Simulation of THz Quantum Cascade Lasers

Viktor Rindert

Master's Thesis, Aug 2021 - Dec 2021
At the Division of Mathematical Physics, Lund.

LUND UNIVERSITY

LTH

FACULTY OF ENGINEERING

Abstract

Quantum Cascade Lasers (QCLs) have since their invention in 1994 become a crucial source for mid-infrared and THz radiation. Mid-infrared QCLs have achieved room temperature operation while THz QCLs have not. As such, simulations are an important tool in order to achieve a THz QCL operational at room temperature. This work uses a simulation package based on a nonequilibrium Green's function model to investigate which Conduction Band Offset (CBO) should be used for the most accurate results and if there is a difference between samples from different labs. These CBOs are then used to simulate current the state-of-the-art THz QCLs, which is done with great success. For example, the thermal degradation of the QCLs was analyzed and showed that thermal backfilling and thermally activated leakage channels were the main cause. Also, new QCLs were designed from the information gained. Two of these samples are predicted according to the simulation package to outperform the current best performing QCLs.

Acknowledgements

First and foremost, I would like to thank my supervisor Andreas Wacker for taking his time to supervise and teach me about the fascinating area of Quantum Cascade Lasers. Secondly, I would also like to thank the other members of the group, Andreas and Ekin, for always being available for discussion and making the project even more enjoyable. Also, I would like to thank Gillis Carlsson for taking his time to examine the project. Last but not least, many thanks to the mathematical physics department for the welcoming atmosphere and the free coffee.

List of Acronyms

ac:	alternating current
CBO:	Conduction Band Offset
dc:	direct current
HLS:	Higher Lying States
LLS:	Lower Laser State
LO-Phonon:	Longitudinal Optical-Phonon
NDC:	Negative Differential Conductance
NEGF:	Non-Equilibrium Green's Function
PDC:	Positive Differential Conductance
QCL:	Quantum Cascade Laser
QW:	Quantum Well
THz:	Terahertz
ULS:	Upper Laser State

Contents

1 Introduction

People have been astounded by the properties of quantum mechanics ever since its inception in the first half of the twentieth century. In short, quantum mechanics describes physics on the microscopic scale and is one of the cornerstones of theoretical physics. The continuous shrinking of electronic components during the last century has led to the birth of so-called quantum devices, i.e. devices that require quantum mechanics to explain their operation [1]. Thus, understanding quantum mechanics is of utmost importance for the development of e.g. transistors and Quantum Cascade Lasers (QCLs).

QCLs are a type of semiconductor laser that emit light in the infrared and THz region of the electromagnetic spectrum. QCLs where first theorized by Kazarinov and Suris in 1971 [2], and realized in 1994 by Faist *et al.* [3]. The advancements since then have been rapid, with mid-infrared QCLs being continuous-wave operational at room temperature [4] and being commercially available. However, designing QCLs that operate in the so-called THz gap(1-10 Thz) has been proven to be more difficult, with a maximum operating temperature of 250 K [5] being achieved in 2020. This text will focus on THz QCLs in particular and the focus on mid-infrared QCLs will be omitted.

THz QCLs have several interesting technical applications. This includes non-invasive medical imaging, gas detection, detection of substances such as explosives and drugs and THz spectroscopy of DNA [1, 6]. To make this a reality, the operating temperature and the wallplug efficiency of the THz QCL has to be increased. Hence, this is the main focus of current THz QCL research. Big improvements have been made lately as cryogenic temperatures were required before recent research [5, 7] enabled the use of thermoelectric cooling, which tremendously facilitates technological applications.

Simulations are a key tool when developing new and better QCLs. Several different approaches exist, e.g. based on rate equations, Monte Carlo techniques, density matrix theory, or Non-Equilibrium Green Function (NEGF) theory [8, 9], which will be deployed in this thesis. Pros and cons exist with all these simulation approaches, whereas the NEGF method is the most sophisticated at the expense of requiring the most computational time [10]. The NEGF method will be deployed by using a simulation program called the NEGF-Package which was developed at Lund University by Wacker, Winge and Franckie [10–12]

The purpose of this thesis is to use the NEGF-package to answer a multitude of questions. First, whether or not there exists uniformity between different groups in how material parameters are defined will be investigated. This will be done by performing simulations with different material parameters and then comparing the simulated results to experimental data. Secondly, simulations will be performed on the QCLs with the highest temperature performance as a means to understand the underlying physics of these devices. Thirdly, the simulation package will be deployed to investigate what causes the thermal degradation of these state-of-the-art QCLs. Lastly, the thesis will round off by presenting a new QCL. The structure of this new sample will be based on what we have learned from the earlier simulations and will then be optimized by using the NEGF-package.

The thesis will begin with an introduction to QCLs in chapter 1. Then in chapter 2 the mechanisms that cause thermal degradation in QCLs will be discussed. In chapter 3, we describe in detail how the simulations are performed and in chapter 4 the results will be presented together with a discussion of the results. Lastly, in chapter 5 the results will be concluded together with a discussion about the outlook of THz QCLs.

1.1 An Introduction to QCLs

Fundamentally, QCLs consist of a *superlattice* placed in a waveguide. Superlattices were first proposed in 1970 [13] and are made by alternating layers of two different materials with different band gaps. The Conduction Band Offset (CBO) between the two materials results in quantum wells where the discrete energy levels spatially overlap, creating energy bands and energy gaps, as shown in figure 1.

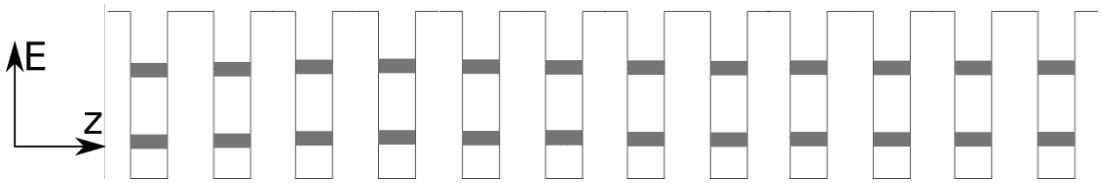

Figure 1: In this figure a semiconductor superlattice is schematically drawn. It shows how the quantum wells and barriers are repeated, resulting in minibands shown as grey bands in the well. The energy is represented by the E-axis and the growth direction by the z-axis.

Using superlattices to amplify light, i.e. achieve gain, was then proposed one year later by Kazarinov and Suris [2]. The idea behind this light amplifying device was to apply an electric field such that the ground state of a quantum well was aligned with the first excited state of the next quantum well. Consequently, the electron would be injected to the first excited state by resonant tunneling, and then transition to the ground state in the same well by the emission of a photon as shown in figure 2.

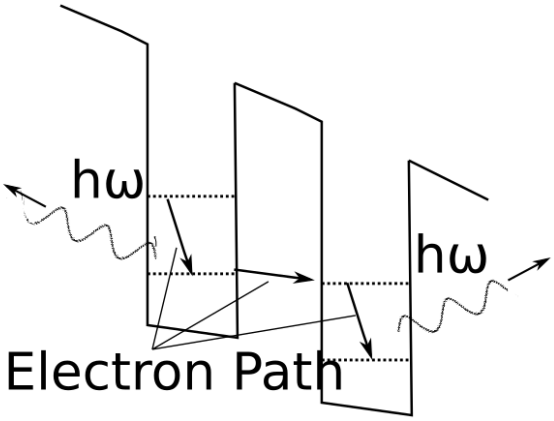

Figure 2: The original idea of a light amplifying superlattice structure as proposed by Kazarinov and Suris in their famous paper [2]. An electric field is applied over the structure such that electrons cascade through the superlattice by tunneling from the ground state to the first excited state of the next quantum well, to then transition down to the ground state in the same quantum well by emitting a photon.

The first realization of this type of cascading superlattice laser was the QCL in 1994 by J. Faist *et al.* [3] which used a more complex structure than what was proposed in 1971. This section thus aims to give the reader sufficient background to understand how a QCL is designed and operates. As a first step, the QCL will be compared to the more well-known traditional semiconductor lasers as a reference.

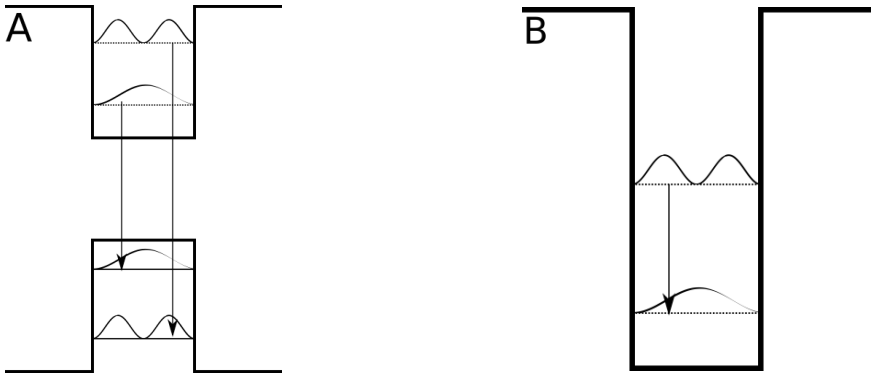

Figure 3: Two figures graphically showing the difference between interband transitions and intersubband transitions. In figure A: two examples of interband transitions, i.e. transitions between two different bands, are shown. In B: a transition between two subbands are shown, i.e. a intersubband transition. These intersubband transitions are relevant for QCLs while interband transitions are relevant for more traditional semiconductor lasers.

The most important feature of the QCL is that it uses *intersubband* transitions instead of *interband* transitions used in more traditional semiconductor lasers. These transitions are drawn schematically in figure 3. The most glaring difference is that intersubband transitions are not restricted by the band gap. Consequently, gain at much lower frequencies can be achieved compared to interband lasers, which has led to the development of QCLs in the THz and mid-infrared region of the electromagnetic spectrum [1]. Another important difference is the joint density of states resulting in very dissimilar gain spectra. For the intersubband case, the gain spectrum is shaped like a delta-function, only allowing absorption/emission of photons with the energy difference of the initial and final state [14]. Consequently, the structure is transparent to photons with higher or lower frequencies. To compare, in the interband case the gain spectrum is wider and looks like a step function [14]. Schematic drawings of the spectra are shown in figure 4.

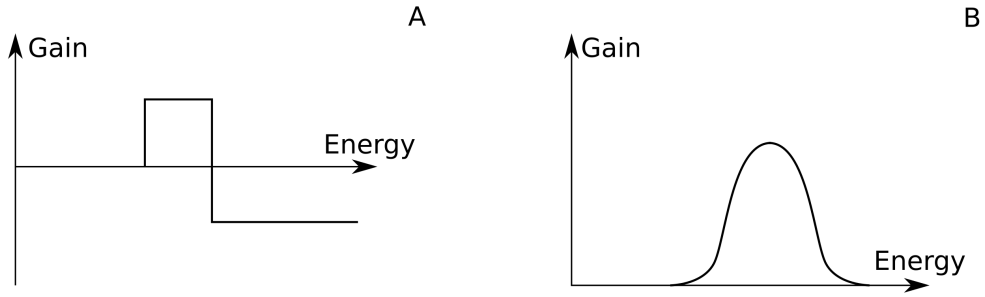

Figure 4: Figures showing the difference in gain spectrum for A: interband transitions and B: intersubband transitions.

Which type of scattering drives the transitions is also widely different between the two types of transition. In the interband case, radiative scattering dominates while non-radiative scattering dominates for intersubband transitions [1]. As a result, the design of the so-called active region, described below, is very important for QCLs.

1.2 Active Region Design

The active region is the region where the stimulated emission takes place to achieve lasing. As with all lasers, population inversion and electrical stability are necessary to achieve this [1, 15]. Here I will briefly go over how this is realized using a two quantum well structure, which is currently both the simplest and best-performing design as of now [5, 7]. In figure 5 two adjacent active regions are shown. Each active region has three levels. The well material is GaAs and the barrier material is $Al_xGa_{1-x}As$ which is standard for THz QCLs. The ground state acts as both the injector and extractor states, the first excited state acts as the Lower Laser State (LLS), and the second excited state acts as the Upper Laser state (ULS). The population inversion is designed to be between the ULS and the LLS. To achieve population inversion, it is necessary to have a long ULS lifetime and short LLS lifetime [1, 15]. This is done by injecting a large number of electrons into the ULS and by rapidly extracting the LLS population.

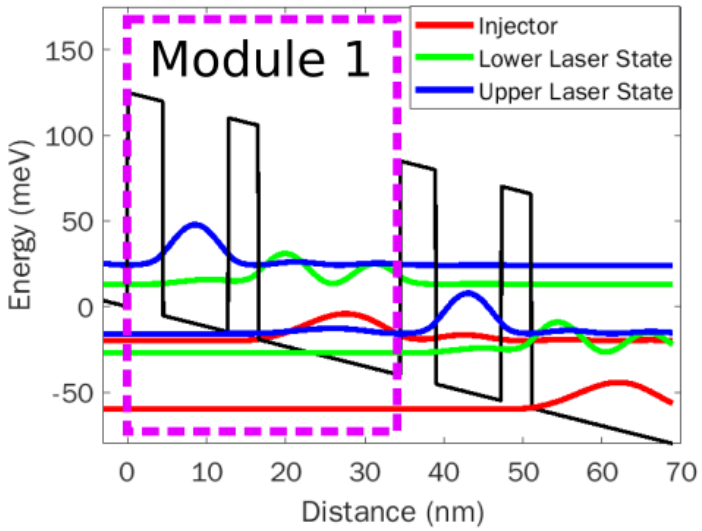

Figure 5: The bandstructure of a typical resonant tunneling injection THz QCL. In the figure, two modules are shown with three states each. In the figure, this can be seen since the ULS and the injector are aligned and the extractor is roughly one Longitudinal Optical-phonon beneath the LLS energy-wise.

The primary ways of injecting electrons are Resonant Tunneling (RT) injection and scattering assisted (SA) injection [15]. RT injection means that electrons tunnel from the injector to the ULS where the injector and the ULS are aligned, as is the case in figure 5. Consequently, the width of the injection barrier is of utmost importance. Decreasing the barrier width will increase the transition rate from the injector to the ULS, as well as from the injector to the LLS [1]. This is relevant for all scattering mechanisms, which will be described in chapter 2. On the other end, increasing the barrier width results in a smaller tunneling rate. This type of tradeoff is typical when designing a QCL as changing parameters will have both beneficial and detrimental effects. However, for the SA injection scheme, the levels are not aligned. Thus the transfer happens by scattering processes [15].

Designs on RT injections are currently performing better [5], but for low-frequency designs (<2 THz), SA injection has been shown to perform better [16]. The authors of Ref. [16] suggest that this is because at low frequencies the ULS and the LLS are separated by an energy close to the width of the level, which results in resonant tunneling to the LLS when using RT injection. As such, *line*

broadening of the LLS can result in detrimental current from the injector directly to the LLS which results in a higher current without the emission of a photon, as well as less population inversion [9, 17].

