Modeling Quantum Cascade Lasers

The Challenge of Infra-Red Devices

Franckie, Martin

Modeling Quantum Cascade Lasers

Modeling Quantum Cascade Lasers:
The Challenge of Infra-Red Devices

by Martin Franckié

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Prof. Andreas Wacker, Dr. Claudio Verdozzi
Faculty opponent: Prof. Gerhard Klimeck

To be presented, with the permission of the Faculty of Science of Lund University, for public criticism in the Rydberg lecture hall (Rydbergsalen) at the Department of Physics on Friday, the 13th of May 2016 at 13:00.

Organization
LUND UNIVERSITY

Department of Physics
Box 118
221 00 LUND
Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2016-05-13

Sponsoring organization

Author(s)
Martin Franckié

Title and subtitle
Modeling Quantum Cascade Lasers: The Challenge of Infra-Red Devices

Abstract
The Quantum Cascade Laser (QCL) is a solid state source of coherent radiation in the terahertz and mid-infrared parts of the electro-magnetic spectrum. In this thesis, I have used a non-equilibrium Green's function (NEGF) model to theoretically investigate the quantum nature of the device operation, and compared simulation results to experiments. As simulations with the NEGF model are time-consuming, especially for mid-infrared QCLs owing to their large state space, the program has been parallelized to run on large computer clusters. Additional challenges of mid-infrared QCLs are increased interface roughness scattering and non-parabolicity effects, which have been included into the existing NEGF model. With this new model, simulation results in close agreement with experimental data has been produced for terahertz and mid-infrared lasers employing various design schemes.

Key words
quantum cascade lasers, non-equilibrium Green's function theory, cluster computations

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title

ISBN
978-91-7623-778-6 (print)
978-91-7623-779-3 (pdf)

Recipient's notes

Number of pages
184

Price

Security classification

Modeling Quantum Cascade Lasers:
The Challenge of Infra-Red Devices

by Martin Franckié

LUND
UNIVERSITY

Cover illustration: Free interpretation of this work by Bernt Franckié Grünewald

Funding information: The thesis work was financially supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 317884, the collaborative Integrated Project MIRIFISENS and the Swedish Research Council (VR).

Faculty of Science, Department of Physics

ISBN: 978-91-7623-778-6 (print)
ISBN: 978-91-7623-779-3 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2016

Till
Fleur

Contents

List of publications . . . iii
List of Acronyms . . . v
Acknowledgements . . . vi
Popular summary in English . . . vii
Populärvetenskaplig sammanfattning på svenska . . . viii

1 Introduction **3**
1.1 Motivation and structure of the Thesis . . . 5
1.2 Operation principles . . . 6
1.3 Material systems . . . 11
1.4 Approaches to QCL modeling . . . 13

2 Theory **15**
2.1 Hamiltonian . . . 15
2.2 Envelope functions . . . 17
2.3 Bloch functions of a periodic heterostructure . . . 18
2.4 Wannier functions . . . 19
2.5 The Wannier-Stark Basis . . . 20
2.6 The NEGF model . . . 20
2.7 Temporal response . . . 23
2.8 Current density . . . 28
2.9 Gain from complex conductivity . . . 29
2.10 Gain from Fermi's Golden Rule . . . 30
2.11 Gain clamping and output power . . . 30

3 The two-band model **33**
3.1 Current density in the two-band model . . . 35
3.2 Interface roughness in the two-band model . . . 37
3.3 Comparison of one- and two-band model simulations . . . 38

4 Parallelization **43**
4.1 Computation scheme . . . 43
4.2 Single node computations . . . 46
4.3 Cluster computations . . . 47

5 Mid-infrared QCLs **51**
5.1 Role of interface roughness scattering 52
5.2 Multi-stack device 54
5.3 Strain compensated InGaAs/InAlAs/AlAs structures . . . 56
5.4 Optimization of short wavelength and low dissipation active region 58

6 Terahertz QCLs **63**
6.1 Injection schemes in THz QCLs 64
6.2 Two-quantum well designs 71

7 Conclusions and outlook **81**

Scientific publications **93**
Author contributions 93
Paper I: Injection Schemes in THz Quantum Cascade Lasers Under Operation 93
Paper II: An indirectly pumped terahertz quantum cascade laser with low injection coupling strength operating above 150 K 93
Paper III: Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions 93
Paper IV: Comparative analysis of quantum cascade laser modeling based on density matrices and non-equilibrium Green's functions 94
Paper V: Impact of interface roughness distributions on the operation of quantum cascade lasers 94
Paper VI: Influence of Interface Roughness in Quantum Cascade Lasers 94
Paper I: Injection Schemes in THz Quantum Cascade Lasers Under Operation 95
Paper II: An indirectly pumped terahertz quantum cascade laser with low injection coupling strength operating above 150 K 107
Paper III: Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions . 123
Paper IV: Comparative analysis of quantum cascade laser modeling based on density matrices and non-equilibrium Green's functions 137
Paper V: Impact of interface roughness distributions on the operation of quantum cascade lasers 143
Paper VI: Influence of Interface Roughness in Quantum Cascade Lasers 157

Appendix: Tables with details for simulated structures **165**

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Injection Schemes in THz Quantum Cascade Lasers Under Operation**

M. Lindskog, D. O. Winge, and A. Wacker
Proc. SPIE, 8846:884603 (2013)

II **An indirectly pumped terahertz quantum cascade laser with low injection coupling strength operating above 150 K**

S. G. Razavipour, E. Dupont, S. Fathololoumi, C. W. I. Chan, **M. Lindskog**, Z. R. Wasilewski, G. Aers, S. R. Laframboise, A. Wacker, Q. Hu, D. Ban, and H. C. Liu
Journal of Applied Physics **113**, p. 203107 (2013)

III **Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions**

A. Wacker, **M. Lindskog**, and D. O. Winge
IEEE Journal of Selected Topics in Quantum Electronics **19**, p. 1200611 (2013)

IV **Comparative analysis of quantum cascade laser modeling based on density matrices and non-equilibrium Green's functions**

M. Lindskog, J. M. Wolf, V. Trinité, V. Liverini, J. Faist, G. Maisons, M. Carras, R. Aidam, R. Ostendorf, and A. Wacker
Applied Physics Letters **105**, p. 103106 (2014)

V **Impact of interface roughness distributions on the operation of quantum cascade lasers**

M. Franckié, D. O. Winge, J. M. Wolf, V. Liverini, E. Dupont, V. Trinité, J. Faist, and A. Wacker
Optics Express **23**, p. 5201-5212 (2015)

VI **Influence of Interface Roughness in Quantum Cascade Lasers**

K. Krivas, D. O. Winge, **M. Franckié**, and A. Wacker
Journal of Applied Physics **118**, p. 114501 (2015)

Publications not included in this thesis:

Non-linear response in Quantum Cascade Structures

D. O. Winge, **M. Lindskog**, and A. Wacker
Appl. Phys. Lett. **101**(21), p. 211113 (2012)

Microscopic approach to second harmonic generation in quantum cascade lasers

D. O. Winge, **M. Lindskog**, and A. Wacker
Optics Express **22**, pp. 18389-18400 (2014)

A phonon scattering assisted injection and extraction based terahertz quantum cascade laser

E. Dupont, S. Fathololoumi, Z. R. Wasilewski, G. Aers, S. R. Laframboise, **M. Lindskog**, S. G. Razavipour, A. Wacker, D. Ban, and H. C. Liu
J. Appl. Phys. **111**(7), p. 73111 (2012)

Superlattice gain in positive differential conductivity region

D. O. Winge, **M. Franckié**, and A. Wacker
preprint arXiv:1601.00915

Simple electron-electron scattering in non-equilibrium Green's function simulations

D. O. Winge, **M. Franckié**, C. Verdozzi, A. Wacker, and M. F. Pereira
preprint arXiv:1602.03997

List of Acronyms

ac	Alternating Current
CBO	Conduction Band Offset
dc	Direct Current
DM	Density Matrix
FCA	Free-carrier absorption
FCG	Free-carrier gain
FGR	Fermi's Golden Rule
IFR	Interface Roughness
IR	Infra-Red
laser	Light Amplification by Stimulated Emission of Radiation
LLS	Lower Laser States
LO	Longitudinal Optical (phonon)
MBE	Molecular Beam Epitaxy
MOCVD	Metal-Organic Chemical Vapour Deposition
MPI	Message Passing Interface
NDR	Negative Differential Resistance
NEGF	Non-Equilibrium Green's function
OMP	OpenMP (Open Multi-Processing)
QCL	Quantum Cascade Laser
RE	Rate Equation
RP	Resonant phonon
SA	Scattering Assisted
THz	Terahertz (10^{12} Hz)
ULS	Upper Laser State

Acknowledgements

Physics has always been the subject that interest me the most, and so doing a PhD in this field has truly been a dream come true. I would like to first thank my supervisor Andreas Wacker for being an inspiration to go into the field of solid state theory and quantum optics. Leading me through my bachelor and master project, and finally supporting me all through my years as a PhD student, Andreas has been both an excellent supervisor, teacher, and friend. Through him, I have not only learned facts and physics, but also how to think as a scientist.

I owe the deepest thanks to my dear colleague and friend David Winge, who has been with me to discuss and exchange knowledge and ideas throughout my PhD project. I also want to thank the other members of our group, past and present: Olof, Fikeraddis, Bahare, Gedeminas, Kasparas, Kevin, Sanna, and my co-supervisor Claudio. Of course, I am very grateful to the rest of the division of Mathematical Physics for the friendly and open atmosphere, not the least thanks to Katarina!

I also would like to send my gratitude to my foreign colleagues Emmanuel Dupont and Ghasem Razavipour, with whom I've had long and interesting discussions and great collaborations. I am also deeply thankful to Jérôme Faist and his group, for accepting me as an intern for three months, a time where I both learned a lot of physics and gained invaluable experience.

Finally, my family has been utterly supportive, not complaining too much about late nights' work and weeks, or even months, of travel. Thank you mother and father, for your support in whatever I engage, for all your love, and for always being there. My parents in-law, Bernt (who also made the cover illustration) and Barbro, have also given me a lot of encouragement and perspectives on life. Thank you!

Thank you Frida for giving me a firm place to stand! I cannot describe how valuable all those things are that you have given me in terms of support, love, friendship, consolation, relaxation, and believing in me. And thank you for the biggest gift of all: in the very same week that I now finish my thesis, we will have our daughter. This work I dedicate to her.

Popular summary in English

We live in a time of rapid scientific and technological advancement. People living 100 years ago could never dream of inventions like those having completely changed our way of life, and our perception of the world; computers, mobile phones, the Internet, space travel, unraveling the mysteries of the early universe and distant galaxies, and our insight into the microscopic world of quantum phenomena. Today we are at the dawn of an era of nanotechnology, with computers components being only tens of nanometers in size, and nano-devices making their entrance into wide industrial use.

One such nanoscopic device is the Quantum Cascade Laser (QCL). Like all as lasers, it emits electro-magnetic waves, which essentially is light. In fact, our eyes can only detect a very narrow range of wavelengths in the electro-magnetic spectrum, which stretches from very long radio waves, followed by microwaves, terahertz and infrared waves, through the wavelengths of visible light, up to ultra-violet UV-A and UV-B radiation (giving us a good tan), and finally X-rays and gamma-rays from radio-active decay and cosmic radiation. In the case of, e. g. conventional laser pointers, the light wave has wavelengths of 500 (blue light) up to 700 (red light) nanometers and is visible to our eyes.

What is special about the QCL is that it does not emit visible light, but light in the terahertz and infrared (IR) regions, and it does so using the same technology as normal light emitting diode (LED) lights. These regions are both interesting for applications in spectroscopy, i. e. the detection of chemical substances by looking at light going through them. For example, QCLs can be used to detect very small quantities of explosive materials, diagnose exhaled air in patients, monitor green house gases and pollution in the atmosphere, and examining the contents of far away stars. But, like IR cameras, they can also be used to make images of that which our eyes cannot see; with terahertz light we can see through clothes and thin materials, which could be used to screen patients in the emergency room without having to remove their clothes, or at the airport to see if they are carrying a weapon.

In this work, I have simulated QCLs using a complicated theoretical model, which in detail accounts for the motion of the electrons inside the device, and their interaction with their surroundings. By improving this model, we are actually able to reproduce real QCL behavior remarkably well. This can help to improve future QCLs, and to understand the basic physical mechanisms underlying their operation. This to me is what physics is all about: To seek to understand nature and make use of this knowledge to help making life better.

Populärvetenskaplig sammanfattning på svenska

Vi lever i en tid av snabb vetenskaplig och teknologisk utveckling. För 100 år sedan hade ingen kunnat drömma om de saker som helt har förändrat vår vardag och hur vi uppfattar omvärlden; datorer, mobiltelefoner, Internet, rymdfart, våra upptäckter om universums barndom och avlägsna galaxer samt vår insyn i kvantfenomenens mikroskopiska värld. Idag står vi på tröskeln till en era formad av nanoteknologi, med datorkomponenter bara tiotals nanometer stora (en tiotusendedel av ett hårstrås tjocklek), och då nya innovationer på nanomenterskalan gör sina intåg i industriellt bruk.

En sådan nanoskopisk apparat är kvantkaskadlasern (eng. Quantum Cascade Laser). Som alla lasrar skickar den ut elektromagnetisk strålning, vilket väsentligen är ljusvågor. Våra ögon kan faktiskt endast uppfatta en liten del av det elektromagnetiska spektrat, vilket sträcker sig från de mycket långa radiovågorna, via mikrovågor, terahertz och infrarött, genom synligt ljus, och vidare till ultraviolett UV-A och UV-B strålning (den som ger en så härlig bränna på sommaren), och slutligen till Röntgen- och gammastrålar från radioaktivt sönderfall och kosmisk strålning. Vanliga laserpekare skickar till exempel ut ljusvågor med våglängder mellan 500 nanometer (blått ljus) och 700 nanometer (rött ljus), vilket är synligt för våra ögon.

Det speciella med kvantkaskadlasern är att den inte skickar ut synligt ljus, utan ljus med terahertz- och infraröda våglängder. Den gör dessutom detta med hjälp av samma teknologi som används för att göra vanliga LED (light emitting diode) lampor som numera finns i vartenda hem. Terahertz och infrarött ljus är intressant att använda inom spektroskopi - alltså detektion av kemiska ämnen genom att titta på det ljus som träffat dem. Till exempel kan kvantkaskadlasrar användas för att upptäcka mycket små mängder sprängämnen, diagnostisera patienter genom utandningsluften, följa koncentrationen av växthusgaser och föroreningar i atmosfären och undersöka innehållet i avlägsna stjärnor. Men det är även intressant att, som med infraröda kameror, avbilda saker i terahertz som våra ögon normalt inte kan se. Terahertzvågor går nämligen igenom textilier och kan användas på flygplatser för att upptäcka om någon döljer ett vapen under sina kläder. Eller också skulle de kunna användas för att inspektera en patient som kommer in på akutavdelningen, utan att behöva klippa upp dennes kläder.

I den här avhandlingen har jag gjort datorsimuleringar av kvantkaskadlasrar genom att använda en komplicerad datormodell som i detalj tar med elektronernas resa genom lasern i beräkningarna, och hur de växelverkar med sin omgivning. Genom att förbättra denna modell, har jag kunnat återskapa hur verkliga kvant-

kaskadlasrar fungerar med anmärkningsvärd precision. Detta kommer att underlätta utvecklingen av, och förbättra framtida, kvantkaskadlasrar samt hjälpa oss förstå de underliggande fysikaliska processerna - hur elektronerna beter sig på sin kvantmekaniska färd genom lasern. Det är just det jag tycker är så fantastiskt med fysik: att försöka förstå naturen och använda denna kunskap för att förbättra människors liv.

Modeling Quantum Cascade Lasers: The Challenge of Infra-Red Devices

"Give me a firm place to stand and I will move the earth."
— Archimedes

Chapter 1

Introduction

In the last century, mankind has made technological advances nobody could ever imagine. Today, nanotechnology is rapidly moving from scientific curiosity into the real lives of people. One great example is the Quantum Cascade Laser (QCL), which is right now transitioning from the science lab into wide industrial use. In this thesis, I will describe the progress I have supplied to this ongoing "revolution". This work is theoretical, but some of the results have been made in close collaboration with experimentalists and institutions pushing this technology to be used in, and to benefit, the general society. The history of the QCL began with the conceptualization by Kazarinov and Suris [1] in 1971. The first operational QCL was reported 22 years ago [2], and lased at a wavelength of 4.2 μm, which is in the mid-infrared (mid-IR) regime. Since then, many technological and theoretical steps have been taken to steadily improve and diversify the usage of QCLs.

Like a conventional atomic laser, the QCL exploits population inversion between two quantum states. In contrast to real atoms, these states are formed in artificial atoms made from coupled quantum wells in a semiconductor heterostructure. Applying an electrical bias to this heterostructure drives the inversion between the two states, and thereby light amplification by stimulated emission of radiation (LASER) can be generated. In addition, the quantum wells can be repeated any number of times, and in each such period of the heterostructure the light-emitting electrons are recycled. The competitiveness of the QCL comes from it's wavelength coverage in combination with it's compactness and high power efficiency, and that it is constructed using the mature fabrication technology of III-V semiconductors [3, 4]. The two important spectral ranges of (i) the mid-IR region (above the Reststrahlen band and below optical frequencies) and (ii)

the terahertz (THz) region (below the Reststrahlen band and above microwave frequencies) are covered. Let us look a little deeper into each of these regions to see why the QCL is such an interesting invention.

The mid-IR spectral range covers frequencies from 40-100 THz, or wavelengths between 3-8 μm. Although there already exists efficient molecular laser sources, they are bulky and have limited spectral ranges [5]. All solid state lasers in this range - except one - are limited by the band gap of the material used [5]. The exception is the QCL which at the same time is compact, energy efficient [6], and has a high output power [7]. The first room-temperature continuous wave operation of a mid-IR QCL was reported in 2002 [8] and over the course of the past 20 years, the QCL has found it's way into industrial applications in a range of fields; environmental monitoring [9], quality control of liquids (from oil platforms as well as food manufacturing) [10], as well as security and defense sensing applications [11] are some examples. Due to the nature of the growth process, multiple frequencies can be emitted from the same QCL, either by a cleverly designed heterostructure [12] or by combining several single-frequency active regions [13]. For sensing applications (e. g. environmental and medical breath analysis [14]), a broad spectrum is important as it allows all desired molecules to be measured and distinguished using the same QCL. A common goal of QCL research is to increase the wall plug efficiency (the ratio of electrical power put in to the device to the power emited as radiation) of mid-IR lasers, in order for portable sensing applications to become reality. To date, some of the best lasers have a wall plug efficiency of about 20-27% at room temperature continuous wave operation [15, 6], although at cryogenic temperatures over 50% wall plug efficiency has been reported [16]. This is also the purpose of the European Union seventh framework project MIRIFISENS [17], which I have been participating as a part of my thesis project; to bring industry and scientific research together and accelerate the development of QCLs for industrial applications.

The THz spectral range stretches from 0.3-10 THz (wavelengths from 30 to 1000 μm). This spectral range is important as it lies between microwave frequencies generated by electronic components (such as Gunn diodes) and optical frequencies. However, available radiation sources (free electron lasers, synchrotron, and photomixing, where non-linearities can be used for down-conversion of IR radiation to THz radiation at room temperature, are some examples) are inefficient, bulky, or very expensive. This has prevented the realization of many applications in this wave length region (hence the so-called "THz gap") [3]. The QCL is a good candidate to fill this gap, and first lased in the THz range in 2002 [18]. While THz radiation can be produced by difference frequency generation in mid-IR QCLs at room temperature [19], to date the maximum operating tem-

perature for single frequency THz QCLs is 199.5 K [20]. The goal would be a QCL operating at a lattice temperature of $T = 240$ K, since then it would be possible to cool the device electronically [3]. Optimizing the wall-plug-efficiency of these lasers [21] is also important for a wider range of applications.

The QCL shows great potential for many applications in a wide wavelength region. Due to their composition, they are relatively cheap, mass-producible and seemingly very durable [22]; recently, the very first laser [2] was operated again and still lased, even at higher temperature (78 K), as we witnessed at the celebration of the 20th anniversary of the QCL [23].

1.1 Motivation and structure of the Thesis

The QCL operates on principles derived from quantum mechanics, and it's performance depends on a complex interaction of the light-generating electrons with many different aspects of the material system and the laser light. The details of what happens on this almost atomic scale is not directly observable, yet they are vital for the development and analysis of these structures. Thus, modeling on the quantum level is essential, as this allows us to understand and predict the operation of these devices in detail. The modeling is done on different levels, from the propagation of the electro-magnetic field inside and outside the QCL to the transport of electrons and their interaction with material impurities, phonons, and the laser light itself.

The motivation of this work is to develop mid-IR QCL technology by careful modeling of electronic transport and light generation. To this end, several extensions to an existing model, adapted for THz QCL simulations, needed to be implemented and tested. One important ingredient in a successful modeling of mid-IR lasers, which have large conduction band offsets, is the inclusion of the effects of the other lattice bands in addition to the conduction band of the semiconductor materials. This gives the conduction band a non-parabolic dispersion [24] and an energy-dependent effective mass. This is done using an effective two-band model, described in Chap. 3. These lasers also host a large number of active states compared to THz QCLs, and together with the large energy-grid required by the high conduction band offset, the computational effort becomes much bigger. In order to make it possible to get useful results in reasonable time, it is essential to run computations in parallel, on several processors on a single machine, or on a computer cluster. The parallelization of the code is described in Chap. 4.

As the limiting factor of continuous wave operation at room temperature is the

large heat generation in the active region compared to the low heat removal possible for such a small device, a big goal in the development of mid-IR QCLs is to have a low heat dissipation and a high wall-plug efficiency. A structure for this purpose is the QCL presented in Ref. [25], emitting at 8.5 μm. Another goal is to reach shorter wavelengths of ~ 3 μm for gas sensing applications. Such short wavelengths require a very large conduction band offset. This can be achieved by the growth of materials with lattice constants which are not matched to the InP substrate [26, 27]. In addition, some layers are very thin, giving a very high interface roughness scattering. In order to reduce this effect, a design based on mono-layers was attempted. Computations for mid-IR QCLs using the above mentioned additions to the program are presented in Chap. 5.

In order to validate the program after the inclusion of the effects in Chaps. 3 and 4, we simulate THz structures, for which the previous version of the program is reliable. These results are presented in Chap. 6. In this section, new proposals for THz QCL designs working up to room temperature in simulations are presented, as well as results and discussions for THz QCLs employing different design schemes.

A wide variety of QCL modeling schemes are on the market [28]. Often, simpler models are convenient to use in order to deliver results quickly. However, it is important to know that the results from these models are reliable. To this end, we compare our advanced modeling scheme to simpler models in Chaps. 5 and 6.

The long-term goal of QCL research is to develop applications in a variety of fields, ranging from diagnostics through exhaled air in medicine [14], to precise spectroscopic measurements in astrophysics [29, 30]. The potential benefit to society is great, once QCLs are commercially available as portable devices, which for THz QCLs require a large increase in operation temperature, and for mid-IR QCLs an increased energy-efficiency and wider wavelength coverage. Regarding the results of this work, an outlook is given in Chap. 7.

1.2 Operation principles

The concept of a QCL utilizes the idea of a quantum well, a very popular example for illustrating quantum mechanics. One of the most common combination of materials used is GaAs as the well material and $Al_xGa_{1-x}As$ as the barrier material (x is the relative amount of aluminum to gallium) which are schematically drawn in Fig. 1.1. When put together, the conduction and valence band edges will bend at the interfaces, creating a conduction band offset depending on the

ratio x. The valence band is completely full, and so we turn our attention to

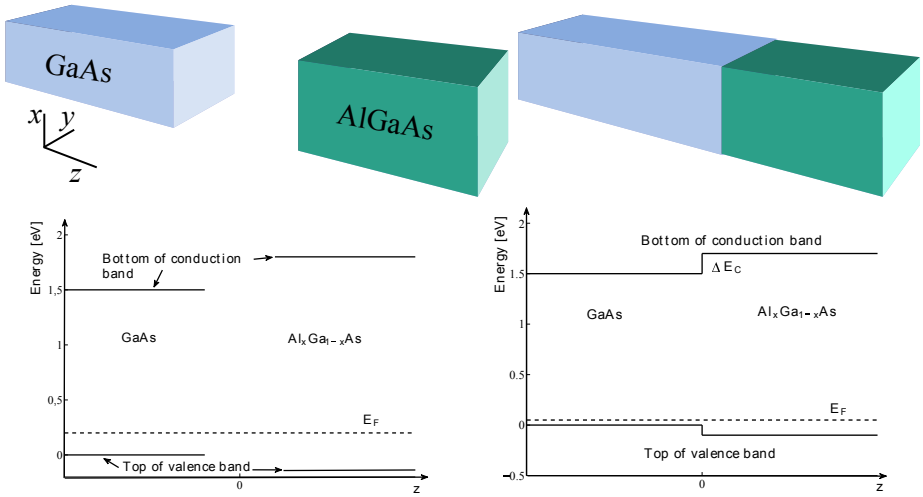


Figure 1.1: Two pieces of bulk GaAs and AlGaAs have different band structure and different band gaps at the Γ-point. Put together, the bands will bend at the interface between the materials, keeping the Fermi energy E_F constant.

what is happening in the conduction band from now on. (The mixing of the valence band will come into play in the calculation of the basis states in the two-band model.) In the z-direction (the growth direction), the electrons incoming from the left in Fig. 1.1 will see a barrier in the GaAs/AlGaAs conduction band offset. In the transverse directions (x and y), however, the electrons can move freely.

Quantum wells in the conduction band are made by alternating between thin (a few Å to a few nm) layers of materials with different band gaps as in Fig. 1.2. The electron energy is quantized according to the depth of the well and the widths of the wells and barriers. This allows for tailoring the eigenenergies and spatial extent of the quantum mechanical states. The energy components in the in-plane directions are not quantized and thus the electrons can have any energy E_k associated with the transverse wave vector. When many quantum wells are placed in sequence, forming a so-called superlattice (or heterostructure), electrons can tunnel across the barriers between neighboring wells. When a bias is applied, see Fig. 1.3, the whole band structure bends in energy and the tunneling electrons will give rise to a tunneling current. For the structure in Fig. 1.3(a), the current as a function of applied bias is shown in Fig. 1.3(b).

The quantum cascade laser exploits inter-subband transitions in the conduction band, where electrons are provided by doping the lattice with donor atoms. The

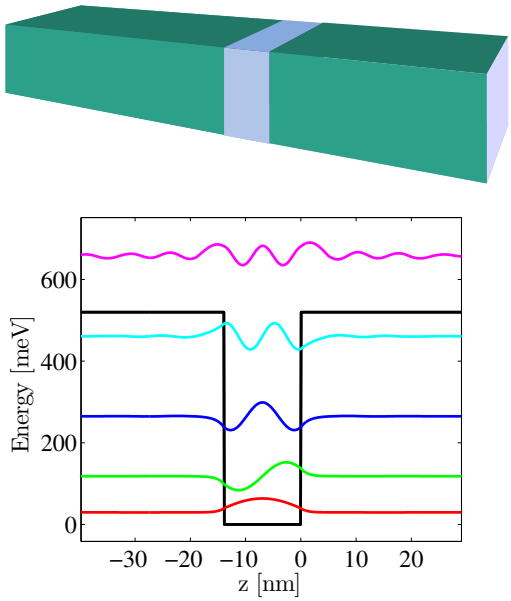


Figure 1.2: Slicing a thin slab of GaAs inside a AlGaAs slab creates a quantum well confinement potential for the electrons in the conduction band.

idea is now to tailor the quantum wells and barriers, so that an applied bias will drive the electronic current in such a way that inversion is created between two subbands. When inversion is present, electrons from the upper laser state (with higher electron density) can be brought into the lower laser state (with less electron density) by stimulated emission. By altering the layer widths, we can get any suitable separation in energy between the states. The energy difference between the lower laser state and some extraction state, e. g., is preferably tuned to the optical phonon energy to get an efficient out-scattering from the lower laser state. Alternatively, the two states are tuned into resonance to get a high tunneling rate, in order to maintain inversion.

One such tailored structure is shown in Fig. 1.4, where the short-hand notation for the lower laser state (LLS), upper laser state (ULS), extraction level (e) and injection level (i) has been introduced. This is actually the simplest possible structure [32], with only three active states and two wells per period.

This is how the laser scheme works: electrons are brought from a state of higher energy (ULS) to a state of lower energy (LLS) by stimulated emission, when a photon with the energy of the level separation interacts with the electrons. This process requires inversion, i. e. more carriers have to be in the ULS than in the LLS. The inversion is driven by the bias voltage, transporting the carriers

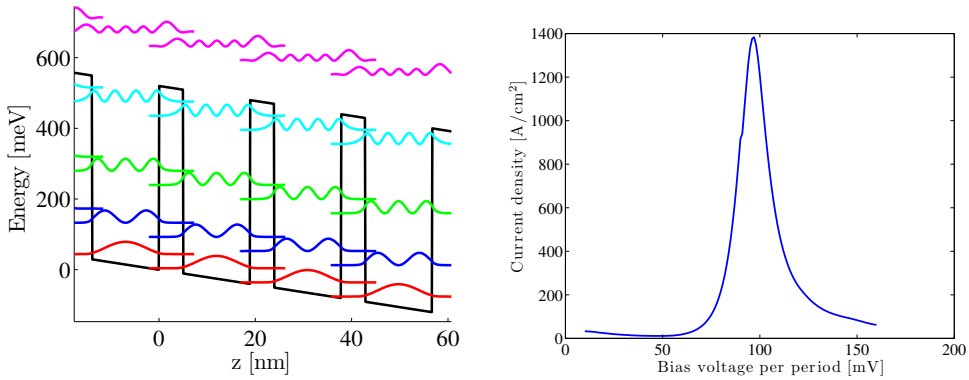


Figure 1.3: Left: Superlattice structure with an applied bias of 40 mV per period. Right: Current vs. applied bias per period. A peak occurs at around 100 mV/period, where the ground state of one well aligns with the first excited state of the neighboring well down stream.

from the LLS to the ULS in the next period via efficient tunneling and phonon scattering.

1.2.1 Mid-infrared designs

QCL designs aimed at emission at widely different wavelengths (from 3 to 200 μm) require different strategies for pumping the upper state and depopulating the lower state. Most importantly, high inversion needs to be established, by having fast injection to the upper laser state, which has a long non-radiative lifetime, and a fast extraction of the lower laser state. Furthermore, this needs to be achieved at a bias with positive differential conductivity in order to have stable operation. Once inversion is guaranteed, different designs can be aimed at specific wavelengths, broad or narrow gain spectra, tunability, etc.

The biggest difference between mid-IR and THz designs, is that they operate at photon energies above and below the longitudinal optical phonon energy E_{LO}, respectively. For mid-IR designs, with $\hbar\omega_{\mathrm{IR}} = E_{\mathrm{ULS}} - E_{\mathrm{LLS}} > E_{\mathrm{LO}}$, injection can be achieved by resonant tunneling, while extraction is dominated by resonant-LO phonon scattering in a band of states separated by energies $\approx E_{\mathrm{LO}}$ below the lower laser state. This is called a bound-to-continuum (BtC) scheme [25] and is shown in Fig. 5.1. Since the states in the lower band all have significant dipole matrix element with the upper laser state, such a design can have a broad gain spectrum. The continuum-to-continuum design is taking this one step further,

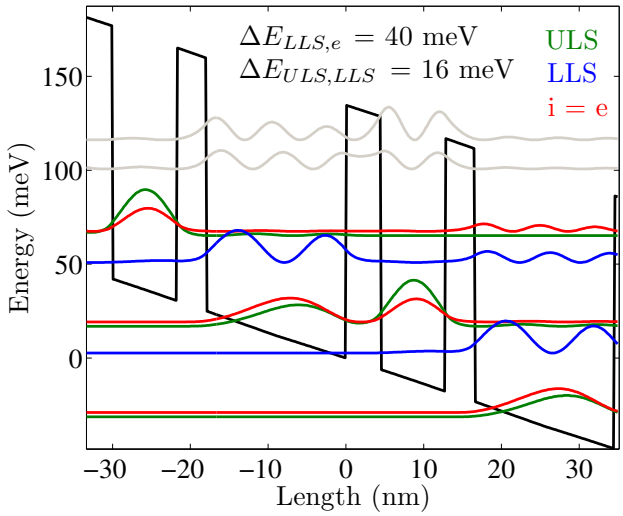


Figure 1.4: Conduction band edge and the Wannier-Stark states of the structure investigated in Ref. [31]. The different colors represent the lower laser state (LLS), upper laser state (ULS), the extraction (e) and injection (i) levels. The injection and extraction levels are the same in this particularly simple structure. The alignment between the injection level and the ULS provides efficient tunneling filling the ULS of the next period.

by pumping the upper laser state also by a band of states, that all have significant dipole matrix elements with the lower band, making the gain even broader[33]. Another pumping scheme is to have a resonant LO phonon scattering channel into the upper laser state, called scattering assisted (SA) injection [34]. This makes the injection very fast.

The above are the most common strategies for inversion between subbands. In addition, several active regions designed for emission at different wavelengths can be grown in sequence, to give one very broad gain spectrum or several narrow spectra centered at separate wavelengths. To date, laser wavelengths as short as 2.6 μm have been achieved [35].

1.2.2 Terahertz designs

THz frequencies have photon energies $\hbar\omega_{\mathrm{THz}} < E_{\mathrm{LO}}$, below the LO phonon energy. As the energy separation $E_{\mathrm{ULS}} - E_{\mathrm{LLS}}$ is small, the ULS lifetime is shorter and it is more difficult to obtain inversion, than for the mid-IR frequencies. For the same reason, the effect of heating is also more detrimental at lower

temperatures than for the mid-IR designs.

A common depopulation scheme is resonant phonon (RP) extraction [36]. This is the scheme shown in Fig. 6.1. Here, the lower state is close in energy to a second state, and both of these states are one E_{LO} above a third state. This gives a very fast depopulation, and reduces thermal backfilling (i. e. thermal excitations from the third state back to the lower laser state) as the LLS and third state can be spatially separated. Designs similar to the mid-IR designs have been tested, although to date it is the RP extraction scheme, with resonant tunneling injection and using three quantum wells per period, which has been most successful [20]. The SA scheme is also promising [37][Paper II], but has so far not been as successful, perhaps due to the small attention it has received compared to the RP schemes. RP designs with four [38] and two [39, 31] wells per period have also been tested. The latter, while having the benefit of a short period and thus large gain per length, has shown higher current densities than expected, possibly due to leakage currents [40]. This effect could be reduced in future designs by using materials with higher conduction band offset [41], something that is explored in Chap. 6.2. Finally, gain in a simple superlattice structure may be possible [42] (although with very limited performance), but has so far not been observed. In conclusion, despite the many QCL designs attempted, the operation is still limited to 200 K. However, several design schemes are promising and further exploration is needed in order to find their fundamental limitations.

1.3 Material systems

The materials comprising the layers of QCLs are usually lattices of semiconductors from the III and V groups, grown by molecular beam epitaxy (MBE) [43] or metal-organic chemical vapor deposition (MOCVD) [44] on a Si, GaAs, or InP substrate. While it is important to use materials with different band-gaps, they still need to have similar lattice constants, in order to form a common lattice and not break.

As seen in Fig. 1.5, it takes some care to find good matching materials. For example, GaAs is often grown on top of a Si substrate since their lattice constants are not too far apart, and due to the technological maturity of both materials. The barrier material can then be chosen to be either pure AlAs, or a mixture of GaAs and AlAs, giving a different conduction band offset without changing the lattice constant much. The GaAs/AlGaAs combination has a high maximum CBO of ~ 1 eV (using pure AlAs barriers) [45]. However, AlAs is an in-direct band gap semiconductor, with an X-valley gap lower than the direct GaAs Γ-valley gap [46]. Thus, for a ratio of AlAs to GaAs of $x > 40\%$, the X-valley

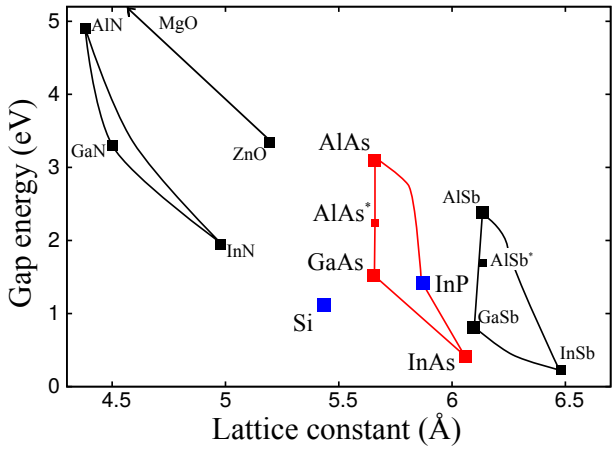


Figure 1.5: Lattice constants and Γ-valley band gaps for common III-V semiconductors, as well as the II-VI compounds ZnO and MgO (the latter lying outside of the plotted area). Red indicates materials used for QCLs in this thesis. Asterisks denote the X-valley gap of indirect bandgap materials. The lines connecting different compounds represent alloys of the connected materials, which have intermediate bandgaps and lattice constants.

the barrier material will be below the Γ-valley of the well material [46]. This may allow carriers to leak into the barrier X-valley [47, 48], and as the basis is typically expanded around the Γ-point with $\mathbf{k} = 0$ (as in eq. (2.5)) it is difficult to model such transitions accurately. This limits the CBO of the GaAs/AlGaAs material system, why it is mainly used for THz devices.

In order to reach higher frequencies, a larger CBO is necessary. From Fig. 1.5 we find InAs has a small band gap, and in combination with both GaAs and AlAs can be lattice matched to an InP substrate. In such a system, the well material would be $Ga_xIn_{1-x}As$ and the barrier material $Al_yIn_{1-y}As$. To reach an even higher CBO, we could use compositions that are not lattice matched to InP, i. e. a smaller x and a larger y. However, this introduces strain in the material [49], which will lead to breaking of the crystal if it is not compensated for.

Other compositions are used, which are interesting for improvement of future QCLs but not used in this thesis. For instance, large CBOs can be achieved using GaSb/AlSb [50] or GaN/AlN/InN [51, 52]. The latter has a very large optical phonon frequency, which might be beneficial for high-temperature THz QCLs. ZnO/MgO heterostructures have also been proposed for this purpose [53, 54], although the growth technique for these materials is not yet fully developed. So far, no QCL has worked utilizing the GaN/AlN/InN or ZnO/MgO material

systems.

1.4 Approaches to QCL modeling

As already discussed, the development of QCL designs is a computational and modeling endeavor. For this purpose, accurate yet computationally efficient models are required and these features are usually trade-offs of each other. A good summary of available modeling schemes is given in Ref. [28]. Advanced design models use genetic algorithms [55, 56, 57][Paper II], and an efficient model is a necessity. The simplest and fastest approach is using rate equations found from a simplified density matrix theory of the localized eigenbasis of the QCL lattice potential [20, 58][Paper II]. In such a model, coherences are neglected and their effect is approximated by density-density tunneling and scattering rates. Generalizing, the coherences can be included in the rate equations, giving a more complete picture of the tunneling and scattering, where the latter is taken into account via Fermi's golden rule type of calculations [59, 60]. The ensemble Monte Carlo (MC) method [61, 62, 63] is used to statistically model scattering by randomly allowing particles in the simulation undergo different scattering processes. This approach has the advantage of relatively easily including electron-electron scattering [61].

Not only do we model QCLs for the purpose of designing new structures, but also to understand basic physical processes and other details of the operating devices. In order to capture these phenomena, more advanced models taking into account the full coherent terms of the density matrix and scattering matrices are used. The most general treatment of scattering is using non-equilibrium Green's function (NEGF) theory [64, 65, 66, 67, 68, 69, 70] [Paper III], where the full statistical results are found by computing an infinite series of scattering processes, which can be visualized by Feynman diagrams. The approach here is to terminate this series at a suitable point, so that sufficient detail is covered in order to model the physics of the system under investigation. Even though the scattering series (self-energy diagrams) is terminated at some suitable point, the theory is fully consistent up to that point, so that in this framework the physics is treated systematically and internally fully consistent.

Chapter 2

Theory

In this section, the theoretical framework required in Chaps. 3 and 4 is outlined. First, the structure of the Hamiltonian for a QCL is discussed. From this Hamiltonian, a suitable set basis functions are derived, where we employ Wannier functions [71, 32]. The NEGF model is then presented, and it is explained how different observables, such as current density and gain, are calculated from the Green's functions. Finally, a brief explanation of the calculation for the scattering part of the Hamiltonian is given, as well as it's relation to the self-energies of the NEGF theory.

2.1 Hamiltonian

The fundamental concept for all quantum systems expressible by a Hermitian Hamiltonian is a set of basis states. For a given Hamiltonian, there exists, in general, a semi-infinite choice of basis functions, which can be found by a rotation of the Hamiltonian matrix. One particular choice of high importance and which often corresponds to the most convenient basis, is the basis of energy eigenstates in which the Hamiltonian is diagonal. The eigenvalues of this basis set are the energy eigenvalues. This is a well defined procedure for closed systems (i. e. systems where the total energy, momentum, and particle number are conserved). However, real systems as e. g. the QCL active region, are connected to other systems and baths, and thus they can exchange energy, momentum, and particles. To model such interacting systems, it would in principle be necessary to include all other systems into the Hamiltonian that are in contact with the reduced system of interest. However, this is impossible in most cases. In our particular situation, our reduced system is constituted by the electrons traveling

through the gain medium of the QCL, and these can exchange momentum and/or energy with impurities, crystal defects, other electrons, and phonons. The latter are treated as baths, i. e. their properties are not altered by the interaction with the reduced system. The particle-bath interaction makes it inconvenient, if not impossible, to find the true energy eigenstates of the extended system. In any case, these true eigenstates would not be suitable for describing the dynamics of the reduced system. For this reason we chose as our basis the energy eigenstates of the single-electron part $H_{\text{s-p}}$, rather than the total Hamiltonian

$$\begin{aligned} H_{\text{el.}} &= H_0 + V_{\text{dc}} + V_{\text{AC}}(t) + H_{\text{scatt}} \\ &= H_{\text{s-p}} + H_{\text{scatt}}. \end{aligned} \tag{2.1}$$

H_0 contains the kinetic part $\hat{p}^2/2m_e$ and the heterostructure potential. V_{dc} contains static electric fields, coming from the applied dc bias as well as the static mean field from ionized impurities and carriers. Terms with explicit time dependence are grouped in $V_{\text{ac}}(t)$, which contains the alternating laser field as well as the responding dynamical mean field of the carriers. Finally, $H_{\text{scatt.}}$ contains the parts non-diagonal in momentum $\mathbf{k}$ which are the complicated scattering mechanisms. These will be described in more detail in Sec. 2.7.1. The most general expression for the single-particle Hamiltonian in the last line of eq. (2.1) in the presence of an electro-magnetic field is

$$\hat{H}_{\text{s-p}} = \frac{(\hat{p} - e\vec{A}(\vec{r},t))^2}{2m_0} + V_L(\vec{r}) + e\Phi(\vec{r},t), \tag{2.2}$$

where $\hat{p} = -i\hbar\vec{\nabla}$ is the momentum operator, $\vec{A}(\vec{r},t)$ is the vector potential, V_L is the atomic heterostructure potential, and Φ is the scalar potential. From now on we will denote 3D vectors by an arrow ($\vec{r}$) and 2D vectors in the x-y plane by bold face ($\mathbf{r}$). Quantum mechanical operators are denoted by a "hat" (as in $\hat{H}$) Here, both the static and dynamical mean field are included in $\Phi(\vec{r},t)$. Alternatively, as described in the following section, the basis can be calculated without electric fields, which are then treated as perturbations of the system.

For convenience, we also define $H_{\text{s-p}}$ without electric fields as

$$H_L = \frac{\hat{p}^2}{2m_0} + V_L(\vec{r}). \tag{2.3}$$

Within a given semiconductor material, $V_L(\vec{r})$ is a lattice periodic function and the eigenstates of H_L are Bloch states $\varphi_{n\vec{k}} = e^{i\vec{k}\vec{r}}u_{n\vec{k}}(\vec{r})$. In a first step, we will now derive the envelope functions of the semiconductor potential by an expansion of the lattice periodic functions $u_{n\vec{k}}(\vec{r})$ around the Γ-point $\vec{k} = 0$. In a second step, we introduce the heterostructure Bloch states originating from the

periodicity of the heterostructure (on a larger scale than the lattice periodicity). Then, we will find another set of states which have the same eigenenergies, but which are maximally localized. These are the Wannier states which are used in the NEGF computations.

2.2 Envelope functions

In the following, the procedure for finding the eigenfunctions to $H_{\text{s-p}}$ of eq. (2.2) is derived. In order to get an idea how to treat applied electric fields, we keep them in this derivation, although they are not used in the actual computations of the basis states. The time-dependent Schrödinger equation for the one-electron wavefunction $\Psi(\vec{r},t)$ is

$$i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t) = \left(\frac{(\hat{p}-e\vec{A}(\vec{r},t))^2}{2m_0} + V_L(\vec{r}) + e\Phi(\vec{r},t)\right)\Psi(\vec{r},t). \tag{2.4}$$

Here, it is convenient to employ envelope functions $\psi_n(\vec{r},t)$, which have no Fourier components outside the first Brillouin zone, by setting

$$\Psi(\vec{r},t) = \sum_n \psi_n(\vec{r},t)u_{n,\vec{k}=0}(\vec{r}). \tag{2.5}$$

Note that $u_{n,\vec{k}=0} \equiv u_n$ satisfies $H_L u_n(\vec{r}) = E_n u_n(\vec{r})$, where n denotes the band index. The above equation thus becomes

$$\begin{aligned} i\hbar\sum_n \dot{\psi}_n(\vec{r},t)u_n(\vec{r}) &= \sum_n \Big[E_n\psi_n(\vec{r},t)u_n(\vec{r}) - \left(\frac{\hat{p}^2}{2m_0}\psi_n(\vec{r},t)\right)u_n(\vec{r}) \\ &+ \frac{1}{m_0}(\hat{p}\psi_n(\vec{r},t))(\hat{p}u_n(\vec{r})) \\ &- \frac{e}{m_0}\vec{A}(t)\Big((\hat{p}\psi_n(\vec{r},t))u_n(\vec{r}) + \psi_n(\vec{r},t)(\hat{p}u_n(\vec{r}))\Big) \\ &+ \left(\frac{e^2A^2}{2m_0} + e\Phi(\vec{r},t)\right)\psi_n(\vec{r},t)u_n(\vec{r})\Big], \end{aligned}$$

where we have assumed that $\vec{A}(\vec{r},t)$ varies slowly with $\vec{r}$ on the scale of $\psi_n(\vec{r},t)$ and $u_n(\vec{r})$. Averaging over one unit cell centered at $\vec{r}_0$, multiplying by $u_m^*(\vec{r})$ from the left and assuming that the envelope functions are constant on this scale[1] we

[1] The validity of this approximation for QCLs can be debated, since the envelope-functions can vary significantly on a scale of a few mono-layers

get

$$\begin{aligned} i\hbar\dot{\psi}_m(\vec{r}_0) &= E_m\psi_m(\vec{r}_0,t) + \frac{\hat{p}^2}{2m_0}\psi_m(\vec{r}_0,t) + \sum_{n\neq m}\frac{1}{m_0}\vec{p}_{mn}\hat{p}\psi_n(\vec{r}_0,t) \\ &- \frac{e}{m_0}\vec{A}(\vec{r}_0,t)(\hat{p}\psi_m(\vec{r}_0,t)) - \sum_{n\neq m}\frac{e}{m_0}\vec{A}(\vec{r}_0,t)\psi_n(\vec{r}_0,t)\vec{p}_{mn} \\ &+ \left(\frac{e^2A^2}{2m_0} + e\Phi(\vec{r}_0,t)\right)\psi_m(\vec{r}_0,t) \end{aligned} \tag{2.6}$$

with the definition $\vec{p}_{mn} \equiv \int_{\text{cell}} \mathrm{d}^3 r u_m^*(\vec{r})\hat{p}u_n(\vec{r})$ which is non-zero only for $m \neq n$. We have also used the orthogonality condition $\int_{\text{cell}} \mathrm{d}^3 r u_m^*(\vec{r})u_n(\vec{r}) = \delta_{m,n}$. We see from eq. (2.6) that the envelope function

$$\bar{\psi}(\vec{r},t) = \begin{pmatrix} \psi_1(\vec{r},t) \\ \psi_2(\vec{r},t) \\ \vdots \\ \psi_m(\vec{r},t) \end{pmatrix} \tag{2.7}$$

satisfies the Schrödinger-like equation

$$i\hbar\frac{\partial}{\partial t}\bar{\psi}(\vec{r},t) = \hat{\bar{H}}_{\text{eff}}\bar{\psi}(\vec{r},t). \tag{2.8}$$

Taking only the conduction band into account, we find

$$\hat{H}^c_{\text{eff}} = E_c + \frac{(\hat{p} - e\vec{A}(\vec{r},t))^2}{2m_0} + e\Phi(\vec{r},t). \tag{2.9}$$

Note the appearance of the bare mass m_0 in eq. (2.9). Taking only the conduction band into account is too crude an approximation. Treating all other bands perturbatively, we get the effective Hamiltonian in the effective mass approximation with parabolic bands, replacing m_0 by the effective mass

$$m_{\text{eff}} = m_0 m^*. \tag{2.10}$$

As we shall see in Chap. 3, including explicitly more bands in $\hat{\bar{H}}_{\text{eff}}$ will lead to an energy-dependence of the effective mass.

2.3 Bloch functions of a periodic heterostructure

In the case of a heterostructure in the z-direction, the lattice periodic functions $u_{m\vec{k}}(\vec{r})$ will be different in different regions along the z-axis. Thus, $E_n = E_n(z)$

and $H_{\text{eff}} = H_{\text{eff}}(z)$ will get a z-dependence as well. Then, the solutions to eq. (2.8) can be decomposed as

$$\psi_n(\vec{r}) \rightarrow \psi_n(z)e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.11}$$

for the lattice wave vector with in-plane components $\mathbf{k}$. Furthermore, $H_{\text{eff}}(z)$ is periodic in the case of a superlattice. This gives finally solutions with z-components $\psi_{nq}(z,t)$ and energies $E_{nq}(z)$, where q is a kind of heterostructure wave vector (associated with the periodicity in the z-direction). The energies E_{nq} are called mini-bands and $\psi_{nq}(z,t)$ Bloch functions [32].

2.4 Wannier functions

In the previous section, the procedure for finding the Bloch functions $\psi_{nq}(z)$ of the periodic heterostructure potential was outlined. However, the Bloch functions are not localized and thus not suitable for describing the dynamics within a few periods of the QCL. In our model, we choose to use Wannier states [71] as basis states instead, as these are localized and they can be obtained via a well defined prescription [32]. Another common choice are Wannier-Stark states [32], that are calculated with a specific applied bias. In contrast, the Wannier states are calculated at zero bias, and thus they have to be calculated only once. Yet another reason for choosing Wannier states is that they are found directly from the Bloch states of the periodic heterostructure, in contrast to the common method of finding the Wannier-Stark states by imposing artificial boundary conditions at the edges of a series of a few periods.

The Wannier states can be obtained by [64]

$$\phi_\nu(z - nd) = \frac{d}{2\pi}\int_{-\pi/d}^{\pi/d} \mathrm{d}q e^{-inqd}\psi_{\nu q}(z). \tag{2.12}$$

from the heterostructure envelope functions $\psi_{\nu q}(z)$, which are defined up to an arbitrary complex phase. By choosing this phase in a certain way, it is possible to construct a set of states which are localized to a few QCL periods. We have considered two approaches: either to find the basis which diagonalizes the position matrix element [64]

$$z_{\alpha\beta} = \int \mathrm{d}z\phi^*_\alpha(z)z\phi_\beta(z), \tag{2.13}$$

or by minimizing the variance of the Wannier states [72]. Comparing the two methods, we find identical results but the latter method is faster and it's implementation more convenient.

2.5 The Wannier-Stark Basis

Let us now add a constant electric field F_{dc}, uniform over the entire heterostructure. This electric field gives rise to a potential energy

$$e\phi(z) = -eF_{\mathrm{dc}}z, \tag{2.14}$$

where $e < 0$ is the electron charge. (In F_{dc} we can also include the static mean field potential.) This will change the Hamiltonian to

$$H_{\nu\nu'} \longrightarrow H_{\nu\nu'}\delta_{\nu\nu'} - eF_{\mathrm{dc}}z_{\nu\nu'} \tag{2.15}$$

which is non-diagonal in the Wannier basis. The Hamiltonian (2.15) gives the Wannier-Stark states as the eigenstates, mixing different Wannier states due to the presence of an electric field. The Wannier-Stark states form a ladder in both space and energy with the same wavefunctions appearing each heterostructure period.

While the Wannier-Stark states are not used in the actual calculations, they are a helpful tool for displaying, discussing and understanding a quantum cascade structure since they are more similar to the eigenstates of $H_{\text{s-p}}$. To this end, it is often also informative to include in an effective way the scattering part of the Hamiltonian before diagonalizing. This procedure will be discussed in more detail in Sec. 2.7.1.

2.6 The NEGF model

In order to get a full description of coherent transport in QCLs, as well as to model scattering in a consistent way, we are using the non-equilibrium Green's function method, which is described in some detail in Paper III. As the Green's function theory is mathematically involved, we only have place here for a brief glimpse at the theory.

The main idea is to start from the non-interacting Hamiltonian with known eigenstates and slowly turn on the interaction. The full interacting result can then be found by integrating from the beginning of time to the current time. The main object is the Green's function, which correlates two events at space-time coordinates $(\mathbf{r_1}, t_1) \equiv (1)$ and $(\mathbf{r_2}, \mathbf{t_2}) \equiv (2)$:

$$G(\mathbf{r_1}, \mathbf{r_2}; t_1, t_2) \equiv G(1,2) = i\langle T\{\hat{\Psi}^\dagger(2)\hat{\Psi}(1)\}\rangle \tag{2.16}$$

where the time-ordering operator $T\{\}$ orders terms with later times to the left. In order to facilitate the mathematics, this integral is performed along some

fictitious time contour, in a plane of complex time τ. In the basis of eigenstates $\Psi(1) = \sum_\alpha a_\alpha(t_1)\psi_\alpha(\mathbf{r}_1) \equiv \sum_\alpha a_\alpha(1)\psi_\alpha(1)$, where it is clear from the context what coordinate (time or space) 1 stands for. Thus,

$$G(1,2) = i\sum_{\alpha\beta}\langle T\{a^\dagger_\alpha(2)a_\beta(1)\}\rangle\psi^*_\alpha(2)\psi_\beta(1). \tag{2.17}$$

This notation will greatly simplify the following equations.

For a system in equilibrium, the NEGF approach also allows for calculating thermodynamical properties of the system, by computing the so called temperature (or Matsubara) Green's function. We are interested in a system out of equilibrium, however, and thus we go along a different contour where it is convenient to split the Green's function into four different parts; the lesser Green's function

$$G^<_{\alpha,\beta}(1,2) = i\langle a^\dagger_\alpha(2)a_\beta(1)\rangle, \tag{2.18}$$

which gives the particle density; the greater Green's function

$$G^>_{\alpha,\beta}(1,2) = -i\langle a_\beta(1)a^\dagger_\alpha(2)\rangle, \tag{2.19}$$

which gives the hole density; The retarded Green's function

$$G^R_{\alpha,\beta}(1,2) = -i\theta(t_1-t_2)\langle\{a_\beta(1),a^\dagger_\alpha(2)\}\rangle, \tag{2.20}$$

which gives the response at a time t_1 later than t_2 and whose imaginary part gives the density of states; and finally the advanced Green's function

$$G^A_{\alpha,\beta}(1,2) = i\theta(t_2-t_1)\langle\{a_\beta(1),a^\dagger_\alpha(2)\}\rangle, \tag{2.21}$$

which gives the response at a time t_1 earlier than t_2, and whose imaginary part gives the density of hole states. As we do not concern ourselves with holes - only electrons -we will be mainly interested in $G^<$ and G^R. From these definitions of the Green's functions, all relevant observables of the system can be found. For instance, the density matrix is

$$\rho_{\alpha,\beta}(\mathbf{k},t) = -iG^<_{\alpha,\beta}(\mathbf{k},t,t^+) \tag{2.22}$$

and the spatial and energetically resolved electron density

$$\rho(z,E) = \frac{2e}{2\pi\mathrm{i}S}\sum_{\mathbf{k}}\sum_{\alpha\beta}G^<_{\alpha\beta}(\mathbf{k},E)\psi^*_\beta(z)\psi_\alpha(z), \tag{2.23}$$

where S is the lateral area of the sample and we have used the space and time Fourier transformations of $G^<(1,2)$ of eq. (2.18). Knowing the density matrix

$\rho_{\alpha,\beta}(\mathbf{k},t)$, it is possible to find expressions for all desired observables by noting that the expectation value of, say, the current density is

$$\langle \hat{J} \rangle = \mathrm{Tr}\{\hat{J}\hat{\rho}\}. \tag{2.24}$$

As the Wannier states or Wannier-Stark states will not be energy-eigenstates to the total Hamiltonian, it is often useful to consider the energy resolved density of states in order to obtain information about the actual energy levels and transition energies. This can be related to the so called spectral function, which in turn is related to the retarded Green's function:

$$A(\mathbf{k},\omega) = -2\mathcal{I}\{G^R(\mathbf{k},\omega)\}. \tag{2.25}$$

(Since we are dealing with a two-dimensional system, the $\mathbf{k}$-integral over the spectral function will give a constant in ω.) So, once the Green's functions are found, we have a good tool box to analyze the system.

The mentioned (complex) time integral leading to the Green's function is very complicated, and indeed impossible to calculate exactly for a non-trivial system. However, it is possible to split the integral in to smaller parts, which can be represented by the famous Feynman diagrams. This diagrammatic expansion makes it easier to group terms into different classes; there is always an infinite amount of diagrams, but they can be summed in different ways. This makes it possible to terminate the summation at some suitable point, giving a well defined approximation and a consistent treatment of the interaction up to that point (see e. g. [73] for details). In addition, from this summation emerges the Dyson equation

$$G(1,2) = G_0(1,2) + G(1,3)\Sigma(3,4)G_0(4,2), \tag{2.26}$$

where $G_0(1,2)$ is the Green's function of the non-interacting system (internal indices like 3 and 4 above follow the normal Einstein summation convention). Σ is called the self-energy, and is given in terms of $G(1,2)$ and the interaction potential. It can also be expanded in diagrams, and the relevant approximation is made at this level. From the Dyson equation, an iterative solution can be found by making an initial guess of Σ and G which, inserted into eq. (2.26), gives an updated value of G and hence Σ. This procedure can be repeated until convergence is reach (the correct solution is found), when the updated G and Σ will be unchanged. This is the most naive way to update the self-energies; in practice we use the more elaborate Broyden scheme [74] which takes into account previous iterations of the self-energy in order to quicker reach convergence.

In our treatment of non-equilibrium Green's functions, the Dyson equation acquires a more complicated form for involving the retarded and advanced quantities of the Green's functions and self-energies. Then, these quantities are related to $G^<$ via the Keldysh relation

$$G^<(1,2) = G^R(1,3)\Sigma^<(3,4)G^A(4,2), \tag{2.27}$$

here written schematically with implied sums over the internal indices 3 and 4. The Dyson equation and Keldysh relation can also be Fourier transformed into functions of momentum and energy, rather than space and time coordinates, and this is the form used in the actual computations.

It should be clear now, that the most complicated task is to compute the self-energies at each iteration of the Dyson equation. This will be described in Sec. 2.7.1.

2.7 Temporal response

In order to capture the time response of the system to the harmonic electric field with frequency ω, we may assume that all observables can be expanded as a Fourier series in ω. Then, we can get the time-dependence of, e. g. , the density matrix as

$$\rho_{\alpha\beta}(\mathbf{k},t) = \sum_{h=-N_h}^{N_h} \rho_{\alpha\beta,h}(\mathbf{k}) \cdot e^{-ih\omega t}, \tag{2.28}$$

or the current density

$$J(z,t) = \sum_{h=-N_h}^{N_h} J_h(z) \cdot e^{-ih\omega t}, \tag{2.29}$$

where the infinite sum is truncated at $h = \pm N_h$. In linear response $N_h = 1$. This implies all Green's functions (and self-energies) can be expanded in terms of their Fourier components $G_{\alpha,\beta,h}(E_k, E)$, where terms with $h > 0$ obey similar equations as those with $h = 0$.

2.7.1 Scattering and self-energies

In the previous section, we have developed our state basis and basic quantum mechanical observables for the pure system with an analytic Hamiltonian. Now, we will consider physical scattering processes which couple this pure system

of electrons to a realistic model of the environment. These effects are lattice vibrations (phonons), material impurities (ionized donors, interface roughness, and alloy disorders), as well as the interaction among the electrons themselves. Note that the electro-magnetic field has already been treated in the classical limit, and thus we will not include the interaction of photons explicitly.

Each scattering process can be described by an electronic scattering potential. In first order perturbation theory, for example, this potential will enter squared in the matrix element in Fermi's golden rule. In the NEGF framework, this matrix element will instead show up in the self-energies.

As mentioned in the beginning of this section, we need to approximate the self-energies in a consistent way, by summing only a chosen subset of Feynman diagrams. The most common choice, which we also use, is the self-consistent Born approximation [Paper III]. In this approximation, the self-energy for elastic scattering takes the form (in momentum and energy space)

$$\Sigma^{</R}_{\alpha\alpha'}(E,\mathbf{k}) = \sum_{\beta\beta',\mathbf{k}'} \langle V_{\alpha\beta}(\mathbf{k}-\mathbf{k}')V_{\beta'\alpha',\mathbf{k}'}(\mathbf{k}'-\mathbf{k})\rangle_{\mathrm{imp}} G^{</R}_{\beta\beta'}(E,\mathbf{k}'), \tag{2.30}$$

where $V(\mathbf{q})$ is the scattering potential and the angled bracket implies an average over a statistical scattering distribution. For in-elastic scattering, there will be additional terms needed for total energy and momentum conservation.

