



Ellis, Aneirin (2025) *Efficiency limiting processes in epitaxially grown mid-infrared lasers on silicon*. PhD thesis.

<https://theses.gla.ac.uk/85153/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

UNIVERSITY OF GLASGOW

Efficiency Limiting Processes in
Epitaxially Grown Mid-Infrared Lasers
on Silicon



University
of Glasgow

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy (Ph.D.)

Aneirin Ellis

College of Science & Engineering
James Watt School of Engineering

April 2025

Abstract

The field of Si photonics is concerned with the development of photonic components that are compatible with highly mature complementary metal-oxide (CMOS) fabrication processes. Recently, optoelectronic integrated circuits (OEICs) operating in the mid-infrared (mid-IR) have garnered interest, driven by their suitability for lab-on-a-chip spectroscopy. Despite significant progress in passive components, the lack of high-quality CMOS compatible light sources is bottlenecking the production of fully integrated circuits. Various methods of incorporating lasers onto Si exist, however, long-term, monolithic epitaxial growth is preferential for scalable, low-cost manufacturing. In this thesis, two different approaches for epitaxial integration of mid-IR light-sources on Si are investigated.

The first approach is to move away III-V lasers, instead utilising direct bandgap group IV materials. The performance of first-generation electrically injected bulk GeSn lasers were probed using high hydrostatic pressure at low temperature. The devices exhibited threshold current densities of 1.2 kA/cm^2 at 85 K which rose with increasing pressure. 93% of the threshold is attributable to large defect related losses in the buffer layer, but even at low temperatures over 1% of carriers are found to occupy states in the L-valley. Temperature-dependent modelling of the carrier distribution was then used to assess the feasibility of GeSn as a bulk active region material. This showed that Sn contents in excess of 20% were required to ensure over 90% of carriers occupy the Γ valley at room temperature (RT). As such, it is likely that the upper operational temperature is limited by indirect carrier losses and a change of heterostructure design is required to achieve electrically injected devices close to RT.

The second approach involves heteroepitaxial growth of mature Sb-based lasers on Si. In this study, devices on Si operating at $2.3 \mu\text{m}$ are compared to those on native GaSb emitting at a similar wavelength. Temperature-dependent measurements were used to confirm heightened defect related recombination in the devices on Si, and illustrated the onset of a highly temperature dependent process above 150 K. A combination of high hydrostatic pressure measurements and modelling suggest that hole leakage and leakage of carriers to confined L-valley states were likely contributing to the observed temperature dependence in the lasers on Si. Adaptations to the active region and waveguiding layers are proposed to improve the performance of future devices.

Declaration of Originality

I certify that the thesis presented here for examination for a PhD degree of the University of Glasgow is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice. The copyright of this thesis rests with the author. No quotation from it is permitted without full acknowledgement. I declare that the thesis does not include work forming part of a thesis presented successfully for another degree. I declare that this thesis has been produced in accordance with the University of Glasgow's Code of Good Practice in Research. I acknowledge that if any issues are raised regarding good research practice based on review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Signature

Date 04/12/2024

Acknowledgements

*"Let me not be so vain to think that I'm the sole
author of my victories and a victim of my defeats."*

— Ze Frank

I have countless people to thank for helping me reach this point in my academic journey. While this list is not exhaustive, it serves as a small reflection of the many individuals whose support has been crucial to my progression.

Firstly, I would like to thank my supervisor, Professor Stephen Sweeney, for providing me with the opportunity to undertake this project and for his continual support throughout. Aside from the invaluable guidance, technical expertise, and opportunities provided, I am especially grateful for his patience and understanding, which culminated in an enjoyable and productive four years of work. I would also like to express gratitude to secondary supervisors, Professor Steve Clowes, Dr Luca Sapienza and Professor Vihar Georgiev, for contributing their time, knowledge and support.

I am deeply indebted to Dr Dominic Duffy for allowing me to utilise his software, `kpymod`, during the final two years of this project, which enabled me to significantly progress my work. My appreciation for the countless technical conversations, support and friendship cannot be understated. I've also been extremely fortunate to benefit from the mentorship of Dr Igor Marko, whose knowledge, time and assistance in all aspects of the experimental lab work were invaluable during this project.

I'd like to also extend gratitude to fellow group members Anoma Yamsiri, Aidas Baltušis, Connor Munro, Jingzhao Liu, Xingyu Zhao and Matthew Robinson, whom I've had the pleasure of sharing an office with at various points over the last few years. I wish you all nothing but the best moving forward. A further mention must go to Dr Adam McKenzie for his assistance with device processing.

Outside of the university, I'd also like to acknowledge the contributions of our collaborators. The GeSn samples were provided by the group of Professor Shui-Qing (Fisher) Yu at the University of Arkansas. GaSb-based QW lasers and ICLs were fabricated by the group of Professor Laurent Cerutti and Professor Eric Tournié at the Université de Montpellier. I am also grateful to Sudip Acharya for a number of useful conversations on electrically injected GeSn lasers at various conferences.

Additionally, I would like to thank the University of Glasgow, the University of Surrey and the Engineering and Physical Sciences Research Council (EPSRC) for providing funding for my project and conference travel.

Finally, I would like to thank my friends and family for their unwavering support over the last four years. In particular, to Mum, Dad, Rich, Ian, and Tom and Ada. They have been there through every challenge, celebrated my small victories, and always encouraged me to keep going. It is to them, with deepest gratitude, that I dedicate this thesis.

Research Outputs

Journal Articles

1. [A. Ellis](#), D. A. Duffy, I. P. Marko, S. Acharya, W. Du, S. Q-. Yu & S. J. Sweeney. *“Challenges for room temperature operation of electrically pumped GeSn lasers”*. Scientific Reports **14**, 10318 (2024).

In Preparation

1. [A. Ellis](#), D. A. Duffy, I. P. Marko, L. Cerutti, E. Tournié & S. J. Sweeney. *“Carrier recombination processes in defect tolerant interband cascade lasers on Si”*.

Conference Presentations

1. [A. Ellis](#), I. P. Marko, T. D. Eales, L. Cerutti, M. R. Calvo, L. M. Bartolome, J. -B. Rodriguez, E. Tournié & S. J. Sweeney. *“Carrier recombination processes in 2.3- μ m epitaxially grown mid-infrared laser diodes on Si(001)”*. Poster presentation, given virtually at ISLC 2021, Potsdam, Germany (online).
2. [A. Ellis](#), I. P. Marko, T. D. Eales, L. Cerutti, M. R. Calvo, L. M. Bartolome, J. -B. Rodriguez, E. Tournié & S. J. Sweeney. *“Temperature and high hydrostatic pressure investigations of epitaxially grown 2.3- μ m GaSb lasers on Si”*. Poster presentation, given virtually at MIOMD 2021, Surrey, UK (online).
3. [A. Ellis](#), D. A. Duffy, I. P. Marko, S. Ojo, W. Du, S. Q-. Yu & S. J. Sweeney. *“The influence of indirect valleys on the performance of electrically injected, bulk GeSn lasers”*. Poster presentation given at UK Semiconductors 2023, Sheffield, UK.
4. [A. Ellis](#), D. A. Duffy, I. P. Marko, S. Acharya, W. Du, S. Q-. Yu & S. J. Sweeney. *“Temperature and band structure dependent properties of GeSn double heterostructure lasers”*. Oral presentation given at MIOMD 2023, Oklahoma, USA.
5. [A. Ellis](#), D. A. Duffy, I. P. Marko, S. Acharya, W. Du, S. Q-. Yu & S. J. Sweeney. *“Challenges for room temperature electrically injected GeSn semiconductor lasers”*. Oral presentation given at ISLC 2024, Orlando, USA.

"Life is nothing but an electron looking for a place to rest."

— Albert Szent-Györgyi

Contents

Abstract	i
Declaration of Originality	ii
Acknowledgements	iii
Research Outputs	v
1 Introduction	1
1.1 Mid-Infrared Lasers	1
1.2 Towards Si Photonics	3
1.3 Mid-Infrared Applications	6
1.3.1 Environmental, Industrial and Defence Applications	8
1.3.2 Medical Applications	9
1.4 Thesis Outline	11
2 Fundamental Theory	13
2.1 Electron Bandstructure and Optical Processes in Semiconductors	13
2.1.1 Carrier Distribution	15
2.1.2 Optical Processes in Semiconductors	17
2.2 Fundamentals of Semiconductor Lasers	18
2.2.1 Gain	19
2.2.2 Optical Feedback	20
2.2.3 The Threshold Gain Condition	22
2.3 Semiconductor Laser Heterostructures	23
2.3.1 Carrier Confinement	23
2.3.2 Quantum Confinement	25
2.3.3 Optical Confinement	27
2.3.4 Strained Layers	29
2.4 Carrier Recombination Mechanisms	33
2.4.1 Defect-related Recombination	34
2.4.2 Radiative Recombination	35

2.4.3	Auger-Meitner Recombination	36
2.4.4	Carrier Leakage	39
2.5	Internal Losses	41
2.5.1	Free Carrier Absorption	41
2.5.2	Intervalence Band Absorption	41
2.6	Summary	42
3	Experimental Techniques	43
3.1	General Set-up	43
3.2	Temperature Dependent Measurements	45
3.3	High Hydrostatic Pressure Measurements	48
3.3.1	Method	49
3.3.2	Low-Temperature, High Pressure Measurements	51
3.4	Light-Current Curves	52
4	Bulk Group-IV Bandstructure Modelling	55
4.1	Introduction	55
4.1.1	Device History	58
4.2	k-p Theory	61
4.2.1	The Single Band Model	62
4.2.2	The 6-band Luttinger-Kohn Model	65
4.2.3	The 8-band Model	69
4.2.4	Pseudomorphic Strain Corrections	72
4.3	Parameterisation	73
4.3.1	Linear Interpolation	74
4.3.2	Bowing	76
4.4	Evaluation of Model Accuracy	78
4.4.1	III-V Materials	78
4.4.2	Group IV Materials	79
4.5	Summary	79
5	Performance Limiting Mechanisms in Electrically Injected GeSn Lasers	81
5.1	Sample Information	81
5.2	High Hydrostatic Pressure Measurements	82
5.3	Temperature Dependent Modelling	88
5.3.1	Bulk Gain Modelling	89
5.3.2	Temperature Dependent Gain in Bulk GeSn	92
5.3.2.1	Residual Strain Optimisation	94
5.3.2.2	Bulk Composition Optimisation	99
5.4	Intervalence Band Absorption	102

5.5	Auger-Meitner Resonances	107
5.6	Summary & Conclusions	110
6	Recombination Processes in Epitaxially Grown GaInAsSb QW Lasers on Si	112
6.1	Introduction	112
6.1.1	Sample Information	114
6.2	Temperature Dependent Measurements	118
6.2.1	Threshold Current Density	118
6.2.2	Spectra	120
6.2.3	Characteristic Temperature	122
6.2.4	Spontaneous Emission	125
6.2.4.1	Z-Analysis	126
6.3	Pressure Dependent Measurements	128
6.3.1	Spectra	128
6.3.2	Threshold Current Density	129
6.3.3	Relative Quantum Efficiency	134
6.4	Modelling	136
6.4.1	Conduction Band Carrier Distribution	138
6.4.2	Valence Band Carrier Distribution	141
6.4.3	Layer Optimisation	144
6.4.3.1	Quantum Well Optimisation	144
6.4.3.2	Barrier and Waveguide Optimisation	146
6.5	Summary & Conclusions	150
7	Conclusions & Future Work	152
7.1	Electrically Pumped GeSn Lasers	152
7.1.1	Future Work	153
7.1.1.1	Alloying & Doping	154
7.1.1.2	Material Growth & Heterostructure Design	155
7.2	GaSb-based QW Lasers on Si	157
7.2.1	Future Work	158
	Bibliography	160
A	Further Notes on Bandstructure Calculations	180
A.1	The 3-band Dresselhaus-Kip-Kittel Model	180
A.2	The 6-band Luttinger-Kohn Model	185
A.3	Python Code	188

Chapter 1

Introduction

1.1 Mid-Infrared Lasers

In 2022 the semiconductor laser market was worth an estimated \$7.7B, with projections reaching \$13.3B by 2029 [1]. This largely reflects their importance in telecommunications networks which are expected to expand significantly to accommodate the requirements of increasingly data heavy machine learning (ML) and artificial intelligence (AI) applications [2]. However, semiconductor lasers are now also fundamental in numerous other sectors, ranging from industrial manufacturing to defence and healthcare. Such devices are favourable over alternative laser types due to long lifetimes, low power consumption, direct modulation capability and small-footprint, making them suitable for integration in compact optoelectronic circuits (OEICs). Additionally, devices can be made to emit over a wide range of wavelengths, spanning from the ultraviolet through to the far infrared (IR), by appropriately engineering the heterostructure design and material compositions. Since the industry has received significant historical investment [3], the manufacturing capabilities today are highly mature and scalable, resulting in relatively low cost lasers.

This work focuses on lasers operating in the mid-IR spectral region. The span of wavelengths encapsulating the mid-IR is somewhat arbitrary and often varies in literature depending on the field or application. In this work it is defined to be $\sim 2 - 12 \mu\text{m}$, which is commonly addressed using the *antimonide* material system. The antimonide family collectively refers to GaSb, InAs, InSb, AlSb and their multinary alloys which are near lattice matched to GaSb. From these materials, bandgaps of between 0.1 - 2.0 eV can be achieved corresponding to emission wavelengths of $\sim 0.6 - 12 \mu\text{m}$ [4]. Such a wide range

of attainable bandgaps permits a variety of band alignments, ranging from standard type I, to type II where electrons and holes are confined in different layers, or even type III broken-gap materials where the conduction band minima is below the valence band maxima in the adjacent layer. As such, this system has become the standard for producing interband cascade lasers (ICLs) and standard quantum well (QW) lasers in the mid-IR, whilst also being widely utilised in quantum cascade lasers (QCLs) [5].

Type I devices are able to cover the 2.0 - 3.7 μm spectral range. Such lasers have demonstrated high power emission up to 2.5 μm , boasting 1.8 W output under continuous wave (CW) pumping at RT with a 28% wall plug efficiency [6], and hundreds of mW output has been achieved for emission above 3 μm [7]. At this time, CW operation has been realised up to 3.44 μm [8], whilst pulsed devices can reach 3.7 μm [9]. A type I design has also been implemented in a cascading scheme, allowing much lower threshold current densities to be achieved in longer wavelength devices as a result of carrier recycling [10]. Above 3 μm , these devices exhibit threshold current densities of just $\sim 100 \text{ A/cm}^2$ compared to over 400 A/cm^2 in non-cascaded structures. However, such lasers necessitate higher turn-on voltages.

In type II devices, electrons and holes are confined in separate layers. This allows each layer to be independently optimised for their respective carrier type. As such, emission energies less than the bandgaps of the constituent materials can be attained, thus greatly extending the range of achievable wavelengths. Typically, type II active regions are implemented in a so-called ‘W’ structure, whereby two electron confining layers are used to sandwich a hole confining layer. Since the layers are thin, the electron wavefunction is coupled between wells resulting in increased electron-hole overlap and thus improved gain when compared to a single electron confining layer. Such active regions are most abundant in ICLs. GaSb-based ICLs have exhibited RT operation in the 2.8 - 6 μm spectral range [11], with lower threshold current densities than standard type I devices above 3 μm [12]. A minimum of $< 200 \text{ A/cm}^2$ is achieved for wavelengths between 3 - 4 μm , above which the value rises consistently to around 1 kA/cm^2 at 6 μm [12].

In QCLs lasing is achieved using only a single carrier type (predominantly electrons) through intraband transitions between quantised sub-band states. In doing so, the emission wavelength is decoupled from the bulk bandgap of the material and is instead predominantly dependent on the confinement and material composition of the active layers [13]. Additionally, their unipolar nature means that Auger-Meitner recombination, which

would otherwise hinder long wavelength operation, is inherently suppressed [14]. This combination of factors has enabled wavelengths ranging from $\sim 3 - 25 \mu\text{m}$ to be achieved using this technology [15]. However, carrier leakage caused by inter-valley scattering commonly degrades the performance of shorter wavelength QCLs lattice-matched to GaAs or InP [16].

In the last decade, group-IV lasers based on the GeSn material system have also emerged as promising candidates for future mid-IR applications. This alloy has gained traction primarily due to the ease with which it could be epitaxially integrated into Si-photonics circuits utilising pre-existing, mature CMOS processes [17]. Whilst Ge is not a strong emitter due to its indirect bandgap, alloying with the semi-metallic α -Sn pulls down the conduction band (CB) at the zone centre faster than the indirect valleys. This results in truly direct gap material for Sn contents above $\sim 8 \%$ [18]. These ideas culminated in the first optically pumped GeSn laser in 2015 [19], with RT lasing recently being reported in highly strained GeSn alloys [20], albeit under very high optical pump powers. Major progression was attained in 2020 with demonstration of the first electrically injected GeSn laser [21], however the operational temperature in such devices is currently limited to $< 140 \text{ K}$ [22].

1.2 Towards Si Photonics

Photonic integrated circuits (PICs) incorporate a number of photonic components, such as modulators, detectors, interferometers and light-sources, on a single chip. Light is guided to different parts of the circuit, where operations are performed, by way of an optical waveguide, which is the photonic equivalent of copper wires in electronics. PICs based on the InP materials system are now commonplace within telecommunications settings, such as data centres, stemming from inherently higher bandwidth density, large data transfer rates and lower power consumption compared to conventional electronics.

For many applications, it is useful to have both photonic and electronic components on a single chip, forming an OEIC. By fully integrating photonic and electronic components on chip, issues with separate packaging are avoided. For example, transmission efficiency can be improved since poor optical coupling is reduced by alleviating the need for precise fibre alignment. Furthermore, the single step production results in a highly scalable process with large yields, thereby reducing costs.

To integrate electronics in a low-cost manner, it is necessary to leverage highly mature CMOS processing technologies. As such, CMOS compatible material systems that can provide optical guiding and manipulation are required. The field concerned with the study of compatible components and their integration is known as Si photonics. In recent years, significant progress has been made in the production of passive CMOS-compatible components, such as photodetectors, high-speed modulators, multiplexors and waveguides, required for photonic circuitry, but integrating light sources on Si has proved to be a bottleneck.

Since Si and other group-IV materials are generally poor emitters, it is highly desirable to incorporate efficient, high-performance lasers formed from III-V multinary alloys, such as those discussed in the previous section. However, direct epitaxial growth of III-V material on group-IV substrates is greatly complicated by three fundamental differences:

1. **Polarity mismatch** - Si is classified as a polar material, whilst III-Vs are non-polar, owing to their two atom basis. This results in a disruption to the lattice periodicity at the interface between the two, leading to the formation of antiphase boundaries (APBs) and threading dislocations (TDs).
2. **Lattice mismatch** - Often the difference in the relaxed lattice constant of III-V and group-IV materials is large. Thus, when growing a layer of one atop the other, it becomes energetically favourable to form large densities of TDs to reduce strain.
3. **Thermal mismatch** - The thermal expansion coefficients of III-V materials are typically larger than that of Si. Hence, when cooled after epitaxial growth, the III-V layer shrinks faster, pulling the layer into tensile strain. This can again result in large dislocation densities, micro cracks or wafer bowing.

TDs and APBs introduce localised energy states into the bandgap, and therefore act as carrier traps. From here, carriers predominantly recombine non-radiatively, leading to significantly degraded device performance. Currently, there are four primary techniques for integrating lasers on Si, each with their own merits and drawbacks, these are:

1. **Flip-chip bonding** - Commonly referred to as hybrid integration, this method involves integrating single, pre-fabricated III-V lasers using an adhesive. In doing so, individual devices can be tested prior to integration, ensuring that only lasers meeting the necessary performance criteria are used. However, coupling light into the
-

circuit is highly dependent on the alignment of the source and waveguide. Thus, very precise processes are needed to attain acceptable in-plane alignment, and variance in the adhesive height can result in vertical misalignment. Hence, despite the use of high quality lasers, the overall efficiency of the produced OEICs may be low, and the requirement for precise positioning of individual devices limits manufacturing scalability.

