commercialization, and defense applications. Operating on the principle of quantum confinement of electron wavefunctions within the conduction band of a heterostructure, a quantum cascade laser can be tuned for output at mid-infrared and terahertz frequencies. In contrast, standard laser diodes have wavelengths largely determined by material bandgap, ranging from visible to near infrared [2]. Interband cascade lasers use the quantum confinement of electron wavefunctions within the conduction and valence bands (the key difference from QCLs) of the device to tune the output wavelength at mid-infrared frequencies [3].

The tunability of QCL frequencies over the mid-infrared spectral region is of particular interest to the field of spectroscopy, as many trace gases have rovibrational transitions in this range [4]. One such application being explored is the ranged detection of explosive material with optical spectroscopy [5]. Other applications of QCLs being explored include infrared countermeasures [6], frequency comb generation [7], and medical diagnostics [8].

With the increased demand for QCLs in these fields, it becomes increasingly necessary that devices can be created with high efficiency and output power. Devices have already been demonstrated with wall plug efficiency exceeding 28% in pulsed operation at multi-watt output [9]. Such designs have been demonstrated to be capable of scaling output power with device geometry in so called broad area quantum cascade lasers [10]. Other designs have successfully demonstrated continuous wave (CW) efficiencies up to 21% [11].
Indeed, high efficiency and high CW output power are highly correlated. A high degree of efficiency is an indication that a greater degree of electrical input power is contributing to output rather than heating the device. CW characteristics are largely determined by temperature, with higher temperatures causing performance degradation through a variety of effects described and studied in this thesis. The focus of this work is the optimization of a QCL device design by altering waveguide and temperature characteristics for the maximization of CW output power. Thus, a large component of this work is the development of a predictive model that can take in active region and waveguide design characteristics to predict output power.

In chapter 2, an overview of the theory of quantum cascade lasers and relevant waveguide physics is provided to form a basis for understanding design motivations and the modeling efforts. Chapter 3 describes the initial growth of QCL devices that were tested to demonstrate CW power scaling with increased active region width. The data generated by these devices are then used to create the model as described in chapter 4, with further discussion provided on potential directions for design optimization. Chapter 5 describes the second growth of QCL devices tested to optimize CW power, where additional thermal effects are evidenced and taken into account in the model. Finally, Chapter 6 provides conclusions to the results of the work to guide further iterations of QCL design for output power optimization.
CHAPTER 2: OVERVIEW OF QUANTUM CASCADE LASER THEORY

Unlike standard semiconductor lasers, which rely on the transition of electrons from the conduction band to the valence band to emit photons, QCLs operate entirely in the conduction band. Bandgap engineering is utilized to construct a superlattice of semiconductor materials to create a potential well structure. This well structure, under electrical bias, is the source of the laser light. The superlattice, often referred to as the active region, is grown on a semiconductor wafer as part of a waveguide. Once the device is ready for use, laser performance characteristics are determined by the properties of the superlattice and waveguide, with modifications due to temperature.

Section 1: Quantum Confinement

At the quantum mechanical scale, a region with lower potential energy than its surroundings can confine wavefunctions localized within the region at discrete energy levels. This is referred to as a quantum well. Wavefunctions describe a probability amplitude of a particle, for the purposes of this work, an electron, being found within a certain region upon observation. Wavefunctions are each associated with a single energy state. Electron transitions between energy levels occur when absorption or emission of energy change the electron energy level to exactly that of the new state, where the electron wavefunction then becomes that of the new state. The wavefunctions and associated energy levels are the eigenfunctions and eigenvalues of the time-independent Schrödinger equation:

\[ \hat{H} |\Psi(x, y, z)\rangle = \left[ -\frac{\hbar^2}{2m} \nabla^2 + V(x, y, z) \right] |\Psi(x, y, z)\rangle = E |\Psi(x, y, z)\rangle \]  

(1)
Transitions between states occur when the energy difference between states is bridged by the electron’s absorption or emission of energy to the exact amount of energy difference. In a QCL, this can be in the form of photons (the quantum of light,) phonons (the quantum of collective vibrational excitation in a periodic structure,) or scattering.

Figure 1: A single, finite, 1D quantum well with two confined states. The vertical axis represents energy level (potential energy in black, eigenenergy for the state as a dashed line.) The horizontal axis is position along the growth direction, $z$. Wavefunctions for each state are plotted as normalized curves anchored with 0 probability density at their energy levels (dashed lines.) The first (red) and second (green) order modes are the two confined states allowed by this quantum well.

A quantum well will feature states with wavefunctions that increase in the number of nodes with higher energy. A representative 1D quantum well is shown above. Transitions between states are mediated with some interaction Hamiltonian, a relevant example being the dipole interaction along the growth direction: $q_e \hat{z}$. With this example, the inner product around position operator $\hat{z}$ between initial and final states will determine the rate that the interaction described by the dipole interaction will result in a transition from the initial to final states as described by Fermi’s Golden Rule [12]:
\[ R_{i\rightarrow f} = \frac{\pi}{2\hbar} \langle \Psi_i | q_e \hat{z} | \Psi_f \rangle \rho(E_f - E_i \pm \hbar \omega) = \frac{\pi q_e Z_{if}}{2\hbar} \rho(E_f - E_i \pm \hbar \omega) \]  

(2)

Here, \( R_{i\rightarrow f} \) is the rate of transition from state \( i \) to state \( f \), \( \hbar \) is Planck’s reduced constant, \( q_e \) is the electron charge, \( \rho \) is a density of states, and \( \hbar \omega \) is an absorbed or emitted photon energy (denoted by \( \pm \), with + for emission and – for absorption.). This approach can describe both absorption and emission, and is used to calculate peak-gain cross-section by simple negation of the absorption. Assuming that the potential well is a perfect replication of design with no additional scattering effects, the density of states is taken to be a delta function, requiring perfect tuning of photon energy to state energy difference. Realistically, the density of states is a lorentzian profile for a device [12].

Figure 2: A double quantum well potential. First and second order states for the coupled wells displayed as wavefunctions anchored at state energy (shown as dashed lines) for 0 probability density. The lower modes in red and green show the coupled first order modes for the two interacting wells, with energy levels separated by a small amount. The upper modes in red and blue are coupled second order modes, slightly more separated in energy.