We will assume that the system is (statistically) rotational invariant in the x-y plane. Thus, we may average all scattering terms over the angles in this plane, giving only an explicit dependence on the modulus $k = |\mathbf{k}|$ of the two-dimensional momentum $\mathbf{k}$. Since the energy associated with this momentum is just $E_k \propto k^2$, we will in the following substitute all dependence on the modulus of the momentum for E_k.

As a side note, it is often desirable to plot the energy levels and wave functions of the QCL structure, as in Fig. 1.4. However, the eigenstates are not directly accessible due to the complicated nature of the Hamiltonian. When transforming to the Wannier-Stark basis (2.15) we completely neglected the scattering part of $\hat{H}$, and as a result they are not eigenstates of the QCL system. For example, their energies do not correspond to the peaks of the spectral function and $\rho(z,E)$ [Paper IV]. In order to get a closer agreement between the Wannier-Stark energies and the energy where, e. g. the spectral function has it's peaks, the self-energies $\Sigma(E_\alpha)$ can be added to $H_{\nu\nu'}$ in eq. (2.15) before diagonalizing, giving

$$H_{\nu\nu'} \to H_{\nu\nu'}\delta_{\nu\nu'} - eF_{\mathrm{dc}} z_{\nu\nu'} + \Sigma_{\nu\nu'}(E_\nu)\delta_{\nu\nu'} \tag{2.31}$$

where E_α is the energy of the Wannier-Stark state computed without scattering. This typically gives a shift down in energy of the Wannier-Stark states, as

compared to the case without scattering, and a much closer match to the peaks of the spectral function.

2.7.2 Elastic Scattering

The terms in $H_{\text{scatt.}}$ of eq. (2.1) from interface roughness, impurity scattering, and alloy disorder constitute the elastic scattering. These are all treated similarly with the self-energy (see Paper III for details)

$$\Sigma_{\alpha\alpha'}(E, E_k) = \sum_{\beta,\beta'} \int dE_{k'} \underbrace{\frac{\rho_0 A}{2} \frac{1}{2\pi} \int_0^{2\pi} d\varphi \langle U_{\alpha\beta}(E_k, E_{k'}, \varphi) U_{\beta'\alpha'}(E_k, E_{k'}, \varphi) \rangle_{\text{imp}}}_{X_{\alpha\alpha'\beta\beta'}(E_k, E_{k'})} G_{\beta\beta'}(E, E_{k'}), \tag{2.32}$$

where $G_{\beta\beta'}(E, E_{k'})$ is the electron Green's function in frequency and momentum space, A is the sample area and $\rho_0 = em_{\text{eff}}/\hbar\pi$. This defines the X-tensors that have to be evaluated for each kind of elastic scattering mechanism. In eq. (2.32), the integral is over momentum angle φ in the x-y plane.

2.7.3 Interface roughness scattering

As we will study the effects of interface roughness in some detail in, for example, Sec. 5.1, this particular scattering mechanism will serve as a thorough example of how different self-energies are calculated. The matrix element for interface roughness scattering contains the potential

$$U_{\alpha\beta}(\mathbf{p}) = \sum_j \frac{1}{A} \int d^2 r e^{-i\mathrm{p}\cdot\mathrm{r}} \eta_j(\mathbf{r}) \Delta E_c \Psi^{\alpha *}(z_j) \Psi^{\beta}(z_j), \tag{2.33}$$

expressed in terms of the deviation from a defined interface $\eta(\mathbf{r})$, see Fig. 2.1, and the conduction band offset ΔE_c. The sum is over all interfaces j. Performing the impurity average over the potential, we have to compute the correlation function

$$\frac{1}{A} \int \mathrm{d}^2 r \int \mathrm{d}^2 r' e^{\mathrm{i}\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} \eta(\mathbf{r}) \eta(\mathbf{r}') = \int \mathrm{d}^2 r e^{\mathrm{i}\mathbf{q}\cdot(\mathbf{r})} \langle \eta(\mathbf{r}) \eta(0) \rangle \equiv f(\mathbf{q}), \tag{2.34}$$

where $f(\mathbf{q})$ is the Fourier transformation of the spatial correlation function $\langle \eta(\mathbf{r}) \eta(0) \rangle$ for the fluctuations. $f(\mathbf{q})$ will depend on the average roughness height Δ and the average island length Λ, though the exact form of the roughness distribution $\eta(\mathbf{r})$ in Fig. 2.1 in real systems is not known.

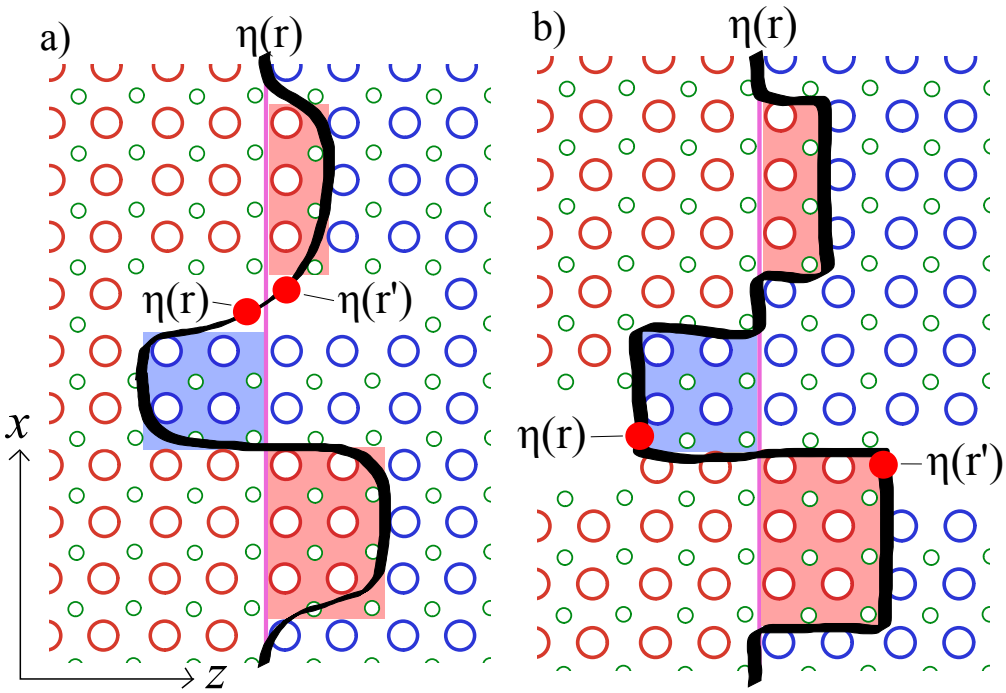


Figure 2.1: Schematic view of the interface roughness across the well (red) and barrier (blue) material. $\eta(\mathbf{r})$ is the z-coordinate of the interface for given in-plane coordinates $\mathbf{r}$ (it may also be taken as the deviation from the average interface). In a), the interface as seen by the incident electrons is smooth, and in b) it is sharp.

Most commonly a Gaussian distribution [75, 76, 77] is used here, with

$$\begin{aligned} \langle \eta(\mathbf{r})\eta(0)\rangle &= \Delta^2 \exp\left(-\frac{|\mathbf{r}|^2}{\Lambda^2}\right) \\ \rightarrow \quad f(\mathbf{q}) &= \pi\Delta^2\Lambda^2 \exp\left(-\frac{\Lambda^2|\mathbf{q}|^2}{4}\right). \end{aligned} \tag{2.35}$$

This is a good choice if $\eta(\mathbf{r})$ is smooth like in Fig. 2.1 (a). Then, the slope of $\langle \eta(\mathbf{r})\eta(0)\rangle$ at $\mathbf{r} \rightarrow 0$ vanishes (clearly, the average slope has to vanish for the mean value to be zero). This is also the case for the Gaussian function.

In our model we mostly use an exponential distribution, based on the findings

in Refs. [78, 79, 80, 81]

$$\begin{aligned} \langle \eta(\mathbf{r})\eta(0)\rangle &= \tilde{\Delta}^2 \exp\left(-\frac{|\mathbf{r}|}{\tilde{\Lambda}}\right) \\ \rightarrow \quad f(\mathbf{q}) &= \frac{2\pi\tilde{\Delta}^2\tilde{\Lambda}^2}{\left(1+\tilde{\Lambda}^2|\mathbf{q}|^2\right)^{3/2}}. \end{aligned} \tag{2.36}$$

This can be further motivated if $\eta(\mathbf{r})$ is a sharp function as in Fig. 2.1 (b). If then $\eta(0)$ is on an island, the product $\eta(\mathbf{r})\eta(0)$ will be constant as long as $\mathbf{r}$ is on the same island, but becomes negative as it crosses the edge. Averaging over all reference points, some points will be precisely on the rim of an island as in Fig. 2.1 (b). Therefore, $\langle \eta(\mathbf{r})\eta(0)\rangle$ will have a finite negative slope for small r, and this is consistent with an exponential distribution. Both roughness distributions will be compared in Sec. 5.1.

Elastic scattering plays an important role in redistributing carriers within subbands. Thus, $f(\mathbf{q})$ for both distributions would give similar results when the functions $f(\mathbf{q})$ as well as their second derivatives coincide at the maximum $q = 0$. This is achieved with the translation

$$\tilde{\Lambda} = \Lambda/\sqrt{6} \quad \text{and} \quad \tilde{\Delta} = \sqrt{3}\Delta \qquad \text{(Expon. 1)}. \tag{2.37}$$

This is probably very well suited for THz QCLs, where the momentum exchange is typically small. For IR QCLs, with much larger momentum transfer, e. g. for transitions between the laser states where substantial in-plane kinetic energy has to be transferred for energy conservation, this transformation would give stronger scattering at large $\mathbf{q}$ for the exponential distribution. A second natural translation is given by requiring an identical average fluctuation height $\langle \eta(\mathbf{r}=0)\eta(0)\rangle$ and identical $f(\mathbf{q}=0)$ for the Gaussian and exponential distribution. This provides

$$\tilde{\Lambda} = \Lambda/\sqrt{2} \quad \text{and} \quad \tilde{\Delta} = \Delta \qquad \text{(Expon. 2)}. \tag{2.38}$$

Fig. 2.2 shows the different distribution functions, translated via the relations (2.37) and (2.38) starting from a Gaussian distribution with $\Delta = 0.1$ nm and $\Lambda = 9$ nm. For typical infrared QCLs with $\hbar\omega = 150$ meV, this implies $q \approx 0.4$ nm^{-1}, which is where the Expon. 2 and Gaussian distributions cross. These would therefore be expected to provide similar results, while the Expon. 1 distribution would result in a larger scattering rate at the laser transition energy.

In general, knowing the form of the auto-correlation function provides the matrix elements

$$\langle U_{\alpha,\beta}(\mathbf{q})U_{\alpha',\beta'}(-\mathbf{q})\rangle = \sum_j \frac{1}{A^2} f_j(\mathbf{q})\Delta E_c^2 \psi_\alpha^{c*}(z_j)\psi_{\alpha'}^{c*}(z_j)\psi_\beta^c(z_j)\psi_{\beta'}^c(z_j), \tag{2.39}$$

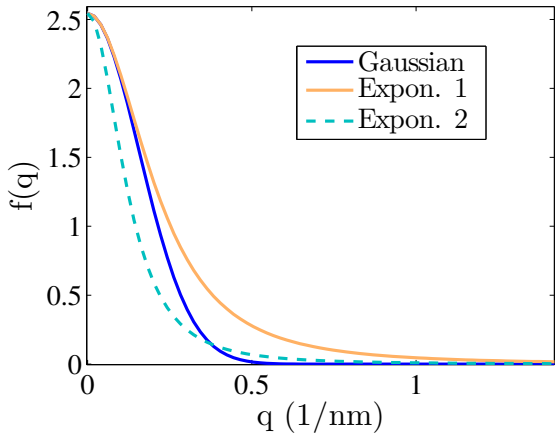


Figure 2.2: Roughness distributions $f(q)$, for a Gaussian (eq. (2.35)) and two different exponential (eq. (2.36) distributions. The Expon. 1 and Gaussian distribution functions are very similar for small values of q, while the Gaussian and the Expon. 2 distributions are of similar magnitude above $q \approx 0.4$/nm. This momentum exchange corresponds to an energy of a typical IR laser transition of 150 meV. The Expon. 1 distribution is significantly larger in this region. Figure taken from Paper V.

where the sum over j takes into account all interfaces in one period.

2.8 Current density

The definition of the current density

$$J(z,t) = e\mathcal{R}\left\{\left\langle \frac{\hat{p}_z - eA_z(z,t)}{m_c(z)} \right\rangle\right\} \tag{2.40}$$

in the effective mass approximation can be directly used to find the current density in terms of the density matrix (the expectation value is given by the trace over the density matrix), and hence the lesser Green's function as described in Paper III. However, this assumes that the effective mass $m_c(z)$ is energy independent. This situation will change when we generalize eq. (2.8) to account also for the valence bands in Chap. 3.

With a time-varying electric field present, eq. (2.29) implies the current density can be written as

$$J(z,t) = J_0(z) + J_1^{\cos}(z)\cos(\omega t) + J_1^{\sin}(z)\sin(\omega t) + \sum_{|h|>1} J_h(z)e^{-ih\omega t}, \tag{2.41}$$

where $J^{\cos}(z)$ (in phase with the ac field) and $J^{\sin}(z)$ (out of phase with the ac field) stem from the terms with $h = \pm 1$.

Frequently, we want to examine the current density resolved in space and energy, just like the electron density in eq. (2.23). We get this from eq. (2.40) in Lorenz gauge (with $A_z = 0$) and straightforward calculations to

$$J(z,E) = \frac{e\hbar}{m_{\text{eff}}} \int \mathrm{d}E_k \sum_{\alpha\beta} G^{<}_{\beta\alpha}(E_k, E) \left(\psi^*_\alpha(z) \frac{\partial \psi_\beta(z)}{\partial z} - \frac{\partial \psi^*_\alpha)(z)}{\partial z} \psi_\beta(z) \right), \tag{2.42}$$

where eq. (2.23) was used.

2.9 Gain from complex conductivity

In a material where Ohm's law $\vec{J} = \sigma \vec{E}$ is valid, the curl of the magnetic field becomes [82]

$$\nabla \times H = -i\omega(\epsilon_r \epsilon_0 + i\sigma/\omega)\vec{E}, \tag{2.43}$$

where ϵ_r is the static dielectric constant and σ is the conductivity. On the other hand we may interpret the ac currents as shifts in the polarization:

$$\nabla \times H = -i\omega\epsilon(\omega)\vec{E}. \tag{2.44}$$

Thus we identify $\epsilon(\omega) = \epsilon_r\epsilon_0 + i\sigma/\omega$, or

$$\sigma = -i\omega(\epsilon(\omega) - \epsilon_r\epsilon_0). \tag{2.45}$$

The complex wave vector of the ac field is written $k = \beta + i\alpha/2$, where α is the absorption coefficient giving rise to a decay of the intensity $I \propto e^{-\alpha z}$ as the wave propagates through the material. The phase velocity relates the wave vector to the dielectric function by

$$v_{\text{phase}} = \frac{\omega}{k} = \frac{1}{\sqrt{\mu\epsilon(\omega)}}, \tag{2.46}$$

giving the system of equations

$$\begin{cases} \beta^2 - \frac{\alpha^2}{4} & = \frac{\omega^2}{c^2} \Re\{\frac{\epsilon}{\epsilon_0}\} \\ \beta \cdot \alpha & = \frac{\omega^2}{c^2} \Im\{\frac{\epsilon}{\epsilon_0}\}. \end{cases} \tag{2.47}$$

Solving (2.47) for α and taking ω large,

$$\alpha(\omega) \approx \frac{\Re\{\sigma(\omega)\}}{c\epsilon_0\sqrt{\epsilon_r}} = \frac{J_1^{\cos}}{F_{ac} c\epsilon_0 \sqrt{\epsilon_r}}, \tag{2.48}$$

where the dependence on the frequency ω is explicit. (The absorption $\alpha(\omega)$ also has a z-dependence since $J_1^{\cos} = J_1^{\cos}(z)$, c. f. eq. (2.41)). We see that the gain $g(\omega) = -\alpha(\omega)$ of the sample can be obtained by calculating the amplitude of the part of the current that is in phase with the applied electric field. While having only an explicit dependence on terms linear in ω, the gain is affected by higher order terms indirectly since the Dyson equation and Keldysh relation couple Green's functions with different h.

2.10 Gain from Fermi's Golden Rule

In many cases, e. g. when the time-dependence of the current density is not available, gain is calculated from an expression derivable from Fermi's golden rule [44, 60]

$$G_{fi}(\omega) = \frac{e^2 \Delta n_{fi} \Delta E_{fi}^2 z_{fi}^2}{2\hbar^2 \omega n_r c \epsilon_0 d} \frac{\Gamma_{\mathrm{w}}}{(\Delta E_{fi} - \hbar\omega)^2 + \Gamma_{\mathrm{w}}^2/4}, \tag{2.49}$$

where Δn_{fi} is the inversion, ΔE_{fi} is the energy difference between the initial and final states, z_{fi} is their dipole matrix element, and Γ_{w} is the full width at half maximum of the transition. The total gain would be the sum of eq. (2.49) over all initial and final states. Here we have the problem of knowing beforehand what the width of the transition is, and second order processes such as dispersive gain [83] are neglected. It is useful also within our model, however, to get a quick estimate of the gain without running computations with $N_h > 1$.

2.11 Gain clamping and output power

As the field strength F_{ac} increases due to stimulated emission between inverted energy levels, eq. (2.48) implies that the gain decreases. This can be understood as being a consequence of diminished inversion on the gain transition, as more electrons undergo stimulated emission and transition from the upper to the lower state. When the gain over the whole structure reaches the same level as the total losses α_{tot} in the system, i. e. $g = -\alpha_{\mathrm{tot}}$, the stimulated emission of photons exactly balances the absorption of photons, and the gain reaches a stable level. This is called gain clamping.

We can arrive at a simple estimate of this effect by the following argument, as outlined in Paper IV: The gain $G_{fi}(\omega)$ of eq. (2.49) is proportional to the inversion Δn_{fi} on a certain transition. But, to first order a field of intensity F_{ac}^2 will reduce the inversion by a factor $\Delta n_{fi} \propto -F_{\mathrm{ac}}^2 G_{fi}(\omega)$. If we know that, at

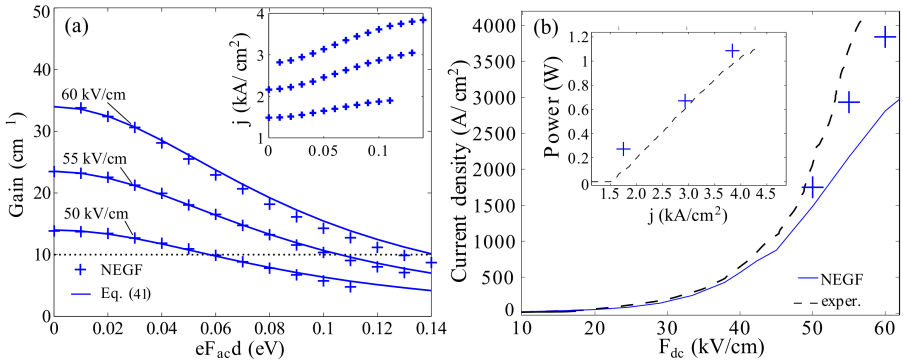


Figure 2.3: (a) Gain clamping simulated by the full NEGF model and by eq. (2.51) and (b) current density simulated using the full NEGF model with and without gain clamping, compared to the experimental data from Paper IV. The figures are taken from Paper IV.

vanishing field strength $G(F_{\mathrm{ac}} = 0) = G_0$, then

$$G(F_{\mathrm{ac}}) = G_0 - b \cdot G(F_{\mathrm{ac}}) F_{\mathrm{ac}}^2. \tag{2.50}$$

Thus,

$$G(F_{\mathrm{ac}}) = \frac{G_0}{1 + bF_{\mathrm{ac}}^2}, \tag{2.51}$$

where $b = \tau e^2 z_{fi}^2 / \hbar\Gamma_{\mathrm{w}}$ is a constant for the structure depending on the transition time through one period τ and the dipole matrix element z_{fi}^2. As seen in Fig. 2.3, this simple expression captures the clamping behavior well. The inset in Fig. 2.3 (a) shows the increase of the current density as the field strength increases. By noting for which F_{ac} the gain reaches the level of the losses (10 cm^{-1} in Fig. 2.3 (a)), it is possible to extract the output power of the device as follows. Consider a cavity with one semi-reflecting mirror with reflectivity R. Then, for a power density P_{in} inside the cavity, the outgoing power density $P_{\mathrm{out}} = (1 - R)P_{\mathrm{in}}$. If we neglect the reflected wave, the power density inside the cavity is given by the time-averaged Poynting vector (see e. g. Ref. [82]) $P_{\mathrm{in}} = |\mathbf{S}| = \frac{1}{2}\epsilon_0 n_r c F_{\mathrm{ac}}^2$. (This provides a lower bound for the output power, and in Ref. [84] the light is additionally modeled as a standing wave, giving an upper bound for the output power.) Some fraction $1/\Gamma_{\mathrm{conf}}$ of the electro-magnetic wave may not cover the gain region of the QCL and does not contribute to the output power. Thus, the total output power is

$$P = (F_{\mathrm{ac}})^2 \frac{n_r c \epsilon_0 A (1 - R)}{2\Gamma_{\mathrm{conf.}}}, \tag{2.52}$$

where A is the mirror area. Using this expression, the output power for the structure and G_0 from Paper IV is shown in the inset of Fig. 2.3 (b). This

is not far from the experimental values. We also see that the current density calculated in the NEGF model agrees better with the experiment when the increased intensity in Fig. 2.3 (a) is considered.

Chapter 3

The two-band model

In simulations using the effective mass approximation of eq. (2.8), the energies of the states high in the conduction band become inaccurate. A cure for this is to take only two bands into account and obtain a closed expression for the conduction band components ψ_c only. While this does give the correct energies, the calculated basis is not orthonormal, giving fallacious results as we shall see in Sec. 3.3. The only consistent solution is to fully take into account the components from the other bands as well, and normalize the total wavefunction. This gives both the correct energies of the Wannier states, and an orthonormal Wannier basis. Here, we restrict to only the conduction and valence bands.

From equations (2.6)-(2.8) we see that, restricting to the conduction band c and valence band v, the two-component envelope function

$$\bar{\psi}(z,t) = \begin{pmatrix} \psi_c(z,t) \\ \psi_v(z,t) \end{pmatrix} \tag{3.1}$$

satisfies the wave equation

$$i\hbar \frac{\partial}{\partial t} \bar{\psi}(z,t) = \hat{\bar{H}}_{\text{eff}} \bar{\psi}(z,t) \tag{3.2}$$

with

$$\hat{\bar{H}}_{\text{eff}} = \begin{pmatrix} E_c(z) + \frac{(\hat{p}-eA(z,t))^2}{2m_0} + e\Phi(z,t) & \frac{p_{cv}}{m_0}(\hat{p} - eA(z)) \\ \frac{p_{vc}}{m_0}(\hat{p} - eA(z)) & E_v(z) + \frac{(\hat{p}-eA(z,t))^2}{2m_0} + e\Phi(z,t) \end{pmatrix}. \tag{3.3}$$

In the absence of an electro-magnetic field and neglecting the $\hat{p}^2$ term, the effective Hamiltonian becomes

$$\hat{\bar{H}}_{\text{eff}} = \begin{pmatrix} E_c(z) & \frac{p_{cv}}{m_0}\hat{p}_z \\ \frac{p_{vc}}{m_0}\hat{p}_z & E_v(z) \end{pmatrix}, \tag{3.4}$$

and this will be the starting point for finding our new basis states in the two-band model.

In order to see the connections between the components of the envelope function (3.1), we solve the time-independent Schrödinger equation for $\psi(z)$ and get separate equations for the conduction and valence band components, respectively:

$$E_c(z)\psi_c(z) + \frac{p_{cv}}{m_0}\hat{p}_z\psi_v(z) = E\psi_c(z), \tag{3.5}$$

$$E_v(z)\psi_v(z) + \frac{p_{vc}}{m_0}\hat{p}_z\psi_c(z) = E\psi_v(z). \tag{3.6}$$

Inserting ψ_c from eq. (3.5) into eq. (3.6) and vice versa, we get

$$\psi_c(z) = \frac{1}{E - E_c}\frac{p_{cv}}{m_0}\hat{p}_z\psi_v(z), \tag{3.7}$$

$$\psi_v(z) = \frac{1}{E - E_v}\frac{p_{vc}}{m_0}\hat{p}_z\psi_c(z) \tag{3.8}$$

which can again be inserted into eqs. (3.5) and (3.6) to yield closed expressions for ψ_c and ψ_v:

$$E_c\psi_c + \hat{p}_z\left[\frac{1}{E_v - E}\frac{|p_{cv}|^2}{m_0^2}\right]\hat{p}_z\psi_c \equiv E_c\psi_c + \hat{p}_z\frac{1}{2m_c(E,z)}\hat{p}_z\psi_c = E\psi_c, \tag{3.9}$$

$$E_v\psi_v + \hat{p}_z\left[\frac{1}{E_c - E}\frac{|p_{cv}|^2}{m_0^2}\right]\hat{p}_z\psi_v \equiv E_v\psi_v + \hat{p}_z\frac{1}{2m_v(E,z)}\hat{p}_z\psi_v = E\psi_v. \tag{3.10}$$

Here we used that $p_{cv} = p_{vc}^*$, and defined the effective masses as

$$m_c(E,z) = -\frac{1}{2}m_0^2\frac{E - E_v(z)}{|p_{cv}|^2} = m_c(E_c,z)\frac{E - E_v(z)}{E_g(z)} \tag{3.11}$$

and

$$m_v(E,z) = -\frac{1}{2}m_0^2\frac{E - E_c(z)}{|p_{cv}|^2} = m_c(E_c,z)\frac{E - E_c(z)}{E_g(z)}, \tag{3.12}$$

where $m_c(E_c,z)$ is the effective mass at the conduction band edge. The two-component enevelope functions are normalized according to

$$\int \mathrm{d}z|\bar{\psi}|^2 = \int \mathrm{d}z(|\psi_c|^2 + |\psi_v|^2) = 1. \tag{3.13}$$

Eq. (3.9) can be used to find the conduction band Bloch functions $\psi_c(z)$ by using the transfer matrix method, as described in Ref. [32]. Then, the valence band components can be found via eq. (3.8), and the Wannier functions are found as in Sec. 2.4 for each band $\nu = \{c, v\}$.

3.1 Current density in the two-band model

This section is a detailed background of the model section in Paper I. The well-known result for the quantum mechanical current for free electrons is

$$J(z,t) = \langle \frac{e}{V}\frac{\partial \hat{z}}{\partial t}\rangle = \langle \frac{e}{V}[\hat{z}, H]\rangle = \tag{3.14}$$
$$e\sum_{\alpha\beta}\Big[\frac{\hbar}{i2m_0}\left(\varphi^*_\beta(z)\frac{\partial\varphi_\alpha(z)}{\partial z} - \frac{\partial\varphi^*_\beta}{\partial z}\varphi_\alpha(z)\right) - eA(z,t)\frac{\varphi^*_\beta\varphi_\alpha(z)}{m_0}\Big]\frac{2}{S}\sum_{\mathbf{k}}\rho_{\alpha\beta}(\mathbf{k},t)$$

where V is the volume and S is the cross-sectional area of the space with charge density $\rho_{\alpha\beta}(\mathbf{k},t)$. In the one-band model, a similar result is obtained for the envelope functions by defining an effective mass, which is z-dependent [64]. This model then requires that the same effective mass can be attributed to all quantum states in the system. In the two-band model this is not the case, since the effective mass in eq. (3.11) is energy-dependent and so each quantum state will have a specific mass. To find the current density operator in this situation, we have to start from more basic principles, such as the continuity equation or the time derivative of the position operator. In the following we will use the former, but the latter gives the same result.

Starting with the Hamiltonian (3.3) and the two component basis states (which may be calculated without external fields)

$$\bar{\psi}_\alpha(z) = \begin{pmatrix}\psi^\alpha_c(z)\\ \psi^\alpha_v(z)\end{pmatrix} \tag{3.15}$$

we get the Hamiltonian in second quantization

$$\hat{H} = \hat{\bar{\Psi}}^\dagger(z)\bar{\bar{H}}_{\text{eff}}(z)\hat{\bar{\Psi}}(z) = \sum_{mn} H_{mn}a^\dagger_n a_m, \tag{3.16}$$

with $H_{mn} = \int \mathrm{d}z\bar{\psi}_m\hat{\bar{H}}_{\text{eff}}\bar{\psi}_n$. Here, we used the field operators

$$\hat{\bar{\Psi}}(z) = \sum_n \bar{\psi}_n(z)\hat{a}_n = \sum_n \begin{pmatrix}\psi^n_c(z)\\ \psi^n_v(z)\end{pmatrix}\hat{a}_n$$
$$\hat{\bar{\Psi}}^\dagger(z) = \sum_n \bar{\psi}^*_n(z)\hat{a}^\dagger_n = \sum_n \begin{pmatrix}\psi^{n*}_c(z)\\ \psi^{n*}_v(z)\end{pmatrix}\hat{a}^\dagger_n.$$

The time-dependent Schrödinger equation $i\hbar\frac{d}{dt}\bar{\psi} = \bar{\bar{H}}_{\text{eff}}\bar{\psi}$ gives the time dependence of the wavefunction components:

$$i\hbar\frac{d}{dt}\psi_c(z) = (E_c(z) + e\Phi(z))\psi_c(z) + \frac{p_{cv}}{m_0}(\hat{p} - eA)\psi_v(z) \tag{3.17}$$

$$i\hbar\frac{d}{dt}\psi_v(z) = (E_v(z) + e\Phi(z))\psi_v(z) + \frac{p_{vc}}{m_0}(\hat{p} - eA)\psi_c(z). \tag{3.18}$$

From the continuity equation for the charge density $\rho(z)$ and the current density $j(z)$, we have that ($e < 0$)

$$e\frac{d}{dt}\rho(z) = e\frac{d}{dt}\bar{\psi}^*(z)\bar{\psi}(z) = e\frac{d}{dt}\begin{pmatrix}\psi_c^*(z) & \psi_v^*(z)\end{pmatrix}\begin{pmatrix}\psi_c(z)\\ \psi_v(z)\end{pmatrix} = e\frac{d}{dt}(\psi_c^*\psi_c + \psi_v^*\psi_v) =$$
$$\frac{e}{i\hbar}\Big[\psi_c^* p_{cv}\frac{\hat{p} - eA}{m_0}\psi_v - (p_{cv}\frac{\hat{p} - eA}{m_0}\psi_v)^*\psi_c + \psi_v^* p_{vc}\frac{\hat{p} - eA}{m_0}\psi_c + (p_{vc}\frac{\hat{p} - eA}{m_0}\psi_c)^*\psi_v\Big]$$
$$= -\frac{e}{m_0}\frac{d}{dz}\left(p_{cv}\psi_c^*\psi_v + p_{vc}\psi_v * \psi_c^m\right) = -\frac{d}{dz}j(z) \equiv -\frac{d}{dz}\bar{\psi}^*\bar{\bar{J}}\bar{\psi},$$

which defines the current density operator in analogy with eq. (3.16)

$$\hat{J} = \hat{\bar{\Psi}}^\dagger(z)\bar{\bar{J}}\hat{\bar{\Psi}}(z) \tag{3.19}$$

with

$$\bar{\bar{J}} = \frac{e}{m_0}\begin{pmatrix}0 & p_{cv}\\ p_{vc} & 0\end{pmatrix}. \tag{3.20}$$

The matrix element $p_{cv} = p_{vc}^*$ has an arbitrary phase that is chosen when calculating the valence band component. In order to get real valence and conduction band components, p_{cv} is chosen to be purely imaginary according to

$$p_{cv} = i|p_{cv}| = im_0\sqrt{\frac{E_g}{2m_c}} = -p_{vc}. \tag{3.21}$$

Thus, the expectation value of the current operator $\hat{J}(z)$ is

$$\langle\hat{J}(z,t)\rangle = \text{Tr}\{\hat{J}\hat{\rho}\} = \sum_{mn} J_{mn}(z)\rho_{nm}(t), \tag{3.22}$$

where

$$J_{mn}(z) = \frac{e}{m_0}im_0\sqrt{\frac{E_g(z)}{2m_c(z)}}(\psi_c^{n*}\psi_v^m - \psi_v^{n*}\psi_c^m). \tag{3.23}$$

The current density (3.22) is in principle independent of z in steady state, but since it is impossible to account for all states in a numerical calculation it has

proved to be more stable to take the mean value of the current over one period[1] [85]. Doing this and collecting the terms in phase ($J_1^{\cos}$) and out of phase ($J_1^{\sin}$), as well as the direct current part (J_{dc}), we write just as we did in eq. (2.41)

$$\langle J(t)\rangle = J_{\mathrm{dc}} + J_{\mathrm{ac}}(t) = J_{\mathrm{dc}} + J_1^{\cos}\cos(\Omega t) + J_1^{\sin}\sin(\Omega t) + \sum_{|h|\leq 2} J_h e^{-ih\Omega t}. \quad (3.24)$$

Here we used eq. (3.22), the fact that $\rho_{\alpha\beta}(t) = \sum_h \rho_{\alpha\beta\rho} e^{ih\Omega}$ from eq. (2.28) and $A = F_{\mathrm{ac}} \cdot (-ie^{-i\Omega t} + ie^{i\Omega t})/2\Omega$. Following Sec. 2.9, we arrive at a similar expression for the gain in terms of $j_1^{\cos}$, but using eq. (3.23) to calculate the current density.

Note that we do not include the extra broadening effects of optical transitions between bands with different dispersion $E(k)$ [86]. This is certainly an interesting effect for mid-IR lasers, and would be an interesting future implementation to the model.

3.2 Interface roughness in the two-band model

As we now involve ourselves with two-component, rather than scalar, wavefunctions, the scattering potential for IFR will take the same matrix form as the effective Hamiltonian. Again, we use interface roughness as an example of the extension of the one-band model potentials to fit the two-band model, but all other scattering matrix elements are affected as well (although in more trivial ways).

For interface roughness, we use the scattering matrix elements of eq. (2.33) in the one-band model where $E_c(z)$ is the Hamiltonian (neglecting the $\hat{p}^2$-term) in absence of external fields. In the two-band model this becomes eq. (3.4) and

$$\Delta V_j = H_{\mathrm{eff}}(z_j^-) - H_{\mathrm{eff}}(z_j^+) = \begin{pmatrix} \Delta E_{c,j} & 0 \\ 0 & \Delta E_{v,j} \end{pmatrix}. \quad (3.25)$$

Inserting this in place of $\Delta E_{c,j}$ in eq. (2.39) and the two-component wavefunctions, we get the scattering matrix elements

$$U_{\alpha\beta}(\mathbf{q}) = \sum_j \frac{1}{A} \int \mathrm{d}^2 r e^{-i\mathbf{p}\cdot\mathbf{r}} \xi_j(\mathbf{r}) \Big(\Delta E_{c,j} \psi_c^{\alpha *} \psi_c^{\beta} + \Delta E_{v,j} \psi_v^{\alpha *} \psi_v^{\beta} \Big). \quad (3.26)$$

[1] Taking $J(t) \approx \int_{1\ \mathrm{per.}} \mathrm{d}z J(z,t)/D_{\mathrm{per}}$, where D_{per} is the length of one QCL period. As indicated by eq. (3.22), if the sum over n and m does not cover the entire basis $\langle \hat{J}(z,t)\rangle$ cannot be guaranteed to be constant in z since $J_{mn}(z)$ is certainly not constant in z.

This provides

$$\begin{aligned}\langle U_{\alpha,\beta}U_{\alpha',\beta'}\rangle &= \sum_j \frac{1}{A^2} f_j(\mathbf{q})\Big(\psi_\alpha^{c*}\psi_{\alpha'}^{c*}\psi_\beta^c\psi_{\beta'}^c\Delta E_c^2 + \psi_\alpha^{v*}\psi_{\alpha'}^{v*}\psi_\beta^v\psi_{\beta'}^v\Delta E_v^2 \\ &+ \big(\psi_\alpha^{c*}\psi_{\alpha'}^{v*}\psi_\beta^c\psi_{\beta'}^v + \psi_\alpha^{v*}\psi_{\alpha'}^{c*}\psi_\beta^v\psi_{\beta'}^c\big)\Delta E_c\Delta E_v\Big). \qquad (3.27)\end{aligned}$$

which is an extension of eq. (2.39) including the valence band offset and valence band components of the wavefunction. As ΔE_c and ΔE_v have opposite signs in the structures studied in this thesis, eq. (3.27) will result in lower scattering than eq. (2.39). This is seen in paper V, while the effect is small for mid-IR devices and negligible for THz devices. Note that simply neglecting the valence band offset $\Delta E_v = 0$ in eq. (3.25), we get back the result of eq. (2.39). However, here one has to be careful, since then only the conduction band wave function should be used and normalized accordingly. If, on the other hand, we view the scattering potential as an electronic one and neglect the two-component nature of ΔV_j, we are left with

$$\langle U_{\alpha,\beta}U_{\alpha',\beta'}\rangle = \sum_j \frac{1}{A^2} f_j(\mathbf{q})\Delta E_c^2\Big(\psi_\alpha^{c*}\psi_{\alpha'}^{c*}\psi_\beta^c\psi_{\beta'}^c + \psi_\alpha^{v*}\psi_{\alpha'}^{v*}\psi_\beta^v\psi_{\beta'}^v\Big), \qquad (3.28)$$

where we can use the same form of the wavefunction and normalization condition as for the full result.

Note that $\rho_0 = em_{\text{eff}}/\hbar\pi$ in eq. (2.32) still contains the energy-independent effective mass m_{eff} in our two-band implementation. We therefore do not capture the additional broadening effect coming from transitions between bands with different parabolic dispersions in $\mathbf{k}$. This is a very small effect compared to the optical broadening.

3.3 Comparison of one- and two-band model simulations

In order to validate the implementation of a two-band model, we show results for one THz and one IR QCL using the previous implementation with only the conduction band [Paper III], and the new implementation with two bands [Paper I]. The first structure, taken from Ref. [37], is shown in Fig. 3.1 at a bias per period of 76 mV, together with the Wannier-Stark states with and without non-parabolicity. The states are very similar, as the bands are all relatively low in energy and the energy-dependent effective masses do not change much. The main difference can be seen for the states with higher energy, being shifted down in energy in the two band model. The lower states, which are the once active in the

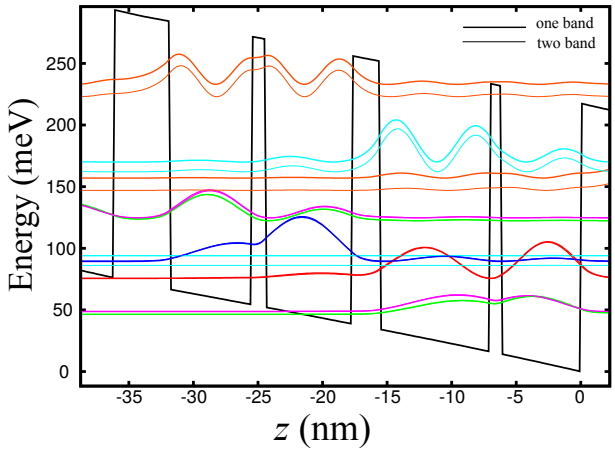


Figure 3.1: Comparison between the Wannier-Stark states for the one-band (thick lines) and two-band (thin lines) implementations, for the structure of Ref. [37]. In the new implementation only the states of high energy are different.

lasing scheme, are not affected to any noticable degree. Indeed, in simulations of the structure from Ref. [37], shown in Fig. 3.2, the difference is very small between the one and the two band model results.

The active states are close to the conduction band edge, well separated from the higher excited states, and so their effective masses do not change significantly with an energy dependent effective mass. Similar results were found for the structure from Paper II (the conduction band structure can be seen in Fig. 6.5), as displayed in Fig. 3.3, although here the shift in peak current is larger. Also, there is a red shift of the peak position of the gain, as well as a small increase. The data is taken at 50 K, but at higher temperature there is actually a rather large decrease in gain at the bias of peak current in the two-band model (see the discussion in Chap. 6). This reduction of gain, and also the shift of peak gain frequency, comes from the lower bias of peak current in the two-band model.

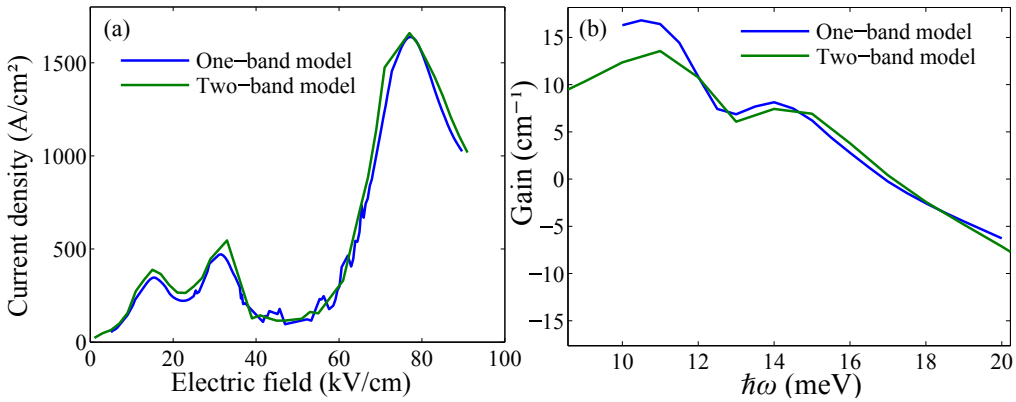


Figure 3.2: Comparison between IV (a) and gain (b) for the one-band model and the two-band model for the THz QCL called V843 [37]. The IV peak positions are unchanged, and the only difference is a slight increase in the overall current density. The gain peaks are shifted towards higher frequencies, but are close in magnitude, except for the peak at 10 meV.

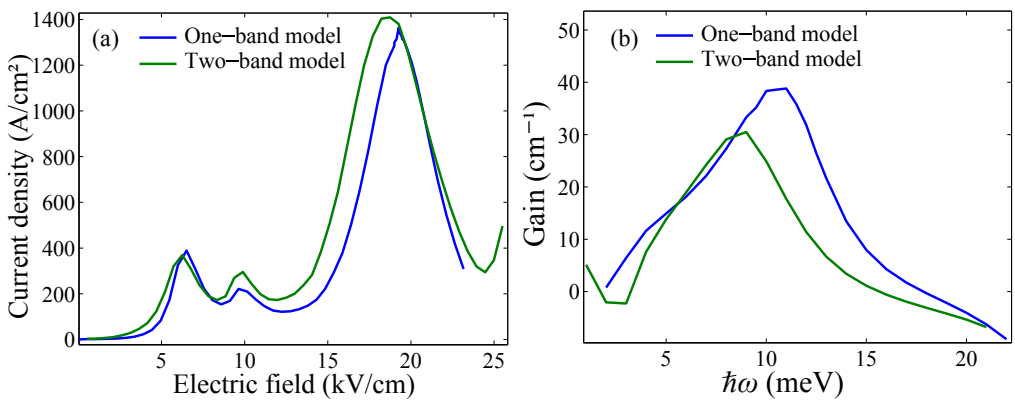


Figure 3.3: Comparison between IV (a) and gain (b) for the one-band model and the two-band model for the THz QCL called V845 [Paper II]. Here, the IV peak positions do shift to a lower bias for the two-band model, and as for Fig. 3.2, a slight increase in the overall current density can be seen. For the gain, which are calculated at the respective peaks in current density, the spectral gain peaks are red-shifted. This is explained by the lower energy difference between the ULS and LLS at the lower bias in the two-band case.

For the IR design [25] in Paper IV and V the results differ significantly between the models, as expected from the large variation in effective mass for high energy levels. The Wannier-Stark states are shown in Fig. 3.4, and there is a large energy difference between the states calculated in the one- and two-band models. As

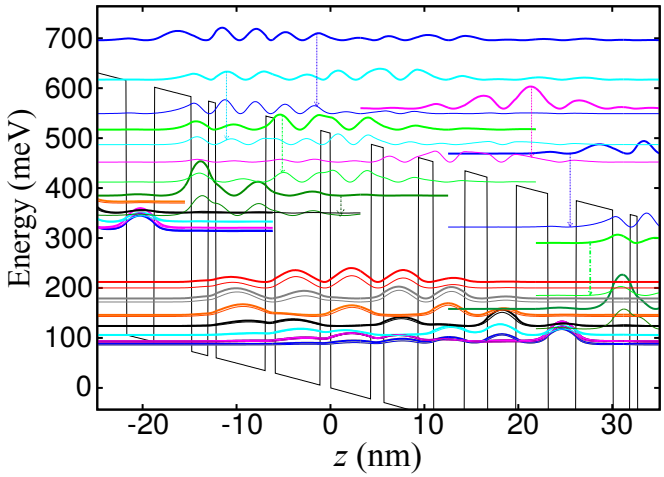


Figure 3.4: Square of the Wannier-Stark wavefunctions calculated in the one-band (thick lines) and two-band (thin lines) models. The dashed arrows indicate the lowering of the level energies for some of the states, when switching from the one- to the two-band model. Note especially the big change in both the energy and wavefunction for the upper laser state (dark green).

the energy of the upper laser level, shown by dark green in Fig. 3.4, is very different using one and two bands, the one-band model would obviously give the wrong emission frequency. Therefore, the current density and gain, shown in Fig. 3.5, are calculated using the energies of the two-band model Wannier states. In this approximation, non-parabolicity is only included as an energy dependent effective mass, while the basis states still have only conduction band components. Thus, the (conduction band) wavefunctions constituting the basis states are not orthogonal, which both leads to difficulties converging the calculations, and that fallacious results are produced. As can be seen in Fig. 3.5 (a), the current density is heavily suppressed in the two-band model as compared to the one-band model. The two-band model also gives results in excellent agreement with the experimental data below the experimental threshold current density of 1.5 kA/cm. Above threshold we expect the lasing experimental device to exhibit larger currents than the non-lasing simulated device, as seen in Fig. 2.3 (b) and discussed in Paper IV. The gain spectra are shown in Fig. 3.5 (b). As expected, the frequency of peak gain is very similar in the two models, and it is overestimated in the one-band model. In these simulations we used the

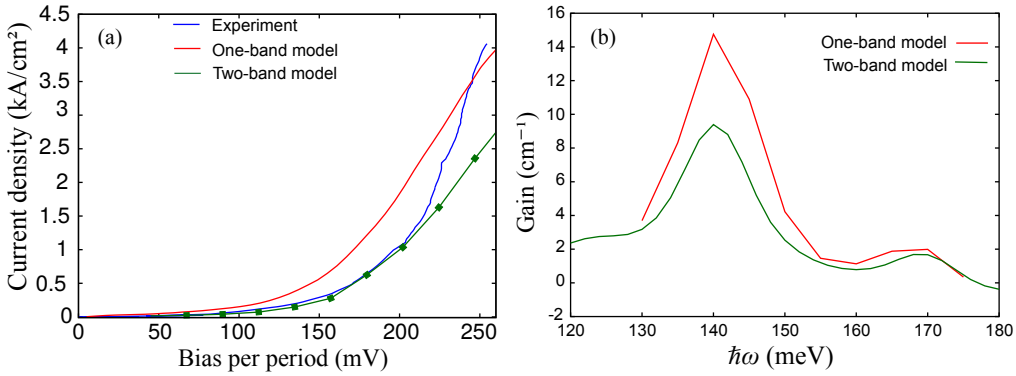


Figure 3.5: Current density (a) of the structure in Ref. [25] with the two- and one-band models compared to the experimental current density, and gain spectra (b) taken at a bias of 224.5 mV/period for the one- and two-band models, respectively. In the one-band model, non-parabolicity is included via an energy renormalization only, so that the energies of the states in the two models are the same but the basis in the one-band model is not orthogonal.

exponential roughness distribution of eq. (2.36) with $\Delta = 0.1$ nm and $\Lambda = 9$ nm. As will be discussed in Sec. 5.1, this value of Λ is probably too large.

Chapter 4

Parallelization

While the NEGF approach provides a general treatment of many-body effects and scattering on a solid theoretical ground, this advantage is also the main limitation of the model. The computations are heavy, both with regard to time and memory consumption, especially for simulations of mid-IR QCLs where a large state space and energy resolution is necessary. In order to be able to analyze, let alone design, such structures, we need something faster than the serial computations, which takes weeks to produce good results. In this section, the parallelization of the code in two steps is explained; first we parallelize the program to use all threads sharing memory on a single machine; secondly we adapt the code to run on a massively parallel scale on computer clusters. The former step is very useful for THz QCL simulations, although it is still too inefficient for simulating infrared devices. This is solved in the second step. In the following we show detailed performance tests of the code on different scales and simulated system sizes, but start with a description of the computational scheme of the NEGF model.

4.1 Computation scheme

The model calculates the Green's functions and self energies by iteratively solving the Dyson equation (2.26) and Keldysh relation (2.27) in discretized momentum and energy space, with N_k and N_E grid points, respectively. The number of levels involved are

$$N_{\text{mat}} = N_\nu(1 + 2N_{\text{per}})$$

where N_ν is the number of states considered per period and N_{per} is the number of periods in the calculation. N_{per} is usually one, but for short structures and

spatially extended states it may be necessary to use a higher number. Thus, in a naive implementation, the Green's functions and self energies would be $N_{\mathrm{mat}}^2 \times N_E \times N_k \times (2N_h + 1)$, where terms with $N_h = 1$ are needed to calculate gain and $N_h > 1$ is needed in order to capture non-linear phenomena and strong ac fields. Now, we use the severe simplification to assume that the self-energies are k-independent, and can be represented by some typical value for E_k. This assumption is eventually made for $X_{\alpha\alpha',\beta\beta'}(E_k, E_k')$ in eq. (2.32). This means we need not store all E_k-elements of $G(E, E_k)$, the computation of the X-tensors becomes very efficient, and the self-energy is simply given by the integral

$$\Sigma^{</R/A}(E)_{\alpha\alpha'} = \sum_{\beta,\beta'} X_{\alpha\alpha',\beta\beta'} \int \mathrm{d}E_k G^{</R/A}_{\beta\beta'}(E, E_k), \tag{4.1}$$

which can be discretized and parallelized easily.

A schematic illustration of the flow of the program is shown in Fig. 4.1, where squares represent code in the main program, ovals represent subroutines, blue are serial computations, and red represents parallelized parts. The program starts with making an initial guess to the self-energies, then $X_{\alpha\alpha',\beta\beta'}(E_k, E_k')$ are calculated once every F_{dc}. For each value of F_{ac} and ω, the convergence loop is iterated until convergence is reached, which typically takes 50-100 iterations. The computational time will therefore mostly be spent inside this loop, containing the loop over E_k from the integral in eq. (4.1) as well as the computation of updated self-energies from eqs. (2.26) and (2.27), as the most time-intense parts. As the program begins each new iteration with a starting point in the previous one, it is more favorable to parallelize the internal loops than only the outer loops for F_{dc}, F_{ac}, and ω (especially if one is interested in only one data point). All timings of the code given below are for a code compiled with the gcc (gfortran) compiler, without any optimization flags.

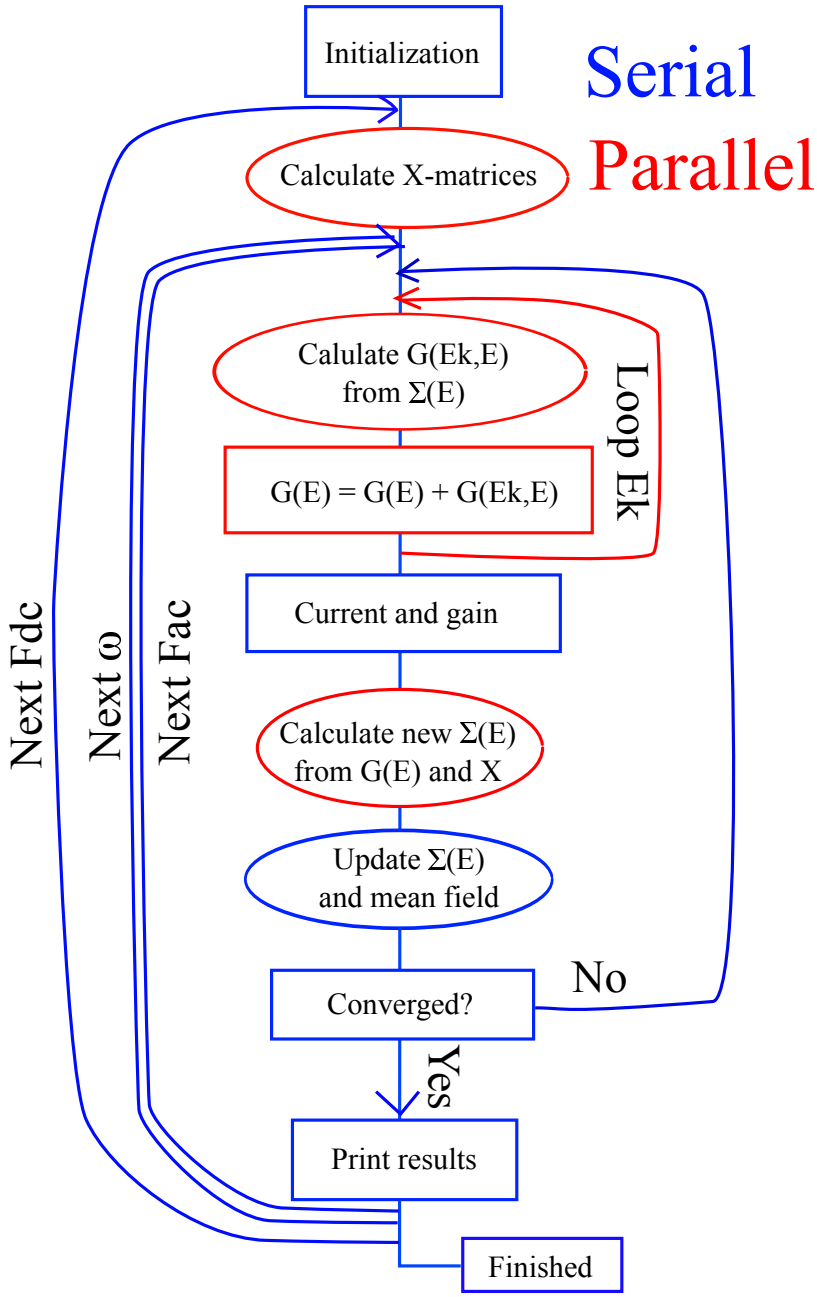


Figure 4.1: Flow chart for the NEGF program. Blue represents serial calculations, while red indicates parallel parts. The calculation of $\Sigma(E)$ is only parallelized in the MPI version of the code.

4.2 Single node computations

For computations on a single node, where all processors share the same memory, parallelization is best done using OpenMP [87] (or similar packages, such as POSIX Threads and Click). The wall clock times for the different parts of the code indicated in Fig. 4.1, as well as the average time for one iteration, are shown in Fig. 4.2. In this test, $N_E = N_k = 1000$ and $N_\nu = 15$, which is a quite large computational space, even for mid-IR lasers. The program is 6 times

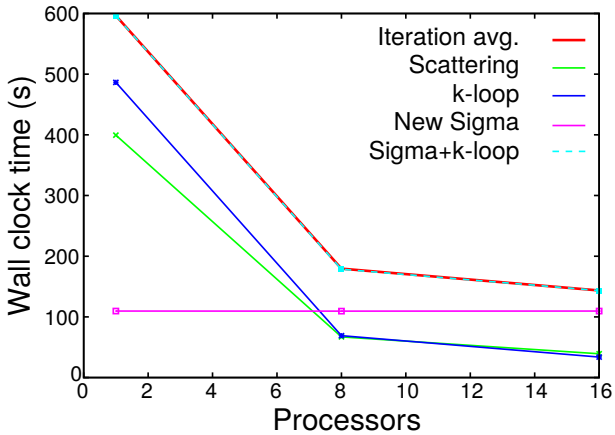


Figure 4.2: Timing of the different parts of the code indicated in Fig. 4.1, parallelized on a single machine using OpenMP [87]. In these simulations, $N_\nu = 15$ and $N_E = 1000$. From these times, we see that the loop over E_k together with the subroutine for evaluating the new self-energies Σ, makes up almost all of the iteration time. Here, the latter is not parallelized, as this is a major bottleneck when going to from 8 to 16 threads.

faster when going from a serial execution with one thread, to parallel execution with eight threads. Going to 16 threads gives a 10 times increase in speed of one iteration. While this is not bad for such a simple and straightforward implementation of parallelization, we see from Fig. 4.2 that the main bottleneck is the subroutine for calculating new self-energies. This was not parallelized since we typically use smaller N_ν for singe machine computations; normally $N_\nu = 15$ requires more threads than exist on a single machine in order to get results within reasonable time (see the next subsection). Another bottleneck is the calculation of the scattering matrix elements, whose speed actually is slower than the k-loop for 16 threads, even though it is faster for serial computations. The $X_{\alpha\alpha',\beta\beta'}$ matrix elements are calculated for all different scattering contributions, inside four nested loops running over the level indices α, α', β, and β'. Here, we simply parallelize the inner-most loop. The bottleneck is dealt with in the

next subsection, where all four nested loops are parallelized. However, since we are limited in the number of threads by the amount of cores available on a single machine, this would most likely give a small improvement in speed for shared memory parallelization. In addition, this calculation lies outside of the convergence loop as shown in Fig. 4.1, and so adds very little to the total computation time of one bias (F_{dc}), field strength (F_{ac}), or frequency (ω) point, so it's speed is not so crucial. It is, though, very convenient to have it parallelized, since this allows the user to more quickly look at the convergence rate and diagnose the numerical parameters chosen.

Finally, we note that the summed times for the k-loop and the calculation of $X_{\alpha\alpha',\beta\beta'}$ in Fig. 4.2, makes up almost all of the computation time per iteration. Thus, we have focused on the most important part for speed improvement for medium system sizes ($N_\nu \lesssim 9$), which is the case for most THz QCLs. However, most mid-IR QCLs have more than 10 states in the active region, and this requires a more elaborate parallelization scheme, as discussed in the following subsection.

4.3 Cluster computations

A single high-performance machine has typically no more than 48 processors, but more common computers usually have 8-16. This limits the speed at which we are able to do computations for structures with more states and finer energy grid, as well as high intensity calculations with $N_h > 1$. The next step in computational power is to migrate to a computer cluster, which consists of many connected machines. This extends the number of possible parallel processes to hundreds [88] or even more. Each process has it's own memory, and to share objects between processes, messages with data are sent between them. This is done through an interface called MPI (Message Passing Interface) [89], which takes much more effort and care than shared memory parallelization for the programmer. Average computation times are shown in Fig. 4.3 (a). Here we have timed the total iteration loop, which is the innermost blue loop in Fig. 4.1, as well as the parallelized section of the code marked as red in Fig. 4.1. First we note that, comparing the sum of times of the parallelized parts within one convergence iteration, most of the computational time is spent in parallel code. The longest time is spent in the loop over E_k (blue curve). Next, we see in Fig. 4.3 (a) that the subroutine for calculating new Σ (purple curve) actually becomes slower for large numbers of processors, completely taking over the time per iteration for 128 cores. To fix this issue, a number of processors are excluded from this calculation, depending on the simple criteria that the number of energy

points calculated per processor

$$N_{E,\text{local}} < C/N_\nu^2, \tag{4.2}$$

where $N_{E,\text{local}} = N_E/(\#\text{ processors})$ is the number of energy points calculated by each processor, and C is some constant that is chosen to give the best performance. The computational times shown by the dashed light blue curve in Fig. 4.3 (a) uses $C = 5000$. Then, the time taken for 64 and 128 processors are both 4.5 s (compared to 5.1 s for 32 processors). Note also that due to the MPI overhead, the time of 670 s with only one processor is higher than the time of 610 s for the serial version of the code.

In order to see how effective the parallelization is for other values of N_E and N_ν, a timing of the code for different system sizes is shown in Fig. 4.3. Here, we have used $N_\nu = 3$, 9, and 15 (red, green, and blue lines, respectively), and $N_E = N_k = 500$ and 1000 (solid and dashed lines, respectively). The thin lines show the expected time if there were no overhead time for sending and waiting to receive messages by the processors. For the smallest computational space (lowest red curve), we see that the time actually increases for a large number of processors, due to the very large overhead compared to the actual computation time. We also see that the relative overhead time (ratio of thick to thin lines) decreases as the system size increases while it increases with the number of processors. While for $N_E = 500$ the gain in efficiency in going from 64 to 128 cores is questionable, this still yields a significant increase of the computation speed for $N_\nu > 3$ and $N_E = 1000$.

As the number of processes and the overhead time is larger for MPI, it becomes more important to have an efficient parallelization. For this reason, the four loops over level index in the calculation of the scattering matrix elements $X_{\alpha\alpha',\beta\beta'}$ is un-rolled into a single loop, so that computation chunks of suitable sizes can always be distributed among the processors. This is evident from that the MPI calculation with 16 processors is faster (22 s) than using OpenMP (39 s), even though the overhead time is longer in the former case.

In conclusion, an efficient parallelization of the code running both on single machines with shared memory, and large computer clusters (the code has also been run successfully on the Brutus cluster of ETH Zürich) has been implemented, enabling computations of infrared devices. Specifically, for a large computational space with $N_E = N_k = 1000$ and $N_\nu = 15$, we can expect a 100-fold increase in the speed for one convergence iteration. As seen in Fig. 4.2, the OpenMP implementation could be made more efficient for large systems by parallelizing also the calculations of new self-energies.