2. **Wafer-to-wafer bonding** - This technique, also known as heterogeneous integration, involves bonding a full wafer containing the III-V laser stack onto the Si substrate. The material is then lithographically patterned and etched to produce the final laser structures. In doing so, precise alignment between the laser and waveguide can be achieved, resulting in enhanced optical coupling into the photonic circuit. Despite this, there are issues pertaining to excessive material wastage. Firstly, Si wafers in standard foundries typically have diameters of 200 - 300 mm. In contrast, the brittleness of III-V materials means GaAs and InP wafers are limited to 150 mm and 100 mm, respectively [23]. This size mismatch means a significant area of the Si wafer cannot be utilised, therefore increasing the cost per unit. These costs are further exacerbated by the amount of III-V material that is not used. Typically, in excess of 95% of the expensive III-V material is etched away during the processing step [24]. Furthermore, the mismatch in thermal expansion coefficients can lead to wafer bowing and delamination [25]. Despite these drawbacks, the high-volume capability of this technique means that it is now widely utilised in the production of Si photonics-based PICs and OEICs, primarily for data centres [26].
 3. **Heteroepitaxial III-V integration** - Heteroepitaxial growth describes the direct deposition of III-V materials onto a silicon substrate through monolithic processes such as molecular beam epitaxy (MBE) or chemical vapour deposition (CVD). The aforementioned issues with growth on Si have thus far limited the integration of III-V devices for Si photonics based PICs and OEICs [27]. However it is widely appreciated that this method has the potential to address the shortcomings of the previous techniques providing these issues can be effectively managed [28].
 4. **Group-IV lasers** - Growth of direct gap, group-IV lasers would be the ideal solution as it retains the benefits of direct epitaxy, while improving on the III-V system due to reduced thermal and polarity mismatch with Si. Since these materials are commonly utilised in CMOS processes, light sources based on these constituents
-

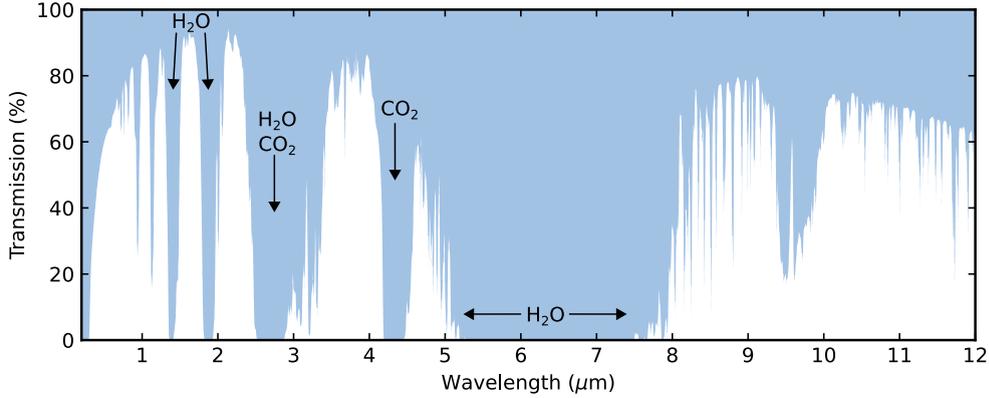


FIGURE 1.1: Atmospheric transmission spectra, indicating wavelength ranges where absorption from water vapour (H_2O) and carbon dioxide (CO_2) is high. Data obtained from [29].

would be easy to integrate into current production flows and likely minimise associated costs. Research on electrically injected group-IV lasers is still in its infancy, with operation limited to cryogenic temperatures.

1.3 Mid-Infrared Applications

Interest in the mid-IR spectral range fundamentally stems from the presence, and absence, of strong molecular absorption resulting in two primary applications: spectroscopic sensing and low-attenuation free space optical communications. In this work, focus is placed on the former, exploring how the lasers discussed in section 1.1 can be utilised in OEIC-based sensors to benefit a number of fields.

Molecules are able to absorb light because they have quantised energy levels corresponding to electronic, vibrational and rotational states. The fundamental vibrational transitions ($v=0$ to $v=1$) occur in the mid-IR and are characteristic of the specific chemical bond. Rotational transitions are much smaller in energy. As such, transitions are nominally roto-vibrational, meaning that both the vibrational and rotational states are altered. This manifests as a band of energies known as the fundamental band, where fine structure is introduced by way of the rotational excitations. Transitions to higher order vibrational states are known as overtones. These typically occur outside of the mid-IR and have orders of magnitude weaker absorption features since such transitions are forbidden in a perfect harmonic oscillator. Thus, the comparative intensity of absorption features in the mid-IR make it especially attractive for spectroscopic sensing applications.

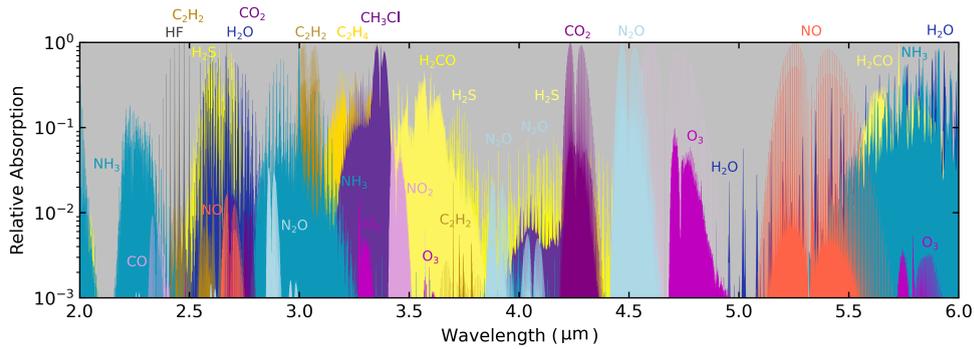


FIGURE 1.2: Absorption lines of important molecules for an array of environmental, medical and industrial applications in the mid-IR. The spectra of each molecule is normalised to it's own maxima. Data obtained using Radis [30].

Measurements of trace gases conducted in atmosphere can be greatly impacted by the presence of competing absorption features, which complicate quantitative analysis. However, the mid-IR has numerous spectral regions where absorption is relatively low for the most abundant atmospheric gases such as water vapour, as shown in Fig. 1.1. These transparency windows coincide with absorption features of numerous molecules important across a wide variety of sectors, from medical to environmental and defence, a selection of which are illustrated in Fig. 1.2.

Currently, Fourier transform IR (FTIR) spectrometers are the most commonly implemented set-ups for performing mid-IR absorption spectroscopy. A standard FTIR is comprised of a Michelson interferometer, broadband light source, sample compartment, detector and various optical components. As a result, a typical set-up required to measure analytes to the appropriate level of sensitivity will often occupy a lab bench, and weigh between 10 - 30 kg. While more portable FTIRs do exist, they often do not exhibit the levels of spectral resolution and sensitivity required for many applications. In addition to issues of manoeuvrability, there is also significant amounts of wasted energy. This stems from the use of a broadband light source, where quantitative analysis of a single molecular species requires only a monochromatic source. However, potentially the most detrimental drawback is the lack of in-situ monitoring which inhibits real-time data collection in highly dynamic situations and may degrade the integrity of the sample when removed from it's original environment.

As such, Si-based OEICs utilising mid-IR semiconductor lasers may provide a number of key benefits for absorption spectroscopy applications. The first of which arises from the ability to leverage highly mature CMOS processing, opening pathways to low cost

and large volume manufacturing of small, lightweight and reliable chips [31], [32]. This scheme would therefore enable cost-effective monitoring of a sample to take place in the original environment, thereby alleviating issues with storage and contamination, whilst simultaneously providing continuous, real-time data in adaptive situations.

While there is significant interest in broadband mid-IR technologies to enable label-free sensing, such as supercontinuum generation [33], diode-based sensors will likely focus on targeting a specific analyte using a single, prominent absorption feature. This is especially true of GeSn-based lasers which aren't expected to yield broadband emission or generate powers high enough to drive supercontinuum generation. Using this sensing scheme, the inherently narrow bandwidth ensures greater energy efficiency when compared to current FTIR-based methods, with less power wasted in generating light at non-essential wavelengths. This feature also provides greater specificity since the source can be tuned, via current or heater, to precisely target a specific molecular absorption line whilst avoiding closely spaced absorption features [34]. Examples of this implementation are explored in the ensuing sections.

Additionally, non-chip-based absorption spectroscopy techniques employing diode lasers are routinely able to detect gas concentrations at parts-per-million and frequently parts-per-billion levels [35], [36]. This highlights the potential for fully integrated OEICs to perform trace gas detection with high levels of sensitivity. The following sections explore how mid-IR absorption spectroscopy is utilised in different areas, and how they may benefit from Si-photonics-based OEICs.

1.3.1 Environmental, Industrial and Defence Applications

Environmental risks can encapsulate a wide array of hazards, including natural disasters, pollution, climate change, and industrial by-products, all of which pose threats to ecosystems, human health, and socio-economic stability. Given the rise in pollution levels, manufacturing processes and the demand for early disaster detection systems, real-time, on-site environmental monitoring is crucial for enabling those responsible to construct informed mitigation strategies. The benefits afforded by “lab-on-a-chip” (LOC) OEICs make them an interesting prospect for such use-cases.

For example, significant progress has been made towards OEICs for two of the primary gases contributing to global warming; carbon dioxide (CO₂) and methane (CH₄). In 2017,

the first silicon photonics chip to detect a greenhouse gas was reported by Tombez et al., who measured absorption using the $\lambda = 1.65 \mu\text{m}$ overtone transition of methane [37]. A transition to the mid-IR was made in 2023, when a Si slot waveguide scheme was utilised to detect methane using the fundamental $\lambda = 3.2704 \mu\text{m}$ line, with a sensitivity in excess of 0.3 parts-per-million [38]. Mid-IR Si-based PICs have also exhibited the ability to assess CO₂ levels using a simple waveguiding system to monitor absorption at $\lambda = 4.2 \mu\text{m}$ [39]. While less widely discussed, some volatile organic compounds (VOCs) also act as greenhouse gases and contribute to global warming effects. Acetylene (C₂H₂) is one such example, which has already been detected using a Si photonic chip at $\lambda = 2.566 \mu\text{m}$ with a sensitivity of 7 parts-per-million [40].

Mid-IR absorption spectroscopy has also found applications in defence, for stand-off detection of chemical, biological and explosive agents. LOC sensing devices may also find uses in this area, for example as wearables for first responders who may encounter hazardous chemicals when dealing with industrial accidents or terrorist attacks. It has been reported that the Department of Homeland Security Science and Technology Directorate in the US are actively pursuing this application. It would entail establishing a system using ML and AI to collate data from numerous sensors and provide the wearer with actionable steps [41]. Work on PIC based sensors for trace detection of hazardous gases has been reported in numerous studies. Typically a sorbent polymer is used to concentrate the specific analyte on the PIC waveguide, resulting in enhanced interaction with the evanescent optical field. A change in the detected field intensity is used to confirm the presence of the targeted substance. For example, Holmstrom et al. used this methodology to enable spectroscopic sensing of organophosphates, compounds present in nerve agents [42]. Tyndall et al. implemented a similar technique to detect trace chemical warfare agent simulants down to sensitivities of parts-per-billion [43].

1.3.2 Medical Applications

LOC devices are also projected to play an important role in the future of healthcare, due to their potential for low-cost, non-invasive, real-time, online monitoring of medically important analytes [44]. At present, the only low-cost, spectroscopic sensors available for medical applications are pulse oximeters, which measure blood-oxygen levels. However, these devices have proved essential for establishments such as the NHS, particularly over

the last half-decade where at home use enabled early detection of hypoxia in COVID-19 patients whilst reducing strain on hospitals [45]. Low-cost spectroscopic solutions for biosensors and glucose monitoring are currently in development [46], which are expected to yield similar improvements to healthcare and diagnostics. Such technologies are likely to benefit significantly from the development of OEICs based on both Si and InP platforms due to reduced manufacturing costs and device footprints.

While subcutaneous measurements of analytes are typically confined to the near-IR and visible spectral regions due to transparency windows for blood, melanin and water at $\sim 600 - 1200$ nm, the mid-IR has become an area of interest for analysis of breath samples. It has long been appreciated that analysis of breath could be used to diagnose a myriad of diseases. For example, in Ancient Greece, sweet scents were noted in exhaled breath from diabetics, and a fish-like odour was observed in individuals who we now understand to have had kidney disease [47]. The seminal 1971 work of Nobel prize winner Linus Pauling later tied such aromas to specific gaseous biomarkers, or profiles of VOCs [48].

There are now hundreds of known VOCs exhaled in breath, and copious studies detailing features of VOC profiles that can be linked to a vast array of diseases. For many VOCs, including those tentatively linked to different cancers [49]–[51], the biochemical process behind their production is not yet understood. As such, they are unable to be used as biomarkers. However, for a small number of VOCs, the mechanism leading to their formation is well-known, and therefore their concentration can be used to indicate the presence of disease. This has resulted in a handful of FDA approved breath tests, such as the “ ^{13}C urea breath test” which is used to determine the presence of the peptic ulcer and gastric cancer causing *h. pylori* bacteria in the stomach.

Moving forward, it is believed that the number of confirmed biomarkers will grow, along with the ability to detect key signatures of disease from VOC profiles. It has therefore been postulated that breath tests may be able to provide significant improvements in healthcare due to a number of key features, including:

1. **Readily available samples** - Unlike blood or urine samples, breath samples can be readily obtained in a non-invasive manner and without compromising patient comfort. As such, testing can be conducted frequently to monitor the effects of intervention in real-time. This may also improve patient compliance with screening programs, for example.
-

Pathology	Biomarker	Wavelength (μm)
Diabetes	$\text{C}_3\text{H}_6\text{O}^\dagger$ (Acetone) [52]	8.2 [53]
Asthma	NO^\dagger (Nitric Oxide) [54]	5.26 [55]
Neonatal jaundice	CO^\dagger (Carbon Monoxide) [56]	4.26 [57]
Breast and lung cancer	CH_2O (Formaldehyde) [58], [59]	3.59 [60]
Oxidative Stress	C_2H_6 (Ethane), C_5H_{12} (Pentane) [61]	3.4 [62]
Hepatic disorders	COS (Carbonyl Sulphide) [63]	4.8 [64]

TABLE 1.1

- 2. Monitoring of dynamic conditions** - By integrating sensors into ventilator systems or anaesthesia equipment, it may be possible to monitor conditions with rapid onset and progression in surgery or ICU settings. For example, during major surgery, extensive physiological trauma leads to increased reactive oxygen and nitrogen species which can overwhelm the body's antioxidant defences and lead to oxidative stress. As illustrated in table 1.1, alkanes are a potential biomarker of oxidative stress. Therefore, continuous monitoring during surgery may enable early detection before the onset of other symptoms, allowing interventions to be made promptly and limiting irreversible damage.
- 3. Reduced reliance on trained staff** - With the rapid growth of ML and AI, it is possible to envisage handheld tests based on OEICs that are able to provide instant feedback and the associated steps to be taken. This would reduce the burden on stretched healthcare services.
- 4. Multi-pathology detection** - There are a large number of diseases with potential exhaled breath biomarkers reported in literature, a selection of which are illustrated in table 1.1. Those marked with (\dagger) already have FDA approved breath tests, others have numerous studies linking the specific biomarkers to the pathology. As such, it may be that this type of test can be utilised for numerous conditions.

1.4 Thesis Outline

The aim of this thesis is to characterise novel mid-IR lasers on Si substrates, focusing primarily on carrier recombination processes that degrade performance. This can be used to inform future heterostructure designs, thereby aiding the production of more efficient and defect tolerant lasers for the applications mentioned above. The thesis is arranged as follows:

Chapter 1 - Introduction - Provides motivation for the studies conducted during this project. Semiconductor materials that are able to cover the mid-IR spectral region are introduced, before discussing routes for their integration into Si photonics platforms. Key applications of these technologies are subsequently explored.

Chapter 2 - Fundamental Theory - Presents the basic theory underpinning semiconductor lasers and different carrier recombination pathways.

Chapter 3 - Experimental Techniques - Outlines the experimental set-ups and operational conditions used when conducting temperature and pressure dependent characterisation of laser diodes.

Chapter 4 - Bulk Group-IV Bandstructure Modelling - Establishes the bulk bandstructure models and parameters used for modelling group-IV materials. To provide background, the history of GeSn laser devices is initially explored, before detailing the theory used to calculate bulk bandstructure in this work. A thorough literature review of parameters for elemental and multinary group-IV alloys is conducted, and used to evaluate the accuracy of the developed models.

Chapter 5 - Performance Limiting Mechanisms in Electrically Injected GeSn Lasers - Utilises the modelling tools detailed in the previous chapter to analyse first generation, bulk GeSn lasers. Device performance is probed using high-hydrostatic pressure at low temperature characterisation, leading to an extensive theoretical study detailing the effect of indirect valley proximity on the performance of bulk lasers at higher temperatures.

Chapter 6 - Recombination Processes in Epitaxially Grown GaInAsSb QW Lasers on Si - Is a comparative study between GaSb-based lasers grown on native substrates and on Si, operating at similar wavelengths. The contributing recombination mechanisms in both sets of devices are explored using temperature and pressure-dependent measurements of threshold before modelling is used to propose improvements to the heterostructure design.

Chapter 7 - Conclusions & Future Work - Summarises the findings of chapters 4 - 6 and presents options for future collaborative work.

Chapter 2

Fundamental Theory

This chapter provides a brief overview of the fundamental principles in semiconductor laser physics, essential for contextualising the results and analyses presented in this work. Section 2.1 describes the concepts of electronic bandstructure, how carriers distribute themselves among these states and how they recombine via optical absorption and emission. In section 2.2, the mechanisms for achieving and maintaining gain are discussed. The key design features of the laser double-heterostructure are detailed in section 2.3, showing how carrier and optical confinement are provided by different layers and how strain can be leveraged to enhance performance. Section 2.4 outlines the predominant carrier recombination mechanisms that act to increase the threshold carrier density. Primary focus is placed on defect-related, radiative, Auger-Meitner and leakage processes. Finally, inter-valence band absorption and free carrier absorption are conceptually outlined in section 2.5.

Further detail on the topics discussed in this chapter can be found in a number of excellent textbooks [65]–[68].

2.1 Electron Bandstructure and Optical Processes in Semiconductors

The formation of electronic bandstructure can be understood by considering the motion of a free-electron in a periodic potential provided by the crystal lattice. In this model, the wavefunction of the electron is approximated as a plane wave that is modulated by a

periodic function using Bloch's theorem [66], [69],

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad (2.1)$$

where \mathbf{r} denotes the electron position, \mathbf{k} is the wavevector and $u_{n\mathbf{k}}(\mathbf{r})$ is a Bloch function that reflects the symmetries of the crystal lattice. Solving the Schrödinger equation leads to a dispersion relation $E(\mathbf{k})$ with multiple solutions for any given wavevector, \mathbf{k} . These form distinct bands of energies separated by a bandgap, E_g , constituting an energetic region in which no states can exist. The band situated immediately below the bandgap is called the valence band (VB). Semiconductors are characterised by a full VB at 0 K, however their small-to-moderate bandgaps (< 2 eV) allow for electrons to be thermally excited to the band immediately above the bandgap, known as the conduction band (CB). The promoted electrons leave behind a positively charged quasi-particle known as a hole. Both VB holes and CB electrons are able to carry a net current and are collectively referred to as charge carriers.