As of writing, the most common method of depopulating the LLS is by using the direct-phonon method. It is also the best performing [5]. The direct-phonon method entails having the extractor, which is the same level as the injector in a three-level system, one Longitudinal Optical (LO) phonon beneath the LLS energy-wise. This is for instance the case in the active region shown in figure 5. As it turns out, LO-phonon scattering is an extremely fast process when possible as detailed in chapter 2.1. Hence, the direct-phonon method results in very fast depopulation of the LLS by the emission of an LO-phonon. On the other hand, the emission of a phonon to transition an electron from the ULS to the injector occurs at a slower rate as the spatial overlap is lower. As such, the lifetime of a state in the LLS is much shorter than in the ULS which results in the desired population inversion between the LLS and the ULS. However, the direct-phonon method has the problem that electrons in the injector can transition back up to the LLS by the absorption of an LO-phonon [15]. This is referred to as thermal backfilling and will be discussed in chapter 2. The problem of thermal backfilling has been solved in mid-infrared QCLs (not considered in this thesis) by adding an additional state one LO-phonon beneath the extractor [8], called double-phonon extraction. Double-phonon extraction is however not currently utilized for THz QCLs.

The active region is then repeated n times which results in a cascading effect where each electron goes through n active regions, thus emitting up to n photons. This has the benefit of reducing the threshold current for lasing J_{th} [1, 15]. In general, a low current is desired as it results in less heating of the device. However, this does not mean that as many modules as possible yield the best performance as the length of each module must be shortened to add more modules to the waveguide, which means that the applied bias must be increased.

The LLS and ULS in figure 5 are not in the same quantum well. Thus a so-called diagonal transition causes the lasing in that structure. Diagonal structures have proven to be the best performing [5, 7, 18]. The reason why diagonal structures are used is to mitigate transitions from the ULS directly to the injector.

Electrical stability requires that there is a positive differential conductance [15]. Negative differential conductance can occur if there is resonant tunneling between two states, and the alignment decreases with increasing bias. A case of negative differential conductance is shown in figure 6. Here when the bias is increased, the alignment of the levels decreases which results in a smaller tunneling rate. As a result, the current decreases with an increasing bias.

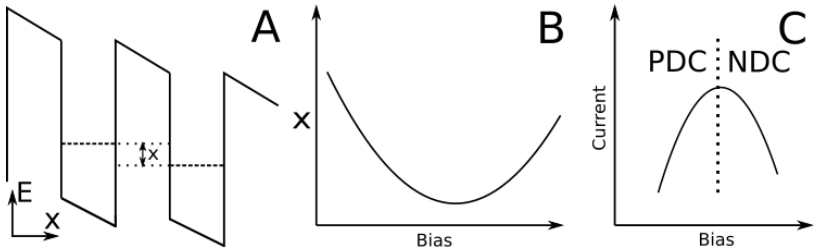

Figure 6: Schematic drawing that shows how negative differential conductance can occur in a QCL which results in electrical instability. In figure A the structure and definition of the energy difference x is shown, in figure B the relation between the absolute value of x and applied bias is shown, and in figure C the resulting IV-curve is shown. The regions of Positive Differential Conductance (PDC) and NDC are also marked.

1.3 Interface Roughness

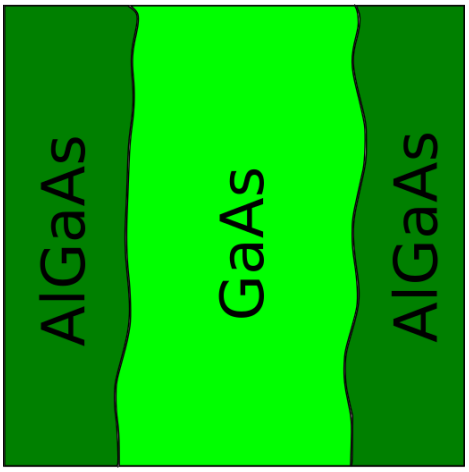

Figure 7: Superlattices have imperfect layer surfaces due to the growth process. This destroys the periodicity and hence causes scattering that is often detrimental to the performance of QCLs. This figure shows schematically how the spatial distribution of the materials could look from a top view.

Interface Roughness (IFR) refers to the imperfect superlattice layers. The interface roughness is thus a result of growth quality and can be seen as fluctuations of the ideally flat layers, as shown in figure 7. In the two studies of Ref. [19] and Ref. [20] the impact of interface roughness scattering was studied and the discussion here is largely based on those two articles. The IFR-scattering is caused by the imperfect layers resulting in a loss of periodicity. What was found in Ref. [19] is that IFR-scattering decreases the gain and increases the current due to the resulting non-radiative transitions from the ULS to the LLS.

Furthermore, in Ref. [20] it was investigated how IFR should be modeled. The fluctuation $\eta(r)$ tells us how far from the perfect surface the actual surface deviates. In Ref. [20], they made the assumption that the position of the fluctuation $\eta(r)$ on the layer surface could be described by different statistical distributions. Therefore, it was assumed that each layer surface looked exactly the same with the same imperfections. The results from the study showed that all tested distributions could be used if the fitting parameters were chosen correctly. For instance, the scattering matrix element squared for the exponential distribution is given by

$$f(\mathbf{q}) = \frac{2\pi\tilde{\Delta}^2\tilde{\Lambda}^2}{(1+\tilde{\Lambda}^2|\mathbf{q}|^2)^{3/2}} \tag{1}$$

where $\mathbf{q}$ is the exchanged momentum and $\tilde{\Delta}$ and $\tilde{\Lambda}$ are fitting parameters with no clear physical interpretations.

It was then found in Ref. [19] where they ran simulations on three samples, that the results were dependent on how the fitting parameters were chosen. For instance, a 50% increase in $\tilde{\Delta}$ resulted in 20% less gain. As a result, it was shown that IFR-scattering indeed plays an important role in THz QCLs.

2 Thermal degradation

Ever since the inception of THz QCLs, the main goal has been to improve the operational temperature of the structures. Still, QCLs operating at room temperature have not yet been achieved. As a consequence, many studies have been doneto investigate why the performance of QCLs degrade with temperature [17, 21–24]. Still, no consensus exists [8, 25]. Later in chapter 4.3, the thermal degradation of three samples will be studied. This chapter should therefore be seen as the background for those tests.

The discrepancy of the findings from these studies might be due to different designs having different degrading mechanisms, and not the same samples were studied in those articles. In chapter 4, the thermal degradation of the new state-of-the-art devices will be investigated. Therefore, the purpose of this chapter is to give an introduction to the detrimental effects temperature has on QCLs. First, some theory about LO-phonons will be given as a background to the Thermally Activated LO-Phonon (TALOP) scattering and the thermal backfilling that follows. Then the chapter ends with a discussion about leakage to the continuum states.

2.1 LO-phonon Scattering

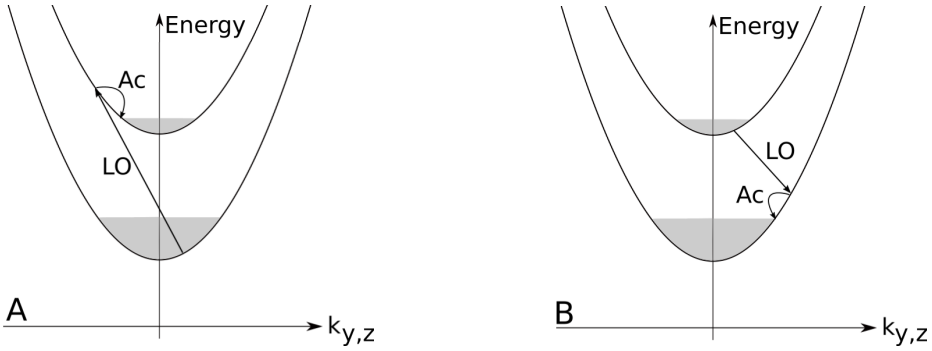

Figure 8: Schematic drawing adapted from Ref. [1] of two subbands where it is assumed that all electrons are relaxed to the lowest possible state. The figure intends to show the difference between a small energy gap in figure A ($E < E_{LO}$) and a large energy gap in figure B (E> E_{LO}) when it comes to the emission and absorption of phonons. The acoustic phonons (Ac in figure) are typically small in energy and results in intrasubband relaxation. The LO-phonons are essentially monoenergetic and can only be emitted from the upper band in the large energy gap-case.

In materials, mechanical vibrations of the atoms can be modeled as quantum quasiparticles called phonons. Phonons are usually categorized into acoustic and optical phonons. If the movement of the atoms in the crystal lattice is coherent they are called acoustic, otherwise they are called optical. Phonons can also be longitudinal, i.e. the atoms are displaced in the same direction as the propagation of the mechanical vibrations, or transversal where the atoms instead are displaced perpendicularly. For III-V semiconductor materials like GaAs LO-phonon scattering is the dominant scattering process. [1]. For instance, the lifetime of a subband is around 1 ps when LO-phonon scattering is energetically possible, and for acoustic phonons, the lifetime is at least 80 ps [26].

As explained in the previous chapter, LO-phonons are of utmost importance when designing a

QCL as they are exploited to get fast relaxation from the LLS. The reason behind this is the monoenergetic nature of LO-phonons since this results in large differences depending on if the energy gap is greater than E_{LO} or not as shown in figure 8. When the energy gap is close to E_{LO} (≈ 36 meV for GaAs), the emission of a photon from the ULS is extremely fast as the scattering rate is proportional to $\frac{1}{q}$ where q is the momentum exchanged in the scattering process [27]. As a consequence, LO-phonon scattering is the dominant scattering process for III-V materials [1], when energetically possible.

2.2 Thermally Activated LO-Phonons

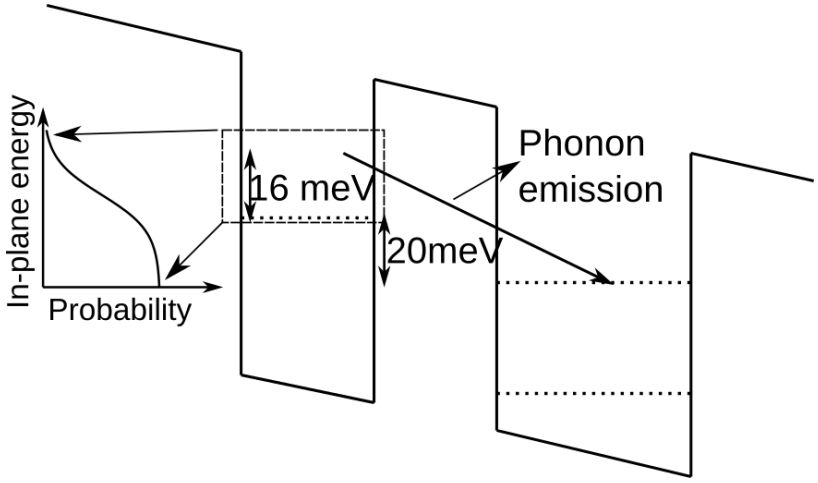

Figure 9: Simplified picture of an active region with three states meant to represent the ULS, the LLS, and the extractor/injector. The figure intends to show how electrons with an in-plane energy greater than 16 meV in the ULS can transition to the LLS by the emission of an LO-phonon if the energy difference between the two states is 20 meV. Furthermore, the higher the temperature, the larger the number of electrons with an energy greater than 16 meV. As such, the rate of LO-phonon scattering from the ULS to the LLS is expected to increase with temperature; it is 'thermally activated'. How the distribution of electrons looks for an arbitrary temperature according to Fermi-Dirac statistics is included in the figure for reference.

As the electrons heat up, their distribution within the subband changes according to Fermi-Dirac statistics, i.e. it is in equilibrium. The distributions, however, change when an external bias is applied. This is due to the electrons gaining energy and thus getting distributed according to a temperature larger than the lattice temperature [8]. The electrons can then lose their excess energy by a multitude of processes including phonon scattering, impurity scattering, IFR-scattering, electron-electron scattering, or by the emission of a photon.

As discussed in the previous section, if LO-phonon scattering is energetically possible then the electrons most likely will lose their excess energy by the emission of a LO-phonon. Hence if the energy difference $\Delta E = E_{\text{up}} - E_{\text{low}}$ between two states $|\text{up}\rangle$ and $|\text{low}\rangle$ is less than E_{LO}, then the electron in the up state needs an *in-plane* energy of $E_{\text{LO}} - \Delta E$. As such, when the temperature is increased more electrons have sufficient in-plane energy to emit an LO-phonon to transition to the lower state. This is referred to as thermally activated LO-phonon emission and causes thermal degradation in QCLs by transitioning electrons from the ULS to the LLS without the emission of a photon. As a consequence, the population inversion is decreased and the current is increased, which

are two detrimental effects. Figure 9 shows how TALOP-scattering works in an active region.

In the literature, different opinions exist on the importance of TALOP-scattering when it comes to the thermal degradation of QCLs. For instance, in Ref. [24] it was found in experimental data that the thermal degradation kicks in at temperatures corresponding to the thermal energy of $E_{LO} - \Delta E$ and therefore, TALOP-scattering is the main mechanism for the degradation of their sample. Experimental analysis in Ref. [23] found the same to be true. However, simulations performed in Ref. [21] found no evidence of TALOP-scattering diminishing the population inversion.

It is not surprising that different samples show signs of TALOP-scattering from the ULS to the LLS as the rate of TALOP-scattering not only depends on energy but also on the spatial overlap. For instance, in Ref. [18] it was found that by increasing the radiation barrier width and in extension increasing the diagonality, the TALOP-scattering between the ULS and the LLS decreased. The same approach has been taken by all recent state-of-the-art devices [5, 7, 28, 29], where high barriers and high diagonality have been utilized to effectively eliminate the detrimental TALOP-scattering.

2.3 Thermal Backfilling

Thermal backfilling is the process of electrons in the extractor absorbing an LO-phonon to transition back to the LLS resulting in less population inversion. The absorption rate is dependent on the LO-phonon population which is only weakly temperature dependent, increasing by roughly 100% when going from 0 K to 400 K [1]. Hence, thermal backfilling is probably only a significant degrading mechanism once other detrimental degrading thermal effects such as TALOP-scattering from the ULS to the LLS are suppressed. Studies on thermal degradation have come to different conclusions about thermal backfilling. For instance Ref. [21] and Ref. [29] found thermal backfilling to be the main mechanism while Ref. [30] and Ref. [31] found it to be negligible. Furthermore, this might explain why the current record-holding QCL [5] has an energy difference E_{ex} between the LLS and the extractor at roughly 46 meV as this would result in less thermal backfilling, albeit at a cost of less effective depopulation of the LLS.

2.4 Leakage Current

Leakage current refers to electrons that get excited up from the lower three states to the Higher Lying States (HLS) or even the continuum over the barriers. This is a detrimental effect as this current heats up the device without emitting any photons. There are mainly two paths for the electrons to get excited this way: by tunneling to the HLS in the next module or by the absorption of an LO-phonon to transition to the HLS in the same module. Both of these paths are shown in figure 10. Only the absorption of LO-phonons should have a temperature dependence which is weak, as in the case with thermal backfilling.