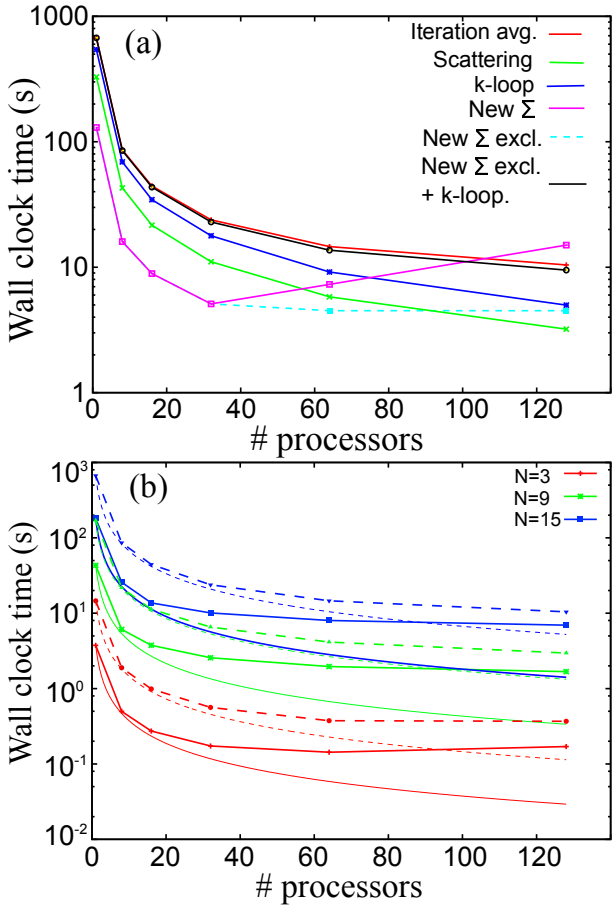


Figure 4.3: (a): Average wall clock time for one full iteration (red), the subroutine for calculating $X_{\alpha\alpha',\beta\beta'}$ (green), the parallelized loop over momentum k (blue), the subroutine for calculating new self-energies using all (purple) or a reduced number of processors (light blue, dashed). We also show the sum of the times for the k-loop and the calculation of new self-energies (black), which are the parallelized parts of the code within one convergence iteration. These times are for a calculation with $N_\nu = 15$ and $N_E = N_k = 1000$, which is typical for mid-IR lasers. (b): Average wall clock time for one full iteration for different N_ν and for $N_E = N_k = 500$ (solid lines) and $N_E = N_k = 1000$ (dashed lines), vs. number of processors. Thin lines show the expected times if no overhead time is added when increasing the number of processors.

For the MPI implementation, as seen in Fig. 4.3 (a) the overhead time for large numbers of processors should be reduced, and the self-energy calculation could be more efficiently parallelized, as shown in Fig. 4.3 (b). The former could be achieved by making use of non-blocking send and receive calls, whereas the current version only implements blocking calls (where the sending and receiving processor both wait until the message has been successfully received). Note also that none of these tests made use of compilation optimization.

Chapter 5

Mid-infrared QCLs

As described in the previous sections, we have built upon the previous model suitable for THz QCL modeling, and included the effects necessary for accurately model mid-infrared devices. We will now apply this model to study a number of such structures.

First, we will investigate the effects of scattering by interface roughness. As mid-IR QCLs have larger band offsets than their THz cousins, the interface roughness scattering will play a much more important role. We will then consolidate our model extensions with experimental measurements and simulations by another model that is well tested for mid-IR devices, for a lattice-matched InGaAs/InAlAs structure. A multi-stack device, where multiple active regions emitting at different wavelengths are stacked together to get a wide laser spectrum, is simulated.

Having consolidated the model for lattice matched structures to experiments and other models, we push it further by simulating a strain-compensated device, which have very high CBO for specific barriers. While a simpler density matrix model predicted the experimental behavior quite well, considering different experimental realizations of the device showed large variations in current density and lasing, we originally found very large interface roughness scattering which degraded the performance of the QCL well below the experimental observed behavior. Including interface correlations, our results agree very well with the experimental data. Finally, we briefly present an optimization of the active region using a mono-layer based design.

5.1 Role of interface roughness scattering

While interface roughness (IFR) scattering is important as an additional source of elastic scattering and provides incoherent tunneling in lower CBO (THz) structures, for mid-IR devices it is more important to be accurate in the modeling of this effect, as it is much stronger [77]. As the actual roughness distribution in real devices is not well known, and many models for interface roughness are used, we here compare different modeling functions and investigate for which parametrization it is possible to get reasonable results.

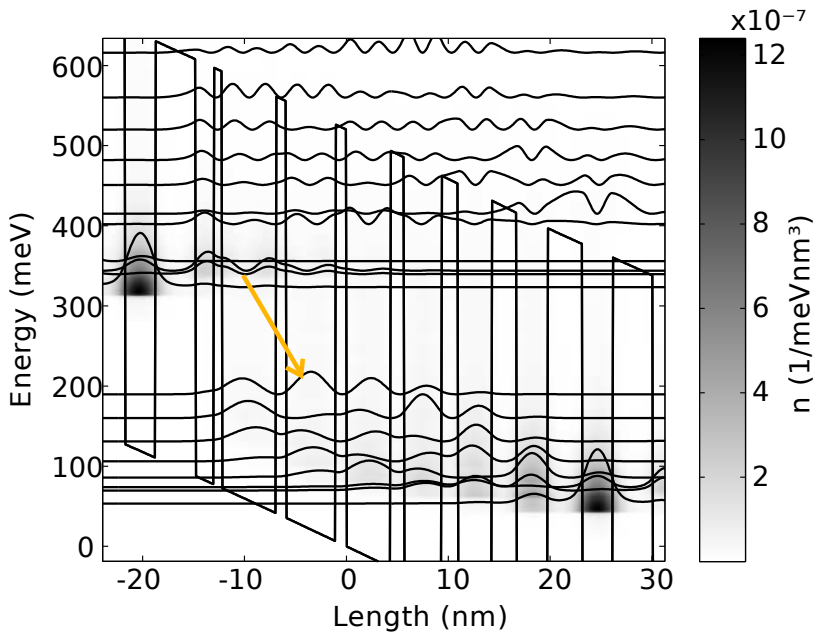


Figure 5.1: Band conduction band structure of the mid-IR QCL of Ref. [25].

The main results were published in Paper V, where we investigate the impact of the exact choice of the statistical distribution of the interface roughness correlation function. We choose either the Gaussian distribution from eq. (2.35), or the exponential distribution from eq. (2.36). The two distribution functions are plotted in Fig. 5.2 for different sets of values for the roughness height Δ and correlation length Λ. These sets of values are chosen systematically in order to match the Gaussian and exponential distributions at low (the set called Expon. 1 and Gauss 1 respectively) or high (Expon. 2 and Gauss 2) momentum transfer, according to eqs. (2.37) and (2.38). The Expon. 1 and 2 values in panel (a) refers to a mid-IR structure [25], while the Gauss 1 and 2 values are taken for a THz structure [37]. The relevant momentum transfer for transport through

the mid-IR structure [25] under study in Paper V is the one corresponding to a transition from the upper to the lower laser level, with $q \approx 0.4$ nm^{-1}. We thus expect the transformation from a Gaussian to an exponential distribution via the Expon. 2 set of parameters to provide similar results. Indeed, looking at the results in Fig. 5 of Paper V, this is exactly what we find.

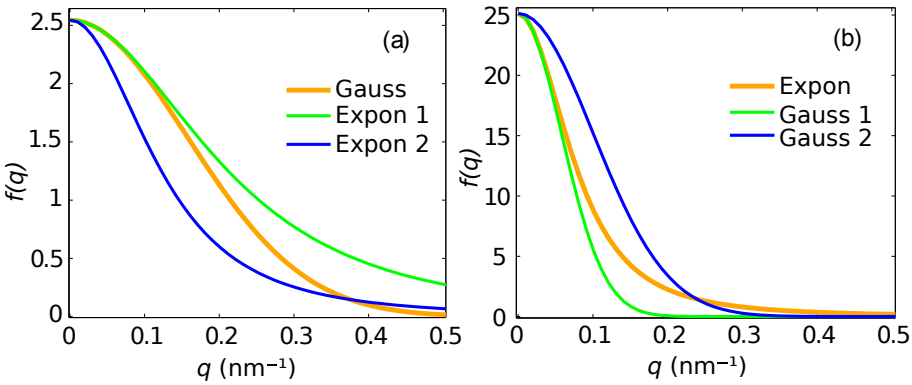


Figure 5.2: Fourier transforms of the auto-correlation function in eq. (2.35) (Gauss) or eq. (2.36) (Expon). In (a), we start from a Gaussian distribution with $\Delta = 0.1$ nm and $\Lambda = 9$ nm adopted for the IR QCL in Ref. [25], and transform to exponential functions by translation (2.37) (Expon 1) and (2.38) (Expon 2). In (b), we start instead from an exponential distribution with $\Delta = 0.2$ nm and $\Lambda = 10$ nm, suitable for the THz QCL in Ref. [37], and apply the reverse transformations of (2.37) (Gauss 1) and (2.38) (Gauss 2). Figure is taken from paper V.

Although this chapter is concerned with mid-IR QCLs, it is interesting here to look at the impact of the choice of interface distribution functions also for THz QCls. For the SA THz structure of Ref. [37], the momentum transfer corresponding to the gain transition is 0.16 nm^{-1}. Here, it is not clear from panel (b) which of the two transformations will provide the most similar results to the original exponential distribution. We find that it is again the second transformation, here Gauss 2, which agrees best with the exponential one, especially for the current density, which can be understood by the large separation of the lower laser state and the extraction state, as well as the injector state and the upper laser state in this design. In contrast, the gain spectral width is more similar when using the Gauss 1 transformation, since the level broadening is mostly affected by scattering with low q. In Paper V, we also study a second THz structure [90], with tunneling injection. As the number of interfaces per period is less and the CBO is lower, the overall impact of interface roughness scattering is smaller in this design. Still, we find once again that the Gauss 2

transformation agrees best with the original exponential distribution, which also fits best with the experimental findings of Ref.[90].

As already discussed in Sec. 3.2, the scattering matrix element for interface roughness contains conduction band as well as valence band components in the two-band model. The effect of making this substitution is very small, especially for the THz QCL, as shown in Paper V.

5.2 Multi-stack device

In order to investigate the possibilities of designing a QCL structure with a very broad and flat gain spectrum, a structure derived from the one presented in Ref. [13] was simulated within the MIRIFISENS project [91]. The structure is composed of five different active regions, each one optimized to lase at specific wavelengths of 7.3, 8.5, 9.4, 10.4, and 12 μm respectively. These active regions will be grown in sequence (symmetrically around the shortest wavelength to get optimal mode propagation), and so the current density through each region will be the same under operation, which puts constraints on the operation bias; it will be limited by the region which reaches $J_{\max}$ at the lowest bias.

The simulated current density vs. bias is shown in Fig. 5.3. The interface roughness parameters in these simulations are those of the Expon. 1 (2.37) transformations in the previous section (from a Gaussian model with $\Delta = 0.1$ nm and $\Lambda =$9 nm, giving $\Delta = 0.17$ nm and $\Lambda = 3.67$ nm) and as we have already seen in Secs. 5.1 and 2.7.2, we therefore have to be careful in the analysis since the laser transition energies are centered around 150 meV. In fact, simulations of the 12 μm design with parameters we conventionally use ($\Lambda = 10$ nm in the exponential model), show currents are slightly higher below the onset of NDR, and a lower peak current density.

The band structures of the individual stacks are all variations of a bound-to-continuum design similar to Ref. [25], which is shown in Fig. 5.1. Here, however, we see that the ground state of the well before the injection barrier is lower in energy compared to Ref. [25].

Gain was simulated at current densities of 3 and 4 kA/cm^2 for each structure, as shown in Fig. 5.4. At 3 kA/cm^2, the total gain profile is below zero for all frequencies. As the gain is higher for higher electric fields, operation at a current density of 4 kA/cm^2 seems to be a better option. At this larger current density the gain is improved, but still does not show the desired broad features. Especially the 7.3 μm structure inhibits large absorption below 130

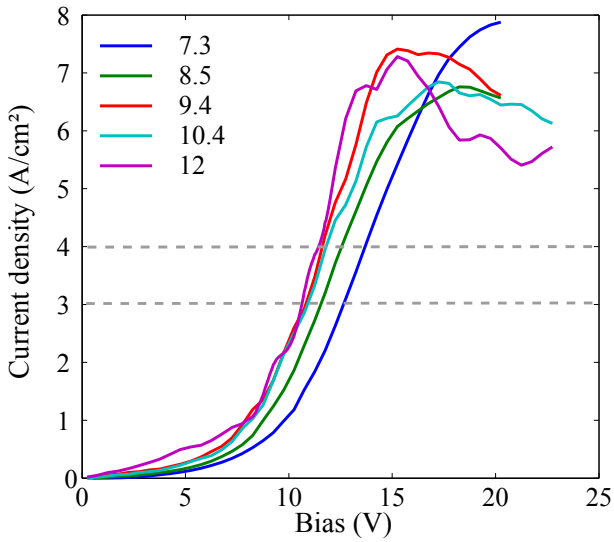


Figure 5.3: Current vs. the total bias over 50 periods for each of the five simulated active regions. The dashed lines indicate at which biases the current density is 3 and 4 kA/cm^2 through all structures.

meV, removing the 12 μm peak in the final spectrum. In Fig. 5.5 (a) we see that for $\hbar\omega = 100$ meV (red), there is absorption to higher states where the density of electrons is the highest before and after the injection barrier. There is gain in the band of lower laser states, but it is too small to compensate for the large absorption in the ground state. For $\hbar\omega = 160$ meV (green curve in Fig. 5.5 (a)) the gain transition is active and the absorption in the ground state is smaller. In order to improve gain at low frequencies, the overlap of the ground state with higher states should be reduced. The 12 μm structure, on the other hand, is absorbing at high frequencies from the lower band of states up to the upper laser state (as well as to higher lying states) as seen in Fig. 5.5 (b). In fact, the 7.3 μm structure is absorbing at frequencies where the 12 μm structure has gain, and vice versa. Omitting the 7.3 and 12 μm structures, the gain actually gets broader in these simulations, as well as larger for all frequencies, as can be seen for the yellow dashed line in Fig. 5.4.

Changing the roughness transformation from the Expon. 1 to the Expon. 2 transformation, or using a Gaussian distribution function, would most likely improve gain, as well as change the bias (with 4 kA/cm^2 current) where the gain should be calculated. As the exact distribution is not known, however, there is no direct way of knowing which results apply best to the experimental situation.

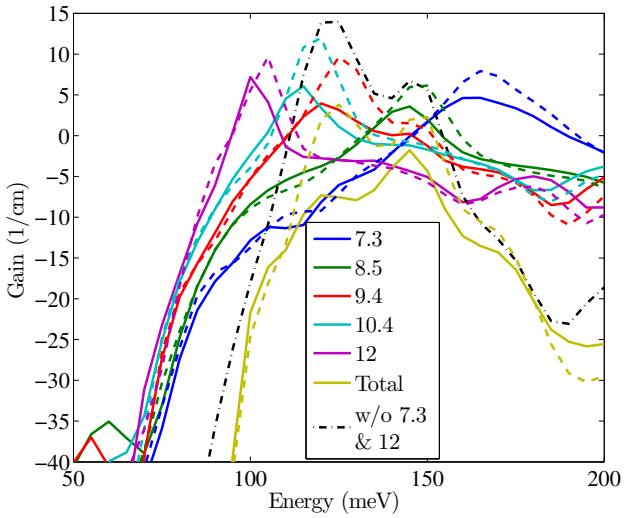


Figure 5.4: Simulated gain at biases corresponding to 3 (solid) and 4 (dashed) kA/cm^2 for each of the five active regions. The dot-dashed black line shows the total gain omitting the 7.3 and 12 μm structures.

5.3 Strain compensated InGaAs/InAlAs/AlAs structures

Within the MIRIFISENS project [17], one goal is to produce a design with a short wavelength of ~ 3.3 μm much shorter than for the usual designs with $\lambda \sim 4-8$ μm. This short wavelength and high photon energy (370 meV) requires a high conduction band offset. Using material compositions with very different conduction band energies requires a relaxation of the lattice-matching condition for growing InP-based InGaAs/InAlAs structures. However, this introduces strain in the system [27]. This strain is compensated by using the combination of the well material $Ga_xIn_{(1-x)}As$ with $x = 0.28$, and the barrier materials AlAs and $Al_xIn_{(1-x)}As$ (with $x = 0.52$). As the number of states is very large, it is challenging to simulate such a structure with the NEGF model. Thus we have done simulations both with the NEGF and a density matrix model developed at ETH Zürich, called SEWLAB [60, 92].

In Fig. 5.6, we show simulation results for EV1695 using nominal data (full lines). The Density matrix simulations, using SEWLAB, produced fairly accurate results, which agreed well with the experimental data below the threshold current of 2 kA/cm^2. However, the NEGF simulations appear to agree worse with experiment. In particular the spectral peak gain position in NEGF is about 50 meV below the experimentally observed lasing frequency, as can be seen in

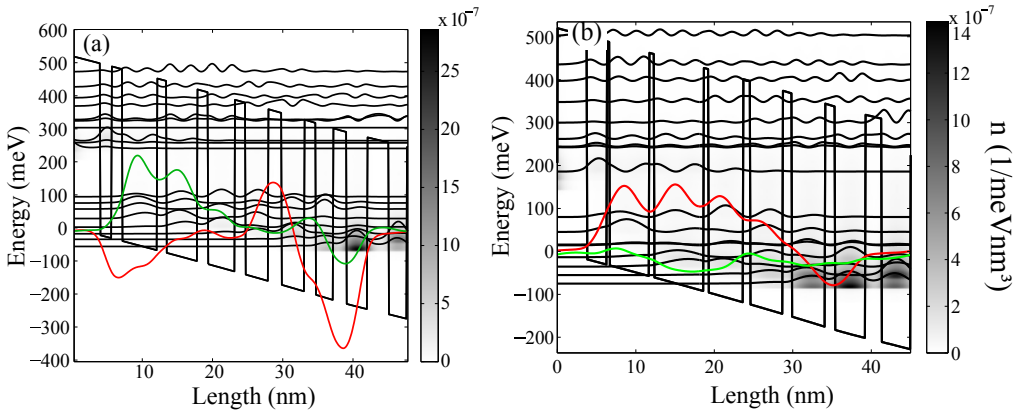


Figure 5.5: Carrier density and spatially resolved gain (a. u.) at a current density of 4 kA/cm^2, at (a) $\hbar\omega = 100$ meV (red) and $\hbar\omega = 160$ meV (green) of the 7.3 μm structure, and at (b) $\hbar\omega = 105$ meV (red) and $\hbar\omega = 140$ meV (green) of the 12 μm structure.

Fig. 5.7. This can be related to the treatment of interface roughness, where the models differ in two respects: Firstly, SEWLAB takes into account vertical interface roughness correlations, which effectivly reduces the scattering. This can be seen in the broadening induced by elastic scattering of 170 meV, in the NEGF model, which is much larger than the width of 50 meV that SEWLAB typically finds. Secondly, the NEGF model includes a higher order energy shift of the energy levels, due to the treatment of scattering via self-energies, which lowers the energy of the upper laser level. We find this self-energy shift is about 50 meV for the upper laser state.

Taking interface correlations into account in a simplified way in the NEGF model, we find a self-energy shift of 30 meV, and elastic scattering rate of 88 meV, which is a dramatic reduction. The reason is that the high AlAs barriers are very narrow, and thus the impact of the correlation (here modeled by a reduction in the interface roughness RMS height $\Delta' = \sqrt{1-c} * \Delta$, where following Ref. [60], $c = \exp(-\Delta z/k)$ and $k = 1.5$ nm,) on the scattering rate is large. In Fig. 5.6 this appears to shift the whole NEGF curve down in current density, agreeing better with the experimental measurements. Also the gain peak shifts to higher frequencies, much closer to the measured laser emission frequency.

A third factor that contributes to the differences between NEGF and SEWLAB results is the treatment of the momentum resolution of scattering processes, where we use a constant value for the scattering rate for all momentum exchanges in the NEGF model. This leads to an over estimation of the scattering strength.

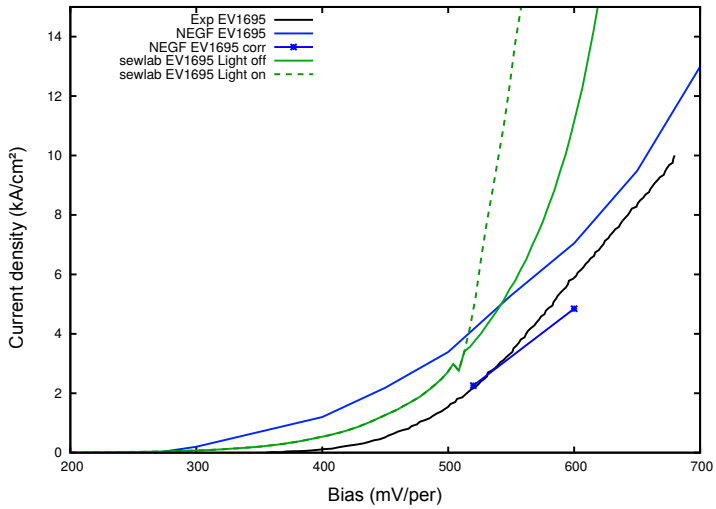


Figure 5.6: Comparison of non-lasing NEGF and SEWLAB simulation results, as well as lasing SEWLAB results, for the original design EV1695. Also shown is the NEGF simulation accounting for interface correlation effects. Experimental data for the device is shown in black.

We conclude that interface roughness scattering plays a crucial role for the performance of this structure, and it is important to model it correctly. As there is expected to be a large uncertainty regarding the effectiveness of vertical interface correlation effects, minimizing the impact of interface roughness scattering may result in a large improvement of the performance. This is important, especially as further active region optimization is not expected to yield significant improvements on its own. Finally, these calculations indicate that our current implementation of the NEGF model has problems with extremely large interface roughness scattering, occurring in these devices with exceptionally high barriers when neglecting correlations.

5.4 Optimization of short wavelength and low dissipation active region

Simulations of the design (EV1695) in the previous section showed that interface correlations are crucial for the operation. However, there is a large uncertainty as to how large this correlation is in a real system, and therefore we think that lowering the interface roughness is a good way to increase the wall plug efficiency in three ways; (i) increased gain due to increased inversion and lower level broadening; (ii) lower current density for a given bias; (iii) decreased absorption due

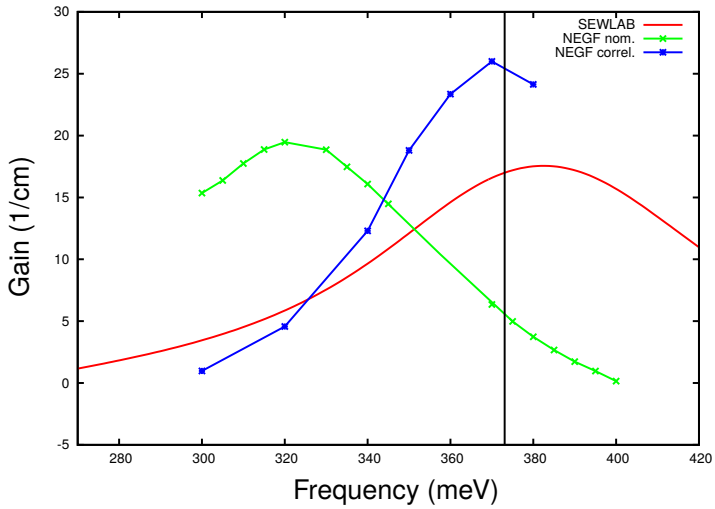


Figure 5.7: Gain calculated for a bias of 600 mV/period using SEWLAB and NEGF. For the NEGF model, data is shown for the cases with (blue) and without (green) vertical interface roughness correlation. The vertical line shows the observed laser frequency of 375 meV (wavelength of 3.3 micron).

to lower width of absorbing levels. Employing lower barriers is dangerous, since the high energy transition requires a large conduction band offset. Instead, if the layers are grown in integer mono-layers, so that each layer becomes completely filled with one material before the growth of the next layer, interface roughness should be minimized. In addition, the variation of period lengths in the sample is expected to be reduced with this approach, providing a narrower gain spectrum with higher output power.

In order to see whether the new layer sequence per se improved the performance, i.e. without assuming a reduction in interface roughness, we use the same roughness parameters as for the original design. The simulated IV is shown in Fig. 5.8. Up to threshold, the two simulation models, NEGF and SEWLAB, provide very similar results, despite the mentioned difference in the treatment of interface roughness scattering. We also find a lower maximum interface roughness scattering matrix element (about 110 meV) in the NEGF model for the same parameters, compared to the original design. We attribute part of this reduction to the fact that the new design has two interfaces less. This indicates that this design is less sensitive to interface roughness.

We find significant gain at 360 meV (3.4 micron) in SEWLAB, and due to the width of the gain spectrum (about 60 meV), the target frequency of 375 meV is still covered. While the wall plug efficiency shown below is calculated at 3.4

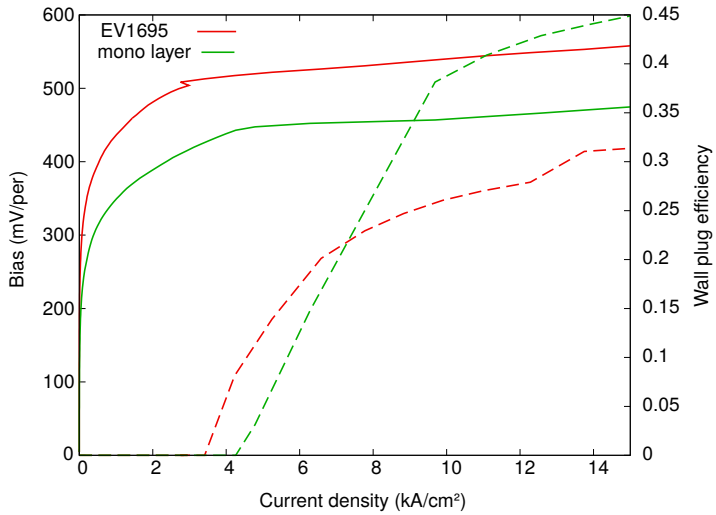


Figure 5.8: SEWLAB simulations of current density vs. bias (solid) and wall plug efficiency (dashed) for the original (EV1695) and new mono layer design.

micron, it will be slightly smaller at 3.3 micron. Again we find a much lower spectral peak gain frequency in the NEGF model (about 300 meV).

In these simulations, the lattice temperature is 300 K, while the electron temperature is set to 360 K, which agrees well with the one found in full NEGF simulations, as well as SEWLAB simulations with kinetic balance.

While the threshold current density is slightly higher for the new design, as can be seen in Fig. 5.8, the corresponding bias is lower, resulting in a similar (but lower) wall plug efficiency as for EV1695 at low current density. However, the wall plug efficiency for a current density of 10 kA/cm^2 is significantly improved from 0.25 to 0.40. Note that these results are found using the same interface roughness as in the case of the non-mono layer sequence, which we believe is a pesimistic assumption.

The wall plug efficiency is much better for the new design at high current densities, which are problematic to reach in a real device due to heat dissipation. In order to reduce the current density, we have simulated the same sample but with half the doping density. The simulated wall plug efficiency as a function of optical output power is shown in Fig. 5.9, and here we see a clear advantage of the new design compared to the original, especially at low output power, where the low doped version is very efficient and has a comparably low heat dissipation.

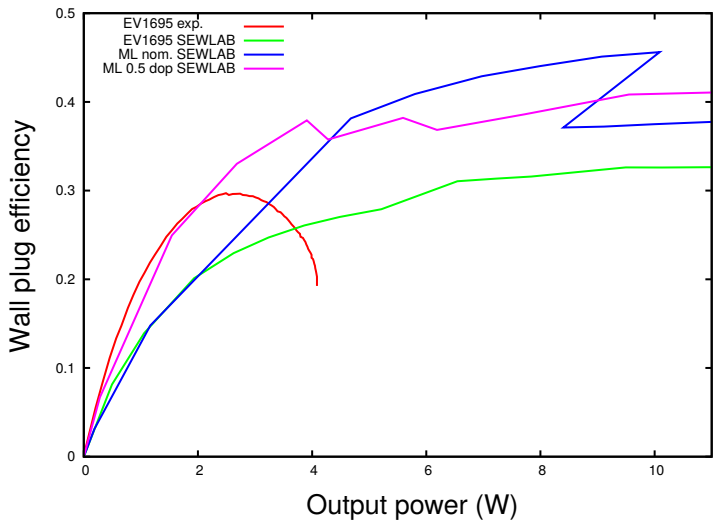


Figure 5.9: Simulated wall plug efficiency vs. optical output power for the original device with nominal doping density (blue) and the new mono layer (ML) design with nominal (red) and half the nominal (green) doping density, as well as measured values (red) for the original design EV1695.

Chapter 6

Terahertz QCLs

Up to this point, we have generalized and expanded the model in order to reproduce and predict mid-IR QCL behavior, and we have shown the model is accurate for a variety of designs in Chap. 5. In this section, the aim is to show that the model still is able to produce reliable results in agreement with experimental data. We will look at a few examples of THz QCL, and see some of the conclusion that can be drawn from such simulations. Here, the emphasis will be on the design schemes, and some specific suggestions for improvements are given. In the last part of this section, new designs will be presented, which out-perform any of the THz QCLs previously simulated with our model.

The main obstacle for the terahertz quantum cascade laser [18] (THz QCL) reaching it's true potential in industrial and scientific applications is the limited operating temperature. To date, the maximum operating temperature is 199.5 K [20], while many practical applications require a minimum temperature of 240 K, where electro-mechanical coolers can be employed. Which effect is mostly responsible for the temperature degradation is not completely settled [93, 94, 95, 96, 40, 41]. It has been speculated that thermal backfilling [3, 97][Paper I] is one major contributor, while thermally activated longitudinal optical (LO) phonon emission, i. e. the emission of phonons by electrons of high in-plane momentum in the upper laser state, is generally accepted as the main cause [96, 95, 40, 98]. In addition, carrier leakage into continuum states of the Γ-valley (for low Al content $x = 0.15$) [41], as well as bound states in the X-valley (for high Al content $x = 0.33$) [99], has been shown to have detrimental effects on the operation on THz QCLs employing the GaAs/$Al_xGa_{1-x}As$ material system. An increased leakage with lower barriers is also indicated by e. g. Ref. [97]. (For even higher Al concentration, $x = 0.45$, Ref. [99] finds that leakage currents are

suppressed by a stronger localization of the bound states.)

6.1 Injection schemes in THz QCLs

The most promising THz QCL design has been to use a tunneling injection with resonant phonon extraction [100], most successfully using a diagonal laser transition [90]. The tunneling injection serves two purposes; it separates the periods in space so that a LO phonon scattering rate between the LLS and the ULS of different periods is smaller, and it prevents negative differential resistance (NDR) at biases lower than that designed for laser operation, by limiting the current through each period when the tunneling resonance is not satisfied. This method has been very successful so far, having reached the highest operating temperature repeatedly [90, 20]. However, this design has inherent drawbacks as we will discuss below. One attempt to improve on these, is the scattering assisted (SA) scheme, which is also discussed with regards to two different SA structures, one an attempted improvement of the other. A more detailed analysis of the structures in this section can be found in Papers I and II.

6.1.1 Tunneling injection

In order to study the tunneling injection scheme, consider the tunneling injection THz QCL of Ref. [90], whose conduction band structure is shown in Fig. 6.1 at a bias of 57 mV/period. In addition to the upper laser state (ULS), lower laser state (LLS) and injector (in) states, this structure features an additional extraction state to aid the extraction process by having a large overlap with the injection state of the next period. The injector and ULS are almost degenerate at the design bias, and share most of the carrier density, which is divided with about half of the population on each side of the injection barrier, as seen in Fig. 6.2. This is due to the nature of the tunneling injection scheme and may give significant thermal backfilling into the LLS. The LLS, on the other hand, is almost empty in comparison, which means this design is close to optimal. Furthermore, the inversion at 50 K is 45% of the total population density, which is very close to the theoretical limit described earlier. Fig. 6.1 (b) shows the band structure and energy resolved current density at a slightly higher bias, and indicates that there is significant leakage into the paracitic current state (PCS), even thought the population of this state is negligible. This leakage would increase the current density, dissipating more power than necessary with resulting reduction of efficiency and lattice heating. We also see bands of current between the ULS and the extractor state at one optical phonon energy below

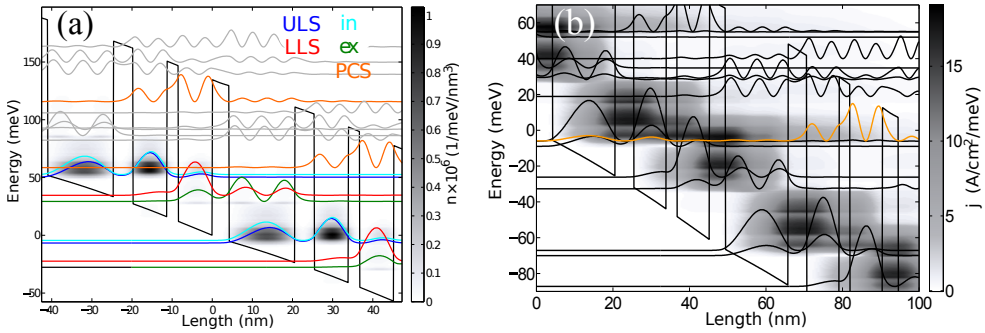


Figure 6.1: Tunneling injection design of Ref. [90] with the upper laser state (ULS), lower laser state (LLS), injector (in) and extractor (ex) states and a parasitic current state (PCS). The electron densities at a bias of 57 mV per period is shown in gray scale in (a). (b) Energy resolved local current density at a bias of 61 mV/period, where the PCS is plotted in orange. The lattice temperature in these simulations is 50 K. This figure is taken from Paper I.

the LLS, caused by LO phonon emission from the bottom of the LLS to a state of high momentum in the injector.

For this structure, gain drops by 70% when increasing the temperature from 50 to 200 K lattice temperature [Paper I]. In order to explain this, Fig. 6.2 shows the change in carrier densities of the different states labeled in Fig. 6.1 with temperature. Inversion (considered as the average population in the injector and ULS, as these states are in resonance) decreases by $\sim 30\%$ over this temperature range, while the width of the transition increases by $\sim 120\%$. From eq. (2.49), this gives an expected reduction of gain by $\sim 73\%$, which explains the loss seen in the simulations. Thus the larger fraction of the loss of gain comes from a radical increase in the transition width, caused by above all optical phonon scattering. This confirms the idea that the main temperature degrading mechanism in structures with tunneling injection is thermal phonon emission. It is thus important to find the correct choice of wavefunction overlap, in order to minimize this rate, but at the same time have an efficient stimulated emission.

Another detrimental temperature effect is the reduction of inversion by thermal backfilling from the ULS to the LLS and extractor. In Fig. 6.2, the dashed lines show the expected amount of carriers removed from the ULS by thermal excitations (if the populations follow a Fermi distribution). This can be seen to account for about half of the increase in the LLS population.

In conclusion, this design is already highly optimized for gain, since the inversion is close to the theoretical limit at low temperatures. However, the high temper-

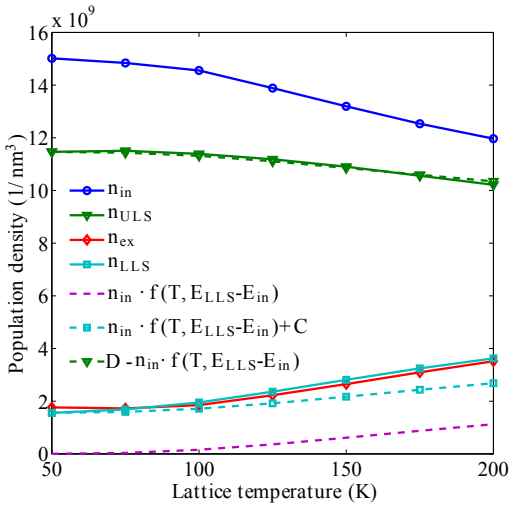


Figure 6.2: Temperature dependence of the densities in each state shown together with an estimate of the thermal backfilling from the injector and extractor states to the LLS as given by $n_{\mathrm{in}} \cdot f(T, E_{\mathrm{LLS}} - E_{\mathrm{in}})$. The estimate is shifted by a constant C in order to coincide the curves at 200 K for easier comparison. The dashed, diamond-marked curve shows the reduction of n_{ULS} due to the backfilling of n_{LLS}, where D is again a constant used for comparing the curves. The data is taken at 57 mV/period. These figures are taken from Paper I.

ature performance could be optimized by lowering the dipole matrix element between the ULS and LLS (to reduce phonon emission ULS→LLS). It is less clear how to reduce thermal backfilling, however this is likely of less importance [96, 95, 40, 98].

6.1.2 Scattering assisted injection

The scattering assisted injection scheme [34] uses an additional optical phonon transition in the current path compared to the tunneling injection scheme, in order to inject carriers into the upper laser level. This allows for a larger bias drop over one period which is supposed to reduce thermal excitations to the lower laser level. However, one would expect there to be twice the occupation of phonon modes in the lattice, providing more lattice heating than for phonon extraction only.

In Ref. [37] we investigated a SA structure labeled V843, shown in Fig. 6.3. The structure was designed using a rate equation model, to optimize the level energies and wavefunction overlaps to match the suggested SA scheme as follows: The

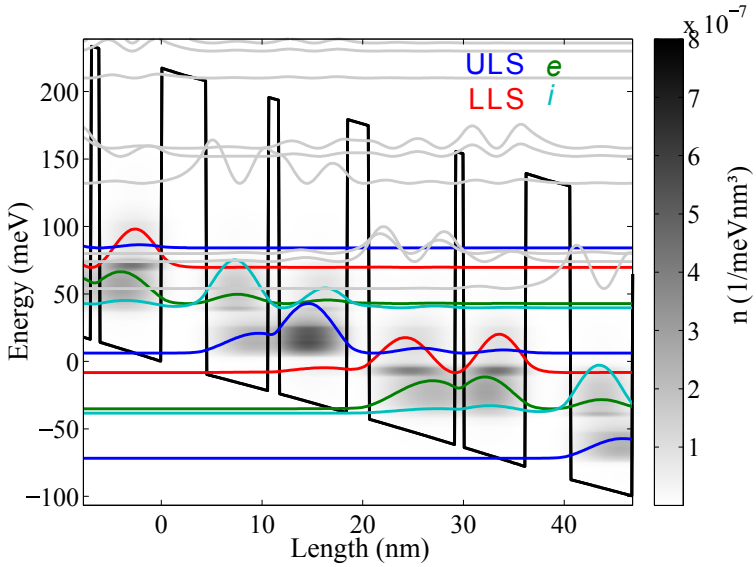


Figure 6.3: Band structure of the QCL in Ref. [37] as well as the carrier density and the Wannier-Stark states, at a bias of 78 mV/period for $T = 125$ K.

injector state (i) has to be one phonon energy $E_{\mathrm{LO}} = 36$ meV above the ULS and below the LLS of the previous period, as shown in Fig. 6.3. This actually means that, theoretically, almost all carriers can be in the ULS, in contrast to the tunneling injection scheme. At the same time, the extraction state (e) needs to be aligned to i. This can only happen at a specific bias, due to the spatial separation of the ULS and LLS. The experimental device did not, however, operate up to a temperature as high as expected from the original simulations, although the simulated current density matched the experimental one well. In order to more deeply investigate the design, we also simulated it using our NEGF model, and found excellent agreement to the experimental data both for current density and gain [37]. In contrast to the simpler model, however, we find that the extraction energy $E_{\mathrm{LLS}} - E_e \approx 27$ meV was not well matched to E_{LO}, as shown in Fig. 6.4 (a). Instead, we find that the transition from the ULS directly down to the extraction state has an energy only slightly larger than E_{LO}. We attribute the different results for the two models to the inclusion of a mean field potential, as well as the shift in energy coming from the real parts of the self-energies. From the densities in Fig. 6.4 (b) it is also evident that the extraction energy is too low, since the extraction state is highly populated at the same time as the LLS population is larger than that of the injector state.

Now we investigate the design presented in Paper II, called V845. This design

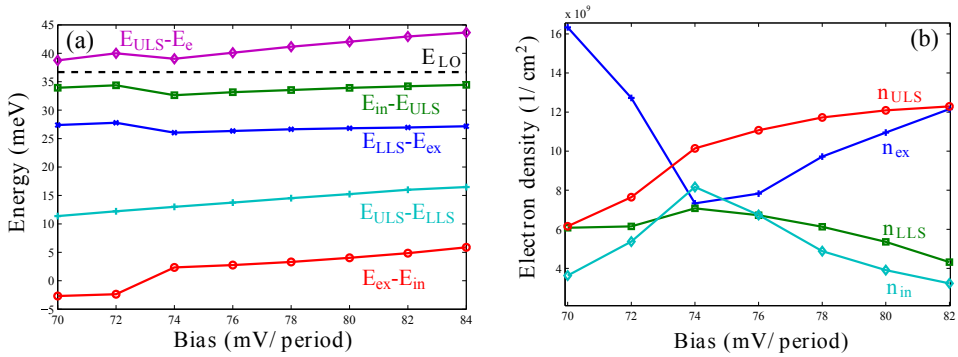


Figure 6.4: Transition energies (a) and populations (b) of the SA design V843 [37], as functions of the applied bias. Since the injection energy $E_i - E_{\mathrm{ULS}}$ is well matched to E_{LO} the injection rate is fast and the injector state population is low, except at the bias of the tunneling alignment where is shares density with the highly populated extraction state. The inversion is increasing even after the tunnel alignment since the leakage channel energy $E_{\mathrm{ULS}} - E_e$ increases away from E_{LO}.

was intended to improve V843 by better matching the extraction energy $E_{\mathrm{LLS}} - E_e$ to E_{LO} at the bias of the tunnel resonance. The NEGF simulations shown in Paper II were performed with the one-band model, and the difference when including the valence band components can be seen in Fig. 3.3. The main effects are a lowering of the overall gain by 5-10 cm^{-1}, a slight red-shift of the spectral peak gain position, and slightly higher current density. The energy levels of the Wannier-Stark states calculated in the two-band model with eq. (2.31) are shown in Fig. 6.6. Indeed this design has a better matching of $E_{\mathrm{LLS}} - E_e$ to E_{LO} compared to V843, and throughout the bias range in Fig. 6.6 the density of the LLS is kept low. However, we find that now the injection energy $E_i - E_{\mathrm{ULS}}$ deviates from E_{LO} instead. As a result, carriers accumulate in the injection state, seen by the high electron density in Fig. 6.5. Another effect from the low injection energy is that the energy difference between the ULS and the extractor state $E_{ULS} - E_e$ is close to E_{LO} (see Fig. 6.6), which works directly against inversion. As for the previous design, we attribute the main part of the differences between the rate equation results in Paper II and the results from the NEGF model to the real parts of the self-energies and the mean field potential included in the NEGF model.

In conclusion, we find that the tunneling injection scheme is operated close to it's theoretical limit, but is still limited to operation below 200 K lattice temperature. This prompts the use of different design schemes in order to reach higher operating temperatures, such as the two scattering assisted injection designs investig-

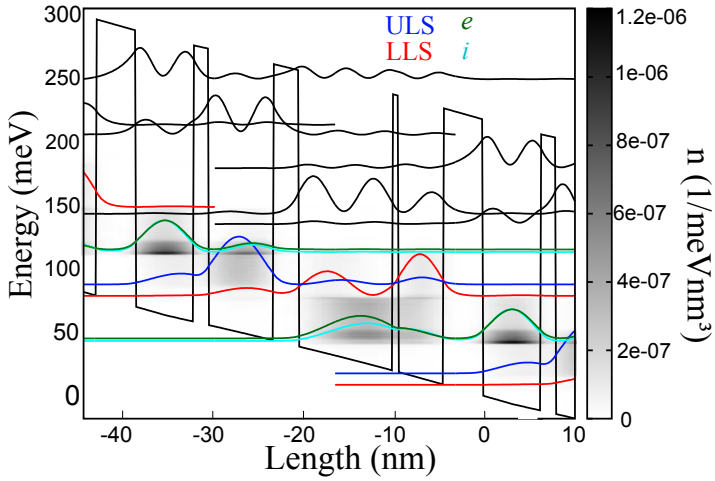


Figure 6.5: Electron density, calculated via eq. (2.23), and square of the Wannier-Stark states, calculated via eq. (2.31), at a bias of 70 mV/period and a lattice temperature of $T = 50$ K.

ated. However, one has a too low extraction energy, which results in a LLS population greater than the population of the injector state. For the other structure, the injection energy is too low instead, resulting in an extractor state population that is greater than the ULS population. Bringing both the injection and extraction energies close to E_{LO} at the same bias (close to the bias of the tunnel resonance), one would hope to achieve a configuration where $n_{\mathrm{ULS}} > n_e/n_i n > n_{\mathrm{LLS}}$. It therefore seems that this scheme has room for improvement by further optimizing the level energies, so that $E_i - E_{\mathrm{ULS}} \approx E_{\mathrm{LLS}} - E_e \approx E_{\mathrm{LO}}$ and $E_i \approx E_e$ at one specific bias. It is important to take into account mean field effects in such an optimization.

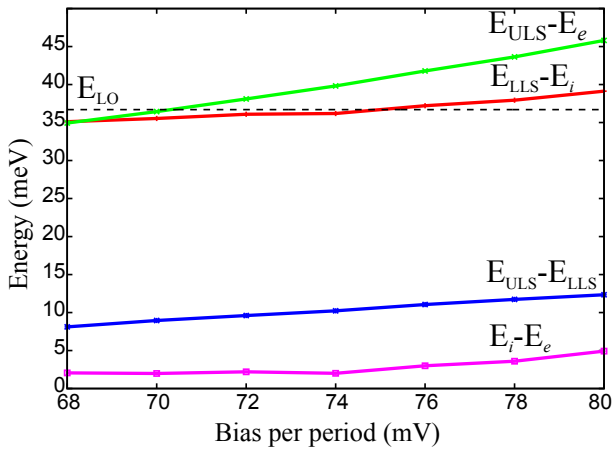


Figure 6.6: Energy differences between the extraction and injection states $E_i - E_e$, the extraction energy $E_{\mathsf{LLS}} - E_e$, the injection energy $E_i - E_{\mathsf{ULS}}$, as well as the leakage channel ULS $\longrightarrow e$ with energy $E_{\mathsf{ULS}} - E_e$. These energy levels are calculated for the Wannier-Stark states with the contribution from the self energy according to eq. (2.31), in contrast to Paper II where the states were calculated according to eq. (2.15). In both cases there is a close match between $E_{\mathsf{LLS}} - E_e$ and E_{LO}.

6.2 Two-quantum well designs

To date, the device with highest operating temperature is based on a resonant tunneling injection and resonant phonon extraction design, with a diagonal lasing transition and $x = 0.15$.[20] The diagonal design aims to limit the wrong injection directly into the lower laser state [93], as well as the emission of thermally activated phonons, by reducing the dipole matrix element z_{ul} between the upper (u) and lower (l) laser levels. This increases the inversion Δn_{ul}, but at the same time reduces the (scaled) oscillator strength $f_{ul} = 2m^*\omega_{ul}z_{ul}^2/\hbar$, where $\omega_{ul} = (E_u - E_l)/\hbar$ is the energy difference between u and l, and m^* is the effective electron mass[1]. Since the modal gain $G(\omega)$ given by eq. (2.49) close to the resonance laser frequency $\hbar\omega \approx \Delta E_{lu}$ is

$$G(\omega) \propto \frac{\Delta n_{ul}\hbar\omega z_{ul}^2}{d\Gamma_{\mathrm{w}}} \propto \frac{f_{ul}\Delta n_{ul}}{dm^*\Gamma_{\mathrm{w}}}, \tag{6.1}$$

where Γ_{w} is the transition line width and d is the period length, the maximum gain is found at some trade-off between oscillator strength and inversion.

The oscillator strength is proportional to the photon energy $\hbar\omega$, although at high energies $\hbar\omega \gtrsim 20$ meV, or frequencies $\nu \gtrsim 5$ THz, optical losses are increased by the absorption of photons by LO phonons, and the probability of thermally activated phonon emission is greater. The previous record-holding design [90] has $f_{ul} = 0.38$, while the corresponding value for the current record-holding design [20] is $f_{ul} = 0.58$. The good temperature performance of the latter design is attributed to the optimized oscillator strength. In contrast, a systematic study of four-well designs with scattering assisted injection [101], oscillator strengths between 0.26-0.47 had a negligible impact on the laser performance.

As the gain $G(\omega)$, given by eqs. (2.49) and (6.1), is inversely proportional to the period length d, a short design is beneficial. Refs. [90, 20] are based on three quantum wells, with period lengths of ~ 45 nm. In order to make the period shorter, a two-well scheme can be employed [39, 31, 102]. The design by Kumar *et al.* [39] has $d = 32.5$ nm, while the one by Scalari *et al.* [31] has $d = 34.5$ nm. The structures presented in this work are even shorter ($d = 31.4 - 31.6$ nm).

The two-well lasers in Refs. [39, 31] lased up to a lattice temperature of $T_L \sim 120$ K. Kumar *et al.* attributes [39] the poor temperature performance to a leakage current from the upper laser level into a highly excited parasitic state. We believe

[1] In the calculation of f we assume $m^* = 0.067m_0$ is constant for simplicity, as non-parabolicity is not expected to significantly alter the results in these types of structures. In the full NEGF calculations, however, non-parabolicity is taken into account in the two-band model described in Chap 3.

it is likely that the discrepancy between simulated ($j_{\mathrm{NEGF}} = 0.3$ kA/cm^2) and measured ($j_{\mathrm{meas.}} = 0.8$ kA/cm^2) current density for the sample of Ref. [31] for $T_L = 100$ K is similarly due to current leakage into the continuum. Compared to three- or four-well structures, the situation is worse with only two quantum wells, since the short period requires a higher electric field to encompass similar laser frequencies, and the wavefunction overlap between the highly occupied states and states of the next period is larger. To prevent such continuum leakage, a material with higher conduction band offset can be used [37, 41, 103, 104]. We find that even though interface roughness scattering is increased, it is of negligible importance (see Paper VI) for the structures studied here. Also, by keeping the Al content moderate $x \approx 0.25$, X-valley leakage is not expected to be a bigger concern than for $x = 0.15$.

Building upon these ideas, we have designed three two-quantum well lasers with varying oscillator strengths, based on the GaAs/$Al_{0.25}Ga_{0.75}$As material system. In addition, the designs have both resonant tunneling and phonon extraction (double extraction). The structures were optimized by iteratively varying the layer sequences to achieve better characteristics, such as gain and oscillator strength. Crude approximations to the final results were found with a density matrix approach [92, 60], while fine-tuning of the energy levels and wavefunctions was accompliced using a non-equilibrium Green's function (NEGF) model [64][Paper III]. The latter has a more elaborate treatment of scattering and coherences, and is therefore more reliable for THz structures, where coherences play an important role [59, 105]. Here it is important to point out that this model includes all of the discussed temperature degrading mechanisms, with the exception of current leakage to continuum states, since the basis states used are localized in nature [64][Paper III] and this would require an unfeasible number of states in the calculations. It also includes free-carrier absorption within the active region.

6.2.1 Design Concepts

Fig. 6.8 shows the three designs, denoted by A, B, and C. Common for all structures is the use of 25% Al barriers, as well as tunneling injection and phonon extraction. In addition, level 4 aligns with the level l of the previous period. This alignment also occurs for the two-well structures of Refs. [39] and [31], albeit at a different bias than the operation point. However, for these earlier designs with lower barriers, the wave-functions of levels 4 and 5 have significant amplitudes in the continuum, and alignments of these states with the highly populated active states is therefore at risk of providing significant leakage currents, reducing in-

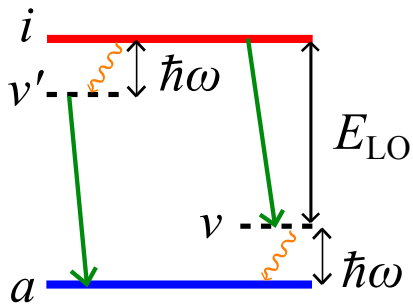


Figure 6.7: Schematic illustration of the process giving free-carrier gain (FCG). Radiative transitions from $i \rightarrow v$, or from $v' \rightarrow a$ are shown by wavy lines, while LO phonon transitions $v \rightarrow a$ and $i \rightarrow v'$ are shown as green arrows. v and v' are virtual states, while a can be chosen to be the ULS. In order to get a large FCG, the dipole matrix element between levels i and a should be large, as well as the population of i (though it does not have to be larger than that of a).

version and increasing current density and heating the active region. In contrast, levels 4 and 5 in Fig. 6.8 are well confined below the barriers, and we are free to use these levels to our benefit. At the operation bias, this alignment provides an additional depopulation of l via resonant tunneling, and then LO phonon emission from 4 directly to u' of the following period.

Another benefit of this scheme can be seen in Fig. 6.8, where all structures are shown at biases still in the region of positive differential conductance (PDC). Here, level u is actually lower in energy than the injector level 1. This is possible since the overlap between levels $u/5$ and $l/4$ provides PDC even after the alignment bias of $u/1$. This allows for higher inversion on the main laser transition than the limit (half the doping density) placed by the resonant tunneling transition, as the carriers will be more localized in the narrow well.

Elastic scattering is dominated by impurity scattering ($\Gamma_{\text{imp}} \gtrsim 10$ meV while $\Gamma_{\text{IFR}} \lesssim 1$ meV), since for these short structures impurities must be placed where they affect either u or l. We have minimized this impact by placing the dopants where the wavefunction l has it's node in the wider well.[31]

In QCL heterostructures, the free-carrier absorption (FCA) is quite different from that of bulk materials [106]; it is much smaller and the biggest contribution comes from the interaction with light via LO phonon scattering, for the structures studied in this work. In fact, if the electron temperature T_e is high enough above T_L, $T_e \geq T_L(1+\hbar\omega/E_{\text{LO}})$, where $E_{\text{LO}} = 36$ meV is the LO phonon energy, the FCA mediated by LO phonon absorption can actually become negative [106, 107]. As is illustrated in Fig. 6.7, this process requires a transition

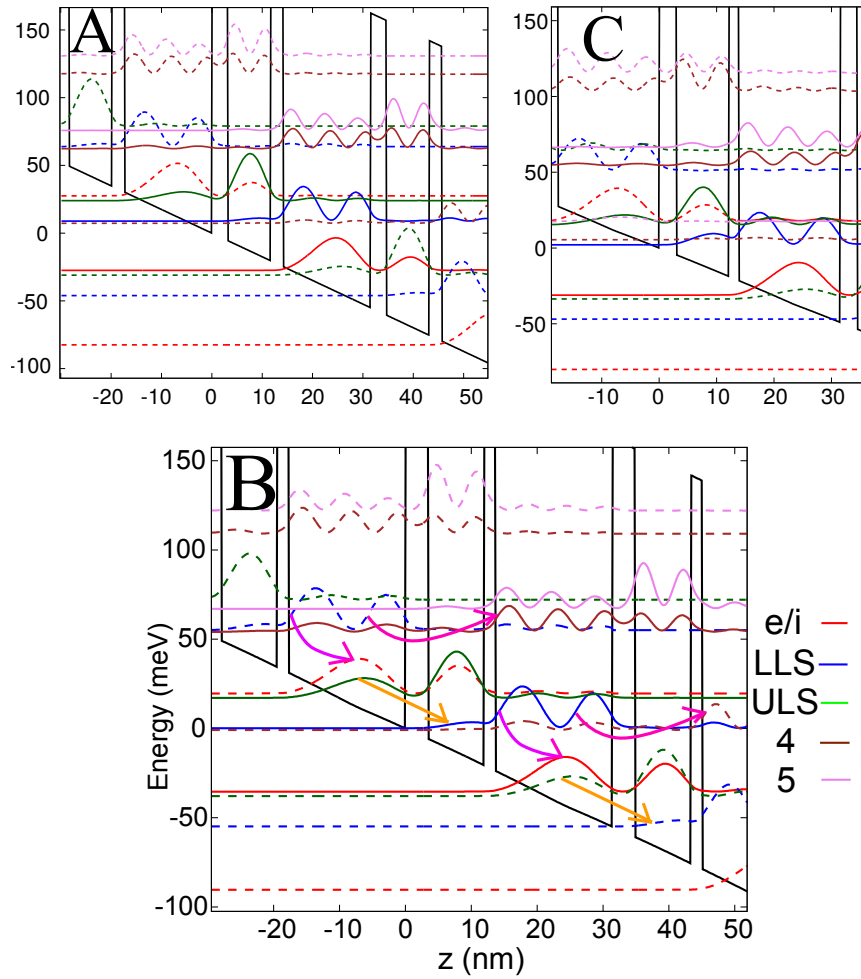


Figure 6.8: Band structure and square of the wavefunctions for the three two-well designs close to the peak current density (at biases of 55, 55, and 49 mV/period respectively for A, B, and C), where A has similar oscillator and coupling strengths, as well as frequency as Ref. [31], B has been optimized for high gain at 200 K, and C is designed to have large negative free-carrier absorption. The layer sequences for the respective structures are A: **33**-85-**26**-172, B: **34**-85-**18**-177, and C: **30**-92-**16**-177, where bold face denotes the barriers. All three lasers are doped to $3 \cdot 10^{10}$ cm^{-2} in the middle of the widest well in order to minimize the impact of impurity scattering.

from a highly populated level i to one of two virtual states v and v' of energy $E_v = E_a + E_{\mathrm{LO}} = E_i - \hbar\omega$ and $E_{v'} = E_i - E_{\mathrm{LO}} = E_a + \hbar\omega$. Here, level a has an energy $E_{\mathrm{LO}} + \hbar\omega$ below level i. In addition, eq. (14) of Ref. [106] suggests that one should choose as low frequency, short period length, high dipole moment z_{ai}, and high electron temperature T_e as possible, in order to improve this free-carrier gain (FCG) [107].

6.2.2 Simulation results

NEGF simulations of the previous two-well designs (Refs. [39, 31]) have shown a large pre-peak at the bias where l aligns with level 1. As can be seen in Fig. 6.9, this is not the case in the present designs. Further simulation results are summarized in Tab. 6.1.

Table 6.1: Summary of NEGF simulations for the two-well designs in Fig. 6.8 as well as Refs. [39] and [31]. Units for doping density and inversion, current density, and gain are 10^{10} cm^{-2}, kA/cm^2, and cm^{-1}, respectively.

Sample	$n_{2\mathrm{D}}$	$j_{\max}^{\mathrm{NEGF}}$ (T_L)	$G_{\max}$ (T_L)	$f_{\mathrm{osc.}}$	Δn_{ul}
Ref. [39]	2.2	1.20 (200 K)	34 (200 K)	0.53	0.65
Ref. [31]	1.5	0.46 (200 K)	13 (200 K)	0.13	0.44
A	3	0.90 (200 K)	32 (200 K)	0.11	1.20
	1.5	0.49 (200 K)	20 (200 K)		0.55
B	3	1.51 (200 K)	57 (200 K)	0.3	0.95
	3	1.53 (300 K)	20 (300 K)		0.59
	2	1.05 (200 K)	43 (200 K)		0.56
	1.5	0.80 (200 K)	35 (200 K)		0.41
C	3	2.41 (200 K)	46 (200 K)	0.47	0.41
	2	1.62 (200 K)	36 (200 K)		0.38[a]
	1.5	1.25 (200 K)	28 (200 K)		0.31

[a] Note that for C, the gain is the sum of the transitions $u \to l$, $5 \to 4$, as well as FCG, which explains the low inversion of the main transition compared to Kumar *et al.* [39].

Laser A

Laser A has been designed to have an oscillator strength $f_{ul} = 0.11$ (0.13 in Ref. [31]), frequency $\hbar\omega = 17$ meV (16 meV in Ref. [31]), and tunnel coupling $\hbar\Omega$ as close as possible to the QCL of Ref. [31]. Evidence of the similarities of the structures is the close match between the current densities and population inversions in Tab. 6.1. Concerning the slight difference in oscillator strengths, it

should be noted that this quantity varies significantly with bias. The optimized structure A has a 65% higher gain than the reference QCL, even though inversion is only 25% higher.

Laser B

Laser B has been optimized to give as high a gain as possible at 200 K. We see from Tab. 6.1 that for the high doped version we find a very high gain of 55 cm^{-1} at 200 K, and still 20 cm^{-1} at 300 K. This is much better than the gain of the previous designs [39, 31]. The efficacy of the double extraction scheme is evident in the electron densities ρ_i of laser B for the simulation parameters of Fig. 6.8 (b), where we find $\rho_u = \rho_1 = 1.3 \cdot 10^{10}$ cm^{-2}, $\rho_l = 3.5 \cdot 10^9$ cm^{-2}, $\rho_4 = 1.3 \cdot 10^9$ cm^{-2}, and $\rho_5 = 2.2 \cdot 10^8$ cm^{-2} at $T_L = 200$ K. This shows that resonant tunneling from level l to level 4 is active, and that inversion is good even at this high lattice temperature (see also Tab. 6.1). Here, care has to be taken as we have actually negative inversion between levels 4 and 5, which might provide large absorption due to the very large $f_{45} = 1.6$. This is avoided by having an energy separation $E_5 - E_4 = 11$ meV between levels 4 and 5 that is smaller than the laser transition energy $\hbar\omega = 17$ meV.