Since the crystal potential is anisotropic, the obtained dispersion relation will vary depending on the k -space direction. Typically, CB minima will occur at high-symmetry points located at the edge of the Brillouin zone, and at the zone centre. For the zinc-blende structures of interest in this body of work, band extrema are generally found at the Γ -, L- and X-points, located at $(0,0,0)$, $(\frac{\pi}{a}, \frac{\pi}{a}, \frac{\pi}{a})$, and $(\frac{\pi}{a}, 0,0)$, respectively. In semiconductor lasers, the band extrema are such that charge carriers will predominantly occupy states close to the Γ -point since this enables efficient radiative emission. A schematic example of a typical CB for a zinc-blende crystal is illustrated at the top of Fig. 2.1 in blue.

Band formation can also be explained by considering the hybridisation of atomic orbitals when atoms are brought close together to form a crystal. Overlapping of nearby orbitals results in the splitting of discrete atomic energy levels to form the continuous bands of states. In the tetrahedrally bonded semiconductors of interest in this study, sp^3 hybridisation is exhibited [70]. For these materials, the CB states at the Γ -point are formed from the overlap of anti-bonding s-orbitals. Since s-orbitals are not degenerate, the CB close to Γ consists of a single band. In contrast, the VB in the vicinity of Γ is a linear combination of the triply degenerate p-orbitals. As such, three VBs are observed at the zone centre, denoted the heavy-hole (HH), light-hole (LH) and spin-orbit (SO) bands. The SO band is named as such since the relativistic spin-orbit interaction lifts the degeneracy of the band, separating it from the degenerate LH and HH bands by the spin-orbit splitting

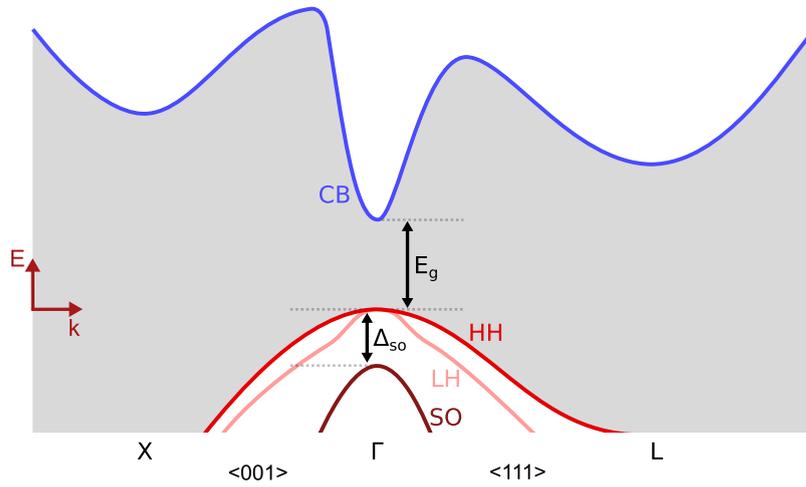


FIGURE 2.1: Schematic depiction of the band dispersion of a zinc-blende crystal lattice along the $\langle 001 \rangle$ and $\langle 111 \rangle$ directions. The CB is highlighted in blue, while the VB is shown in shades of red.

energy, Δ_{so} . The structure of the VB is depicted at the bottom of Fig. 2.1 in shades of red. The CB close to the X and L points have mixed s- and p- like character, stemming from coupling between an excited s state and the p-orbitals [71].

For small \mathbf{k} , the conduction and valence bands are often well approximated by a parabola,

$$E = \frac{\hbar^2 \mathbf{k}^2}{2m^*}, \quad (2.2)$$

where m^* is the effective carrier mass which determines the curvature of the bands. In general, the effective mass is dependent on the \mathbf{k} -direction, and as such, it takes the form of a tensor. However, if the bands are assumed to be isotropic, the effective mass can take a single scalar value.

2.1.1 Carrier Distribution

The number of states per unit volume at a given energy is given by the density of states (DoS), which takes the general form

$$\rho(E) = \frac{1}{V} \int_{\mathbb{R}^d} \frac{d^d \mathbf{k}}{(2\pi)^d} \delta(E - E(\mathbf{k})), \quad (2.3)$$

where E is the energy, V is the volume, and d denotes the spatial dimensionality of the system. For bulk semiconductors $d = 3$, whilst quantum confined materials will take values

between 0-2 depending on the number of confined dimensions. Under the assumption of isotropic, parabolic bands, the DoS for a given band in a bulk semiconductor is given by

$$\rho_{3D}(E) = \frac{1}{2\pi^2} \left(\frac{2m_{dos}^*}{\hbar^2} \right)^{3/2} \sqrt{E} \quad (2.4)$$

where m_{dos}^* is the DoS effective mass, and E denotes the energy above (below) the conduction (valence) band extrema.

With knowledge of the DoS, it is important to understand how charge carriers are distributed between these states. Both electrons and holes possess half-odd-integer spin and are therefore classified as fermions. Such particles obey the Pauli-exclusion principle, which states that no two indistinguishable particles may occupy the same quantum state simultaneously. As such, the energetic distribution of carriers is described by Fermi-Dirac statistics. In equilibrium, the distribution of electrons is given by the Fermi function

$$f(E) = \frac{1}{1 + \exp\left(\frac{E-E_f}{k_B T}\right)}, \quad (2.5)$$

where k_B is the Boltzmann constant, T is the temperature and E_f is the Fermi level, which describes the energy at which the probability of occupation is 50%. This manifests as a perfect step function at 0 K, which thermally broadens with increasing temperature. As a hole is the absence of an electron, the probability of hole occupation is therefore given by $1 - f(E)$.

Since the number of states in an energy range $(E, E + dE)$ is given by $\rho(E)dE$, and $f(E)$ gives the probability of an electron occupying these states, the total number of electrons within the energy range is $f(E)\rho(E)dE$. The density of electrons in the CB can therefore be calculated by integrating over the band as,

$$n_{3D} = \int_{E_{cb}}^{\infty} f(E) \rho_{cb}(E) dE. \quad (2.6)$$

Similarly, the density of holes in the VB is given by

$$p_{3D} = \int_{-\infty}^{E_{vb}} [1 - f(E)] \rho_{vb}(E) dE. \quad (2.7)$$

In the parabolic band model, utilising Eqn. 2.4 results in simplified expressions of

$$n_{3D} = N_c \mathcal{F}_{1/2} \left(\frac{E_f - E_{cb}}{k_B T} \right), \quad (2.8)$$

and

$$p_{3D} = N_v \mathcal{F}_{1/2} \left(\frac{E_{vb} - E_f}{k_B T} \right), \quad (2.9)$$

for the electron and hole densities, respectively. Here, $N_{c,v} = 2 \left(\frac{k_B T m_{dos}^*}{2\pi\hbar^2} \right)^{3/2}$ is the effective DoS for the CB and VB [65]. Within the framework of the Boltzmann approximation, the Fermi-Dirac integral of order 1/2 can be approximated as $\exp(x)$. This is only valid providing E_f is energetically distant from the band of interest, i.e. for low levels of doping or injection [72]. Otherwise, the full Fermi-Dirac integral must be conducted. This is defined as

$$\mathcal{F}_{1/2}(x) = \int_0^\infty \frac{\varepsilon^{1/2}}{1 + \exp(\varepsilon - x)} d\varepsilon, \quad (2.10)$$

which must be evaluated numerically.

2.1.2 Optical Processes in Semiconductors

There are three distinct mechanisms by which optical interband transitions can occur in semiconductor materials. These are absorption, spontaneous emission and stimulated emission, which are illustrated schematically in Fig. 2.2.

In an interband absorption process, a photon with energy greater than the bandgap is used to promote an electron from the VB to an unoccupied state in the CB, leaving behind a hole. The rate of absorption is therefore dependent on the density of unoccupied CB states, of electrons in the VB and of photons with energy greater than the bandgap. A semiconductor material will be transparent to photons with energy less than the bandgap due to energy conservation laws.

Interband spontaneous emission occurs when an electron-hole pair recombine. The produced photon has an energy equal to the electron and hole state splitting, and is emitted in a random direction, with a random phase and polarisation. Since the transition is not mediated by a photon, the rate of spontaneous emission is dependent only on the density of electrons in the CB and holes in the VB.

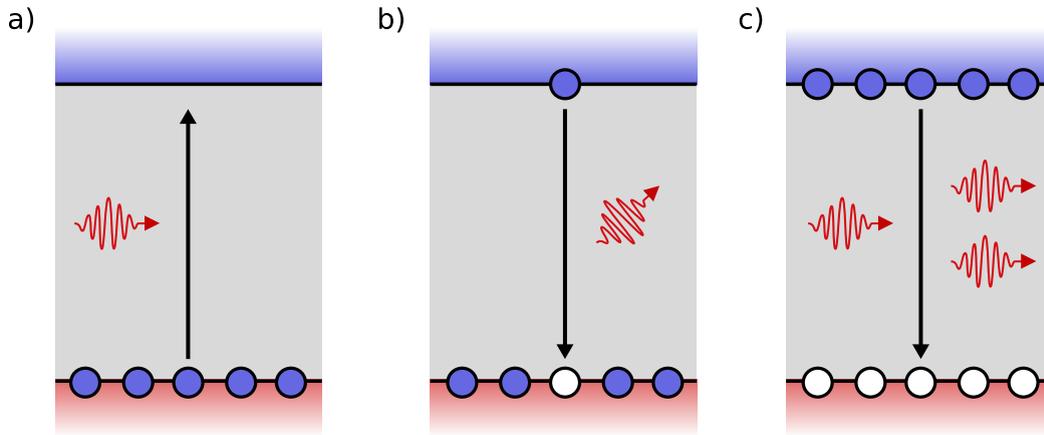


FIGURE 2.2: Illustration of a) absorption, b) spontaneous emission and c) stimulated emission processes. Blue and white markers represent electrons and holes, respectively. Red wavepackets denote photons.

The stimulated emission process requires a photon with energy greater than the bandgap to perturb an electron in the CB and induce recombination with a VB hole. The emitted photon again has an energy equal to the difference between the electron and hole energies. However, in contrast to spontaneous emission, the emitted photon retains the phase, polarisation and momentum of the incident photon. This amplification of light is the basis of lasing. The rate of stimulated emission is dependent of the density of photons with energies greater than the bandgap, and the densities of electrons in the CB and holes in the VB.

2.2 Fundamentals of Semiconductor Lasers

Lasing occurs when light is coherently amplified by stimulated emission during a round-trip within a cavity. A semiconductor laser is therefore underpinned by three primary, fundamental requirements

1. A material that is able to produce stimulated emission and contribute to gain
 2. A mechanism for injecting carriers into the gain medium to attain population inversion
 3. A cavity that is able to produce optical feedback, which describes the recycling of stimulated emission within the gain medium
-

The roles of each element within the lasing process are explored below.

2.2.1 Gain

A net positive value of gain may only be achieved if the rate of stimulated emission is greater than the rate of absorption. For this to be attained, the carrier population must be ‘inverted’, such that a greater number of electrons are occupying CB rather than VB states. As such, the system must be driven away from equilibrium by means of optical pumping or electrical injection of carriers. In this body of work the primary concern is the latter implementation. In electrically injected devices, a p-n junction is utilised such that holes from the p-type region and electrons from the n-type region are swept into the intrinsically doped gain material when a bias is applied.

Under non-equilibrium conditions, the Fermi level given by Eqn. 2.5 is no longer sufficient for describing the carrier distribution. However, since thermalisation of carriers occurs on timescales (~ 10 ps) much shorter than that of interband transitions (~ 1 ns), the carriers reach a local equilibrium with their respective bands. Fermi-Dirac statistics may therefore be applied, but with distinct quasi-Fermi levels describing the electron distributions in the conduction and VBs. These are given by

$$f_c(E_{cb}) = \frac{1}{1 + \exp\left(\frac{E_{cb} - E_{fc}}{k_B T}\right)}, \quad f_v(E_{vb}) = \frac{1}{1 + \exp\left(\frac{E_{vb} - E_{fv}}{k_B T}\right)}, \quad (2.11)$$

respectively. Here, E_{fc} and E_{fv} are the quasi-Fermi energies of the CB and VB, and the hole distribution in the VB can be determined as $f_h(E_{vb}) = 1 - f_v(E_{vb})$.

The condition for inversion, and therefore net positive gain, is then $f_c(E_{cb}) > f_v(E_{vb})$. By rearranging Eqns. 2.11, and noting that gain may only be achieved for photon energies, $\hbar\omega$, exceeding the bandgap, the Bernard-Duraffourg condition for gain is obtained as [73]

$$\Delta E_f > \hbar\omega > E_g, \quad (2.12)$$

where $\Delta E_f = E_{fc} - E_{fv}$ is the quasi-Fermi splitting energy and E_g is the bandgap energy. Fig. 2.3 a) depicts this condition schematically.

Fig 2.3 b) depicts the gain spectra as a function of increasing injection illustrating the movement of the peak gain as the carrier density is raised. Photon energies marginally

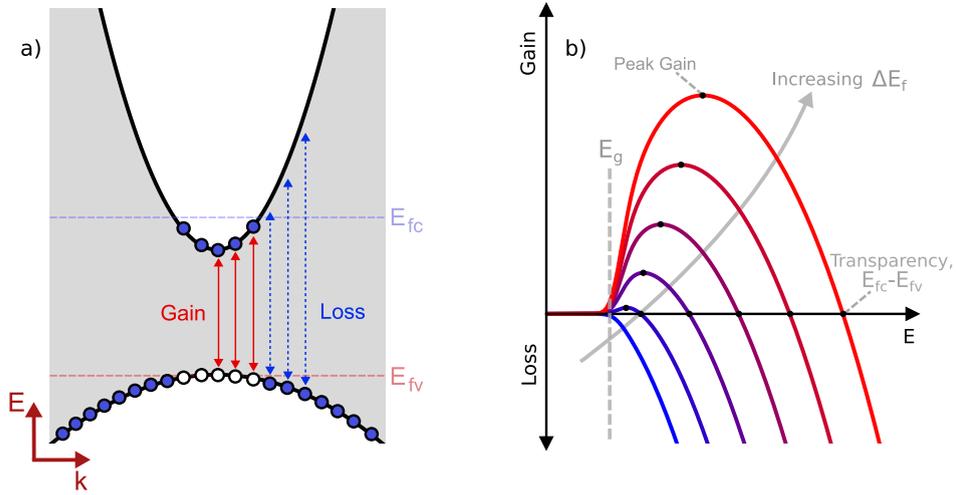


FIGURE 2.3: a) Bandstructure diagram schematically illustrating the Bernard-Duraffourg condition, where only transitions with energies between ΔE_f and E_g produce gain. b) Typical gain spectral as a function of increasing carrier quasi-Fermi splitting (blue to red).

below the bandgap energy are able to contribute to gain as a result of both homogeneous and inhomogeneous broadening which act to smooth the gain spectra. The transparency point is the photon energy at which there is net zero gain/absorption for a given injection, which occurs at ΔE_f . The material is therefore effectively transparent to photons of this energy since there is no change in field intensity.

From the rate equations described by Fermi's golden rule, the energy dependent material gain can be derived as [67]

$$g(\hbar\omega) = \frac{\pi q^2}{n\varepsilon_0 m_0^2 c \omega} \rho(\hbar\omega) |M|^2 [f_c(\hbar\omega) - f_v(\hbar\omega)] \quad (2.13)$$

where q is the electron charge, n is the refractive index of the gain media, ε_0 is the permittivity of free space, m_0 is the electron rest mass, c is the speed of light in a vacuum, ρ is the photon density and $|M|^2$ is the interband transition matrix element.

2.2.2 Optical Feedback

For lasing to be sustained, photons must be recycled within a cavity to induce further stimulated emission. For the ridge lasers of interest in this work, optical feedback is provided by a Fabry-Pérot (FP) cavity, schematically depicted in Fig. 2.4. Here, photons are reflected at either end of the device by cleaving along a crystallographic plane parallel

to the ridge width. The reflectivity is governed by the contrast in refractive indices at the semiconductor-air interface and can be expressed as

$$R = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|^2. \quad (2.14)$$

For devices in this body of work (where $n_1 \sim 3.5$), the reflectivity of an uncoated facet typically take values around 30%. Since emission is often only required from a single facet, coatings can be used to independently increase or decrease the facet reflectivities accordingly.

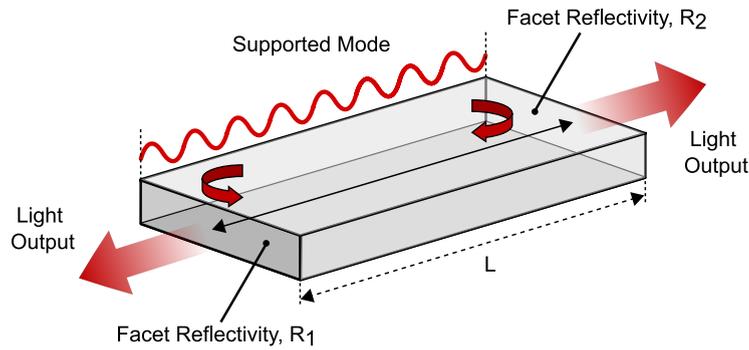


FIGURE 2.4: Schematic illustration of a FP cavity.

The cavity will form longitudinal standing waves for light satisfying

$$m \frac{\lambda}{2} = n_1 L, \quad (2.15)$$

resulting in a mode spacing of [74]

$$\Delta\lambda = \frac{\lambda^2}{2n_1 L}. \quad (2.16)$$

Thus for very short cavities ($\sim \mu\text{m}$), the modes transmitted by the cavity have an appreciable energetic separation as shown by the blue resonances in Fig. 2.5. The resultant modal gain spectra, is therefore the convolution of the calculated gain spectra and the cavity modes. This property can be important for devices with short cavities, such as vertical cavity surface emitting lasers (VCSELs) and photonic crystal surface emitting lasers (PCSELs), since the mode closest to the gain peak will undergo the largest amplification, allowing for devices with increased mode selectivity. For the FP-based edge-emitting lasers discussed in this thesis, cavity lengths are nominally in excess of 1 mm, resulting in

mode spacings of < 0.2 meV. Thus, the resultant gain spectra is not significantly modified by the cavity mode spacing, as illustrated by the red line in Fig. 2.5.

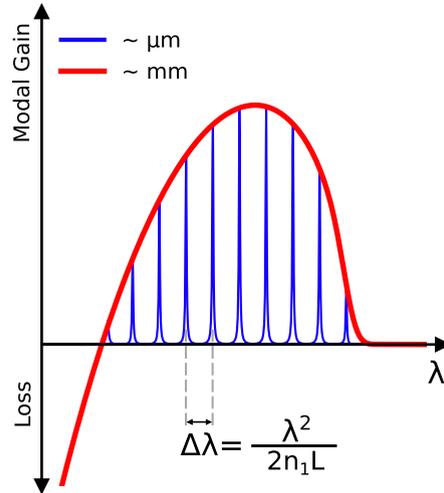


FIGURE 2.5: Schematic depiction of typical gain spectra for devices with micron (blue) and millimetre (red) cavity lengths. For longer cavities the modal spacing is small, resulting in a smooth gain spectra.

2.2.3 The Threshold Gain Condition

In Eqn. 2.13, it is assumed that photons are only removed from the cavity by means of interband absorption. In reality there are numerous processes by which this may happen, including intervalence band absorption (IVBA), free carrier absorption (FCA), and scattering, which are discussed further in section 2.5. These are difficult to quantify individually via experimental methods, and are therefore normally grouped as ‘internal losses’, denoted by α_i . Additionally, as shown in the subsequent section, not all of the optical mode will overlap with the gain producing layers. This means that only a fraction of photons are available to contribute to gain. As such the material gain is weighted by Γ , which describes the fraction of the modal energy that is confined to the gain producing region, giving the so-called modal gain, $G = \Gamma g$.