Two or more well in close proximity will couple, forming wave functions across the multiple well structure. Degenerate states appear as similar states at energy levels separated by a
small margin. This energy splitting becomes more severe with less separation between coupled wells, and is described with a coupling factor $\Omega$:

$$\Delta E = 2\hbar\Omega$$ (3)

This coupling factor will play an important role in determining electron transport characteristics in a biased multiple well system (the QCL superlattice.)

Section 2: Lasers

LASER is an acronym for light amplification by stimulated emission of radiation. Laser light is generated when an incoming photon causes a transition of an electron from an upper laser energy level to a lower laser level, resulting in the stimulated emission of an identical photon with identical phase. For this to happen, the photon must have an energy level equal to the difference of the upper and lower laser energy levels.

In a large population of electrons, the population of electrons in the upper laser level must be greater than that of the lower laser level for a net output of laser light to be statistically favorable. This is referred to as population inversion. Without population inversion, an incoming photon will likely be absorbed by an electron in a lower laser level, causing a transition to an upper laser level.

A standard approach for calculating population inversion and threshold condition for a laser is a rate equation model. Light interaction, state lifetimes, and scattering sources are taken into account to describe the rate of population change for each energy level. The system of differential equations is then solved for the upper and lower laser level populations. This approach is often insufficient to describe QCL populations, as the effects of tunneling and the coupling strength between adjacent wells separated by a barrier are not accounted for. A QCL operates with periodically repeating stages in the superlattice, where electrons transport from the first to last
stage through tunneling and scattering effects, and the current at which the electrons can cascade through the structure is an important figure in the modelling of QCL performance.

Figure 3: Illustration of the simplified two level model for a quantum cascade. The potential, shown in black, is a series of quantum wells, coupled by proximity. The slope is introduced as a voltage bias contributing to potential energy for the electron experiencing the potential. Representations of wavefunction locations (horizontal position) and energy levels (vertical position) are shown as the red and blue upper and lower energy levels. Lasing is modelled as stimulated emission from an upper to lower laser level within a well. Electrons tunnel through barriers from lower levels in one well to the upper level in the next well with lower potential energy due to the voltage bias.

A closer approximation of electron transport can be made with a density matrix approach. Approximating the superlattice as a two-level system that is periodically repeated, with the two levels repeating for each period (active region stage.) This is the approach used by Kazarinov and Suris to determine maximum current density [12, 13]:

---
\[ J_{\text{max}} = \frac{q_e n_s \Omega^2 \tau_\parallel}{1 + 4 \Omega^2 \tau_\parallel \tau_{\text{lower}}} \] (4)

Here, \( J_{\text{max}} \) is the maximum current density. \( n_s \) is the 2D sheet density of carriers in the superlattice. \( \Omega \) is the coupling factor between the lower laser level of one stage and the upper laser level of the next, as described in equation (3), at a bias where the two levels are as close as possible in energy level. \( \tau_\parallel \) is the coherence time for two states accounting for elastic and inelastic effects. \( \tau_{\text{lower}} \) is the lower level lifetime [12].

**Section 3: Superlattice Design and Characteristics**

The superlattice is a periodic heterostructure. Lower bandgap materials serve as quantum wells when adjacent to higher band gap materials, which serve as barriers. Each period, or stage, is composed of an active region and an injector region. The active region serves to localize laser levels within relatively wide quantum wells. The injector region transports electrons out of lower laser level toward the next stage of the superlattice to maintain population inversion. The devices are driven at a DC bias, giving a linear dipole contribution, \( q_e \hat{Z} \), to the potential, giving the well structure a slope. This bias alters the eigenmodes and eigenenergies of the structure, giving rise to the electrical performance characteristics of the QCL.

In the lasing scheme used in this work, the upper laser level is labelled as 4, where the lower laser level is level 3. An electron is injected into level 4, where, through spontaneous or stimulated emission, it transitions to level 3. Phonon scattering depopulates level 3 into the band of states beneath it, which are separated in energy level by the \(~30\) meV approximately matching phonon energy. The injector (ground) state of the period, level g, transports the electron into the proceeding period of the superlattice. Level g has high overlap with the proceeding level 4, and
very close energy level correspondence, leading to a high injection efficiency, $\eta_i$. Injection efficiency is the probability that an electron will be injected from a previous injector state into the proceeding upper laser level and then stimulated into releasing a photon.

Figure 4: The active region design used for the first iteration of devices tested in this work. The potential energy structure making up the quantum well superlattice is shown as the thin black lined background pattern. The superlattice is a series of Aluminum Indium Arsenide (barriers) and Indium Gallium Arsenide (wells) layers biased by a voltage to introduce the slope in the potential profile. Curves represent wavefunctions generated by this potential. Wavefunctions are anchored at their state energy, with the flat regions representing areas of zero probability density, while high amplitude portions of the wavefunctions show locations of high probability density for electron location. Key states are shown as more thickly drawn curves. Upper laser level, 4, and lower laser level, 3, in red. State g is the injector state leading from the previous period. State 2 is a state of particularly long lifetime, affecting lower level population in state 3. State 5 is a parasitic state where electrons may transition without the probability of lasing, generating heat and reducing efficiency.
The downward net transition energy from a lower laser level of one period to the upper laser level of the next is referred to as the voltage defect. These transitions are non-radiative, and thus, release energy as heat. The voltage defect is necessary for transport, as leading directly into a proceeding upper laser level without enough injector length and energy level difference would allow strong backscattering to the previous lower laser level.

Injection into the proceeding period is not guaranteed, and electrons will sometimes non-radiatively decay from injector to lower levels, transition into higher laser levels (so called parasitic states,) or escape into continuum. Figure 4 shows a parasitic state, level 5, highlighted above the upper laser level. Electrons in such a state will not be able to lase off the 4-3 transition in the lasing scheme. The effect is additional heat buildup due to non-radiative transitions to the next stage in lieu of the conversion to photon energy.

Section 4: Waveguides

Optical waveguides are structures designed to confine light within a region. This is done by engineering refractive index to be higher in or near the region light is to be confined, and lower around said region. The surrounding layers are referred to as the cladding.

A given waveguide may have several eigenmodes, which are electromagnetic distributions with stable transverse profiles throughout propagation within the waveguide. Each eigenmode has an effective refractive index that describes the propagation of the mode through the waveguide. For a 2D waveguide, such solutions can be found numerically. This work utilizes COMSOL Multiphysics’ Electromagnetic Waves, Frequency Domain module to calculate and visualize eigenmodes and effective indices.
Figure 5: Schematic of the waveguide used for the devices in this work. The Active Region is composed of a superlattice of Indium Gallium Arsenide and Aluminum Indium Arsenide. Cladding layers are composed of Indium Phosphide doped with Silicon to carry charge and modify index. Pure Indium Phosphide sidewalls confine current and optical mode to a fixed width.