As can be seen in Tab. 6.1, the gain does not scale linearly with doping; we find significantly more than half the gain for half the doping density for all samples, which can be explained partly by an increased line broadening Γ_w due to increased impurity scattering.

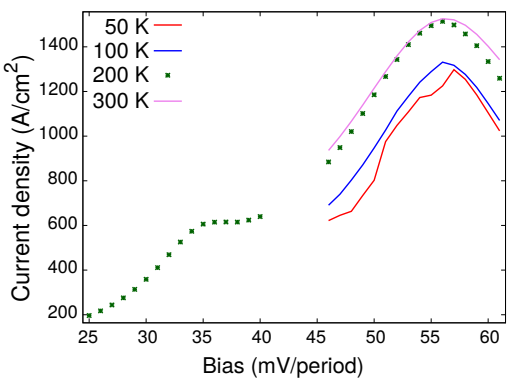


Figure 6.9: Current density at different lattice temperatures for laser B. The effect of electron-electron scattering is a reduction of the peak current density by less than 10%.

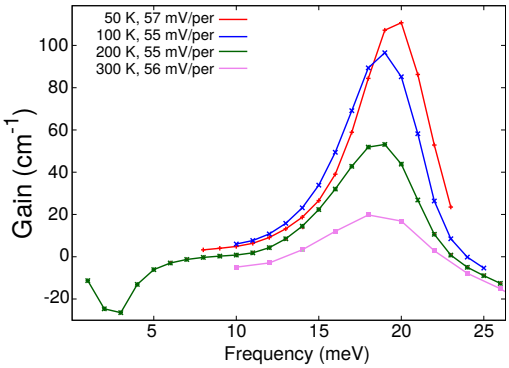


Figure 6.10: Gain at different lattice temperatures for laser B. Including electron-electron scattering, the gain decreases by ~ 5 cm^{-1}. There is a slight red-shift with increasing temperature owing to an increased impurity scattering rate as the screening diminishes.

Laser C

Finally, laser C has gain from three contributing channels; (i) the normal gain transition from u to l, with $\Delta n_{ul} = 0.38{\cdot}10^{10}$ cm^{-1}; (ii) the transition 5 to 4 with a very large dipole matrix element ($f_{45} = 1.75$) and inversion of $\Delta n_{45} = 0.1{\cdot}10^{10}$ cm^{-1}, which is possible since the alignment of u and level 5 occurs at a lower bias than the alignment of l and 4; (iii) free-carrier gain (FCG) between level 5 and virtual states one E_{LO} above u and below l, as explained above. The levels a and i are here levels u and 5, respectively, where the dipole moment z_{u5} has been maximized. This was achieved by swapping energies of levels 4 and 5 compared to lasers A and B; it can be seen in Fig. 6.8 c) that level 4 is now the first excited state of the narrow well, while level 5 is the second excited state of the wide well. In addition, the laser barrier was made thinner, which simultaneously gives a large oscillator strength of $f_{ul} = 0.47$. For a lattice temperature of $T_L = 200$ K, Fig. 6.11 shows the calculated FCG using either $T_e = 280$ or 300 K in eq (14) of Ref. [106]. The behavior at the laser frequency $\hbar\omega = 14$ meV is very sensitive to electron temperature, and we need $T_e > 300$ K in order not to have absorption, in which case the contribution to the gain is expected to be a few cm^{-1} for a doping concentration of $3 \cdot 10^{10}$ cm^{-2}. In fact, in laser C, the upper level 5 is highly populated as it is aligned with u of the previous period, and the effective temperature T_e is therefore significantly higher. Although $E_5 - E_u$ is not perfectly matched to $\hbar\omega + E_{\mathrm{LO}}$, further optimization is not expected to yield a FCG comparable to the total gain.

The gain for laser C is lower than that of B, mainly due to the large oscillator

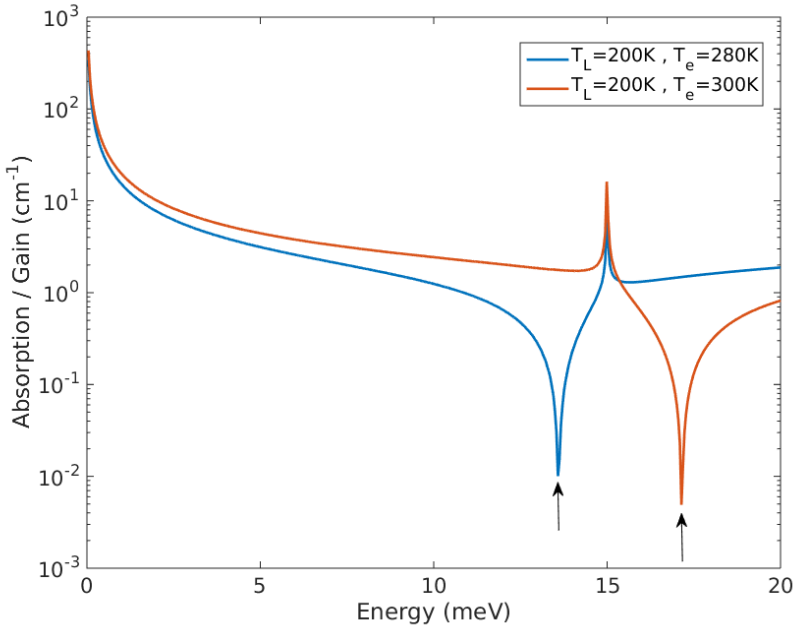


Figure 6.11: Logarithm of the modulus of the free carrier absorption due to LO phonon absorption, for laser C. Below the arrows there is gain and above there is absorption. At the frequency $(E_5 - E_u)/\hbar$, close to the laser frequency, there is a gain peak for $T_e = 300$ K, and an absorption peak for $T_e = 280$ K. The doping density is $n_{2D} = 3 \cdot 10^{10}$ cm^{-2}. Figure and simulations made by Camille Ndebeka-Bandou, ETH Zürich 2015.

strength, which gives a low inversion. Compared to the laser by Kumar *et al.*, the design is quite similar, as evident from the similar f_{ul}, current density, and gain. They differ significantly in Δn_{ul}, however, which is explained by the separation of the total gain of 36 cm^{-1} into the two main channels (i) and (ii); channel (i) has a product of $\Delta n_{ul} \cdot f_{ul}$ =0.175 cm^{-1}, while the corresponding number for channel (ii) is 0.179 cm^{-1}. As their transition line widths are similar, this shows that the transition 5 to 4 contributes to the total gain only slightly less than the conventional gain transition.

6.2.3 Conclusions

We have presented three (A, B, and C) two-quantum well THz QCL designs, with tunneling injection and both LO phonon and resonant tunneling extraction. The latter extraction mechanism is made possible by the use of taller barriers as compared to previous two-well designs [39, 31], a feature which is also expected to reduce parasitic currents into the continuum [41]. Compared to the design of Ref. [31], which is similar to but with lower barriers than A, we find almost twice the gain, and only a slight increase in the simulated current density. Furthermore, we expect that higher barriers lead to a better agreement between measurements and simulations, reducing the current density as well as increasing the gain in a real device. Another sample, B, has a very high gain (57 cm^{-1}) at 200 K, and the NEGF simulations predict that the gain is possibly sufficient to overcome the optical losses even at room temperature. Finally, we have designed a laser, C, which has three distinct gain channels which together provides a gain of 46 cm^{-1} at 200 K. The most contributing channel after the conventional one, exploits a transition with low inversion but very large oscillator strength between highly excited states. While such transitions decay rapidly with increased optical field strength, it is promising for enhancement of the maximum operating temperature, especially if the inversion of the main transition can be improved.

Finally, it is worth to note that the only effect of the specific Al content used is on the band offset $\Delta E_c = 217.5$ meV and the barrier effective mass $m_b = 0.0877$. These designs were optimized using these values, and since it is not entirely clear which offset to use for a specific Al contents, a growth may require adjustments of this value in order to give the required ΔE_c.

Chapter 7

Conclusions and outlook

This thesis describes the extension of an NEGF approach used to simulate THz QCLs to great success, to give accurate results also for mid-IR lasers for the first time. This has paved the way for future improvements of mid-IR designs, using a model that accounts for the complex physics in the most fundamental and sophisticated way and at the same time gives results in close agreement with experimental data for real devices. To achieve these results, the model has now been generalized to a two-band model as non-parabolicity is important for mid-IR devices. In order to perform simulations within reasonable time, the program has been parallelized in two stages; to be able to run locally with multiple processors sharing memory, as well as massively parallel cluster computations using message passing communications. This was absolutely necessary in order to obtain the results for mid-IR QCLs presented in this thesis. After these implementations, the model was used to successfully reproduce experimental results for several mid-IR QCLs, while THz QCL behavior can still be modeled with great accuracy. Thus, we now have a computer model that can accurately reproduce experimental behavior of a wide range of samples; from two-well based, as well as conventional, structures for THz radiation, to structures covering all of the mid-IR spectrum from 3-12 μm.

The impact of interface roughness scattering was investigated, as this is more important in mid-IR devices due to their higher conduction band offsets. In detail, we study relevant interface roughness auto-correlation functions (exponential and Gaussian to be precise), the inclusion of the valence band, as well as vertical interface correlations. We find that it is important to consider all of these points, while acknowledging that the details of the actual roughness distribution of experimental devices remain unknown. In future versions of the NEGF

program, it would therefore be motivated to include more thoroughly vertical interface correlations.

In the course of this work, the NEGF model has been compared to several simpler models. We find that these results agree well, and thus we have asserted the confidence in the accuracy of ours and others' current state-of-the art models of varying degree of complexity. Some investigations into THz samples have also been conducted, as these devices are highly interesting for the under-developed THz spectral range. The most important results are the assertion that the SA THz structures in the published literature are not perfectly designed, and suggestions of improvements are found, as well as the affirmation of the success of the resonant phonon depopulation scheme. We also propose three new THz designs that show a promising gain at lattice temperatures above 240 K (and which can therefore be cooled electronically). These new devices will be grown and tested in the future, and such measurements may indicate further improvements to bring THz QCLs closer to room temperature operation.

As an outlook, a new time-resolved density matrix model is under development, and therefore not presented in this thesis. This model treats the time evolution of all coherences of the density matrix, and includes all scattering between coherences and densities using a Boltzmann distribution for the in-plane momenta. While the model shows promising results for simple test cases, due to the choice of basis states, QCLs have not yet been accurately modeled. The strength of such a model would be faster calculations and time-resolved simulations (enabling the study of electro-magnetic pulses, multiple simultaneous frequencies and much more), which are of high interest in many QCL applications. It would also be a great tool in conjunction with the NEGF model, since the latter could be used to validate the former and cross-check results between the models.

In conclusion, we now have an NEGF model capable of accurately and reasonably fast simulating the whole range of QCLs; from the longest wavelengths of THz QCLs, to the shortest IR ones. The model can be used for designing new structures, or to give in-depths analysis and explanations of the performance of experimental devices. This will be an important tool for developing future QCL devices, in an exciting field that is rapidly growing and transitioning from scientific curiosity to real world applications. And perhaps in the future, we will see something like what is envisioned in the upcoming computer game "Star Citizen": "Quantum cascade lasers fire electromagnetic bolts that convert any material contacted into plasma, giving it a reasonable 'bite' that is useful in most space combat encounters." [108] However, this requires a significant increase in the output power.

References

[1] R. F. Kazarinov and R. A. Suris, *Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice*, Sov. Phys. Semicond. **5**, 707 (1971)

[2] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, *Quantum cascade laser*, Science **264**, 553 (1994)

[3] B. S. Williams, *Terahertz quantum-cascade lasers*, Nat. Photonics **1**, 517 (2007)

[4] F. Capasso, *High-performance midinfrared quantum cascade lasers*, J. Opt. Eng. **49**, 111102 (2010)

[5] A. Godard, *Infrared (2-12 μm) solid-state laser sources: a review*, Comptes Rendus Physique **8**, 1100 (2007)

[6] Y. Yao, A. J. Hoffman, and C. F. Gmachl, *Mid-infrared quantum cascade lasers*, Nature photonics **6**, 432 (2012)

[7] M. Razeghi, S. Slivken, Y. Bai, B. Gokden, and S. R. Darvish, *High power quantum cascade lasers*, New Journal of Physics **11**, 125017 (2009)

[8] M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior, *Continuous wave operation of a mid-infrared semiconductor laser at room temperature*, Science **295**, 301 (2002)

[9] J. B. McManus, M. S. Zahniser, D. D. Nelson, Jr., J. H. Shorter, S. Herndon, E. Wood, and R. Wehr, *Application of quantum cascade lasers to high-precision atmospheric trace gas measurements*, Optical Engineering **49**, 111124 (2010)

[10] *Website of QuantaRed Technologies*, last acquired April 2016

[11] A. Tsekoun, A. Lyakh, R. Maulini, M. Lane, T. Macdonald, R. Go, and C. K. N. Patel, *High power and efficiency quantum cascade laser systems for defense and security applications*, in *Proc. SPIE*, volume 7325, pp. 73 250L–73 250L–7 (2009)

[12] A. Tredicucci, C. Gmachl, F. Capasso, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, *A multiwavelength semiconductor laser*, Nature **396**, 350 (1998)

[13] A. Hugi, R. Terazzi, Y. Bonetti, A. Wittmann, M. Fischer, M. Beck, J. Faist, and E. Gini, *External cavity quantum cascade laser tunable from 7.6 to 11.4 μm*, Applied Physics Letters **95**, 061103 (2009)

[14] R. Lewicki, A. A. Kosterev, D. M. Thomazy, T. H. Risby, S. Solga, T. B. Schwartz, and F. K. Tittel, *Real time ammonia detection in exhaled human breath using a distributed feedback quantum cascade laser based sensor*, in *SPIE OPTO*, pp. 79 450K–79 450K (International Society for Optics and Photonics, 2011)

[15] Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, *Room temperature quantum cascade lasers with 27% wall plug efficiency*, Appl. Phys. Lett. **98**, 181102 (2011)

[16] Y. Bai, S. Slivken, S. Kuboya, S. R. Darvsih, and M. Razeghi, *Quantum cascade lasers that emit more light than heat*, Nat. Photonics **4**, 99 (2010)

[17] *Web page of the MIRIFISENS project*, last acquired April 2016

[18] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, *Terahertz semiconductor-heterostructure laser*, Nature **417**, 156 (2002)

[19] M. A. Belkin, F. Capasso, F. Xie, A. Belyanin, M. Fischer, A. Wittmann, and J. Faist, *Room temperature terahertz quantum cascade laser source based on intracavity difference-frequency generation*, Appl. Phys. Lett. **92**, 201101 (2008)

[20] S. Fathololoumi, E. Dupont, C. Chan, Z. Wasilewski, S. Laframboise, D. Ban, A. Mátyás, C. Jirauschek, Q. Hu, and H. C. Liu, *Terahertz quantum cascade lasers operating up to $\sim$ 200 K with optimized oscillator strength and improved injection tunneling*, Opt. Express **20**, 3866 (2012)

[21] M. S. Vitiello, G. Scamarcio, V. Spagnolo, S. S. Dhillon, and C. Sirtori, *Terahertz quantum cascade lasers with large wall-plug efficiency*, Applied Physics Letters **90**, 191115 (2007)

[22] D. D. Nelson, J. B. McManus, S. C. Herndon, J. H. Shorter, M. S. Zahniser, S. Blaser, L. Hvozdara, A. Muller, M. Giovannini, and J. Faist, *Characterization of a near-room-temperature, continuous-wave quantum cascade laser for long-term, unattended monitoring of nitric oxide in the atmosphere*, Opt. Lett. **31**, 2012 (2006)

[23] M. Lindskog, D. Winge, and A. Wacker (2014), eye witness account at the 20th anniversary of the QCL workshop, ETH Züric

[24] C. Sirtori, F. Capasso, J. Faist, and S. Scandolo, *Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells*, Phys. Rev. B **50**, 8663 (1994)

[25] A. Bismuto, R. Terazzi, M. Beck, and J. Faist, *Electrically tunable, high performance quantum cascade laser*, Applied Physics Letters **96**, 141105 (2010)

[26] A. Bismuto, M. Beck, and J. Faist, *High power Sb-free quantum cascade laser emitting at 3.3 μm above 350 K*, Applied Physics Letters **98**, 191104 (2011)

[27] J. M. Wolf, S. Riedi, M. J. Süess, M. Beck, and J. Faist, *3.36 μm single-mode quantum cascade laser with a dissipation below 250 mW*, Opt. Express **24**, 662 (2016)

[28] C. Jirauschek and T. Kubis, *Modeling techniques for quantum cascade lasers*, Appl. Phys. Rev. **1**, 011307 (2014)

[29] P. Krötz, D. Stupar, J. Krieg, G. Sonnabend, M. Sornig, F. Giorgetta, E. Baumann, M. Giovannini, N. Hoyler, D. Hofstetter, and R. Schieder, *Applications for quantum cascade lasers and detectors in mid-infrared high-resolution heterodyne astronomy*, Applied Physics B **90**, 187 (2007)

[30] D. Turcinková, *Terahertz Quantum Cascade Lasers for Astronomical Applications*, Ph.D. thesis, ETH Zürich (2014)

[31] G. Scalari, M. I. Amanti, C. Walther, R. Terazzi, M. Beck, and J. Faist, *Broadband THz lasing from a photon-phonon quantum cascade structure*, Opt. Express **18**, 8043 (2010)

[32] A. Wacker, *Quantum cascade laser: An emerging technology*, in *Nonlinear Laser Dynamics*, edited by K. Lüdge (Wiley-VCH, Berlin, 2012)

[33] L. Le, X. Wang, J.-Y. Fan, M. Troccoli, D. L. Sivco, and C. F. Gmachl, *Ultra-broadband (3.3-12.5 μm) single stack quantum cascade gain medium*, in *CLEO: 2015*, p. STu2G.4 (Optical Society of America, 2015)

[34] M. Yamanishi, K. Fujita, T. Edamura, and H. Kan, *Indirect pump scheme for quantum cascade lasers: dynamics of electron-transport and very high T_0-values*, Opt. Express **16**, 20748 (2008)

[35] O. Cathabard, R. Teissier, J. Devenson, J. C. Moreno, and A. N. Baranov, *Quantum cascade lasers emitting near 2.6 μm*, Appl. Phys. Lett. **96**, 141110 (2010)

[36] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno, *Terahertz quantum-cascade laser operating up to 137 K*, Appl. Phys. Lett. **83**, 5142 (2003)

[37] E. Dupont, S. Fathololoumi, Z. R. Wasilewski, G. Aers, S. R. Laframboise, M. Lindskog, S. G. Razavipour, A. Wacker, D. Ban, and H. C. Liu, *A phonon scattering assisted injection and extraction based terahertz quantum cascade laser*, J. Appl. Phys. **111**, 073111 (2012)

[38] M. I. Amanti, G. Scalari, R. Terazzi, M. Fischer, M. Beck, J. Faist, A. Rudra, P. Gallo, and E. Kapon, *Bound-to-continuum terahertz quantum cascade laser with a single-quantum-well phonon extraction/injection stage*, New J. Phys. **11**, 125022 (2009)

[39] S. Kumar, C. W. I. Chan, Q. Hu, and J. L. Reno, *Two-well terahertz quantum-cascade laser with direct intrawell-phonon depopulation*, Appl. Phys. Lett. **95**, 141110 (2009)

[40] A. Albo and Q. Hu, *Investigating temperature degradation in THz quantum cascade lasers by examination of temperature dependence of output power*, Appl. Phys. Lett. **106**, 131108 (2015)

[41] A. Albo and Q. Hu, *Carrier leakage into the continuum in diagonal GaAs/Al0.15GaAs terahertz quantum cascade lasers*, Appl. Phys. Lett. **107**, 241101 (2015)

[42] D. O. Winge, M. Franckié, and A. Wacker, *Superlattice gain in positive differential conductivity region*, preprint arXiv:1601.00915 (2016)

[43] A. Cho and J. Arthur, *Molecular beam epitaxy*, Progress in Solid State Chemistry **10**, 157 (1975)

[44] J. Faist, *Quantum Cascade Lasers* (Oxford University Press, Oxford, 2013)

[45] C. G. Van de Walle, *Band lineups and deformation potentials in the model-solid theory*, Phys. Rev. B **39**, 1871 (1989)

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

[47] M. Hosoda, N. Ohtani, H. Mimura, K. Tominaga, P. Davis, T. Watanabe, G. Tanaka, and K. Fujiwara, *Evidence for $\Gamma - X$ transport in type-I GaAs/AlAs semiconductor superlattices*, Phys. Rev. Lett. **75**, 4500 (1995)

[48] H. Mimura, N. Ohtani, M. Hosoda, K. Tominaga, T. Watanabe, G. Tanaka, and K. Fujiwara, *Delayed photocurrent affected by X-resonance in GaAs/AlAs type-I short-period superlattices*, Applied Physics Letters **67**, 3292 (1995)

[49] C. G. Van de Walle and R. M. Martin, *Theoretical study of band offsets at semiconductor interfaces*, Phys. Rev. B **35**, 8154 (1987)

[50] L. D. Shvartsman and B. Laikhtman, *Bipolar THz-lasing structures based on InAs-GaSb coupled quantum wells as an alternative to intersubband lasing*, pp. 88 460K–88 460K–14 (2013)

[51] W. Terashima and H. Hirayama, *Molecular beam epitaxy growth of GaN/AlGaN quantum cascade structure using droplet elimination by thermal annealing technique*, physica status solidi (a) **208**, 1187 (2011)

[52] H. Yasuda, T. Kubis, I. Hosako, and K. Hirakawa, *Non-equilibrium greenâs function calculation for gan-based terahertz-quantum cascade laser structures*, Journal of Applied Physics **111**, 083105 (2012)

[53] H. Chou, J. Zeller, T. Manzur, and M. Anwar, *$ZnO/Zn_{1-x}Mg_xO$ QCL: A high power room temperature THz source*, in *Lester Eastman Conference on High Performance Devices (LEC), 2012*, pp. 1–4 (2012)

[54] H. C. Chou, A. Mazady, J. Zeller, T. Manzur, and M. Anwar, *Room-temperature quantum cascade laser: $ZnO/Zn_{1-x}Mg_xO$ versus $GaN/Al_xGa_{1-x}N$*, Journal of Electronic Materials **42**, 882 (2013)

[55] A. Matyas, R. Chashmahcharagh, I. Kovacs, P. Lugli, K. Vijayraghavan, M. A. Belkin, and C. Jirauschek, *Improved terahertz quantum cascade laser with variable height barriers*, Journal of Applied Physics **111**, 103106 (2012)

[56] A. Gajić, J. Radovanović, V. Milanović, D. Indjin, and Z. Ikonić, *Genetic algorithm applied to the optimization of quantum cascade lasers with second harmonic generation*, Journal of Applied Physics **115**, 053712 (2014)

[57] A. Bismuto, R. Terazzi, B. Hinkov, M. Beck, and J. Faist, *Fully automatized quantum cascade laser design by genetic optimization*, Applied Physics Letters **101**, 021103 (2012)

[58] E. Dupont, S. Fathololoumi, and H. C. Liu, *Simplified density-matrix model applied to three-well terahertz quantum cascade lasers*, Phys. Rev. B **81**, 205311 (2010)

[59] H. Callebaut and Q. Hu, *Importance of coherence for electron transport in terahertz quantum cascade lasers*, J. Appl. Phys. **98**, 104505 (2005)

[60] R. Terazzi, *Transport in quantum cascade lasers*, Ph.D. thesis, ETH Zürich (2011)

[61] S. M. Goodnick and P. Lugli, *Effect of electron-electron scattering on nonequilibrium transport in quantum-well systems*, Phys. Rev. B **37**, 2578 (1988)

[62] M. Manenti, F. Compagnone, A. Di Carlo, and P. Lugli, *Monte carlo simulations of THz quantum-cascade lasers*, J. Comput. Electron. **2**, 433 (2003)

[63] X. Gao, D. Botez, and I. Knezevic, *X-valley leakage in GaAs-based midinfrared quantum cascade lasers: A Monte Carlo study*, Journal of Applied Physics **101**, 063101 (2007)

[64] S.-C. Lee and A. Wacker, *Nonequilibrium Green's function theory for transport and gain properties of quantum cascade structures*, Phys. Rev. B **66**, 245314 (2002)

[65] N. Vukmirovic, D. Indjin, Z. Ikonic, and P. Harrison, *Electron transport and terahertz gain in quantum-dot cascades*, IEEE Photonics Tech. Lett. **20**, 129 (2008)

[66] T. Schmielau and M. Pereira, *Nonequilibrium many body theory for quantum transport in terahertz quantum cascade lasers*, Appl. Phys. Lett. **95**, 231111 (2009)

[67] T. Kubis and P. Vogl, *Predictive quantum theory of current and optical gain in quantum cascade lasers*, Laser Physics **19**, 762 (2009)

[68] H. Yasuda, T. Kubis, P. Vogl, N. Sekine, I. Hosako, and K. Hirakawa, *Nonequilibrium Green's function calculation for four-level scheme terahertz quantum cascade lasers*, Appl. Phys. Lett. **94**, 151109 (2009)

[69] G. Haldaś and, A. Kolek, and I. Tralle, *Modeling of mid-infrared quantum cascade laser by means of nonequilibrium Green's functions*, Quantum Electronics, IEEE Journal of **47**, 878 (2011)

[70] S. Steiger, M. Povolotskyi, H. H. Park, T. Kubis, and G. Klimeck, *NEMO5: A parallel multiscale nanoelectronics modeling tool*, IEEE Transactions on Nanotechnology **10**, 1464 (2011)

[71] G. H. Wannier, *The structure of electronic excitation levels in insulating crystals*, Phys. Rev. **52**, 0191 (1937)

[72] A. Bruno-Alfonso and D. R. Nacbar, *Wannier functions of isolated bands in one-dimensional crystals*, Phys. Rev. B **75**, 115428 (2007)

[73] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, 1996)

[74] C. G. Broyden, *A class of methods for solving nonlinear simultaneous equations*, Math. Comp. **19**, 577 (1965)

[75] T. Ando, A. B. Fowler, and F. Stern, *Electronic properties of two-dimensional systems*, Rev. Mod. Phys. **54**, 437 (1982)

[76] Y. V. Flores, S. S. Kurlov, M. Elagin, M. P. Semtsiv, and W. T. Masselink, *Leakage current in quantum-cascade lasers through interface roughness scattering*, Appl. Phys. Lett. **103**, 161102 (2013)

[77] A. Bismuto, R. Terazzi, M. Beck, and J. Faist, *Influence of the growth temperature on the performances of strain-balanced quantum cascade lasers*, Applied Physics Letters **98**, 091105 (2011)

[78] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, *Surface roughness at the Si(100)-SiO_2 interface*, Phys. Rev. B **32**, 8171 (1985)

[79] R. M. Feenstra, D. A. Collins, D. Z. Y. Ting, M. W. Wang, and T. C. McGill, *Interface roughness and asymmetry in InAs/GaSb superlattices studied by scanning tunneling microscopy*, Phys. Rev. Lett. **72**, 2749 (1994)

[80] A. Y. Lew, S. L. Zuo, E. T. Yu, and R. H. Miles, *Correlation between atomic-scale structure and mobility anisotropy in InAs/$Ga_{1-x}In_xSb$ superlattices*, Phys. Rev. B **57**, 6534 (1998)

[81] K.-J. Chao, N. Liu, C.-K. Shih, D. W. Gotthold, and B. G. Streetman, *Factors influencing the interfacial roughness of InGaAs/GaAs heterostructures: A scanning tunneling microscopy study*, Applied Physics Letters **75**, 1703 (1999)

[82] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, New York, 1998), 3rd edition

[83] D. G. Revin, M. R. Soulby, J. W. Cockburn, Q. Yang, C. Manz, and J. Wagner, *Dispersive gain and loss in midinfrared quantum cascade laser*, Appl. Phys. Lett. **92**, 081110 (2008)

[84] D. O. Winge, M. Lindskog, and A. Wacker, *Microscopic approach to second harmonic generation in quantum cascade lasers*, Opt. Express **22**, 18389 (2014)

[85] A. Wacker, *Coherence and spatial resolution of transport in quantum cascade lasers*, phys. stat. sol. (c) **5**, 215 (2008)

[86] M. Załużny, *Intersubband absorption line broadening in semiconductor quantum wells: Nonparabolicity contribution*, Phys. Rev. B **43**, 4511 (1991)

[87] *OpenMP official web site, www.openmp.org* (2016)

[88] *Lunarc cluster at Lund University, web page www.lunarc.lu.se* (2016)

[89] *OpenMPI official web site, www.open-mpi.org* (2016)

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

[91] (2013), internal MIRIFISENS document

[92] R. Terazzi and J. Faist, *A density matrix model of transport and radiation in quantum cascade lasers*, New Journal of Physics **12**, 033045 (2010)

[93] D. Indjin, P. Harrison, R. W. Kelsall, and Z. IkoniÄ, *Mechanisms of temperature performance degradation in terahertz quantum-cascade lasers*, Applied Physics Letters **82**, 1347 (2003)

[94] R. Nelander and A. Wacker, *Temperature dependence of the gain profile for terahertz quantum cascade lasers*, Appl. Phys. Lett. **92**, 081102 (2008)

[95] H. Li, J. C. Cao, Z. Y. Tan, Y. J. Han, X. G. Guo, S. L. Feng, H. Luo, S. R. Laframboise, and H. C. Liu, *Temperature performance of terahertz quantum-cascade lasers: experiment versus simulation*, Journal of Physics D: Applied Physics **42**, 025101 (2009)

[96] Y. Chassagneux, Q. Wang, S. Khanna, E. Strupiechonski, J. Coudevylle, E. Linfield, A. Davies, F. Capasso, M. Belkin, and R. Colombelli, *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 (2012)

[97] S. Khanal, L. Zhao, J. L. Reno, and S. Kumar, *Temperature performance of terahertz quantum-cascade lasers with resonant-phonon active-regions*, J. Opt. **16**, 094001 (2014)

[98] M. S. Vitiello, G. Scalari, B. Williams, and P. D. Natale, *Quantum cascade lasers: 20 years of challenges*, Opt. Express **23**, 5167 (2015)

[99] X. Gao, D. Botez, and I. Knezevic, *X-valley leakage in GaAs/AlGaAs quantum cascade lasers*, Appl. Phys. Lett. **89**, 191119 (2006)

[100] C. Jirauschek, G. Scarpa, P. Lugli, M. S. Vitiello, and G. Scamarcio, *Comparative analysis of resonant phonon THz quantum cascade lasers*, J. Appl. Phys. **101**, 086109 (2007)

[101] S. Fathololoumi, E. Dupont, Z. R. Wasilewski, C. W. I. Chan, S. G. Razavipour, S. R. Laframboise, S. Huang, Q. Hu, D. Ban, and H. C. Liu, *Effect of oscillator strength and intermediate resonance on the performance of resonant phonon-based terahertz quantum cascade lasers*, J. Appl. Phys. **113**, 113109 (2013)

[102] A. Wacker, *Extraction-controlled quantum cascade lasers*, Appl. Phys. Lett. **97**, 081105 (2010)

[103] C. W. I. Chan, Q. Hu, and J. L. Reno, *Tall-barrier terahertz quantum cascade lasers*, Applied Physics Letters **103**, 151117 (2013)

[104] T. T. Lin, L. Ying, and H. Hirayama, *Significant reduction of threshold current density of GaAs/AlGaAs terahertz quantum cascade lasers by using high-Al-content AlGaAs barrier*, in *Infrared, Millimeter and Terahertz Waves (IRMMW-THz), 2011 36th International Conference on*, pp. 1–2 (2011)

[105] C. Weber, A. Wacker, and A. Knorr, *Density-matrix theory of the optical dynamics and transport in quantum cascade structures: The role of coherence*, Phys. Rev. B **79**, 165322 (2009)

[106] F. Carosella, C. Ndebeka-Bandou, R. Ferreira, E. Dupont, K. Unterrainer, G. Strasser, A. Wacker, and G. Bastard, *Free-carrier absorption in quantum cascade structures*, Phys. Rev. B **85**, 085310 (2012)

[107] C. Ndebeka-Bandou, M. Rösch, K. Ohtani, M. Beck, and J. Faist, *Negative free carrier absorption in terahertz quantum cascade lasers*, Applied Physics Letters **108**, 091102 (2016)

[108] *Web page of the Star Citizen computer game https://robertsspaceindustries.com/comm-link/transmission/13152-galactic-guide- hurston-dynamics*, last acquired March 2016

Scientific publications

Author contributions

Paper I: Injection Schemes in THz Quantum Cascade Lasers Under Operation

In this proceeding, I did all simulations for the scattering assisted structure in section 4, data analysis, as well as the writing of the manuscript (except for section 3 which was written by David Winge)

Paper II: An indirectly pumped terahertz quantum cascade laser with low injection coupling strength operating above 150 K

I wrote the appendix, did all the NEGF simulations therein, as provided the relevant figures, as well as actively discussed and interpreted the results with the co-authors.

Paper III: Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions

I was involved in writing the whole manuscript, and I wrote the parts regarding the 2-well structure presented in section IV A and the QC detector in section V, for which I also did all the simulations and provided the figures. I also discussed and analyzed - together with the co-authors - all of the presented results.

Paper IV: Comparative analysis of quantum cascade laser modeling based on density matrices and non-equilibrium Green's functions

Being the corresponding author, I wrote the manuscript, provided all the simulation data of the NEGF model, made all the figures, as well as did the data analysis. I was leading and coordinating the cooperation leading to the published results.

Paper V: Impact of interface roughness distributions on the operation of quantum cascade lasers

For this publication, I was driving the work, wrote the main part of the manuscript, and provided the NEGF data for all structures, except the one with resonant tunneling injection, which was simulated by David Winge. He also supplied Fig. 8 and 9 as well as the text for this structure. All the other figures were produced by myself. I also implemented the valence band components in the calculation of interface roughness, as described in Chap. 3.

Paper VI: Influence of Interface Roughness in Quantum Cascade Lasers

For this paper, I supported Kasparas Krivas in performing the NEGF simulations, had a significant contribution to the discussion and analysis of the results and conclusions, as well as participating in producing the final manuscript.

Paper I

Invited Paper

Injection schemes in THz quantum cascade lasers under operation

M. Lindskog, D. O. Winge, and A. Wacker

ABSTRACT

The two main design schemes for Terahertz quantum cascade lasers, based on tunnelling and scattering injection, respectively, are theoretically compared. We apply our simulation package based on the non-equilibrium Green's function technique. Our results provide a good description of the gain degradation with temperature. Thermal backfilling contributes to decrease of population inversion in both cases. However, the dropping inversion cannot account for the total reduction of gain.

1. INTRODUCTION

Terahertz (THz) radiation is an important part of the electro-magnetic spectrum for high-technological applications in the fields of medicine and astronomy, amongst others.[1,2] In order to fully develop these applications, compact, high-power devices are needed and the quantum cascade laser[3] (QCL) is a promising device capable of combining these features. Much effort has been put into improving the temperature performance of these lasers, the best to date working up to 199.5 K,[4] which still requires cryostatic cooling.

The tunnelling injection (TI) design[5] has been very successful, setting new records in temperature performance repeatedly by using a diagonal lasing transition instead of a direct one.[4,6] It has not been able, however, to operate at temperatures with thermal energies significantly above the energy of the lasing transition, why it is thought that there is some fundamental limitation of the performance of QCLs of this design. It has been pointed out[7] that this limitation might come from the temperature broadening of the active QCL states making the tunnel injection to the lower laser state significant compared to the injection to the upper laser state. It is also a well known fact that thermal backfilling into the lower laser level worsens inversion at higher temperatures. Indeed, the TI design has the fundamental limitation that the population inversion cannot exceed 50 % of the total population.[8,9]

For the above given reasons, the indirect pump scheme,[8] or scattering assisted (SA) injection design was proposed, initially in the mid-infra-red, where the injection into the upper laser state is assisted by longitudinal optical (LO) phonon scattering, thus eliminating the problem of tunnel injection into the lower laser level. In addition, the population of the extractor level is supposedly lower in this design, reducing the thermal backfilling into the lower laser level. In the THz regime, SA designs[7,10,11] have operated at thermal energies exceeding the lasing energy by up to almost two times, however still at cryostatic temperatures (a maximum temperature of 163 K has been achieved[7]).

In this work we investigate the two designs described above using the non-equilibrium Green's function (NEGF) method, which gives detailed information about carrier distributions as well as current densities and gain characteristics. The stability of population inversion with respect to bias drop and temperature is studied, as well as the main reasons for the loss of population inversion and gain.

Terahertz Emitters, Receivers, and Applications IV, edited by Manijeh Razeghi, Alexei N. Baranov, John M. Zavada,
Proc. of SPIE Vol. 8846, 884603 · © 2013 SPIE · CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2024030

2. MODEL

Our model was recently described in some detail[12] with the possibility to include higher harmonics of the external ac field. The model takes into account the superlattice potential, an external classical electro-magnetic field (with one dc part and one ac part), and the mean field potential from the carrier and doping densities. In addition, phonon, impurity and interface roughness scattering with electrons are included by the means of self-energies. We here briefly present our model extended to take into account the non-parabolicity of the conduction band within the effective two-band model, in which the effective Hamiltonian is written[13]

$$H = \begin{pmatrix} E_c + e\Phi & \frac{p_{cv}}{m_0}(\hat{p} - e\vec{A}) \\ \frac{p_{vc}}{m_0}(\hat{p} - e\vec{A}) & E_v + e\Phi \end{pmatrix} \tag{1}$$

and the Schrödinger equation

$$H\psi_\alpha = E_\alpha \psi_\alpha, \tag{2}$$

with the two-component wave functions

$$\psi_\alpha = \begin{pmatrix} \psi_c^\alpha \\ \psi_v^\alpha \end{pmatrix} \tag{3}$$

provides the valence band components ψ_v^α in terms of the conduction band components ψ_c^α

$$\psi_v^\alpha = -\hbar \sqrt{\frac{2m_c(E_c, z)}{E_g(z)}} \frac{1}{m_c(E_\alpha, z)} \frac{\partial \psi_c^\alpha(z)}{\partial z} \tag{4}$$

in the absence of the electro-magnetic field. Here, the effective mass is defined as

$$m_c(E, z, t) = -\frac{1}{2} m_0^2 \frac{E - E_v(z)}{|p_{cv}|^2} = m_c(E_c, z) \frac{E - E_v(z)}{E_g(z)}, \tag{5}$$

where $m_c(E_c, z)$ is the effective mass at the conduction band edge. For simplicity, we assume that the lateral effective masses are constant. Thus, we can easily obtain the valence band components from our previously used basis states $\psi_c^\alpha(z)$, to obtain the two-component basis states (3).

The two-component formalism requires new matrix elements for the desired observables and scattering terms to be calculated using the wave functions (3), which is straightforward in most cases. However, the current density can no longer be expressed simply as[12]

$$J(z, t) = e\Re\left\{ \left\langle \frac{\hat{p} - e\vec{A}}{m_c(z)} \right\rangle \right\} \tag{6}$$

since the effective mass is now energy dependent; we have to find the current operator from more basic principles, such as the continuity equation or the time derivative of the position operator, both giving the same result

$$\hat{J} = \frac{1}{A} \sum_k \sum_{mn} \begin{pmatrix} \psi_c^{n*} & \psi_v^{n*} \end{pmatrix} J(z) \begin{pmatrix} \psi_c^m \\ \psi_v^m \end{pmatrix} a_n^\dagger(\mathbf{k}) a_m(\mathbf{k}) \equiv \frac{1}{A} \sum_k \sum_{mn} J_{mn} a_n^\dagger(\mathbf{k}) a_m(\mathbf{k}), \tag{7}$$

where A is the cross-sectional area and

$$J(z) = \frac{e}{m_0} \begin{pmatrix} 0 & p_{cv} \\ p_{vc} & 0 \end{pmatrix}. \tag{8}$$

Interestingly, the current density operator, and consequently its expectation value, no longer explicitly contain the vector potential $\vec{A}$, and the basis states can be chosen to not depend on $\vec{A}$ either.

We model the QCL system by an infinite repetition of one period of the structure. This approximation allows for neglecting the leads as well as calculating the Wannier states of the sub-bands $\psi_\alpha^c(z, t)$ using periodic boundary conditions. The latter gives more reliable eigen-energies than the conventionally used Wannier-Stark

states, if sufficiently many basis states are included, while the former is hardly expected to make any difference for the central periods of THz QCLs, which typically consists of more than 200 periods.

In order to describe the system as physically accurate as possible, we use non-equilibrium Green's functions to express the observables of interest. This method, used by various groups simulating QCL systems,[14–17] keeps the coherences between the states and allows for a consistent treatment of scattering mechanisms and many-body interactions. For instance, one Green's function of interest is the lesser Green's function $G^<$, defined as:

$$G^<_{\alpha,\beta}(\mathbf{k}, t_1, t_2) = i\langle a^\dagger_\beta(\mathbf{k}, t_2) a_\alpha(\mathbf{k}, t_1)\rangle. \tag{9}$$

In order to account for non-linear response to the ac electro-magnetic field, we allow the system to oscillate at the different harmonics of the field frequency Ω. This allows for Fourier expanding in terms of the fundamental frequency Ω, yielding the Fourier transform for all self-energies and Green's functions

$$F(\mathbf{k}, t_1, t_2) = \frac{1}{2\pi}\int dE \sum_h e^{-iE(t_1-t_2)/\hbar} F_h(\mathbf{k}, E) e^{-ih\Omega t_1} \tag{10}$$

Using the above definitions, the expectation value of the current density becomes (with $t_1 = t_2 = t$)

$$\langle \hat{J}(z,t)\rangle = \frac{1}{A}\sum_k \sum_{mn} J_{mn}(z)\langle a^\dagger_n(\mathbf{k}, t) a_m(\mathbf{k}, t)\rangle \tag{11}$$

$$= -i\frac{1}{A}\sum_k \sum_{mn}\int \frac{dE}{2\pi}\sum_h J_{mn}(z) G^<_{mn,h}(\mathbf{k}, E) e^{-ih\Omega t}. \tag{12}$$

Finally, we point out that all temperatures given refer to the phonon distributions in the lattice, while the carrier distribution is in non-equilibrium.

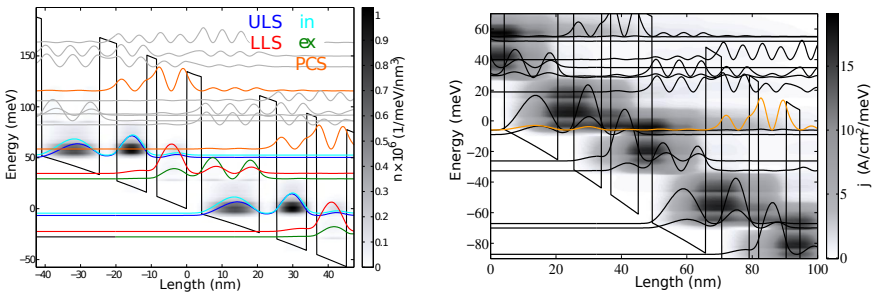
3. RESONANT PHONON DESIGN

The tunnelling injection design has repeatedly broken the terahertz temperature record.[4,6] A resonant tunnelling current serves as a fast way of filling the upper laser state while the lower laser state is placed at an optical phonon energy above the injector state. The robust injection and very fast extraction builds up the inversion at the design bias. One period of the sample of Ref. 6 is shown in Fig. 1(a). The bias is 57 mV/period and it is around this bias that the crossing of the injector and the upper laser state (ULS) occurs. The transition energies of this design are shown in Fig. 2 as functions of applied bias. This figure shows a robust performance over a wide bias range, also seen experimentally[6] where lasing was observed over a 13 mV wide bias range per period. The injector and ULS are clinging together and the extraction energy is close to the optical phonon energy $E_{\mathrm{LO}} = 36.7$ meV in GaAs.

From the current-voltage characteristics displayed in Fig. 3(a), we see that the region of negative differential resistance (NDR) of this design begins at a bias well above the bias of the crossing of the injector and ULS in Fig. 2. This can be understood by two effects. Firstly, despite the level crossing at about 57 mV per period the densities in Fig. 1(a) show that the electrons just right of the injector well still have higher energy than those in the left one at that bias. This effect helps the system keep a positive differential resistivity and can only be seen when resolving the energy distribution of the electrons. Secondly, the aligning of the parasitic current state (PCS) to the injector and ULS also enhances the current density, which can be observed by resolving also the local current density in energy as shown in Fig. 1(b).

Another detail in Fig. 3(a) is the peak at around 37 mV per period. This has no counterpart in measurements made[6] and we currently attribute it to a long range tunnelling effect that might survive due to the lack of electron-electron scattering in our model,[12] though we have no clear cut interpretation of this phenomenon at this time. The same effect might cause an overestimate of the parasitic current from the injector and ULS to the PCS at higher biases.

As temperature increases, electrons gain thermal energy and occupy states of large in-plane momentum k to a higher degree, which results in an increase of the LO phonon emission rate in the ULS as the transition



(a) (b)

Figure 1. (a) Tunnelling injection design of Ref. 6 with the upper laser state (ULS), lower laser state (LLS), injector (in) and extractor (ex) states and a parasitic current state (PCS), as well as the electron densities at a bias of 57 mV per period. (b) Energy resolved local current density at a bias of 61 mV/period, where the PCS is plotted in orange.

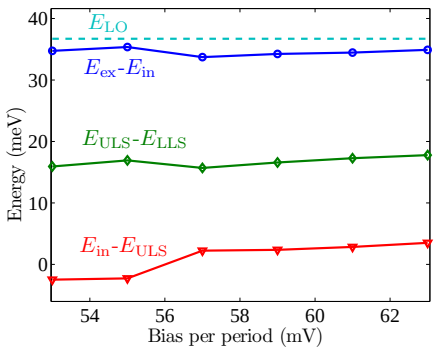
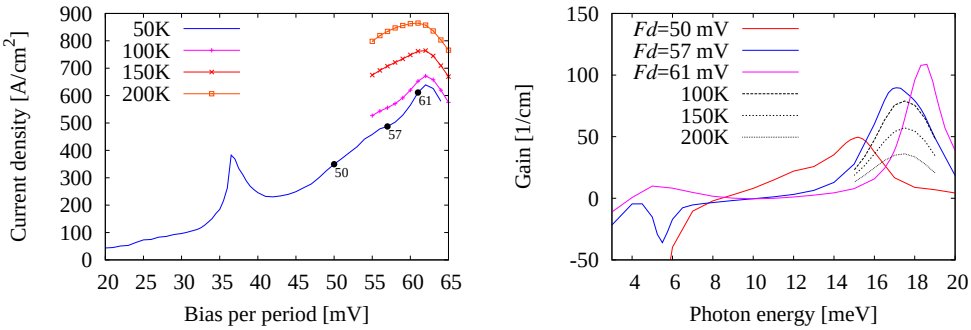


Figure 2. Transition energies for the energy levels shown in Fig. 1(a) of the tunnelling injection design as functions of applied bias. The dashed line shows the LO phonon energy E_{LO} in GaAs.



(a) (b)

Figure 3. (a) Current-voltage characteristics of the tunnelling injection design. The change in peak current density with temperature is displayed. Gain simulations were carried out at the marked bias points for 50 K, see (b). For the 57 mV per period point, gain degradation with temperature is also shown, where the dashed lines represents different temperatures.

energies $E_{\mathrm{ULS}} + E_k - E_{\mathrm{LLS}}$ and $E_{\mathrm{ULS}} + E_k - E_{\mathrm{ex}}$ of the electrons with a significant E_k are closer to E_{LO}. This explains the strong enhancement in current with temperature.

Fig. 3(b) shows the gain spectra at the bias points marked in Fig. 3(a). This shows, in addition to Fig. 2, that the lasing range is wide, since gain sufficiently large to overcome the level of losses (assumed to be significantly below 40 cm^{-1}) can be observed for a bias range of more than 13 mV per period (at 48 mV per period calculations show gain around 40 cm^{-1} for $\hbar\omega = 14$ meV). This is in good agreement with the experimental measurements of a bias range of $\sim$ 13.5 mV/period.[6]

The experimentally observed laser peak is at a photon energy of 16 meV, and a small shift in energy is also seen when going from lower to higher biases. Fig. 3(b) shows a drift of the frequency of peak gain, and at the highest current point it is substantially higher than what was measured. This might indicate that experimentally lasing has already stopped from entering the NDR, while in our simulations this has not happened due to an overestimate of the parasitic leakage current.

Our simulations show that threshold is reached at about 350 A/cm^2 which reasonably matches the experimental results. Peak current is found to be 600 A/cm^2 at 50 K, whereas measurements suggest 850 A/cm^2. However, this high current density is measured under lasing conditions, whereas the current density in Fig. 3(a) is calculated in the off-state of the laser.

As temperature increases we find that current rises and gain drops, as can be seen in Fig. 3(b) for the bias point 57 mV per period. At 200 K lattice temperature, the laser has sufficiently large gain in order to overcome the level of the losses, but this is expected to stop at a somewhat higher temperature. This is consistent with the fact that experimentally, lasing stops at 186 K heat sink temperature.[6] The degradation of inversion with temperature can be seen in Fig. 4 where the densities are plotted against temperature. Assuming a Fermi distribution for the injector state (in) we can explain half of the increase of n_{LLS} by means of thermal backfilling. Furthermore, we see no indication of a parasitic current from the injector to the LLS. A possible additional effect, which might further degrade the performance, is hot phonons out of equilibrium,[18] which is not considered here.

Despite having a drop in population inversion of about 30 % over the temperature range shown in Fig. 4, gain drops by 70 % and so the decrease in population inversion, and thus thermal backfilling, can only account for about half of the reduction of the gain. This effect was also noted in Ref. 19, and more research into the origin of the remaining reduction of gain is necessary for a complete understanding of this phenomenon.

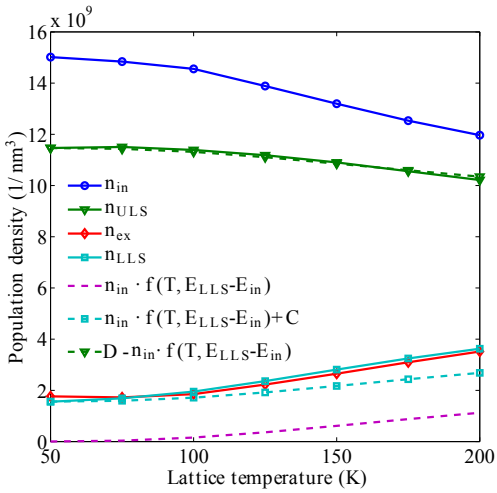


Figure 4. Temperature dependence of the densities in each state shown together with an estimate of the thermal backfilling from the injector and extractor states to the LLS as given by $n_{\text{in}} \cdot f(T, E_{\text{LLS}} - E_{\text{in}})$. The estimate is shifted by a constant C in order to coincide the curves at 200 K for easier comparison. The dashed, diamond-marked curve shows the reduction of n_{ULS} due to the backfilling of n_{LLS}, where D is again a constant used for comparing the curves. The data is taken at 57 mV/period.

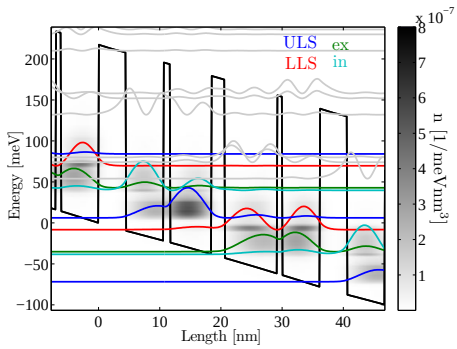


Figure 5. Band structure of the QCL in Ref. 10 as well as the carrier density and the Wannier-Stark states, at a bias of 78 mV/period for $T = 125$ K.

4. SCATTERING ASSISTED DESIGN

The scattering assisted injection scheme[8] uses an additional optical phonon transition in the current path compared to the tunnelling injection scheme, in order to inject carriers into the upper laser level. This allows for a larger bias drop over one period which is supposed to reduce thermal excitations to the lower laser level. However, one would expect there to be twice the occupation of phonon modes in the lattice, providing more lattice heating than for phonon extraction only.

We investigate here the structure, shown together with the carrier densities in Fig. 5, from Ref. 10. There we discussed the importance of taking into account the mean field in the optimization process, since even small

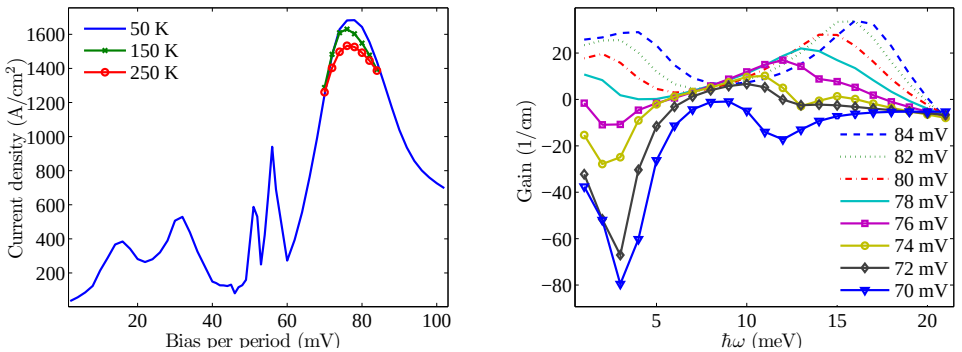


Figure 6. (a) Simulated current density of the SA QCL[10] as a function of applied dc bias per period for three different lattice temperatures and (b) gain profile at different biases for the same structure, for a lattice temperature of $T = 50$ K.

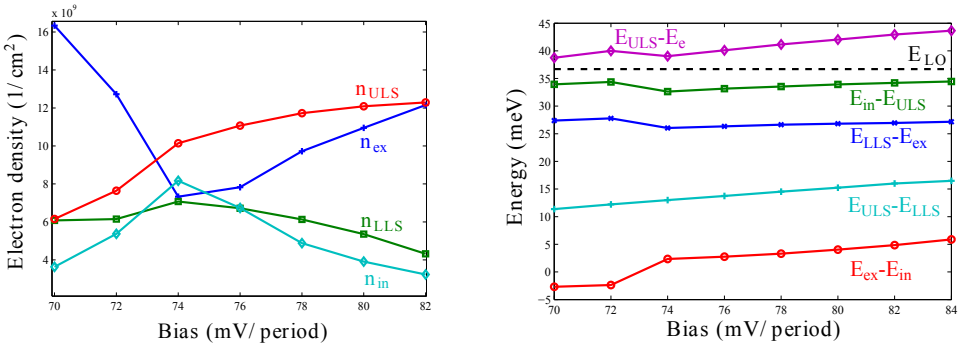
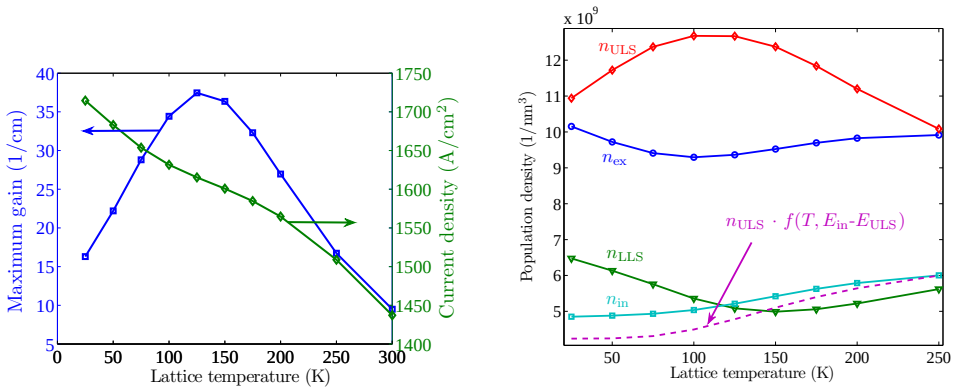


Figure 7. (a) population densities of the extraction (e), injection (i), lower (LLS), and upper laser states (ULS) as well as (b) transition energies of the SA design, as functions of applied bias.

differences in the conduction band profile might put the tunnel resonance at a bias where unwanted effects occur, for instance that the phonon resonances are not satisfied.

Figure 6(a) shows the simulated current density as a function of applied bias per period. For low temperatures, the maximum current ($J_{\max}$) occurs at a bias of 78 mV/period, whereas for slightly higher temperatures it occurs at 74 mV/period, and then slowly shifts back to 78 mV/period at 300 K. There are two smaller current peaks at 16 and 32 mV/period, which arise when the injector and extractor levels align with the lower and upper laser levels respectively. These peaks are also seen in the experiment and our model reproduces them well. In addition there are two very sharp peaks at 51 and 56 mV/period, which occur due to long range tunnelling as the high energy state in the ULS well aligns with the lower and upper laser level respectively. These peaks have not been seen in experiments[10,11] and we currently attribute them to unexplained artefacts of our model. The gain, shown in figure 6(b) for different biases, exceeding the level of losses (assumed to be 20 cm^{-1}) is peaked at 12-14 meV, corresponding to frequencies of 3-3.5 THz, in good agreement with experimental data.[10]

Fig. 7(a) indicates that the tunnel resonance where $n_{\text{in}} \approx n_{\text{ex}}$ occurs around a bias of 74 mV/period. At



(a) (b)

Figure 8. (a) Maximum gain and current density as well as (b) the population densities as functions of the lattice temperature at the bias of J_{max}. $n_{ULS} \cdot f(T, E_{in} - E_{ULS})$ shows the appreciated change in ULS population due to thermal backfilling into the injector state.

70 mV/period, Fig. 7(b) shows that the extraction and injection states are already near resonance and nothing drastic happens with the energy levels as the bias increases. What affects the current flow is rather that both the injection and extraction energies approach E_{LO} as the bias increases above the tunnel resonance, at roughly the same rate as the energy of the leakage transition ULS $\rightarrow$ ex increases away from E_{LO}. This ultimately leads to the peak current density J_{max} occurring at a bias of 78 mV/period, where the tunnelling rate starts to limit the current flow.

Figure 6(b) shows the gain for T=50 K and different applied biases. The gain is increasing with bias, even far into the region of NDR. We also see from figure 7(a) that inversion is increasing over the same range of biases, and that the tunnel resonance occurs at a bias of 74 mV/period. The reason is that the injection and extraction energies become better matched to E_{LO} at higher biases. Another feature of Fig. 6(b) is that there is an additional gain peak at $\hbar\omega \sim 4$ meV, prominent at high biases, coming from the population inversion between the extraction and injection states.

We see that even though a better matching of the biases where the tunnel and phonon resonances occur would increase inversion for the particular QCL studied here, the operation is not very sensitive to the detuning of the tunnel resonance, but depends more on the injection and extraction energies. Even though the NDR begins before the phonon resonances for injection and extraction are reached, this is at a 4 mV/period higher bias than the bias of the tunnel resonance. Designing the heterostructure so that the phonon resonances occur at a slightly higher bias than the tunnel resonance might therefore allow for a reliable way to delay an NDR caused by any parasitic current channel.

Turning our attention to the temperature performance we see from Figs. 8(a) and 8(b) that inversion and gain decreases with temperature above 125 K. Here the gain is lower than for the tunnelling injection structure in accordance with the worse temperature performance in the experiment. On the contrary, gain increases with temperature up to 125 K in our simulation. The reason for this increase can be attributed to the extraction energy $E_{LLS} - E_{ex}$ being about 10 meV lower than E_{LO}. As temperature increases, carriers in the LLS acquire more thermal energy, and at $T = 125$ K the thermal energy of $k_B T = 10.8$ meV has exceeded the activation energy of 10 meV, and the depopulation of the lower laser state becomes more efficient than the relaxation rate from the upper to the lower laser state. This effect can be seen in Fig. 8(b) for n_{LLS} going to a minimum at 150 K and n_{ULS} a maximum at 125 K. On the contrary, the density of the injection state n_i has its minimum at $T = 25 - 50$ K, corresponding to $k_B T \approx 2 - 4$ meV, since the injection energy only differ from E_{LO} by ~ 3 meV. Another effect that can be seen in Fig. 8(b) is the decrease of n_{ex} with temperature, indicating a

diminishing detuning of the tunnel resonance. This detuning comes from the mean field of the carriers in the LLS, and decreases as n_{LLS} decreases. Our model does not include electron-electron scattering, and in a realistic situation this scattering mechanism would likely increase the energy of the carriers in the LLS enough to allow for efficient LO phonon emission, and thus the increase of inversion and gain at low temperatures would be suppressed. At higher temperatures other scattering channels such as acoustic phonons dominate and the lack of electron-electron scattering becomes less important. At temperatures above 125 K the population of the upper laser state n_{ULS} decreases, and the population of the lower laser state n_{LLS} increases. Both the population of the injection and extraction states increase, which means the ULS loose carriers to all these states with increasing lattice temperature (the remaining states have negligible population densities).

Fig. 8(a) also shows that the current density decreases with temperature, in contrast to the TI design, as the carriers get more thermalized and the tunnelling current goes down since less current channels are available. This is also the reason for the shift of J_{max} to lower biases as temperature increases; the tunnel resonance becomes more selective as carriers get more thermalized towards the bottom of the sub-bands. This is quite the opposite to the situation for the TI QCL, where the increase in temperature opens up new current channels and the total current density increases with temperature.

In Fig. 8(b) we see that we have significant thermal backfilling from the ULS to the extraction and injection levels, even though an additional filling effect at low temperatures is necessary to fully account for the increase of n_{in}. The backfilling effect on the LLS from i is less clear due to the decrease of n_{LLS} below $T = 125$ K, however it is certainly present as a contributing factor and both thermal backfilling channels lower the population inversion of the lasing transition. This should be contrasted to the TI QCL, where thermal backfilling is filling the LLS whereas in this case it is mainly emptying the ULS. Even though thermal backfilling seem to be the most important effect preventing the QCL[10] from lasing at high temperatures, again we see that the reduction in population inversion from 125 K to 300 K of ~ 40 % cannot completely account for the reduction of gain over the same temperature range of ~ 70 %.