For lasing to occur, the modal gain must be equal to the total incurred loss in the system. This is referred to as the threshold condition, above which amplification becomes self-sustaining. The condition is derived by considering the change in field intensity incurred during a round-trip and expressed mathematically as [75]

$$R_1 R_2 e^{2L(\Gamma g_{th} - \alpha_i)} = 1, \quad (2.17)$$

such that the material gain required to reach threshold is given by

$$g_{th} = \frac{1}{\Gamma} \left[\frac{1}{2L} \ln \left(\frac{1}{R_1 R_2} \right) + \alpha_i \right]. \quad (2.18)$$

2.3 Semiconductor Laser Heterostructures

Acknowledging the key requirements for lasing, noted above, the following section discusses how the modern FP semiconductor laser heterostructure is designed to optimise performance.

2.3.1 Carrier Confinement

The most fundamental basis of a semiconductor laser is the p-n junction, created by doping materials with impurity atoms that have either fewer valence electrons (p-type) or greater valence electrons (n-type). For n-type material, the excess valence electrons (not used in bonding) are promoted from the defect state into the CB since the donor depth is typically $< k_B T$. n-type dopants are therefore referred to as ‘donors’, since they donate electrons that can be used as charge carriers. Similarly, for p-type materials, excess electrons in the VB are promoted to fill the vacant defect states, thereby increasing the hole density in the VB. p-type dopants are therefore known as ‘acceptors’, since they accept electrons

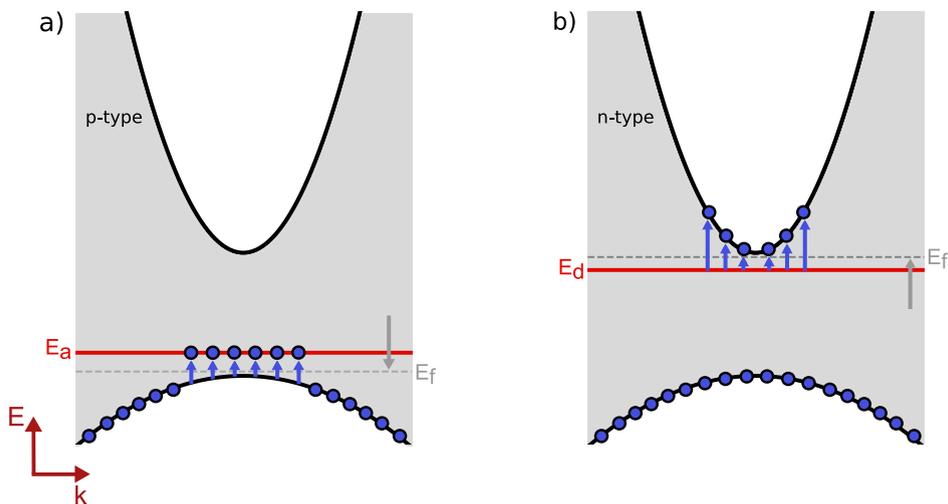


FIGURE 2.6: Diagram illustrating the effect of incorporating a) p-type and b) n-type dopants into a zinc-blende crystal lattice. Electrons are depicted by blue markers, and the acceptor and donor levels are given by solid red lines. Dotted grey lines show which direction the Fermi level would move as a result of the doping.

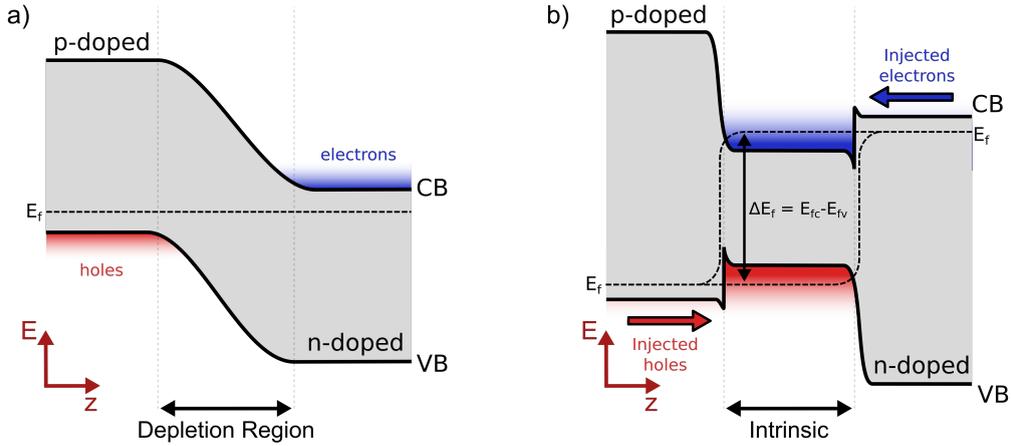


FIGURE 2.7: band edge diagrams for a) a p-n junction and b) a p-i-n junction under bias. Blue and red gradients are used to depict the spatial concentration of electrons and holes, respectively.

from the host material. The doping induced alteration to the carrier distribution results in a downward shift of the Fermi-level for p-type material and an upward shift for n-type material.

When p- and n- doped material are placed in contact, free electrons from the n-region diffuse across the interface and recombine with holes in the p-region, leaving behind negative impurity ions. Similarly, holes diffuse into the n-region, resulting in a layer of positive ions. The electric field produced between oppositely charged ions inhibits further charge flow across the so-called depletion region, thereby reaching equilibrium and a continuous Fermi level is found throughout the structure. This situation is indicated in Fig. 2.7. When a forward bias is applied, the Coulomb barrier between the p- and n-regions is reduced and, above a certain turn-on voltage, charge carriers are able to flow into the depletion region where they can recombine. This is the basic premise of both LEDs and lasers.

Improvements can be made by placing intrinsically doped gain media between p- and n-type material, forming a p-i-n structure. The increased free charge carrier density in the doped layers means that carriers can be swept into the active region by applying a bias to deposited contact layers. The choice of material for the doped layers stems from the need to confine the carriers to the active region. By choosing wider bandgap materials than the intrinsic layer, electrons and holes encounter large potential barriers at the p-type and n-type layer interfaces, respectively. This prevents carrier escape and therefore greatly

enhances the probability of radiative recombination. Fig. 2.7 b) schematically depicts a p-i-n structure, biased to levels where lasing can be achieved.

2.3.2 Quantum Confinement

The active region discussed in the previous section can be improved upon by utilising quantum confined structures which provide advantages for laser performance stemming from enhanced strain incorporation, emission wavelength tuning, and preferable DoS. In this study, quantum confined devices are in the form of quantum wells (QWs). QWs are formed when the layer thickness of the smaller bandgap material is reduced to scales comparable with the de-Broglie wavelength of the charge carriers. Carriers are then confined by a potential barrier at the layer interfaces. For motion solely in the growth-direction, z , the allowed energy states are quantised. Recalling the solutions for a particle in an infinite, 1D square potential well from basic quantum mechanics, the energies are given by [76]

$$E_z = \frac{n_z^2 \pi^2 \hbar^2}{2m^* L_z^2} \quad (2.19)$$

where n_z is the energy level index, m^* is the effective mass of the particle and L_z is the layer thickness. As such, the thinner the layer, the larger the energetic shift of the state away from the bulk band edge. In reality, the potential barrier is finite which acts to reduce the confinement energy and is calculated using numerical methods. The first

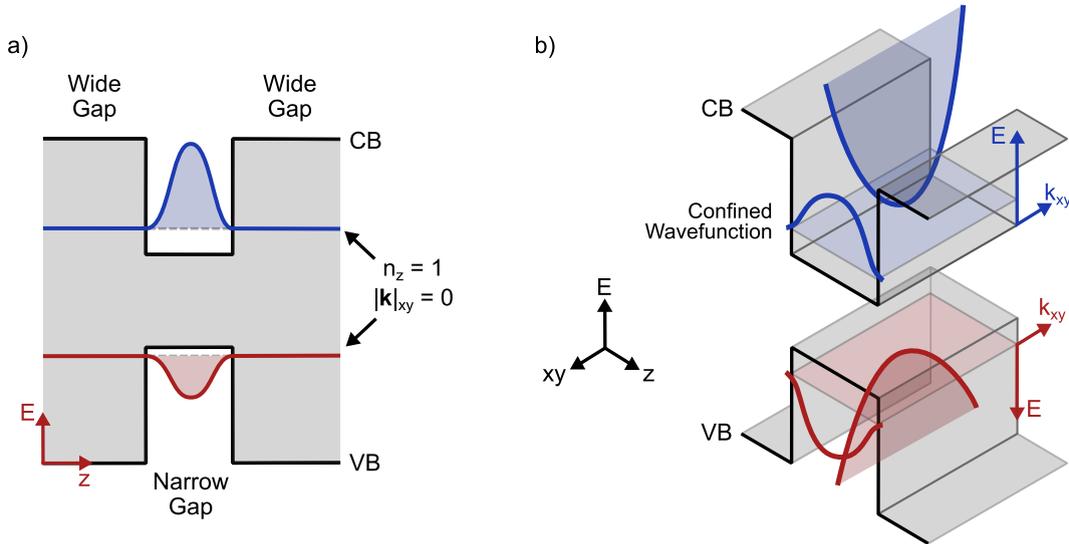


FIGURE 2.8: a) Illustration of the lowest CB and highest VB states in a quantum well, with their associated carrier wavefunctions. b) The parabola of sub-band states for the energy levels in a), arising from non-confined in-plane motion.

confined state for the CB and VB is illustrated for an arbitrary type-I QW structure in Fig. 2.8 a).

In contrast, the motion in the x, y plane remains unrestricted and hence, within the confines of the parabolic, isotropic model, the energy is given by

$$E_{x,y} = \frac{\hbar^2 |\mathbf{k}_{xy}|^2}{2m^*}. \quad (2.20)$$

The allowed energies are then

$$E = E_z + E_{x,y} = E_z + \frac{\hbar^2 |\mathbf{k}_{xy}|^2}{2m^*}, \quad (2.21)$$

hence the lowest possible energy is E_z when $n_z = 1$, which occurs for $k_{x,y} = 0$. For the same value of n_z , higher energy states are then realised for non-zero wavevectors $|\mathbf{k}_{xy}|$. The result is a parabola raised by an energy E_z from the bulk band edge, this is depicted in Fig. 2.8 b). For wider wells, there may be multiple confined states, i , each raised by an energy $E_z(n_z = i)$, and each with an associated parabola of extended states. The states associated with a given n_z are therefore said to belong to a ‘sub-band’ [67].

As discussed in section 2.1.1, the bulk DoS is proportional to $E^{1/2}$. Therefore in bulk materials, the DoS at the band edge is low. As such, the CB quasi-Fermi level must penetrate deep into the band, where the number of available electron states is higher, in

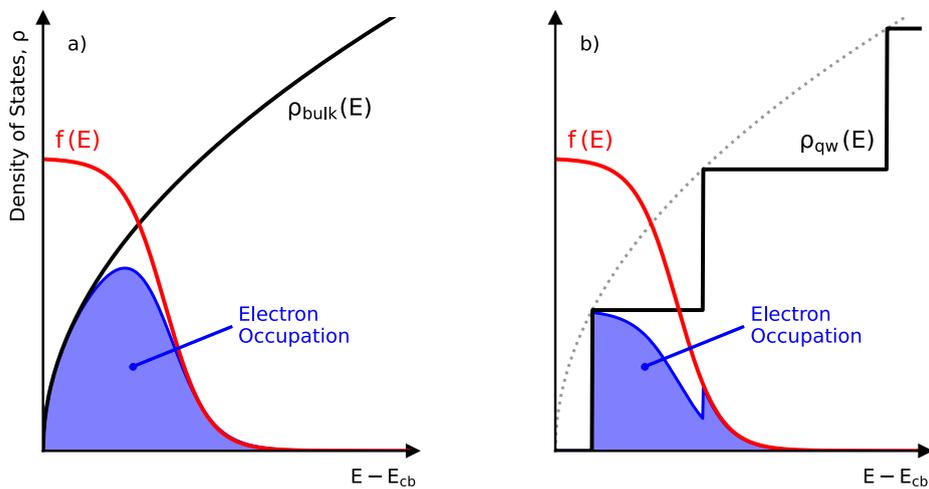


FIGURE 2.9: Schematic comparison between the density of states (black lines) in a) bulk material and b) a QW. In both cases, the solid red line is the Fermi distribution for a given penetration into the bulk band edge, while the blue shaded regions represent the resulting electron distributions. Adapted and reproduced from [77].

order to satisfy the Bernard-Duraffourg condition and reach transparency. This requires a large carrier density and results in highly broadened emission spectra.

For QW structures, it can be shown that the 2D DoS for a single sub-band is given by

$$\rho_{qw}(E) = \frac{m_{\parallel}^*}{\pi \hbar^2} \quad (2.22)$$

dependent only on the effective mass, m_{\parallel}^* . If there are multiple confined states, the total DoS at an energy, E , is the total sum of the DoS for all states below that energy, written as [68]

$$\rho_{qw}^{total}(E) = \rho_{qw}(E) \sum_{i=0}^n \Theta(E - E_i), \quad (2.23)$$

where Θ is the Heaviside step function. Assuming parabolic bands, this results in steps of equal height that are not evenly spaced in energy, as shown in 2.9 b). The 3D DoS can be obtained as $\rho_{qw,3D} = \rho_{qw}^{total}/L_z$ which converges to the bulk case for large thicknesses. In the quantum well, states are therefore available at the sub-band edge. Thus, for a given level of inversion, it can be seen in Fig. 2.9 that a greater proportion of carriers fill states at the sub-band edge and can therefore contribute to gain.

Since $E_z \propto 1/m^*$, the sub-bands for the LH and HH bands are no longer degenerate in a quantum confined structure. This degeneracy lifting further reduces the DoS in the vicinity of the VB edge. These two effects combine to result in heightened internal quantum efficiency, and lower transparency and threshold carrier densities when compared to bulk devices.

2.3.3 Optical Confinement

To minimise the threshold gain, it is essential to maximise the overlap of the optical field with the gain-producing region, as shown in Eqn. 2.18. Optical confinement in the vertical direction is achieved by utilising the properties of index guiding. Here, the field is confined via total internal reflection which occurs at the boundary between a high refractive index region, known as the core, and lower index surrounding layers. Fortunately, the refractive index of a material scales inversely with bandgap, meaning that both carrier and optical confinement can be achieved by sandwiching low bandgap material with higher bandgap material.

Typically, the core of the waveguiding region is designed to be $\sim \mu\text{m}$, such that it is similar to the emission wavelength. However, the QWs need to be ~ 10 nm or below to capitalise on the effects of quantum confinement. As such, the optical confinement factor for a single QW is low. To enhance this, multiple QWs are often utilised in the active region. These are surrounded by thick layers of higher bandgap material referred to as separate confinement heterostructure (SCH) layers, which serve a dual purpose, providing both carrier and optical confinement. These layers form the core of the waveguide, and are intrinsically doped to avoid large free carrier absorption losses. The refractive index contrast is provided between the SCH and larger bandgap layers known as cladding, which are heavily doped such that the aforementioned p-i-n junction is formed. The effect of this double heterostructure design on carrier and optical confinement is illustrated in Fig. 2.10 a) and b), respectively.

To reduce the number of lateral modes and reduce the overall device threshold current, the mode should also be confined laterally. In the FP lasers discussed in this study, lateral confinement is most commonly attained through the use of stripe contacts or by processing ridge waveguide structures. By narrowing the stripe width of the top contact, the lateral width in which there is sufficient carrier density to lase is reduced. The optical mode is then referred to as being gain-guided, reflecting that the largest gain is experienced in the central region where the carrier density is highest, rather than the tails. A secondary effect occurs as a result of minor current-induced changes to the refractive index, which also provides a weak guiding effect.

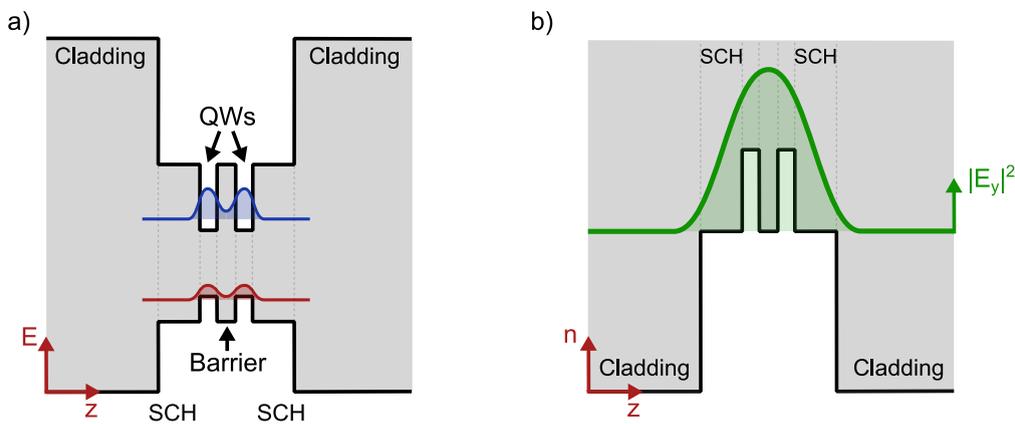


FIGURE 2.10: a) Illustration of electron (blue) and hole (red) confinement in a typical QW double heterostructure laser. b) Optical confinement (green) provided by the same structure as a result of high index contrast between the active region/SCH and the cladding.

Stronger lateral confinement is achieved when using a ridge waveguide structure. Here, a mask is used to define a narrow ridge, typically $\sim 10 \mu\text{m}$, before etching through the structure. In the laterally injected devices studied here, the material is etched through to the lower cladding layer. This injection scheme is often adopted in lasers grown on Si as it allows carriers to bypass the highly defective buffer layers. The ridge itself is then often coated in a low index material such as SiO_2 to provide electrical isolation and enhance the lateral mode confinement.

2.3.4 Strained Layers

Strain is introduced when a thin epilayer is grown on top of a thick, bulk-like material with a different lattice constant. The thin layer deforms and adopts the in-plane lattice constant of the underlying material. In doing so, elastic strain energy builds and beyond a certain critical thickness it becomes energetically preferable for dislocations to form, allowing the layer to relax back to its natural lattice constant. Such defects can act as carrier traps or scattering centres, therefore hindering the performance of bulk lasers. However, the typical dimensions of QWs allow moderate strains of $\sim 1 - 2\%$ to be realised without introducing dislocations. The bandstructure of semiconductors are closely tied to their lattice spacing, and as such, it was postulated that growing purposely strained

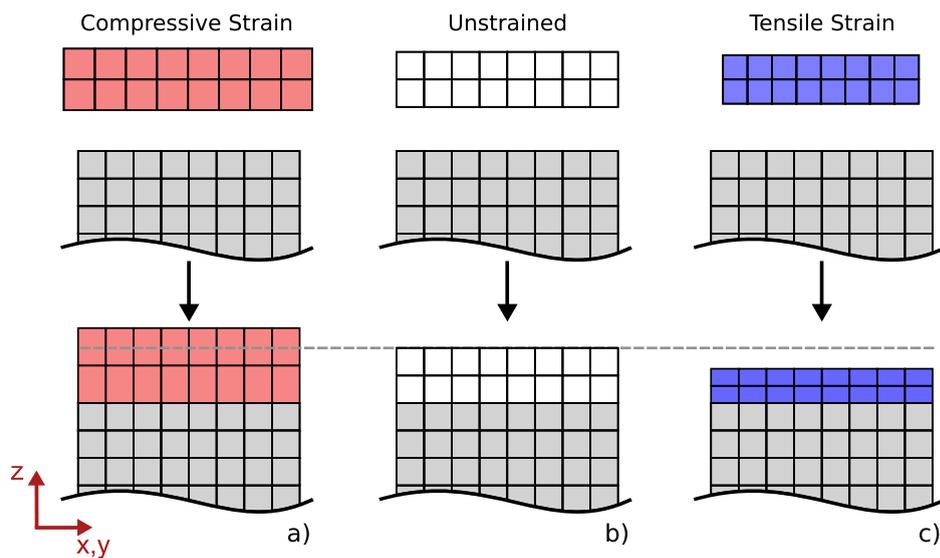


FIGURE 2.11: The strain induced in epilayers of material with a lattice constant that is a) larger than the substrate (compressive strain), b) lattice matched to the substrate (unstrained) and c) smaller than the substrate (tensile strain).

layers may produce changes that would be advantageous for QW-based laser performance [78], [79].