The effective or modal index, \( n \), is a complex value. The real portion describes the phase velocity, while the complex portion describes the losses associated with travelling through the waveguide, \( \alpha_{WG} \). These relations are summarized below, where \( \lambda_0 \) is the vacuum wavelength of light and \( \beta \) is the propagation constant:

\[
\begin{align*}
    n &= n' + i n'' \quad (5) \\
    k_0 &= \frac{2\pi}{\lambda_0} \quad (6) \\
    \beta &= k_0 n' \quad (7) \\
    \alpha_{WG} &= 2k_0 n'' \quad (8)
\end{align*}
\]

In the case of a Fabry-Perot device, the waveguide is cut to length \( L \) and the flat facets provide interfaces for transmission and reflection. The light lost through transmission at the facets
from the waveguide to some external material is referred to as mirror loss, and is a function of reflection $R$:

$$R = \left| \frac{n - n_{\text{external}}}{n + n_{\text{external}}} \right|^2 \quad (9)$$

$$\alpha_M = -\frac{\ln(R_1 R_2)}{2L} \quad (10)$$

Here, $R_1$ and $R_2$ refer to reflection at the first and second interfaces of the Fabry-Perot device. For uncoated devices, $n_{\text{external}}$ is treated as the refractive index of air, 1. When a highly reflective (HR) coating is applied to a facet, say facet 2, $R_2$ is assumed to be 1.

In semiconductor lasers, the region where laser light is generated is the region in which it is desirable to confine light, and is referred to as the active region. A greater proportion of light confined to the active region promotes more stimulated emission and fewer losses in the cladding regions, creating more laser light. In the context of QCL waveguides, the active region refers to the entire superlattice structure grown within the cladding layers of the waveguide.

An important quantity for the modelling of QCL device performance is the confinement factor, which describes the percent of the mode confined to the active region, and thus, able to participate in lasing. It is calculated by integrating the square of the electric field amplitude over the active region and dividing by the integral of the entire distribution ($y$ is taken as the axis of propagation) [12]:

$$\Gamma = \frac{\iint_{\text{Active Region}} |E|^2 dxdz}{\iint |E|^2 dxdz} \quad (11)$$
Figure 6: Fundamental TM mode profile of a representative device, taken from COMSOL calculation. Electric field magnitude is plotted, with red representing high field magnitude, and blue signifying little or no field magnitude. The single lobe structure indicates that it is a fundamental mode. The majority of the intensity exists evanescently in the cladding, with roughly 20% existing the active region.

Important to note is that only the portion of the electric field along the growth (z) direction interacts with the superlattice. The in plane continuity of material in the active region results in a symmetry where interactions along the transverse direction, i.e. interactions with a transverse electric (TE) polarized wave, results in 0 dipole matrix element ($x_{if}$ or $y_{if}$) and 0 interaction rate [12]. Only the transverse magnetic (TM) wave will “see” the structure of the superlattice, allowing the field to “push” electrons through the structure via stimulated emission.

The waveguide structure utilized in this work is the buried heterostructure (BH) waveguide. This is structure composed of a ridge waveguide structure with cladding deposited as side walls. The side walls allow for lateral heat extraction. This is critical for QCL cooling, as the superlattice has a highly anisotropic thermal conductivity. In the lateral direction, the superlattice acts as a single continuous material. In the growth (vertical) direction, the superlattice has potentially
hundreds of interfaces in which thermal transport is inhibited due to scattering effects, resulting in a cross-plane thermal conductivity an order of magnitude lower than the in-plane (lateral) thermal conductivity [14]. Broad area QCLs have less relative lateral surface area to extract heat in-plane, forcing a greater portion of heat to be extracted vertically. The side walls are also insulating, guiding the current through a fixed width active region, as opposed to an outward diffusion through a slab waveguide structure. This ensures that the entire active region is biased in an approximated equal manner at all points.

Specific to this work, the cladding materials are indium phosphide layers with variable doping. Doping adds a Drude component to their permittivity, effecting their refractive indices. The doping also lowers resistance to allow for current flow. The side walls are left undoped for electrical insulation. Being a binary material, indium phosphide has a much higher thermal conductivity than the ternary materials composing the active region (aluminum indium arsenide and indium gallium arsenide.)

Section 5: Pulsed Characteristics

In pulsed operation, a QCL is driven by a rectangular current and voltage signal with low duty cycle. This gives a rectangular optical signal with similar duty cycle. The plateau of optical power for a pulsed signal is referred to as the peak power, and is the focus of an L-I, or light-current, plot.

The L-I characteristic shows lasing beginning at threshold current and continuing in an approximately linear fashion until reaches a maximum point, the pulsed rollover current. This rollover current corresponds to the maximum current density supported by the superlattice, where detuning between injector and proceeding upper laser level energies are minimized. At higher
currents, optical power rapidly drops. The range between the threshold and rollover currents is the dynamic range of the device.

Threshold current density is calculated as:

\[ J_{th} = \frac{l_{th}}{wL} = j_{tr} + \frac{\alpha_M + \alpha_{WG}}{\Gamma g} \tag{12} \]

Here, \( j_{tr} \) is the transparency current density and \( \Gamma g \) is the differential gain (the product of confinement factor \( \Gamma \) and differential gain coefficient \( g \)). Transparency current density is the current density at which the net gain accounting for intersubband absorption of the active region itself becomes 0. A beam traversing a pure active region material driven at \( j_{tr} \) would experience neither gain nor loss. Differential gain describes the growth of the gain coefficient (units of inverse length) per unit current density.

Slope efficiency is the power increase per unit current. As its name implies, it describes a linear behavior:

\[ \eta_S = N_s \left( \frac{\hbar \omega}{q_e} \right) \left( \frac{\alpha_M}{\alpha_M + \alpha_{WG}} \right) \eta_i = \frac{dP}{dI} \tag{13} \]

\( N_s \) is the number of active region stages, \( \hbar \omega \) is the photon energy, the losses \( \alpha_M \) and \( \alpha_{WG} \) are as described in the previous section, and \( \eta_i \) is the injection efficiency described in section 3 taking into account state lifetimes.