5. CONCLUSIONS

We have investigated the tunnelling injection and scattering assisted injection schemes for terahertz quantum cascade lasers using a model based on the non-equilibrium Green's function formalism.[12] Correct maximum current and gain peaks with respect to experiments can be reproduced, but for the TI scheme smaller pre-peaks due to long-range tunnelling in the current spectrum are grossly overestimated for reasons that are not yet clear. In contrast, for the SA scheme these pre-peaks are in good agreement with experimental data.

We conclude that thermal backfilling from the extraction to the LLS can explain about half of the decrease in population inversion with temperature for the TI scheme, while for the SA scheme it is backfilling from the ULS to the injection and extraction levels that is responsible for a large part of the decrease instead. We see no indication of other processes such as hot phonon emission or injection tunnelling directly to the LLS to have a large impact on the population inversion for the TI QCL.[6] However, thermal backfilling cannot by itself explain the reduction in gain with temperature, as was also seen in Ref. 19. We leave the question open as to what effect is responsible for further reducing the gain, be it re-absorption of carriers into higher energy levels, thermal broadening of the injection and laser levels, the thermal distributions of the population densities, or something not addressed here. Furthermore, we see that the populations of the TI QCL energy levels are less sensitive to a change in bias than those of the SA QCL, why the latter is more sensitive to placing the tunnel and LO phonon resonances at the same bias. We have also shown that parasitic (for the TI QCL) and LO phonon (for the SA QCL) current channels have a greater impact on the current flow than the tunnel resonances, something that might be utilized to avoid an early NDR.

6. ACKNOWLEDGEMENTS

This work has been partly financed by the EU FP7 project MIRIFISENS and the Swedish Research Council. The authors like to thank Emmanuel Dupont, Ghasem Razavipour and Dayan Ban for helpful discussions.

REFERENCES

[1] Tonouchi, M., "Cutting-edge terahertz technology," *Nat. Photonics* **1**(2), 97–105 (2007).

[2] Williams, B. S., "Terahertz quantum-cascade lasers," *Nature Phot.* **1**, 517 (2007).

[3] Faist, J., Capasso, F., Sivco, D. L., Sirtori, C., Hutchinson, A. L., and Cho, A. Y., "Quantum cascade laser," *Science* **264**(5158), 553–556 (1994).

[4] Fathololoumi, S., Dupont, E., Chan, C., Wasilewski, Z., Laframboise, S., Ban, D., Matyas, A., Jirauschek, C., Hu, Q., and Liu, H. C., "Terahertz quantum cascade lasers operating up to $\sim$ 200 K with optimized oscillator strength and improved injection tunneling," *Opt. Express* **20**(4), 3866–3876 (2012).

[5] Williams, B. S., Kumar, S., Callebaut, H., Hu, Q., and Reno, J. L., "Terahertz quantum-cascade laser operating up to 137 K," *Appl. Phys. Lett.* **83**(25), 5142 (2003).

[6] Kumar, S., Hu, Q., and Reno, J., "186 K operation terahertz quantum-cascade lasers based on a diagonal design," *Appl. Phys. Lett.* **94**, 131105 (2009).

[7] Kumar, S., Chan, C. W. I., Hu, Q., and Reno, J., "A 1.8 THz quantum cascade laser operating significantly above the temperature of $\hbar\omega/k_{\mathrm{B}}$," *Nat. Phys.* **7**, 3866–3876 (2011).

[8] Yamanishi, M., Fujita, K., Edamure, T., and Kan, H., "Indirect pump scheme for quantum cascade lasers: dynamics of electron-transport and very high T_0 values," *Opt. Express* **16**(25), 20748–20758 (2008).

[9] Wacker, A., "Quantum cascade laser: An emerging technology," in [*Nonlinear Laser Dynamics*], Lüdge, K., ed., Wiley-VCH, Berlin (2012).

[10] Dupont, E., Fathololoumi, S., Wasilewski, Z. R., Aers, G., Laframboise, S. R., Lindskog, M., Razavipour, S. G., Wacker, A., Ban, D., and Liu, H. C., "A phonon scattering assisted injection and extraction based terahertz quantum cascade laser," *J. Appl. Phys.* **111**, 073111 (2012).

[11] Razavipour, S. G., Dupont, E., Fathololoumi, S., Chan, C. W. I., Lindskog, M., Wasilewski, Z. R., G.Aers, Laframboise, S. R., Wacker, A., Hu, Q., Ban, D., and Liu, H. C., "An indirectly pumped terahertz quantum cascade laser with low injection coupling strength operating above 150 K," *J. Appl. Phys.* **113**, 203107 (2013).

[12] Wacker, A., Lindskog, M., and Winge, D. O., "Nonequilibrium Green's function model for simulation of quantum cascade laser devices under operating conditions," *IEEE J. Sel. Top. Quantum Electron.* **113**, 203107 (2013).

[13] Sirtori, C., Capasso, F., Faist, J., and Scandolo, S., "Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells," *Phys. Rev. B* **50**, 8663 (Sep 1994).

[14] Lee, S.-C. and Wacker, A., "Nonequilibrium Green's function theory for transport and gain properties of quantum cascade structures," *Phys. Rev. B* **66**, 245314 (2002).

[15] Schmielau, T. and Pereira, M., "Nonequilibrium many body theory for quantum transport in terahertz quantum cascade lasers," *Appl. Phys. Lett.* **95**, 231111 (2009).

[16] Kubis, T., Yeh, C., Vogl, P., Benz, A., Fasching, G., and Deutsch, C., "Theory of nonequilibrium quantum transport and energy dissipation in terahertz quantum cascade lasers," *Pys. Rev. B* **79**, 195323 (2009).

[17] Haldaś and, G., Kolek, A., and Tralle, I., "Modeling of mid-infrared quantum cascade laser by means of nonequilibrium Green's functions," *IEEE J. Quantum Electron.* **47**(6), 878 (2011).

[18] Iotti, R. C., Rossi, F., Vitiello, M. S., Scamarcio, G., Mahler, L., and Tredicucci, A., "Impact of nonequilibrium phonons on the electron dynamics in terahertz quantum cascade lasers," *Applied Physics Letters* **97**(3), 033110 (2010).

[19] Nelander, R. and Wacker, A., "Temperature dependence of the gain profile for terahertz quantum cascade lasers," *Applied Physics Letters* **92**(8), 081102 (2008).

Paper II

JOURNAL OF APPLIED PHYSICS **113**, 203107 (2013)

An indirectly pumped terahertz quantum cascade laser with low injection coupling strength operating above 150 K

S. G. Razavipour,[2,a)] E. Dupont,[1,b)] S. Fathololoumi,[1,2] C. W. I. Chan,[3] M. Lindskog,[4] Z. R. Wasilewski,[1,2] G. Aers,[1] S. R. Laframboise,[1] A. Wacker,[4] Q. Hu,[3] D. Ban,[2] and H. C. Liu[5,c)]

[1]*National Research Council, Ottawa, Ontario K1A0R6, Canada*
[2]*Department of Electrical and Computer Engineering, Waterloo Institute for Nanotechnology, University of Waterloo, 200 University Ave W., Waterloo, Ontario N2L3G1, Canada*
[3]*Department of Electrical Engineering and Computer Science, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[4]*Division of Mathematical Physics, Lund University, Box 118, Lund 22100, Sweden*
[5]*Key Laboratory of Artificial Structures and Quantum Control, Department of Physics, Shanghai Jiao Tong University, Shanghai 200240, China*

(Received 19 March 2013; accepted 8 May 2013; published online 29 May 2013)

We designed and demonstrated a terahertz quantum cascade laser based on indirect pump injection to the upper lasing state and phonon scattering extraction from the lower lasing state. By employing a rate equation formalism and a genetic algorithm, an optimized active region design with four-well $GaAs/Al_{0.25}Ga_{0.75}As$ cascade module was obtained and epitaxially grown. A figure of merit which is defined as the ratio of modal gain versus injection current was maximized at 150 K. A fabricated device with a Au metal-metal waveguide and a top n^+ GaAs contact layer lased at 2.4 THz up to 128.5 K, while another one without the top n^+ GaAs lased up to 152.5 K ($1.3\hbar\omega/k_B$). The experimental results have been analyzed with rate equation and nonequilibrium Green's function models. A high population inversion is achieved at high temperature using a small oscillator strength of 0.28, while its combination with the low injection coupling strength of 0.85 meV results in a low current. The carefully engineered wavefunctions enhance the quantum efficiency of the device and therefore improve the output optical power even with an unusually low injection coupling strength. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4807580]

I. INTRODUCTION

Terahertz (THz) quantum cascade lasers (QCLs), one of the most important sources of coherent THz radiation, can cover a spectral range from ~1.2 to ~5.2 THz.[1–3] Since their invention,[4] even though many efforts have been made to improve the performance of THz QCLs in terms of operating temperature, it is not high enough to work under thermoelectric cooling systems. Recently, a design based on a three-well resonant tunneling (RT) structure, implemented in the $GaAs/Al_{0.15}Ga_{0.85}As$ material system, has reached a maximum operating temperature T_{max} of 199.5 K.[5] The limitations of RT-QCLs were addressed by Yasuda *et al.*,[6] Kubis *et al.*,[7] Kumar *et al.*,[8] and Dupont *et al.*[9] Many carriers are stationed in the injector state, ready to be transferred via resonant tunneling to the long-lived, hence heavily populated, upper lasing state (ULS). In this configuration, the bidirectional nature of resonant tunneling limits the maximum possible population inversion of RT-QCLs to 50%.[10] An efficient injector barrier must be thick enough to suppress wrong carrier injection to the lower lasing state (LLS) or other states lower than ULS and to prevent early negative differential resistance (NDR). On the other hand, it should be thin enough to reduce the tunneling time and increase the maximum current, thereby the dynamic range of the laser. The constraint on the injector barrier becomes even worse when the device lasing frequency approaches 2 THz which corresponds to a photon energy of less than 8.5 meV.[11] All the aforementioned RT-QCL issues impel designers to find novel approaches to overcome the bottlenecks of THz RT-QCL.

The majority of high-performance devices are based on RT-QCL structures.[5,12–14] In contrast, there is the indirectly pumped (IDP) scheme, well implemented in mid-IR QCL,[10] and a promising approach to improve the performance of THz QCL,[6] especially for low lasing frequencies. To date, several designs based on IDP scheme have been theoretically presented and avowed to have enough gain at higher temperatures to improve the temperature performance and overcome the fundamental limitation of designs based on RT structures.[6,15,16] Three groups have demonstrated THz structures based on the IDP scheme,[8,9,17,18] and the best performance THz QCL in terms of $k_B T_{max}/\hbar\omega$ was achieved in the $GaAs/Al_{0.15}Ga_{0.85}As$ material system by Kumar *et al.*[8]

The structures in the $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ material system, presented by Yamanishi *et al.*,[17–19] while not showing the highest operating temperature, exhibited advantages of a high peak output power and smooth current density-voltage (J-V) characteristics with no tunneling resonance before the designed electric field. However, the light-current

a)Electronic mail: sgrazavi@uwaterloo.ca.
b)Electronic mail: emmanuel.dupont@nrc-cnrc.gc.ca
c)Electronic mail: h.c.liu@sjtu.edu.cn

density characteristics (L-J) showed an optical power roll-over that reduces the temperature performance of the devices. It was proposed that the optical roll-over in power, might come from (i), the excess energy effect of the hot carriers in the injection region and (ii), fast tunneling rate to the next module that could frustrate the thermalization of carriers in the injector. This roll-over effect was not observable in the next generation of IDP structure with an extended tunneling time, which supports this excess energy hypothesis.[19]

Different from all aforementioned IDP designs, we are presenting an IDP four-well structure with an unusually low injection coupling strength. This new design demonstrates better performance than that of Ref. 9 in terms of the output power, threshold current density, resonance before the designed electric field, and $k_B T_{\max}/\hbar\omega$. Among all reported approaches for modeling the charge transport in THz QCL,[20–30] a simplified rate equation (RE) model, which is easy to implement and fast to compute, is employed to predict the carrier transport and estimate the optical gain in this paper. A nonequilibrium Green's function (NEGF) approach, presented in the Appendix, was used to confirm and complement the RE results.

This reported IDP four-well THz QCL is based on phonon-photon-phonon (3P) successive emissions within two neighboring thick injection barriers. Figure 1 shows the schematic energy level diagram of a 3P-QCL structure. Since both carrier injection into the ULS and extraction from the LLS are mediated by LO-phonon scattering, and since no tunneling is involved in carrier transport within a module, an efficient injection and extraction are crucial for this design, which are achieved here by wavefunction engineering through a genetic algorithm (GA). This control of wavefunctions is different from energy-band engineering and gives us an opportunity to control the desired and undesired scattering

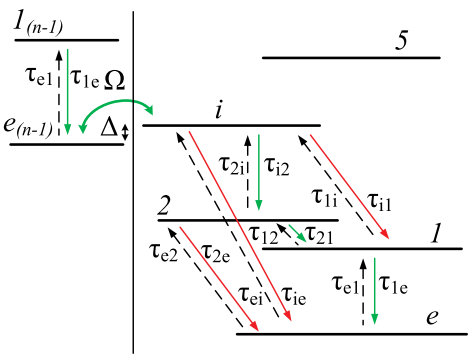


FIG. 1. Schematic diagram of a scattering-assisted QCL active region based on a phonon-photon-phonon configuration. Throughout this paper and whatever the electric field, the states within a module are labeled in energy ascending order *e*, *1*, *2*, *i*, and *5*. The solid lines show the forward scatterings, while the dashed lines indicate the back scatterings. Δ and Ω are the detuning and the coupling between states *i* and *e*, respectively. The green lines indicate the correct injection and extraction, while the red lines show the wrong injection and extraction in each module. *2* and *1* are the ULS and LLS, respectively.

rates through the shapes of wavefunctions and their relative overlaps.[31] The RE model presented in Ref. 9 is employed again during the design stage with a modified figure of merit. In Ref. 9 in order to maximize the gain in a module, a figure of merit was defined as the product of population inversion, oscillator strength, and the inverse of the superperiod length and resulted in the design of a QCL structure named V843. The design details of another THz 3P-QCL (second generation), including its specific figure of merit, will be discussed in the next section.

II. DEVICE DESIGN

A. Wavefunction engineering challenges

To understand how this scheme works, we can assume that the four-well structure consists of two phonon double wells separated by a radiative barrier. The more spatially overlapped wavefunctions in each double well lead to a faster phonon scattering process. In addition, the energy spacing of each double well should be as close as possible to the LO-phonon energy of GaAs (36.7 meV). The thickness of the radiative barrier strongly affects the oscillator strength, which, in turn, directly affects the gain and the population inversion. Even though there is no tunneling for carrier injection to the ULS and carrier depopulation from the LLS, the wrong injection and extraction channels still exist. To overcome this problem, those undesired scattering rates must be minimized to decrease the chance of the wrong injection to the LLS or non-radiative relaxation from the ULS. That is, the scattering times (red arrows in Fig. 1) must be increased. From the point of view of population inversion, the optimum structure of this scheme is a design that has a short τ_{i2} and τ_{1e} to maximize the correct injection and extraction, and second, a long τ_{i1} and τ_{2e} to minimize the wrong injection and extraction, respectively. In addition, it needs to have a fairly long relaxation time between the lasing states to keep the population inversion high enough even at higher temperatures. Therefore, in 3P structures, all four eigenenergies and their corresponding wavefunctions have to be carefully and simultaneously tailored to efficiently inject carriers into ULS and extract them from LLS. To satisfy those requirements, a GA was employed to optimize the design candidates.

Two minor issues arise when the GA is employed to find the optimum structure.

1. The fifth energy state, which was not an issue in the RT structure, may play an adverse role in 3P designs. In a conventional THz RT-QCL, the total potential across one module, which equals the sum of a THz photon energy and a LO-phonon energy, is typically less than 57 meV (considering 36 meV for LO-phonon energy and maximum 21 meV for THz photon energy[3]). In a THz 3P-QCL, this energy spacing increases to almost 90 meV so as to put the fifth energy state closer to the ULS and LLS of the previous (upstream) module at an electric field lower than the designed electric filed. This situation can substantially enhance the leakage current through tunneling to this state and decrease the population inversion. In

addition, this state can be a reason for early NDR if the injector barrier is thinned too much in order to reach a high maximum current. Moreover, as the injector state cannot be totally depleted, the dipole moment between levels *i* and *5* should be kept low in order to minimize the free-carrier absorption;[32,33] this is typically achieved by a thick radiative barrier.

2. The second issue that must be noted is the effect of tunneling between level *e* of the left module and levels *2* or *1* of the right one. If this tunneling is stronger than that of levels *e* and *i* at the desired threshold electric field, we may confront an early NDR, which could block the lasing operation of the structure or cause voltage instabilities.[34] It has been demonstrated that an IDP structure with a two-well injector can substantially reduce all resonances prior to the threshold.[18]

B. Figure of merit

Considering the aforementioned challenges, in the second generation of 3P-QCL, we targeted to maximize the ratio of the gain versus injection current at a lattice temperature of 150 K. The product of the modal gain, transit time, and inverse of the superperiod length was thus defined as the figure of merit, the details of which are provided in Ref. 35. A GA was employed to optimize the design candidates. The three-dimensional doping concentration, the injector barrier, the desired electric field, and the material ($GaAs/Al_{0.25}Ga_{0.75}As$) were fixed while the quantum well and barrier widths were free to change. The electron temperature was assumed to be 50 K higher than the lattice temperature for all subbands. Even though only the LO-phonon scattering was included in the GA process, the electron-LO-phonon, electron-impurity, and interface roughness (IR) intersubband scatterings are considered in all simulation results presented in this paper. As shown in Fig. 1, both forward and backward scattering channels were computed in our design. The quantum wells and barriers of the converged solution from the GA program are (starting with the injector barrier): **44**/64.5/**16.2**/71.5/**27.9**/104.45/**6**/49.65 Å, where the bold font indicates the barriers. Fixing the 3D doping of the structure and obtaining the quantum wells and barriers widths determine the superlattice length and the two-dimensional doping concentration. The first well after the injection barrier was delta-doped with Si to $n_s = 3.45 \times 10^{10}\,cm^{-2}$ near the center. This structure was grown by molecular beam epitaxy as a wafer named V845.

The energy spacings of the first four energy states contributing in our rate equation model are 36.6, 13.9, and 30.3 meV at 21 kV/cm, respectively. The oscillator strength and the injection coupling strength of this structure are lower than those of V843. Such a low tunnel coupling ($\hbar\Omega = 0.85$ meV at 21 kV/cm) makes the transport through the injection barrier incoherent and limits the maximum current in the lasing regime. This low value of tunnel coupling strength is attributed to the specifically defined figure of merit, whose target is to maximize the modal gain over the current ratio. Both the threshold current and the maximum current are lower than those of V843. The energy spacing of 36.6 meV between levels *1* and *e* results in a faster relaxation time of 0.21 ps at 150 K compared to 0.41 ps in V843. A longer scattering time between the lasing states (*2* and *1*) can increase the population inversion at higher temperatures. This result is achieved by the spatially separated wavefunctions of the lasing states of the structure. The conduction band diagram and squared wavefunctions of the corresponding energy states in two adjacent modules at 21 kV/cm are shown in Fig. 2. Since we have defined a specific figure of merit to find a structure with higher gain at lower current and also forced the algorithm to set the injector barrier thickness at 44 Å, the two minor issues, presented in Sec. II A, may not affect our optimization process. If we decrease the thickness of the injector barrier and enhance the coupling strength between levels *e* and *i* to reach a high current dynamic range and operating temperature, the effect of the fifth energy state and tunneling before the threshold should be considered. In this paper, both four and five-level RE were implemented and the simulation results were almost the same, suggesting the 5th state of this design has a marginal impact on transport.

C. Rate equation modeling assumptions

In this section, we present the assumptions of the RE model used for the calculations of the coherence term,

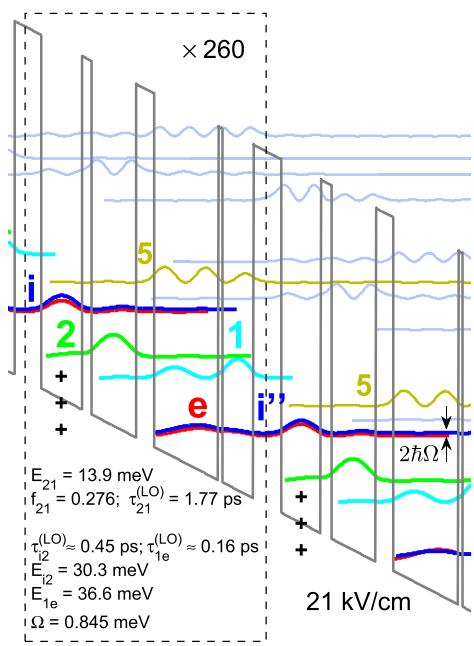


FIG. 2. Conduction band diagram and the moduli squared of wavefunctions of the THz 3P-QCL, V845, at 21 kV/cm. The "+" signs denote the position of Si doping in each module. The intersubband lifetimes by LO-phonon emission are given at the resonant in-plane kinetic energy.

tunneling time, dephasing rate during tunneling and optical transition. The RE model is based on the density matrix formalism and it is computationally very efficient. However, unlike a full density matrix model, the RE model greatly simplifies the expression of the coherence term between two tunneling states, ρ_{ab}, by involving only the coupling strength $\hbar\Omega_{ab}$, detuning energy $\hbar\Delta_{ab}$, dephasing time $\tau_{\parallel ab}$, and populations (ρ_{aa} and ρ_{bb}) of this particular pair of states a and b

$$(\Delta_{ab} - \jmath\tau_{\parallel ab}^{-1})\rho_{ab} = \Omega_{ab}(\rho_{aa} - \rho_{bb}). \quad (1)$$

The model ignores that the coherences can be interlinked: for instance, when two states b and c are coupled to the same state a by tunneling, a coherence term ρ_{bc} develops and also intervenes in the expression of ρ_{ab}. This simplification ignores the indirect resonances between states that are two modules apart.[34] Our RE model for transport is a simplified version of the one described in chapter 7 of Ref. 36 in the sense it is based on first-order resonant tunneling and it does not solve self-consistently the electronic temperature and the coupled Schrödinger-Poisson equations.

All tunneling channels between two neighbor modules were calculated based on a first-order approximation and included in our RE model. In the first-order approximation, the tunneling time between the extractor and injection states is defined as a Lorentzian function of the detuning energy, $\tau_{\text{tun}} = (1 + \Delta^2\tau_{\parallel ei}^2)/2\Omega^2\tau_{\parallel ei}$, where $\hbar\Delta$ is the energy detuning between the two extractor and injection states. The tunneling rate depends sensitively on the dephasing time, easily obtained by $\tau_{\parallel ei} = (\frac{1}{\tau^*} + \frac{1}{2\tau_e} + \frac{1}{2\tau_i})^{-1}$. It consists of lifetime terms ($\frac{1}{2\tau_e} + \frac{1}{2\tau_i}$) due to intersubband scattering and a component commonly named the "pure" dephasing time ($\frac{1}{\tau^*}$) due to intrasubband scattering. By calculating each individual broadening that comes from different scattering mechanisms, the total broadening (Γ_{tun}) will be determined by having $\Gamma_{\text{inter}}^{e(i)} = \sum_m \Gamma_{\text{inter}}^{m-e(i)}$ and $\Gamma_{\text{intra}}^{ei} = \sum_m \Gamma_{\text{intra}}^{m-ei}$, where m denotes a specific scattering mechanism, and using $\Gamma_{\text{tun}}^{ei} = \frac{1}{2}\Gamma_{\text{intra}}^{ei} + \frac{1}{2}(\Gamma_{\text{inter}}^{e} + \Gamma_{\text{inter}}^{i})$.[37] A pure dephasing time was obtained by calculating the intrasubband scattering between levels e and i at an electric field right before the two states were aligned ($\tau^* = 2\hbar/\Gamma_{\text{intra}}^{ei}$). A constant pure dephasing time of 0.2 ps was employed for all temperatures to calculate the tunneling time between e and i. This value of 0.2 ps was estimated by including the intrasubband interface roughness ($\hbar/\Gamma_{\text{intra}}^{\text{IR}-ei} \sim 0.25$ ps) and impurity ($\hbar/\Gamma_{\text{intra}}^{\text{ION}-ei} \sim 0.18$ ps) scatterings in our model and using the aforementioned equations.

In this paper, we did not model the bandwidth, $\Delta\nu$, of the gain. The peak gain of the design was simply written as

$$g_{\text{peak}} = \frac{q^2 f_{21}}{2m^*\epsilon_0 n_r c} \times \Delta N_{21}^{3D} \times \frac{C}{\Delta\nu}, \quad (2)$$

where C is a normalization constant of the gain profile, i.e., $1/\pi$ for a Lorentzian (assumed in most cases) or $\sqrt{\ln 2/\pi}$ for a Gaussian, f_{21} is the oscillator strength between the lasing states, ΔN_{21}^{3D} the population inversion 3-D density averaged over one superperiod, m^* is the effective mass in GaAs. Since we decided to keep the bandwidth as an unknown parameter, in reality throughout this work, we only modeled the product between peak gain and bandwidth, $g_{\text{peak}} \times \Delta\nu$ (gain-bandwidth product).

D. Evaluation of selected design

To investigate the effectiveness of our wavefunction engineering, the scattering times between the five states in one module, included in our RE model, and the tunneling rate between two modules, should be calculated. The following parameters were calculated: the scattering times among the first five states in one module, the population of each state, the current density, and the gain-bandwidth product based on the rate equation model at each electric field, starting from 0.5 kV/cm to 23 kV/cm at different operating temperatures. Figure 3 shows the simulation results at two electron temperatures: 70 K (lattice temperature = 20 K) and 200 K (lattice temperature = 150 K) versus electric field.

The most relevant time constants are shown in Fig. 3(a) for $T_e = 70$ K, where τ_i is the lifetime of the injection state and τ_{tun} is the tunneling time between level e and i. The variables $\tilde{\tau}_{2\text{eff}}$, $\tau_{\text{tr}}^{<}$, and $\tau_{\text{tr}}^{<}$ are the *modified* effective lifetime,

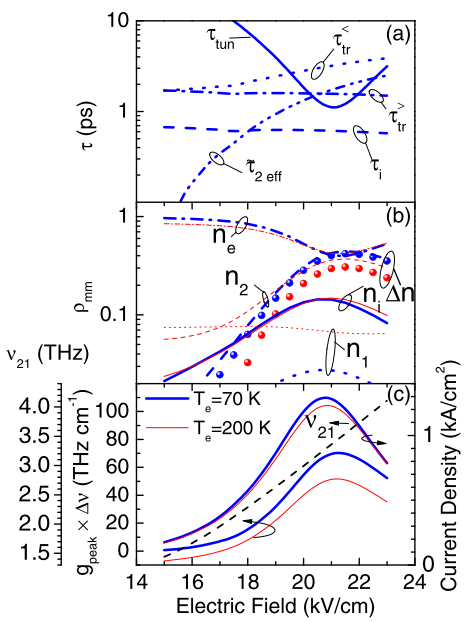


FIG. 3. The 4-level RE simulation results of the structure presented in Fig. 2. (a) Different characteristic times at 20 K ($T_e = 70$ K, thick blue lines). The scattering time presented in figure are defined as follows: τ_{tun} is tunneling time (solid line), $\tau_{\text{tr}}^{<}$ (dotted line), and $\tau_{\text{tr}}^{>}$ (dashed dotted line) are the transit times—excluding the tunneling time—across the four wells before and after threshold, respectively; τ_i is injection state lifetime (dashed line); and $\tilde{\tau}_{2\text{eff}}$ is the *modified* effective lifetime (dashed dotted dotted line). (b) Normalized populations of the four states at 20 K (thick blue lines) and 150 K (thin red lines) lattice temperatures and the population inversion ($n_2 - n_1$) at 20 K (blue solid circles) and 150 K (red solid circles), (c) Current density, lasing frequency (dashed line), and optical gain-bandwidth product vs electric field at 20 and 150 K lattice temperatures.

transit time—excluding the tunneling time—through the four quantum wells, *without* stimulated emission, and transit time—excluding the tunneling time—through the four quantum wells, *with* stimulated emission. The last three parameters are defined in Eqs. (2), (3), and (8) of Ref. 9, respectively. Since the injector barrier is thick (44 Å), and due to the specifically defined figure of merit, the coupling between the wavefunctions of level e and i is small; the tunneling between these states is incoherent. This incoherency in transport could result in carrier accumulation at level e, increasing the backfilling, specially at higher temperatures. The faster scattering from the injector state i (τ_i), compared to the tunneling time τ_{tun} shown in Fig. 3(a), under various electric fields suggests the population on the extractor state will be significantly higher than that of the injector state. $\tau_{tr}^{<}$ and τ_{tun} are longer in V845 than in V843, which is due to higher diagonality of the structure and smaller coupling strength. Therefore, the population on the extractor state, n_e, does not change very much between the two structures as $n_e \sim \tau_{tun}/\tau_{tr}^{<}$. Even though Ω is quite small, the device is not penalized dramatically because the transport by carrier hopping between two 44 Å injection barriers (i.e., when excluding tunneling) has also been slowed down by the strong diagonality. Since by design, the lifetime of the injection state, τ_i, is short, and the transit time without stimulated emission is rather long, there is no need to reach coherent transport through the injection barrier, i.e., $\tau_{tun} \ll \tau_i$. However, we would recommend to have τ_{tun} comparable to τ_i, i.e., $\tau_{tun} \gtrsim \tau_i$, to lower significantly the population on the extractor state, which can be achieved by increasing the tunnel coupling strength. At 21 kV/cm, the energy spacing between states i and 2 (E_{i2}) in V845 is 6.4 meV below the GaAs phonon energy, as a result the injection of carriers on ULS is slowed down: $\tau_{i2} \sim 0.56$ ps vs 0.33 ps for V843 at 150 K. This increase in injector state lifetime explains why n_i is even worse (i.e., larger) in V845 as $n_i \sim \tau_i/\tau_{tr}^{<}$. The current at e-i alignment, calculated as the product of $n_e - n_i$ and the inverse of τ_{tun} will be lower than that of V843 due to lower $n_e - n_i$ and higher τ_{tun} simultaneously.

Figure 3(b) shows the carrier density of each state at different electric fields. As expected, the carrier density at level e, n_e, is dominant almost over the entire bias range. This shows that most of the carriers are piled up at level e, even near the designed electric field of 21 kV/cm. The same effect of carrier accumulation on level e is predicted by NEGF simulations (see Appendix). Consequently, due to the backfilling from level e, the density of carriers at level 1 (LLS) increases dramatically when the temperature increases from 20 K to 150 K. The population inversion (solid circles) decreases when the temperature varies from 20 K to 150 K but it is still higher than that presented in Ref. 9 due to a longer *modified* effective lifetime $\tilde{\tau}_{2eff}$. At lower temperatures since the backward scattering is not fast, level 1 is almost empty.

The gain-bandwidth product, current density, and lasing frequency are shown in Fig. 3(c). The value of the gain-bandwidth product changes modestly, from 20 K to 150 K. The maximum value of the gain-bandwidth product is 60.4 THz cm^{-1} at 20 K, while it decreases to 48.1 THz cm^{-1} and 43.1 THz cm^{-1} at temperatures of 125 K and 150 K, respectively. The backfilling to the level 1 at higher temperatures, due to the piling-up at level e, is the main reason for the population inversion reduction and hence gain reduction. Even though the tunneling time τ_{tun} is exactly minimized at 21 kV/cm, i.e., at the electric field when e-i are perfectly aligned, the current density is peaked at ~20.7 kV/cm due to a decline of $n_e - n_i$ close to the resonance. Figure 3(c) shows that the lasing frequency will vary from 2.8 THz near the threshold to 3.2 THz at the electric field near the NDR by assuming the product of the gain bandwidth ($\Delta\nu$) and the cavity loss (α_{cav}) is $\Delta\nu \times \alpha_{cav} \sim 42$ THz cm^{-1}, which will be discussed in Sec. IV E.

III. EXPERIMENTAL RESULTS

The whole QCL structure V845 consists of 260 repeats of the module presented in Fig. 2 and is grown on a semi-insulating GaAs substrate using molecular beam epitaxy with a total thickness of 10 μm. The active region is sandwiched between a 100 nm of 3×10^{18} cm^{-3} bottom n$^+$ GaAs and a top stack of 40 nm of 7×10^{17} cm^{-3}, 50 nm of 5×10^{18} cm^{-3}, 10 nm of 5×10^{19} cm^{-3} n$^+$ and 3.5 nm of low-temperature (LT) grown GaAs. The first layer of the top stack is meant to adjust the Fermi level so as to align with the injector state of the first module, and the last two layers are used to form a non-alloyed ohmic contact.

We should point out that this V845 wafer was grown using a nearly depleted Ga cell. Indeed, the cell ran out of gallium during the flux measurement procedure conducted on the next day. Since the evaporation in such a situation is often taking place from several remaining droplets of gallium, the evaporation surface area can change in a random fashion. Thus, despite attempt to stabilize the flux by appropriate ramp of the cell temperature during the growth, the average Ga flux decreased by more than 2% during the active region deposition, while the analysis of X-ray Diffraction (XRD) data revealed step-like changes in the Ga flux which resulted in three distinct regions of different periodicity: the main region with 1.4% shorter period than the target value, and the other two regions with about 30 repeats each in which the periods were 0.4% shorter and 0.6% longer than the target value.

Two different fabrication processes were employed to compare the effect of waveguide loss on device performance of this structure. Since the oscillator strength of this 3P structure is lower than that of V843, the maximum operating temperature may be enhanced more substantially by lowering waveguide loss.[34,38] The first fabrication process, device A, used a Au-Au ridge waveguide with a 144 μm width and 1.07 mm length, while the second fabrication process, device B, has a wider (159 μm) and longer (1.79 mm) waveguide. In addition, the 100 nm thick top contact layer was removed to lower the waveguide loss. The Ti/Au metalization process and In-Au bonding technique were employed for device A, while device B was fabricated using a Ta/Au metallization process and an Au-Au bonding process.

Figure 4(a) shows the pulsed light-current density-voltage (L-J-V) characteristics of device A from 10 K to 128.5 K, with a pulse duration of 250 ns and repetition rate

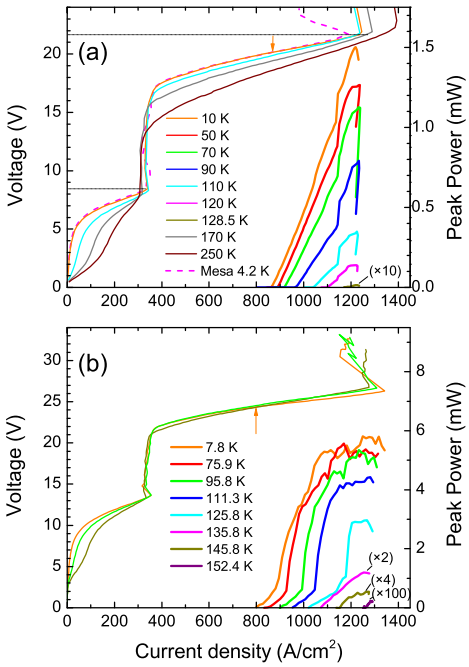


FIG. 4. Left axis: The bias voltage of THz 3 P-QCL V845 versus the current density, (a) device A (b) device B. The short vertical arrows show the change in the slope of the V-J curves at laser threshold and the lowest temperature (10 K for device A or 7.8 K for device B). Right axis: Collected THz light (optical output power) versus current density at different heat sink temperatures. Since the measurement set-up and the waveguide properties are different, the collected light, the maximum current density, and the threshold current are different in plots (a) and (b). Drop voltage on device B is higher than on device A, the latter having the top 100 nm n^+ contact GaAs layer hence, a top Schottky contact with a short depleted region (~18 nm).

of 1 kHz. The threshold current density of 0.87 (0.8) kA/cm^2 was measured for device A (device B) package, while the maximum current density was 1.25 (1.34) kA/cm^2. The lower cavity loss results in a lower threshold current density (i.e., 0.8 kA/cm^2 at 7.8 K for device B vs 0.87 kA/cm^2 for device A at 10 K) and a higher maximum current density which enhances the dynamic range of device B and improves the maximum operating temperature. The maximum operating temperatures of 128.5 K and 152.5 K were achieved with devices A and B, respectively. The maximum collected optical power in devices A and B was 1.5 mW and 5.8 mW at lattice temperature of 10 K and 7.8 K, respectively (The optical set up and the injected electric power were different in devices A and B).

The solid orange V-J curve (device A) in Fig. 4(a) shows the first NDR at 8.5 V and the final NDR at 21.8 V. By considering a 0.75 V Schottky drop voltage from the top contact,[39] we will reach 21.05 V as the main NDR of the design, which nicely matches with our design electric field (21 kV/cm). The first NDR at 8.5 V (7.75 kV/cm) comes

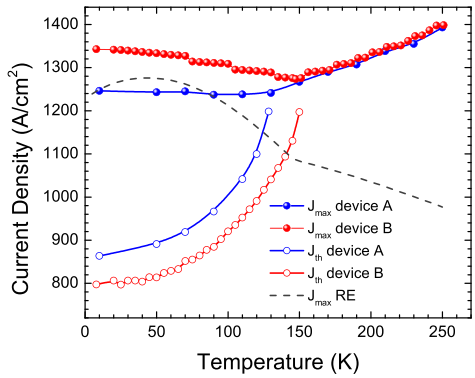


FIG. 5. Maximum current density and threshold current density as functions of heat sink temperature for devices A and B. The dashed line shows the result of a 5-level rate equation simulation assuming a constant product $\Delta\nu \times \alpha_{cav} = 42\,\mathrm{THz\,cm^{-1}}$.

from the resonance tunneling between levels *e* and *1*. The Schottky drop voltage of top contact can be accurately obtained by comparing the V-J measurements of the lasing and non-lasing devices. The width and length of the non-lasing device were reduced to 248 μm and 174 μm, respectively, to increase the loss of the cavity, hence suppress the stimulated radiation. In addition, the non-lasing device was annealed to form PdGeTiPtAu ohmic contacts, which had the advantage—for this particular application—to increase the waveguide loss.[39] Besides, these ohmic contacts reduce the potential drop across the metal/semiconductor interfaces and help to achieve accurate V-J curve measurements. The dashed-magenta curve in Fig. 4(a), showing the V-J characteristic of the non-lasing device, is also plotted, and the difference between the V-J curves of the lasing and non-lasing devices becomes distinguishable only after the threshold voltage indicated by a vertical arrow. The dashed-magenta curve was shifted upward along the voltage axis by a Schottky drop voltage of 0.75 V so as to overlap with the lasing V-J curve of the device A. At the NDR, the difference in current density between the lasing ridge and the non-lasing mesa is small, only ~60 A/cm^2. We also note that, above 130 K, a small resonance in the J-V characteristics develops slightly above 2 V. As shown in Fig. 5, the maximum current density (J_{max}) of both devices A and B slightly decreases from 8 K to a temperature close to T_{max}. This reduction of current density is 10 A/cm^2 from 10 K to 110 K in device A and 70 A/cm^2 from 7.8 K to 145.8 K in device B. Above 110 K for device A, or 145.8 K for device B, the J_{max} starts to increase slowly. Above 150 K, the J_{max} of devices A and B are very similar since both devices are not lasing anymore. At a high temperature of 250 K, the J_{max} of device A is only ~150 A/cm^2 higher than at 10 K.

The light measurement of device A shows that the maximum amount of the light was collected at a current density near the NDR. The roll-over of output optical power in current ranges below maximum currents, observed in Ref. 18,

does not exist in our device. Since, on the one hand, the coupling injection strength of our device is lower than that of the one presented in Ref. 18 (0.85 meV compared to 2 meV), and on the other hand, the excess energy (voltage drop per module minus $2E_{LO}$) of our device is low (7.3 meV at design electric field that could compensate for the small injection energy E_{i2}), our device does not suffer from the roll-over effect that can degrade the performance of THz IDP-QCL.

The spectral measurements of this structure at different current injections and different temperatures are illustrated in Fig. 6. At 10 K, the lasing frequency started from ~2.4 THz at near threshold voltage and blue-shifted to ~2.8 THz at 21.7 V considering the highest amplitude longitudinal mode. At a current injection of 1.25 kA/cm^2, corresponding to 21.7 V, the structure acts as a multiple-wavelength source that lases from 2.32 THz to 2.94 THz at 10 K. Even though the Schrödinger equation helped us to estimate the lasing frequency of the first generation of 3P-QCL, it cannot predict the spectrum of V845 accurately. For comparison, the simulations (the dashed line in Fig. 3(c)) predict that the lasing frequency is 2.9 THz at 19.7 kV/cm (an actual device bias of 20.45 V); experimentally, 2.4 THz was observed. At 21 kV/cm, the lasing frequency is 3.36 THz (simulation) vs 2.32–2.94 THz (experiment). At 150 K and near the J_{max}, the spectral measurement shows a single lasing frequency of ~2.4 THz, which seems to be the dominant frequency range (2.4–2.5 THz) over the lasing operating temperatures.

The theoretical study of laser frequency versus bias and temperature is beyond the scope of this paper. We will mention only three mechanisms that can change the peak gain frequency, and which were not taken into account in our RE model. Many-body interactions, in particular the depolarization (intersubband plasmon), can red-shift the optical resonance in an inverted two-level system.[40,41] This depolarization effect could be weak though due to the small overlap between the lasing wavefunctions. If the population on LLS is substantial, for instance at high temperature by backfilling, the occurrence of Bloch gain cannot be excluded and it would tend to red-shift the peak position.[42,43] Finally, our model does not solve self-consistently the Schrödinger, Poisson and rate equations. In reality, due to charge separation the electric field is not uniform within one module. Since the section between the expected positions of the ULS and LLS wavefunctions is more conductive when stimulated emission occurs, the electric field can be configured differently when device is lasing, thereby reducing the Stark effect on the lasing transition.[44] However, considering the small carrier concentration in THz QCLs, this effect should be weak too. We have not verified yet numerically these three hypotheses. It is worth mentioning the NEGF simulations were more successful than RE in predicting the laser frequency (see Appendix).

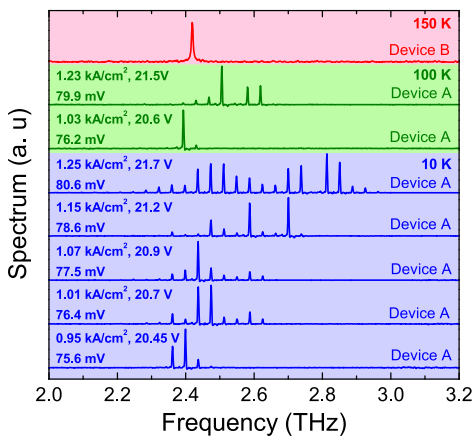


FIG. 6. THz spectra recorded for different biases and temperatures. The current density, the applied voltage bias, and voltage drop per module are reported in the figure. Spectrum at 150 K was collected from device B while all other spectra were measured from device A.

IV. RATE EQUATION ANALYSIS

A. Electrical characteristics

For the sake of simplicity, the pure dephasing time, τ^*, was left temperature independent (0.2 ps) in our rate equation-based modeling. We did not try to match for all temperatures the simulated J_{max} with the experimental values. To investigate the performance of this design, the current density of the structure in a full range of electric fields at different temperatures was calculated and the simulation results of lasing and non-lasing devices at 10 K are plotted in Fig. 7. A threshold electric field of 19.4 kV/cm was derived and matched the experimental value by assuming a gain bandwidth × cavity loss product of 38.5 THz cm^{-1}. Two pre-threshold current peaks occur before the final NDR: the effect of tunneling of states *e* to *1* and *e* to *2* aligned at

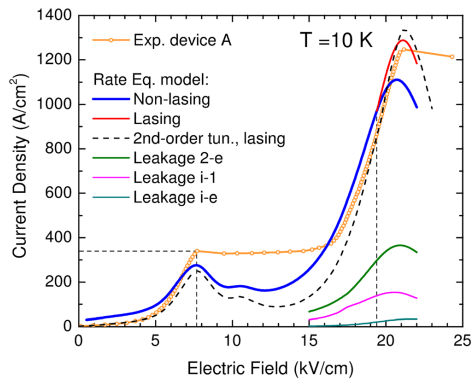


FIG. 7. The current density vs electric field were calculated by using a 5-level first-order rate equation formalism at 10 K for lasing (red) and non-lasing (blue) devices. The green, pink, and cyan lines represent the leakage currents from the wrong extraction *2-e*, and the wrong injections *i-1*, *i-e*, respectively. The vertical dashed lines were plotted to determine the first NDR and threshold voltage of the device at 10 K. The black dashed line shows the current density by using the second-order model of tunneling. The experimental curve of device A, shown as an orange dotted line, was measured at 10 K for comparison.

electric fields of 7.7 and 10.5 kV/cm, and giving rise to current peaks at ~7.7 and ~10.5 kV/cm, respectively. The current calculation with stimulated emission (red) and without stimulated emission (blue) is different after the threshold voltage. The peak current density of the lasing curve increases by ~177 A/cm^2, and is slightly shifted by +0.3 V in voltage compared to the non-lasing curve, while the two curves overlap before the threshold. From Fig. 3, two observations should be recalled: (i) above threshold, the total transit time—excluding the tunneling time—across the four quantum wells, $\tau_{tr}^{>}$, is much faster than the counterpart without stimulated emission, $\tau_{tr}^{<}$, this being mostly due to the high diagonality of the laser transition; (ii) the tunnel coupling strength is too low (0.85 meV) to ensure a coherent transport through the tunneling barrier ($2\Omega\tau_{\parallel ei} = 0.44$, instead of $2\Omega\tau_{\parallel ei} > 1$ if tunnel transport were coherent). As a result, the transport through the structure is essentially limited by the transit time within the period without stimulated emission ($\tau_{tun} \ll \tau_{tr}^{<} \approx 2\tau_{i2} + \tau_{21} + \tau_{1e}$), while the tunneling time becomes relevant with stimulated emission ($\tau_{tun} \lesssim \tau_{tr}^{>} \approx 2(\tau_{i2} + \tau_{1e})$), hence the maximum current with (without) stimulated emission will (not) be sensitive to the dephasing time, $\tau_{\parallel ei}$.[25] We came out with this estimation $\tau^* \sim 0.2$ ps in order to bring the simulated lasing V-J curve close to that of non-lasing and we found that it was roughly consistent with the intrasubband scattering rates by IR and e-impurity potentials (Sec. II C).

As already mentioned in Sec. II D, our 5-level rate equation model suggests a slight decrease of voltage (0.3 kV/cm in Fig. 7) at J_{max} for the non-lasing devices because of the decline of the population difference between the tunneling states, $n_e - n_i$, when these levels approach the resonance. This effect is not likely to happen on lasing devices because, above threshold, the population inversion is clamped, hence the population difference $n_e - n_i$ does not change very much above threshold and J_{max} occurs when the tunneling time, τ_{tun}, is at a minimum, i.e., when the levels *e* and *i* are aligned (21 kV/cm).[45] This voltage shift effect for non-lasing devices would not be easy to measure experimentally, unless a larger coupling strength is used to empty more drastically the extractor state. Recently, the electrical characteristic of a non-lasing device from a wafer with a large coupling ($\hbar\Omega = 1.54$ meV) strongly suggests that the peak current of the *e-i* channel is shifted by −1.5 V from the design voltage.[46]

The leakage current from the wrong extraction channel *2-e*, shown by a green line in Fig. 7, is lower than that of V843, as is its fractional contribution to the total current. Lower spatial overlap and higher energy spacing (50.5 meV in V845 compared to 42 meV in V843) between states *2* and *e* are the two main reasons for this lower leakage current and simply result from the specific wavefunction engineering where the net gain per electron injected was approximatively optimized through our GA approach. Two leakage currents from the wrong injection channels (*i-1* and *i-e*) were also plotted in Fig. 7. As expected, they are lower than the wrong extraction current, showing the good injection efficiency to the ULS in this IDP-QCL. The lower wrong injection could result in a higher internal quantum efficiency (longer $\tilde{\tau}_{2eff}$) and hence a higher output optical power.

The observed small shoulder in the J-V curves around 2 V that slowly develops above 130 K (Fig. 4(a)) is related to tunneling between levels 2_{n-1} and i_n and is very well simulated by the NEGF model (see Fig. 11 in the Appendix for NEGF simulated J–V at high temperatures). At such high temperatures, level *1* is more populated and the channel $1 \rightarrow i$ (resonant at ~6.6 kV/cm) is more activated, which, when combined with the $e \rightarrow 1$ channel (resonant at 7.7 kV/cm), results in slightly shifted, slightly less intense and broader peaks of the J–V characteristics (lower conductance) before the first NDR. Such alterations of the electrical characteristics before the first NDR have been experimentally observed here at high temperatures. The non-vanishing current below the *e-1* resonance comes from the first-order approximation of tunneling and a good fit was obtained in this voltage range by using the second-order approximation of tunneling (black dashed line in Fig. 7).[47]

B. Maximum current density versus temperature

When the RE model is employed with a constant pure dephasing time τ^*, a constant product $\Delta\nu \times \alpha_{cav}$, and five states (including the different scatterings between them), the simulated maximum current density stays roughly constant up to 85 K (with even a slight increase below 45 K) and then decreases more rapidly with temperature due to back scattering, whereas the experiment shows a reduction in J_{max} as devices are lasing (Fig. 5). The result of this simulation is plotted as a dashed line in Fig. 5. For a constant product $\Delta\nu \times \alpha_{cav} = 42$ THz cm^{-1}, the back scattering, simulated by RE model, would result in a reduction of J_{max} by ~260 A/cm^2 from 10 to 250 K, which is contrary to the experimental observation of an increase ~150 A/cm^2 over the same temperature range. For a phonon population in thermal equilibrium at 250 K, the different scatterings involving the fifth state result in a 4% raise in the simulated J_{max}, which is insufficient to compensate the back scattering effects.

Since on the one hand, the J_{max} of devices A and B decrease with temperature until the laser effect dies and eventually match their values close to and above the highest T_{max}, and on the other hand, the increase in current is unambiguous above T_{max}, this could suggest that two (or more) mechanisms are involved in the temperature dependence of J_{max}, for instance a mechanism specific to transport in the lasing regime, and another one specific to transport at high temperatures. The disagreements between experiments and RE simulations as shown in Fig. 5 are not well understood yet. The same behavior in maximum current density was also observed by NEGF simulations (see Appendix). The discrepancies may come (i) from employing a constant product $\Delta\nu \times \alpha_{cav}$ while this parameter might increase with temperature, (ii) from assuming a constant pure dephasing time while this parameter could decrease with temperature, thereby affecting the J_{max} in the lasing regime, and (iii) from omitting several possible leakage mechanisms, such as the excitation to the fifth level by reabsorption of nonequilibrium optical phonons[48–51] followed by thermionic emission to the continuum.[50]

The initial raise of the simulated J_{max} at low temperature is related to the energy gap between levels *i* and *2* that is

6.4 meV smaller than the phonon energy; however, if the subband i is more populated at high momentum than the assumed Boltzmann distribution, the increase of J_{max} (and gain too) at low temperature would partially—if not completely—disappear. Modeling accurately the maximum current density, in particular its temperature dependence, is the subject of intense research in order to understand all aspects of transport in THz QCL.[48,50,52,53]

C. Intermediate resonances

At low temperature, there are two important anticrossing resonances before the main resonance between states e and i. Since both simulated current peaks at electric fields of 7.7 and 10.5 kV/cm are far less than the threshold current, the two pre-threshold tunneling resonances (*e-1* and *e-2*) impose a minimum impact on device lasing performance. This calculation should be performed for all structures based on the 3P scheme to make sure the currents at resonances of e to *1* and e to *2* are substantially lower than the threshold current, without sacrificing the dynamic range of the laser. To show how the wavefunctions couple when the leakage current density due to the *e-1* tunneling resonance is peaked, the conduction band diagram of V845 is calculated and plotted in Fig. 8. The lowest energy state of the left module (e_{n-1}) is in resonance with the second energy state of the right module (1_n) at an electric field of 7.7 kV/cm, which is lower than the threshold electric field. The coupling strength between these states is $\hbar\Omega_{e1} = 0.235$ meV. Since the tunneling between level e and *1* is incoherent ($4\Omega_{e1}^2\tau_{\|e1}\tau_1 = 0.35$ at 20 K), the dephasing time constant can affect the tunneling current.[8,25] The carriers passing through the injector barrier (tunnel from e_{n-1} to 1_n) will be relaxed to the next extraction state (e_n). The second resonance should be observed at 10.5 kV/cm,

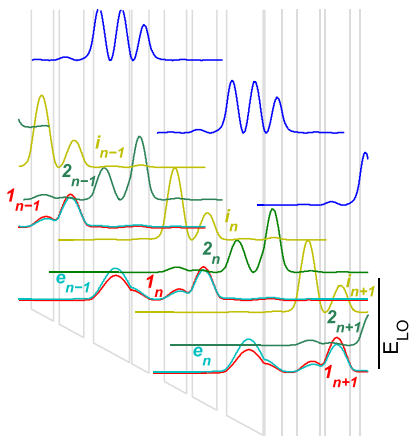


FIG. 8. Conduction band diagram and the moduli squared of wavefunctions of V845 at 7.7 kV/cm. States in left module (upstream), middle module, and right module (downstream) are represented by subscripts $n-1$, n, and $n+1$, respectively. The extraction state (e) of each module is in resonance with state (1) of next module at an electric field of 7.7 kV/cm.

where the states e and *2* are aligned. For this resonance, the transport is clearly incoherent, with a low coupling strength between the states e to *2* ($\hbar\Omega_{e2} = 0.147$ meV) and a short relaxation time of level *2* ($\tau_{2e} = 0.25$ ps) that result in very low coherence in tunneling ($4\Omega_{e2}^2\tau_{\|e2}\tau_{2e} \sim 6\times10^{-3}$ at 20 K). Therefore, the current through the *e-2* channel is smaller than that of *e-1* channel because of its smaller coupling strength and, to a lesser extent, due to its shorter dephasing time. We note that the relaxation times of levels *2* and *1* are inverted at these low electric fields, i.e., fast for level *2* (0.18 ps) and slow for level *1* (3.48 ps) as the two lasing states did not anticross yet (*2* and *1* are aligned at 13.3 kV/cm).

The measured peak leakage current at the first NDR, i.e., 7.7 kV/cm, is $J_{res} = 340\text{A/cm}^2$, while the simulation result is 276 A/cm^2. Even though the maximum current density of the lasing device can be reasonably well predicted by our simulation (at least at low temperature), the experimental value of J_{max} on the non-lasing device is 78 A/cm^2 higher than our simulation result. A similar vertical shift (64 A/cm^2) was also observed in the peak leakage current density at 7.7 kV/cm (340 A/cm^2 measured value vs a 276 A/cm^2 simulation result). The ratio of the maximum current over the resonance peak current before threshold (J_{max}/J_{res}) at 10 K in V845 is slightly better than that of V843 (3.7 in V845 compared to 3.1 in V843), due to the weaker *e-2* resonance in V845 because of the wider radiative barrier ($\hbar\Omega_{e2} = 0.147$ meV in V845 vs 0.24 meV in V843).

D. Differential resistance at threshold

The differential resistance of lasing device A (at different temperatures) versus current density, plotted in Fig. 9, shows a clear discontinuity at the laser threshold, whereas the same parameter is displayed as a smooth curve (dashed-magenta line) for the non-lasing mesa device (shown only at 4 K). The ~16% fractional change of differential resistance at threshold, observed in measurement at 10 K, is lower than what the rate equation model predicts (32% extracted from

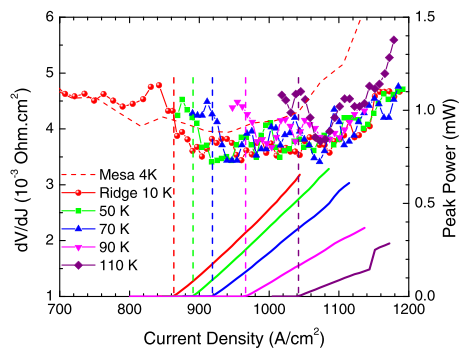


FIG. 9. Left axis: The differential resistance of non-lasing (the red dashed line) and lasing (solid lines with symbols) device A versus current density at different temperatures. The L-J measurement results are also plotted (right scale) to determine the threshold current at each temperature.

Fig. 7 by assuming $\Delta\nu \times \alpha_{cav} = 38.5\,\mathrm{THz\,cm^{-1}}$). To more accurately determine the position of the discontinuity which shifts to higher currents with temperature, the L-J curves of device A are plotted with vertical dashed lines to denote the corresponding threshold points. A better internal efficiency of the second generation of THz 3P-QCL due to the longer *modified* effective lifetime of the ULS and shorter lifetime of the LLS results in a clear discontinuity of differential resistance and a higher output power (1.5 mW in device A compared to 0.8 mW in Ref. 9 with the same collecting optics). In Ref. 9, the discontinuity of differential resistance was estimated and appears to depend on two relaxation times associated with the extractor state, the wrong extraction lifetime (τ_{2e}), and the depopulation (τ_{1e})

$$\frac{\Delta R_{th}}{R_{th}} = \Delta\rho_{th}\frac{\tau_{2e} - \tau_{1e}}{\tau_{2e} + \tau_{1e}}, \tag{3}$$

where $\Delta\rho_{th}$ is the normalized population inversion at threshold. The discontinuity of differential resistance in V845 is improved compared to the first iteration 3P design (V843) in a two-fold strategy: (i) τ_{2e} is longer because of the thicker radiative barrier ($\tau_{2e} = 5.6$ vs 2.55 ps in V843 at electron temperature of 70 K) which results in a higher *modified* effective ULS lifetime and (ii) τ_{1e} is shorter due to the higher energy spacing to make the depopulation more efficient. The observation of a discontinuity in differential resistance on V845 tends to support our hypothesis about the vanishing discontinuity on V843, which was attributed to a slow depopulation rate as the energy spacing between the states *1* and *e* was 9 meV below the LO-phonon energy.[9] This discontinuity in differential resistance on V845 is consistent with the higher emission power from this wafer, since $\Delta R_{th}/R_{th}$ is closely related to the internal efficiency of the QCL.

E. Cavity loss estimation

Even though this RE model is not very accurate, for instance in predicting J_{max}, it was used here for estimating the product between the gain bandwidth, $\Delta\nu$, and the cavity loss, α_{cav}. Figure 10 shows the simulated $\Delta\nu \times \alpha_{cav}$ as a function of lattice temperature. The gain of the structure was calculated at all electric fields and different temperatures and the cavity loss and/or the gain bandwidth were adjusted to match the simulated and experimental voltage thresholds. Simulations were performed under two conditions: (i) the electron temperature is 50 K($\equiv\Delta T_e$) higher than lattice temperature, (ii) the electron temperature is 100 K higher than lattice temperature. By having the voltage threshold of device A at different temperatures (the experimental results range from ~19.4 V at 10 K to 20.5 V at 128.5 K after substration of the Schottky barrier) and mapping the gain value at each temperature, a product $\Delta\nu \times \alpha_{cav} \sim 42\,\mathrm{THz\,cm^{-1}}$ was inferred for $\Delta T_e = 50$ K. This inferred value remains almost constant (~42 $\mathrm{THz\,cm^{-1}}$) at temperatures of 50-120 K, but slightly drops to ~38.5 $\mathrm{THz\,cm^{-1}}$ at 10 K. This drop is likely attributed to the underestimated gain calculation at lower temperatures because the slow-down in carrier injection from level *i* to *2* at low temperatures (e.g., below

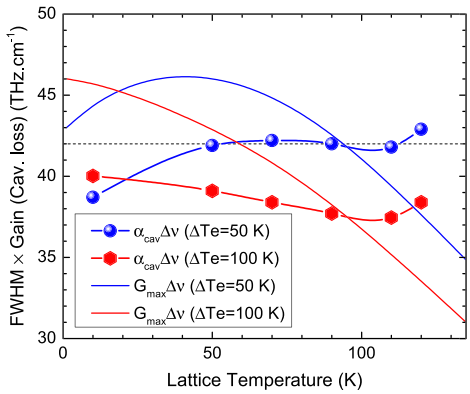


FIG. 10. Symboled lines are the cavity loss × gain bandwidth products of device A, calculated by a 4-level RE model, at different lattice temperatures (T_L) and for two electron temperatures, T_e, such as $T_e - T_L = 50$, 100 K. The solid lines are the peak gain × gain bandwidth products vs lattice temperature at 19.7 kV/cm calculated by the RE model.

20-30 K) might be exaggerated in the modeling. This underestimation of the gain leads to a lower inferred cavity loss. Indeed, especially at low temperatures, the injection of carriers into the ULS is very sensitive to the actual electron distribution of carriers in the injector subband as the energy spacing E_{i2} is fairly detuned from the phonon energy (7.45 meV at 19.4 kV/cm, the threshold electric field at 10 K); in other words, the Boltzmann distribution with a characteristic temperature of 60 K (used for the simulation at 10 K lattice temperature) might be too far from reality.