Recently, strides have been made to get rid of the leakage to the continuum by using higher barriers. In Ref. [28] they analyzed a highly diagonal $Al_{0.15}Ga_{0.85}As$ structure as to eliminate TALOP-scattering from the ULS to the LLS, and they found that thermally activated leakage channels were the main cause for the thermal degradation. In Ref. [32] the same authors studied a highly diagonal structure again, but this time using an $Al_{0.3}Ga_{0.7}As$, i.e. significantly higher barriers. What they found is that NDC was kept when using high barriers, i.e. leakage to the continuum had been suppressed as it usually counteracts the expected NDC. They found, however, that TALOP-scattering to the HLS still was present in this high barrier sample. In Ref. [28] these principles were used to achieve record high-temperature performance of a 2-QW THz QCL. In Ref. [29] and then in Ref. [7]

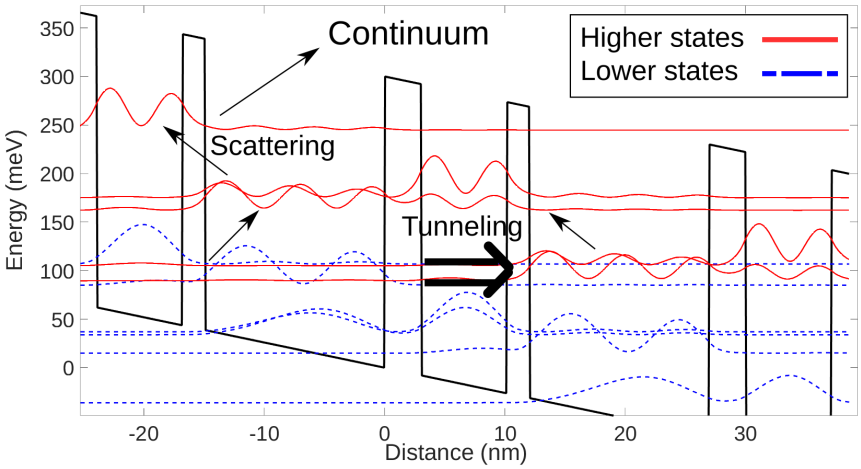

Figure 10: Figure showing the two main paths for leakage current to HLS and in extension leakage to the continuum. The electrons can tunnel from the lower states to the HLS in the next module or by scattering to the HLS in the same module. Roughly two modules are shown in the figure. However, not all states are included in the figure to make it more readable.

the temperature record for 2-QW THz QCLs was once again beaten, using a similar highly diagonal high barrier structure. The authors of those articles, however, found that thermal backfilling and reabsorption were the main causes for the thermal degradation by simulating the structure. Thus, it would seem like the leakage currents presented in those articles had been suppressed. However, in Ref. [5], they beat the record and still have it as of writing, and the authors of that article attributed their better performance compared to Ref. [29] and [7] to better suppressing the leakage currents. This overview of the literature about 2-QW THz QCLs from the period 2015-2021 shows two things. Firstly, the performance has improved by suppressing leakage to HLS and the continuum. Secondly, the remaining mechanisms that are relevant for thermal degradation are not agreed upon.

3 The NEGF-Package

In this chapter, the NEGF-package will be introduced. The chapter is split into two parts. In the first part, the underlying mathematical models will be described briefly. The intent behind this is to give the reader an understanding of what the program actually calculates. In the second part, it is discussed how the NEGF-package is used to study QCLs. The purpose behind this is two-fold. First, it will effectively function as a method part for the results part of chapter 4. Secondly, it will explain how the NEGF-package is structured.

3.1 Modelling

3.1.1 Hamiltonian and Eigenstates

In this section the relevant Hamiltonian used to model the QCL in the NEGF-package will be discussed. The discussion is largely based on chapters 2 & 3 of the thesis of Franckié [12].

The Hamiltonian is given by

$$H = H_0 + V_{dc} + V_{ac}(t) + H_{scatt} \tag{2}$$

where $H_0 = \frac{p^2}{2m_e} + V_L(z)$ where $V_L(z)$ is the potential of the superlattice. Furthermore, V_{dc} is the potential from the applied bias over the structure, $V_{ac}(t)$ is the time dependent potential caused by a electromagnetic light field, and H_{scatt} contains the scattering mechanisms discussed in chapter 2. The complicated scattering mechanisms H_{scatt} is omitted when calculating the eigenstates which leaves us with the Hamiltonian for a single particle in an electromagnetic field in a superlattice

$$H_{sp} = H_0 + V_{dc} + V_{ac}(t) = \frac{(p - eA(r,t))^2}{2m_0} + V_L(z) + e\Phi(r,t). \tag{3}$$

The scattering mechanisms are then entered into the system by the self-energies, using the self-consisten Born approximation [10].

To get the eigenstates of eq. 3, the eigenstates are expanded in envelope functions ψ_n, such that the single electron state becomes

$$\Psi(r,t) = \sum_n \psi_n(r,t) u_{n,k0}(x) \tag{4}$$

where n denotes bands (not to be confused with the subbands of the QCL) such as the conduction band and valence band, $u_{n,k0}$ are the periodic Bloch functions of the periodic structure, and x is the growth direction. In the NEGF-package, two bands at are considered for the envelope functions for a fixed Bloch vector with $k = k_0$: the conduction band and the valence band. This is called the two-band model, and is widely used for the modeling of QCLs [1]. As it turns out, if eq. 4 is applied on the Hamiltonian eq. 3, you get an effective Schrödinger equation for the envelope function.

In the two band model, the envelope function can be written as

$$\bar{\psi}(r,t) = \begin{pmatrix} \psi_c(r,t) \\ \psi_v(r,t) \end{pmatrix} \tag{5}$$

where ψ_c and ψ_c is the envelope function for the conduction band and valence band respectively. Inserting this into eq. 3 and eq. 4, while neglecting any electromagnetic fields, results in two closed

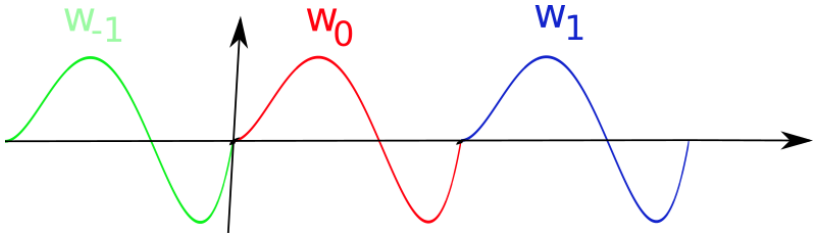

Figure 11: Schematic drawing illustrating how Wannier states are localised in space and repeated once in each region. A Bloch state would look like the sum of the Wannier states.

expressions for ψ_c and ψ_v

$$E_c\psi_c + p_z \frac{1}{2m_c(E,z)} p_z\psi_c = E\psi_c, \tag{6}$$

$$E_v\psi_v + p_z \frac{1}{2m_v(E,z)} p_z\psi_v = E\psi_v, \tag{7}$$

where E_c is the energy of the conduction band, E_v is the energy of the valence band, and E is the energy of the state ψ. Full calculations can be found in Ref. [12] and Ref. [1]. However, these envelope functions pose the problem of not being localised in space. As such, the writers of the NEGF-package choose to use Wannier-states [10, 12, 33]. The Wannier states are an alternative basis for the Hamiltonian and can be chosen in different ways. As such, several methods exist to change the basis, but the one used in the NEGF-package is to find the basis that diagonalizes the position matrix element

$$z_{\alpha,\beta} = \int \mathrm{d}z \phi^*_\alpha z \phi_\beta. \tag{8}$$

where ϕ_α and ϕ_β are Wannier states. The Wannier states are localized in space and are repeated each period as shown in figure 11 but do not take the external bias into account. Therefore, the energy values can be unreliable when you want to know the energy levels under an applied bias. To solve this, the NEGF-package can also calculate the Wannier-Stark states, e.g. when you want to plot the band structure, more on this in chapter 3.2. The Wannier-Stark states are the mix of Wannier states (perturbation) that you get when you apply an external bias on the Wannier states.

Now that we have a basis for the Hamiltonian, it is straightforward to calculate the current density J through the QCL, which is defined as

$$J(z,t) = \left\langle \frac{e}{V} \frac{\partial z}{\partial t} \right\rangle. \tag{9}$$

However, to evaluate eq 9, the density matrix is required. It is in this part that the NEGF-package utilizes the NEGFs it is named after. The theory behind NEGFs is mathematically advanced and thus outside the scope of this text. Nonetheless, the reason why the NEGF-method yields better results than less sophisticated simulation methods is that it takes the line broadening caused by the non-equilibrium distribution of electrons within a subband into account. The NEGF-method, therefore, excels in calculating the coherences, i.e. the non-diagonal elements in the density matrix.

3.1.2 Optical Response

In this chapter the optical response (specifically the gain) of an active region of a QCL will be discussed. The discussion is largely based on chapter 2 of the thesis of Winge [11].

The polarisation $\mathbf{P}$, which can be assumed to be constant over one active region due to the wavelengths of the electric field being many times larger than the length of the active region, describes the response of a system that is subjected to an external electric field $\mathbf{E}$, and is given by

$$\mathbf{P}(t) = \epsilon_r \epsilon_0 \int_{-\infty}^{t} \mathrm{d}t' \chi^{rel}(t-t')\mathbf{E}(t') \tag{10}$$

where χ^{rel} is the relative susceptibility that describes the history of the system. The expression becomes more useful when it is Fourier transformed into the frequency space

$$\mathbf{P}(\omega) = \epsilon_r \epsilon_0 \chi^{rel} \mathbf{E}(\omega). \tag{11}$$

The complex refractive index can be written as $\tilde{n}(\omega) = c\tilde{k}/\omega$ [34], where $\tilde{k}$ is the complex wave vector. If we write $\tilde{k} = k + i\alpha/2$ and plug it into the electric field traveling along x

$$\mathbf{E}(x) = \hat{\mathbf{z}} F_{\mathrm{ac}} e^{i\tilde{k}\cdot x}, \tag{12}$$

it is easy to see that α corresponds to absorption per unit length. Furthermore, it is well known that $\tilde{n} = \sqrt{\epsilon(\omega)/\epsilon_0}$, and as a result $\alpha = \omega \,\mathrm{Im}\{\epsilon(\omega)\}/(\epsilon_0 c)$. Finally, we have the definition of the dielectric field tensor $\epsilon(\omega) = \epsilon_0 \epsilon_r (1 + \chi^{rel}(\omega))$ which yields the expression we will use

$$\alpha = \frac{\omega\sqrt{\epsilon_r}}{c} \,\mathrm{Im}\{\chi^{rel}(\omega)\} \tag{13}$$

From this point it would be possible to calculate the susceptibility by using time-dependent perturbation theory. However, the NEGF-package expands the current frequency, following Ref. [35], to get $\mathbf{J}(\omega)$ and can therefore relate it to the susceptibility to the current by the expression

$$\mathbf{J}(\omega) = \sigma(\omega)\mathbf{E}(\omega) = -i\omega\mathbf{P}(\omega) \tag{14}$$

which results in

$$\alpha = -\frac{1}{c\epsilon_o\sqrt{\epsilon_r}} \,\mathrm{Im}(\sigma(\omega)). \tag{15}$$

The gain $G(\omega) = -\alpha$ can thus be extracted directly from the frequency-dependent current.

3.2 Simulations Using the NEGF-Package

In this chapter, the different subprograms in the NEGF-package will be presented as well as how to use them to get quantitative results.

3.2.1 Wannier States

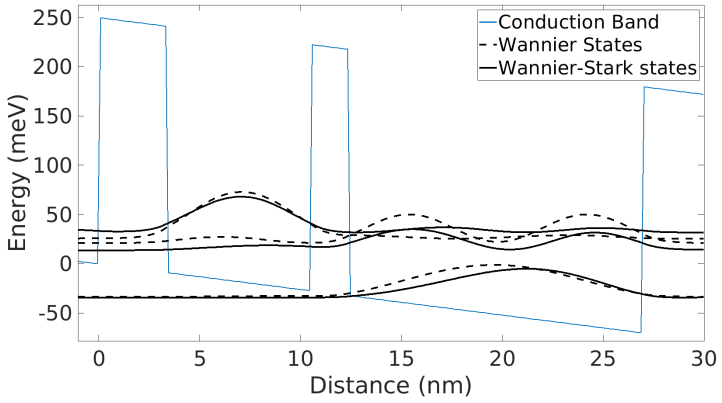

Figure 12: The band structure of a sample generated with the bandplot program. In the plot, there are also three Wannier states calculated using the Wannier8 program as well as three Wannier-Stark states calculated using the calcWS program.

The first step when simulating a device with the NEGF-package is to calculate the Wannier states. The program used for this is called Wannier8 and works by writing all the simulation parameters in a text file. These simulation parameters include the studied sample's CBO, layer sequence, aluminum content in the barriers, effective mass of the electron in the well and in the barrier, the doping levels, and the IFR parameters.

Once you have the Wannier states, you can plot the band structure using the bandplot program. However, often it is more interesting to plot the Wannier-Stark states since the Wannier states do not take the external bias into account. Therefore a subprogram called calcWS exists that estimates the Wannier-Stark states by giving the applied bias as an input. The calcWS program leaves out the mean-field and is therefore not as exact as the hdiag program described below. In figure 12 the bandstructure of a sample is shown that was generated using the bandplot program. In the figure, three calculated Wannier states and three Wannier-Stark states are plotted to show the difference between them. The energy difference between the two upper-most Wannier-Stark states is greater than the two upper-most Wannier states due to the introduced coupling between the states.

3.2.2 Current Densities

When the Wannier-states have been generated it is possible to use the negft8 program which can be seen as the main program of the NEGF-package. There are in essence three things that can be simulated with the negft8 program: IV-curves without lasing, gain, and IV-curves with lasing. In this subchapter we will go over how this is done and also some new theory about gain saturation will be introduced in order to explain IV-curves under lasing conditions.

First, IV-curves without lasing are simulated by entering at which intervals the current is calculated. Other important simulation parameters that are entered are the phonon temperature T, the amount of Wannier states N_{nu}, and how many neighboring periods each state interacts with N_{per}. Increasing N_{nu} and N_{per} increases in general the accuracy at the expense of computing time as the number of states to be calculated is $N = N_{\mathrm{nu}}(1 + 2N_{\mathrm{per}})$ [12]. As such, in general you want to use

as few Wannier states as you can while getting acceptable results, as this can save weeks of time.

Simulating the same structure under the same conditions using a different N_{nu} gives insight into how important HLS states are for the operation of that sample. For instance, if the simulated maximum current increases drastically when going from three to five states, then it is evident that current through the HLS is present. By the same token, if the results change when going from $N_{\mathrm{per}} = 1$ to $N_{\mathrm{per}} = 2$, then it is evident that intermodular tunneling through two periods plays an important role in the sample.