The L-I-V, or light-current-voltage, characteristic overlays voltage with optical power out as functions of current. The voltage shows diode-like behavior at low current, flattening out to approximately linear behavior over the dynamic range before diverging toward infinity after rollover.
Figure 7: Pulsed L-I-V for a representative narrow ridge device used in this work. The blue curve is voltage as a function of current. Voltage demonstrates a diode-like turn on voltage, followed by a linear regime over the operational range of the device, and finally a runaway effect at rollover current. The red curve is pulsed peak power as a function of current. Power demonstrates linear behavior after threshold current, developing sublinearity and rolling over at rollover current.

Wall Plug Efficiency (WPE) is the percent of input power converted to optical signal. It is a significant figure of merit for QCL performance, and can be calculated by dividing the optical power out by electrical power in:

\[ WPE = \frac{P}{IV} \]  

(14)
Section 6: 1/L Measurements

A given active region can be characterized with L-I measurements across a series of representative devices with varying lengths. Starting with the standard threshold condition for a round trip in a laser cavity:

\[ 1 = R_1 R_2 e^{\left(\Gamma g (J_{th} - J_{tr}) - \alpha_{WG}\right) L} \]  \hspace{1cm} (15)

Algebraic manipulations can give a relationship for threshold current density that is linear with the inverse of cavity length \( L \):

\[ J_{th} = \left( J_{tr} + \frac{\alpha_{WG}}{\Gamma g} \right) + \left( -\frac{\ln(R_1 R_2)}{\Gamma g 2} \right) L^{-1} \]  \hspace{1cm} (16)

A similar relationship can be found from the inverse of slope efficiency:

\[ \eta_s^{-1} = \left( \frac{q_e}{N_s \hbar \omega \eta_I} \right) \left( 1 + \frac{\alpha_{WG}}{\alpha_M} \right) \]  \hspace{1cm} (17)

Remembering the definition of mirror loss:

\[ \eta_s^{-1} = \left( \frac{q_e}{N_s \hbar \omega \eta_I} \right) \left( 1 + \frac{-\alpha_{WG} 2}{\ln(R_1 R_2)} L \right) \]  \hspace{1cm} (18)

Figure 8: 1/L measurement for first iteration of active region design. Data in red shows the linear relationship between the reciprocal of slope efficiency and cavity length. Data in blue shows the linear relationship between inverse cavity length and threshold current density.
Thus linear relationships are found between length and two features of the L-I characteristic; slope efficiency and threshold current density. This 1/L measurement can be done with at least two representative devices of different length, allowing the retrieval of injection efficiency $\eta_i$, differential gain $\Gamma g$, transparency current density $J_{tr}$, and waveguide loss $\alpha_{WG}$.

Section 7: Thermal Effects

Changes in average temperature of the active region cause changes in the pulsed characteristics of the QCL. Both threshold current density and slope efficiency follow exponential curves as functions of temperature, with threshold condition increasing with temperature while slope efficiency decreases.

Figure 9: Characteristic temperature measurement for first iteration device. Data in red shows the exponential growth of threshold current density with increasing active region temperature. Data in blue shows the exponential decay of slope efficiency with increasing active region temperature.

\[
J_{th}(T) = J_{th}(T_i) e^{\frac{T-T_i}{T_0}}
\]  
(19)  

\[
\eta_S(T) = \eta_S(T_i) e^{-\frac{T-T_i}{T_1}}
\]  
(20)
Here, $T_i$ indicates a reference temperature (20 °C in this work) from which to anchor the distribution. $T_0$ is the characteristic temperature of threshold current density and $T_1$ is that of slope efficiency.

**Section 8: Continuous Wave Operation**

Driving the laser with a continuous direct current (DC) provides bias for lasing, just as in pulsed operation. However, a low duty cycle in pulsed operation deposits a small amount of power in the active region, followed by a longer off cycle that allows the heat waste to dissipate. In CW, the laser has no cooling cycle, causing average temperature of the active region to raise toward a steady state solution balanced by heat extraction through cladding and substrate into a heat sink.
CHAPTER 3: INITIAL DESIGN MEASUREMENTS

A QCL design based on a high efficiency design [9] was created with BH ridge widths varying from 10 μm to 30 μm to observe CW power scaling with increased BH ridge width. The 40 stage design was reduced to 15 stages and grown and processed into QCL chips for testing.

Section 1: Quantum Cascade Laser Growth and Processing

A QCL active region is grown on a semiconductor wafer based on strain considerations. Each material has a lattice constant describing the atomic spacing of the crystal lattice. When growing a heterostructure, mismatched lattice constants cause a strain on the overall structure that can lead to bending, deformities, and even a separation of layers. For this reason, the indium phosphide based waveguide is grown on an indium phosphide substrate, as opposed to a cheaper silicon or gallium arsenide substrate.

The cladding layers of the waveguide and the superlattice are grown by molecular beam epitaxy (MBE,) slow growth process where material targets are heated to sublimate material in vacuum onto the target. The slow growth rate and vacuum pressures allow for fewer errors in the creation of the superlattice relative to other growth processes such as metalorganic chemical vapor deposition (MOCVD.)

The wafer is grown as a slab waveguide. The BH waveguide is formed by etching material away to the bottom of the superlattice, leaving a ridge of a designed width with an intact waveguide structure. Indium phosphide is deposited into the etched area, filling in the sidewall cladding. A gold ohmic contact layer is deposited over the center of the ridge as a thin ridge epitaxy side and over the substrate. Chips are then cleaved to size; roughly 400 μm wide by 2-10 mm long (propagation axis.)
The QCL chips are tested with point contacts to test for good electrical behavior (voltage characteristic.) Chips are then mounted on aluminum nitride submounts, epitaxy side down on the positive voltage contact, with gold wire bonds connecting the substrate contact to the negative contact. The submount and wire bonds allow current to spread evenly across the device to prevent localized high current density to burn the chip from contact points. Chips at this level are tested for L-I-V characteristics.

To test chip at carrying temperatures in pulsed or CW operation, a heat sink is needed. The chip on submount is baked onto a gold plated copper heat sink with indium solder. Temperature is read by a thermistor, a resistor with a well-defined temperature sensitivity, to aid in temperature control with a thermoelectric cooler (TEC.) Chips in pulsed operation can thus be tested at fixed temperature for characteristic temperature measurement. CW operation requires active cooling to prevent a thermal runaway from damaging the active region.