It is worth mentioning that the inferred product $\Delta\nu \times \alpha_{cav} \sim 42\,\mathrm{THz\,cm^{-1}}$ is for an assumed Lorentzian gain profile ($C = 1/\pi$ in Eq. (2)); if a Gaussian gain spectrum had been chosen, this value would be multiplied by 1.48 (62 THz $\mathrm{cm^{-1}}$). Close to T_{max} of device A, the peak gain and bandwidth simulated by NEGF are 33 $\mathrm{cm^{-1}}$ and 1.56 THz, respectively (see Appendix), resulting in a product of 51.5 $\mathrm{THz\,cm^{-1}}$, in agreement with the estimation from the RE model (42-62 $\mathrm{THz\,cm^{-1}}$). Under the second condition, when the electron temperature is 100 K higher than the lattice temperature, the derived $\Delta\nu \times \alpha_{cav}$ decreases with temperature which is unlikely to happen.[13] This behavior is due to a too fast degradation of the gain versus temperature when $\Delta T_e = 100$ K as this assumed temperature difference probably overestimates the carrier distribution at high momentum in the subbands. Globally, it seems the heating is less than 100 K.[54] Repeating the same exercise with the lasing device B, a lower $\Delta\nu \times \alpha_{cav}$ product of ~38 $\mathrm{THz\,cm^{-1}}$ (with Lorentzian profile) was estimated for this waveguide, which is consistent with previous reports.[5,13,55]

V. CONCLUSION

In this paper, we present a THz indirectly pumped QCL based on phonon-assisted injection and extraction. A simple rate equation formalism was employed and a genetic

algorithm was used to optimize the structure to have the maximum gain while keeping the current injection as low as possible. The product of net optical gain and transit time was defined as a figure of merit to be maximized at the lattice temperature of 150 K in the design optimization process. The fifth energy state was engineered to be far from the first four states and makes sure its carrier density is negligibly low.

Based on the design optimization, the second generation of THz 3 P-QCL was demonstrated and a higher dynamic range ($(I_{max} - I_{th})/I_{max} = 0.3$ (0.4) in device A (device B) compared to 0.24 in V843), higher operating temperature (152.5 K compared to 138 K (Ref. 56)) and lower lasing frequency (2.4 THz at 150 K compared to 3.2 THz at 138 K) were obtained. The lower loss waveguide plays an important role in pushing the maximum temperature to 152.5 K and a high performance THz QCL with $k_B T_{max}/\hbar\omega = 1.3$ was demonstrated. Despite these improved performances, the intrasubband scattering contribution to the dephasing of the tunneling process has been identified as a crucial parameter, which limited the dynamic range of this device. Therefore, future optimization schemes should include a realistic model of the tunnel dephasing and linewidth of the lasing transition.

To further improve the performance of THz 3P-QCL, one can design a structure with a thinner injector barrier to increase the gain and maximum current. A higher injection tunnel coupling strength might be tolerable as long as the early NDR, here observed at 7.7 kV/cm, can effectively be suppressed. As a result, the lasing dynamic range can be expanded and the maximum lasing temperature could be further enhanced. Inserting one more quantum well in the upstream phonon double well could effectively minimize the leakage current arising from tunneling resonances prior to lasing threshold.[11] In such QCL structures, a higher injection tunnel coupling strength might, therefore, be chosen to further improve device performance.

It is fair to recall the other very efficient—and experimentally proven—strategy to suppress the early NDR, which employs a design of a two-state extractor.[18,19,57] This solution would imply dealing with five and possibly six levels per module because the third state of the double-well extractor will likely be close in energy with the lasing states. In such a case, this third state should be employed in the depopulation mechanism via tunneling, resulting in the abandonment of the "3P" nature of the designs we have explored because a two-level extractor IDP-QCL would work according to a 2PTP sequence (phonon-photon-tunnel-phonon) between two injection barriers.

The main message conveyed by this paper is to demonstrate that, despite its disarming simplicity, the 3P-QCL device can be optimized in some sophisticated ways and perhaps, could reach performances suitable for practical applications.

ACKNOWLEDGMENTS

The authors thank Dr. Marek Korkusinski from NRC for providing the genetic algorithm and Pietro Patimisco from the Università and Politecnico di Bari for helpful stimulating discussions. They also would like to acknowledge the financial supports from Natural Science and Engineering Research Council (NSERC) of Canada, Canadian Foundation of Innovation (CFI), the CMC Microsystems, and Ontario Research Fund (ORF). H.C.L. was supported in part by the National Major Basic Research Project (2011CB925603) and the Natural Science Foundation of China (91221201 and 61234005).

APPENDIX: SIMULATIONS BY NONEQUILIBRIUM GREEN'S FUNCTION MODEL

Current and carrier densities as well as gain spectra of the V845 device were simulated with the NEGF method,[28] where the recent implementation described in Ref. 58 is used. As basis states, we use the Wannier functions of the envelope hamiltonian of the perfect QCL structure at zero bias. For this approach, the barrier height was ~12.5 meV smaller than in the RE model and a zero non parabolicity was assumed. Both features widely compensate and the resulting level spectrum is almost identical. Scattering effects from interface roughness, impurities, and phonons are treated with self-energies within the self-consistent Born approximation, while electron-electron interactions are appreciated by a self-consistently calculated mean field potential.

The current density has been calculated at different lattice temperatures as shown in Fig. 11. In these simulations, J_{max} is found at a total bias which is lower than the experimental bias of J_{max} (shifted by the Schottky bias drop) by ~1.8 V. The difference is slightly smaller for the first resonance peak. This contrasts with the RE method used in the main text, where no such shift was observed. We attribute the main part of the shift to the real parts of the self-energies and the mean field potential included in the NEGF model. Taking into account the real parts of the self-energies and the mean field potential provides different Wannier-Stark states as shown in Fig. 12(a). Here, the tunneling resonance between states e and i occurs at a field of 18.7 kV/cm, but the

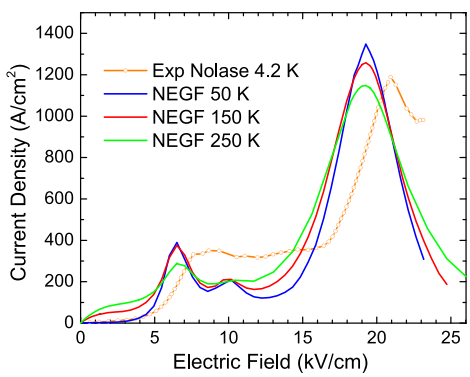


FIG. 11. Current densities at different lattice temperatures, calculated with the NEGF method. J_{max} is located at 19.3 kV/cm. The experimental curve of the non-lasing device, shown as an orange dotted line, was measured at 4.2 K for comparison.

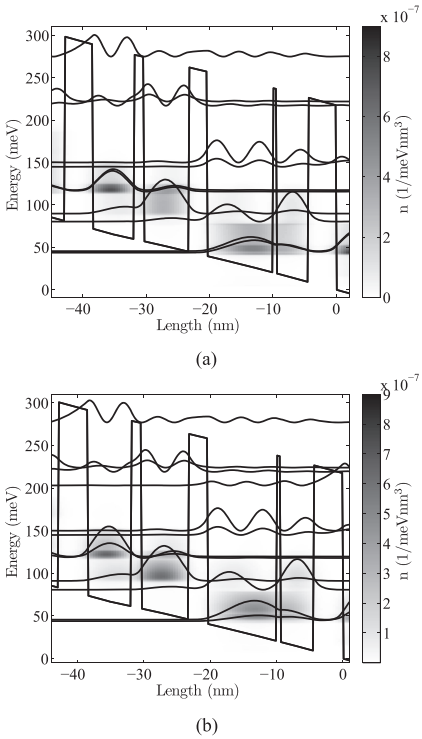


FIG. 12. Carrier densities at (a) 18.7 kV/cm and (b) 19.3 kV/cm. Current is peaked at the bias in (b), although the tunneling resonance is greater in (a).

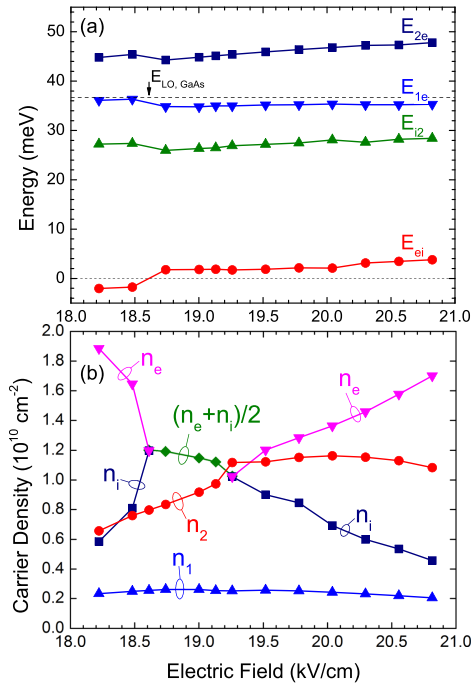


FIG. 13. (a) The detuning energy between extraction and injection states E_{ei}, the energy differences of the extraction E_{1e}, the injection E_{i2} and the energy spacing E_{2e} of the main leakage channel $2 \rightarrow e$. (b) Population densities of the injection n_i, upper laser n_2, lower laser n_1 and extraction n_e states. At electric fields where i and e are almost degenerate, the average value of n_e and n_i is shown.

injection energy $E_{i2} = 25.97$ meV deviates strongly from the phonon energy $E_{LO} = 36.7$ meV. (A detailed summary of the values is given in Fig. 13.) As a result, carriers accumulate in the injection state. Correspondingly, the current peak occurs for a slight detuning of E_i and E_e at 19.3 kV/cm (Fig. 12(b)). At higher fields, the detuning between E_i and E_e becomes significant and the current drops. Note that the population of the ULS is asymmetric with respect to the current peak, as the transition energy E_{2e} of the leakage channel $2 \rightarrow e$ becomes more detuned from E_{LO} with increasing field. This also results in a stronger inversion on the gain transition. The relevance of the mean-field can be quantified by solving Poisson's equation including the electron densities in Fig. 12(b) and the doping density. We find a maximum mean field potential of $\sim$1.8 meV (in the LLS well) and a minimum of ~ -0.7 meV in the doped well. This potential effectively raises e and the LLS, while it lowers i. This requires a lower bias for the $e-i$ tunneling resonance as compared to the case without mean field.

As the temperature increases, inversion decreases as a result of thermal backfilling from e to the LLS, as well as from the ULS to i. The first backfilling effect shows the importance of emptying the extraction state efficiently at the operating bias, whereas the latter shows the importance of matching the injection energy to E_{LO}. At fields below the first resonance peak, the current increases with temperature, whereas the peak itself decreases, in good agreement with the experiment (Fig. 4). In the region between the first resonance peak and J_{max}, the current also increases as in the experimental measurements. However, J_{max} decreases with temperature, whereas for the lasing device there is a decrease in current at low temperatures, but for higher temperatures the current drastically increases. The same effect is seen in the RE simulations with 5 states per period, and the NEGF with its 7 states per period thus confirms that this increase in J_{max} does not come from thermal excitations to higher states.

The gain spectra computed at different temperatures and for the electric field of J_{max} ($\sim$19.3 kV/cm) are shown in Fig. 14. Peak gain is achieved for $\nu = 2.5$ THz at J_{max}, which agrees much better with the experimental data than the RE results. This frequency corresponds to the difference in energy of the Wannier-Stark states responsible for lasing, and we thus conclude that the dispersive gain and depolarization shift included in the NEGF model do not play an

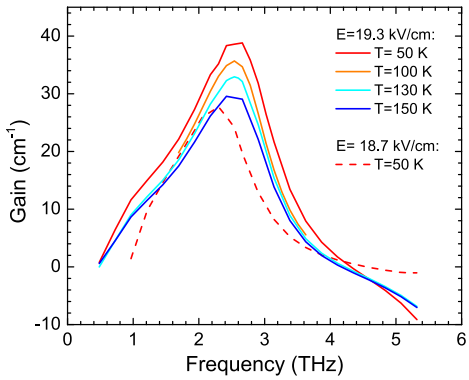


FIG. 14. The gain spectrum of the structure near the NDR and for different lattice temperatures. At 50 K, the gain spectrum at lower electric field is also plotted to show the red shift in the spectrum.

important role. The peak value of the gain decreases from 39 cm^{-1} at 50 K to 33 cm^{-1} at 130 K and finally, down to 30 cm^{-1} at 150 K. The full width at half maximum of the calculated gain ($\Delta\nu$) at 130 K is ~1.56 THz. At lower fields (see dashed line of Fig. 14), the gain goes down in magnitude, and the peak gain shifts to lower frequencies with a Stark shift ~0.45 THz/V, which is slightly higher than the experimental observation at 10 K (Fig. 6).

[1]B. S. Williams, Nature Photon. **1**, 517 (2007).
[2]C. Walther, M. Fischer, G. Scalari, R. Terazzi, N. Hoyler, and J. Faist, Appl. Phys. Lett. **91**, 131122 (2007).
[3]C. W. I. Chan, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **101**, 151108 (2012).
[4]R. Kohler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, Nature **417**, 156 (2002).
[5]S. Fathololoumi, E. Dupont, C. Chan, Z. Wasilewski, S. Laframboise, D. Ban, A. Mátyás, C. Jirauschek, Q. Hu, and H. C. Liu, Opt. Express **20**, 3866 (2012).
[6]H. Yasuda, T. Kubis, P. Vogl, N. Sekine, I. Hosako, and K. Hirakawa, Appl. Phys. Lett. **94**, 151109 (2009).
[7]T. Kubis, S. R. Mehrotra, and G. Klimeck, Appl. Phys. Lett. **97**, 261106 (2010).
[8]S. Kumar, C. W. I. Chan, Q. Hu, and J. L. Reno, Nat. Phys. **7**, 166 (2011).
[9]E. Dupont, S. Fathololoumi, Z. R. Wasilewski, G. Aers, S. R. Laframboise, M. Lindskog, S. G. Razavipour, A. Wacker, D. Ban, and H. C. Liu, J. Appl. Phys. **111**, 073111 (2012).
[10]M. Yamanishi, K. Fujita, T. Edamura, and H. Kan, Opt. Express **16**, 20748 (2008).
[11]S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **88**, 121123 (2006).
[12]S. Kumar, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **94**, 131105 (2009).
[13]M. A. Belkin, J. A. Fan, S. Hormoz, F. Capasso, S. P. Khanna, M. Lachab, A. G. Davies, and E. H. Linfield, Opt. Express **16**, 3242 (2008).
[14]M. Belkin, Q. J. Wang, C. Pflugl, A. Belyanin, S. Khanna, A. Davies, E. Linfield, and F. Capasso, IEEE Sel. Top. Quantum Electron. **15**, 952 (2009).
[15]A. Wacker, Appl. Phys. Lett. **97**, 081105 (2010).
[16]T. Liu, T. Kubis, Q. J. Wang, and G. Klimeck, Appl. Phys. Lett. **100**, 122110 (2012).
[17]M. Yamanishi, K. Fujita, N. Yu, T. Edamura, K. Tanaka, and F. Capasso, paper presented at Conference on Lasers and Electro-Optics, Baltimore, MD, USA, 1–6 May 2011.
[18]K. Fujita, M. Yamanishi, S. Furuta, K. Tanaka, T. Edamura, T. Kubis, and G. Klimeck, Opt. Express **20**, 20647 (2012).
[19]M. Yamanishi, K. Fujita, S. Furuta, K. Tanaka, T. Edamura, T. Kubis, and G. Klimeck, paper presented at International quantum cascades laser school and workshop, Baden, Austria, 2–6 September 2012.
[20]D. Indjin, P. Harrison, R. W. Kelsall, and Z. Ikonić, Appl. Phys. Lett. **82**, 1347 (2003).
[21]H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **83**, 207 (2003).
[22]H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **84**, 645 (2004).
[23]C. Jirauschek, G. Scarpa, P. Lugli, M. S. Vitiello, and G. Scamarcio, J. Appl. Phys. **101**, 086109 (2007).
[24]G. Scalari, R. Terazzi, M. Giovannini, N. Hoyler, and J. Faist, Appl. Phys. Lett. **91**, 032103 (2007).
[25]S. Kumar and Q. Hu, Phys. Rev. B **80**, 245316 (2009).
[26]E. Dupont, S. Fathololoumi, and H. C. Liu, Phys. Rev. B **81**, 205311 (2010).
[27]H. Callebaut and Q. Hu, J. Appl. Phys. **98**, 104505 (2005).
[28]S.-C. Lee and A. Wacker, Phys. Rev. B **66**, 245314 (2002).
[29]T. Kubis, C. Yeh, P. Vogl, A. Benz, G. Fasching, and C. Deutsch, Phys. Rev. B **79**, 195323 (2009).
[30]T. Schmielau and J. M. F. Pereira, Appl. Phys. Lett. **95**, 231111 (2009).
[31]H. Sakaki, IEEE J. Quantum Electron. **22**, 1845 (1986).
[32]F. Carosella, C. Ndebeka-Bandou, R. Ferreira, E. Dupont, K. Unterrainer, G. Strasser, A. Wacker, and G. Bastard, Phys. Rev. B **85**, 085310 (2012).
[33]C. Ndebeka-Bandou, F. Carosella, R. Ferreira, A. Wacker, and G. Bastard, Appl. Phys. Lett. **101**, 191104 (2012).
[34]S. Fathololoumi, E. Dupont, Z. R. Wasilewski, C. W. I. Chan, S. G. Razavipour, S. R. Laframboise, S. Huang, Q. Hu, D. Ban, and H. C. Liu, J. Appl. Phys. **113**, 113109 (2013).
[35]The figure of merit used during the GA optimization of this work was defined as a product of (gain(cm^{-1})-23), average transit time per period, and inverse of the superperiod length at 150 K. The value of 23 cm^{-1} for the waveguide loss might been slightly underestimated. The gain was calculated by using a 4-level rate equation model and by assuming $\Delta\nu =$ 1 THz for the full width at half maximum. A pure dephasing time of 0.35 ps was used. Even though during the GA optimization the gain was estimated for a 3D doping of 7×10^{15} cm^{-3}, at the last minute, before the MBE growth, it was finally increased to 9×10^{15} cm^{-3}.
[36]R. Terazzi, Ph.D. dissertation, ETH Zurich, 2012.
[37]T. Ando, J. Phys. Soc. Jpn. **54**, 2671 (1985).
[38]A. Mátyás, M. A. Belkin, P. Lugli, and C. Jirauschek, Appl. Phys. Lett. **96**, 201110 (2010).
[39]S. Fathololoumi, E. Dupont, S. G. Razavipour, S. R. Laframboise, G. Parent, Z. Wasilewski, H. C. Liu, and D. Ban, Semicond. Sci. Technol. **26**, 105021 (2011).
[40]M. V. Kisin, M. A. Stroscio, G. Belenky, and S. Luryi, Appl. Phys. Lett. **73**, 2075 (1998).
[41]T. Liu, K. E. Lee, and Q. J. Wang, Phys. Rev. B **86**, 235306 (2012).
[42]H. Willenberg, G. H. Döhler, and J. Faist, Phys. Rev. B **67**, 085315 (2003).
[43]R. Terazzi, T. Gresch, M. Giovannini, N. Hoyler, N. Sekine, and J. Faist, Nat. Phys. **3**, 329 (2007).
[44]Y. Yao, Z. Liu, A. Hoffman, K. Franz, and C. Gmachl, IEEE J. Quantum Electron. **45**, 730 (2009).
[45]In the limit of infinitely short lifetime of states i and 1, $n_2 + n_e = n_s$, which suggests that the population difference between the tunneling states above threshold, $n_e - n_i$, would be weakly voltage dependent in order to take into account the voltage dependence of the oscillator strength.
[46]S. G. Razavipour, E. Dupont, S. Fathololoumi, Z. R. Wasilewski, M. Lindskog, A. Wacker, S. R. Laframboise, and D. Ban, paper presented at International quantum cascades laser school and workshop, Baden, Austria, 2–6 September 2012.
[47]R. Terazzi, T. Gresch, A. Wittmann, and J. Faist, Phys. Rev. B **78**, 155328 (2008).
[48]R. C. Iotti, F. Rossi, M. S. Vitiello, G. Scamarcio, L. Mahler, and A. Tredicucci, Appl. Phys. Lett. **97**, 033110 (2010).
[49]M. S. Vitiello, L. Viti, L. Romeo, D. Ercolani, G. Scalari, J. Faist, F. Beltram, L. Sorba, and A. Tredicucci, Appl. Phys. Lett. **100**, 241101 (2012).
[50]S. Kumar, C. W. I. Chan, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **95**, 141110 (2009).
[51]S. Kumar, Proc. SPIE **8496**, 849603 (2012).
[52]I. Bhattacharya, C. W. I. Chan, and Q. Hu, Appl. Phys. Lett. **100**, 011108 (2012).

[53]W. Freeman and G. Karunasiri, Phys. Rev. B **85**, 195326 (2012).
[54]P. Patimisco, G. Scamarcio, M. V. Santacroce, V. Spagnolo, M. S. Vitiello, E. Dupont, S. R. Laframboise, S. Fathololoumi, G. S. Razavipour, and Z. Wasilewski, Opt. Express **21**, 10172 (2013).
[55]A. Mátyás, R. Chashmahcharagh, I. Kovacs, P. Lugli, K. Vijayraghavan, M. A. Belkin, and C. Jirauschek, J. Appl. Phys. **111**, 103106 (2012).
[56]Device V843 was also fabricated at MIT with the same process as device B and it lased up to 141 K.
[57]B. Williams, S. Kumar, Q. Hu, and J. Reno, Opt. Express **13**, 3331 (2005).
[58]A. Wacker, M. Lindskog, and D. O. Winge, IEEE J. Sel. Top. Quantum Electron. **19**, 1200611 (2013).

Paper III

Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions

Andreas Wacker, Martin Lindskog, and David O. Winge

(Invited Paper)

***Abstract*—A simulation scheme based on nonequilibrium Green's functions for biased periodic semiconductor heterostructure devices is presented in detail. The implementation can determine current and optical gain both for small and large optical fields. Specific results for superlattices, quantum cascade lasers, and quantum cascade detectors are shown which demonstrate the capabilities of the approach.**

***Index Terms*—Nonequilibrium Green's functions (NEGF), quantum cascade laser, simulation.**

I. Introduction

QUANTUM cascade lasers (QCLs) [1] have become an important source of radiation for infrared spectroscopy [2]–[4]. In addition, lasing in the terahertz range [5] opens up a wide range of further applications [6]–[10]. These devices rely on a precise design of the underlying semiconductor heterostructure, so that the combination of electron tunneling and scattering provides inversion in the population of the subbands. While a first idea about the operation can be obtained by relatively simple principles such as level alignments and scattering rates [1], a full understanding of the operating requires a more detailed quantum description.

Typically the operation of QCLs is modeled by *rate equations* [11] between the levels of the active regions, while the current flow through the injector is taken into account phenomenologically. The transition rates are evaluated microscopically within Fermi's golden rule for phonon scattering [12], [13] and partially for electron–electron scattering [14], [15]. In addition, confined phonon modes [16] and hot phonon effects [17] have been studied. If one includes the injector states in such a simulation, one obtains a self-consistent simulation of the entire structure [18], [19] within the semiclassical carrier dynamics.

While rate equations only take into account the electron density of subbands, *Monte Carlo simulations* of the Boltzmann equation [20], [21] allow for a study of nonequilibrium distributions within the subbands [22]. Here, the importance of electron–electron scattering [23] is debated. In [24], Iotti and Rossi show that the impact of electron–electron scattering is strong if no elastic scattering mechanism is taken into account. In contrast the authors of [25] find that elastic impurity scattering gives stronger effects than electron–electron scattering.

Density matrices include the correlations $\rho_{\alpha\beta}$ between different quantum states $\alpha \neq \beta$. These are of particular importance for the tunneling through the injection barrier, where their neglect provides the wrong result, that is the peak tunnel current does not drop with the barrier width [26], [27]. In a more phenomenological way, this can be done on the level of densities [28]–[30], which is very cost effective. Taking into account the k-resolution, the equations for the density matrix $\rho_{\alpha\beta}(k_x, k_y)$ become much more involved [24], [27], [31]. Here, it is a well-known problem that unphysical negative occupations may occur in frequently used approximation schemes; see [27] for a thorough discussion. A possible solution by using further approximations is outlined in [32] on the level of densities.

Nonequilibrium Green's functions (NEGF) constitute a higher step of sophistication. They can be viewed as energy-resolved density matrices and allow for a consistent perturbative treatment of scattering and quantum evolution. On the other side, their numerical implementation is heavy, in particular if electron–electron scattering is considered. They have been used for QCL simulations by various groups in different types of implementations [33]–[38].

The purpose of this paper is to present a detailed account of our implementation of NEGF in Section II. On the one hand, this complements earlier descriptions [34], [35] for the stationary transport. Furthermore, we present our extension to include the lasing field beyond linear response, which we recently applied in [39]. We demonstrate the strength of the simulation tool by applications for superlattices, QCLs, and quantum cascade detectors (QCDs) in Sections III–V, respectively.

Manuscript received December 3, 2012; revised January 9, 2013; accepted January 9, 2013. Date of publication January 18, 2013; date of current version May 13, 2013. This work was supported by the Swedish Research Council.

The authors are with the Division of Mathematical Physics, Lund University, Lund 22100, Sweden (e-mail: andreas.wacker@fysik.lu.se; martin.lindskog@teorfys.lu.se; david.winge@teorfys.lu.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTQE.2013.2239613

II. Theoretical Formulation

We consider a general sequence of heterostructure layers, where z is the growth direction. For an ideal structure, the system is translational invariant in the x and y directions and we use a

complete set of states

$$\Psi_{\alpha,\mathbf{k}}(x,y,z)=\varphi_\alpha(z)\frac{\mathrm{e}^{\mathrm{i}(k_x x+k_y y)}}{\sqrt{A}} \tag{1}$$

where $\mathbf{k}=(k_x,k_y)$ is a vector within the x,y plane of normalization area A. The states $\varphi_\alpha(z)$ can be chosen in different ways. Typical examples are eigenstates of the pure Hamiltonian $\hat{H}_0$ (also called Wannier–Stark states) as used by most groups, or a site representation as used in [37] and [38]. We usually apply Wannier States; see [34, Appendix A], which allow for a consistent description of periodic extended structures. In particular, their construction does not require any artificial boundary conditions (such as the Wannier–Stark states) and they provide reliable energy levels (while these energies are approximate in a site representation with a manageable number of grid points). Nevertheless, the following description can be used for any kind of orthonormal states $\varphi_\alpha(z)$.

With in this basis, the total Hamiltonian can be written as follows:

$$\hat{H}=\hat{H}_0+\hat{H}_{\mathrm{ac}}(t)+\hat{H}_{\mathrm{MF}}(t)+\hat{H}_{\mathrm{scatt}}. \tag{2}$$

Here, $\hat{H}_0$ contains the kinetic energy, the heterostructure potential, and the constant electric field F due to the applied bias.

$\hat{H}_{\mathrm{ac}}(t)$ is due to a laser field with the electrical component $F_{\mathrm{ac}}\vec{e}_z\cos(\omega t-kx)$. Here, $k=\sqrt{\epsilon_r}\omega/c$ with the refractive index $\sqrt{\epsilon_r}$. Neglecting terms of order k (just as in the common dipole approximation), we obtain the potentials [33]

$$\text{Coulomb gauge}\quad \vec{A}(t)=-\frac{F_{\mathrm{ac}}}{\omega}\vec{e}_z\sin(\omega t)\quad \phi(z,t)=0$$

$$\text{Lorenz gauge}\quad \phi(z,t)=-F_{\mathrm{ac}}z\cos(\omega t)\quad \vec{A}(t)=0$$

which allow us to construct $\hat{H}_{\mathrm{ac}}(t)$ in the standard way.

$\hat{H}_{\mathrm{MF}}(t)$ is the mean field due to the charge distribution $\rho(z,t)$ in the heterostructure, which is obtained by solving Poisson's equation. (Its time-independent part is frequently incorporated into $\hat{H}_0$.) Finally, $\hat{H}_{\mathrm{scatt}}$ describes the interaction with impurities, phonons, and other sources destroying the translational invariance in x and y directions. Thus, this term is not diagonal in $\mathbf{k}$, in contrast to all other contributions to $\hat{H}$.

Most physical observables can be expressed by the (reduced) density matrix

$$\rho_{\alpha\beta}(\mathbf{k},t)=\left\langle a^\dagger_{\beta\mathbf{k}}(t)a_{\alpha\mathbf{k}}(t)\right\rangle=\mathrm{Tr}\left\{\hat{\rho}a^\dagger_{\beta\mathbf{k}}(t)a_{\alpha\mathbf{k}}(t)\right\} \tag{3}$$

where the creation/annihilation operators $a^\dagger_{\alpha\mathbf{k}}(t)/a_{\alpha\mathbf{k}}(t)$ of the single particle level $\Psi_{\alpha\mathbf{k}}(x,y,z)$ have a time dependence in the Heisenberg picture and $\hat{\rho}$ is the density operator. Note that averaging over impurity positions and phonon mode occupations typically renders zero expectation values of $\langle a^\dagger_{\beta\mathbf{k}'}(t)a_{\alpha\mathbf{k}}(t)\rangle$ for $\mathbf{k}'\neq\mathbf{k}$. These density matrices $\rho_{\alpha\beta}(\mathbf{k},t)$ allow for the calculation of the relevant physical properties. For example, the electron sheet density in level α is given by $n_\alpha(t)=\bar{\rho}_{\alpha\alpha}(t)$, where

$$\bar{\rho}_{\alpha\beta}(t)=\frac{2}{A}\sum_{\mathbf{k}}\rho_{\alpha\beta}(\mathbf{k},t) \tag{4}$$

sums over the lateral degrees of freedom and the spin. Even more important is the current density in the growth direction

$$J(z,t)=e\Re\left\{\left\langle\frac{\hat{p}_z-eA_z(z,t)}{m_c(z)}\right\rangle\right\}$$
$$=\frac{e}{\hbar}\sum_{\alpha,\beta}\left[\mathrm{i}W_{\beta\alpha}(z)-\frac{eA_z(z,t)d}{\hbar}F_{\beta\alpha}(z)\right]\bar{\rho}_{\alpha\beta}(t) \tag{5}$$

where $A_z(z,t)$ is the z component of the electrodynamic vector potential (we tacitly assume that its x,y dependence is not of relevance), $e<0$ the charge of the electron, and

$$W_{\beta\alpha}(z)=-\frac{\hbar^2}{2m_c(z)}\left(\varphi^*_\beta(z)\frac{\partial\varphi_\alpha(z)}{\partial z}-\frac{\partial\varphi^*_\beta(z)}{\partial z}\varphi_\alpha(z)\right)$$
$$F_{\beta\alpha}(z)=\frac{\hbar^2}{d}\frac{\varphi^*_\beta(z)\varphi_\alpha(z)}{m_c(z)}.$$

In practice, $J(z,t)$ depends on the number of states used [40], as the completeness relation is not entirely satisfied for a finite basis. However, the spatial average over one period $J(t)$ converges much better and is used for almost all practical purposes. In the presence of an ac field with amplitude $F_{\mathrm{ac}}\cos(\omega t)$, we find a time-periodic current density

$$J(t)=\sum_h J_h\mathrm{e}^{-\mathrm{i}h\omega t}. \tag{6}$$

Here, J_0 is the dc current and the material gain can be obtained via the real part of the dynamical conductivity $\sigma(\omega)$ as

$$G(\omega)=-\frac{\Re\{\sigma(\omega)\}}{c\epsilon_0\sqrt{\epsilon_r}}=-\frac{1}{c\epsilon_0\sqrt{\epsilon_r}}\frac{J_1+J_{-1}}{F_{\mathrm{ac}}}. \tag{7}$$

Thus, the main task is to evaluate the density matrix $\rho_{\alpha\beta}(\mathbf{k},t)$ in the presence of Hamiltonian (2). (In fact, for most purposes it is sufficient to consider $\bar{\rho}_{\alpha\beta}(t)$ as used in many recent density matrix approaches [28]–[30], [32]). In this context, it is a major problem that a standard perturbative treatment of the scattering Hamiltonian(see [24] and [27]) provides negative densities, which is a general feature [41]. An outcome is to modify the scattering terms in a clever way; see, e.g., [32]. However, such an artificial procedure might mask relevant physical effects, where details of the k-distribution matter [38]. Thus, it is of general interest to have a formalism at hand, which is entirely based on a systematic treatment of perturbation theory and avoids unphysical outcomes. This justifies the far more evolved method of NEGF as discussed here.

A. Green's Functions

The formalism of NEGF [42]–[44] can be based on the following Green's functions: the correlation function (or "lesser"

Green's function)

$$G^{<}_{\alpha_1,\alpha_2}(\mathbf{k};t_1,t_2) = \mathrm{i}\langle \hat{a}^{\dagger}_{\alpha_2}(\mathbf{k},t_2)\hat{a}_{\alpha_1}(\mathbf{k},t_1)\rangle \tag{8}$$

which can be regarded as an extension of the density matrix (3) to different times, as well as the retarded and advanced Green functions

$$\begin{aligned} G^{\mathrm{r}}_{\alpha_1,\alpha_2}(\mathbf{k};t_1,t_2) &= -\mathrm{i}\Theta(t_1-t_2)\left\langle\left\{\hat{a}_{\alpha_1}(\mathbf{k},t_1),\hat{a}^{\dagger}_{\alpha_2}(\mathbf{k},t_2)\right\}\right\rangle \\ G^{\mathrm{a}}_{\alpha_1,\alpha_2}(\mathbf{k};t_1,t_2) &= \mathrm{i}\Theta(t_2-t_1)\left\langle\left\{\hat{a}_{\alpha_1}(\mathbf{k},t_1),\hat{a}^{\dagger}_{\alpha_2}(\mathbf{k},t_2)\right\}\right\rangle \\ &= \left[G^{\mathrm{r}}_{\alpha_2,\alpha_1}(\mathbf{k},t_2,t_1)\right]^{*} \end{aligned} \tag{9}$$

where $\{\hat{A},\hat{B}\} = \hat{A}\hat{B}+\hat{B}\hat{A}$ is the anticommutator. Using matrix notation, where matrices like $G_{\alpha\beta}$ are denoted by bold capitals $\mathbf{G}$, differentiating in time provides the equations of motion

$$\begin{aligned} &\left(\mathrm{i}\hbar\frac{\partial}{\partial t_1} - E_k\right)\mathbf{G}^{\mathrm{r/a}}(\mathbf{k};t_1,t_2) - \mathbf{U}(t_1)\mathbf{G}^{\mathrm{r/a}}(\mathbf{k};t_1,t_2) \\ &\quad = \hbar\delta(t_1-t_2)\mathbf{1} + \int\frac{\mathrm{d}t}{\hbar}\mathbf{\Sigma}^{\mathrm{r/a}}(\mathbf{k};t_1,t)\mathbf{G}^{\mathrm{r/a}}(\mathbf{k};t,t_2) \end{aligned} \tag{10}$$

where the matrix $\mathbf{U}$ refers to the k-diagonal part of the Hamiltonian

$$\hat{H}_0 + \hat{H}_{\mathrm{ac}}(t) + \hat{H}_{\mathrm{MF}}(t) = \sum_{\alpha,\beta,\mathbf{k}} U_{\alpha,\beta}(\mathbf{k},t)a^{\dagger}_{\alpha}(\mathbf{k})a_{\beta}(\mathbf{k}) \tag{11}$$

and $E_k = \hbar^2\mathbf{k}^2/2m^*$ is the in-plane energy, where we use the effective mass m^* of the well for simplicity. (Nonparabolicity has been recently considered in the Green's function model of [45].) The self-energies $\mathbf{\Sigma}$ take into account $\hat{H}_{\mathrm{scatt}}$ perturbatively. A similar equation of motion for the lesser Green's function is solved by the Keldysh relation [44]

$$\begin{aligned} &\mathbf{G}^{<}(\mathbf{k};t_1,t_2) \\ &\quad = \int\frac{\mathrm{d}t}{\hbar}\int\frac{\mathrm{d}t'}{\hbar}\mathbf{G}^{\mathrm{r}}(\mathbf{k};t_1,t)\Sigma^{<}(\mathbf{k};t,t')\mathbf{G}^{\mathrm{a}}(\mathbf{k};t',t_2). \end{aligned} \tag{12}$$

In earlier work [34], [35], we focused on the stationary behavior, where the Green's functions only depend on the time difference $t_1 - t_2$. Now, we consider a periodic solution, where the system is allowed to oscillate at a frequency ω and its higher harmonics as a consequence of the time periodicity of $\hat{H}_{\mathrm{ac}}(t)$ (and consequently $\hat{H}_{\mathrm{MF}}(t)$). Setting

$$\mathbf{U}(t) = \sum_h \mathbf{U}_h e^{-\mathrm{i}h\omega t} \tag{13}$$

and using the Fourier decomposition similar to [66]

$$F(\mathbf{k};t_1,t_2) = \frac{1}{2\pi}\int \mathrm{d}E\sum_h \mathrm{e}^{-\mathrm{i}E(t_1-t_2)/\hbar}F_h(\mathbf{k},E)\mathrm{e}^{-\mathrm{i}h\omega t_1} \tag{14}$$

both for self-energies and Green's functions, we find the Dyson equation

$$\begin{aligned} \mathbf{1}\delta_{h,0} &= (E+h\hbar\omega-E_k)\mathbf{G}^{\mathrm{r/a}}_h(\mathbf{k},E) \\ &\quad - \sum_l\left(\mathbf{U}_{h-l} + \Sigma^{\mathrm{r/a}}_{h-l}(\mathbf{k},E+l\hbar\omega)\right)\mathbf{G}^{\mathrm{r/a}}_l(\mathbf{k},E) \end{aligned} \tag{15}$$

and the Keldysh relation

$$\begin{aligned} \mathbf{G}^{<}_h(\mathbf{k},E) &= \sum_{l,l'}\mathbf{G}^{\mathrm{r}}_{h-l-l'}(\mathbf{k},E+(l+l')\hbar\omega) \\ &\quad \times\Sigma^{<}_{l'}(\mathbf{k},E+l\hbar\omega)\mathbf{G}^{\mathrm{a}}_l(\mathbf{k},E) \end{aligned} \tag{16}$$

providing a set of equations to determine the Green's functions for given self-energies within the truncation $|h| \le h_{\max}$. The density matrix is then element wise reproduced by

$$\rho_{\alpha\beta}(\mathbf{k},t) = -\mathrm{i}\int\frac{\mathrm{d}E}{2\pi}\sum_h G^{<}_{\alpha\beta,h}(\mathbf{k},E)\mathrm{e}^{-\mathrm{i}h\omega t} \tag{17}$$

allowing for the evaluation of current and gain as outlined previously.

B. Self-Energies

For the self-energies, we use the self-consistent Born approximation. In Appendix B, details on the scattering processes are given. Here, we restrict to the main concept. For example, for elastic scattering

$$\hat{H}^{\mathrm{elast}}_{\mathrm{scatt}} = \sum_{\alpha\mathbf{k}\beta\mathbf{k}'} V_{\alpha\beta}(\mathbf{k}-\mathbf{k}')a^{\dagger}_{\alpha\mathbf{k}}a_{\beta\mathbf{k}'} \tag{18}$$

we have

$$\Sigma^{</\mathrm{r}}_{\alpha\alpha',h}(E,\mathbf{k}) = \sum_{\beta\beta',\mathbf{k}'}\langle V_{\alpha\beta}(\mathbf{k}-\mathbf{k}')V_{\beta'\alpha'}(\mathbf{k}'-\mathbf{k})\rangle\, G^{</\mathrm{r}}_{\beta\beta',h}(E,\mathbf{k}') \tag{19}$$

where $\langle\rangle$ averages over different realizations of the scattering potential, which restores the spatial homogeneity. For the same reason, the Green's function and self-energy do only depend on the absolute value $|\mathbf{k}|$, which is conveniently expressed by E_k. Then, (19) can be rewritten as

$$\Sigma^{</\mathrm{r}}_{\alpha\alpha',h}(E,E_k) = \sum_{\beta\beta'}\int_0^{\infty}\mathrm{d}E_{k'}\,X_{\alpha\alpha',\beta\beta'}(E_k,E_{k'})G^{</\mathrm{r}}_{\beta\beta',h}(E,E_{k'}). \tag{20}$$

Assuming that $G^{</\mathrm{r}}_{\beta\beta',h}(E,E_{k'})$ peaks at $E_{k'} \approx E-(E_\beta+E_{\beta'})/2 = E'_{\mathrm{typ}}$, and $\Sigma^{</\mathrm{r}}_{\alpha\alpha',h}(E,E_k)$ is mostly needed at $E_k \approx E-(E_\alpha+E_{\alpha'})/2 = E_{\mathrm{typ}}$, we may approximate $X_{\alpha\alpha',\beta\beta'}(E_k,E_{k'})$ by $X_{\alpha\alpha',\beta\beta'}(E_{\mathrm{typ}},E'_{\mathrm{typ}})$, which effectively corresponds to delta-like scattering potentials. Here, we usually chose the value of E such that $\mathrm{Min}\{E_{\mathrm{typ}},E'_{\mathrm{typ}}\} = 1\mathrm{meV}+0.3k_BT$. This *constant k approximation* provides an essential simplification of the numerics, as the self-energies do not depend any longer on $\mathbf{k}$.

The k-dependence of Σ has been shown to be of relevance for longitudinal optical phonon scattering [46], [47], so the results have to be taken with some care. However, with our choice of a level-dependent $E_{\mathrm{typ}}, E'_{\mathrm{typ}}$, we obtained reasonable results for all cases studied so far. A comparison with an exact treatment of intersubband absorption for roughness scattering [48] showed that our approach reproduced the peak very well, but overestimated the absorption tails at high energies, where the momentum dependence of the matrix elements is crucial.

C. Boundary Conditions and Iteration Scheme

We consider a periodic system, where the same sequence of layers is repeated infinitely (d is the length of the period). We tacitly assume that the solution follows this periodicity.[1] Taking into account a potential drop eFd per period due to the external dc bias, this implies in Coulomb gauge

$$\Sigma_{\alpha\beta,h}(E) = \Sigma_{\alpha'\beta',h}(E + eFd)$$

if the states $\alpha'\beta'$ are shifted by one period with respect to $\alpha\beta$. (For the Lorenz gauge, the transformation is more complex and mixes terms with different h.) Thus, it is sufficient to evaluate the self-energies $\Sigma_{\alpha\beta,h}(E)$, where the state α is located in a given central period.

In order to obtain a solution, we determine $G^{\mathrm{r/a}}$ by solving (15) for a given $\Sigma_{\alpha\beta,h}(E)$ taking into account the states of the central period as well as N_{per} periods on either side. Then, we determine $G^{<}$ via (16). Based on these Green's functions, we calculate the new self-energies $\Sigma_{\alpha\beta,h}(E)$ for α being located in the central period. If the new self-energies do not agree with the starting ones, we choose new values by the Broyden algorithm [49] and repeat the iteration. Convergence with an accuracy of typically 5×10^{-4} (or even smaller) is obtained after 30–100 iterations. Finally, the quantities of interest are evaluated from the Green's functions in the central region. By increasing $h_{\max}$ and N_{per}, the quality of the respective truncations can be verified.

D. Meanfield

The electron charge density $\rho_{el}(z,t)$ is directly obtained from the density matrix. Solving Possion's equation together with the doping density for periodic boundary conditions $\phi(z) = \phi(z + d)$ provides the potential $\phi(z,t)$, which determines $\hat{H}_{\mathrm{MF}}(t)$.

III. Comparison with Analytical Results for Superlattices

Semiconductor superlattices represent the simplest periodic structures and provide an excellent testing ground for the code. Here, we consider the structure of [50], which has a rather large miniband width $\Delta = 26.4$ meV. Fig. 1(a) shows the calculated current-bias relation using the nominal system parameters and neglecting interface roughness (which should be small, as wells and barriers are binary systems and carefully chosen to consist of an integer number of monolayers). All calculations are done at a lattice temperature of 300 K and we use $N_{\mathrm{per}} = 3$. The peak height and position agree reasonably well with the measured characteristics. The data can be fitted very well by a simple expression $J(F) = 2J_{\max} eFd\Gamma/((eFd)^2 + \Gamma^2)$ based on the Esaki–Tsu characteristics [51] where $\Gamma = 18$ meV is the average scattering rate multiplied by $\hbar$. This is expected, as the miniband model is good for $\Delta \gtrsim \Gamma$ and $\Delta \gtrsim |eFd|$ [52]. Indeed, minor deviations are seen for larger bias drops Fd per period. Following the work of [53], analytical expressions can

[1]Recently nonperiodic solutions have been discussed in [65]. As this effect appears in direct connection to the injecting contact, it is not clear whether this is of relevance far inside the periodic structure.

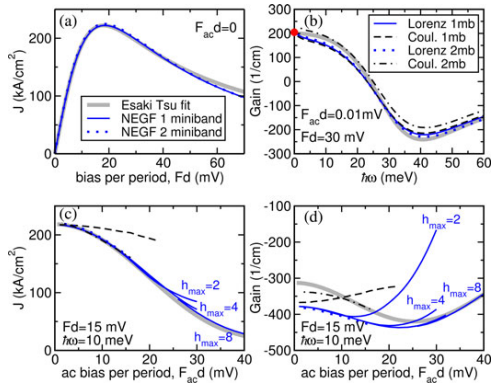


Fig. 1. Results of the NEGF simulation for the superlattice of [50]. The thick gray line gives the analytic solution of the simplified miniband model for comparison. The NEGF solutions are displayed for one miniband (mb) (full/dashed line) and two minibands (dotted/dash–dotted line) for Lorenz/Coulomb gauge, respectively. In panels (c) and (d), different choices of $h_{\max}$ are shown for Lorenz gauge and one miniband, while all other curves are solutions with sufficiently high $h_{\max}$. In panel (b), the gain value at $\omega \to 0$ obtained from the NEGF conductance in panel (a) is marked by a dot.

be obtained within the Esaki–Tsu model even for a finite ac field (see [54, Section VI.2] for details), which can be used for comparison.

In Fig. 1(b) the gain is shown for a weak ac field at a dc bias dropping over one period of $Fd = 30$ mV. Here, we are in the region of negative differential conductivity of the superlattice, and consequently, gain occurs for $\hbar\omega \lesssim |eFd|$. The NEGF solutions are in reasonable agreement with the simple model, which can be seen as a validation for both approaches. We note that the choice of gauge does not affect the result much. However, it can be seen that the inclusion of the second subband (i.e., letting α run over two states per well) changes the result slightly for the Coulomb gauge. (The current–voltage characteristics are not affected as the second miniband, located 300 meV above the first miniband, is essentially empty.) Finally, we see that the gain approaches the value $G_0 = -\frac{\sigma_{\mathrm{dc}}}{c\epsilon_0\sqrt{\epsilon_r}} = 205/\mathrm{cm}$ for $\omega \to 0$, where the dc conductivity σ_{dc} is obtained from the slope of the current–voltage characteristics.

Fig. 1(c) shows the dc current at $Fd = 15$ mV as a function of the ac-field strength. We find that the current drops as a result of dynamical localization [55]. The result in Lorenz gauge and the analytic miniband model agree very well. For the Coulomb gauge, the agreement is only good if two minibands are taken into account. For the Lorenz gauge, we have shown simulations with different numbers $h_{\max}$. As expected, with increasing $eF_{\mathrm{ac}}d$ a larger number of h-components is needed in order to reach convergence.

Finally, Fig. 1(d) shows the gain for $\hbar\omega = 10$ meV at a dc bias of $Fd = 15$ mV as a function of the ac-field strength. The behavior is less dramatic. Here, the agreement between the NEGF solution and the analytic miniband model is only of qualitative

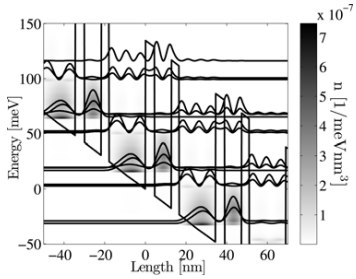


Fig. 2. Energetically and spatially resolved electron density $\rho(E,z)$ and the Wannier–Stark states of the QCL in [56] for a bias drop per period $Fd = 48.25$ mV.

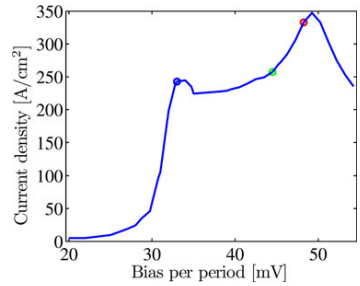


Fig. 3. Current-bias relation for the QCL in [56]. The circles mark the operating points where the gain is simulated and the conductivity is calculated.

nature. A better agreement would be surprising, as the high fields $F + F_{\rm ac}$ in combination with the high frequency ω are limiting the validity of the miniband model. The disagreement between the gauges is stronger and again, the result in Coulomb gauge depends on the number of minibands included.

These findings indicate that the NEGF model provides reasonable results in Lorenz gauge. In Coulomb gauge, the results depend strongly on the number of states included, even, if they are not directly involved. This may be attributed to the A_z term in the current (5). Technically, the equivalence between the two gauges requires the completeness relation for the basis used. For any finite basis, the completeness relation is only approximate, and thus, the equivalence between the approaches is not guaranteed. The data shown here indicate that the Lorenz gauge is more reliable.

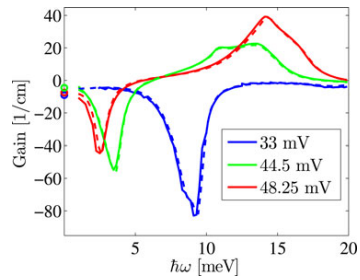


Fig. 4. Gain spectrum for $F_{\rm ac}d = 1$ mV at different biases for the QCL in [56] at two different gauges (solid lines: Lorenz; dashed lines: Coulomb). Circles mark the gain coefficients at $\omega \to 0$ calculated from the conductivity.

IV. QCL Results

A. Characterization of a QCL

In order to test the validity of our model, we present simulation results on the QCL structure of [56]. The structure is shown in Fig. 2, with the Wannier–Stark states and the energetically and spatially resolved electron density $\rho(E,z)$, directly found from the Green's function by

$$\rho(E,z) = \frac{2e}{2\pi i A}\sum_{\mathbf{k}}\sum_{\alpha\beta} G^{<}_{\alpha\beta,h=0}(E,\mathbf{k})\varphi^{*}_{\beta}(z)\varphi_{\alpha}(z). \quad (21)$$

Fig. 3 shows the current-bias relation without an external ac field at a lattice temperature of 50 K. The peak current $\sim$350 A/cm^2 is reached at 49 mV/period, which corresponds well with the experimentally observed peak at lasing at about 48 mV/period, where the current for a nonlasing device was measured to be $\sim$ 400 A/cm^2. The relatively large peak around 34 mV/period is of much smaller magnitude in the measurements. It corresponds to tunneling over two barriers. We think that dephasing is underestimated in our model here, probably due to the neglect of electron–electron scattering.

The simulated gain spectra for two gauges of a weak applied ac field are shown in Fig. 4. As discussed previously, the gauge transformations should in principle leave the results unchanged, but the incompleteness of the basis used gives rise to minor numerical differences. However, the results for different gauges agree fairly well for low ac bias as in the case for the superlattice discussed previously.

Again, we validate our model by comparing the theoretical value of the gain coefficients with the simulated gain spectra. In Fig. 4, the gain coefficients at $\omega \to 0$ are marked with circles. These have been calculated from the slope of the *IV* curve at the bias points indicated in Fig. 3 (using a quadratic fit) to be -9.09, -4.59, and -7.523/cm for increasing bias. The gain at low ω can be seen to approach the calculated values. Experimentally, lasing is observed at $\hbar\omega = 11.5$ and 13.3 meV at lower bias and at $\hbar\omega = 16$ meV for higher bias, which agrees well with our gain spectrum.

B. Integrated Absorption

The physics of gain in a QCL is essentially the physics of intersubband absorption and stimulated emission. The total absorption of all states in the conduction band, which we will here denote the integrated absorption, can be estimated

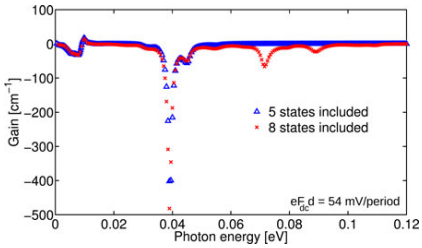


Fig. 5. Simulated absorption on a structure presented in [59]. Results from two different simulations are compared, one using 5 states per period and one using 8 states per period. As a reference, the conduction band offset in this structure is about 135 meV.

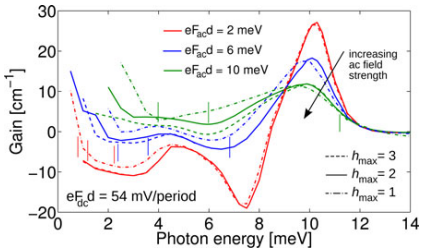


Fig. 6. Simulations with different $h_{\max}$ for three ac field strengths. An estimate of the critical frequency for each calculation is shown with a line.

following [57] and [58] to

$$\int_0^\infty d\omega\, \alpha_{\text{intersubband}}(\omega) = n_{\text{av}} \frac{\pi e^2}{2m^* c\epsilon_0 \sqrt{\epsilon_r}} \tag{22}$$

with $n_{av} = \sum_i n_i/d$ as the average 3-D carrier distribution and $\sqrt{\epsilon_r}$ again the refractive index. Note that this expression depends neither on the structure nor on the applied bias.

The constant property of the integrated absorption provides an excellent opportunity to test the model. To do this, we investigate a structure presented in [59]. This is a four-well QCL where the main transport properties can be retrieved from simulations by taking the first five states in each period into account. For this structure, simulations were carried out over a wide range of frequencies, and the results are shown in Fig. 5. The integrated absorptions for the two cases are found via numerical integration over the frequency interval to yield 2.23 eV/cm (1.62 eV/cm) when including 8 (5) states. Here, the errors in the numerical calculation can be severe as the number of points in the critical regions are few, and then exaggerated by the numerical integration. The simulated integrated absorption should be compared to the theoretical value of 2.41 eV/cm from (22).

As seen in the figure, the inclusion of states not contributing to transport is crucial in order to reproduce the correct integrated absorption. It is worth noting that the difference between a noncomplete basis set of five states and one almost complete of eight states is large. This shows again the importance of the completeness relation in the model, as we see here that an incomplete basis does not nearly reproduce the theoretically predicted result. The ninth state is in the continuum, and its importance is so far neglected.

Looking closer at (22), we note that the effective mass m^* is in fact z-dependent, as it is not the same in the barriers as in the wells. However, absorption will occur where the electrons are localized, as this is the argument for approximating the mass in these calculations with the effective mass of the wells of the heterostructure. Therefore, we expect the theoretical estimate to slightly overshoot the real value. Although the electrons are most probably found in one of the wells, we would find a certain percentage in the barriers, effectively raising the average mass and thus contributing to lower integrated absorption. This should also be taken into account when interpreting the results given previously.

C. High Intensity

Recently, we showed the relevance of nonlinear effects on the gain spectrum of a QCL [39]. In this context, it is important to study the convergence of the solution depending on the truncation parameter $h_{\max}$. Obviously, for low intensity of the ac field, a sharp truncation is very reasonable but when the photon energy is comparable to the ac field strength higher harmonics are needed in order to expand the space of allowed response functions of the system.

This criterion can be expressed as $eF_{\text{ac}}d \lesssim h_{\max}\hbar\omega$, for $N_{\text{per}} = 1$ used here. More orders of the response function must be taken into account when calculating with low photon energies $\hbar\omega$. This can be understood as including higher orders of absorption phenomena, as several photons of low energy can give the same absorption effect in the electronic system as one photon at the exact transition energy. These effects are naturally more important as intensity is increased.

The lowest truncation possible is $h_{\max} = 0$ which cannot treat any external light field at all but gives the stationary behavior of the system. Second lowest would then be $h_{\max} = 1$, where response at the fundamental frequency is seen. This gives the main part of the system dynamics, making it possible to retrieve quantities such as gain and absorption. Higher orders $h_{\max} > 1$ provide higher accuracy as more dynamical effects are taken into account as more Fourier components of the Green's function enter the system. This is important for low photon energies as discussed before. For terahertz QCLs, which operate at typical photon energies of 10 meV, large parts of the gain spectra are actually in the regime where the higher harmonics play an important role. In order to illustrate the importance, simulations of a QCL described in [59] are displayed in Fig. 6. The main effect, as discussed in detail in [39], is the reduction of gain with the intensity of the optical field due to bleaching.

The simulations show that it is not trivial to approach the low-frequency part of the spectra with a finite ac field strength. At the energy scale of the transport dynamics, the photon energies are now small, and a number of photons would be required in order to affect the system on the same premises as in the cases

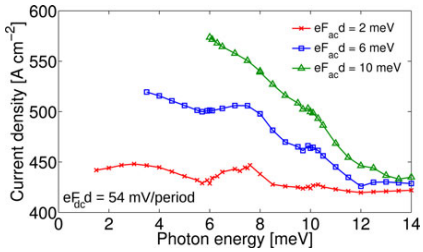


Fig. 7. Current at the same points as in Fig. 6 for $h_{\max} = 2$ plotted for the three different intensities. Data only taken at points where we estimate the simulations to be accurate.

of higher photon energy. This is why $h_{\max}$ is such an important parameter in the results shown in Fig. 6.

Here, the simulations with $h_{\max} = 3$ are the best cases and closest to the would-be exact calculation. For this system, it is on the limit of what is at the moment possible for us to do and only a few points are shown for the $h_{\max} = 3$ case, plotted as dashed lines in the figure. The simulations for different cases of $h_{\max}$ eventually coincide at higher photon energies as expected, early for low ac field strength and later for high ac field strength. It is clear from the results that the higher harmonics are of importance at operating intensities above $eF_{\rm ac}d = 6$ meV. In the plot we also show, for each ac field strength and $h_{\max}$, an estimation of the critical frequency. At this frequency, one should start being restrictive in trusting the results of the simulations.

In a QCL under operating conditions, the laser field will drive the current as the frequency of the photons is matched to the energy difference of the upper and lower laser state, stimulating emission and thus destroying the inversion of the carriers. This will result in a higher current through the system, and this can be seen in Fig. 7 where current is plotted at the same points as in Fig. 6 for $h_{\max} = 2$. The general trend is that for a given ac field strength the current increases with decreasing photon energies. This can be attributed to the fact that with decreasing photon energy, the number of photons increases for a fixed intensity. Therefore, more stimulated transitions occur, which provide extra channels for the electron transport.

Characteristic features can also be seen at the absorption peak at $\hbar\omega \sim 7$ meV and at the gain peak at $\hbar\omega \sim 10$ meV, compared with Fig. 6. We attribute the first peak in current to a redistribution of carriers from localized states to more mobile ones. At the gain peak at $\hbar\omega \sim 10$ meV, the stimulated emission provides an enhanced current along the common current path through the structure. This is commonly seen experimentally as a change of conductance at the onset of lasing.

Finally, above $\hbar\omega > 13$ meV, the current is hardly affected by the laser field. This is just the region without any significant absorption/gain, as can be seen in Fig. 5.

V. Quantum Cascade Detector

A suitable application of our model other than the QCL is the QCD [60], which is an alternative to the quantum well

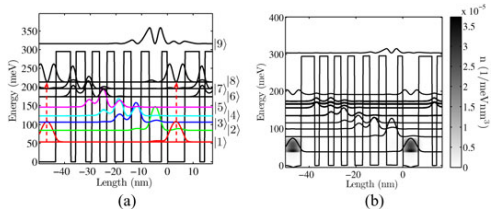


Fig. 8. (a) Basis states of the QCD, numbered from bottom to top. (b) Maxima of the density of states and the electron density.

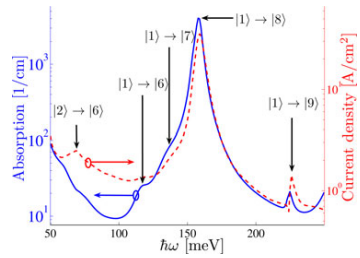


Fig. 9. Absorption (blue) and response current (red) as functions of the applied ac field frequency. The optical transitions causing the absorption and response peaks are indicated.

infrared photodetector (QWIP) for detecting infrared radiation. In contrast to the QWIP and many other photodetectors, the QCD operates at zero (or very small) bias, and therefore has minimal dark currents, in addition to having a narrow detection wavelength width.

We will now study the QCD presented in [61], shown in Fig. 8(a) with the calculated basis states. In Fig. 8(b), we show the density of states (lines) at the energies where it has maxima. The reason for doing this is that the Wannier–Stark states, conventionally used as basis states that here contain mean field, are not defined at zero bias, and that the Wannier states do not include these effects. The density of states allows for determining which transitions are most probable with respect to populated and empty states.

Light is detected when an electron in the ground state of the widest, heavily doped well, $|1\rangle$, absorbs a photon with an energy matching the transition $|1\rangle \to |8\rangle$. The electron will thereby occupy state $|8\rangle$. Tunneling and scattering then allow for a cascading current from $|8\rangle$ down to the ground state of the neighboring period, to the right in Fig. 8. This gives an overall current flowing through the device (consisting of 50 periods or so) that is detected at the terminals.

Fig. 9 shows the simulated absorption and the response current arising from the photon-absorbing carriers. This current is a second-order effect, since it arises from carriers excited by the applied ac field. The main peaks are located at (going from low to high $\hbar\omega$) 69, 119, 138, 159, and 226 meV. Using the labeling in Fig. 8 (a) and taking the energy differences between the density of states maxima in Fig. 8 (b), the response peaks are

attributed to the transitions shown in Fig. 9. The QCD current response agrees almost exactly with the absorption peaks.

The experimentally measured QCD [61] showed the same structure for both absorption and response current as the presented simulations. The simulated peaks are shifted to higher frequencies by about 5 meV (more for larger transitions, and less for smaller transitions) with respect to the measured peaks. We attribute this to the neglect of nonparabolicity in our current scheme. In general, nonparabolicity reduces the energy of high energy states.

VI. Conclusion

We have presented our implementation of NEGF for the simulations of heterostructure devices in detail. Our model allows for a reliable determination of the current-field relation in most cases. However, some problems exist with extra peaks due to long range tunneling, which indicates that some dephasing channels are still missing. The main advance compared to previous implementations is the possibility to treat the radiation field even beyond linear response. This allows for studying a variety of important effects such as dynamical localization in superlattices, gain bleaching in quantum cascade lasers, and photon-assisted currents in QCDs. For small optical fields, the choice of the gauge does not matter much. However, the Coulomb gauge seems to require the inclusion of higher minibands in a superlattice even if they do not contribute explicitly. This effect is particularly strong for large optical fields.