Epitaxial growth of semiconductor layers with different lattice constants results in biaxial strain. Here, the epilayer adopts the in-plane lattice constant of the bulk-like substrate layer, whilst the out-of-plane lattice constant deforms conversely to compensate this effect, resulting in tetragonal distortion. In the case of biaxial tensile strain, the natural lattice constant of the epilayer is less than that of the substrate. As such, it is stretched in-plane and compressed in the growth direction. In contrast, if the natural lattice constant of the epilayer is greater than that of the substrate, the material is compressed in-plane and stretched along the growth direction. This is known as a compressive strain.

The in-plane biaxial strain is calculated as

$$\varepsilon_{\parallel} = \frac{a_s - a_e}{a_s} \quad (2.24)$$

where a_e is the natural lattice constant of the epilayer and a_s is that of the substrate layer. It should be noted that the ordering in the numerator is arbitrary. Here, tensile strains take positive values and compressive strains are negative. This convention is used throughout the entirety of the thesis. The out-of-plane strain is related to the in-plane strain through the Poisson ratio, ν , as

$$\varepsilon_{\perp} = -2\nu\varepsilon_{\parallel} = -\frac{2C_{12}}{C_{11}}\varepsilon_{\parallel} \quad (2.25)$$

where C_{12} and C_{11} are the elastic constants.

A general stress can be separated into a hydrostatic and a shear component [82]. The hydrostatic strain results in a change in the lattice volume, whilst preserving the crystal symmetries. In the case of a compressive strain, the lattice constant is reduced in-plane resulting in a stronger interaction perturbation, thereby increasing the material bandgap. Conversely, a tensile strain decreases the bandgap.

A shear stress acts to break the cubic crystal symmetry, in this case through a uniaxial strain in the growth direction which produces tetragonal symmetry. This results in degeneracy lifting between the LH and HH bands at the zone centre. Under compressive strain, the HH band energy increases while the light-hole LH band energy decreases. In contrast, tensile strain elevates the LH band energy and lowers the HH band energy. In addition,

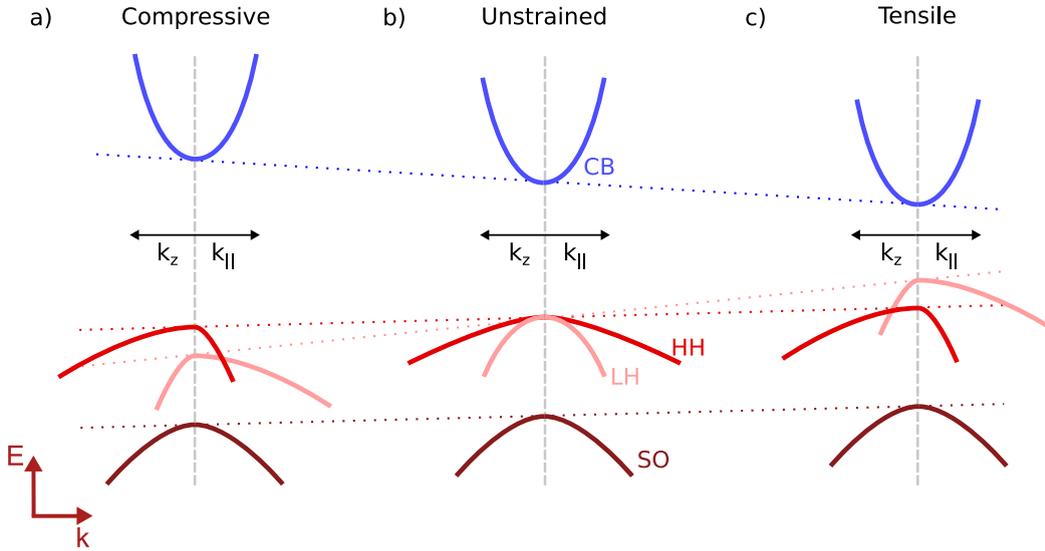


FIGURE 2.12: Illustration of the effect of biaxial strain on the bandstructure of a bulk zinc-blende semiconductor. b) shows the standard, unstrained bandstructure, whilst a) and c) highlight the effects of compressive and tensile strain respectively. Under compressive strain the bandgap is increased and the highest VB has a light in-plane mass. For tensile strained epilayers, the opposite is true. Adapted from [80], [81].

the shear stress introduces strong anisotropy to the VB dispersion such that the in-plane and out-of-plane effective masses are no longer equal. It is useful to note that the name assigned to the upper-most band in a strained semiconductor is a matter of preference. In this work the naming convention where bands are referred to by their growth direction character is adopted, since this is the component used when contemplating confinement effects. As such, for a compressive strain, the upper-most VB is termed the HH band, but exhibits a light in-plane mass. These effects are schematically illustrated in Fig. 2.12.

The latter effect is particularly important for improving device performance. Since the heavy hole effective mass is $\sim 5\times$ greater than the effective electron mass, the change in the VB quasi-Fermi level with injection will be far less than the CB quasi-Fermi level. Typically, at transparency, the VB quasi-Fermi level is situated just above the VB edge. Thus, large carrier densities are required to force the CB quasi-Fermi level far enough into the CB to reach the transparency condition, $E_{fc} - E_{fv} > E_g$. The introduction of compressive strain removes the VB degeneracy and decreases the in-plane effective mass of the highest band. As such, the DoS is significantly reduced, allowing the transparency condition to be met at lower carrier densities. In turn, this reduces the current associated with loss processes such as Auger-Meitner and radiative recombination that raise the overall threshold current of the device. This will be discussed in the ensuing section.

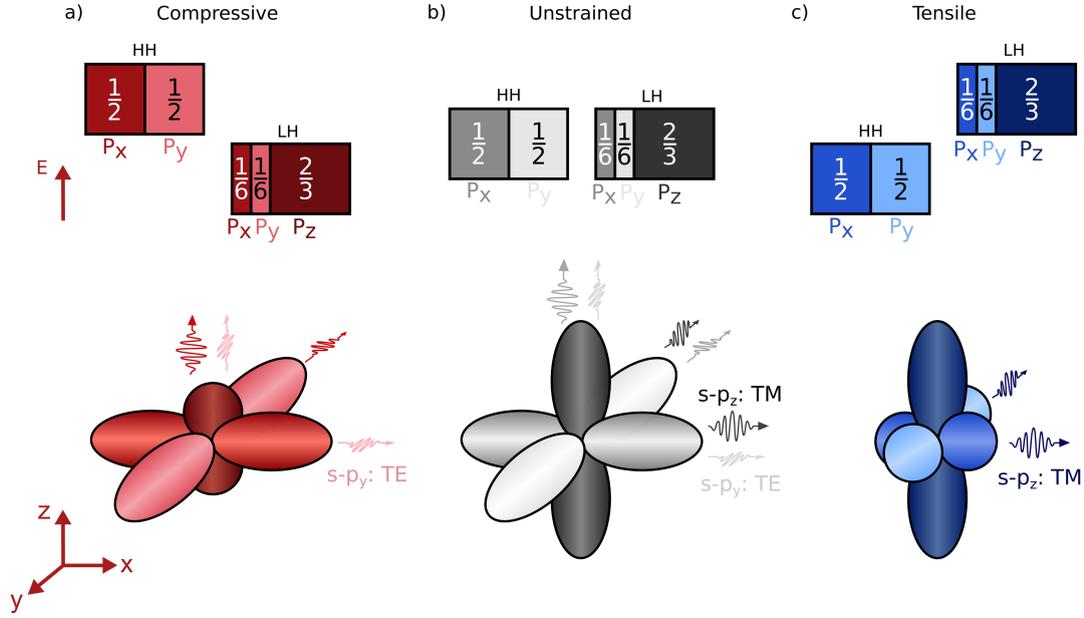


FIGURE 2.13: Depiction of the impact of biaxial strain on the distribution of holes amongst the p-orbitals. Larger orbital lobes constitute higher carrier occupation. In b) unstrained, bulk material, holes are distributed equally between the three p orbitals. In a) biaxially strained epilayers, holes are split equally between p_x and p_y , resulting in TE emission along the cavity direction, x . For c) tensile strained epilayers, holes are predominantly concentrated in p_z such that emission along the cavity is primarily TM.

The degeneracy lifting also acts to change the orbital characteristic of the states at the top of the VB, which proves beneficial for increasing gain. The HH band character is $1/2 p_x$ and $1/2 p_y$, whereas the LH band is $2/3 p_z$, $1/6 p_x$ and $1/6 p_y$ [83]. Radiative transitions between s and p-like states produce photons that are polarised along the direction of the p-like orbital. As such, a transition to a state with p_y symmetry will produce a y -polarised photon that may propagate in either x or z . Photons produced in this manner that travel along the cavity length in x are TE in nature and contribute to the modal gain. In contrast, transitions to a p_x -like state will result in x -polarised light that is unable to contribute to the lasing mode. Thus, in unstrained material, only $1/3$ of the holes available at threshold can produce useful gain.

In compressively strained materials, the HH band is raised to higher energies. As a result, a majority of the holes at threshold will occupy the p_x and p_y orbitals in equal measure [84]. The fraction of holes able to contribute to TE gain is therefore increased from $1/3$ in unstrained material to $1/2$ in compressively strained layers. Under tensile strain, the LH band is shifted upwards, and thus a majority of carriers occupy the p_z and p_y orbitals and can produce photons in the lasing mode. Since $2/3$ of holes are in the p_z lobe compared

to only 1/6 in p_y -states, the majority of emission is into the TM mode. These processes are outlined schematically for compressive, unstrained and tensile material in Fig. 2.13 a), b) and c), respectively.

2.4 Carrier Recombination Mechanisms

As discussed previously, a certain transparency carrier density is required to satisfy the Bernard-Duraffourg condition. This is purely dependent on the bandstructure of the device. To achieve lasing, the carrier density must be raised further above the transparency value such that the modal gain is greater than the total incurred optical losses. To maintain this carrier density, a certain current must be provided to compensate for the rate at which carriers are lost from the Γ -valley of the active region. The minimum value above which lasing becomes sustained is known as the threshold current. Assuming Boltzmann conditions and charge neutrality ($p = n$), the threshold current may be expressed using a simple power law known as the ABC model, given by [85]

$$I_{th} = eV(An + Bn^2 + Cn^3) + I_{leak}. \quad (2.26)$$

Here e is the electron charge, V is the active region volume, and A , B , and C are coefficients associated with the rate of defect-related, radiative and Auger-Meitner recombination, respectively. I_{leak} pertains to carriers that do not recombine within the active region, for example due to inefficient carrier injection or leakage of carriers into different layers of the heterostructure. The term also accounts for carriers leaking from the direct-valley to indirect valleys such as L or X, where they can no longer radiatively recombine and contribute to gain. Each mechanism is discussed more extensively in the following section.

To evaluate the performance of semiconductor lasers with varying heterostructure designs, the threshold current density, J_{th} , is utilised. This constitutes the required current per unit area and is written as

$$J_{th} = \frac{I_{th}}{w \times l}, \quad (2.27)$$

where w and l are the cavity width and length, respectively.

2.4.1 Defect-related Recombination

A defect pertains to an imperfection that disrupts the crystal or compositional symmetry of the lattice. Some common examples include interstitial atoms, vacancies and surface states. These defects introduce highly spatially localised states at energies within the bandgap that are consequently delocalised in k -space. As such, these ‘trap’ states are able to facilitate carrier recombination across the bandgap. This is known as Shockley-Read-Hall (SRH) recombination, in honour of the scientists who first reported on carrier recombination statistics for defect mediated transitions in 1952 [86], [87].

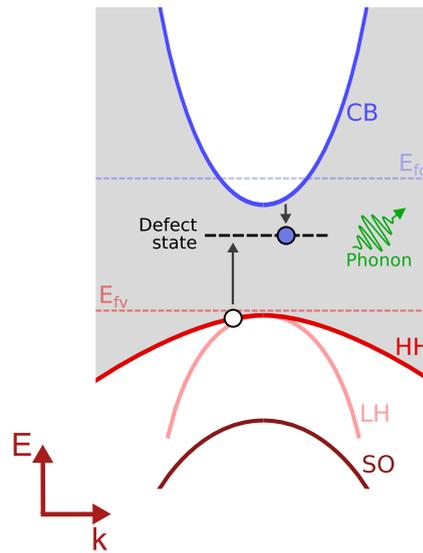


FIGURE 2.14: Illustration of carrier recombination via a trap level. Dashed blue and red lines illustrate the CB (blue) and VB (red) quasi-Fermi levels above transparency, showing that the defect state would be filled and subsequently recombine with a hole. Here, the energy is dissipated through a phonon (green), but photon emission is also viable.

If the state is located close to the CB edge, there is a high probability that it will be occupied by an electron. Typically, a hole will then rapidly recombine with the electron at the defect site, resulting in the emission of a phonon, as illustrated in Fig. 2.14, or low energy photon. However, it has been shown that trap sites can also assist other non-radiative processes such as Auger-Meitner recombination [88]. Since the mechanism nominally describes the recombination of a single carrier via a trap state, the associated current is described as monomolecular, and rises linearly with carrier density as

$$I_{def} = eVAn \quad (2.28)$$

2.4.2 Radiative Recombination

During radiative recombination, an electron close to the CB minima recombines with a hole near the VB maxima, as discussed in section 2.1.2. Spontaneous emission is initially required in order to provide the photons necessary to trigger stimulated emission events. However, spontaneous recombination produces photons with a broad range of energies that are emitted in random directions. Consequently, a majority of carriers involved in spontaneous emission events are effectively wasted, as they produce light which cannot couple to the lasing mode. These lost carriers therefore contribute to the overall threshold current of the device. As illustrated in section 2.3.4, the incorporation of strained layers can act to reduce spontaneous emission into unwanted modes, thereby decreasing the radiative current.

As radiative recombination is dependent on the availability of both electrons and holes, the associated current is proportional to both the n and p . In the Boltzmann regime where $n = p$, the radiative current is given by

$$I_{rad} = eVBn^2, \quad (2.29)$$

where the radiative recombination coefficient B for a single, ideal QW is [89]

$$B = \frac{e^2 d}{\varepsilon_0 m_0^2 c^3 k_B T} \frac{n_r E_g \langle M_{av}^2 \rangle}{(m_c + m_v)}. \quad (2.30)$$

Here, e is the electron charge, d is the thickness of the well, n_r is the refractive index, E_g is the bandgap, $\langle M_{av}^2 \rangle$ is the square of the momentum matrix element averaged over all polarisations, and m_c and m_v are the electron and hole effective masses, respectively. All other symbols not explicitly defined herein follow their conventional definitions. It can be seen that, in this regime, the radiative recombination coefficient is independent of carrier density. However, under high injection, this approximation no longer holds and B typically decreases with n . Also, since the threshold carrier density n_{th} increases as $n_{th} \propto T$ in an ideal QW, it can be shown that the temperature dependence of the radiative current is

$$I_{rad} \propto T. \quad (2.31)$$

2.4.3 Auger-Meitner Recombination

An Auger-Meitner process occurs when the energy of an electron-hole recombination is imparted to a third carrier, rather than emitting a photon. The carrier subsequently relaxes by radiating phonons which act to heat the device and deteriorate performance. In the 2-3 μm spectral region, two different Auger-Meitner processes are typically considered, known as CHSH and CHCC recombination.

In CHSH Auger-Meitner processes, the energy from a recombining electron-hole pair (CH-) is imparted to an electron in the SO band which is subsequently promoted to the HH band (-SH). It should be noted that this may also be referred to as CHHS recombination in literature, since the excitation of a HH to the SO band (-HS) is effectively equivalent. During CHCC recombination, the energy is imparted to a CB electron which is promoted further into the CB (-CC). Since these processes are dependent on three carriers, within the Boltzmann framework ($p = n$), it can be shown that the Auger-Meitner current is

$$I_{Aug} = eVCn^3, \quad (2.32)$$

where the Auger-Meitner coefficient is

$$C = C_0 \exp\left(-\frac{E_a}{k_B T}\right). \quad (2.33)$$

In this expression, C_0 is a material dependent parameter, and E_a is the activation energy of the specific Auger-Meitner process. The activation energy mathematically reflects the requirement for both energy and quasimomentum to be conserved. It is defined as the total energy difference between the carriers and their respective band edges for the lowest energy configuration in which the conservation laws can be adhered to. The activation energy of the CHCC process is given as [90]

$$E_a(CHCC) = \frac{m_{cb}}{m_{cb} + m_{vb}} E_g, \quad (2.34)$$

where m_{cb} and m_{vb} are the electron and hole effective masses, and E_g is the bandgap. It is therefore evident that CHCC processes will be particularly deleterious in mid-IR devices since the bandgap is small. This is experimentally supported by numerous studies of devices operating between 2 - 3 μm , which are succinctly summarised in [91].

Since CHSH recombination involves carriers being excited between the SO and HH bands, separate activation energies exist for material where the bandgap is greater than the SO split off energy, and vice versa. They are given by [90]

$$E_a(CHSH) = \begin{cases} \frac{m_{so}}{2m_{hh} + m_{cb} - m_{so}} (E_g - \Delta_{so}), & \text{if } E_g > \Delta_{so} \\ \Delta_{so} - E_g, & \text{if } E_g < \Delta_{so} \end{cases} \quad (2.35)$$

where m_{so} and m_{hh} are the effective SO and HH masses, and Δ_{so} is the SO split off energy. A resonance therefore exists for $E_g = \Delta_{so}$ since carriers at the zone centre are able to contribute, resulting in large CHSH currents.

The activation energies discussed above are derived by assuming simple parabolic, isotropic bands and Boltzmann statistics. For mid-IR devices, smaller bandgaps result in higher non-parabolicity due to increased CB-VB coupling. Thus, this model is likely to be inaccurate for such lasers under high injections, however it remains useful for assessing how the activation energy generally varies with bandgap.

Thus far, the considered Auger-Meitner processes have satisfied quasimomentum and energy conservation laws as transitions occur in the plane of the QW, or in bulk material. However, in devices where the transition energy is greater than the barrier height, it is possible for carriers to be excited to bulk-like continuum states above the barrier. This is able to occur because the heterobarrier acts as a momentum-nonconserving scattering

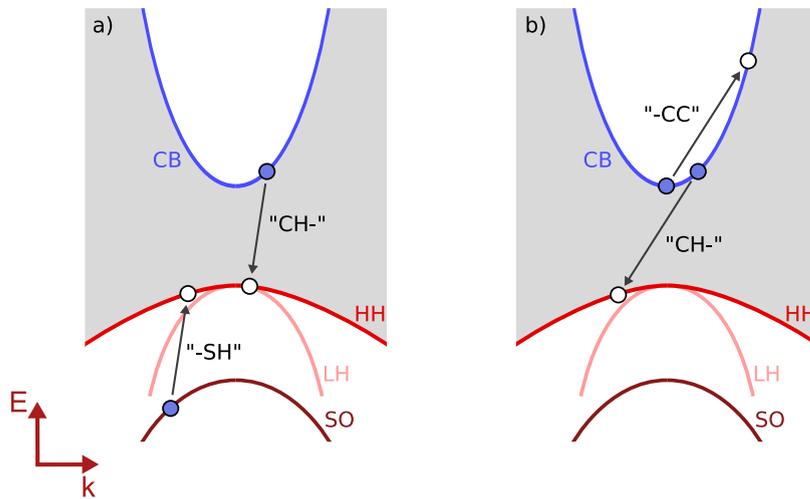


FIGURE 2.15: Typical recombination pathways for the a) CHSH and b) CHCC Auger-Meitner processes. Electrons are depicted with blue markers, while white markers indicate holes.