Section 2: Device Measurements

Devices grown to designed BH ridge widths of 10, 20, and 30 μm were measured to have ridge widths of 10.4, 21.2, and 31.0 μm, respectively [10]. All three devices had lengths of 3.15 mm and were designed for an output wavelength of 5.5 μm (5.7 μm measured.) L-I-V characteristics were taken in pulsed operation, showing an increase of maximum peak power with ridge width. Previous results for \( T_0 \) and \( T_1 \) were already measured to be 140 K and 710 K for the unaltered active region design [9].

The three devices, from narrowest to widest, had threshold currents measured to be 0.99 A, 1.79 A, and 2.63 A, and slope efficiencies measured to be 1.44 W/A, 1.41 W/A, and 1.45 W/A, respectively. 1/L analysis of this iteration of devices measured rollover current density to be 11.4
kA/cm$^2$, transparency current density to be 1.61 kA/cm$^2$, a differential gain of 2.4 cm/kA, and an injection efficiency of $\sim$72% [10].

The improvement in power with chip dimension in pulsed operation is a natural consequence of increased confinement to the active region. The increasing threshold current is a reflection of the fixed transparency current density and similar loss for each device. The increasing dimensions requiring greater total current to match roughly equivalent threshold current densities. The current limit of 10 A of the pulsed laser driver prevented reaching the predicted $\sim$11 A rollover current for the 31 $\mu$m device.

Figure 10: Pulsed L-I characteristics for three devices of varying ridge widths. Peak pulsed power is plotted against driving pulsed current. The 10.4 $\mu$m (blue,) 21.2 $\mu$m (purple,) and 31.0 $\mu$m (red) devices show standard linear threshold and sublinear rollover behavior. Rollover current density is fixed, so rollover current travels as a function of device width. Rollover current for the 31.0 $\mu$m device is predicted to be $\sim$11 A, beyond the maximum capability of the driver used to gather the data.
To test the devices in CW, the TEC was set to keep the thermistor (in thermal contact with the aluminum nitride submount) at 20 °C. While the active region cannot be held to a fixed, uniform temperature while being driven in CW, keeping the submount at a fixed 20 °C will act as an appropriate heat sink for the active region to keep it at a steady state.

The three devices, from narrowest to widest, had CW threshold currents measured to be 1.12 A, 2.40 A, and 3.85 A, and slope efficiencies measured to be 1.62 W/A, 2.34 W/A, and 1.70 W/A, respectively.

Wider devices retain more heat in the active region, leading proportionally larger increases in CW threshold current. A maximum CW power of 1.62 W was measured for the 10 μm device, compared to 2.34 W for the 20 μm device [10]. This is the demonstration of power scaling with increased ridge width.

Figure 11: CW L-I characteristics for three devices of varying ridge widths. CW power is plotted against driving current. The 10.4 μm (blue,) 21.2 μm (purple,) and 31.0 μm (red) devices show a general curvature throughout their range, and experience a thermal rollover before the rollover current shown in pulsed operation.
Figure 10 shows the measured CW power for the 10, 20, and 30 μm devices. The peaks at 5 A for the 20 μm device and 6.5 A for the 30 μm device are indicative of multiple transverse modes contributing to total power. This sudden change in power is referred to as “mode-hopping.” This is an effect determined by both geometry and temperature of the active region, and can also appear as part of the temporal pulse profile. Mode-hopping is consistent on a per chip basis (the effect always occurs at the same currents in CW, or same time in pulsed operation for a given driving current.) Mode-hopping in QCLs is not yet well understood and is beyond the scope of this work.
CHAPTER 4: THE SEMIEMPIRICAL MODEL

A model was developed to predict pulsed and CW power for a given QCL design from first principles. To prevent overcomplicating the model with superlattice simulations, various approximations and fittings are used to allow for timely simulations of adjustments to superlattice design and device geometry.

Section 1: Waveguide Simulations

The waveguide is simulated in COMSOL to find the fundamental TM mode. This can be found programmatically by creating a redefinition of confinement factor restricted to the field in the growth direction, the TM confinement factor:

$$\Gamma_{TM} = \frac{\iint_{Active\,Region} |E_z|^2 dydz}{\iint |E|^2 dydz}$$  \hspace{1cm} (21)

The eigenmode with maximum TM confinement factor will correspond to the fundamental TM mode. Higher order TM modes will feature nodes in the active region and more overlap with cladding features, reducing confinement to the active region. This approach eliminated TE modes, as their growth direction components will be negligible relative to the total field integral in the denominator. The confinement factor of this mode is calculated as equation (9).

The linear portion of the voltage characteristic is then found via linear regression. To account for design alterations such as variable doping and number of stage, the line is taken to be:

$$V = N_s(V_0 + JR_d)$$  \hspace{1cm} (22)

Here, $V_0$ is a voltage offset, and $R_d$ is a differential resistance expressed as voltage per unit current density. Voltage scales with number of stages. Differential resistance is inversely
proportional to doping level (sheet density,) reflecting greater carrier transport with increased doping.

This linear voltage model is used to simulate the active region temperature with the assumption that 100% of electrical power (the product of current and voltage) is deposited uniformly into the active region. The device with submount and thermistor is simulated for a steady state thermal solution, where it is assumed that the thermistor is kept at 20 °C. A distribution of average active region temperatures is taken from no current to rollover current at 100 equally spaced points.

Section 2: Pulsed Power

Using 1/L measurements, waveguide simulation results, and design specifications, the pulsed power is projected from threshold current and slope efficiency. Because slope efficiency only describes a linear trend, an envelope must be developed to conform the line to the rollover phenomenon at the end of the dynamic range. The dynamic range is parametrized from 0 to 1 as:

\[ x = \frac{J - J_{th}}{J_{roll} - J_{th}} \]  \hspace{1cm} (23)

An envelope is empirically fitted to recreate the curvature of a representative pulsed L-I characteristic introduced by sublinearity and rollover current. In the case of the first iteration of chips, the following envelope was used:

\[ C_{env} = \left( 0.7 + 0.3\sqrt{1 - x^{0.7}} \right) \]  \hspace{1cm} (24)

Power is found by integrating slope efficiency from threshold to a current.

\[ P_{pulse}(I) = C_{env} \int_{I_{th,pulse}}^{I} \eta_{S,pulse} \, dI = C_{env} \eta_{S,pulse} (I - I_{th}) \]  \hspace{1cm} (25)
Section 3: Continuous Wave Threshold

Characteristic temperature describes the changing of threshold current with temperature as in equation (19). When driving a QCL in CW, the ohmic heating and active cooling of the heat sink leads to a steady state temperature distribution. This gives a direct 1-1 relation from current to average active region temperature, thus, temperature as a function of current.