Appendix A
Angle Averaged Scattering Matrix Elements

Using the self-consistent Born approximation in (19), we already argued for the use of only the absolute value $|\mathbf{k}|$ for the potentials. This means that when transforming the self-energies into a function of $E_k = \hbar^2\mathbf{k}^2/2m^*$ and $E_{k'}$ instead of $\mathbf{k}$ and $\mathbf{k}'$, we have to write the self-energy as

$$\Sigma^{</\mathrm{ret}}_{\alpha\alpha',h}(E,E_k) = \sum_{\beta\beta'}\int_0^\infty \mathrm{d}E_{k'}\frac{A}{(2\pi)^2}\frac{m^*}{\hbar^2} \times\int_0^{2\pi}\mathrm{d}\phi\langle V_{\alpha\beta}(E_k,E_{k'},\phi)V_{\beta'\alpha'}(E_k,E_{k'},\phi)\rangle \times G^{</\mathrm{ret}}_{\beta\beta',h}(E,E_{k'})$$

where we now have converted the sum over $\mathbf{k}'$ into a 2-D integral of E'_k and the angle $\phi = \angle(\mathbf{k},\mathbf{k}')$. Now, we identify the second-rank tensor $X_{\alpha\alpha',\beta\beta'}$ as the angle averaged part of the scattering matrix elements as a function of $E_k, E_{k'}$ in the following way:

$$\Sigma^{</\mathrm{r}}_{\alpha\alpha',h}(E,E_k) = \sum_{\beta\beta'}\int_0^\infty \mathrm{d}E_{k'}\,X_{\alpha\alpha',\beta\beta'}(E_k,E_{k'}) \times G^{</\mathrm{r}}_{\beta\beta',h}(E,E_{k'}) \quad (23)$$

where also the constants are included in $X_{\alpha\alpha',\beta\beta'}$. As discussed previously, these tensors are evaluated at certain typical energies E_{typ} and $E_{\mathrm{typ}'}$ which allow us to take the tensors outside the integral leaving an expression where we only need the k-integrated Green's function for the evaluation of the self-energies.

Appendix B
Details for the Scattering Matrix Elements Used

Here, we describe in detail how the different scattering processes are included in the formalism. The alloy scattering matrix elements are evaluated to describe the procedure of these calculations, and then the other scattering mechanisms are shown for completeness. These can be derived by the interested reader in the spirit of [34] with only minor changes with respect to the formalism presented here.

A. Elastic Scattering

Alloy scattering is treated here as the disturbing potential arising from alternating one of the atoms in a binary alloy, and the treatment follows [62]. The strength of this potential will depend on the mixing parameter x and the difference in the conduction-band edge ΔV between the compounds with $x = 0$ and $x = 1$. The potential from one fluctuating species located at $\vec{r}_i = \mathbf{r}_i + z_i\vec{e}_z$ is given by

$$V^i(\vec{r}) = C^x_i\,\delta(\vec{r}-\vec{r}_i)$$

where C^x_i is a random variable such that

$$\langle C^x_i\rangle_{\mathrm{alloy}} = 0$$
$$\langle C^x_i C^x_j\rangle_{\mathrm{alloy}} = \delta_{ij}x(1-x)(\Delta V V_{\mathrm{mol}})^2$$

and V_{mol} is the volume of one dimer (i.e., a fourth of the cubic unit cell for zincblende lattices). These properties determine the averages for the matrix elements in our basis, which read

$$\langle V^i_{\alpha\mathbf{k}_\alpha,\beta\mathbf{k}_\beta}V^j_{\gamma\mathbf{k}_\gamma,\delta\mathbf{k}_\delta}\rangle_{\mathrm{alloy}} = \frac{\delta_{ij}x(1-x)(\Delta V V_{\mathrm{mol}})^2}{A^2} \times\varphi^*_\alpha(z_i)\varphi_\beta(z_i)\varphi^*_\gamma(z_i)\varphi_\delta(z_i)\mathrm{e}^{-\mathrm{i}(\mathbf{k}_\alpha-\mathbf{k}_\beta+\mathbf{k}_\gamma-\mathbf{k}_\delta)\cdot\mathbf{r}_i} \quad (24)$$

showing that this is a local scattering process. Summing over all unit cells, effectively turning the sum into an integral

$$\sum_i \to \frac{1}{V_{\mathrm{mol}}}\mathrm{d}^3 r_i$$

gives

$$\langle V^i_{\alpha\mathbf{k}_\alpha,\beta\mathbf{k}_\beta}V^j_{\gamma\mathbf{k}_\gamma,\delta\mathbf{k}_\delta}\rangle_{\mathrm{alloy}} = \frac{V_{\mathrm{mol}}}{A}\delta_{\mathbf{k}_\alpha-\mathbf{k}_\beta,\mathbf{k}_\delta-\mathbf{k}_\gamma}\int\mathrm{d}z\,\varphi^*_\alpha(z)\varphi_\beta(z)\varphi^*_\gamma(z)\varphi_\delta(z) \times x(z)(1-x(z))\Delta V^2(z) \quad (25)$$

which provides us with the squared matrix element used in (19) for the alloy scattering. The ϕ-integral provides a factor of 2π giving the final expression

$$X_{\alpha\alpha',\beta\beta'} = \frac{V_{\mathrm{mol}}m^*}{2\pi\hbar^2}\int\mathrm{d}z\,\varphi^*_\alpha(z)\varphi_\beta(z)\varphi^*_{\beta'}(z)\varphi_{\alpha'}(z) \times x(z)(1-x(z))\Delta V^2(z). \quad (26)$$

Now we proceed with the tensors for the other elastic scattering matrix elements. For interface roughness scattering, we assume the correlation functions

$$\langle \xi_i(\mathbf{r})\rangle = 0 \quad \langle \xi_i(\mathbf{r})\xi_j(\mathbf{r}')\rangle = \delta_{ij}\eta^2 \mathrm{e}^{-|\mathbf{r}-\mathbf{r}'|/\lambda_i}$$

for the deviation $\xi_i(\mathbf{r})$ of the interface i from its nominal position. Here, η is the average (RMS) height of the interface roughness and λ lateral correlation. Then, we obtain approximately (analogously to [34, Appendix B])

$$X^{\mathrm{rough}}_{\alpha\alpha'\beta\beta'}(E_k, E_{k'}) = \sum_j \frac{\Delta E_c^2 \eta_j^2}{2E_{\lambda_j}} \frac{1}{(a_j - b_j)\sqrt{a_j + b_j}} \times \varphi^*_\alpha(z_j)\varphi_\beta(z_j)\varphi^*_{\beta'}(z_j)\varphi_{\alpha'}(z_j) \quad (27)$$

where

$$a = 1 + \frac{E_k + E_{k'}}{E_\lambda} \qquad b = 2\frac{\sqrt{E_k E_{k'}}}{E_\lambda}$$

ΔE_c is the conduction band offset in energy, η is the average (RMS) height of the interface roughness, and $E_\lambda = \hbar^2/2m^*\lambda^2$.

Scattering at ionized dopants is treated as scattering from a number of delta-doping layers at positions z_i and areal density N^i_{2D}:

$$X^{\mathrm{imp}}_{\alpha\alpha',\beta\beta'}(E_k, E_{k'}) = \sum_i N^i_{2D} \frac{e^4}{16\pi\epsilon_s^2\epsilon_0^2} \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi \left\{ \frac{1}{E_k + E_{k'} - 2\sqrt{E_k E_{k'}}\cos\phi + E_\lambda} \times \int \mathrm{d}z_1\, \varphi^*_\alpha(z_1)\varphi_\beta(z_1) e^{-q(\phi)|z_1 - z_i|} \times \int \mathrm{d}z_2\, \varphi^*_{\beta'}(z_2)\varphi_{\alpha'}(z_2) e^{-q(\phi)|z_2 - z_i|} \right\}$$

where $\hbar^2 q^2(\phi)/2m^* = E_k + E_{k'} - 2\sqrt{E_k E_{k'}}\cos\phi + E_\lambda$ and $E_\lambda = \hbar^2\lambda^2/2m^*$. The inverse screening length λ is determined by an interpolation between Debye and Tomas–Fermi screening for the average 3-D electron density. Even for very inhomogeneous distributions of the electrons along the period, this approximation is good as long as $\lambda d \lesssim 1$ [63].

B. Inelastic Scattering

In our current implementation, inelastic scattering processes are only due to the phonons. For the longitudinal optical phonons, we assume a constant frequency ω_{LO}. We consider the perturbation potential for the electron due to phonons with wave vector $\vec{q} = \mathbf{q} + q_z\vec{e}_z$:

$$V_{\mathrm{phon}}(\vec{r}, t) = \sum_{\mathbf{q}q_z} g(\vec{q}) \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r} + \mathrm{i}q_z z} \left[b(\vec{q}, t) + b^\dagger(-\vec{q}, t)\right] \quad (28)$$

where $b(\vec{q}, t)$ and $b^\dagger(\vec{q}, t)$ are the bosonic annihilation and creation operators in the Heisenberg picture. For the polar interaction of the longitudinal optical phonons, we have the standard expression

$$|g(\vec{q})|^2 = g(\vec{q})g(-\vec{q}) = \frac{1}{AL}\frac{e^2}{|\vec{q}|^2}\frac{\hbar\omega_{LO}}{2\epsilon_0}\left(\frac{1}{\epsilon(\infty)} - \frac{1}{\epsilon(0)}\right)$$

where L is the normalization length for the phonons in the z-direction. Essentially following [44] and [64], we find in Born approximation (see also [54, Section IV.1.2])

$$\Sigma^<_{\alpha\alpha'}(E, \mathbf{k}) = \sum_{\beta\beta'}\sum_{q_z\mathbf{q}} |g(\vec{q})|^2 M_{\alpha\beta}(q_z) M_{\beta'\alpha'}(-q_z) \times \Big[G^<_{\beta\beta'}(E - \hbar\omega_{\mathrm{LO}}, \mathbf{k} - \mathbf{q}) n_B + G^<_{\beta\beta'}(E + \hbar\omega_{\mathrm{LO}}, \mathbf{k} - \mathbf{q})(n_B + 1)\Big]$$

where n_B is the Bose distribution for the optical phonons at lattice temperature and

$$M_{\alpha\beta}(q_z) = \int \mathrm{d}z\, \varphi^*_\alpha(z)\mathrm{e}^{\mathrm{i}q_z z}\varphi_\beta(z).$$

The retarded self-energy reads

$$\Sigma^r_{\alpha\alpha'}(E, \mathbf{k}) = \sum_{\beta\beta'}\sum_{q_z\mathbf{q}} |g(\vec{q})|^2 M_{\alpha\beta}(q_z) M_{\beta'\alpha'}(-q_z) \times \Big[G^r_{\beta\beta'}(E - \hbar\omega_{\mathrm{LO}}, \mathbf{k} - \mathbf{q})(n_B + 1) + G^r_{\beta\beta'}(E + \hbar\omega_{\mathrm{LO}}, \mathbf{k} - \mathbf{q}) n_B + \frac{1}{2}G^<_{\beta\beta'}(E - \hbar\omega_{\mathrm{LO}}, \mathbf{k} - \mathbf{q}) - \frac{1}{2}G^<_{\beta\beta'}(E + \hbar\omega_{\mathrm{LO}}, \mathbf{k} - \mathbf{q})\Big]$$

where a term containing a principal value integral [(127) of [54]] has been neglected. Using the fact that the Green's functions and self-energies do only depend on $|\mathbf{k}|$, the individual terms can be rewritten in the form

$$\Sigma_{\alpha\alpha'}(E, E_k) = \sum_{\beta\beta'}\int_0^\infty \mathrm{d}E_{k'} X^{\mathrm{opt}}_{\alpha\alpha'\beta\beta'}(E_k, E_{k'}) \times G_{\beta\beta'}(E \pm \hbar\omega_{\mathrm{LO}}, E_{k'}) \times [n_B \text{ or } (n_B + 1)] \quad (29)$$

with

$$X^{\mathrm{opt}}_{\alpha\alpha',\beta\beta'}(E_k, E_{k'}) = \frac{e^2\hbar\omega_{LO}}{16\pi^2\epsilon_0}\left(\frac{1}{\epsilon(\infty)} - \frac{1}{\epsilon(0)}\right) \times \int \mathrm{d}q_z \frac{M_{\alpha\beta}(q_z) M_{\beta'\alpha'}(-q_z)}{\sqrt{(\hbar^2 q_z^2/2m + E_k + E_{k'})^2 - 4E_k E_{k'}}}. \quad (30)$$

As in the elastic case, we apply the constant k approximation setting $E_k \approx E - (E_\alpha + E_{\alpha'})/2 = E_{\mathrm{typ}}$ and $E_{k'} \approx E \pm \hbar\omega_{\mathrm{LO}} - (E_\beta + E_{\beta'})/2 = E'_{\mathrm{typ}}$ in the Green's functions with the corresponding energy argument $E \pm \hbar\omega_{\mathrm{LO}}$. Again, the typical energy E is chosen such that $\mathrm{Min}\{E_{\mathrm{typ}}, E'_{\mathrm{typ}}\} = 1\mathrm{meV} + 0.3k_BT$.

For the deformation potential scattering with longitudinal acoustic phonons, we have

$$|g(\vec{q})|^2 = \Xi_c^2 \frac{\hbar\omega(\vec{q})}{2\rho_m AL c_L^2}$$

in (28), where Ξ_c is the deformation potential of the conductance band, ρ_m is the mass density of the crystal, and c_L is

the longitudinal sound velocity. In general, these matrix elements are much smaller than the ones for optical phonon scattering or impurity scattering. However, we want to take acoustic phonons into account at least phenomenologically, so that there is some energy dissipation if optical phonon emission is energetically not allowed for. We make the essential simplifications to set $\hbar\omega(\vec{q}) \to \hbar\omega_{\rm ac} = k_B T$ in the energy arguments of the Green's functions and $n_B(\vec{q}) \to n_B(\hbar\omega_{\rm ac}) * \hbar\omega_{\rm ac}/\hbar\omega(\vec{q})$ as well as $n_B(\vec{q}) + 1 \to (n_B(\hbar\omega_{\rm ac}) + 1) * \hbar\omega_{\rm ac}/\hbar\omega(\vec{q})$ to mimic the q dependence of the occupation functions. This prefactor actually cancels with $\omega(\vec{q})$ in $|g(\vec{q})|^2$ and we find an expression completely analogously to (29) with

$$X^{\rm ac}_{\alpha\alpha'\beta'\beta} = \hbar\omega_{\rm ac} \frac{\Xi_c^2 m_e m^*}{4\pi\hbar^2 \rho_m c_L^2} \frac{1}{L} \sum_{q_z} M_{\alpha\beta}(q_z) M_{\beta'\alpha'}(-q_z). \tag{31}$$

This expression does not depend on E_k and $E_{k'}$, so the constant k approximation is already included in the aforementioned approximations.

ACKNOWLEDGMENT

The authors would like to thank S.-C. Lee and R. Nelander for their earlier contributions to the model development.

REFERENCES

[1] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser," *Science*, vol. 264, pp. 553–556, 1994.

[2] R. F. Curl, F. Capasso, C. Gmachl, A. A. Kosterev, B. McManus, R. Lewicki, M. Pusharsky, G. Wysocki, and F. K. Tittel, "Quantum cascade lasers in chemical physics," *Chem. Phys. Lett.*, vol. 487, pp. 1–18, 2010.

[3] A. Müller and J. Faist, "The quantum cascade laser: Ready for take-off," *Nat. Photon.*, vol. 4, p. 291, 2010.

[4] F. Capasso, "High-performance midinfrared quantum cascade lasers," *Opt. Eng.*, vol. 49, 111102 (9 pp.), 2010.

[5] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, "Terahertz semiconductor-heterostructure laser," *Nature*, vol. 417, pp. 156–159, 2002.

[6] J. Darmo, V. Tamosiunas, G. Fasching, J. Kröll, K. Unterrainer, M. Beck, M. Giovannini, J. Faist, C. Kremser, and P. Debbage, "Imaging with a terahertz quantum cascade laser," *Opt. Exp.*, vol. 12, pp. 1879–1884, 2004.

[7] H.-W. Hübers, S. G. Pavlov, H. Richter, A. D. Semenov, L. Mahler, A. Tredicucci, H. E. Beere, and D. A. Ritchie, "High-resolution gas phase spectroscopy with a distributed feedback terahertz quantum cascade laser," *Appl. Phys. Lett.*, vol. 89, 061115 (3 pp.), 2006.

[8] B. S. Williams, "Terahertz quantum-cascade lasers," *Nature Phot.*, vol. 1, pp. 517–525, 2007.

[9] M. Lee and M. C. Wanke, "APPLIED PHYSICS: Searching for a solid-state terahertz technology," *Science*, vol. 316, pp. 64–65, 2007.

[10] M. Periera and O. Shulika, Eds., *Terahertz and Mid Infrared Radiatio*. (NATO Science for Peace and Security Series B Physics and Biophysics):. Dordrecht, The Netherlands: Springer-Verlag, 2011.

[11] F. Capasso, J. Faist, and C. Sirtori, "Mesoscopic phenomena in semiconductor nanostructures by quantum design," *J. Math. Phys.*, vol. 37, pp. 4775–4792, 1996.

[12] D. Paulavičius, V. Mitin, and M. A. Stroscio, "Hot-optical-phonon effects on electron relaxation in an AlGaAs/GaAs quantum cascade laser structure," *J. Appl. Phys.*, vol. 84, pp. 3459–3466, 1998.

[13] S. Slivken, V. I. Litvinov, M. Razeghi, and J. R. Meyer, "Relaxation kinetics in quantum cascade lasers," *J. Appl. Phys.*, vol. 85, pp. 665–671, 1999.

[14] P. Hyldgaard and J. W. Wilkins, "Electron–electron scattering in far-infrared quantum cascade lasers," *Phys. Rev. B*, vol. 53, pp. 6889–6892, 1996.

[15] P. Harrison, "The nature of the electron distribution functions in quantum cascade lasers," *Appl. Phys. Lett.*, vol. 75, pp. 2800–2802, 1999.

[16] C. Becker, C. Sirtori, H. Page, A. Robertson, V. Ortiz, and X. Marcadet, "Influence of confined phonon modes on the thermal behavior of AlAs/GaAs quantum cascade structures," *Phys. Rev. B*, vol. 65, 085305 (4 pp.), 2002.

[17] F. Compagnone, A. Di Carlo, and P. Lugli, "Monte Carlo simulation of electron dynamics in superlattice quantum cascade lasers," *Appl. Phys. Lett.*, vol. 80, pp. 920–922, 2002.

[18] K. Donovan, P. Harrison, and R. W. Kelsall, "Comparison of the quantum efficiencies of interwell and intrawell radiative transitions in quantum cascade lasers," *Appl. Phys. Lett.*, vol. 75, pp. 1999–2001, 1999.

[19] D. Indjin, P. Harrison, R. W. Kelsall, and Z. Ikonic, "Self-consistent scattering theory of transport and output characteristics of quantum cascade lasers," *J. Appl. Phys.*, vol. 91, pp. 9019–9026, 2002.

[20] S. Tortora, F. Compagnone, A. Di Carlo, P. Lugli, M. T. Pellegrini, M. Troccoli, and G. Scamarcio, "Theoretical study and simulation of electron dynamics in quantum cascade lasers," *Physica B*, vol. 272, pp. 219–222, 1999.

[21] R. C. Iotti and F. Rossi, "Carrier thermalization versus phonon-assisted relaxation in quantum cascade lasers: A Monte Carlo approach," *Appl. Phys. Lett.*, vol. 78, pp. 2902–2904, 2001.

[22] C. Jirauschek, G. Scarpa, P. Lugli, M. S. Vitiello, and G. Scamarcio, "Comparative analysis of resonant phonon THz quantum cascade lasers," *J. Appl. Phys.*, vol. 101, 086109 (3 pp.), 2007.

[23] O. Bonno, J. Thobel, and F. Dessenne, "Modeling of electron–electron scattering in Monte Carlo simulation of quantum cascade lasers," *J. Appl. Phys.*, vol. 97, 043702 (7 pp.), 2005.

[24] R. C. Iotti and F. Rossi, "Nature of charge transport in quantum-cascade lasers," *Phys. Rev. Lett.*, vol. 87, 146603 (4 pp.), 2001.

[25] H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, "Importance of electron-impurity scattering for electron transport in terahertz quantum-cascade lasers," *Appl. Phys. Lett.*, vol. 84, pp. 645–647, 2004.

[26] H. Callebaut and Q. Hu, "Importance of coherence for electron transport in terahertz quantum cascade lasers," *J. Appl. Phys.*, vol. 98, 104505 (11 pp.), 2005.

[27] C. Weber, A. Wacker, and A. Knorr, "Density-matrix theory of the optical dynamics and transport in quantum cascade structures: The role of coherence," *Phys. Rev. B*, vol. 79, 165322 (14 pp.), 2009.

[28] S. Kumar and Q. Hu, "Coherence of resonant-tunneling transport in terahertz quantum-cascade lasers," *Phys. Rev. B*, vol. 80, 245316 (14 pp.), 2009.

[29] E. Dupont, S. Fathololoumi, and H. C. Liu, "Simplified density-matrix model applied to three-well terahertz quantum cascade lasers," *Phys. Rev. B*, vol. 81, 205311 (18 pp.), May 2010.

[30] R. Terazzi and J. Faist, "A density matrix model of transport and radiation in quantum cascade lasers," *New J. Phys.*, vol. 12, 033045 (10 pp.), 2010.

[31] I. Waldmueller, W. W. Chow, E. W. Young, and M. C. Wanke, "Nonequilibrium many-body theory of intersubband lasers," *IEEE J. Quantum Electron.*, vol. 42, no. 3, pp. 292–301, Mar. 2006.

[32] A. Gordon and D. Majer, "Coherent transport in semiconductor heterostructures: A phenomenological approach," *Phys. Rev. B*, vol. 80, 195317 (6 pp.), 2009.

[33] A. Wacker, "Gain in quantum cascade lasers and superlattices: A quantum transport theory," *Phys. Rev. B*, vol. 66, 085326 (7 pp.), 2002.

[34] S.-C. Lee and A. Wacker, "Nonequilibrium Green's function theory for transport and gain properties of quantum cascade structures," *Phys. Rev. B*, vol. 66, 245314 (18 pp.), 2002.

[35] S.-C. Lee, F. Banit, M. Woerner, and A. Wacker, "Quantum mechanical wavepacket transport in quantum cascade laser structures," *Phys. Rev. B*, vol. 73, 245320 (6 pp.), 2006.

[36] T. Schmielau and M. Pereira, "Nonequilibrium many body theory for quantum transport in terahertz quantum cascade lasers," *Appl. Phys. Lett.*, vol. 95, 231111 (3 pp.), 2009.

[37] T. Kubis, C. Yeh, P. Vogl, A. Benz, G. Fasching, and C. Deutsch, "Theory of nonequilibrium quantum transport and energy dissipation in terahertz quantum cascade lasers," *Phys. Rev. B*, vol. 79, 195323 (10 pp.), 2009.

[38] G. Haldaś, A. Kolek, and I. Tralle, "Modeling of mid-infrared quantum cascade laser by means of nonequilibrium Green's functions," *IEEE J. Quantum Electron.*, vol. 47, no. 6, pp. 878–885, Jun. 2011.

[39] D. O. Winge, M. Lindskog, and A. Wacker, "Nonlinear response of quantum cascade structures," *Appl. Phys. Lett.*, vol. 101, 211113 (4 pp.), 2012.

[40] A. Wacker, "Coherence and spatial resolution of transport in quantum cascade lasers," *Phys. Stat. Sol. (c)*, vol. 5, pp. 215–220, 2008.

[41] H.-P. Breuer and F. Petruccione, *Open Quantum Systems*. Oxford, U.K.: Oxford Univ. Press, 2006.
[42] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*. New York, USA: Benjamin, 1962.
[43] L. V. Keldysh, "Diagram technique for nonequilibrium processes," *Sov. Phys. JETP*, vol. 20, p. 1018–1026, 1965 [Zh. Eksp. Theor. Fiz. **47**, 1515 (1964)].
[44] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors*. Berlin, Germany: Springer-Verlag, 1996.
[45] A. Kolek, G. Haldaś, and M. Bugajski, "Nonthermal carrier distributions in the subbands of 2-phonon resonance mid-infrared quantum cascade laser," *Appl. Phys. Lett.*, vol. 101, 061110 (4 pp.), 2012.
[46] T. Schmielau and M. F. Pereira, "Impact of momentum dependent matrix elements on scattering effects in quantum cascade lasers," *Phys. Stat. Sol. (b)*, vol. 246, pp. 329–331, 2009.
[47] T. Kubis and P. Vogl, "Assessment of approximations in nonequilibrium Green's function theory," *Phys. Rev. B*, vol. 83, 195304 (12 pp.), May 2011.
[48] C. Ndebeka-Bandou, F. Carosella, R. Ferreira, A. Wacker, and G. Bastard, "Relevance of intra- and inter-subband scattering on the absorption in heterostructures," *Appl. Phys. Lett.*, vol. 101, 191104 (4 pp.), 2012.
[49] V. Eyert, "A comparative study on methods for convergence acceleration of iterative vector sequences," *J. Comp. Phys.*, vol. 124, pp. 271–285, 1996.
[50] F. Klappenberger, K. F. Renk, P. Renk, B. Rieder, Y. I. Koshurinov, D. G. Pavelev, V. Ustinov, A. Zhukov, N. Maleev, and A. Vasilyev, "Semiconductor–superlattice frequency multiplier for generation of submillimeter waves," *Appl. Phys. Lett.*, vol. 84, pp. 3924–3926, 2004.
[51] L. Esaki and R. Tsu, "Superlattice and negative differential conductivity in semiconductors," *IBM J. Res. Develop.*, vol. 14, pp. 61–65, 1970.
[52] A. Wacker and A.-P. Jauho, "Quantum transport: The link between standard approaches in superlattices," *Phys. Rev. Lett.*, vol. 80, pp. 369–372, 1998.
[53] A. A. Ignatov and Y. A. Romanov, "Nonlinear electromagnetic properties of semiconductors with a superlattice," *Phys. Stat. Sol. (b)*, vol. 73, pp. 327–333, 1976.
[54] A. Wacker, "Semiconductor superlattices: A model system for nonlinear transport," *Phys. Rep.*, vol. 357, pp. 1–111, 2002.
[55] D. H. Dunlap and V. M. Kenkre, "Dynamic localization of a charged particle moving under the influence of an electric field," *Phys. Rev. B*, vol. 34, pp. 3625–3633, 1986.
[56] G. Scalari, M. I. Amanti, C. Walther, R. Terazzi, M. Beck, and J. Faist, "Broadband THz lasing from a photon-phonon quantum cascade structure," *Opt. Exp.*, vol. 18, pp. 8043–8052, 2010.
[57] M. Helm, W. Hilber, T. Fromherz, F. M. Peeters, K. Alavi, and R. N. Pathak, "Infrared absorption in superlattices: A probe of the miniband dispersion and the structure of the impurity band," *Phys. Rev. B*, vol. 48, pp. 1601–1606, 1993.
[58] A. Wacker, G. Bastard, F. Carosella, R. Ferreira, and E. Dupont, "Unraveling of free-carrier absorption for terahertz radiation in heterostructures," *Phys. Rev. B*, vol. 84, 205319 (6 pp.), 2011.
[59] D. Burghoff, T.-Y. Kao, D. Ban, A. W. M. Lee, Q. Hu, and J. Reno, "A terahertz pulse emitter monolithically integrated with a quantum cascade laser," *Appl. Phys. Lett.*, vol. 98, 061112 (3 pp.), 2011.
[60] L. Gendron, M. Carras, A. Huynh, V. Ortiz, C. Koeniguer, and V. Berger, "Quantum cascade photodetector," *Appl. Phys. Lett.*, vol. 85, 2824 (3 pp.), 2004.
[61] L. Gendron, C. Kroeniguer, V. Berger, and X. Marcadet, "High resistance narrow band quantum cascade photodetectors," *Appl. Phys. Lett.*, vol. 86, 121116 (3 pp.), 2005.
[62] T. Ando, "Self-consistent result for a GaAs/$Al_x Ga_{1-x}$As heterojunction. ii. low temperature mobility," *J. Phys. Soc. Jpn.*, vol. 51, pp. 3900–3907, 1982.
[63] R. Nelander and A. Wacker, "Temperature dependence and screening models in quantum cascade structures," *J. Appl. Phys.*, vol. 106, 063115 (6 pp.), 2009.
[64] D. C. Langreth, "Linear and nonlinear response theory with applications," in *Linear and Nonlinear Electron Transport in Solids*, J. T. Devreese and V. E. van Doren, Eds. New York, USA: Plenum, 1976.
[65] T. Kubis and P. Vogl, "How periodic are quantum cascade lasers?" *J. Phys., Conf. Ser.* vol. 193, 012063 (4 pp.), 2009.
[66] T. Brandes, "Truncation method for Green's functions in time-dependent fields," *Phys. Rev. B*, vol. 56, pp. 1213–1224, 1997.

Andreas Wacker was born in Neuwied, Germany, and studied physics at RWTH Aachen. He received the Ph.D. degree from Technical University of Berlin, Berlin, Germany, under the supervision of Prof. E. Schöll.

After stays in Lyngby, Denmark, Madrid, Spain, and Berlin, he is currently a Professor for theoretical physics at Lund University, Lund, Sweden. His research focuses on nonequilibrium quantum transport in nanosystems.

Martin Lindskog was born in 1988 in Lund, Sweden. He received the Master of Science degree in physics from Lund University, Lund, in 2012, where he is currently working toward the Ph.D.degree at the Division of Mathematical Physics.

His work mainly focuses on simulation of quantum transport in quantum cascade structures.

David O. Winge was born in Lidhult, Sweden, 1987. He received the M.Sc. degree from Lund University, Lund, Sweden, in 2012, where he is currently working toward the Ph.D. degree at the Division of Mathematical Physics.

His research is in the field of quantum transport, especially in quantum cascade structures.

Paper IV

APPLIED PHYSICS LETTERS **105**, 103106 (2014)

Comparative analysis of quantum cascade laser modeling based on density matrices and non-equilibrium Green's functions

M. Lindskog,[1,a)] J. M. Wolf,[2] V. Trinite,[3] V. Liverini,[2] J. Faist,[2] G. Maisons,[3] M. Carras,[3] R. Aidam,[4] R. Ostendorf,[4] and A. Wacker[1]
[1]*Mathematical Physics, Lund University, Box 118, 22100 Lund, Sweden*
[2]*ETH Institute for Quantum Electronics, ETH-Zürich, 8093 Zürich, Switzerland*
[3]*III-V Lab, 1 Avenue Augustin Fresnel, 91767 Palaiseau, France*
[4]*Fraunhofer-Institut für Angewandte Festkörperphysik, Tullastrasse 72, 79108 Freiburg, Germany*

(Received 18 July 2014; accepted 28 August 2014; published online 11 September 2014)

We study the operation of an 8.5 μm quantum cascade laser based on GaInAs/AlInAs lattice matched to InP using three different simulation models based on density matrix (DM) and non-equilibrium Green's function (NEGF) formulations. The latter advanced scheme serves as a validation for the simpler DM schemes and, at the same time, provides additional insight, such as the temperatures of the sub-band carrier distributions. We find that for the particular quantum cascade laser studied here, the behavior is well described by simple quantum mechanical estimates based on Fermi's golden rule. As a consequence, the DM model, which includes second order currents, agrees well with the NEGF results. Both these simulations are in accordance with previously reported data and a second regrown device. © *2014 AIP Publishing LLC*.
[http://dx.doi.org/10.1063/1.4895123]

Quantum cascade lasers[1] (QCLs) have become an important source of infra-red radiation for spectroscopy applications,[2] and different modeling techniques of varying level of detail have been used to simulate the performance of such structures.[3] The density matrix (DM) scheme,[4–6] where the transport is governed by scattering transitions and selected tunneling rates calculated from the off-diagonal elements of the density matrix, provides fast calculations in good agreement with experimental data.[7] This makes the DM model suitable for predicting the behavior of well-known QCL types, and allows for layer sequence optimization.[8] However, such models rely on a number of approximations, such as scattering mechanisms based on Fermi's golden rule and thermalized subbands, in order to form a consistent and efficient model.

In contrast to the DM scheme, the method of non-equilibrium Green's functions[9–12] (NEGF) takes into account the full coherences between the off diagonal elements of the density matrix, as well as scattering between states of different in-plane momentum k. Scattering effects are treated by matrices of self-energies,[9] which provides a self-consistent solution with the coherences mentioned above. Furthermore, the energetic widths of the states are fully included. However, the calculations are very resource-demanding, and thus put limitations on the structure optimization that is possible with such an approach. Instead, it is well suited to model well-defined problems in deep detail, since information about carrier, current, and state densities resolved in energy and space can be extracted from the Green's functions. The NEGF approach can thus be used to validate simpler models, as well as structures that have been optimized using these.

In this letter, we present simulations of a QCL[7] based on a GaInAs/AlInAs heterostructure lattice matched to InP under operation. We consider two implementations of the DM model, where the first (DM simp.) uses the first order current only, while the second one (DM 2nd), described in Refs. 6 and 13, includes the second order current following Ref. 14. These are directly compared with the NEGF model described in Ref. 15. The results are compared with experimental data from Ref. 7 and a regrowth reported here. Going beyond Ref. 16, where a similar comparison was done for another sample, we present simulations under lasing conditions as well.

The models used differ essentially with respect to the main concepts applied. For the DM models, the quantum electronic states are calculated for each period separately. The scattering transition rates between the states within the period are calculated by Fermi's golden rule. In addition, the tunneling rates between pairs of states of different periods are derived from a density matrix model. In contrast, the NEGF model uses a basis of Wannier functions, and treats all states on equal footing based on the microscopic Hamiltonian. The DM models have the areal electron densities of the specific levels and the coherences between them as principal variables. In order to evaluate the total transition rates, they assume a simple Fermi-Dirac distribution function with the same temperature for all sub-bands which is provided as an input. The full distribution with respect to k is resolved within the NEGF model, which is determined self-consistently, where the lattice temperature determines the occupation of the phonon modes. The only approximation is the use of momentum-independent effective scattering matrix elements, which highly simplifies the numerical scheme.[15] In the DM models, the optical transitions in the laser field are calculated via Fermi's golden rule where the energy-conserving delta-function is replaced by a Lorentzian. DM simp. uses an empirical broadening of 20 meV for the gain and the tunneling rates as an input to the program, whereas DM 2nd evaluates them with the Ando

[a)]Electronic mail: martin.lindskog@teorfys.lu.se

model.[13,17] In the NEGF model, gain is calculated from the dynamical conductance, and calculations in a strong ac field follow the procedure outlined in Ref. 18.

We consider scattering from interface roughness, longitudinal optical phonons, alloy disorder, acoustic phonons approximated by a single frequency (not used in DM 2nd), and ionized dopants (not used in DM simp.). Non-parabolicity in the band structure is treated via an effective two-band or three-band (for DM simp.) model.[19,20] All basic parameters, such as band gaps, effective masses and optical properties of the QCL materials, are the same in all models. For instance, all models use the same Gaussian interface roughness correlation length of 9 nm and height of 0.1 nm and a lattice temperature of 300 K. Additionally, in the DM models, the electron temperature was fixed to 430 K in contrast to the self-consistent calculation reported in Ref. 6. Inter-carrier Coulomb interaction is treated on a mean-field level.

We have simulated the structure described in Ref. 7 using the three models described above. The structure with the Wannier-Stark states is shown in Fig. 1, together with the carrier densities, which have a shift with respect to the Wannier-Stark states caused by impurity scattering.[21] These shifts mainly result from the real parts of the self-energies, which are of the order of 14 meV. They are similar for all states, so that they hardly affect the tunneling resonances.

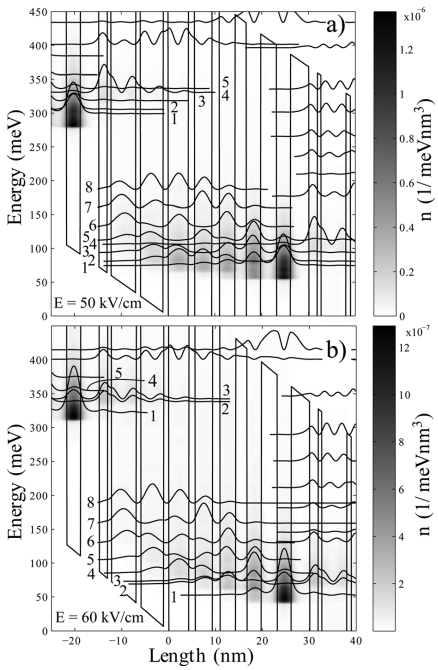


FIG. 1. Carrier densities and the square of the wave functions, calculated in the NEGF model, for different electric fields near threshold (a) and far above threshold (b). The numbers label the injection level (1), the upper laser levels (4 in (a), 2 and 3 in (b)), and the lower laser level (8).

Fig. 2 shows the current densities vs. applied electric field. The dashed lines show the experimental data for the original device and our regrowth, which agree until threshold (J_{th} = 1.5 kA/cm^2 and 2 kA/cm^2 of the original and regrown device, respectively). This reflects the reproducible growth quality as verified by X-ray measurements determining the actual period of 44.6 nm for the original and 44.7 nm for the regrown device (nominal 44.9 nm). In addition, the peak currents are comparable. The full lines give the simulation results of the different models without lasing. We see a good agreement between the NEGF model and the DM 2nd model for fields below ≈52 kV/cm, which both reproduce the experimental data below threshold. In contrast, as expected,[6] the DM simp. model provides a much larger current density, which shows the importance of including the second order current in the calculations.

The simulated gain spectra, taken in the limit of a vanishing lasing field, are shown in Fig. 3, near threshold (a) and far above threshold (b). The results of the DM 2nd and NEGF models agree near threshold while the DM simp. model shows a slightly larger gain with a blue-shift, coming from the approximation of constant effective mass in the gain calculation within this model. This trend is also seen in Fig. 3(c) for a wide range of electric fields, where the peak positions for both the NEGF and DM 2nd models agree qualitatively with the measurements. As the electric field increases, the gain evaluated by the two DM models increases strongly, while the NEGF model provides a significantly lower gain. This discrepancy most likely comes from the approximation of constant sub-band temperatures used in the DM models, while increasing electron temperatures provide detrimental occupation of higher levels at larger fields in the NEGF simulations. Another contributing factor is the restriction of the basis states to one period, which explains that the 1→8 side peak in Fig. 3(b) is not visible in the DM models. As can be seen in Fig. 1(a), the relevant upper laser level at threshold is level 4. At the higher field, however, the levels responsible for lasing are the two resonant levels

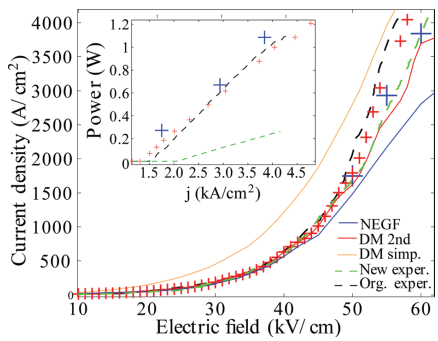


FIG. 2. Current-field characteristics for the different simulation models and experimental measurements. The field of the experimental data is obtained from the bias divided by the nominal length. Simulations under lasing conditions are given by large blue crosses for the NEGF and small red crosses for the DM 2nd model. The inset shows the measured and calculated output power as a function of the current density.

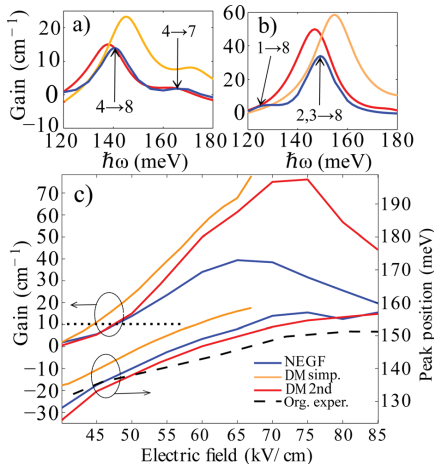


FIG. 3. Simulated spectral gain at electric fields of (a) 50 kV/cm and (b) 60 kV/cm. The gain is calculated in the off-state of the laser. In (a) and (b), the transitions of the respective gain peaks in the NEGF simulations are indicated. (c) Peak values and positions vs. electric field. The horizontal dotted line shows the computed threshold gain G_{th} for the original experiment. The experimental peak position (black dashes) is taken from the electroluminescence spectrum of Ref. 7.

2 and 3, whose coherence is only fully taken into account in the NEGF model.

The waveguide losses α_w for the device in Ref. 7 are reported to be 2.8 cm^{-1}. With a mirror loss of $\alpha_m = 3.3$ cm^{-1} and a mode confinement factor of $\Gamma_{\text{conf.}} = 0.63$, a gain $G_{\text{th}} \approx 10$ cm^{-1} is required in the QCL active region for achieving lasing. From Fig. 3(c), it is seen that in the DM 2nd and NEGF models, G_{th} is reached at $E_{\text{th}} = 47.3$ kV/cm and $E_{\text{th}} = 47.6$ kV/cm, respectively. This is in very good agreement with the experimental value of 48 kV/cm. For this field, the NEGF model provides a threshold current density of 1.2 kA/cm^2, and the DM 2nd model gives 1.3 kA/cm^2 (Exp. 1.5 kA/cm^2). Again, the DM simp. model differs, giving a threshold field of 44 kV/cm with the corresponding current density of 1.7 kA/cm. As the new device was processed using a double trench waveguide instead of the buried heterostructure technique used in the original device, the losses for this sample are higher. The observed threshold field of 52 kV/cm and threshold current of ≈2 kA/cm^2 reflect the trend for the gain simulations.

Under laser operation, an output power of $P = 0.5 - 1$ W was reported in Ref. 7. Using the relation for a traveling wave[18]

$$P = (F_{\text{ac}})^2 \frac{n_r c \varepsilon_0 A (1-R)}{2\Gamma_{\text{conf.}}}, \qquad (1)$$

where n_r is the refractive index of the gain medium, A is the facet area, and R is the reflectivity, we obtain an ac field inside the active region of strength $F_{\text{ac}}d \approx 100 - 130$ mV, where d is the period length. This is a significant amount compared with the dc field under operation $F_{\text{dc}}d \sim 250$ mV. This ac field is affecting the transport by increasing the current density above threshold, as expected in the case of photon-driven transport.[22] The simulated transport under an applied ac field is shown in Fig. 4, where an increasing $F_{\text{ac}}d$ results in an increased current density and a decreasing gain. For $F_{\text{dc}} = 50$ kV/cm, the NEGF results agree well with the DM 2nd model. Again, for $F_{\text{dc}} = 60$ kV/cm, the gain is higher for the DM model. The dotted line in Fig. 4 indicates G_{th} for the original device. Its intersection with the gain determines the ac field where gain and losses compensate. For these ac fields, Fig. 2 shows with crosses the NEGF and DM 2nd current densities under operation. This is in much better agreement with the experimental current density as compared to the simulations without a laser field. The corresponding power output (inset of Fig. 2) calculated from Eq. (1) for the NEGF model also agrees well with this experimental data as well as the DM 2nd model.

Now, we show that the gain in the NEGF model follows simple estimates, which demonstrates that the behavior can be understood in conventional terms. A simple calculation using Fermi's golden rule provides (see, e.g., Refs. 13 and 23)

$$G(\omega) = \frac{e^2 \Delta n_{fi} \Delta E_{fi}^2 z_{fi}^2}{2\hbar^2 \omega n_r c \epsilon_0 d} \frac{\Gamma_{\text{w}}}{\left(\Delta E_{fi} - \hbar\omega\right)^2 + \Gamma_{\text{w}}^2/4}, \qquad (2)$$

where Δn_{fi} is the inversion, $z_{fi} = 2.2$ the dipole matrix element, and ΔE_{fi} the energy difference for the final (f) and initial (i) states for the main gain transition. $\Gamma_{\text{w}} = 14$ meV is the FWHM of the gain spectrum from Fig. 3(a). Extracting the values Δn_{fi}, ΔE_{fi}, and z_{fi} from the NEGF simulations, we calculate for electric fields of 50 and 60 kV/cm a peak gain of 16 and 43 cm^{-1}, respectively, agreeing reasonably well with Figs. 3(a) and 3(b). Thus, the gain can be solely explained by the relationship (2), where Δn_{fi} accounts for the largest fraction of the variation in gain. This, together with the fact that the frequency of peak gain agrees with the energy difference between the peaks in the spectral function, shows that in this particular case, complex effects such as dispersive gain[24,25] and depolarization shifts[26,27] are not of

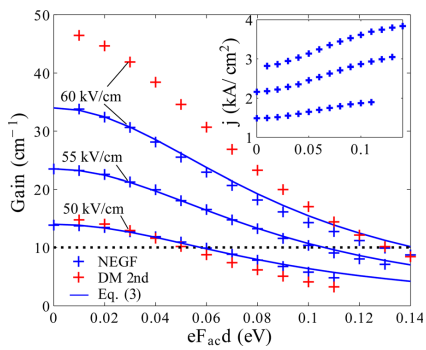


FIG. 4. Gain vs. applied ac field strength for different dc fields, simulated in the NEGF and DM 2nd models, compared to the relationship (3). The inset shows the current density in the NEGF model. The horizontal dotted line shows the threshold gain G_{th} of the original sample.

relevance. The differences in gain between the NEGF and the DM models for high dc fields can be attributed to the differences in Δn_{fi} addressed above.

Regarding the F_{ac} dependence in Fig. 4, it is reasonable to assume that the inversion drops proportionally to the product of gain times lasing intensity. Thus, we expect $G(F_{ac}) = G_0 - bG(F_{ac}) \times (F_{ac}d)^2$, where G_0 is the gain at zero field. The parameter b can be evaluated using Eq. (2) and standard kinetics, providing

$$G(F_{ac}) = \frac{G_0}{1 + b(F_{ac}d)^2} \quad \text{with} \quad b = \frac{\tau e^2 z_{fi}^2}{\hbar \Gamma_w d^2}. \tag{3}$$

We obtain $b = 120\,\mathrm{V}^{-2}$ for $z_{fi} = 2.2$ nm and $\tau = 0.47$ ps (approximately the lifetime of the upper laser state). Fig. 4 shows that Eq. (3) reproduces the full NEGF calculations.

From the NEGF modeling results, the carrier densities for each level α and in-plane momentum k can be extracted. Fitting to a Fermi distribution function, the subband temperatures T_α can be extracted. With $E = 55$ kV/cm without a lasing field, the upper laser state temperature $T_{ULS} = 398$ K is close to the electron temperature of 430 K used in the DM simulations, while the lower laser state temperature $T_{LLS} = 345$ K is significantly lower. When the laser field is turned on, with $F_{ac}d = 110$ mV, $T_{ULS} = 512$ K is now larger, while $T_{LLS} = 411$ K is close to the DM temperature instead.

In conclusion, we have presented simulations of a QCL both with and without a laser field, based on DM and NEGF models, and compared these results to experimental data from two separate growths. The transport and gain characteristics are well explained by simple relations, and thus, the DM 2nd model reproduces the experimental data as well as the NEGF model. However, the NEGF model predicts significantly lower gain and current for large dc fields close to the current density peak. We also find that the DM simp. model overestimates the current density in the devices, confirming the importance of taking the second order current into account. Finally, electron temperatures similar to those assumed in the DM models have been calculated using the NEGF model.

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 317884, the collaborative Integrated Project MIRIFISENS.

[1]J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, Science **264**, 553 (1994).
[2]R. F. Curl, F. Capasso, C. Gmachl, A. A. Kosterev, B. McManus, R. Lewicki, M. Pusharsky, G. Wysocki, and F. K. Tittel, Chem. Phys. Lett. **487**, 1 (2010).
[3]C. Jirauchek and T. Kubis, Appl. Phys. Rev. **1**, 011307 (2014).
[4]S. Kumar and Q. Hu, Phys. Rev. B **80**, 245316 (2009).
[5]E. Dupont, S. Fathololoumi, and H. C. Liu, Phys. Rev. B **81**, 205311 (2010).
[6]R. Terazzi and J. Faist, New J. Phys. **12**, 033045 (2010).
[7]A. Bismuto, R. Terazzi, M. Beck, and J. Faist, Appl. Phys. Lett. **96**, 141105 (2010).
[8]A. Bismuto, R. Terazzi, B. Hinkov, M. Beck, and J. Faist, Appl. Phys. Lett. **101**, 021103 (2012).
[9]S.-C. Lee, F. Banit, M. Woerner, and A. Wacker, Phys. Rev. B **73**, 245320 (2006).
[10]T. Schmielau and M. Pereira, Appl. Phys. Lett. **95**, 231111 (2009).
[11]T. Kubis, C. Yeh, P. Vogl, A. Benz, G. Fasching, and C. Deutsch, Phys. Rev. B **79**, 195323 (2009).
[12]G. Haldaś, A. Kolek, and I. Tralle, IEEE J. Quantum Electron. **47**, 878 (2011).
[13]R. Terazzi, *Transport in Quantum Cascade Lasers*, Ph.D. thesis, ETH Zürich, 2011.
[14]R. Terazzi, T. Gresch, A. Wittmann, and J. Faist, Phys. Rev. B **78**, 155328 (2008).
[15]A. Wacker, M. Lindskog, and D. Winge, IEEE J. Sel. Top. Quantum Electron. **19**, 1200611 (2013).
[16]P. Friedli, H. Sigg, A. Wittmann, R. Terazzi, M. Beck, A. Kolek, and J. Faist, Appl. Phys. Lett. **102**, 012112 (2013).
[17]T. Ando, J. Phys. Soc. Jpn. **54**, 2671 (1985).
[18]D. O. Winge, M. Lindskog, and A. Wacker, Opt. Express **22**, 18389 (2014).
[19]C. Sirtori, F. Capasso, J. Faist, and S. Scandolo, Phys. Rev. B **50**, 8663 (1994).
[20]M. Lindskog, D. O. Winge, and A. Wacker, Proc. SPIE **8846**, 884603 (2013).
[21]C. Ndebeka-Bandou, A. Wacker, F. Carosella, R. Ferreira, and G. Bastard, Appl. Phys. Express **6**, 094101 (2013).
[22]H. Choi, L. Diehl, Z.-K. Wu, M. Giovannini, J. Faist, F. Capasso, and T. B. Norris, Phys. Rev. Lett. **100**, 167401 (2008).
[23]J. Faist, *Quantum Cascade Lasers* (Oxford University Press, Oxford, 2013).
[24]R. Terazzi, T. Gresch, M. Giovannini, N. Hoyler, N. Sekine, and J. Faist, Nat. Phys. **3**, 329 (2007).
[25]D. G. Revin, M. R. Soulby, J. W. Cockburn, Q. Yang, C. Manz, and J. Wagner, Appl. Phys. Lett. **92**, 081110 (2008).
[26]P. Hyldgaard and J. W. Wilkins, Phys. Rev. B **53**, 6889 (1996).
[27]M. F. Pereira, Jr., S.-C. Lee, and A. Wacker, Phys. Rev. B **69**, 205310 (2004).

Paper V

Impact of interface roughness distributions on the operation of quantum cascade lasers

Martin Franckié,[1,*] David O. Winge,[1] Johanna Wolf,[2] Valeria Liverini,[2] Emmanuel Dupont,[3] Virginie Trinité,[4] Jérôme Faist,[2] and Andreas Wacker[1]

[1]*Mathematical Physics, Lund University, Box 118, 22100 Lund, Sweden*
[2]*ETH Institute for Quantum Electronics, ETH-Zürich, 8093 Zürich, Switzerland*
[3]*National Research Council, 1200 Montreal Road, Ottawa, Ontario K1A 0R6, Canada*
[4]*III-V Lab, 1 Avenue Augustin Fresnel, 91767 Palaiseau, France*

[*]*martin.lindskog@teorfys.lu.se*

Abstract: We study the impact of interface roughness on the operation of mid-IR and THz quantum cascade lasers. Particular emphasis is given towards the differences between the Gaussian and exponential roughness distribution functions, for which we present results from simulation packages based on nonequilibrium Green's functions and density matrices. The Gaussian distribution suppresses scattering at high momentum transfer which enhances the lifetime of the upper laser level in mid-IR lasers. For THz lasers, a broader range of scattering transitions is of relevance, which is sensitive to the entire profile of the interface fluctuations. Furthermore we discuss the implementation of interface roughness within a two band model.

OCIS codes: (140.5965) Semiconductor lasers, quantum cascade; (000.6800) Theoretical physics.

References and links

1. J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser," Science **264**, 553–556 (1994).
2. R. F. Curl, F. Capasso, C. Gmachl, A. A. Kosterev, B. McManus, R. Lewicki, M. Pusharsky, G. Wysocki, and F. K. Tittel, "Quantum cascade lasers in chemical physics," Chem. Phys. Lett. **487**, 1–18 (2010).
3. T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akiyama, "Intersubband absorption linewidth in GaAs quantum wells due to scattering by interface roughness, phonons, alloy disorder, and impurities," J. Appl. Phys. **93**, 1586–1597 (2003).
4. S. Tsujino, A. Borak, E. Muller, M. Scheinert, C. V. Falub, H. Sigg, D. Grutzmacher, M. Giovannini, and J. Faist, "Interface-roughness-induced broadening of intersubband electroluminescence in p-SiGe and n-GaInAsAlInAs quantum-cascade structures," Appl. Phys. Lett. **86**, 062113 (2005).
5. A. Leuliet, A. Vasanelli, A. Wade, G. Fedorov, D. Smirnov, G. Bastard, and C. Sirtori, "Electron scattering spectroscopy by a high magnetic field in quantum cascade lasers," Phys. Rev. B **73**, 085311 (2006).
6. A. Vasanelli, A. Leuliet, C. Sirtori, A. Wade, G. Fedorov, D. Smirnov, G. Bastard, B. Vinter, M. Giovannini, and J. Faist, "Role of elastic scattering mechanisms in GaInAs/AlInAs quantum cascade lasers," Appl. Phys. Lett. **89**, 172120 (2006).
7. Y. Chiu, Y. Dikmelik, P. Q. Liu, N. L. Aung, J. B. Khurgin, and C. F. Gmachl, "Importance of interface roughness induced intersubband scattering in mid-infrared quantum cascade lasers," Appl. Phys. Lett. **101**, 171117 (2012).
8. J. B. Khurgin, "Inhomogeneous origin of the interface roughness broadening of intersubband transitions," Appl. Phys. Lett. **93**, 091104 (2008).
9. J. B. Khurgin, Y. Dikmelik, P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, and C. F. Gmachl, "Role of interface roughness in the transport and lasing characteristics of quantum-cascade lasers," Appl. Phys. Lett. **94**, 091101 (2009).

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

10. T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," Rev. Mod. Phys. **54**, 437–672 (1982).
11. S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, "Surface roughness at the Si(100)-SiO_2 interface," Phys. Rev. B **32**, 8171–8186 (1985).
12. R. M. Feenstra, D. A. Collins, D. Z. Y. Ting, M. W. Wang, and T. C. McGill, "Interface roughness and asymmetry in inas/gasb superlattices studied by scanning tunneling microscopy," Phys. Rev. Lett. **72**, 2749–2752 (1994).
13. A. Y. Lew, S. L. Zuo, E. T. Yu, and R. H. Miles, "Correlation between atomic-scale structure and mobility anisotropy in $InAsGa_{1-x}In_xSb$ superlattices," Phys. Rev. B **57**, 6534–6539 (1998).
14. K.-J. Chao, N. Liu, C.-K. Shih, D. W. Gotthold, and B. G. Streetman, "Factors influencing the interfacial roughness of InGaAs/GaAs heterostructures: A scanning tunneling microscopy study," Appl. Phys. Lett. **75**, 1703–1705 (1999).
15. P. Offermans, P. M. Koenraad, J. H. Wolter, M. Beck, T. Aellen, and J. Faist, "Digital alloy interface grading of an InAlAs/InGaAs quantum cascade laser structure studied by cross-sectional scanning tunneling microscopy," Appl. Phys. Lett. **83**, 4131–4133 (2003).
16. T. Saku, Y. Horikoshi, and Y. Tokura, "Limit of electron mobility in AlGaAs/GaAs modulation-doped heterostructures," Jpn. J. Appl. Phys. **35**, 34–38 (1996).
17. F. Lopez, M. R. Wood, M. Weimer, C. F. Gmachl, and C. G. Caneau (2013). Abstract presented at the ITQW, Bolton Landing.
18. Y. Chiu, Y. Dikmelik, Q. Zhang, J. Khurgin, and C. Gmachl, "Engineering the intersubband lifetime with interface roughness in quantum cascade lasers," in "Lasers and Electro-Optics (CLEO), 2012 Conference on," (2012), pp. 1–2.
19. S. A. Cohen, "The Fourier transform asymptotic behavior theorem," IEEE Transactions on Education **12**, 56–57 (1969).
20. C. Sirtori, F. Capasso, J. Faist, and S. Scandolo, "Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells," Phys. Rev. B **50**, 8663–8674 (1994).
21. A. Wacker, M. Lindskog, and D. Winge, "Nonequilibrium Green's function model for simulation of quantum cascade laser devices under operating conditions," IEEE J. Sel. Topics Quantum Electron. **19**, 1200611 (2013).
22. M. Lindskog, D. O. Winge, and A. Wacker, "Injection schemes in THz quantum cascade lasers under operation," Proc. SPIE **8846**, 884603 (2013).
23. E. Dupont, S. Fathololoumi, Z. R. Wasilewski, G. Aers, S. R. Laframboise, M. Lindskog, S. G. Razavipour, A. Wacker, D. Ban, and H. C. Liu, "A phonon scattering assisted injection and extraction based terahertz quantum cascade laser," J. Appl. Phys. **111**, 073111 (2012).
24. A. Bismuto, R. Terazzi, M. Beck, and J. Faist, "Electrically tunable, high performance quantum cascade laser," Appl. Phys. Lett. **96**, 141105 (2010).
25. R. Terazzi and J. Faist, "A density matrix model of transport and radiation in quantum cascade lasers," New Journal of Physics **12**, 033045 (2010).
26. M. Lindskog, J. M. Wolf, V. Trinite, V. Liverini, J. Faist, G. Maisons, M. Carras, R. Aidam, R. Ostendorf, and A. Wacker, "Comparative analysis of quantum cascade laser modeling based on density matrices and nonequilibrium Green's functions," Appl. Phys. Lett. **105**, 103106 (2014).
27. S. Kumar, Q. Hu, and J. L. Reno, "186 K operation of terahertz quantum-cascade lasers based on a diagonal design," Appl. Phys. Lett. **94**, 131105 (2009).
28. S. Fathololoumi, E. Dupont, C. W. I. Chan, Z. R. Wasilewski, S. R. Laframboise, D. Ban, A. Mtys, C. Jirauschek, Q. Hu, and H. C. Liu, "Terahertz quantum cascade lasers operating up to ∼200 K with optimized oscillator strength and improved injection tunneling," Opt. Express **20**, 3866–3876 (2012).
29. D. O. Winge, M. Lindskog, and A. Wacker, "Nonlinear response of quantum cascade structures," Appl. Phys. Lett. **101**, 211113 (2012).
30. C. Jirauschek and P. Lugli, "Monte-Carlo-based spectral gain analysis for terahertz quantum cascade lasers," J. Appl. Phys. **105**, 123102 (2009).
31. S. Fathololoumi, E. Dupont, Z. R. Wasilewski, C. W. I. Chan, S. G. Razavipour, S. R. Laframboise, S. Huang, Q. Hu, D. Ban, and H. C. Liu, "Effect of oscillator strength and intermediate resonance on the performance of resonant phonon-based terahertz quantum cascade lasers," J. Appl. Phys. **113**, 113109 (2013).

1. Introduction

Quantum cascade lasers (QCLs) [1] have become an important source of infrared radiation for spectroscopy applications [2]. They consist of a vast number of specifically designed semiconductor layers. As the interfaces of these layers are never entirely perfect, the lateral translational invariance is broken and interface roughness scattering becomes inevitable. Several studies have focused on its relevance for the lifetime of the upper laser level [3–7], which is a key element for the lasing performance. Furthermore, interface roughness is relevant for the broadening of

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

tunneling transitions [8, 9] in QCLs. The deviations from an ideal interface are treated statistically, where the spatial correlation function of the height fluctuations contains the relevant information to evaluate the scattering matrix elements. (Correlations between different interfaces are of minor importance as they are washed out under typical growth conditions unless barriers are very thin [4].) It is common to model this correlation function by a Gaussian with two fit parameters "although there is no definite physical ground" [10]. In this paper such a Gaussian correlation function is compared to an exponential fit and calculations for different QCLs are presented under both assumptions. While there are some specific differences, the calculated current-voltage characteristics and gain spectra are comparable if the respective fit parameters are correctly transformed.