When simulating the gain it is necessary to enter which photon energies $h\omega$ the gain should be calculated for, the electric field strength of the oscillating light field F_{ac} and the number of harmonics used n_h, in addition to the simulation parameters needed to simulate the IV-curve under no lasing conditions. To simulate the maximum gain of a sample we thus simulate the gain in an interval of the applied external bias F_{dc} and in an interval of $h\omega$. Then by using the curve-fitting tool in MATLAB, a Lorentzian curve shape is fitted onto the data for each F_{dc}, as the gain spectra generally have a Lorentzian shape [1, 15]. However, here NDC might complicate things as if the max gain is in a region of NDC, then the max gain is at an unstable point and thus experimentally uninteresting. Therefore, the interval of F_{dc} is restricted to the upper bound of F_{dc} at the maximum current $F_{\mathrm{dc}}^{\mathrm{max}}$ plus 4 mV/module. The reasoning behind this is that even though we expect NDC after $F_{\mathrm{dc}}^{\mathrm{max}}$ in the simulations, in the experiments leakage to the continuum will counteract the drop in current at this point.

Finally, calculating the IV-curve under *lasing conditions* is a bit more complicated and requires some iterative process. This is due to the alternating field giving rise to stimulated emission from the ULS to the LLS. Three effects follow: increased current since the rate of transitions increase, a larger F_{ac}-field due to more photons in the system, and less population inversion. As such, if there was no optical loss in the medium then F_{ac} would increase until the population inversion was gone and equilibrium would be met. In reality, there is a waveguide loss α that is typically around 20 cm^{-1}. So, to calculate the current under lasing conditions, the program follows the following algorithm for each F_{dc}:

1. For the initial F_{ac}, calculate the gain for a given interval of photon energies.
2. Fit a lorentzian curve to the gain spectrum to calculate the maximum gain.
3. If the maximum gain is higher/lower than α, increase/decrease F_{ac}.
4. If the maximum gain still higher/lower than α, increase/decrease F_{ac} until it is lower/higher.
5. Once you have a point where $G > \alpha$ and a point where $G < \alpha$, use interpolation to find the point where $G = \alpha$ extract the current at that point.
6. Repeat for the next F_{dc}

One relevant quantity that is often included in scientific papers about QCLs is the threshold current for lasing J_{th}. This current can be estimated by simply simulating the IV-curve under lasing conditions and then extracting the point with the lowest current where $G > \alpha$. However, this poses the problem that the temperature entered into the simulation parameters is the *LO-phonon temperature*, and the temperature given in experimental data is the *heatsink temperature*. The heatsink temperature is normally much lower than the LO-phonon temperature which corresponds to the temperature inside the sample, whereas the heatsink temperature corresponds to the temperature on the boundary of the sample. As a result, the difference in temperature between simulated results

and experimental results can vary drastically. In addition, J_{th} has the strong empirical temperature dependence

$$J_{\text{th}} = a + be^{T/c} \tag{16}$$

where a,b and c are fitting parameters. It can therefore be difficult to compare the simulated J_{th} to the experimental one as e.g. the assumption $T_{\text{Heatsink}} = T_{\text{LO-phonon}} - 100$ K will yield a much lower J_{th} than the assumption $T_{\text{Heatsink}} = T_{\text{LO-phonon}} - 50$ K

3.2.3 Resolved Densities

After the negft8 program has been run, it is possible to use a program called hdiag which calculates the Wannier-Stark states including the mean-field (as opposed to calcWS). With these Wannier-Stark states, a program called resolve can be used which generates the electron and current densities as a function of energy and distance. This information can then be used to get information about the population densities of each state and thus also the population inversion. Furthermore, it can give insight into how much current flows through each level, and as a consequence, it can be used to deduce the magnitude of the leakage to HLS.

4 Simulation Results

In this chapter, the results from the simulations are presented. In part one, we test which simulation parameters should be used. Then in part two, these parameters were used to simulate and compare the operation of different samples. In the third part, we investigate what causes the thermal degradation of the samples. The chapter then rounds off by introducing a new sample, optimized by using what was learned in the first parts of the chapter. Important to note here is that the temperature used in the simulations is the LO-phonon temperature, while the temperature from experimental results is the heatsink temperature. The LO-phonon temperature of an operating device is higher than the heatsink temperature and we assume that the LO-phonon temperature is roughly 50-100 K higher than the heatsink temperature, as it was found in Ref. [29]. The reason for this is that the inside of the QCL is hotter than the cooler. Furthermore, it is important to note that x throughout this section refers to the aluminum content of the $Al_xGa_{1-x}As$ barriers.

4.1 Conduction Band Offset

Knowing the Conduction Band Offset (CBO) is crucial for accurate simulations since many important properties such as tunneling rate and energy levels depend on it. The CBO refers to which expression should be used to define the barrier height as a function of the aluminum content x. Hence, it is of great interest that no consensus exists on what value to use [36]. For instance, two widely used values for the CBO are $E_c = 0.97x$ eV suggested by Vurgaftman *et al.* [37] and $E_c = 0.831x$ eV suggested by Yi *et al.* [38]. However, the latter is more often used for QCL simulations, including the authors of the NEFG-package [36, 39].

In a study from 2016 Winge *et al.* [36] tested samples from several different labs using the NEGF-package. In the simulations they used the widely accepted effective mass of $m_e = 0.067 + 0.083x$ and Kane energy $E_p = 22.7$ eV, but also the same CBO of $E_c = 0.831x$ eV for all simulations. The results of the study showed clearly that the calculations either overestimate, underestimate or correctly estimate the current, depending on which lab the sample was manufactured in. Consequently, if two structures that are identical on paper were to be grown by two different labs, the outcome would not be the same. Therefore different definitions of material parameters and/or calibration of molecular beam epitaxy growth could be the cause. The two most likely candidates are either the CBO or the IFR parameters.

Firstly, the impact of the IFR parameters was studied extensively in two separate studies by Franckié *et al.* [20] and Krivas *et al.* [19]. In these studies, they concluded that changing the IFR parameters alters the current since IFR scattering results in non-radiative transitions from the ULS to the LLS, as explained in chapter 3.1. This also increases the current at the expense of gain as non-radiative transitions to the LLS increase in rate. However, it was found in Ref. [36] that IFR parameters can not on their own account for the discrepancy between samples from different labs, which means that the actual barriers should differ between different groups for the same nominal parameters, as reflected by the use of different values for the CBO.

It is important to understand why different CBOs gives the most accurate result for different structures. In communication with my supervisor [39], he explained that one reason could be due to different groups defining the width of a barrier differently, shown schematically in figure 13. And even though the barriers are modeled as perfect step functions, there is a finite gradient between the well and the barrier. Thus different heights have to be used to model the imperfect barrier in figure 13, depending on if the width is defined as width 1 or width 2. This follows from that increasing

the barrier width has a similar result to increasing the barrier height. Another reason could be the nominal x-content differing between different groups. It should be stressed that no one truly knows what is going on and this is just speculation.

In this section I therefore investigate which CBO yields the best results for two samples from ETH and three samples from Sandia/MIT. All simulations in this subsection use a lattice temperature of 100K, waveguide loss of 20/cm and $N_{\text{nu}} = 5$, $N_{\text{per}} = 2$.

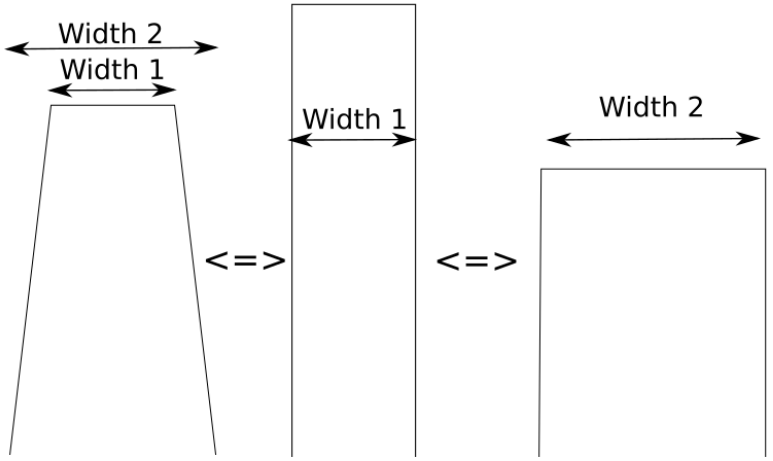

Figure 13: Schematic figure highlighting two different possible definitions of barrier width. The leftmost barrier represents the actual barrier, while the two rightmost barriers show that the height must be adjusted depending on which definition of width being used.

4.1.1 ETH Samples

The sample ScalariOE2010 was presented in Ref. [40] and has a RT-injection 2 QW design. It has a relatively low barrier of $x = 0.15$, a maximum operating temperature $T_{\text{max}} = 125$ K, and a J_{max} of roughly 800 A/cm^2. More detailed device properties are listed in table 1. Firstly, the current was simulated without lasing for three different CBO: $0.97x$ to represent a high barrier, $0.83x$ to represent a medium barrier, and $0.7x$ to represent a low barrier. The results are shown in figure 14. As expected, lower barriers results in higher current due to higher tunneling rate and more coupling between interwell states. Furthermore, it is safe to conclude that $E_c = 0.97x$ eV is too high as we do not expect such large increase in current when a lasing field is applied. Hence, J_{max} under lasing condition was only simulated with the CBOs $E_c = 0.7x$ eV and $E_c = 0.831x$ eV. J_{max} under lasing conditions was determined to be 750 A/cm^2 for the former and 1000 A/cm^2 for the latter. For both CBOs J_{th} was heavily overestimated, probably due to the high temperature dependence of the sample together with a simulation lattice temperature higher than the experimental one. However, J_{th} was much closer for $E_c = 0.831x$ eV as the simulated J_{th} for $E_c = 0.7x$ eV was more than 800 A/cm^2. So for the first sample, $E_c = 0.831x$ eV is the best choice.

In general, only small changes could be seen between different gain spectra with different CBO. In figure 15 two simulated gain spectra from ScalareOE2010 with two different CBO are shown as a reference. In general, the photon energy which yields the maximum gain for all CBOs was the same, albeit with a different value. More important is what F_{dc} we are looking at as the gain spectrum for a device could change considerably for a small change in F_{dc}. Furthermore, as it is hard to know the exact bias/module that was present in the experiment, it becomes difficult to compare simulated

gain spectra to experimental results. Hence, gain spectra will be omitted when assessing which CBO produces the most accurate results, and only IV-curves will be considered.

Wafer	CBO	$J_{\max}^{\text{sim}}$	$J_{\max}^{\text{exp}}$	$J_{\text{th}}^{\text{sim}}$	$J_{\text{th}}^{\text{exp}}$	$T_{\max}^{\text{exp}}$	x	Layer sequence
ScalariOE2010 [40]	0.831x	0.75	0.8	0.5	0.25	125	0.15	**45**/83/**38**/<u>179</u>
BoscoAPL2019 [7]	0.831x	3.7	3.5	2.4	2.0	210	0.25	**32.6**/79.9/**19.0**/<u>164.6</u>

Table 1: Simulated results using the most accurate CBO. CBO is given in eV, current densities in kA/cm^2, temperatures in K, and layer sequences in Å with bold values being Al_xGa_xAs barriers, and the underlined well is doped in a 30 Å region in the center. Experimental values are taken from their references respectively. Simulated values were obtained using the NEGFT-package with $T_{\text{L}} = 100K$, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$ and $\alpha = 20$/cm.

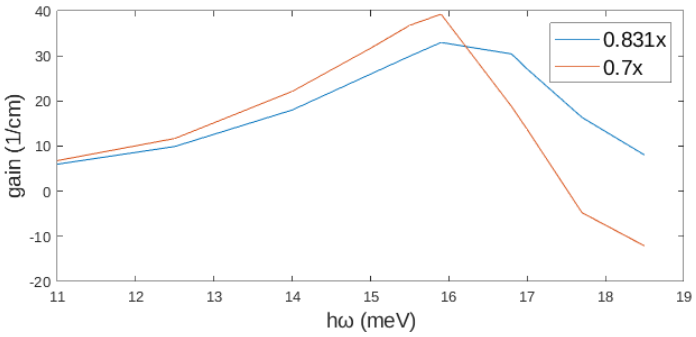

Figure 15: Gain as a function of h obtained using the NEGF-package with $F_{\text{ac}} = 52$mV/period, $T_{\text{L}} = 100$ K, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$ for two different CBO. The simulated structure is ScalariOE2010.

The second sample studied, in this text called BoscoAPL2019, was presented in Ref. [7] and has also a 2-QW RT-injection design. BoscoAPL2019 has an aluminum content of $x = 0.25$, $T_{\max} = 210$ K and $J_{\max} = 3.5$ kA/cm^2. More detailed properties are presented in table 1. For this sample, the typical CBO $E_c = 0.831x$ eV was used together with the CBO $1.01x$ eV to compare to the results

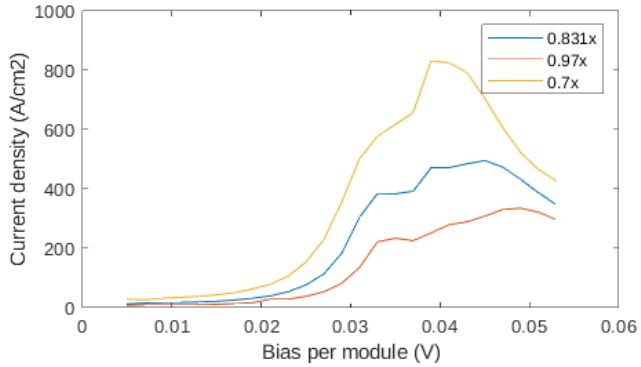

Figure 14: Non-lasing IV-curves obtained using the NEGFT-package using $T_{\text{L}} = 100$ K, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$ for three different CBO. The red line for $E_c = 0.97x$ eV for a high barrier, the blue line for $E_c = 0.831x$ eV for a medium barrier height, and the yellow line $E_c = 0.7x$ eV for a low barrier height.

from the MIT samples, see below. The results are shown in figure 16. Also here $E_c = 0.831x$ eV produced results closer to the experimental results with a simulated $J_{\max}$ of 3.7 kA/cm^2. On the other hand, $E_c = 1.01x$ eV resulted in a $J_{\max}$ of 2.8 kA/cm^2.

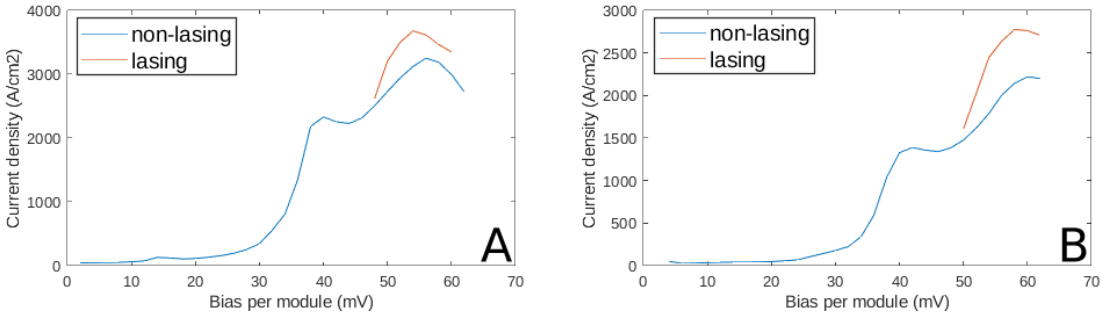

Figure 16: IV-curves for BoscoApl2019 obtained using the NEGFT-package for two different CBO. The blue line represents current with lasing turned off and the red line represents current density with lasing turned on. Figure A shows the low barrier case with $E_c = 0.831x$ eV with a max current density of 3.7 kA/cm^2 and figure B shows the high barrier case with $E_c = 1.01x$ eV with a max current density of 2.8 kA/cm^2. Simulation parameters used was $T_{\mathrm{L}} = 100$ K, $\alpha = 20$/cm, $N_{\mathrm{nu}} = 5$ and $N_{\mathrm{per}} = 1$. $E_c = 0.831x$ eV yielded simulation results closer to the experimental values.