Figure 12: Threshold current and temperature as a function of CW driving current for a hypothetical narrow ridge device with threshold current density characteristic temperature $T_0 = 75 \text{ K}$. Temperature (red) increases as a smooth curve with driving current. Threshold current (blue) is calculated as a function of temperature, itself a function of driving current. Driving current is represented as a dashed line of slope 1 to indicate where threshold current becomes equal to its corresponding driving current. A device can operate in CW in the driving current range between the two interception points of the driving current line and the threshold current curve, provided it does not first reach rollover current.
The CW threshold condition is the driving current that self consistently heats the device to the temperature required to raise pulsed threshold to be equal to driving current:

\[
J_{th,cw} = J_{th} \left(T(I_{th,cw})\right) \rightarrow I_{th,pulse} e^{\frac{T(I_{th,cw}) - T_i}{T_0}} = I_{th,cw}
\]  

(26)

When observed graphically against driving current, the threshold current as a function of driving current creates a range where CW operation is thermally plausible. The first intersection with the driving current line of slope 1 is the CW threshold current, while the second intersection is the current at which device heating causes the laser to no longer output power in CW.

Section 4: Continuous Wave Power

When CW operation is considered, the effects of temperature must be considered. While pulsed slope efficiency at variable temperature is adequately explained with a characteristic temperature, more factors must be considered to predict the distribution of CW power.

In pulsed operation, the derivative of power with respect to current can be taken as slope efficiency. However, this makes an assumption of constant temperature. The amount of energy introduced to the active region during CW operation requires temperature considerations. By introducing a temperature varying with current, we redefine the derivative of power with respect to current, temporarily neglecting the correction due to the effects of rollover current. First, the total derivative of power with respect to current must be considered:

\[
\frac{dP}{dl} = \frac{\partial P}{\partial I} + \frac{\partial P}{\partial T} \frac{dT}{dl}
\]  

(27)

Temperature shifted slope efficiency (equation (20)) is taken as the partial derivative with respect to current, while the partial derivative of power with respect to temperature requires consideration of the characteristic temperatures of slope efficiency and threshold current. The
partial derivative of temperature with respect to current is taken from simulation data numerically. The form of power is assumed as an expression similar to the linear case, but with temperature dependence.

\[
\frac{\partial P}{\partial T} = \frac{\partial}{\partial T} \left( \eta_s(T)(I - I_{th}(T)) \right)
\]  

(28)

The derivative is computed using equations (19) and (20) for the temperature dependent quantities. The end result is the expression:

\[
\frac{\partial P}{\partial T} = \eta_s(T) \left( \frac{1}{T_1} - \frac{1}{T_0} \right) I_{th}(T) - \frac{1}{T_1} I
\]  

(29)

The envelope is reparametrized along the new threshold and applied to the integral of the total derivative:

\[
x_{cw} = \frac{J - J_{th,cw}}{J_{roll} - J_{th,cw}}
\]  

(30)

\[
P_{cw}(I) = C_{env}(x_{cw}) \int_{I_{th,cw}}^{I} \left( \frac{\partial P}{\partial I} + \frac{\partial P}{\partial T} \frac{\partial T}{\partial I} \right) dI
\]  

(31)

Section 5: Replication of Experimental Data

With L-I-V characteristics for representative devices of the three design ridge widths, the design parameters were input into the model to recreate the experimental data. For the three devices, from narrowest to widest, predicted pulsed threshold currents were 0.91 A, 1.80 A, and 2.63 A, with slope efficiencies of 1.61 W/A, 1.60 W/A, and 1.60 W/A, respectively. These values were calculated using equations (12) and (13). The model simulated pulsed power for each device, displaying good agreement with experiment [10].
Figure 13: Comparison of projected and measured peak pulsed power as a function of pulsed driving current for first iteration devices. 10.4 μm (blue), 21.2 μm (purple), and 31.0 μm (red) device peak pulsed power are shown as data points for measured results and solid and dashed lines for projected results in the model.

CW L-I characteristics were similarly simulated using the model. From narrowest to widest, predicted pulsed CW currents were 0.91 A, 1.80 A, and 2.63 A, with slope efficiencies of 1.61 W/A, 1.60 W/A, and 1.60 W/A for the three devices, respectively. The predicted maximum power for the 21.0 μm device was 2.01 W [10].
Figure 14: Comparison of projected and measured CW power for first iteration devices. 10.4 μm (blue), 21.2 μm (purple), and 31.0 μm (red) device CW power are shown as data points for measured results and solid and dashed lines for projected results in the model. Spikes in power are due to multi-mode contributions (mode-hopping) beyond the scope of this work.

Of note is the rollover behavior predicted by the model in CW operation, earlier than the proper rollover current. This is a thermally induced effect. This will be referred to as a “thermal rollover.” The steady increase of the calculated threshold at a given temperature (not the literal threshold point where lasing starts) and decrease of slope efficiency give a gradual, roughly symmetric shape to the distribution of CW power around thermal rollover. Contrast this with the sudden decline visible on a pulsed L-I characteristic at the rollover current.
Section 6: Design Optimization for the Base Active Region Stage Design

The model was then swept through a three dimensional space of design parameters to optimize the base active region design. Keeping the cavity length of 3.15 mm, the number of active region stages, BH ridge width, and relative doping to original design were varied to optimize maximum projected CW power. The optimization converged on a device with 21 stages, half relative doping in the active region, and with a ridge width of 25 μm. The predicted power of 2.24 W is an 11% increase from the projected 2.01 W for the 21 μm wide device. The linear voltage model to calculated input power, and thus, WPE. The 21 μm was projected to have 5.7% efficiency, but the new optimization projected 6.7%. This is to be expected, as a greater efficiency is indicative of less heat buildup, heat being the primary obstacle to CW output [10].

![Projected pulsed and CW power of 21.2 μm device (blue,) and the optimized device (red.)](image)

Figure 15: Projected pulsed and CW power of 21.2 μm device (blue,) and the optimized device (red.) The optimized device is a 21 stage device with an active region 25 μm wide with half the active region doping of the 21.2 μm design. The lower active region doping results in a compressed dynamic range (earlier threshold and rollover currents) with higher slope efficiency for a net improvement in CW power.
Section 7: Characteristic Temperature Analysis

To analyze the sensitivity of CW power to characteristic temperature, the 31 μm device was simulated over a 2D space of characteristic temperatures, leaving all other input parameters identical to that used to simulate the device. The widest device was chosen, as it is the most prone to heat buildup.