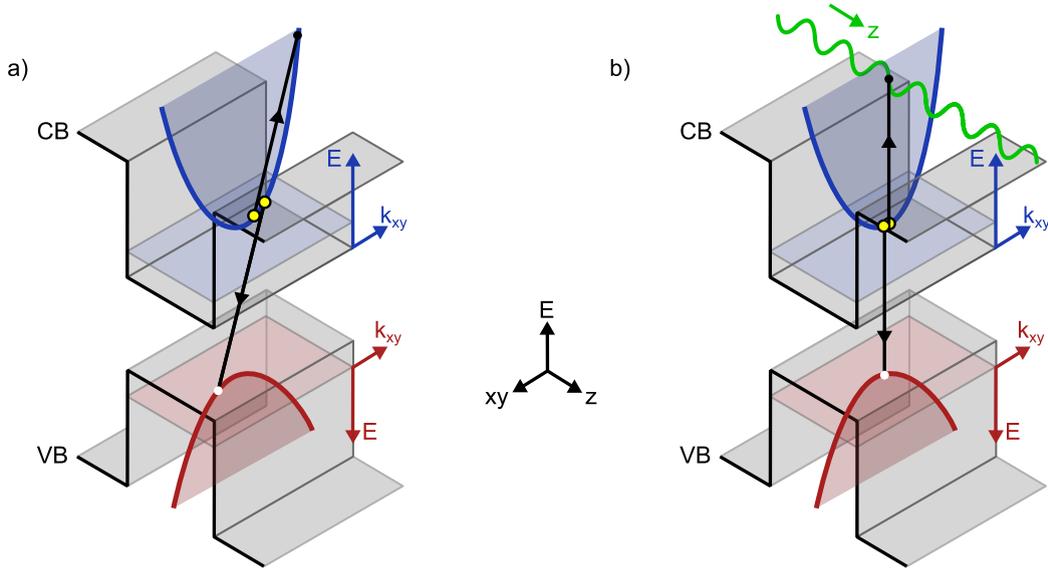


FIGURE 2.16: CHCC recombination pathways for the a) activated process where quasi-momentum and energy are conserved due to transitions in the plane and b) thresholdless processes, where scattering at the QW/barrier interface relaxes quasi-momentum conservation laws, allowing vertical ejection of the electron from the active region. Reproduced and adapted from [95].

site [92]. Here, the lattice symmetries are broken at the interface between the well and barrier which means, whilst the in-plane quasi-momentum remains conserved, the out-of-plane quasi-momentum may change [93]. The ‘hot’ Auger-Meitner carrier may therefore be ejected vertically from the well. This is referred to as a thresholdless Auger-Meitner process since an activation energy is no-longer required, and carriers at the zone centre may contribute. As a result, the thresholdless processes are far less temperature sensitive than activated Auger-Meitner recombination [94].

The activated and thresholdless processes are illustrated for CHCC recombination in Fig. 2.16 a) and b), respectively. It should be noted that the same may apply for CHSH processes, with holes being ejected vertically. The strength of these processes are dependent on the mismatch between the emission energy and the barrier offsets. Thresholdless processes are inherently suppressed for heterostructures in which the barrier offset, $\Delta E_{c,v}$ is greater than the transition energy, E_0 . However, if $E_0 \gg \Delta E_{c,v}$, a large change in the growth direction momentum is required. As the Coulomb interaction required to mediate the transition is low for large momentum transfer, the strength of this process is weakened [96]. Therefore, a resonance exists as the thresholdless Auger-Meitner coefficient is maximised when $\Delta E_{c,v} = E_0$.

2.4.4 Carrier Leakage

The loss stemming from carrier leakage can be divided into two separate processes, both of which act to remove carriers from the Γ -valley of the active material.

The first mechanism involves carriers escaping from the active region into surrounding heterostructure layers. The carriers are then able to recombine in these layers, thus not contributing to gain and raising the overall threshold current. This spatial carrier leakage occurs when the tails of the Fermi distribution extends beyond the barrier energy, resulting in a fraction of the injected carriers occupying unconfined states above the barrier edge. If the leakage is dominated by diffusion and subject to Boltzmann statistics, a simple model can be used to calculate the number of electrons able to contribute to the leakage current. This is given by [67]

$$n_{leak} = \int_{E_{bar}^{cb}}^{\infty} f_{cb}(E)\rho_{cb}(E) dE = N_c \mathcal{F}_{1/2} \left(\frac{E_{fc} - E_{bar}^{cb}}{k_B T} \right). \quad (2.36)$$

Here, $f_{cb}(E)$ is the Fermi distribution for a conduction band quasi-Fermi level denoted E_{fc} , $\rho_{cb}(E)$ is the conduction band density of states, and E_{bar}^{cb} is the energy of the conduction band edge in the barrier. Similarly for holes,

$$p_{leak} = \int_{\infty}^{E_{bar}^{vb}} f_{vb}(E)\rho_{vb}(E) dE = N_v \mathcal{F}_{1/2} \left(\frac{E_{bar}^{vb} - E_{fv}}{k_B T} \right). \quad (2.37)$$

From these equations, it can be seen that the thermal broadening of the Fermi distribution will increase both n_{leak} and p_{leak} , meaning that leakage will be suppressed at low temperature, but could be particularly unfavourable at RT. In reality, the total number of carriers contributing to the leakage current will be lower since a fraction may be recaptured by adjacent wells, and others may have trajectories that aren't conducive to leakage [97]. The leakage current associated with holes is also generally far lower than electrons, stemming from low carrier mobility. However, if the VB offset is particularly low, hole leakage may become appreciable [98].

The second variant involves carriers occupying states in indirect valleys. Since the L and X valleys have large effective masses and higher degeneracies than the Γ valley, their associated DoS can be orders of magnitude higher. Thus, in materials with insufficient energetic separation between Γ and the indirect minima, the fraction of injected carriers

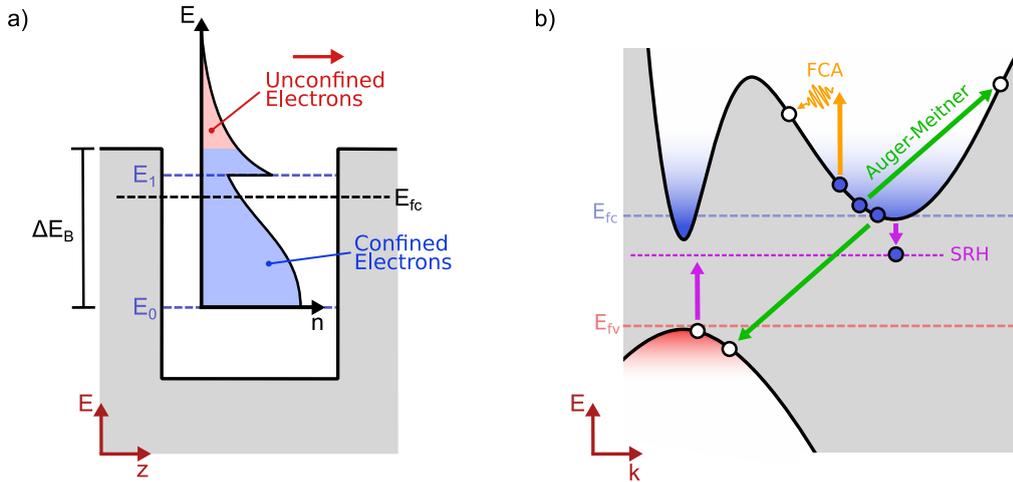


FIGURE 2.17: Schematic of a) spatial carrier leakage, stemming from carriers occupying states above the barrier edge (red) and b) indirect carrier leakage from the tails of Fermi-distribution extending into satellite valleys. Three important non-radiative recombination processes that may occur in indirect valleys are highlighted.

that reside in the L or X valleys can be substantial. In effect, this is therefore a depopulation of the Γ -valley, dictated by the Fermi-function and the proximity of the indirect valleys. For both concision and historical reasons, this is hereafter referred to as indirect valley leakage [99]. As with the spatial leakage discussed above, this leakage in ‘k-space’ is exacerbated by increasing temperature, since a greater fraction of the Fermi distribution overlaps with the indirect states.

Indirect carriers are unable to radiatively recombine in an efficient manner since the transition must be mediated by a high energy phonon in order to conserve momentum. This results in long radiative lifetimes for indirect carriers. For example, in Si, the radiative lifetime is on the order of tens of milliseconds [100]. In this time it is far more likely for carriers to recombine via a non-radiative mechanism, for example via a defect, or Auger-Meitner processes involving the indirect valleys [101]. The latter results in increased device temperatures since phonons are produced when the hot carrier relaxes, leading to further indirect-valley population in a thermal runaway effect. Alternatively, the carriers may act as a source of optical loss through FCA. A similar negative feedback loop is found here as the optical loss means larger carrier densities are required to reach threshold, thereby increasing the lattice temperature and raising the indirect valley population further [102].

2.5 Internal Losses

Free carrier absorption (FCA) and intervalence band absorption (IVBA) are the two main optical loss mechanisms in semiconductor lasers, aside from mirror losses. These losses are detrimental to device performance as it necessitates a larger carrier density to attain net modal gain. In turn, this heightens non-radiative processes and increases J_{th} .

2.5.1 Free Carrier Absorption

FCA involves the absorption of a lasing mode photon by an electron in the CB or a hole in the valence band VB, resulting in an intraband transition. The latter is a special case of FCA known as intervalence band absorption and will be discussed in more detail in the following section. Since photons transfer negligible momentum, the carrier is excited vertically in k -space and thus, for electrons in the CB, FCA must be mediated by a phonon. Absorption is therefore stronger in longer wavelength devices, scaling roughly with $\sim \lambda^2$ [103], so is a particularly important consideration in mid-IR devices.

This process is inherently dependent on the density of available free carriers and is therefore particularly detrimental in highly doped layers. As such, it is vital that the waveguiding region of the heterostructure is designed to minimise the overlap between the tails of the optical mode and the cladding material. Additionally, as mentioned previously, materials with low levels of directness can also provide a large concentration of free carriers due to long radiative recombination lifetimes. If such materials exist in the proximity of the active region there will also be a large photon density which could exacerbate FCA losses.

2.5.2 Intervalence Band Absorption

In IVBA, electrons occupying lower valence band states can absorb photons resulting in transitions between the SO-HH, SO-LH and LH-HH bands. Since the VB consists of three distinct bands, transitions between states of the same k -vector are possible, and phonon mediation is not a necessity. Since the VB quasi-Fermi energy does not typically penetrate into the VB, only states near the band edge are populated by holes. Therefore, intervalence band transitions are heightened when the emission energy is close to resonance with the energy separation at Γ , since carriers in the vicinity of $k=0$ can contribute.

In mid-IR devices, the bandgap can be similar to the SO splitting energy, resulting in a large SO-HH IVBA contribution. Δ_{so} is therefore a key consideration when designing the active region material. As the wavelength increases further, LH-HH transitions become increasingly detrimental. Since the hole density in the LH band is significantly lower than the HH band for unstrained and compressively strained materials, IVBA from LH-SO transitions is typically weakest. Methods for calculating IVBA theoretically were outlined in [104] and are reproduced in chapter 5 for completeness.

2.6 Summary

This chapter detailed the fundamental theory of semiconductor lasers required to follow the material in the thesis. Firstly, the concept of electronic bandstructure was introduced, focusing specifically on tetrahedrally bonded, zinc-blende crystal lattices. The DoS and the notion of a Fermi-level were discussed, leading to an understanding of carrier concentrations and equations for calculating the carrier density in bulk materials. The three primary optical processes occurring in semiconductor materials were briefly outlined, before the three fundamental requirements for producing lasing were detailed. Subsequently, the typical heterostructure design for a double-heterostructure laser was discussed, noting how each layer is chosen to optimise both carrier and optical mode confinement. The final section pertains to how carriers recombine across the bandgap and optical loss processes which may hinder device performance, paying particular attention to defect, radiative, Auger-Meitner and leakage effects.

Additional theoretical concepts, expanding upon those discussed in this chapter, are included in subsequent chapters and appendices, where required.

Chapter 3

Experimental Techniques

3.1 General Set-up

All electroluminescence-based experimental studies conducted in this body of work utilise the same underlying set-up, illustrated in Fig. 3.1. The measurements typically involve varying the injected current and recording an aspect of the emitted light, for example, the total light emitted from a single facet or an emission spectra.

Devices are mounted and contacted as per the requirements of the specific technique (to be discussed below), before being biased using square-voltage pulses. Internal self-heating can obfuscate trends in the experimental data and complicate interpretation. Consequently, lasers were always operated in pulsed mode, allowing the device temperature to stabilise to that of the heat sink between consecutive pulses. In this vein, a key metric to be considered is the duty cycle,

$$D = \frac{w}{T} = wf, \quad (3.1)$$

which specifies the fraction of time for which the device is biased. Here, w is the width of the pulse, T is the time elapsed between pulses and f is the pulse frequency. This value should be optimised to reduce internal heating whilst ensuring sufficient output power can be obtained. Nominally, 500 ns or 1 μ s pulse widths were supplied at 1-10 kHz during this work, resulting in duty cycles of between 0.05 - 1 % which have proven adequate in previous measurements [105]. The choice of pulse generator was determined by the current required to reach threshold in the given device. The GaSb-based QWs (chapter 6) had thresholds below 2 A. For these devices, an AVTECH AV-1011A-C was used, capable of driving currents of 2 A with up to 100 V. The current requirements were significantly

higher for the bulk GeSn devices (chapter 5), and therefore an AVTECH AVRZ-5W-B unit was employed, capable of driving 10 A with up to 500 V. Current through the device was monitored using a current probe (BNC6040+202H), connected to a Tektronix TDS 3012 two-channel oscilloscope that was triggered by the TTL output of the respective pulse generator. Since the impedance of a laser diode under forward bias is typically 2-3 Ω , a 47 Ω resistor is connected in series for impedance matching to the standard 50 Ω BNC and mains cables.

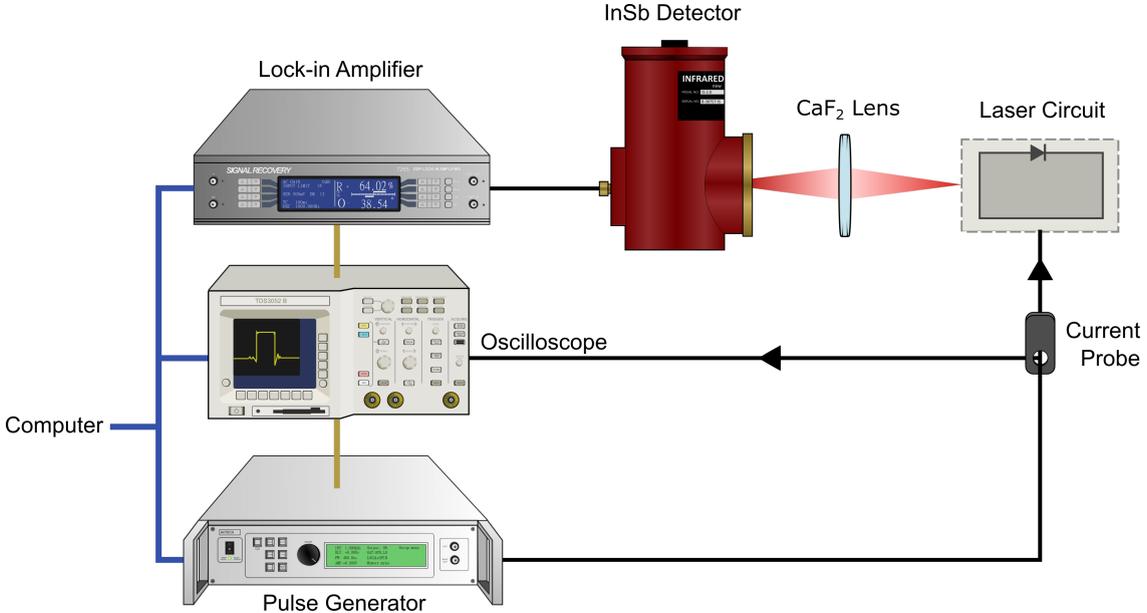


FIGURE 3.1: Schematic of the core elements utilised in electroluminescence measurements. Connections from the pulse generator to the lock-in amplifier and oscilloscope are shown in gold. Blue lines represent GPIB connections.

When measuring the light output from a single facet as a function of current, emission was focused onto a detector using calcium fluoride (CaF_2) lenses, since standard glass is highly absorbing in the mid-IR, as illustrated in Fig. 3.2. For measurements of spontaneous emission from GaSb-based QWs in chapter 6, light was coupled from the Si substrate into a Thorlabs FG105LCA low hydroxyl ion, multi-mode fiber and onto the detector. This specific model was utilised as it extended the wavelength range to 400-2400 nm, ensuring high transmission for these devices. The detector utilised for these measurements was an InfraRed Associates IS-2.0 liquid nitrogen cooled InSb detector. Signal from the detector was amplified using a Signal Recovery 7265 DSP lock-in amplifier. It does so by comparing the triggering signal from the pulse generator to the output of the detector. Any signals arriving at the triggering frequency will be amplified, whilst other signals are attenuated, enabling significantly improved signal-to-noise ratios. GPIB cables from the lock-in amplifier and pulse generator were connected to the oscilloscope, and a

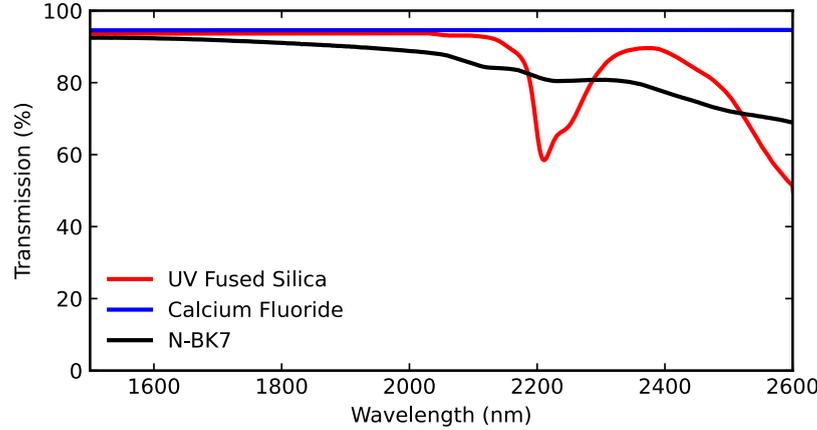


FIGURE 3.2: Transmission spectra for 10mm of UV fused silica (red), CaF_2 , and N-BK7 glass. Raw data available through Thorlabs [106]–[108]

National Instruments GPIB-USB-HS was used to connect the hardware to a computer via standard USB. Device control and acquisition automation was then achieved using LabView programs developed internally by Dr Igor Marko.

Spectral measurements of the facet emission described in chapter 6 were conducted using an OceanOptics NIRQuest 512-2.5 with a detection range of 900 - 2500 nm and spectral resolution of 6.3 nm. Here, emission was focused by a CaF_2 lens into an OceanOptics bifurcated optical fiber, with one output connected to the spectrometer and the other placed in front of the InSb detector such that measurements could be conducted simultaneously.