4.1.2 MIT Samples

The devices G552, G605 and G652 were presented in 2021 by Khalatpour *et al.* [5] and are as of writing the devices with the highest operational temperature. In chapter 4.2 we will have a closer look at their operation while we in this chapter focus solely on the CBO. The CBOs that will be tested are $E_c = 1.01x$ eV as it was used by the authors to get the most accurate results, together with the conventional $E_c = 0.831x$ eV. Device properties are summarized in table 2. All wafers have $x = 0.3$.

Wafer	$J_{\max}^{\mathrm{exp}}$	$J_{\mathrm{th}}^{\mathrm{exp}}$	$T_{\max}$	Layer sequence
G552	2.67	1.48	235	**33.0**/76.3/**17.5**/153.8
G605	2.87	1.75	225	**31.4**/76.3/**17.5**/153.1
G652	2.60	1.54	250	**33.7**/72.0/**18.7**/144.9

Table 2: Experimental results taken from Ref. [5]. Current densities are given in kA/cm^2, temperatures in K, and layer sequences in Å with bold values being $Al_{0.3}Ga_{0.7}As$ barriers, and the underlined well is doped in a 30 Å region in the center.

First, simulations of $J_{\max}$ under lasing conditions was run using $E_c = 0.831x$ eV. This yielded $J_{\max}$ values 30% too large on average and J_{th}'s 60% to large on average. The simulated values are shown in table 3. The overestimated currents suggest that the barrier height is too high and the relatively larger overestimation of J_{th} is probably due to the discrepancy between experimental and simulation temperature.

Wafer	CBO	$J_{\max}^{\mathrm{sim}}$	$J_{\mathrm{th}}^{\mathrm{sim}}$	Max gain
G552	0.831x	3.2	2.0	116
G605	0.831x	3.6	2.4	113
G652	0.831x	3.5	2.5	102

Table 3: CBO is given in eV, current densities in kA/cm^2, temperatures in K, and gain in 1/cm. Simulated values were obtained using the NEGFT-package with $T_{\mathrm{L}} = 100K$, $N_{\mathrm{nu}} = 5$, $N_{\mathrm{per}} = 1$ and $\alpha = 20$/cm

Then the same simulations were run using $E_c = 1.01x$ eV. The results were in excellent agreement with the experimental results, at most differing by 10%. There is less accuracy however for J_{th} as expected. The simulated values are shown in table 4. As such, it is evident that $1.01x$ yields substantially more accurate results for all three MIT samples.

Wafer	CBO	$J_{\text{max}}^{\text{sim}}$	$J_{\text{th}}^{\text{sim}}$	Max gain
G552	1.01x	2.4	1.3	101
G605	1.01x	2.7	1.5	102
G652	1.01x	2.6	1.5	71

Table 4: CBO is given in eV, current densities in kA/cm^2, temperatures in K, and gain in 1/cm. Simulated values were obtained using the NEGFT-package with $T_{\text{L}} = 100K$, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$ and $\alpha = 20/\text{cm}$

The current is overestimated for $E_c = 0.831x$ eV and this could be due to the IFR parameters being too large. It would also result in an underestimation of gain, which the heavily overestimated J_{th} is a sign of. However, J_{th} was simulated with great accuracy for $E_c = 1.01x$ eV and since the IFR should decrease with barrier height, the IFR should be underestimated for $E_c = 0.831x$ eV. The max gain was simulated for both CBOs to further investigate this which yielded a higher maximum gain for $E_c = 0.831x$ eV, which is more evidence that it is the CBO that is the most crucial parameter here, and not the IFR. This validates the approach taken in this thesis where only the CBO is fitted on to the sample.

4.1.3 Conclusion

The simulations agreed tremendously well with the experimental data provided that the CBO $E_c = 0.831x$ eV was used for the ETH samples and $E_c = 1.01x$ eV for the MIT samples and will thus be used in the proceeding simulations. This is consistent with previous studies where the NEGF-package has overestimated currents from MIT when using $E_c = 0.831x$ eV. Furthermore, it is clear that inconsistently defined CBO is the reason why the NEGF package has failed to produce accurate results for samples from MIT.

4.2 Simulation of the Best Performing QCLs

In this part the already introduced samples G552, G605 and G652 will be studied in depth using the NEGF-package with the most fitting CBO found in the previous section. The layer sequence and experimental values are presented in table 2. These three QCLs are currently the three with the highest operational temperatures. As such, simulating the structures might give insight into how these samples operate and why they managed to achieve such high operational temperatures.

The authors behind Ref. [5] aimed to improve the 'direct phonon-extraction 2-QW QCL' with RT-injection as these were already the best performing designs. They argued that the previous designs fail at high temperatures due to carrier leakage channels and they aim to minimize these with their design choices. They did so by increasing the energy separation $E_{i,pp}$ between the injector state and parasitic states, by increasing the anticrossing $\Omega_{u,i}$ to decrease the coupling from the three-level system to the HLS, and by optimizing channels such as $|u_n\rangle$ to $|p_{2,n+2}\rangle$ and $|l_n\rangle$ to $|p_{1,n+1}\rangle$. These and other relevant states are shown in figure 17 where the band structure of G652 is shown.

Experimentally, G652 had the highest operating temperature followed by G552 and then G605 at 250K, 235K and 225K respectively. Ref. [5] attributed the good performance of G652 to its low coupling to HLS quantified by a low oscillator strength $f_{i,pp}$, even though G552 performed better

than G605 while having a larger $f_{i,pp}$ than G605. The reason for the latter, they argue, is that G605 has a small anticrossing gap between $|l_n\rangle$ and $|p_{1,n+1}\rangle$ resulting in a leakage channel to the HLS. It is also mentioned that the growth quality is worse for G605 and this might play a role.

To get further insight into the operation of these structures, simulations under lasing conditions were made using $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$, $\alpha = 20/\text{cm}$, $E_c = 1.01x$ eV at various T_{L}. G652 will first be compared to G552 as they have the same injection barrier and well but differs in radiation barrier. Then G605 and G552 will be compared as they only differ in injection barrier width. Lastly, simulations will be made with $N_{\text{nu}} = 7$ to investigate wheter more states is necessary to capture the relevant physics of these QCLs, as well as simulations with $N_{\text{per}} = 2$ to analyse the leakage channels across one module.

4.2.1 G552 vs G652

The active region design of G652 is different from the active region designs of G605 and G552, which are more similar. G652 differs by having a wider radiation barrier and radiation well but has roughly the same injection barrier width as G552. Hence a comparison between G652 and G552 can give insight into how the radiation barrier width and radiation well width should be chosen for a preferred performance. How the energy levels are affected by these widths can be explained with the help of figure 18, where a simplified version of the active region is shown. Here E_1 is the

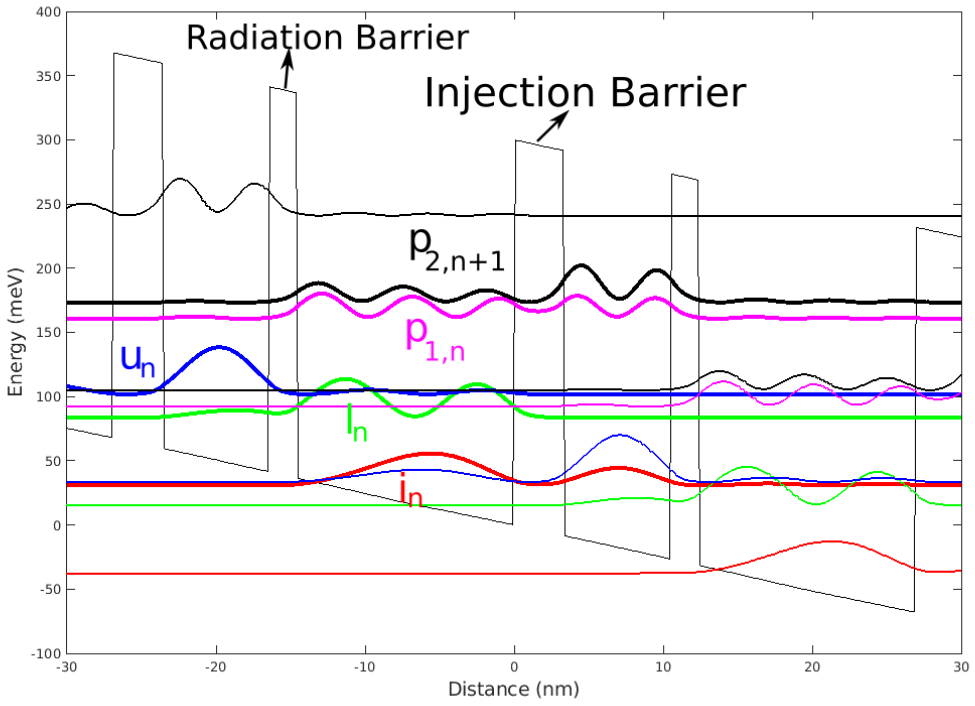

Figure 17: Band structure of G652 together with probability density functions given by Wannier-Stark states for two modules. The states are color coded with the injector state being red, LLS being green, ULS blue, first parasitic state being magenta, and the second parasitic state being black. The bold lines are denoted with the name of the states.

energy difference between the LLS l and the injector i, and E_2 is the energy difference between the ULS u and the HLS p_1 and p_2. Increasing the well width W has a similar effect to decreasing the barrier width B, which is decreasing E_1 and E_2. The reason why B affects the energy levels is that the spatial overlap and thus the coupling between the states is altered by the thickness of the barrier.

However, changing B has roughly the same effect on E_1 and E_2 while W mostly changes E_2. As a result, the larger widths of G652 results in a 13 % greater E_1 and a 17 % greater E_2 compared to G552. Thereby, for G652 more thermal energy is required for LO-phonon scattering to the parasitic states while the spatial overlap between the lower states and the HLS is lower. Although, this is at the expense of less tunneling injection because of the thicker radiation barrier, as well as less effective depopulation of LLS as E_1 is further from the LO-phonon energy.

In table 5 the results from the simulations are shown. By using max gain as a performance metric the conclusion is that the performance of G552 is better at lower temperatures, probably due to the more effective depopulation of the LLS. At higher temperatures, the performance of G552 is degrading faster than G652 as indicated by the higher $T_{\max}$ of G652 together with a similar max gain at 285K. This could be due to either TALOP-absorption between $|i\rangle$ and $|l\rangle$ with the energy separation E_1, or TALOP-absorption between $|u\rangle$ and $|p_1\rangle$ with the energy separation E_2, or both. A more detailed analysis of the temperature degradation is given in chapter 4.3.

Compared to the experimental results, the difference in $T_{\max}$ between the two wafers shows excellent agreement. Furthermore, $J_{\max}$ at 100K is also predicted with high accuracy, albeit somewhat underestimated for G552 which might be due to the low number of Wannier states being used. At 285K $J_{\max}$ is underestimated for both wafers, which is probably due to leakage to the continuum not being taken into account in the simulations.

Wafer	G552	G652
E_1	46	52
E_2	53	62
Max gain (100K)	101	93
Max gain (285K)	26	26
$T_{\max}^{\text{sim}}$	320	335
$T_{\max}^{\text{exp}}$	235	250
$J_{\max}^{\text{sim}}$(100K)	2.37	2.64
$J_{\max}^{\text{sim}}$(285K)	2.04	2.32
$J_{\max}^{\text{exp}}$	2.67	2.6

Table 5: Simulated results using the NEGF-package using a $E_c = 1.01x$ eV, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$, $\alpha = 20$/cm. Energies are given in meV, gain in 1/cm and current density in (kA/cm^2). The definitions of E_1 and E_2 are given in figure 18.

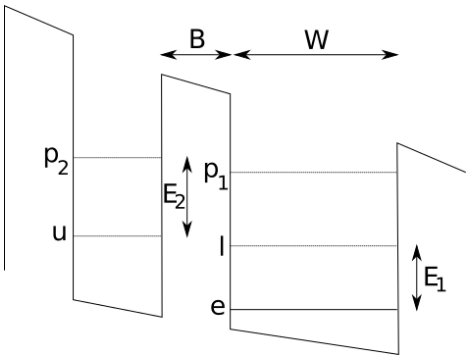

Figure 18: Simplified drawing of the active region meant to show the important states, energy separations and widths discussed in chapter 7.2.1.

4.2.2 G605 vs G552

The only major difference between G552 and G605 is the injection barrier width. Hence, a comparison between these two structures presents the possibility to investigate the tradeof between RT injection strength and coupling to leakage channels. The simulated results are presented in table 6. The population difference between the LLS and ULS for both structures are similar at 100K and

285K. However, the current is higher for G605 as a result of the thinner injection barrier. Therefore it seems like that the simulation predicts that a thinner injection barrier leads to a higher current without affecting the population inversion. This explains why G605 is predicted to have higher gain.

The simulated result does not entirely agree with the experimental results. For instance, G552 had a $T_{\max}$ 10K higher than G605. The authors of Ref. [5] attributed the difference in performance to the fact that even though the increase in current is beneficial for the gain, the thinner injection barrier results in a larger coupling between $|i_{n-1}\rangle$ and $|p_{1,n}\rangle$, $|p_{2,n+1}\rangle$, and between $|l_{n-1}\rangle$ and $|p_{1,n}\rangle$ as well, see figure 18. The first will result in LO-scattering rates to leakage channels that increase with temperature, while the latter will result in resonant tunneling to the leakage channels that is not particularly temperature dependent. So to predict the temperature performance of the devices, the first one has to be taken into account. Since $N_{\text{per}} = 1$ has been used so far, the coupling between i_{n-1} and $p_{2,n+1}$ has not been included in the simulations. Accordingly, simulations with $N_{\text{per}} = 2$ is done in the next section to investigate whether or not it is necessary to include more neighbors to accurately predict the behavior of the structures at high temperature.

Wafer	G605	G552
ΔE_1	47	46
ΔE_2	54	54
Max gain (100K)	108	101
Max gain (285K)	32	26
$T_{\max}^{\text{sim}}$	340	320
$T_{\max}^{\text{exp}}$	225	235
$J_{\max}^{\text{sim}}$(100K)	2.7	2.4
$J_{\max}^{\text{sim}}$(285K)	2.4	2.0
$J_{\max}^{\text{exp}}$	2.87	2.67

Table 6: Simulated results using the NEGF-package using $E_c = 1.01x$ eV, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$, $\alpha = 20$/cm. Energies are given in meV, gain in 1/cm and current density in (kA/cm^2). The definitions of E_1 and E_2 are given in figure 18.