![Surface of maximum CW power produced by given characteristic temperatures in the 31.0 μm device.](image)

Characteristic temperatures for threshold current density ($T_0$) and slope efficiency ($T_1$) are varied from 100 K to 1000 K. Purple indicates no power, with dark red indicating highest power. Particular sensitivity to $T_0$ is shown as the steep gradient in the direction of increasing threshold current characteristic temperature.

The original values of 140 K for $T_0$ and 710 K for $T_1$ were allowed to vary from 100 K to 1000 K. Characteristic temperature for slope, $T_0$, was shown to be a limiting factor for performance. The wide device was projected not to generate CW power if $T_0$ was allowed to reduce to 100 K. While performance suffered for low slope efficiency characteristic temperature, lasing was still possible. The literature shows that threshold characteristic temperature near 250 K are
achievable [11]. Such an increase would result in a projected power increase from 1.6 W to 4.5 W [10].

This significant shift shows that efforts to increase $T_0$ should be made for broad area active region designs. Reference [9] attributed the low $T_0$ of this design to backscattering from level 2 (refer to Figure 4,) which is confined to the active region and does not promote good transport to the injector. $T_1$ is sufficiently high, and increases to it would result in diminishing returns.

Section 8: Separate Confinement Heterostructure

In the standard BH configuration used in this work, the high refractive index of the active region is used as the primary guiding layer in the waveguide. Greater confinement of the mode to the active region can be achieved by surrounding it with higher index material “spacer layer” in a so called separate confinement heterostructure configuration. In this work, doped indium gallium arsenide is considered as it has a sufficiently high refractive index and the material is already being used in the growth process. However, using a ternary material with low thermal conductivity around the heat producing active region would result in a thermal blanket effect, possibly endangering CW performance.

300 nm layers of doped indium gallium arsenide on both the substrate and epitaxy side of the active region were simulated as additions to the 31 μm device. Modal confinement to the active region increased from 20% to 30%, while heating at thermal rollover increased by 5 K. The end result was an increase in projected maximum CW power from 1.6 W to 2.0 W [10]. This design addition requires no modification to the active region, and is a good method to increase both pulsed and CW output power for a broad area design. A patent is currently pending for this technique.
Figure 17: Temperature distribution at thermal rollover for 31 μm design without (left) and with (right) additional 300 nm indium gallium arsenide spacers added on both the substrate and epitaxy side of the active region. Average temperature is shown to increase in the case of additional indium gallium arsenide spacers as a darker red throughout the active region. However, the difference is small, relative to the increased confinement, leading to increased CW power.

Section 9: Short Injector Design

The primary obstacle in CW power optimization is effective heat removal from the active region. Broad area designs allow for power scaling, but make lateral heat extraction less effective with more material to travel through before leaving the active region. As already discussed, the superlattice geometry gives a cross plane thermal conductivity an order of magnitude lower than that of the in plane. A potential solution to this problem that would still allow a broad ridge is a thinner superlattice stage. This can be accomplished by reducing the thickness of the injector region. This can allow for more stages to be grown for given superlattice thickness, or a reduced thickness at fixed number of stages. The former increases slope efficiency and power input (voltage scales with stage number,) while the latter reduces temperature build up (and confinement by virtue of the thinner superlattice layer.)
Figure 18: 1D cross section of predicted maximum power in the 6D sweep. Shown are the power distributions with variable relative doping (top left,) stage thickness (top right,) number of stages (bottom left,) and ridge width (bottom right) around the optimal configuration. All plots encompass the optimal design of a 40 μm wide, half doped active region with 16 stages. A trend of increased CW power with decreasing stage thickness/height is observed, showing that a shorter injector, other parameters being equal, is a preferred design.

To simulate a short injector design based on the active region design to be modified, the linear voltage model assumes a fixed voltage drop per stage, independent of individual stage thickness. Gain coefficient is assumed inversely proportional to stage thickness to keep overall gain fixed for a fixed number of stages. Simulations were run over six dimensional parameter space. Relative doping, active region stage thickness, number of stages, ridge width, and the
individual thicknesses of added spacer layers at both substrate and epitaxy side were allowed to vary independently. A cavity length of 3.15 mm and the same characteristic temperatures as the base design were assumed to allow for comparison. The model converged on an optimal result of 50% relative doping (corresponding to a rollover current density of ~ 6 kA/cm$^2$,) active region stage thickness of 30 nm (45.3 nm in the unaltered design,) 16 stages, 40 μm wide ridge, and spacer thicknesses of 350 nm and 250 nm for substrate and epitaxy side, respectively. The resulting maximum power was projected to be 2.83 W with a WPE of 6.7 %. Relative to the 21 μm device projections, this optimization predicts a 41% increase in maximum CW power.
CHAPTER 5: SECOND DESIGN EXPERIMENTAL RESULTS AND THE REFINING OF THE MODEL

A second set of devices were grown and processed in parallel to the first design. These devices utilized a shorter injector design with a greater number of stages, and ridge widths ranging from design widths of 9 to 40 μm. These devices were characterized and tested.

Section 1: Design Changes

The devices were grown and processed as described in Chapter 3 Section 1. Stage height was reduced from 45.3 nm to 41.6 nm by reducing the thickness of the injector region. Laser wavelength was reduced from 5.7 μm to 4.4 μm (an increase in photon energy.) Number of stages was increased from 15 to 20, leading an overall superlattice layer height increase from 679.5 nm to 832 nm. The new geometry greatly increased active region confinement, i.e. for a 10 μm design, confinement would increase from 18.4% to 37.0%.

The design of this device focused on increasing T₀, reducing injector thickness, and increasing slope efficiency. The decreased injector thickness and increased photon energy led to a higher voltage needed for device operation, providing a steeper “slope” for the quantum well superlattice to provide the energy difference needed for both the photon transition and voltage defect.

Measurement of the L-I characteristics showed diminished CW performance for ~20 μm wide devices relative to ~10 μm, and no CW lasing for larger ridge widths. It was anticipated that the ~30 μm devices would have CW performance exceeding smaller ridge devices of his design. Thermal effects not yet present in the model had become prominent in this design and had to be accounted for.
Figure 19: Simulated and projected power for a 8.5 μm x 3 mm device with second iteration active region design. Peak pulsed power is blue while CW power is red. Measured data is shown as points while projections from the model are shown as solid lines.