3.2 Temperature Dependent Measurements

Analysis of the dominant recombination mechanisms is typically achieved by assessing how certain performance metrics change as the bandstructure is varied. As discussed in chapter 2, the bandstructure is intrinsically linked to the lattice-spacing and may therefore be altered by changing the device temperature. In this work, two different cryostats are utilised, enabling measurements to be performed at either low or high-temperatures.

The first is a Leybold closed cycle helium cryostat, depicted in Fig. 3.3, capable of achieving temperatures between 20 - 300 K. During these measurements, bar devices were mounted to a clip developed by Professor Stephen Sweeney, illustrated in Fig. 3.4. The mount is formed of a copper base plate separated from a copper clip by an insulating spacer, and serves to secure the position of the laser device whilst also providing electrical contact. The clip segment has a small hole in the end such that a microscope can be used

to align to the p-contact of the device. A $10\ \mu\text{m}$ thick wire is threaded through the hole and along the underside of the clip to make contact. For devices with a bottom contact, the copper base plate then acts as the n-contact. The devices measured in this work are all top contacted (laterally injected) to avoid losses in defective buffer layers. As such, the circuit is completed by placing silver conductive paint between the bottom plate and an n-contact on the laser bar. For measurements of the spontaneous emission, a hole is made in the base-plate, into which an optical fibre can be inserted and glued in place.

The laser clip is mounted at the end of a cold finger which is situated in vacuum so to prevent condensation and provide thermal isolation. The cold-finger is cooled using a Coolpak 2000 compressor which pressurises He gas into its liquid state and pumps it into the cold head. At this point, the He is allowed to expand which enables heat to be extracted from the cold finger, which is in thermal contact with the device. The He is then returned to the compressor where excess heat is transferred to an external water-cooling loop before it is re-compressed.

In the case of good vacuum, at pressures of 10^{-5} - 10^{-6} mbar, temperatures as low as 20 K can be attained. To access temperatures between 20 K and RT, the cooling is countered using a coil of nichrome wire situated at the end of the cold finger which acts as a heater. The temperature is then monitored using a thermistor placed at the base of

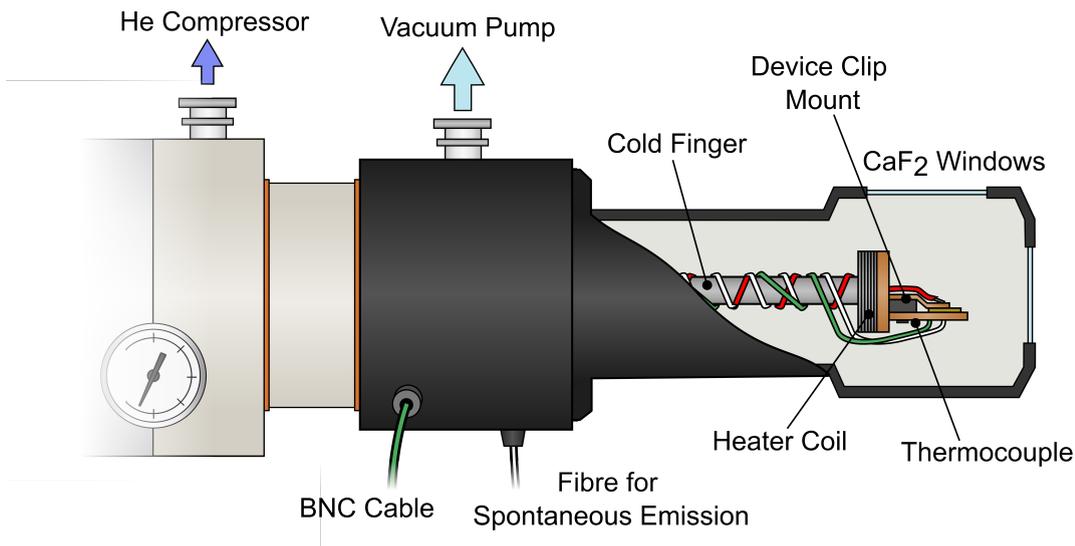


FIGURE 3.3: Diagram of the Leybold closed cycle He cryostat. In this depiction, the clip mount is illustrated, however alternative fixtures can be fitted for measuring devices on TO headers.

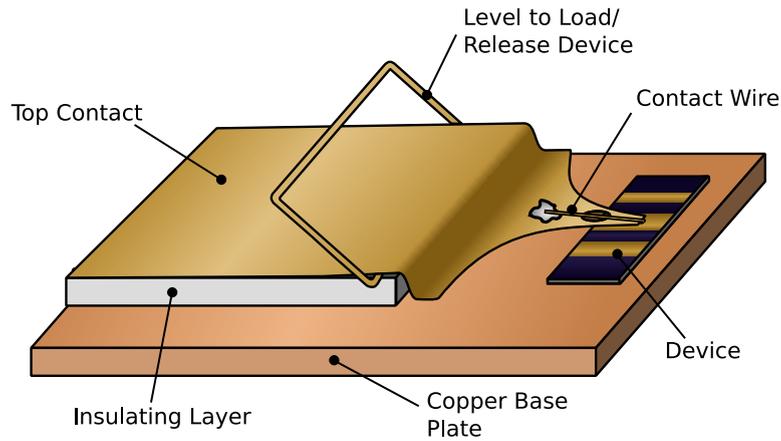


FIGURE 3.4: Illustration of the clip mount. To attain a complete circuit, positive and negative wires are connected to screws on the clip and copper plate, respectively.

the clip which is connected, along with the heater, to an Oxford Instruments ITC 502S temperature controller.

For devices pre-mounted on TO headers, or for those requiring measurements above RT, an Oxford Instruments Optistat DN2 gas exchange cryostat was employed. Here, an outer and inner chamber are separated by a liquid nitrogen (LN_2) reservoir. The sample is mounted to an arm and inserted into the inner chamber, which is evacuated to ensure no condensation occurs, since this may affect facet reflectivity and alter the device performance. Dry He gas is then pumped into the inner chamber to provide thermal contact with the laser, and the reservoir is supplied with liquid nitrogen. The two then interact at the heat exchanger which removes heat from the device, enabling temperatures to reach that of the LN_2 at 77 K. The outer chamber is evacuated to ensure the inner chamber and nitrogen reservoir are thermally isolated from the environment. As per the closed-cycle cryostat, the cooling is countered using a heater coil, enabling temperatures of up to 450 K to be attained. This convection based temperature control is often preferable as it promotes more uniform temperature distributions and faster heat dissipation when compared to conduction-based systems.

In all temperature-dependent measurements conducted in this body of work, a 20 minute wait was implemented subsequent to each temperature step to ensure that the device has reached equilibrium with the surrounding environment.

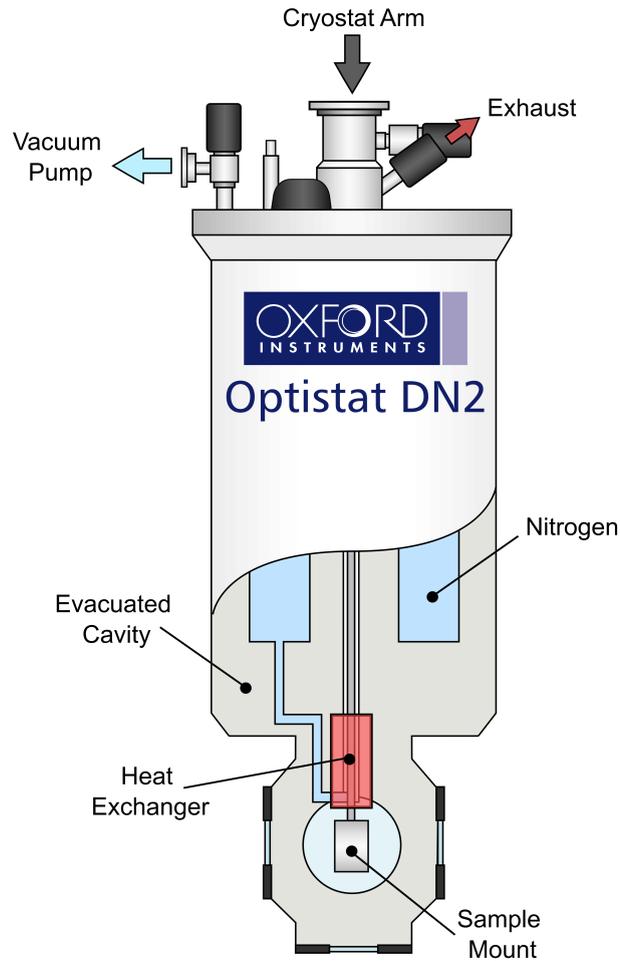


FIGURE 3.5: Schematic of the Oxford Instruments Optistat DN2 gas exchange cryostat, adapted and reproduced with permission from [95].

3.3 High Hydrostatic Pressure Measurements

The application of hydrostatic pressure to a semiconductor material can provide rich information about underlying physical processes. For example, this technique has previously been used to explore phase transitions, defects and impurities, lattice dynamics, and assess the performance of emitters and detectors [109]. The key merits of the method stem from the ability to tune the bandgap of the material, without altering the thermal distribution of carriers, as per temperature-dependent techniques. Under hydrostatic pressure the lattice constant is reduced equally and reversibly along all three axes. The resulting change in bandgap with pressure at the Γ -, X- and L-points, known as the pressure coefficients, are described by William Paul's empirical rule, which states [110]:

” Independent of the family member studied, the pressure coefficient of the

direct energy gap at Γ falls in the range 10-15 meV/kbar, that of the indirect gap at L near 5 meV/kbar, and that of the indirect gap at X between -1 and -2 meV/kbar.”

This holds true for both III-V and group-IV materials. Since the pressure coefficients take nominally similar values across a wide range of materials, the bandgap can be engineered without altering key heterostructure features, such as band offsets. Minor variation in the compressibility of different layers will generally result in changes to the in-plane strains with increasing pressure. This is typically ignored since the compressibility is similar for III-V materials, resulting in shifts in strain on the order of 0.01% over 10 kbar [95]. However, marked material differences between III-V alloys and group IV materials mean that this effect should be considered explicitly when studying III-V devices on Si.

The technique also allows the band ordering to be tuned from direct to indirect, and vice versa, which can give valuable information about the bandstructure in more novel materials and devices such as GeSn. Since the effect of alloying and pressure on the CB minima are similar, the performance of different alloys can be probed without the need to grow numerous structures of varying compositions, thus reducing potential complications the observed trends arising from differences in doping, heterostructure quality etc. Such comparisons are made in chapter 5.

3.3.1 Method

High hydrostatic pressure is applied using a Unipress U11 helium gas pressure system connected to a Unipress gas optical cell housing the device. The cell consists of a monoblock of copper-beryllium (CuBe) alloy which is cylindrically hollowed. This allows a multitude of CuBe plugs to be inserted for various applications. In this work, facet emission is collected through an optical plug with a sapphire window, providing $> 80\%$ transmission for wavelengths between 400-5000 nm. Devices are mounted to a similar plug fitted with a sample holder and electrical feed-through. For bar devices, samples are mounted using the clip design discussed in section 3.2. A specific mount is also available for TO headers which requires the pins to be bent through 90° ; this is the plug most frequently utilised in this work.

The U11 system consists of three hydraulic stages connected to a He gas cylinder and is schematically illustrated in Fig. 3.7. The device is placed under pressure by compressing

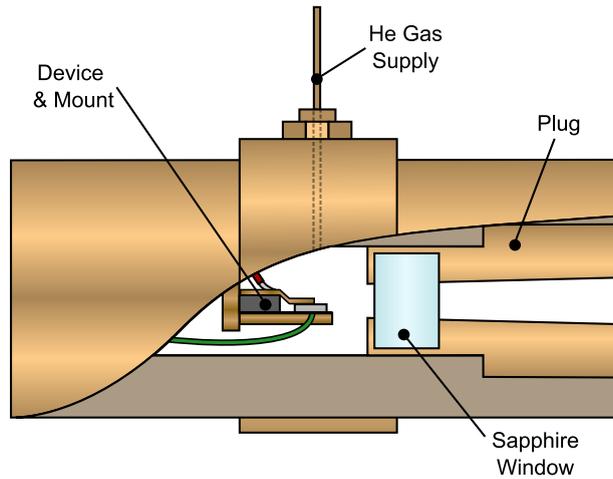


FIGURE 3.6: Cross-section schematic of the pressure cell. Here, the cell is fitted with plugs containing the clip mount and a sapphire window. These can be swapped out for a number of alternatives depending on the application.

a fixed quantity of gas into a smaller total volume. He gas acts as an excellent medium for this purpose since its refractive index is almost exactly 1 in the mid-infrared, whilst also boasting a thermal conductivity $5\times$ greater than air.

During set up, the stages and lines are filled with He, up to a pressure of 0.12 kbar before the input valve is closed to ensure a fixed volume of gas. Hydraulic oil is introduced beneath the arm to lift the piston in the first stage. This forces gas from the first stage into the lines and chambers of the second and third stages. At this point, the valve to the first stage is closed, and the hydraulic oil is drained from the first chamber. This enables the pressure on the hydraulic pump oil to remain below its quoted maximum of 0.7 kbar when further increasing pressure on the device using the second and third stages. The procedure is repeated for the second and third stages, attaining 3.5 kbar and ~ 9 kbar, respectively, whilst measurements are taken at intermediate pressures. While 15 kbar is the quoted maximum from the manufacturer, due to the high pressures involved, actual performance is limited by leakages which can occur from the CuBe cell plugs and/or the stages as a result of small scratches or debris. As such, great care must be taken when cleaning the cell and third stage prior to measurement. Upon reaching the maximum required pressure, the cell is slowly returned to atmosphere by opening the valves in the reverse order. Measurements are taken at multiple points during this process to ensure that recorded threshold values are repeatable and the device was not damaged or degraded.

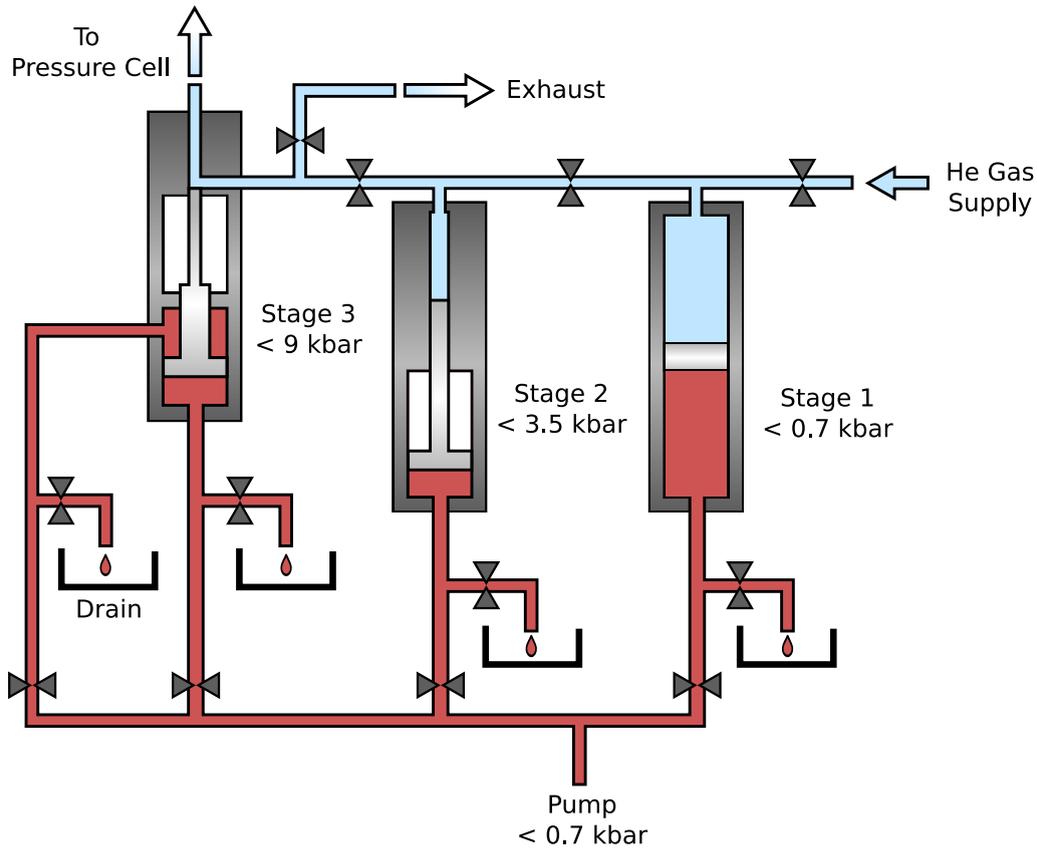


FIGURE 3.7: Depiction of the Unipress U11 three-stage He compression system. He is represented in blue, whilst hydraulic oil is shown in red. Bowtie symbols schematically illustrate the positions of valves.

3.3.2 Low-Temperature, High Pressure Measurements

For some more novel devices, cryogenic temperatures are required to attain lasing. In this instance, determination of dominant recombination mechanisms is even more vital to inform new heterostructure design and growth. Additionally, thermally activated Auger-Meitner and leakage processes can be reduced or even suppressed at low temperature such that information about other recombination pathways can be obtained. As such, performing high pressure measurements at low temperature is highly desirable.

In this work, this is achieved by placing the CuBe cell inside a bespoke cryostat set-up, built upon a Leybold Coolpower 140 T cold head. The apparatus is displayed schematically in Fig. 3.8. Here, the cold head operation is as described for the Leybold closed cycle helium cryostat in the previous section. The CuBe cell is clamped into a brass mount at the end of the cold finger to provide thermal contact. The unit is evacuated and an Oxford Instruments ITC 502S temperature controller is used to monitor the temperature.

It should be noted that the temperature must be set prior to application of pressure, since changing the temperature of the He gas in the cell will alter the volume and therefore the applied pressure. Due to the vast size of the unit, it is difficult to achieve high levels of vacuum. As such, the minimum achievable temperature currently sits at ~ 60 K, but could potentially be improved with a more powerful pump and new seals.

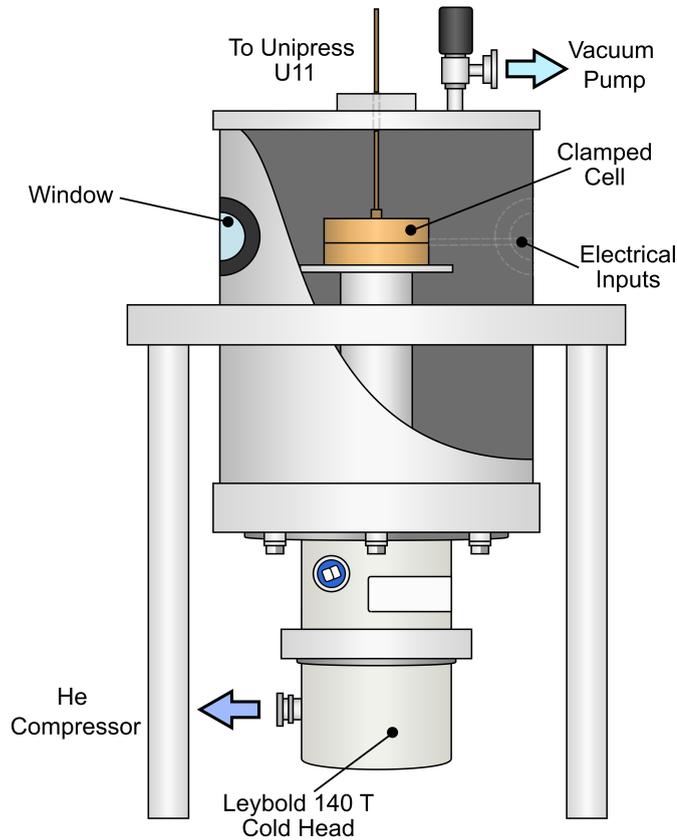


FIGURE 3.8: Schematic diagram of the bespoke cryostat used from low-temperature, pressure dependent device characterisation.