4.2.3 Simulations with $N_{\text{per}} = 2$

The results from using $N_{\text{per}} = 2$ are presented in table 7. All other simulation parameters were the same as in section 4.2.2. The decrease in $T_{\max}$ for G605 and the increase for G552 and G652 brings the result closer to the trend seen in the experiments. However, the changes are marginal and probably within the margin of error considering the resolution being used. Thus $N_{\text{per}} = 1$ suffices for the simulations. The simulated results, therefore, suggest that the leakage channel i_{n-1} to $p_{2,n+1}$ cannot explain the difference in simulated and experimental results in the last part. Hence, the thinner injection barrier of G605 might actually be beneficial for performance due to the resulting higher currents, and the experimental results must then be caused by the discrepancy in growth quality. Our results do not, therefore, agree with the discussion provided by the authors in Ref. [5],

Wafer	G552	G605	G652
Max gain (285K)	26 (26)	33 (32)	25 (26)
$T_{\max}^{\text{sim}}$	325 (320)	335 (340)	340 (335)

Table 7: Simulated results using the NEGF-package with $E_c = 1.01x$ eV, $N_{\text{nu}} = 5$, $N_{\text{per}} = 2$, $\alpha = 20$/cm. Energies are given in meV, gain in 1/cm and current density in (kA/cm^2). The values given in parenthesis are from earlier simulations with $N_{per} = 1$ instead.

where the leakage channels discussed in section 4.2.2 were seen as the reason why G552 outperformed G605.

4.2.4 Simulations with $N_{\mathbf{nu}} = 7$

The simulated results for G652 agreed remarkably well with the experimental ones, while G605 and G552 underestimated the currents by about 10% which is also a great result. The reason why the currents were underestimated in G605 and G552 and not in G652 could be due to the number of Wannier states being used was too low. Using a low number of Wannier states would underestimate the current through the HLS and the leakage to the continuum. Simulations were thus run using the same parameters as before, but with $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 1$. Simulations with $N_{\mathrm{nu}} = 3$ were also used as a reference. The results are shown in table 8.

First, G552 and G605 have an interesting behavior where the increase in current when going from 5 to 7 Wannier states is greater at low temperatures. This could be due to the following: the main way electrons transition to the HLS at low temperature is by resonant tunneling from the LLS, see figure 17. Furthermore, at high temperatures electrons can transition to the HLS by LO-phonon scattering due to the higher thermal energy, which contributes to the current. This is why we see a greater impact when going from 3 to 5 Wannier states at higher temperatures. However, introducing more Wannier states affect the energy levels of all states. For G552 and G605, the energy separation between $|u_n\rangle$ and $|p_{2,n+1}\rangle$, and between $|l_n\rangle$ and $|p_{1,n+1}\rangle$ increases, which counteracts the expected increase in current by decreasing the rate of these tunneling transitions. The opposite is the case for G652.

Another problem when calculating $T_{\mathrm{max}}^{\mathrm{sim}}$ is the heavily underestimated current at temperatures above 300 K. Five Wannier states seem sufficient at 285K considering the $J_{\mathrm{max}}^{\mathrm{sim}}$ values shown in table 8, but this might not be the case for temperatures ≈320-340K. For instance, at 340K the value for $J_{\mathrm{max}}^{\mathrm{sim}}$ is simulated to be 2.2 kA/cm^2 for both G652 and G605, i.e. a drop from 285K of 5% and 10% respectively. This is not in agreement with the experiment as J_{max} increases with temperature. Consequently, leakage to the continuum plays a much greater role at these temperatures and is not taken into account by using only 5 Wannier states. This leakage has detrimental effects on the performance as it heats up the electron gas, and will probably affect G605 the most as its simulated $J_{\mathrm{max}}^{\mathrm{sim}}$ drops the most with increased temperature. Consequently, using more than 5 Wannier states might more accurately predict the T_{max} of G605 contra G552 and G652. Thus the T_{max} with 7 Wannier states was simulated. The difference in $T_{\mathrm{max}}^{\mathrm{sim}}$ was negligible when going from 5 to 7 Wannier states which mean that 5 Wannier states is sufficient to simulate these samples, even at temperatures in the region of 300-350K.

Wafer	J_{max}^{3}(100K)	J_{max}^{5}(100K)	J_{max}^{7}(100K)	J_{max}^{3}(285K)	J_{max}^{5}(285K)	J_{max}^{7}(285K)	$T_{\mathrm{max}}^{\mathrm{sim}}$
G552	2.31	2.37	2.53	1.65	2.04	2.12	320
G605	2.63	2.68	2.87	2.03	2.42	2.54	335
G652	2.33	2.64	2.67	1.98	2.32	2.40	335

Table 8: Simulated currents under lasing conditions at 100K and 285K using 3, 5 and 7 Wannier states. Current densities are given in kA/cm^2 and temperatures in K. The simulated max temperature was simulated with 7 Wannier states. Relevant simulation parameters are $\alpha = 20$/cm and $N_{\mathrm{per}} = 1$. The simulations estimates how important the HLS are for simulating the operation of the samples.

4.2.5 Conclusion

The simulation results clearly show that using $E_c = 1.01x$ eV as CBO yields results that agree well with the experimental results. $J_{\max}^{\text{sim}}$ at 100K was much closer to the experimental value when compared to $J_{\max}^{\text{sim}}$ at 285K, which is due to background current not being considered in the simulations. Deploying $N_{\text{nu}} = 5$ and $N_{\text{per}} = 1$ yielded $T_{\max}^{\text{sim}}$ of 340K, 335K and 320K for G605, G652 and G552 respectively. G605 thus deviates from the experimental result. By simulating with $N_{\text{nu}} = 7$ and $N_{\text{per}} = 2$ respectively, it was ruled out that the deviation was due to previous simulation parameters being incapable of capturing the underlying physics. Furthermore, the analysis of the value of N_{nu} shows that the leakage current to the higher lying states is small for the MIT samples. Thus, the authors of Ref. [5] managed to achieve a clean 3-level system, according to the simulation results, which were what they set out to do. The analysis of the value of N_{per} shows that tunneling through modules is not as important an issue as previously thought.

4.3 Thermal Degradation for High-Barrier QCLs

When the sample BoscoAPL2019 (discussed in chapter 4.1) was introduced in 2019 by Bosco *et al.*, it broke the temperature record at that time, with a $T_{\max}$ of 210K [7]. They did so by using unconventionally high $Al_{0.25}Ga_{0.75}As$ barriers. One year later in 2020, G652 beat the record with a $T_{\max}$ of 250K by following the same high barrier approach [5]. Khalatpour *et al.* used even higher $Al_{0.3}Ga_{0.7}As$. As such, they claim that they managed to eliminate most of the detrimental current to higher lying states by using this design choice together with their optimization algorithm.

In this section, I try to answer why the device from Bosco *et al.* had more temperature degradation than the devices from Khalatpour *et al.*. Additionally, the difference in temperature degradation in the devices from the article by Khalatpour *et al.* will be investigated as they have slightly different designs.

4.3.1 Devices Studied

In this section, the two devices G605 and G652 will be studied further together with the device BoscoAPL2019. In the previous analysis of chapter 4.1, it was found that a CBO of 1.01x eV = 300 meV yielded the most accurate result for G605 and G652 while a CBO of 0.831x eV = 210meV yielded the most accurate result for BoscoAPL2019. Some device properties are summarized and repeated in table 9.

Wafer	CBO	E_1	E_2	$J_{\max}^{\text{sim}}$	$J_{\max}^{\text{exp}}$	$T_{\max}^{\text{exp}}$	Layer sequence	x
BoscoAPL2019	210meV	40	40	3.7	3.5	210	**32.6**/79.9/**19.0**/<u>164.6</u>	0.25
G605	300meV	48	50	2.7	2.9	225	**31.4**/76.3/**17.5**/<u>153.1</u>	0.3
G652	300meV	51	58	2.6	2.6	250	**33.7**/72.0/**18.7**/<u>144.9</u>	0.3

Table 9: The samples that will be studied in this text. The samples G605 and G652 comes from Ref. [5] and BoscoAPL2019 comes from Ref. [7]. CBO gives which CBO was used for simulating $J_{\max}^{\text{sim}}$ under lasing conditions. The simulation was run with a lattice temperature of 100K, $N_{\text{nu}} = 5$, $N_{\text{per}} = 1$ and $\alpha = 20$/cm. $J_{\max}^{\text{exp}}$ and $T_{\max}^{\text{exp}}$ are taken from respective reference. Layer sequences are given in Å with bold values being $Al_xGa_{1-x}As$ barriers, and the underlined well is doped in a 30 Å region in the center. The definitions of E_1 and E_2 are given in figure 18.

4.3.2 Leakage Current

First, how the current densities are distributed in the band structure was simulated. The simulation was run with $T_{\mathrm{L}} = 300K$, $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ under no lasing conditions. The results are shown in figure 19. From these figures, it is evident that BoscoAPL2019 suffers significantly more leakage current through the HLS at 300K compared to G605 and G652. Furthermore, G605 has more leakage current compared to G652. This suggests that higher E_2 in fact results in less leakage current through the HLS. However, it could also be the thicker radiation barrier of G652 which results in less coupling between $|i_{n-1}\rangle$, $|u_n\rangle$ and $|p_{1,2,n}\rangle$ since it decreases the spatial overlap. Nonetheless, Khalatpour *et al.* managed to eliminate the detrimental current through the higher lying states which seems to result in better performance at higher temperatures.

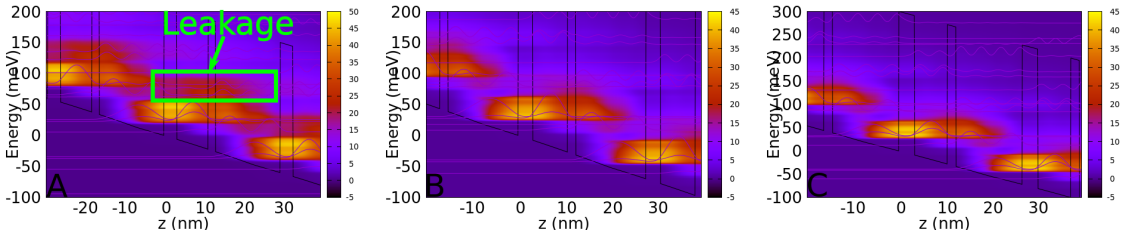

Figure 19: Current densities at the point of maximum current for three different devices simulated using the NEGF-package. Figure A shows BoscoAPL2019, figure B shows G605, and figure C shows G652. Simulation parameters used was $T_{\mathrm{L}} = 300$ K, $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ was used and no lasing was assumed.

4.3.3 Maximum Gain

The maximum gain was simulated to get data on how the performance of the devices degrade with temperature. The simulation was run with $T_{\mathrm{L}} = 100$ K to 350 K in steps of 50, $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ under no lasing conditions. The applied bias used was the applied bias at the point of maximum current $F_{\mathrm{dc}}^{\mathrm{op}}$, $F_{\mathrm{dc}}^{\mathrm{op}} + 2$ mv/module and $F_{\mathrm{dc}}^{\mathrm{op}} + 4$ mv/module as this is the interval where the maximum gain could be seen. This region is thus a region of NDC in the simulations, but we expect the background current that increases with bias will counteract the NDC in this region. As a consequence, this region will not show NDC in the experiments.

If thermal backfilling to the LLS is causing thermal degradation, then it should start to have a big effect once the electrons have enough in-plane energy to transition from $|i_n\rangle$ to $|l_n\rangle$ needing $E_1 - E_{LO} = 11.5$, 14.5 and 3.5meV for G605, G652 and BoscoAPL2019 respectively. These correspond to lattice temperatures of 140K, 170K and 40K. Hence, at these temperatures, the thermal degradation should increase. Here the relative change in gain $\Delta G_R = \frac{1}{\mathrm{Gain}(T)} \frac{\mathbf{d}Gain}{\mathbf{d}T}$ is used to quantify temperature degradation. Using the same logic for leakage to HLS, changes in ΔG_R would be seen at 150K, 250K and 40K for G605, G652 and BoscoAPL2019 respectively, since those temperatures correspond to the energy of $E_2 - E_{LO}$. The results are shown in table 10, where it is clear that the temperature degradation decreases with increasing E_1 and/or increasing E_2.

Firstly, since BoscoAPL2019 has both of these temperatures at 40K, we cannot use it to see the hypothesized change in ΔG_R. However, it can be used as a benchmark for G605 and G652. For instance, if G652 starts to get a lower ΔG_R at around 250K relative to BoscoAPL2019, then it is evidence that a leakage channel has opened for G652 at around 250K. This works since we don't expect any noticeable change for BoscoAPL2019 at around 250K.

Wafer	G605	G652	Bosco
E_1	48	51	40
E_2	50	58	40
Bias	66-70	68-72	56-60
Max Gain(100K)	108	97	116
Max Gain(150K)	90	83	93
Max Gain(200K)	66	62	64
Max Gain(250K)	46	44	42
Max Gain(300K)	29	28	23
Max Gain(350K)	17	17	9
$\frac{G(100)-G(350)}{250}$	0.36	0.32	0.43

Table 10: Simulated results together with relevant energy separations. Simulated using the NEGF-package using simulation parameters $T_{\mathrm{L}} = 300$ K, $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ was used and no lasing was assumed. Energies are given in meV, bias in mV and gain in 1/cm. The biases choosen are the biases of maximum current up to +4mV, since this is the interval where the max gain was seen.

We do not see an increase in ΔG_R at 250 K for G652. Thus leakage to higher lying states as the degrading mechanism can be ruled out for G652, consistent with figure 20. Furthermore, G605 diverges from G652 in ΔG_R at around 150K and then starts to converge again right before 200 K. As discussed previously, this is around the temperature we expect to see an increase in ΔG_R for G605 and G652 respectively. As such, the conclusion is that thermal backfilling to the LLS is a bigger cause of the thermal degradation of G652 compared to thermal excitations to the HLS. For G605 the same type of analysis is not possible as $E_1 \approx E_2$.

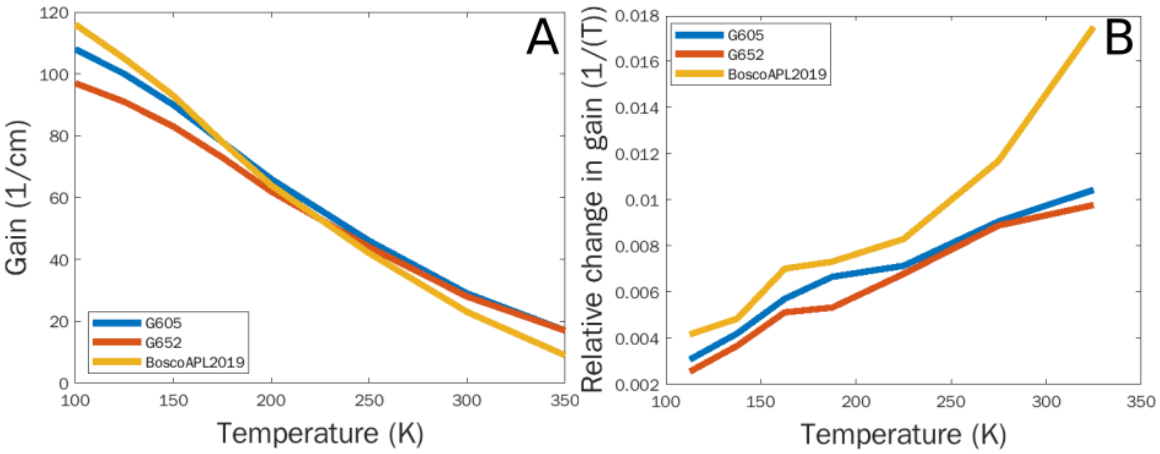

Figure 20: Figure A shows Gain and figure B shows ΔG_R as a function of temperature. Simulated using the NEGF-package using simulation parameters $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ was used and no lasing was assumed. Simulations were run in the temperature interval of 100-350K.