Section 2: Characteristic Temperature

L-I characteristics were measured from 20 to 120 °C to determine characteristic temperatures for threshold current density and slope efficiency. Data was taken using a room temperature mercury cadmium telluride (MCT) detector from 20 to 70 °C on the TEC (whose maximum sustainable temperature is 80 °C.) To capture higher temperature data, the device was placed on a hot plate with thermal grease to promote good thermal contact. Data was taken with the MCT from 40 to 120 °C to ensure that any scaling factor needed to compensate for the different alignment would be obtainable from the 40 to 70 °C overlapping data sets.
The measurements revealed that the characteristic temperatures produced significantly different results for temperatures below and above 80 °C. $T_0$ decreased from 212.8 K to 194.9 K when heating past this critical temperature, showing a discontinuity in the threshold current relationship with temperature. $T_1$ shows nearly flat behavior below 80 °C, meaning that slope efficiency remains nearly constant at this temperature range. Slope efficiency begins a sudden decline past 80 °C.

Once the dual characteristic temperature behavior had accounted for in the model, the CW projections produced more accurate results for narrow ridge (~10 μm) devices. A point of inflection appears in both experimental and projected figures at a current associated with ~80 °C heating.

Section 3: Rollover Current

When taking variable temperature data for the characteristic temperature measurement, it was observed that rollover current appeared to be negatively correlated with increasing temperature. This differed from the first iteration design, which had approximately constant rollover current except when approaching cryogenic temperatures.

The narrow ridge device showed anomalously high rollover current, attributed to device to device variation. The 21.2 μm and 31.0 μm devices showed very close correspondence for rollover current densities, showing an approximately linear trend with temperature. This is attributed to the increased bias and raised upper laser level of the active region. Carriers in the upper laser level are more easily thermally excited to parasitic and continuum states in this design, as they have less energy separation in this design. The escape into non-lasing states at high bias prevents lasing,
lowering the voltage bias (and driving current) where maximum optical output can be achieved at elevated temperatures.

Figure 20: Rollover current density as a function of temperature for devices of variable ridge width. Three devices of different widths were tested; 8.5 μm (black), 21.2 μm (red), and 31.0 μm (blue). The 8.5 μm device has anomalously high rollover current. The 21.2 μm and 31.0 μm devices both show a similar linear trend, with decreasing rollover current density with increasing temperature.

Section 4: Model Refinements

To accurately reproduce the experimental data from the second iteration design, the additional effects had to be accounted for. The dual ranges for characteristic temperatures were implemented, along with the temperature varying rollover current. The envelope for the new
design also had to be changed to more accurately reflect the curvature of the new design L-I characteristic.

Figure 21: Envelope (measured peak pulsed power divided by linear projected power) as calculated from four representative devices, and a best fitting curve. The value $x$ is the parametrization from threshold ($x = 0$) and rollover ($x = 1.0$) current. An envelope value of 1.0 indicates direct correspondence to linear projected power. The envelope decreases near rollover current and rapidly decays at and after rollover current.

The measured pulsed powers of four representative devices were divided by the linear power predicted by a naïve interpretation of the slope efficiency (without rollover/sub-linearity effects.) A best fit function for this distribution was found to be used as the envelope for this structure.
The variable rollover current provided a new dynamic behavior for the parametrization of the envelope. Again, the relationship between temperature and current density are taken from a COMSOL thermal simulation.

With these model refinements and a linear fit for the higher voltage characteristic, the broad area device performance could be more accurately projected. However, the addition of these new factors added additional sensitivity of model accuracy to parameters that determine temperature distribution, such as the measurement of ridge width.

The behavior of the shrinking rollover current is visible in the data for ~20 μm devices. Instead of the roughly symmetric thermal rollover expected of a thermally induced degradation of slope efficiency and threshold condition, there is an abrupt drop-off similar to a pulsed L-I characteristic at rollover.

The end result is that dynamic range is reduced from the left by the increased bias causing heating to increase the threshold current, and from the right by the reduction of rollover current. The reduced dynamic range leaves little room for the growth of power, even with increased slope efficiencies (~ 3 W/A.) Previous projections were able to account for the dynamic range reduction from the left, but not the right.
Figure 22: Projected power for two 21.2 \( \mu \text{m} \) wide devices of second iteration design. Peak pulsed power is blue while CW power is red. Measured data is shown as points while projections from the model are shown as solid lines.

Due to the inclusion of new temperature effects, increased sensitivity to temperature projections and width measurements cause CW projections to properly approximate overall power with a shifted dynamic range.
CHAPTER 6: CONCLUSIONS

The goal of increased QCL CW power will require the consideration of many different variables during the design stage. It is clear that the path forward lies with broad area configurations and ensuring a good dynamic range over which to build up power.

The second design sacrificed part of the CW dynamic range from the left by allowing temperature build up to increase threshold current. The benefit was a greater slope efficiency, which would have ensured a greater CW power output had the rollover current remained stationary. The temperature dependence of rollover current reduced dynamic range from the right, giving an overall smaller range over which to build up CW power. This destroyed broad area device performance through greater heat buildup due to the increased bias for the design.

Characteristic temperature in the second iteration improved upon that of the first, with very strong stability at sub-80 °C temperatures. As shown in the analysis in Chapter 4 Section 7, high characteristic temperatures are critical for CW power optimization. However, the unanticipated dual range of characteristic temperatures created an inflection point on CW L-I characteristics not already significantly affected by the thermally dependent rollover. Threshold condition fared well with the dual range, with $T_0$ in both portions near ~200 K. Slope efficiency experienced a strong negative effect after the critical temperature. This may be attributed to the short injector more easily allowing thermally activated backscattering at elevated temperatures.

Future work for CW power maximization requires the balancing the requirements of low temperature CW operation, large dynamic range, high slope efficiency aided by good confinement, increasing threshold characteristic temperature, and the prevention of carrier escape and backscattering in the superlattice. Problems introduced by the short injector design on rollover
current may be mitigated by lowering photon energy, pushing the upper laser level deeper into the superlattice to prevent carrier escape, with the added benefit of a smaller required bias for a reduced thermal load. A smaller number of stages coupled with spacer layers would allow for a compromise of reduced number of stages to aid in heat extraction while maintaining good confinement. The effects of the dual range characteristic temperature would be less significant if operating temperature can be kept below critical temperature.
LIST OF REFERENCES