2. Interface roughness models

The central assumption for modeling interface roughness is that the position of the interface between two materials is fluctuating by $\eta(\mathbf{r})$, where $\mathbf{r}$ is a two-dimensional vector in the $x-y$ plane of the heterostructure layers. Averaging over a large area, the statistical properties of $\eta(\mathbf{r})$ become important. In order to quantify roughness scattering, the square of the matrix element for momentum transfer $\mathbf{q}$ between different subband states is required. Next to prefactors, it contains the integral

$$\begin{aligned}\frac{1}{A}\int \mathrm{d}^2 r \int \mathrm{d}^2 r' e^{\mathrm{i}\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\eta(\mathbf{r})\eta(\mathbf{r}') &= \int \mathrm{d}^2 r\, e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \int \mathrm{d}^2 r_0 \eta(\mathbf{r_0}+\mathbf{r})\eta(\mathbf{r_0}) \\ &= \int \mathrm{d}^2 r e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \langle \eta(\mathbf{r})\eta(0)\rangle \equiv f(\mathbf{q}), \end{aligned} \tag{1}$$

where A is the sample area. This is the Fourier transformation of the spatial correlation function $\langle \eta(\mathbf{r})\eta(0)\rangle$ for the fluctuations, which we denote by $f(\mathbf{q})$. It is common to assume a Gaussian distribution [10] with

$$\begin{aligned}\langle \eta(\mathbf{r})\eta(0)\rangle &= \Delta^2 \exp\left(-\frac{|\mathbf{r}|^2}{\Lambda^2}\right) \\ \rightarrow \quad f(\mathbf{q}) &= \pi\Delta^2\Lambda^2 \exp\left(-\frac{\Lambda^2|\mathbf{q}|^2}{4}\right). \end{aligned} \tag{2}$$

Alternatively, the idea, that there is a constant likelihood to be at the rim of a roughness plateau, suggests an exponential distribution:

$$\begin{aligned}\langle \eta(\mathbf{r})\eta(0)\rangle &= \tilde{\Delta}^2 \exp\left(-\frac{|\mathbf{r}|}{\tilde{\Lambda}}\right) \\ \rightarrow \quad f(\mathbf{q}) &= \frac{2\pi\tilde{\Delta}^2\tilde{\Lambda}^2}{\left(1+\tilde{\Lambda}^2|\mathbf{q}|^2\right)^{3/2}}. \end{aligned} \tag{3}$$

The average fluctuation height Δ and the spatial correlation length Λ are not directly measurable, but should be seen as fit parameters. Thus it is meaningless to compare the Gaussian and an exponential distribution with the same set of parameters, i.e. setting $\tilde{\Delta}=\Delta$ and $\tilde{\Lambda}=\Lambda$. In this case the exponential distribution would result in twice the scattering of the Gaussian one at $\mathbf{q}\approx\mathbf{0}$. Rather, Λ and Δ should be changed as to achieve similar scattering rates in a wide range of $\mathbf{q}$, in order to quantify the difference using either distribution. Provided the dominating scattering events have small $\mathbf{q}$, both distributions are expected to provide similar results with the

$$\text{Translation 1} \quad \tilde{\Lambda}=\Lambda/\sqrt{6} \quad \text{and} \quad \tilde{\Delta}=\sqrt{3}\Delta \tag{4}$$

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

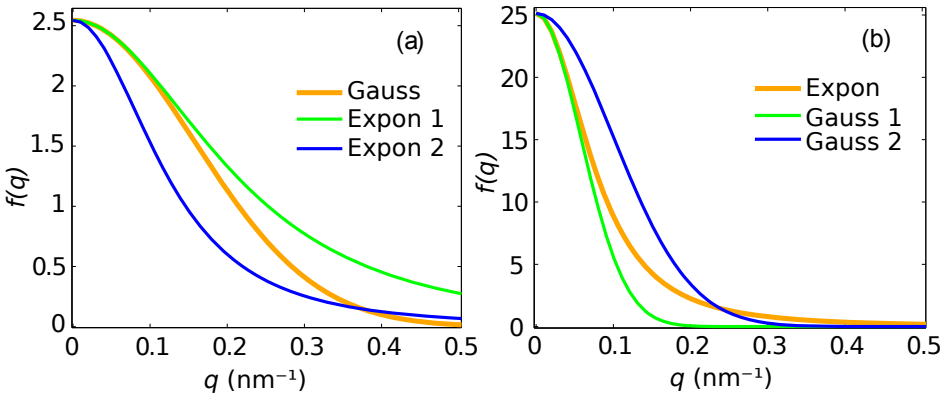


Fig. 1. Fourier transforms $f(q)$ of the correlation functions for different roughness distribution functions. In panel (a), the Gaussian distribution, Eq. (2), has the parameters $\Lambda = 9$ nm and $\Delta = 0.1$ nm. The parameters for the exponential distributions, Eq. (3), Expon 1 ($\tilde{\Lambda} = 3.6$ nm, $\tilde{\Delta} = 0.17$ nm) and Expon 2 ($\tilde{\Lambda} = 6.3$ nm, $\tilde{\Delta} = 0.1$ nm), are transformed via Eq. (4) and (5), respectively. These distributions are used for the InGaAs/InAlGaAs based IR QCL. Panel (b) shows the exponential distribution with $\tilde{\Lambda} = 10$ and $\tilde{\Delta} = 0.2$ nm, and the parameters for the two Gaussian distributions, Gauss 1 ($\tilde{\Lambda} = 24.5$ and $\tilde{\Delta} = 0.115$ nm) and Gauss 2 ($\tilde{\Lambda} = 14.1$ and $\tilde{\Delta} = 0.2$ nm), which are transformed via Eqs. (4) and (5), respectively. These distributions are used for the GaAs/AlGaAs based THz-QCLs.

so that the functions $f(\mathbf{q})$ as well as their second derivatives coincide at the maximum $|\mathbf{q}| = q = 0$. The corresponding functions $f(\mathbf{q})$ are displayed in Fig. 1 for different roughness parameters. We find, that the exponential distribution provides much stronger scattering for large $\mathbf{q}$ with this translation. A second natural translation is given by requiring an identical average fluctuation height $\langle \eta(\mathbf{r} = 0)\eta(0)\rangle$ and identical $f(\mathbf{q} = 0)$ for the Gaussian and exponential distribution. This provides

$$\text{Translation 2} \quad \tilde{\Lambda} = \Lambda/\sqrt{2} \quad \text{and} \quad \tilde{\Delta} = \Delta \tag{5}$$

which better reproduces the fall-off at larger q-values but agrees less well for small q as can be seen in Fig. 1.

There are actually experimental indications for an exponential distribution for several material systems such as Si/SiO_2 [11]; InAs/GaSb [12], InAs/GaInSb [13]; GaAs/InGaAs [14]. For GaAs/AlGaAs or InGaAs/InAlAs interfaces, relevant for QCL structures, less information is available. In [15], Offermans *et al.* report lateral fluctuations of 10 nm for an InGaAs/InAlAs QCL but do not provide a distribution function, while [16] reported values of 10-20 nm for a GaAs/AlGaAs superlattice. In [5], Leuliet *et al.* used a Gaussian distribution with $\Lambda = 6$ nm and $\Delta = 0.15$ nm, to fit their data for a $GaAs/Al_{0.33}Ga_{0.67}As$ QCL. Recently, $\langle \eta(\mathbf{r})\eta(0)\rangle$ has been measured for a GaAs/InGaAs QCL, where three different Gaussians were required to fit the data [17].

Considering the case of an IR QCL, the out-scattering from the upper laser level requires a particular large momentum transfer q, as the large energy mismatch between initial and final states (typically the lasing energy) must be transferred into in-plane kinetic energy. For typical infrared QCLs with a lasing energy of $\hbar\omega = 150$ meV, this implies $q \approx 0.4\ \text{nm}^{-1}$ (using the in-plane effective mass $m_e^* = 0.043$ for InGaAs lattice matched to InP). For a typical value of $\Lambda \sim 10$ nm, the Gaussian thus provides a strong suppression of the scattering matrix element. In contrast, the exponential distribution of Eq. (3) shows a power law suppression which allows

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

for some out-scattering, in particular with the Translation 1, see Fig. 1(a). For THz lasers, on the other hand, $\hbar\omega \approx 10$ meV corresponds to $q \approx 0.1$ nm^{-1} (using the in-plane effective mass $m_e^* = 0.067$ for GaAs) and scattering at intermediate q values decreases the lifetime. In this case the level spacings for the extraction process (as well as the injection for a scattering assisted design) match the optical phonon energy (~ 36 meV) with $q \approx 0.25$ nm^{-1}. Thus, large q scattering might even improve performance in these structures. Similar interface engineering has been proposed in [18], where barriers are inserted to decrease the lifetime of the lower laser state.

The power law for large q can actually be related to the behavior of $g(\mathbf{r}) = \langle \eta(\mathbf{r})\eta(0)\rangle$ for $r \to 0$. Standard rules of Fourier transformation provide the second derivative

$$\frac{\partial}{\partial \mathbf{r}} \cdot \frac{\partial}{\partial \mathbf{r}} g(\mathbf{r}) = -\frac{1}{4\pi^2} \int \mathrm{d}^2 q\, q^2 f(\mathbf{q}) \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}.$$

Assuming, that $g(\mathbf{r}) = g(r)$ and $f(\mathbf{q}) = f(q)$ are rotational invariant, the asymptotic behavior $f(q) \sim 1/q^3$ is thus related to $\partial^2 g(\mathbf{r})/\partial \mathbf{r}^2 \to \infty$ for $r \to 0$, which means, that the gradient of $g(\mathbf{r})$ is discontinuous at the origin (this corresponds to the Fourier transform asymptotic behavior theorem in the one dimensional case discussed in [19]). Such a discontinuity naturally occurs, if $g(r)$ has a finite slope at $r = 0$ (which is the case for the exponential distribution in contrast to the Gaussian).

A (negative) slope of $g(r)$ at $r = 0$ can be motivated by the following argument: We consider the product $\eta(\mathbf{r})\eta(\mathbf{r}_0)$ for a fixed reference point $\mathbf{r}_0$. This product is positive for $\mathbf{r} = \mathbf{r}_0$ and maintains its value as long as $\mathbf{r}$ is on the same plateau as $\mathbf{r}_0$. Crossing the rim of the plateau, $\eta(\mathbf{r})\eta(\mathbf{r}_0)$ changes, more likely to a negative value, as the average elongation $\langle \eta(\mathbf{r})\rangle = 0$. As one averages over all reference points $\mathbf{r}_0$, there are some points, which are precisely on the rim of a plateau and thus $\langle \eta(\mathbf{r})\eta(0)\rangle$ is expected to have a finite negative slope in the direction of $\mathbf{r}$ for small r. Note, that this argument requires a sharp drop of the scattering potential for the conduction band electrons at the rims between the plateaus. On the other hand, if the rims of the plateaus result in a smooth change of the potential landscape, there would be no such negative slope for small r, and the Gaussian distribution would be a viable choice. To determine which behavior is the most accurate in a real situation would therefore require the precise measurement and interpretation of the actual potential landscape.

In order to demonstrate the relevance of the roughness distribution, we provide simulation results for a mid-IR QCL, a THz QCL with scattering injection, and a THz QCL with tunneling injection. We apply the interface roughness parametrizations shown in Figs. 1(a) and (b) for the mid-IR QCL and THz QCLs, respectively. In all cases we apply identical distributions for all interfaces.

3. Two-band model: including the valence band offset

For IR QCLs non-parabolicity in the conduction band is relevant as the electronic states cover a large range of energies. This can be effectively implemented by mixing the conduction band wave-function with at least one component from the valence band [20]. As the valence band offset (VBO) differs from the conduction band offset, roughness may act differently, whether one restricts to the conduction band or implements a two-band model. In a phenomenological approach, where little is known about the actual roughness distribution, appropriate choices of the roughness parametrization can to a large extent compensate for the difference between these concepts. Nevertheless the magnitude of the difference is of interest for *a priori* calculations.

The Hamiltonian for interface roughness scattering is written in second quantization as

$$\hat{H} = \sum_{\alpha\beta}\sum_{\mathbf{k},\mathbf{p}} U_{\alpha\beta}(\mathbf{p}) a^\dagger_{\alpha\mathbf{k}+\mathbf{p}} a_{\beta\mathbf{k}} \quad \text{with} \quad U_{\alpha\beta}(\mathbf{p}) = \sum_j \int \mathrm{d}^2 r \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{A} \eta_j(\mathbf{r}) \Psi^*_\alpha(z_j) \Delta E \Psi_\beta(z_j) \quad (6)$$

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

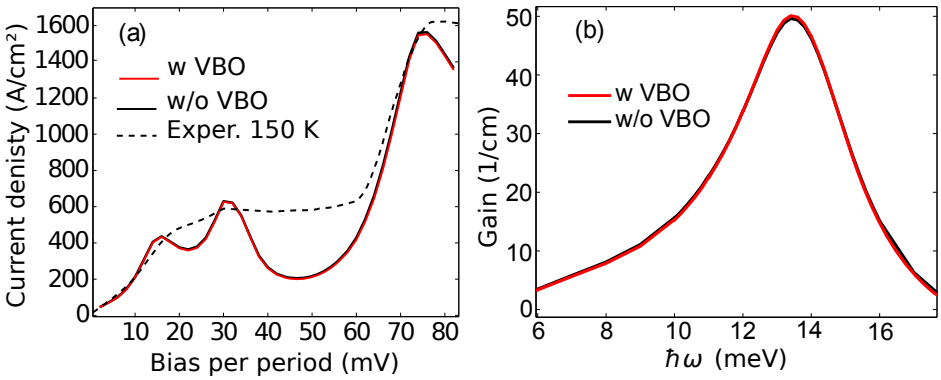


Fig. 2. (a) Current-field simulations of the THz QCL [23] using the NEGF model with and without inclusion of the valence band offset (VBO) in the roughness scattering. The simulation temperature is 140 K and the dashed line shows the experimental data at 150 K. Here, the exponential distribution function from Fig. 1(b) is used. (b) Gain simulations for the same sample at a bias of 74 mV/period.

where z_j denote the interface positions. In the two-band model,

$$\Psi_\beta(z) \to \begin{pmatrix} \psi^c_\beta(z) \\ \psi^v_\beta(z) \end{pmatrix} \quad \text{and} \quad \Delta E \to \begin{pmatrix} \Delta E_c & 0 \\ 0 & \Delta E_v \end{pmatrix}, \tag{7}$$

where c and v denote the conduction and valence band components, respectively, and $\Delta E_{c/v}$ are the respective band offsets. This gives terms in the matrix elements squared as

$$\langle U_{\alpha,\beta}(-\mathbf{p}) U_{\alpha',\beta'}(\mathbf{p})\rangle = \sum_j \frac{f_j(\mathbf{p})}{A} (\Delta E_c \psi^{c*}_\alpha \psi^c_\beta + \Delta E_v \psi^{v*}_\alpha \psi^v_\beta)(\Delta E_c \psi^{c*}_{\alpha'} \psi^c_{\beta'} + \Delta E_v \psi^{v*}_{\alpha'} \psi^v_{\beta'}), \tag{8}$$

where the wave functions are taken at the respective z_j. If only the conduction band offset is taken into account, then the resulting expression is instead

$$\langle U_{\alpha,\beta}(-\mathbf{p}) U_{\alpha',\beta'}(\mathbf{p})\rangle = \sum_j \frac{f_j(\mathbf{p})}{A} \Delta E_c^2 (\psi^{c*}_\alpha \psi^c_\beta + \psi^{v*}_\alpha \psi^v_\beta)(\psi^{c*}_{\alpha'} \psi^c_{\beta'} + \psi^{v*}_{\alpha'} \psi^v_{\beta'}). \tag{9}$$

Since for type I heterostructures ΔE_c and ΔE_v have opposite signs, the former case will exhibit lower interface roughness scattering.

Using our Nonequilibrium Green's function (NEGF) simulation scheme [21, 22] we performed simulations for different QCLs in order to compare the expressions (8) and (9). For the THz QCL of [23], Fig. 2 shows that the neglect of the valence band offset in the roughness scattering term neither changes the current density nor the gain. In contrast, a slight increase in the current density at the peak and a slight reduction of the gain is visible for the IR QCL of [24] in Fig. 3. This follows the expected trend, that nonparabolicity becomes more relevant with reduced band gap and increased state energies. Consequently, the difference between both approaches may become more prominent for QCLs at shorter wavelengths.

In both cases the simulations agree quantitatively with experimental data. As the simulations shown do not include the lasing field they cannot reproduce the experimental currents above the threshold current density of 1.5 kA/cm^2 in Fig. 3.

In the rest of this work, all simulations of the mid-IR QCL are performed using Eq. (9), whereas all simulations for the THz QCLs employ Eq. (8). In the calculations of the energy

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

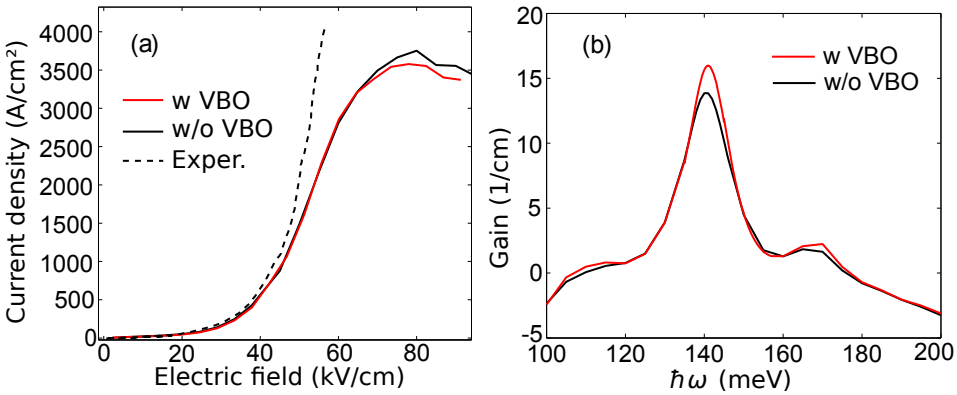


Fig. 3. (a) Current-field simulations of the mid-IR QCL [24] using the NEGF model with and without inclusion of the valence band offset (VBO) in the roughness scattering. The Gaussian interface roughness from Fig. 1(a) is applied. (b) Results for the gain at an electric field of 50 kV/cm and a lattice temperature of 300 K.

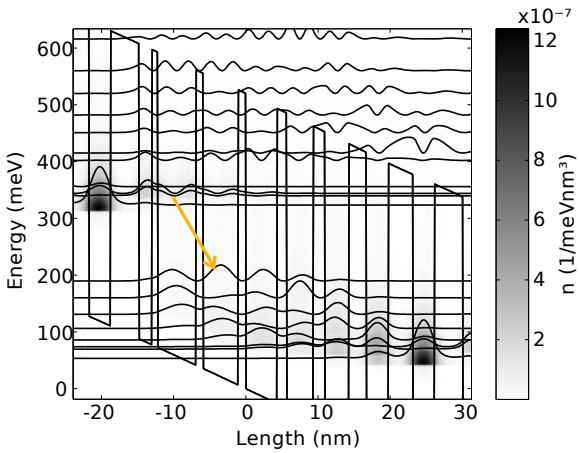


Fig. 4. Band structure with the square moduli of the wavefunctions, together with the carrier density from the NEGF simulation, for the mid-IR structure of [24]. The arrow indicates the main laser transition.

levels as well as the matrix elements of all other scattering mechanisms, i.e. acoustical and optical phonons, alloy and impurity scattering which do not depend explicitly on the band offsets, we use the two-component wavefunctions within the effective two-band model [20].

4. Results

4.1. Mid-IR laser

In order to quantify the impact of different interface roughness distributions on the QCL performance, we simulate the IR device [24] shown in Fig. 4 with the DM model from [25] and the NEGF model. A detailed discussion of the model differences is given in [26], where we applied the Gaussian roughness model with $\Lambda = 9$ nm and $\Delta = 0.1$ nm. Here we compare these results with the two exponential models applying different translations of the parameters as shown in

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

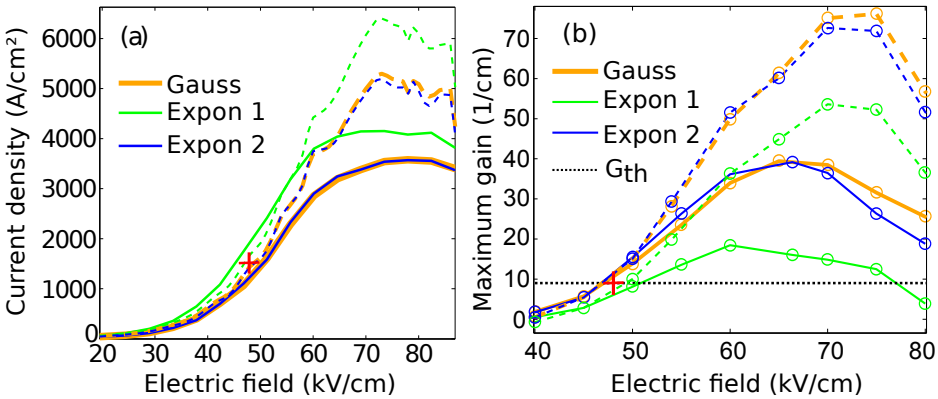


Fig. 5. (a) Current-field characteristics of the QCL in [24] for the DM (dashed lines) and NEGF (full lines) simulation schemes and different roughness distributions given in Fig. 1(a). (b) Peak gain vs. electric field. The dotted line denotes the gain required to compensate the losses. The red crosses show the experimental threshold data in both panels.

Fig. 1(a).

A comparison of the current-field characteristics is shown in Fig. 5(a). We find that the Expon 1 roughness distribution provides generally higher currents than Expon 2 and Gauss. This can be attributed to the shorter lifetime of the upper laser state, as shown in Table 1, due to enhanced roughness scattering with large momentum transfer, which facilitates the transfer of carriers through the device. The trend is the same in both simulation schemes, albeit the currents from the NEGF model are generally smaller than the DM results. Note that the field for the experimental data does not take into account any possible bias drop in contacting regions, which would reduce the field slightly. Thus, comparison with experiment cannot clearly support a certain model.

Now we consider the simulated peak value of the gain as a function of applied electric field, displayed in Fig. 5(b). Here we find the highest gain for the Gaussian roughness, somewhat lower values for Expon 2 and relatively low gain for Expon 1. Again this can be directly attributed to the lifetime of the upper laser state, which is central for the inversion. Quantitatively, the NEGF model provides an inversion Δn of 2.54, 1.14, and 2.33 $\times 10^9 \mathrm{cm}^{-2}$ and a linewidth Γ of 13.8, 13.1, and 11.6 meV for the Gauss, Expon 1 and Expon 2 distribution, respectively, at an electric field of 50 kV/cm. Thus the key contribution $\Delta n/\Gamma$ suggests a reduction in gain by 53% (Expon 1) and an increase by 9% (Expon 2) of the peak gain compared to the Gaussian distribution. This reflects the trend in the full NEGF calculations at 50 kV/cm, where the cor-

Table 1. Scattering times in the DM model for the upper and lower laser state at a bias of 50 kV/cm. The NEGF simulations show the same trend but due to the intricate treatment of coherences, it is less straightforward to extract a single time.

		Upper laser state			Lower laser state		
		Gauss 1	Expon 1	Expon 2	Gauss 1	Expon 1	Expon 2
$\tau_{\text{LO phonon}}$	(ps)	0.549	0.549	0.549	0.158	0.158	0.158
$\tau_{\text{interface roughness}}$	(ps)	2.611	0.739	2.439	0.777	0.458	1.123
$\tau_{\text{alloy disorder}}$	(ps)	2.283	2.283	2.283	2.270	2.270	2.270
τ_{sum}	(ps)	0.378	0.277	0.375	0.124	0.112	0.130

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

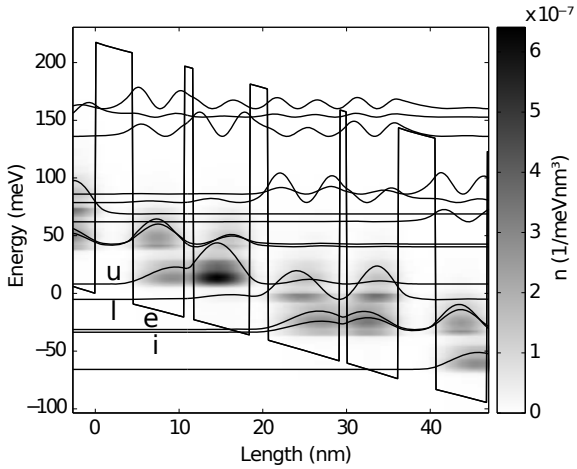


Fig. 6. Electron density and square of the wavefunctions moduli at a bias of 74 mV/period, for the structure of [23], based on resonant phonon injection and extraction. The upper and lower laser levels are labeled by u and l, respectively, i label the injector level and e the extractor level. The simulations are carried out for a lattice temperature of 140 K.

responding relative changes are a reduction by 42% (Expon 1) and an increase by 10% (Expon 2), respectively. Γ is dominated by intra-subband scattering with low q, hence the similar Γ for the Gauss and Expon 1 distributions. Expon 2 has a lower Γ, as expected from the lower $f(q)$ at low q. From this reasoning, the Expon 1 distribution is expected to have the largest Γ, however we find that Gauss results in a somewhat higher value, for which we currently have no clear explanation.

A gain of $\sim$9 cm^{-1} is required in order to overcome the total losses of the experimental sample [26], and this is observed at the experimental threshold field of 48 kV/cm. All three roughness distributions agree reasonably with the experimental threshold current density, the Gauss and Expon 2 requiring slightly higher losses and the Expon 1 slightly lower, and both the DM and NEGF models provide the same threshold field when the same distribution is employed. Finally, we note that the DM model provides significantly larger gain than the NEGF model at higher fields. The reasons are not yet fully understood, however similar output powers are found if gain saturation is considered [26].

4.2. Scattering assisted injection THz QCL

For the THz structures studied below, the exponential distribution function with $\tilde{\Lambda} = 10$ nm and $\tilde{\Delta} = 0.2$ nm provide results in close agreement with experimental data [23, 27]. Employing translations (4) and (5) thus provides the corresponding Gaussian distributions displayed in Fig. 1(b). Here Gauss 1 agrees with the exponential at low q while Gauss 2 agrees better at higher q.

First, we study the scattering assisted injection design presented in [23]. For this structure, see Fig. 6, the laser level separation is $\sim$ 14 meV, which corresponds to a momentum transfer of $q \sim 0.16$ nm^{-1}. One would therefore expect that in this case the intermediate-q is of high relevance. Furthermore, the electron transport through the device is also relying on two different optical phonon resonances (with an energy difference matching $q \approx 0.25$ nm^{-1}) and a tunneling resonance at the main operation point. This explains the differences in the current densities shown in Fig. 7(a), where the current essentially increases with the size of the scattering matrix

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

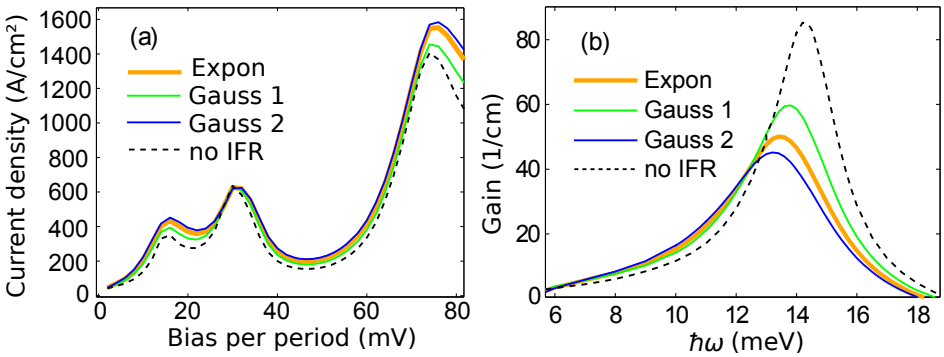


Fig. 7. (a) Current-field simulations of the THz QCL [23] using the NEGF model with different roughness distributions given in Fig. 1(b). The simulations are carried out for a lattice temperature of 140 K. (b) Gain at 74 mV/period for the THz QCL [23], using different roughness models. The linewidths Γ of the gain peaks are without interface roughness (IFR): 2.8 meV; for Gauss 1: 3.6 meV; for Expon: 3.9 meV; and for Gauss 2: 4.3 meV.

elements in the range of $0 < q \lesssim 0.25$ nm^{-1} as determined by $f(q)$ displayed in Fig. 1(b).

These effects are more pronounced at lower temperature, as phonon scattering becomes weaker while roughness scattering is less temperature dependent. However, at low temperature, the NEGF simulations provide an extremely non-thermalized carrier distribution in the upper laser state, which appears to be an artifact due to the neglect of electron-electron scattering. Thus we consider our results as not reliable under these conditions.

Figure 7(b) shows the gain evaluated near the current peak. As expected, the peak gain decreases with the scattering strength around $q = 0.16$ nm^{-1}, which dominates the scattering from upper to the lower laser level. Furthermore, the linewidth is affected, which enhances the impact of the different roughness models. Here, the difference in linewidth between the Gauss 1 and Expon roughness distribution is actually the least, as low q-scattering dominates the width. Neglecting interface roughness altogether (dashed lines) gives a large and sharp gain peak, showing that roughness plays a vital role, deteriorating the performance of the QCL at 140 K for this design.

4.3. Resonant Tunneling Injection THz QCL

As the resonant-tunneling designs have received much attention and have also repeatedly broken the temperature record, it is of interest to expand this study to cover those designs as well. In this work we focus on the three well structure shown in Fig. 8 investigated by Kumar [27] which was later improved to reach operation temperatures of ~ 200 K [28].

A collection of simulation results is presented in Fig. 9. Let us first consider the validity of the simulations by comparing the *ab initio* calculations to the experimental results of [27]. Threshold current is reported to increase from 410 A/cm^2 at 5 K *heatsink* temperature to 800 A/cm^2 at 180 K. At the design bias of 56 mV/period the low temperature simulations provide a high gain of the order of 100 cm^{-1}and a current density above the experimental threshold (except for Gauss 1), consistent with the onset of gain at a lower current. While the current peak without laser operation, only reaches ≈ 550 A/cm^2, a significantly higher current is found employing simulations under laser operation [29], shown by crosses in Fig. 9(a). This is consistent with experimentally measured currents for this sample. For higher temperatures gain drops and currents around ≈ 800 A/cm^2 match well the experimental threshold current.

The results for different roughness distributions follow the same trend as for the scattering-

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

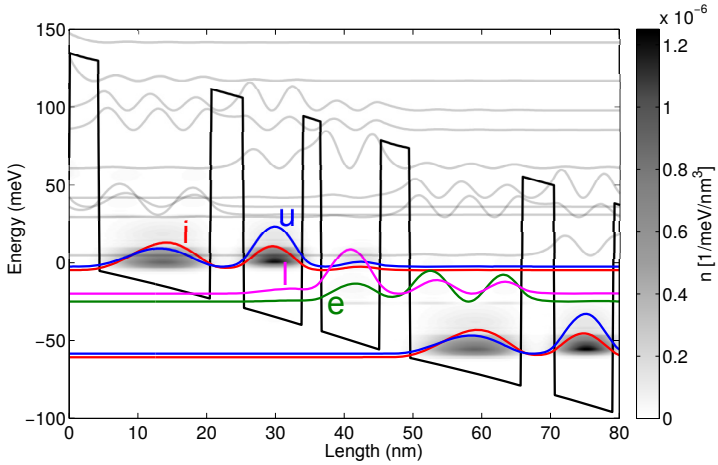


Fig. 8. Electron density and square of the wavefunctions moduli at a bias of 56 mV/period, for the structure of [27], based on resonant tunneling injection from the injector level (i) into the upper laser level (u) and phonon extraction from the extraction level (e). The lower laser level is labeled by l.

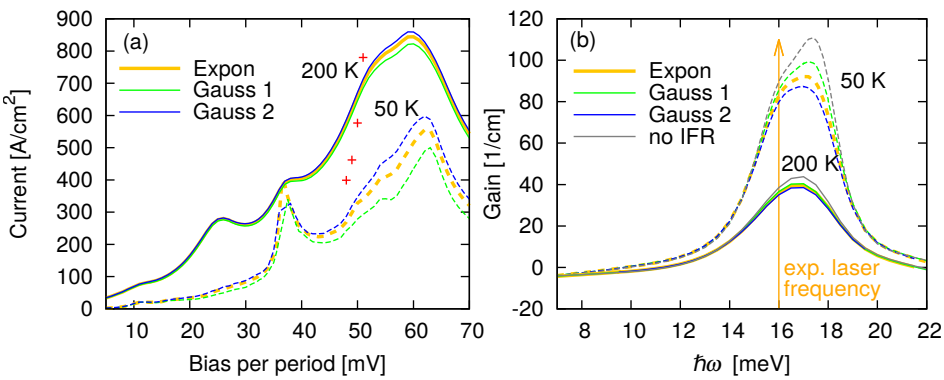


Fig. 9. (a) Current field characteristics of the resonant-tunneling design demonstrated in [27] at both high and low temperatures for different roughness distributions given in Fig. 1(b). Crosses show current densities under laser operation, assuming total losses of 30 cm^{-1}. (b) Gain simulations at the same temperatures where in addition the case of no interface roughness is also included.

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

assisted design. Comparing the low temperature threshold current indicates that roughness scattering in accordance with the Expon or Gauss 2 distribution is adequate. In general the differences between the roughness distributions are slightly less significant than for the scattering-assisted design addressed above, as there are fewer interfaces per length present in this design. For low temperatures, switching on the scattering given by the exponential correlation function leads to a decrease of peak gain with 20%, which is also consistent with Monte-Carlo studies for resonant-tunneling designs [30].

Below threshold, the current-field characteristics in Fig. 9(a) show a pre-peak around 37 mV/period. This is due to the intermediate resonance between the injector state and the extraction state as shown in Fig. 8, in accordance with the discussion in [27]. At low temperatures the experiment observes a plateau region for a bias range of ≈ 1.6 V with a current of ≈ 400 A/cm^2 after which lasing sets in. This is consistent with our calculated current-field relation, if one takes into account domain formation for biases surpassing the pre-peak [31]: In this case a high-field domain forms with a field around 50 mV/period, while the current and the low-field domain is locked at the pre-peak. Within the high-field domain we calculate a gain of 50/cm and thus lasing sets in if approximately half of the device (222 periods in total) are covered by the high-field domain. The difference between the fields in the low- and high-field domain, thus suggests a bias range of 1.4 V for the current plateau before lasing sets in.

5. Conclusions

We have studied the relevance of different distribution functions for the interface roughness in Quantum Cascade Lasers. In principle, exponential distribution functions provide a slower decay of scattering with increasing transition wave-vector q, compared to Gaussians, which can be related to a finite slope of the spatial correlation function at the origin. We find that for IR QCLs, the scattering at large q values is most important and that the results for Gaussian and exponential distribution functions are comparable, if they provide similar matrix elements in this region. These findings are recovered by different simulation schemes, which demonstrates that they hold beyond specific approximations in the respective models. In contrast, for THz QCLs scattering at intermediate q is more relevant and we showed that the width of the gain spectrum is most sensitive to low q scattering. For all structures studied here, increased roughness scattering enhances current and deteriorates gain. However, scattering matrix elements with different ranges of q are of relevance for IR and THz structures, which has to be taken into account, when choosing a model distribution functions.

Furthermore we have shown that corrections due to the valence band components used for the treatment of non-parabolicity are negligible for THz QCLs and are of minor importance for IR-QCLs in the 8 μm region.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 317884, the collaborative Integrated Project MIRIFISENS and the Swedish Research Council (VR).

#229159 - $15.00 USD Received 4 Dec 2014; revised 15 Jan 2015; accepted 15 Jan 2015; published 20 Feb 2015

Paper VI

JOURNAL OF APPLIED PHYSICS **118**, 114501 (2015)

Influence of interface roughness in quantum cascade lasers

K. A. Krivas, D. O. Winge, M. Franckié, and A. Wacker[a)]
Division of Mathematical Physics, Lund University, Box 118, Lund 221 00, Sweden

(Received 18 June 2015; accepted 27 August 2015; published online 15 September 2015)

We use a numerical model based on non-equilibrium Green's functions to investigate the influence of interface roughness (IFR) scattering in terahertz quantum cascade lasers. We confirm that IFR is an important phenomenon that affects both current and gain. The simulations indicate that IFR causes a leakage current that transfers electrons from the upper to the lower laser state. In certain cases, this current can greatly reduce gain. In addition, individual interfaces and their impact on the renormalized single particle energies are studied and shown to give both blue- and red-shifts of the gain spectrum. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4930572]

I. INTRODUCTION

Quantum cascade lasers (QCLs)[1,2] have proven to be useful devices with important applications, as they can be designed to emit in the region of 5–14 μm, crucial to molecular spectroscopy.[3] These lasers are solid state devices that employ mini-bands to achieve population inversion and thereby lasing in semiconductor heterostructures. Due to their high possible wall-plug efficiency[4] and capability to operate in the mid-infrared (mid-IR) and terahertz regions of the electromagnetic spectrum that are poorly covered by other coherent radiation sources, these devices have attracted a lot of attention.

While mid-IR QCLs are already industrialized, possible applications[5] of THz QCLs are not reached due to a lack of room temperature operation. Difference frequency generation using mid-IR QCLs has proven to be a way, although with limited power.[6] In order to reach high temperature operation of THz QCLs, it is necessary to understand the underlying mechanisms that govern the operation of these devices.

The two main causes of elastic scattering in QCLs are impurities, where electrons are scattered by the dopant ions, and interface roughness (IFR), which provides scattering due to imperfections of the interfaces between two semiconductor layers. The interface roughness scattering[7,8] is dominant in mid-IR QCLs,[9] and is also relevant in THz QCLs.[10] It affects the occupation of states by scattering electrons from one mini-band into another, due to the lack of lateral symmetry at the interfaces. It was shown by Deutsch *et al.*,[11] by producing symmetrical lasers from materials, in which IFR depends on growth direction, that interface roughness scattering strongly affects the operation of THz QCLs.

Using our non-equilibrium Green's function (NEGF) model,[12] we investigate the influence of interface roughness in detail both with respect to growth direction and roughness fluctuations of individual interfaces. The use of a complicated and computationally expensive model is motivated by the access to several important features. Among these effects is the renormalization of the single particle energies from the scattering potentials. Thus, altering IFR can provide shifts of the energy levels. Furthermore, the approach is able to capture effects such as dispersive gain[13] and gain linewidth reductions due to correlations, which are crucial in THz QCLs.[14] These effects are unique to the NEGF scheme, thanks to the use of two times in the evaluation of the density matrix. A summary of different methods for modeling QCLs can be found in Ref. 15.

II. THEORY

In our numerical computation method documented in Ref. 12, the IFR scattering enters the equations through the self-energy, a physical quantity that describes the interaction of the particle with its surroundings. The real part of the self energy gives a shift in the energies of the single particle states, while the imaginary part is related to its lifetime.

IFR is characterized by the autocorrelation function for the spatial distributions of the deviation from a perfect interface. In this work, we chose an exponential autocorrelation function, given by

$$\langle f(\mathbf{r}) f(\mathbf{r}') \rangle = \eta^2 \exp\left(- \frac{|\mathbf{r} - \mathbf{r}'|}{\lambda} \right). \tag{1}$$

The Fourier transform of this function is

$$f(q) = 2\pi \frac{\eta^2 \lambda^2}{(1 + (q\lambda)^2)^{3/2}}, \tag{2}$$

where q is the absolute change in wavevector, and the two variables λ and η describe the correlation length and root mean square deviation from a perfect interface, respectively.

A Gaussian distribution function is another common choice of autocorrelation function. However, as shown in Ref. 10, it is possible to obtain similar results by an appropriate transformation of parameters between both autocorrelation functions. Since there is no convincing physical argument for either of them, we normally use the exponential type.

In Fig. 1, we display Eq. (2) for the reference case and also for the case when η and λ are separately increased. It is possible to see that increasing the parameters λ and η have different effects on the scattering. While η increases the

a)Electronic mail: andreas.wacker@fysik.lu.se

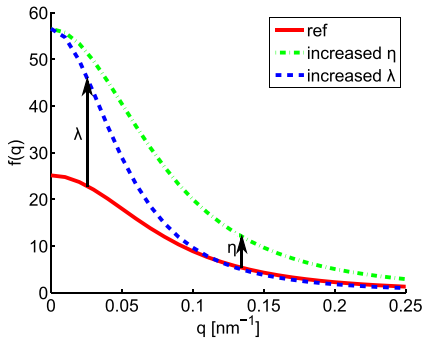


FIG. 1. Different forms of the autocorrelation function, Eq. (2). The full line is the reference case with $\eta = 0.20$ nm and $\lambda = 10.0$ nm, the dashed line shows the case when λ is increased by 50% and the dotted-dashed shows the behavior for the same increase of η.

scattering over the entire range of q values, λ tends to emphasize low momentum transfer. Comparing results when these parameters are increased separately could give insight into how important scattering with large momentum transfer is.

As our non-equilibrium Green's function model applies the self-consistent Born approximation in the calculation of self-energies, multiple-scattering events with a single imperfection are neglected. Thus, we cannot reproduce any bound states due to disorder, which might cause distinct effects of inhomogeneous broadening. However, such effects only become of relevance for larger spatial correlation lengths as discussed in Section IV of Ref. 16. In order to quantify this, we consider the energy balance for localization at an island of size λ and thickness η. The possible gain in energy at an island with a locally enlarged well width is about $\eta \Delta E_c |\Psi(z_i)|^2$, where ΔE_c is the conduction band offset and $\Psi(z_i)$ the wave function at the interface. However, the lateral localization costs an energy larger than $\hbar^2/(m_c \lambda^2)$, where m_c is the effective mass. Thus, we can exclude any localization effects as long as

$$\lambda^2 \eta < \frac{\hbar^2}{m_c \Delta E_c |\Psi(z_i)|^2}. \tag{3}$$

For all interfaces considered in this study, the right hand side is at least 100 nm^3 (for the thin barrier in the four-well laser). Thus, the inequality holds even for the enlarged values $\lambda = 15$ nm and $\eta = 0.3$ nm.

An alternative approach to study these issues is the use of exact eigenstates.[7] We could actually show that the lineshape of our model agrees with such calculations very well,[17] which justifies the Born approximation for interface roughness.

III. DEVICES STUDIED

In this work, the influence of IFR scattering is investigated using three different terahertz QCL designs, namely, a two-well,[18] a three-well,[19] and a four-well[20] structure. The first one employs three states per period for electron transport: an upper lasing state (ULS), a lower lasing state (LLS), and an injector-extractor state (i-e), as shown in Fig. 2. This laser operates over a range of frequencies from 2.8 to 4.1 THz, with a maximum reported operating temperature of 125 K. The second (three-well) laser is of resonant phonon design and therefore has separate states for injection (i) and extraction (e). The reported lasing frequency is 3.9 THz, and the reported maximum temperature of operation is 186 K. The band diagram of this laser is displayed in Fig. 3. The last investigated QCL employs a scattering assisted design. It relies on 4 mini-bands distributed over four wells per period. This laser operates at 3.2 THz at the maximum temperature of 138 K. The band diagram is shown in Fig. 4.

IV. RESULTS

We investigate the influence of IFR scattering by altering the interface roughness parameters in the simulations. These results are then compared to simulations with unaltered IFR. As a reference, we use the parameters $\lambda = 10$ nm and $\eta = 0.20$ nm. The IFR of the altered interfaces is chosen to have one of these two parameters increased by 50%. The interfaces are also assumed to be uncorrelated, so that one interface distribution does not depend on the others.[21] All simulations are performed for a lattice temperature of 200 K.

It is known that interface roughness can depend on growth direction.[11] Therefore, increasing IFR on every second interface would recreate the naturally occurring configuration. The applied bias tilts the potential wells and lets us distinguish between two different cases: first, when the altered interfaces are on the lower potential side of the wells (*wb*), and second, when the altered interfaces are on the higher potential side of the wells (*bw*).

First, we investigate the effect of IFR scattering on the current density. The results when either half of the interfaces or all of them are altered are shown in Fig. 5. If the changes

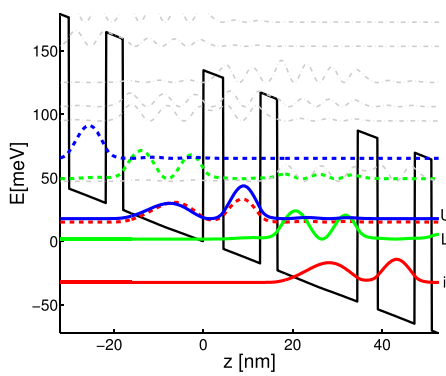


FIG. 2. Band diagram of the two-well laser[18] at 47.5 mV per period with respect to the growth direction z. The conduction band profile is shown together with the probability density for the most important subbands at their respective energies.

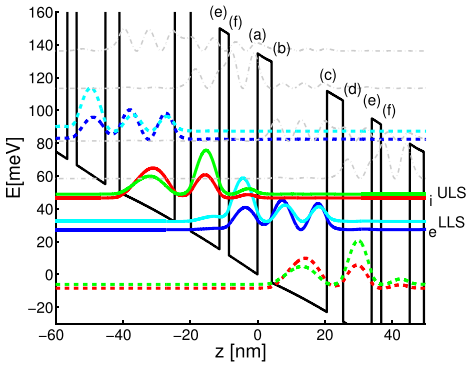


FIG. 3. Band diagram of the three-well laser,[19] in the same way as Fig. 2 with labels added to distinguish the specific interfaces. The bias is 55 mV per period.

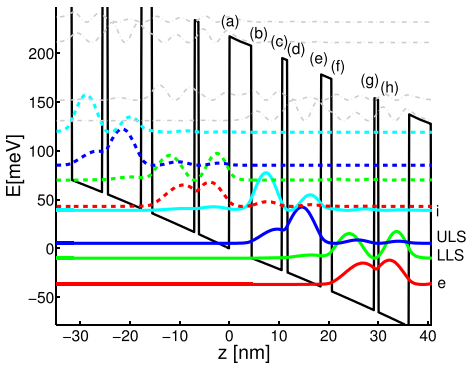


FIG. 4. Band diagram of the four-well laser[20] plotted in the same way as in Fig. 3, at a bias of 80 mV per period.

in current due to a change in IFR are small, we expect that if we increase the interface roughness on two interfaces simultaneously, we obtain the same change in current density as if we would add the changes in current densities from simulations when these two interfaces have roughness increased separately. As can be seen from Fig. 5, it is actually possible to relate the magnitude of the increase in current density to how well this superposition holds. For instance, the two-well laser exhibits the lowest increase in current (4%) and shows the best agreement between simulated relative current and the sum, while the four-well laser shows the least agreement, and the increase in current is the highest (21%). The three-well laser is an intermediate case, having 11% increase in current density when η is increased by 50% on all interfaces.

It can be seen in Fig. 5 that altering *wb* interfaces causes a larger increase in current density than *bw* interfaces. This can be understood by the effect that the wavefunctions tend to shift to the lower potential side of the well when a bias is applied, as can be seen in Figs. 2–4. This results in higher wavefunction values at the interfaces at the lower potential side of a well. Since IFR scattering is proportional to the product of the wavefunction values at the interfaces, changing IFR on interfaces with high wavefunction values has a larger impact on the transport. This observation confirms the results shown in Ref. 11.

For the two- and three-well lasers, the results of Fig. 5 can be understood using the reasoning above. Here, the (*wb*) interfaces dominates the IFR scattering. However, in the four-well case, the changes due to *wb* and *bw* are approximately the same. This is an effect of the thin barriers, as the value of the wavefunction of the ground state is actually lower on the high potential side of the barrier (or, equivalently, on the low potential side of the foregoing well) as seen in Fig. 4. This is because thin barriers are placed where the ground state wavefunctions have their maximum, rather than their minimum, value. However, the thick barriers act as in the cases of the other two lasers.

The simulated relative gain spectra compared to unaltered IFR, for the respective devices are shown in Fig. 6, and all devices show an overall decrease in gain as a result of increasing IFR. However, the magnitude of the effect differs widely. Again, the two-well laser is the most insensitive to changes in IFR, with a decrease in gain of 7.2%. The three-well laser displays a 19.1% reduction, while the gain of the four-well laser shows a decrease by 50.6% when roughness parameter η was increased by 50% on all interfaces.

In order to determine in detail the influence of individual interfaces on the current density and gain, we now change only one interface at a time and compare these results to the reference case. These results are shown in Fig. 7 for the three-well laser. The effects of the changes are twofold; we

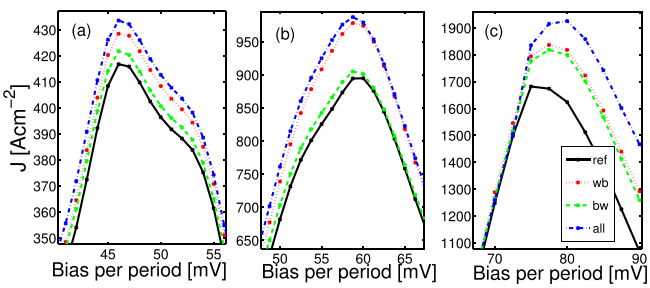


FIG. 5. Results of current density simulations. (a) Two-well laser,[18] (b) three-well laser,[19] and (c) four-well laser.[20] The parameter η is increased by 50% with respect to the reference calculation (ref) for different sets of interfaces.

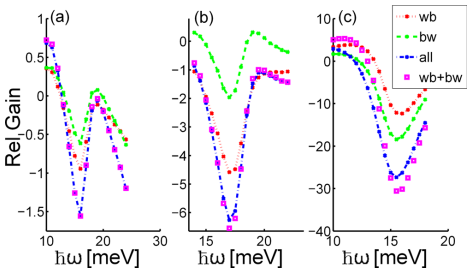


FIG. 6. Relative gain at operation bias *per period*. In (a), the two-well laser[18] at 47.5 mV, (b) the three-well laser[19] at 55 mV, and (c) four-well laser[20] at 80 mV. In this study, η is increased.

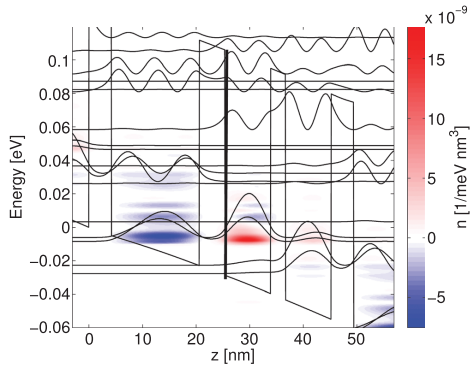


FIG. 8. Changes in spatially and energetically resolved charge density due to increase of η on one interface (marked by black line). This configuration results in shift in gain as shown in Fig. 7(d). Results for the three-well laser for a bias of 57.5 mV per period.

see both reductions and shifts in the gain spectra. There is an overall decrease in magnitude due to the enhanced depopulation of the ULS, being the main effect in (a) and (e). This is also reflected in the significant increase in current. These two are *wb* interfaces and, following the above discussion, crucial to the operation. In contrast, the main effect in (b), (c), and (f), is a blue-shift. As the scattering potential is increased, there will be a renormalization of the single particle energies. In our model, this is taken into account by the real part of the self energies, but it can be understood in general as a level shift in second order perturbation theory. The intraband scattering will shift the levels down in energy, and for the interfaces (b), (c), and (f), this effectively lowers the LLS, which gives a blue-shift. This interpretation is strengthened by the remaining case, Fig. 7(d) where the ULS is mostly affected. Here, in combination with a gain reduction, a red-shift is observed. In all cases, the impact of changing η and λ is similar, albeit λ has less impact due to the limited strength at higher q-values.

One of the possible causes for changes in gain and current density might be a leakage into high kinetic energy states of lower energy mini-bands. This leakage can be observed by investigating the change in spatially and energetically resolved electron densities. The change in electron density for the simulation of Fig. 7(d), shown in Fig. 8, shows an increase in the occupation near the bottom of the ULS. Thus, in accordance with Fig. 7(d), this is not expected to result in a significant decrease in gain. In contrast, the change in densities shown in Fig. 9, which correspond to Fig. 7(e), shows an increase in charge density at higher energies in the LLS. This indicates that electrons are scattered elastically and then relaxes by optical and acoustic phonon scattering. This lowers the inversion between the ULS and LLS, and consequently reduces the gain.

Comparing the values in Fig. 7, one finds that interfaces at the lower side of the well affect gain more, in agreement with the results from the current density simulations. Large decreases in gain are observed at the interfaces where the ULS and LLS have high wavefunction values.

For the case of the four-well QCL, shown in Fig. 4, the thicker barriers affect gain similarly to the other structures. However, the effect of the interfaces of the thin barriers is dominating the reduction in gain, as displayed in Fig. 10. For

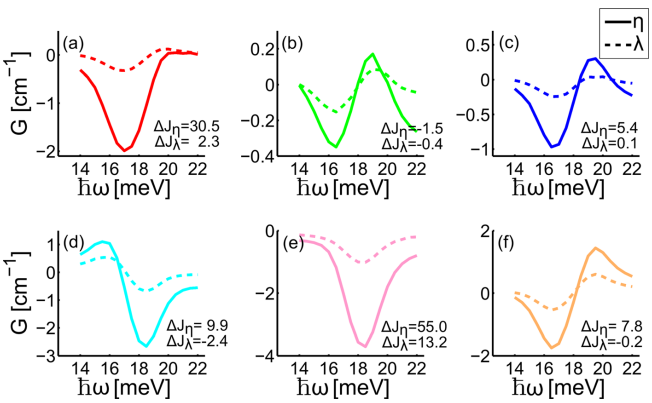


FIG. 7. Changes (relative gain) due to alteration of a single interface in each period of the three-well laser.[19] Also included are the changes in current in units of A/cm^2. The respective interfaces are denoted in Fig. 3. Either η or λ were increased by 50%. The bias is 57.5 mV per period.

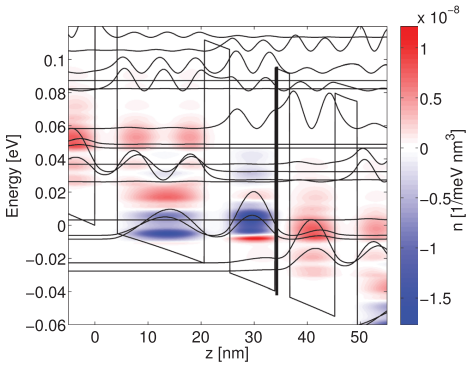


FIG. 9. Changes in spatially and energetically resolved charge density due to increase of η on one interface (marked by black line). Simulation shows strong reduction in gain as shown in Fig. 7(e). Results for the three-well laser for a bias of 57.5 mV per period.

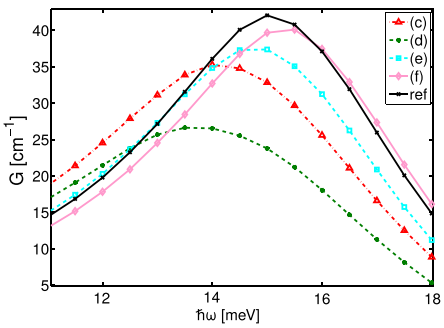


FIG. 10. Gain spectrum of the four-well laser. η is increased on the indicated interfaces, as labeled in Fig. 4, by 50%. Only the most important interfaces are shown for clarity.

these barriers, the aforementioned rule, that the (*wb*) interfaces are the most important, does not hold. This can be understood by examining the probability density of the ULS in Fig. 4. Here, it is clear that it has a larger value at interface (d) compared to (c).

V. CONCLUSION

In this work, we numerically investigate the influence of interface roughness on the operation of several THz QCL designs. We confirm that IFR scattering is an important phenomenon that may greatly affect both current density and gain. The most sensitive interfaces are the ones, where the wavefunctions have significant values. This makes interfaces at the lower potential side of wells more important than those on the side of higher potential. Thin barriers work differently, impacting transport and gain significantly more since they are usually placed where the ground state wavefunctions have their maximum values. Increased current and decreased gain indicate that a certain leakage current forms at the interfaces. At certain interfaces, this leakage mechanism strongly reduces population inversion. Due to alterations in the scattering potential, a shift of the gain peak frequency is observed for some interfaces. This can be explained by changes in the real parts of the self-energy.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under Grant Agreement No. 317884, the collaborative Integrated Project MIRIFISENS and the Swedish Research Council (VR).

[1]J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, Science **264**, 553 (1994).
[2]J. Faist, *Quantum Cascade Lasers* (Oxford University Press, Oxford, 2013).
[3]R. F. Curl, F. Capasso, C. Gmachl, A. A. Kosterev, B. McManus, R. Lewicki, M. Pusharsky, G. Wysocki, and F. K. Tittel, Chem. Phys. Lett. **487**, 1 (2010).
[4]Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, Appl. Phys. Lett. **98**, 181102 (2011).
[5]B. S. Williams, Nat. Photonics **1**, 517 (2007).
[6]M. A. Belkin, F. Capasso, F. Xie, A. Belyanin, M. Fischer, A. Wittmann, and J. Faist, Appl. Phys. Lett. **92**, 201101 (2008).
[7]J. B. Khurgin, Appl. Phys. Lett. **93**, 091104 (2008).
[8]Y. Chiu, Y. Dikmelik, P. Q. Liu, N. L. Aung, J. B. Khurgin, and C. F. Gmachl, Appl. Phys. Lett. **101**, 171117 (2012).
[9]Y. V. Flores, S. S. Kurlov, M. Elagin, M. P. Semtsiv, and W. T. Masselink, Appl. Phys. Lett. **103**, 161102 (2013).
[10]M. Franckié, D. O. Winge, J. Wolf, V. Liverini, E. Dupont, V. Trinité, J. Faist, and A. Wacker, Opt. Express **23**, 5201 (2015).
[11]C. Deutsch, H. Detz, T. Zederbauer, A. M. Andrews, P. Klang, T. Kubis, G. Klimeck, M. E. Schuster, W. Schrenk, G. Strasser, and K. Unterrainer, Opt. Express **21**, 7209 (2013).
[12]A. Wacker, M. Lindskog, and D. O. Winge, IEEE J. Sel. Top. Quantum Electron. **19**, 1200611 (2013).
[13]A. Wacker, Nat. Phys. **3**, 298 (2007).
[14]F. Banit, S.-C. Lee, A. Knorr, and A. Wacker, Appl. Phys. Lett. **86**, 041108 (2005).
[15]C. Jirauschek and T. Kubis, Appl. Phys. Rev. **1**, 011307 (2014).
[16]T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akiyama, J. Appl. Phys. **93**, 1586 (2003).
[17]C. Ndebeka-Bandou, F. Carosella, R. Ferreira, A. Wacker, and G. Bastard, Appl. Phys. Lett. **101**, 191104 (2012).
[18]G. Scalari, M. I. Amanti, C. Walther, R. Terazzi, M. Beck, and J. Faist, Opt. Express **18**, 8043 (2010).
[19]S. Kumar, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **94**, 131105 (2009).
[20]E. Dupont, S. Fathololoumi, Z. R. Wasilewski, G. Aers, S. R. Laframboise, M. Lindskog, S. G. Razavipour, A. Wacker, D. Ban, and H. C. Liu, J. Appl. Phys. **111**, 073111 (2012).
[21]S. Tsujino, A. Borak, E. Muller, M. Scheinert, C. Falub, H. Sigg, D. Grutzmacher, M. Giovannini, and J. Faist, Appl. Phys. Lett. **86**, 062113 (2005).

Appendix: Tables with details for simulated structures

Table 1 contains the material data used for structures with well of the specified material compositions. This data is mainly used for calculating the phonon scattering matrix elements and gain.

Table 1: Material parameters for the well materials of the QCLs examined in this thesis. GaAs is used for all THz structures. $In_{0.53}Ga_{0.47}As$, which is lattice matched to InP is used for the structure of Ref. [25] studied in Papers IV and V, as well as the structures in Sec. 5.2. $In_{0.71}Ga_{0.29}As$ is used in Secs. 5.3 and 5.4. m^* is the effective mass as defined in eq. (2.10). ϵ_r and ϵ_∞ are respectively the static and high frequency relative dielectric functions, and E_{LO} is the optical phonon energy. $V_{\mathsf{def.}}$ is the deformation potential, v_l the lognitudinal sound velocity, ρ_{mass} the mass density, and V_{mol} the mol volume, used for accoustic phonon scattering.

Material	m^*	ϵ_r	ϵ_∞	E_{LO}	$V_{\mathrm{def.}}$	v_l	ρ_{mass}	V_{mol}
GaAs	0.067	13.0	10.89	36.7	7.2	4730	5317	0.04517
$In_{0.53}Ga_{0.47}As$	0.043	13.9	11.6	34	7.79	4253	5500	0.08610
$In_{0.71}Ga_{0.29}As$	0.0336	14.3	11.9	34	7.79	4082	5576	0.0525

Table 2 shows simulation parameters used for simulating the mid-IR samples in this Thesis. Their layer sequences are shown in Tables. Table 3 shows parameters for the THz structures simulated in Secs. 3.3 and 6. The layer sequences of the structures are given in their respective referenced publications or in the sections referenced.

Table 2: Simulation parameters for the mid-IR QCLs discussed in Chap. 5 and Papers IV and V.

Ref.	ΔE_c	E_g/m^*	m_{well}	m_{barr}	IFR distr.	η	λ	$n_{2\text{D}}$
[25]	0.520	18.35	0.043	0.08	Gauss	0.1	9	7.8
					Expon 1	0.17	3.6	
					Expon 2	0.1	6.3	
Sec. 5.2	0.520	18.35	0.043	0.08	Expon 1	0.17	3.6	5.46
Sec. 5.3	0.63/1.41[a]	20.91	0.039	0.0978/0.15[a]	Gauss	0.1	9	64.2

[a]For InAlAs/AlAs barriers, respectively

Table 3: Simulation parameters for the THz QCLs discussed in Chap. 6 and Papers I-III and VI.

Ref.	ΔE_c	E_g/m^*	m_{well}	m_{barr}	IFR distr.	η	λ	$n_{2\text{D}}$
[37]	0.2175	22.67	0.067	0.0877	Expon	0.2	10	3.25
Paper II	0.2175	22.67	0.067	0.0877	Expon	0.2	10	3.46
[90]	0.1347	22.67	0.067	0.07945	Expon	0.2	10	3.0
Chap. 6.2, [39], [31]	0.1347	22.67	0.067	0.07945	Gauss	0.1	9	1.5-3