However, these trends are small so the conclusion must be taken cautiously. It is however clear that G652 degrades the least with temperature and it is probably due to less thermal backfilling as we found in the previous section that leakage to higher lying states is small for G605 and G652. To get more insight the population densities were simulated in the next section.

4.3.4 Population Densities

The population densities were simulated over a temperature interval of 100-350K to see how the electron densities change with temperature. The simulations were carried out using the same input as in the last section.

First, lets take a look at the results for the ULSs, shown in figure 21. For all devices, it is evident that the population of the ULS is diminishing with temperature resulting in less population inversion. All devices have roughly the same slope, suggesting that the thermal degradation effect of the ULS has similar strength for all devices. This is not surprising as the main way the population density of the ULS decreases is by TALOP-scattering, which depends mostly on the radiation barrier width, which is similar for all three devices.

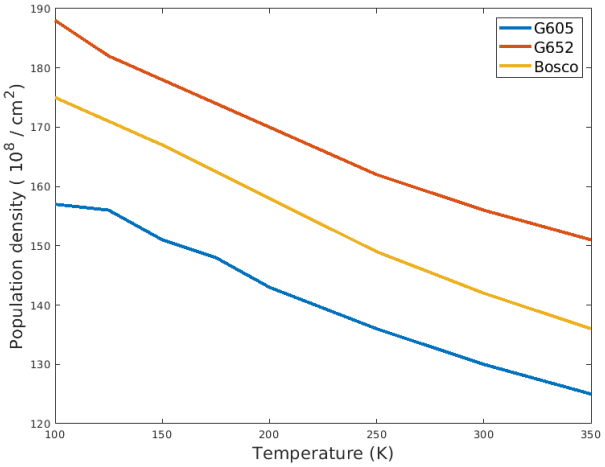

Figure 21: Electron populations in the ULS for the devices BoscoAPL2019, G605, and G652. Simulated using the NEGF-package using simulation parameters $N_{\text{nu}} = 7$ and $N_{\text{per}} = 2$ under no lasing conditions. Simulations were run in the temperature interval of 100-350K.

Secondly, the LLS will be considered. If we assume that all electrons in the LLS are extracted to the injector and all electron to the LLS comes from the injector (i.e thermal backfilling is the main contributor the the LLS population density), then we get the following expression

$$\frac{\partial n_{\text{LLS}}}{\partial t} = -R_{\text{LLS}\to\text{I}} \cdot n_{\text{LLS}} + R_{\text{I}\to\text{LLS}} \cdot n_{\text{I}} \tag{17}$$

for the population density of the LLS n_{LLS}, where $R_{\text{LLS}\to\text{I}}$ and $R_{\text{I}\to\text{LLS}}$ are the rate coefficients. As such, $R_{\text{I}\to\text{LLS}} \cdot n_{\text{I}}$ is an expression for the thermal backfilling from the injector to the LLS. Since the probability per atom is the same for absorption and stimulated emission, we get $R_{\text{I}\to\text{LLS}} = R_{\text{LLS}\to\text{I}} \equiv R = e^{-E_1/(k_B T)}$ and we thus get

$$\frac{1}{R} \cdot \frac{\partial n_{\text{LLS}}}{\partial t} = \frac{I}{R} = n_{\text{LLS}} - n_{\text{I}} e^{-E_1/(k_B T)} \tag{18}$$

where I is the current through the LLS. As a consequence, if the difference $n_{\text{LLS}} - n_{\text{I}} e^{-E_1/(k_B T)}$ is constant, then the reason for the change in LLS population is due to thermal backfilling. In figure 22, the simulated population densities together with the expression for thermal backfilling $n_I \cdot e^{-E_1/(k_B T)}$

are plotted for all three structures. For all three structures the two curves run parallel, meaning that our assumption about most of the electrons in the LLS coming from the injector is correct. Consequently, thermal backfilling is present in all three structures, and the amount is given by E_1. The greater E_1 is, the smaller is the slope of the curve and in extension the thermal backfilling. As a result, G652 has the least thermal backfilling, followed by G605 and BoscoAPL2019.

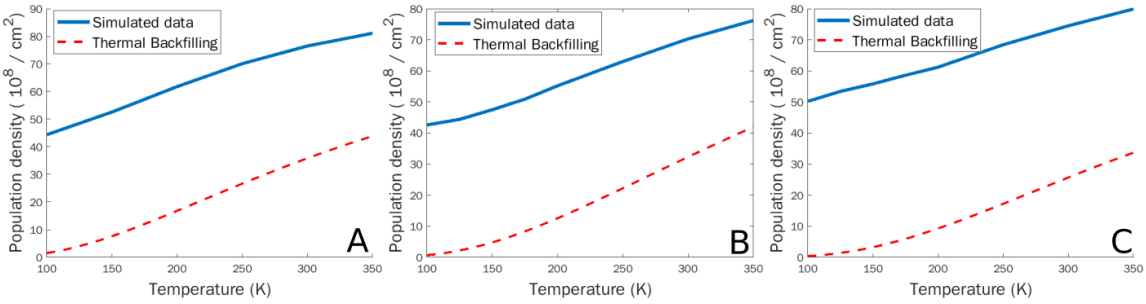

Figure 22: Electron populations in the LLS for the devices BoscoAPL2019 in figure A, G605 in figure B, and G652 in figure C. A curve representing the thermal backfilling from a derived expression is included and plotted besides the simulated data. The curves are parallel, which shows that thermal backfilling is the main cause for the increasing population of the LLS. Simulated using the NEGF-package using simulation parameters $N_{\text{nu}} = 7$ and $N_{\text{per}} = 2$ under no lasing conditions. Simulations were run in the temperature interval of 100-350K.

Thirdly, the parasitic states $|p_{1,n}\rangle$ and $|p_{2,n}\rangle$ will be discussed. Here their combined population densities are shown in figure 23. It is evident once more that G652 manages leakage to parasitic currents the best while BoscoAPL2019 performs the worst in this aspect, in agreement with chapter 4.1.

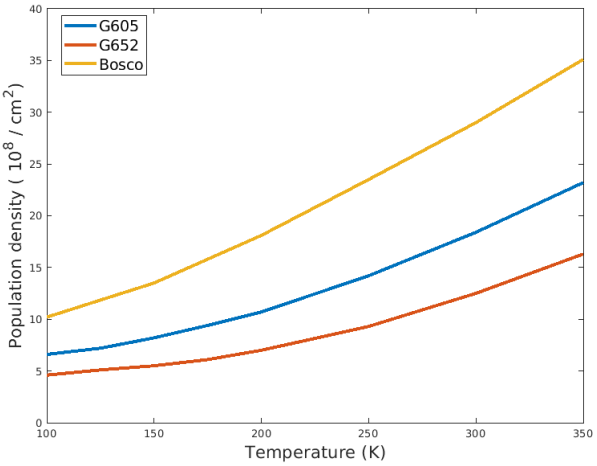

Figure 23: The sum of the electron densities from the two parasitic states $|p_{1_n}\rangle$ and $|p_{2_n}\rangle$. Simulated using the NEGF-package using simulation parameters $N_{\text{nu}} = 7$ and $N_{\text{per}} = 2$ was used and no lasing was assumed. Simulations were run in the temperature interval of 100-350K.

Lastly, the injector state will be considered. The results are shown in figure 24. For G605 and

BoscoAPL2019, the injector level loses an electron population of $20{\cdot}10^8/\mathrm{cm}^2$ and $30{\cdot}10^8/\mathrm{cm}^2$. This corresponds to 60% and 80% of the increase in LLS population respectively, see figure 21a. Hence, for these structures, thermal backfilling might be the main cause for thermal degradation as opposed to TALOP-emission from ULS to LLS. For G652 the situation is different. Here the electron population actually increases until around 175K, in good agreement with when thermal backfilling should become a factor. The decrease in population is roughly $5{\cdot}10^8/\mathrm{cm}^2$ which corresponds to 20% of the increase in the LLS population. G652, therefore, seems to suffer the least thermal backfilling, most likely due to it having the largest E_1. This agrees with what we found in the two earlier sections.

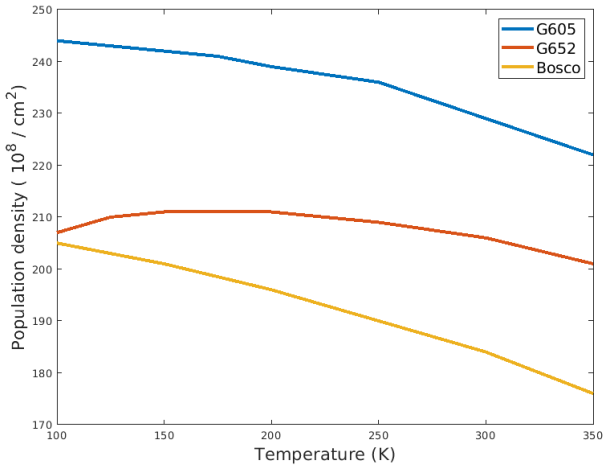

Figure 24: Total electron population for the injector state. Simulated using the NEGF-package using simulation parameters $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ no lasing was assumed. Simulations were run in the temperature interval of 100-350K.

4.3.5 Conclusion

According to the simulation results, the authors of Ref. [5] are correct about their assertion that their devices manage to prohibit parasitic current through the higher lying states, compared to earlier devices such as BoscoAPL2019. As a result, thermal backfilling plays a greater role when it comes to temperature degradation for these devices. Furthermore, G652 has the least thermal backfilling as well as the least parasitic current and hence also the greatest population inversion. As such, G652 seems to have the most favorable characteristics, which may explain why it performed the best.

Also, it has been shown that the gain more rapidly decreases with temperature once thermal backfilling is possible, i.e. $k_B T \approx E_1 - E_{\mathrm{LO}}$. Even larger E_1 might therefore be the way forward to increase the operating temperatures even more.

4.4 Simulation of New Devices

Inj. Barrier \ Rad. Barrier	17.5 (G605 value)	18.7 (G652 value)	19.7	20.7
33.7 (G652 value)	≈23 (G552)	29	-	-
31.4 (G605 value)	30	35	33	32
30.4	-	36	35	33
29.4	-	36	33	32

Table 11: Simulated max gain of nine new samples. The simulations were performed using $T_{\mathrm{L}} = 300$ K, $N_{\mathrm{nu}} = 5$, $N_{\mathrm{per}} = 2$ under no lasing conditions. Barrier widths are given in Å and max gains are given in 1/cm. These new samples were simulated to try to improve on the performance of G605, which has a simulated max gain of 30/cm at 300 K.

The analysis so far has shown that the NEGF-package succeeds to simulate the samples with good accuracy. Consequently, the NEGF-package can be deployed to develop new samples. In this part, the MIT samples will be tried to be improved upon, and thus the CBO $E_c = 1.01x$ eV will be used. In chapter 4.2 it was found that the injection barrier width of G605 was better than the injection barrier width of G552, and it was found that the radiation barrier width of G652 was better than the radiation barrier width of G552. As G605 and G552 had the same radiation barrier width, and G652 and G552 had the same injection barrier width, the logical choice for a starting point is a structure with the injection barrier width of G605 and the radiation barrier width of G652.

Furthermore, the parameter space around this starting point was tested. Three injection barrier widths and three radiation barrier widths were simulated, resulting in nine different samples. Since a thinner injection barrier width was found to be beneficial, the chosen injection barrier widths were the injection barrier width of G605 31.4 Å, together with 30.4 Å and 29.4 Å. By the same token, as a thicker radiation barrier width was found to be beneficial, the radiation barrier widths of 18.7 Å (G652 value), 19.7 Å and 20.7 Å were used.

The simulations of these nine samples were performed assuming no lasing conditions, $T_{\mathrm{L}} = 300$ K, $N_{\mathrm{nu}} = 5$ and $N_{\mathrm{per}} = 2$. The well widths were chosen such that the extractor energy $E_1 = 51$ meV, and the injector and ULS align at an applied bias of $F_{\mathrm{dc}} = 70$ mV/period. This was achieved by using the calcWS program, described in chapter 3.2, to calculate the energy levels. Then the well widths were adjusted step by step until the preferred characteristics were achieved. However, calcWS does not yield as exact values as the hdiag program but was used since it is more time-efficient. The results are presented in table 11. All samples had a higher max gain than G605, which was the best performing MIT sample at 300 K with a max gain of 30/cm.

Finally, we chose to look closer at two of the nine samples, the one with injection barrier width 18.7 Å and radiation barrier width 30.4 Å, as well as the sample with injection barrier width 29.4 Å and radiation barrier width 20.7 Å. These two were chosen as they had a high gain in proportion to current density. More details about these samples are shown in table 12, together with G605 as reference. To get higher precision, $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$ was used to simulate the max gain and $J_{\mathrm{max}}^{\mathrm{sim}}(300K)$ for these samples. The gain decreased at 300 K by around 10% for the two samples. Nonetheless, both still outperform G605 at 300 K and 350 K. LU2 is interesting because it has a relatively high gain at a much lower current than G605 and can thus expect less heating when operating. LU1 is interesting on the other hand as it has a significantly higher gain, albeit at the cost of a higher current.

Wafer	E_1	E_2	$J^{\mathrm{sim}}_{\mathrm{max\ gain}}$(300 K)	Max Gain(300 K)	Max Gain(350 K)	Layer sequence
LU1	53	55	2.6	33	22	**30.4**/71.5/**19.7**/<u>146</u>
LU2	51	53	2.1	30	19	**31.4**/71.5/**20.7**/<u>149</u>
G605	48	50	2.4	29	17	**31.4**/76.3/**17.5**/<u>153.1</u>

Table 12: Simulated results for two of the new samples together with the simulated results for G605 as reference. The two new samples were chosen from a trial of nine samples that were simulated with $N_{\mathrm{nu}} = 5$ and $N_{\mathrm{per}} = 2$. To investigate their performance further, the max gain at 300 K and 350 K was simulated as well as the current, using $N_{\mathrm{nu}} = 7$ and $N_{\mathrm{per}} = 2$. Furthermore, the assumed waveguide loss used was 20/cm. Current densities are given in kA/cm^2, energies in meV, and gain in 1/cm. Layer sequences are given in Å with bold values being $Al_{0.3}Ga_{0.7}As$ barriers, and the underlined well is doped in a 30 Å region in the center with a doping density of $1.5{\cdot}10^{17}$/cm^3.

Furthermore, the IV-curves, gain spectra, and resolved current densities for LU1 and LU2 are shown in figure 25. The properties of the two samples are similar to the properties of G552, G605 and G652. This is not surprising as the design of LU1 and LU2 are based on those structures. The maximum current is however at slightly higher biases and the maximum gain is at slightly higher photon energies. Also, the leakage to HLS is low, meaning that these leakage channels are under control with these new types of structures.

These simulations suggest that there is room for improvement when it comes to THz QCL design. Of course, the samples must be fabricated and tested in a lab environment before one can celebrate. Nevertheless, this is evidently an interesting area of the parameter space.

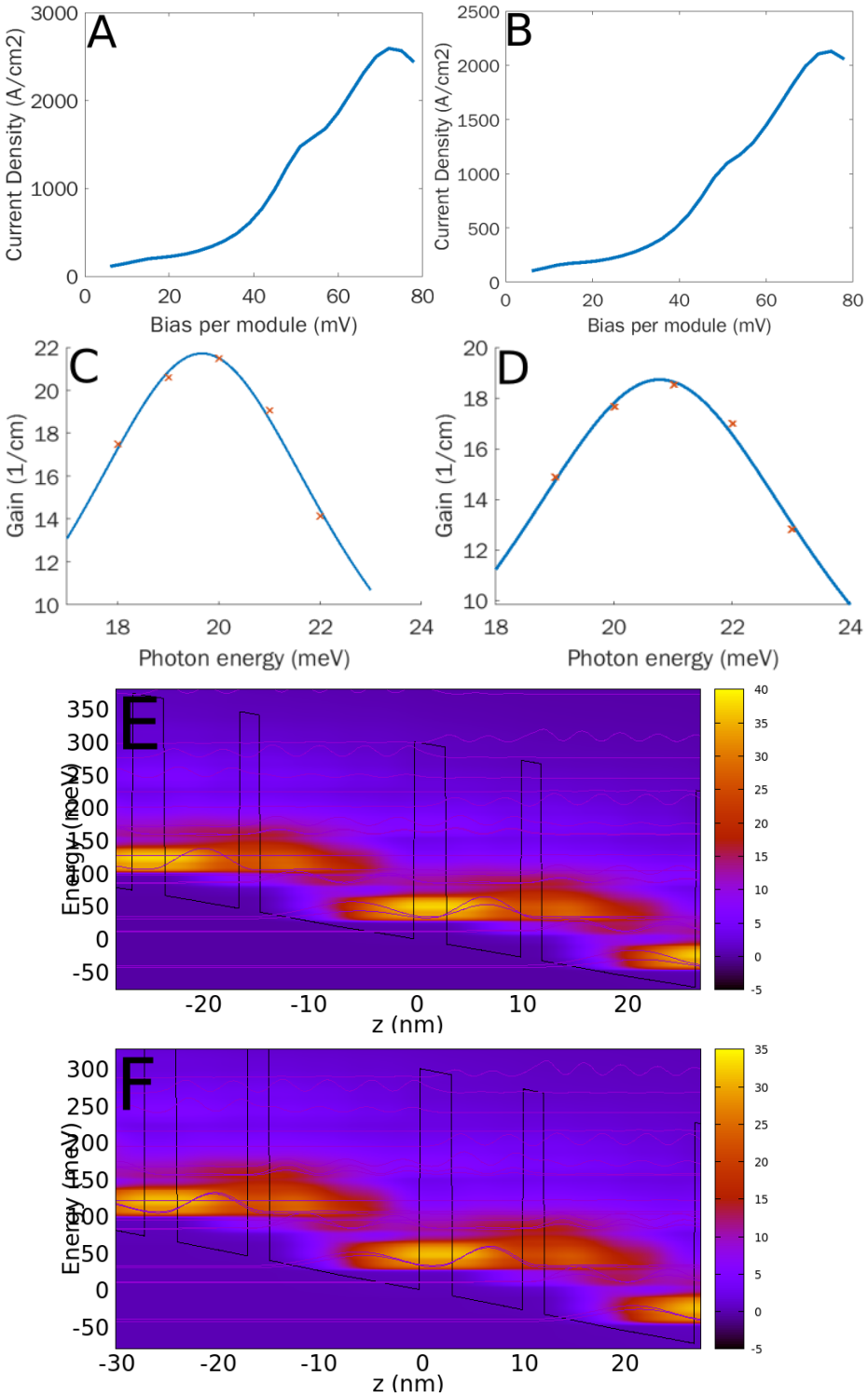

Figure 25: Simulated data using $N_{\text{nu}} = 7$, $N_{per} = 2$ under no lasing conditions. Figure A shows the IV-curve for LU1 and figure B shows the IV-curve for LU2. Figure C and Figure D show the gain spectra for LU1 and LU2 at 350K. The red x's show simulated data and the blue line represents a Lorentzian fit to the data. Figure E and figure F show the resolved current densities for LU1 and LU2 respectively at 350 K.

5 Conclusion & Outlook

In this thesis, detailed simulations of 2 QW THz QCLs have been performed. From the results, it is clear that a different CBO is necessary for samples from different labs. This is useful information for those who want to simulate QCLs in the future. Furthermore, it was shown that by using a CBO of $E_c = 1.01x$ eV we were able to simulate the performance of the new record-holding THz QCLs with tremendous accuracy, even though the CBO is unconventional. Thereby, it was shown that the structures BoscoAPL2019 and G605 had thermal degradation due to thermal backfilling while G652 had limited thermal backfilling. As such, G652 seems to have achieved the highest operating temperature for a THz QCL by successfully suppressing leakage to HLS and thermal backfilling. What was learned from the simulations could then be applied to design new samples that according to the simulation package outperform the current record-holding samples when it comes to max gain at temperatures over 300 K. Therefore it is entirely possible that further optimization could result in better samples. However, if this part of the parameter space contains a QCL operational at room temperature is too early to tell, but the future for THz is looking bright (even at room temperature) as many parameters such as doping density, doping width, and barrier height are yet to be optimized for this new generation of high barrier and highly diagonal 2-QW THz QCLs.

References

1. Faist, J. *Quantum Cascade Lasers* (Oxford University Press, Oxford, 2013).
2. Kazarinov, R. F. & Suris, R. A. Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice. *Sov. Phys. Semicond.* **5,** 707 (1971).
3. Faist, J. *et al.* Quantum Cascade Laser. *Science* **264,** 553–556. ISSN: 0036-8075. http://www.sciencemag.org/cgi/doi/10.1126/science.264.5158.553 (Apr. 1994).
4. Beck, M. *et al.* Continuous Wave Operation of a Mid-Infrared Semiconductor Laser at Room Temperature. *Science* **295,** 301 (2002).
5. Khalatpour, A., Paulsen, A. K., Deimert, C., Wasilewski, Z. R. & Hu, Q. High-power portable terahertz laser systems. *Nature Photonics* **15,** 16–20. https://doi.org/10.1038/s41566-020-00707-5 (2021).
6. Sirtori, C. Terahertz race heats up. *Nat. Photonics* **15,** 1–2 (2021).
7. Bosco, L. *et al.* Thermoelectrically cooled THz quantum cascade laser operating up to 210K. *Appl. Phys. Lett.* **115,** 010601. https://doi.org/10.1063/1.5110305 (2019).
8. Vitiello, M. S. & Tredicucci, A. Physics and technology of Terahertz quantum cascade lasers. *Advances in Physics: X* **6,** 1893809. eprint: https://doi.org/10.1080/23746149.2021.1893809. https://doi.org/10.1080/23746149.2021.1893809 (2021).
9. Jirauschek, C. & Kubis, T. Modeling techniques for quantum cascade lasers. *Appl. Phys. Rev.* **1,** 011307. ISSN: 1931-9401. http://aip.scitation.org/doi/10.1063/1.4863665 (Mar. 2014).
10. Wacker, A., Lindskog, M. & Winge, D. O. Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions. *IEEE J. Sel. Top. Quant.* **19,** 1200611. ISSN: 1077-260X. http://ieeexplore.ieee.org/document/6415241/ (Sept. 2013).
11. Winge, D. *Quantitative Modeling of Gain in Quantum Cascade Lasers under Operational Intensities* PhD thesis (Lund University, Nov. 2016). ISBN: 978-91-7753-046-6. http://portal.research.lu.se/portal/files/16302081/dw_lucris.pdf.
12. Franckié, M. *Modeling quantum cascade lasers : the challenge of infra-red devices.* ISBN: 9789176237793. http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cat07147a&AN=lub.4829872&site=eds-live&scope=site (Division of Mathematical Physics, Department of Physics, Lund University, 2016).
13. Esaki, L. & Tsu, R. Superlattice and Negative Differential Conductivity in Semiconductors. *IBM J. Res. Dev.* **14,** 61–65. ISSN: 0018-8646. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5391729 (Jan. 1970).
14. Davies, J. H. *The Physics of Low-dimensional Semiconductors: An Introduction* (Cambridge university press, 1997).
15. Wacker, A. *Quantum Cascade Laser: An Emerging Technology* in *Nonlinear Laser Dynamics* (ed Lüdge, K.) (Wiley-VCH, Berlin, 2012).
16. Khanal, S., Reno, J. L. & Kumar, S. 2.1 THz quantum-cascade laser operating up to 144 K based on a scattering-assisted injection design. *Opt. Express* **23,** 19689–19697. http://www.opticsexpress.org/abstract.cfm?URI=oe-23-15-19689 (July 2015).
17. Nelander, R. & Wacker, A. Temperature dependence of the gain profile for terahertz quantum cascade lasers. *Appl. Phys. Lett.* **92,** 081102 (2008).

18. Kumar, S., Hu, Q. & Reno, J. L. 186 K operation of terahertz quantum-cascade lasers based on a diagonal design. *Appl. Phys. Lett.* **94,** 131105 (2009).

19. Krivas, K. A., Winge, D. O., Franckié, M. & Wacker, A. Influence of interface roughness in quantum cascade lasers. *J. Appl. Phys.* **118.** http://scitation.aip.org/content/aip/journal/jap/118/11/10.1063/1.4930572 (2015).

20. Franckié, M. *et al.* Impact of interface roughness distributions on the operation of quantum cascade lasers. *Opt. Express* **23,** 5201–5212. http://www.opticsexpress.org/abstract.cfm?URI=oe-23-4-5201 (Feb. 2015).

21. Lindskog, M., Winge, D. O. & Wacker, A. Injection schemes in THz quantum cascade lasers under operation. *Proc. SPIE* **8846** (eds Razeghi, M., Baranov, A. N. & Zavada, J. M.) 884603. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2024030 (Sept. 2013).

22. Chassagneux, Y. *et al.* Limiting Factors to the Temperature Performance of THz Quantum Cascade Lasers Based on the Resonant-Phonon Depopulation Scheme. *Terahertz Science and Technology, IEEE Transactions on* **2,** 83–92. ISSN: 2156-342X (Jan. 2012).

23. Li, H. *et al.* Temperature performance of terahertz quantum-cascade lasers: experiment versus simulation. *Journal of Physics D: Applied Physics* **42,** 025101. http://stacks.iop.org/0022-3727/42/i=2/a=025101 (2009).

24. Albo, A. & Hu, Q. Carrier leakage into the continuum in diagonal GaAs/Al0.15GaAs terahertz quantum cascade lasers. *Appl. Phys. Lett.* **107.** http://scitation.aip.org/content/aip/journal/apl/107/24/10.1063/1.4937455 (2015).

25. Vitiello, M. S., Scalari, G., Williams, B. & Natale, P. D. Quantum cascade lasers: 20 years of challenges. *Opt. Express* **23,** 5167–5182. http://www.opticsexpress.org/abstract.cfm?URI=oe-23-4-5167 (Feb. 2015).

26. Curl, R. F. *et al.* Quantum cascade lasers in chemical physics. *Chem. Phys. Lett.* **487,** 1–18 (2010).

27. Price, P. Two-dimensional electron transport in semiconductor layers. I. Phonon scattering. *Annals of Physics* **133,** 217–239. ISSN: 0003-4916. https://www.sciencedirect.com/science/article/pii/0003491681902505 (1981).

28. Albo, A., Flores, Y. V., Hu, Q. & Reno, J. L. Two-well terahertz quantum cascade lasers with suppressed carrier leakage. *Applied Physics Letters* **111,** 111107. eprint: https://doi.org/10.1063/1.4996567. https://doi.org/10.1063/1.4996567 (2017).

29. Franckié, M. *et al.* Two-well quantum cascade laser optimization by non-equilibrium Green's function modelling. *Appl. Phys. Lett.* **112,** 021104. https://doi.org/10.1063/1.5004640 (2018).

30. Albo, A. & Hu, Q. Investigating temperature degradation in THz quantum cascade lasers by examination of temperature dependence of output power. *Appl. Phys. Lett.* **106.** http://scitation.aip.org/content/aip/journal/apl/106/13/10.1063/1.4916961 (2015).

31. Williams, B. S., Kumar, S., Qin, Q., Hu, Q. & Reno, J. L. Terahertz quantum cascade lasers with double-resonant-phonon depopulation. *Appl. Phys. Lett.* **88,** 261101 (2006).

32. Albo, A., Qing, H. & Reno, J. L. Room temperature negative differential resistance in terahertz quantum cascade laser structures. *Applied Physics Letters* **109,** 081102-1 - 081102–5. ISSN: 00036951. http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=117752425&site=eds-live&scope=site (2016).

33. Wannier, G. H. The structure of electronic excitation levels in insulating crystals. *Phys. Rev.* **52,** 191 (1937).

34. Snoke, D. W. *Solid State Physics: Essential Concepts* (Addison-Wesley, 2009).

35. Brandes, T. Truncation method for Green's functions in time-dependent fields. *Phys. Rev. B* **56,** 1213. http://link.aps.org/doi/10.1103/PhysRevB.56.1213 (3 July 1997).

36. Winge, D. O., Franckié, M. & Wacker, A. Simulating terahertz quantum cascade lasers: Trends from samples from different labs. *Journal of Applied Physics* **120,** 114302. http://dx.doi.org/10.1063/1.4962646 (2016).

37. Vurgaftman, I., Meyer, J. R. & Ram-Mohan, L. R. Band parameters for III-V compound semiconductors and their alloys. *J. Appl. Phys.* **89,** 5815 (2001).

38. Yi, W., Narayanamurti, V., Lu, H., Scarpulla, M. A. & Gossard, A. C. Probing semiconductor band structures and heterojunction interface properties with ballistic carrier emission: GaAs/$Al_xGa_{1-x}As$ as a model system. *Phys. Rev. B* **81,** 235325. http://link.aps.org/doi/10.1103/PhysRevB.81.235325 (23 June 2010).

39. Wacker, A. personal communication. Oct. 15, 2021.

40. Scalari, G. *et al.* Broadband THz lasing from a photon-phonon quantum cascade structure. *Opt. Express* **18,** 8043. ISSN: 1094-4087. https://www.osapublishing.org/oe/abstract.cfm?uri=oe-18-8-8043 (Apr. 2010).