3.4 Light-Current Curves

Many of the measurements discussed in this body of work focus on collecting light as a function of increasing current under a variety of conditions. This is known as a light-current (LI) measurement. In this work, light is collected from a single facet or directly through the Si substrate, as illustrated in Fig. 3.9 a). The corresponding LI curves for facet emission and emission through the substrate are illustrated in Fig. 3.9 b).

For facet emission, a typical LI curve has a low-gradient region where there are small increases in light output with current stemming from heightened spontaneous emission (SE). Above a critical threshold current density, the light output increases rapidly owing to the onset of lasing. From this data, there are two key parameters to extract, these are the external differential efficiency above threshold, $\eta_d \propto dL/dI$, and the threshold current density J_{th} . There are typically four methods that are readily utilised for determining threshold from an LI curve [111]. In this work, straight lines are fitted through the spontaneous emission and stimulated emission segments respectively, using a least-squares fitting approach. The point at which these lines intersect is taken to be threshold. This is schematically by dotted lines in Fig. 3.9 b).

For SE measurements, light is collected through the substrate of the material. Since Si is highly transparent between 2-3 μm and all devices investigated are laterally contacted, emission can be collected with no further processing required. By collecting light along the growth direction, it cannot undergo any appreciable levels of gain or loss due to thin (~ 10 nm) wells, and wider bandgap cladding and buffer layers [85]. As such, only SE should be collected. SE transitions between the CB and HH states produce light polarised in the x-y plane which propagates in the growth direction. However, CB-LH transitions are predominantly z-polarised and propagate in the plane, hence if these transitions are

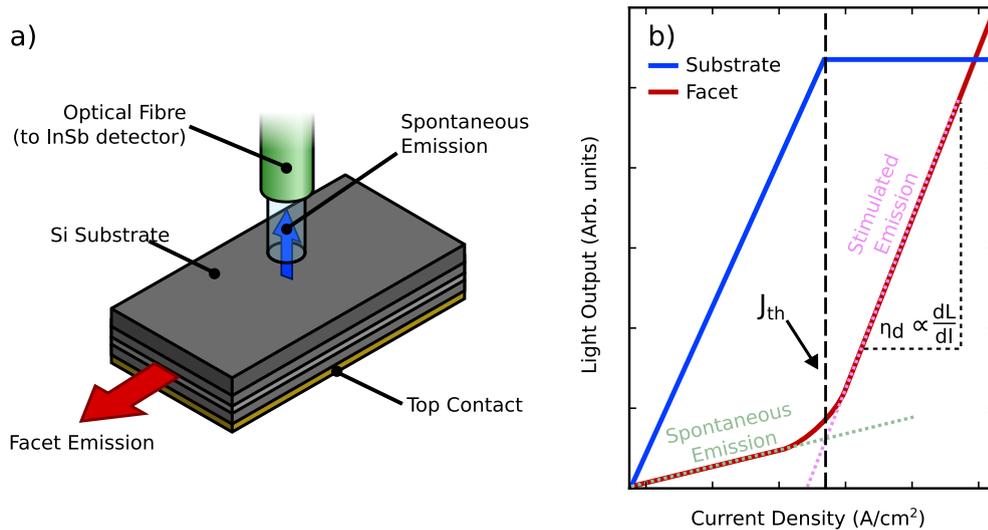


FIGURE 3.9: a) Schematic of collection geometry for facet (red) and substrate emission (blue). Adapted and reproduced with permission from [95]. b) Normalised LI curves for emission from the substrate (SE) and facet (lasing), illustrating the changes in behaviour above and below threshold, J_{th} . The intersection of pink and green dotted lines show how threshold is determined.

prevalent the total SE may be underestimated [80]. For the compressively strained active layers in these studies, LH states are energetically remote and are therefore not expected to contribute significantly to the total SE.

The LI curve of a SE measurement initially rises with current due to increased SE, as depicted in Fig. b). In a real device, the portion of the graph below threshold may be curved, which can give important information about the dominant current path. Above threshold, in an ideal device, the light output pins. This reflects that any further carriers injected into the device recombine rapidly via stimulated emission, which pins the carrier concentration and therefore the SE [112].

Chapter 4

Bulk Group-IV Bandstructure Modelling

4.1 Introduction

A laser capable of leveraging current complimentary metal-oxide-semiconductor (CMOS) technologies is considered the ‘holy grail’ for Si-based photonic and optoelectronic integrated circuits (PICs & OEICS). Whilst hybrid integration of III-V lasers on Si has proved successful [113], [114], this process suffers from limited component integration densities and high cost stemming from the use of more expensive native substrates [115]. Epitaxial growth on CMOS compatible substrates, known as monolithic growth, is therefore preferable. However, III-Vs grown directly on group IVs suffer from high-defect densities due to lattice, thermal-expansion and polarity mismatches, which greatly hinder performance [116]. Thus, a true direct-gap group-IV light source remains highly desirable. Such a device promises to decrease the effective footprint and cost associated with PICs by utilising highly mature facilities and processes afforded by current electronics foundries [117].

From a fabrication standpoint, the most ideal light source would be a Si laser, since it would offer seamless integration into current Si PIC and OEIC processes. However, due to its indirect bandgap, Si is an inherently poor emitter. From the bandstructure illustrated in Fig. 4.1, it can be seen that conduction band (CB) minima is located $0.85 \frac{2\pi}{a}$ along the Δ direction, close to the X-point, and is around 3 eV below the CB edge at the Γ -point. Since carriers will primarily occupy these indirect states, radiative transitions must be assisted by phonons in order to conserve momentum. The requirement for a third particle

to mediate the transition increases the radiative lifetime to the point where non-radiative recombination pathways become preferable.

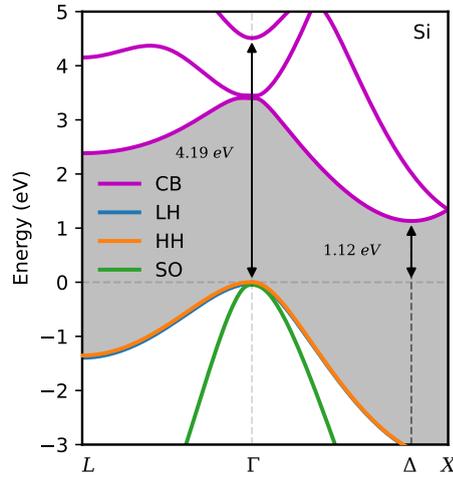


FIGURE 4.1: The bandstructure of Si calculated using an EPM, available on nanoHUB [118].

Ge is another material widely utilised in CMOS foundaries, and has a bandstructure which is better suited to emission when compared to Si. The bandstructure of Ge is illustrated in Fig. 4.2 a), showing that the material is indirect, with a CB minima at the L-point. However, the energetic splitting between the L- and Γ -valley minima, known as the *directness*, is only 136 meV [119]. As such, it is easier to populate the direct valley and induce radiative transitions.

The next heaviest group IV element is Sn, which can be found in two crystalline phases, α and β . In this study α -Sn is the phase of interest since it shares the same diamond crystal structure of Si and Ge. Fig. 4.2 b), shows that this is a zero gap material. In α -Sn, the LH band and CB are inverted, such that the HH and CB are degenerate at the Γ -point and a bandgap opens between the HH and LH bands. This occurs as, unlike Si and Ge, the CB is formed from p-like orbitals, and the LH band from s-orbitals [120].

It can be seen that the group-IV materials span from very wide gap through to semi-metallic, which suggests the possibility of large-scale bandgap tunability through alloying. Due to its minimally indirect bandgap, strain engineering of Ge is also an attractive route for achieving a truly direct gap group-IV laser; this formed much of the early work on direct-gap group IV lasers. In the last decade there has been intense interest in alloying Ge with relatively low amounts of Sn to produce direct-gap material that also provides good

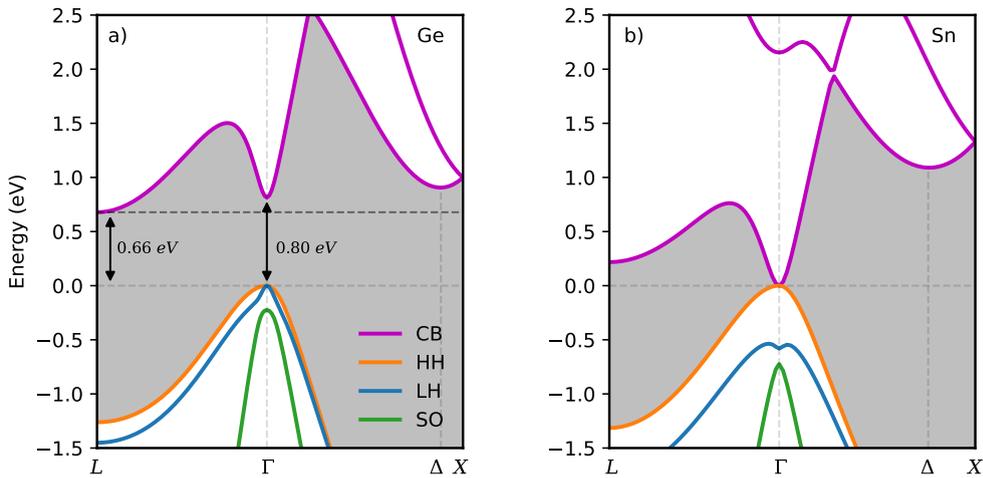


FIGURE 4.2: The bandstructure of a) Ge and b) α -Sn, showing a nominally indirect and semi-metallic behaviours, respectively. Ge bandstructure was calculated using an EPM [118], whilst the Sn bandstructure was calculated using DFT, with the inclusion of SO interactions.

carrier transport properties [121]. This was enabled by advances in growth techniques and the expansion of the available precursors. Growth of good quality GeSn is difficult since Sn has less than 1% solubility in Ge. Despite this, GeSn/Ge superlattices with 26% Sn were grown as early as 1995 using molecular beam epitaxy (MBE). This technique is inherently slow, making the process difficult to scale for manufacturing, which limited impact. Early chemical vapour deposition (CVD) techniques utilised the high-cost SnD_4 precursors, but progress was made in IMEC and KTH who were able to grow layers using SnCl_4 and Ge_2H_6 , and later by al-Kabi et al. who implemented SnCl_4 and GeH_4 , all of which are commercially available [122], [123].

The effect of alloying with the semi-metal is to reduce the bandgap at the zone centre faster than the indirect gaps. Initial estimates of the indirect-direct crossover point were placed at $\sim 20\%$ Sn, however experimental measurements and computational modelling now estimate a gradual change in band character occurring between 6-10% [18], as shown in Fig. 4.3 a). Since Sn incorporation is $\leq 1\%$, this discovery made truly direct group-IV material significantly more feasible. Despite this, as shown in Fig. 4.3 b), incorporating such levels of a relatively heavy element results in a steep increase in the natural lattice constant of the alloy, resulting in high levels of compressive strain when grown directly on Si and Ge. This can yield highly defective material. As such, the tuning of the bandstructure and reduction of defects, amongst other things, are vital to producing commercially viable devices.

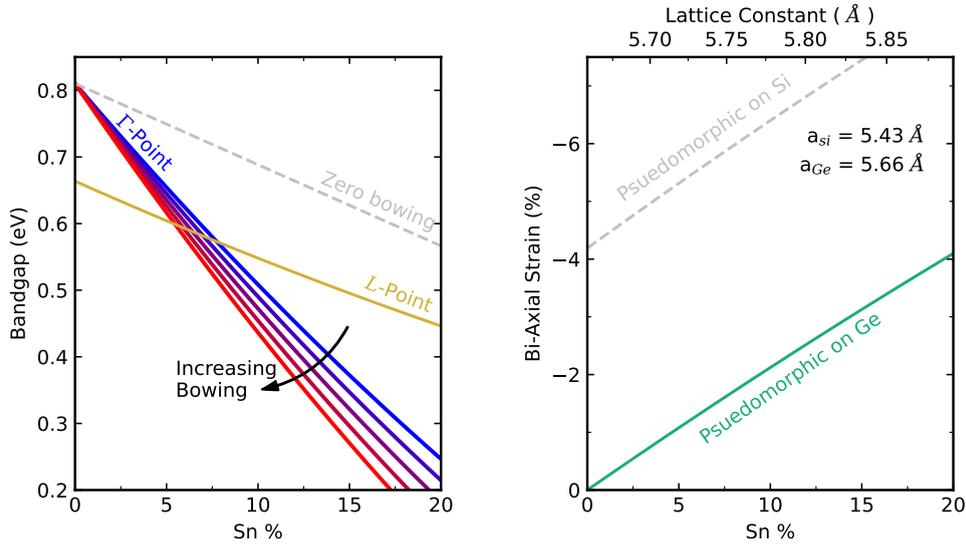


FIGURE 4.3: a) The bandgap at the Γ and L-points as a function of increasing Sn. Blue-to-red coloured lines depict the Γ energy calculated while increasing the Γ bowing parameter between 1.94-2.92 eV, as per the values quoted in literature. b) Shows the level of biaxial strain when growing GeSn alloys pseudomorphically on Ge (green, solid) and Si (grey, dashed).

In the following sections, the history of GeSn based lasers are outlined, highlighting key findings which should be implemented during the device design stage. 6-band and 8-band k-p Hamiltonians are then established to allow the bandstructure of bulk group IV material to be accurately modelled. A substantial review of the parameters utilised in the modelling of Si, Ge, Sn and their multinary alloys is undertaken, before the validity of bandstructure simulations are compared to notable III-V materials and finally group-IVs. This chapter therefore represents the theoretical framework required to understand the material system, and analyse data from electrically injected GeSn lasers presented in the ensuing chapter.

4.1.1 Device History

Production of direct group-IV materials initially centred around the engineering of Ge, since the Γ -valley is situated only 136 meV above the L-valley minima. Preliminary attempts utilised highly doped Ge, such that L-valley states were highly populated, increasing the density of carriers generated in the Γ -valley under optical pumping. Between 2010 and 2012, lasing was reported in both optically [124] and electrically [125] pumped devices doped to $1 \times 10^{19} \text{ cm}^{-3}$ and $4 \times 10^{19} \text{ cm}^{-3}$, respectively. However, subsequent studies were unable to replicate these findings using similar devices, citing that net material gain could not be achieved as a result of heightened intervalence band absorption [126],

[127]. Focus subsequently switched to methods of bandstructure engineering through the incorporation of tensile strain in the form of microbridges [128], SiN stressor layers [129] and ion implantation of thin-films [130].

The first GeSn based laser was reported by members of the Peter Grünberg Institute 9 (PGI9) in their seminal 2015 paper [19]. The study consisted of five samples of CVD grown GeSn, with Sn compositions between 8-13%, grown on Ge virtual substrates (VS). By etching Fabry-Pérot (FP) cavities of differing lengths, optically-pumped lasing was demonstrated in one sample (13% Sn) above an excitation power density of 325 kWcm^{-2} at 20 K using the variable stripe length technique. This was affirmed by observation of FP oscillations in the photoluminescence spectra. The lasing peak was lost above 90 K, which was attributed by the authors to increased loss, stemming from the onset of Shockley-Read-Hall (SRH) recombination and the proximity of the Γ and L-valleys.

Subsequent progress in GeSn-based devices has primarily been driven by PGI9 and a research group at the University of Arkansas (UA), with other notable mentions being the CEA, based at the University Grenoble Alpes. In 2016, PGI9 followed their initial paper by demonstrating an optically pumped, under-etched microdisk (MD) laser which offered performance improvements in both maximum operational temperature (130 K) and threshold power density (220 kWcm^{-2} @ 50 K) [131]. Such large improvements were attributed to the under-etching of the Ge VS, which resulted in enhanced optical mode confinement when compared to the previous FP structures, and provided strain relaxation to improve band alignment for gain production. The under-etching was also successful in removing large amounts of defective VS material which likely also contributed to the observed improvements, however this was not explicitly reported by the authors. In the same year, UA reported an optically pumped GeSn FP laser, akin to that of the initial PGI9 paper, operating at $2.5 \mu\text{m}$ up to a maximum of 110 K. Here, the authors used a GeH_4 precursor rather than Ge_2H_6 , allowing for the production of high quality, intrinsically relaxed layers. Similar to the under-etched microdisk, this reduction in defect density allowed the maximum operational temperature to be raised.

In 2017, the maximum operational temperature for optically pumped devices was further increased, with both CEA and UA reporting operation up to 180K [132], [133]. Both groups did so by increasing the Sn% of the active material, with CEA adopting a MD structure with 16% Sn and UA opting for a FP heterostructure with 17.5% Sn. The high Sn compositions resulted in a substantial increase in operational temperature due to

increased separation between the Γ and L-valleys, meaning carriers preferentially populated the direct states where they could contribute to gain production. Such compositions were previously thought to be beyond the Sn incorporation limit when using the adopted deposition techniques, regardless of the choice of precursor, due to the low temperature required to prevent Sn diffusion/segregation [134]. However, UA established that compressive strain was also a factor limiting incorporation. They therefore utilised a GeSn VS, enabling compressive strain to be reduced in the active material and leading to heightened Sn compositions. A similar approach was implemented by CEA, where graded GeSn buffer layers were grown on Ge strain relief buffers, also resulting in reduced compressive strains in the active region.

An important milestone was achieved in 2018 when PGI9 reported the first instance of lasing in multi quantum well (MQW) devices [135], [136]. As demonstrated in III-V lasers, confinement provides a pathway to lower thresholds when compared to bulk devices due to the 2D density of states which improve gain characteristics and thermal-stability of the quasi-Fermi levels. Such improvements are also expected for group-IV based MQW heterostructures. The best reported device featured 22 nm GeSn_{0.13} wells and 22 nm Si_{0.05}GeSn_{0.13} barriers, which operated up to 120 K under 1550 nm optical pumping at 45 kW/cm². This illustrates a ten-fold improvement in threshold power density when compared to a double-heterostructure design with 14.5% Sn, however no increase in the maximum operational temperature was observed. This was primarily attributed to the difference in effective mass between the Γ - and L-valleys. Since the direct valley is around 3 \times lighter than the L-valley, confinement results in a decreased directness, meaning the enhanced characteristics observed at low temperatures are offset by increased indirect-state occupation at higher temperatures. Nevertheless, this remains an important step in lowering the operational temperature of future devices.

At the same time, CEA reported an optically pumped micro-disk structure with a threshold of 134 kW/cm² at 15 K, competitive with other bulk active region devices at the time, but with a maximum operational temperature of 230 K [137]. This built upon the groups previous 16% Sn device operating up to 180 K. Whilst the former consisted of a single layer of GeSn, the improved device utilised a GeSn_{0.138} / GeSn_{0.16} / GeSn_{0.138} double heterostructure design. This acted to provide greater optical mode and carrier confinement which, when paired with the graded strain relief layers, under-etching and high Sn% active region, enabled operation at temperatures achievable using commercially

available coolers. By applying a SiN stressor to a similar device geometry, lasing was achieved under optical pumping at 300 K, becoming the first RT group-IV laser in 2022 [20].

During this time, UA reported lasing at 270 K in an optically pumped FP structure, without use of stressor layers, by attaining bulk material at 20% Sn. Improvements were also attributed to the use of broad area ridges which increased heat dissipation capabilities and lowered surface recombination effects [138]. This was followed by a major milestone in 2020, with the first reported electrically injected GeSn laser, based on a bulk double-heterostructure design and an 11% Sn active region which elicited lasing up to 100 K [21]. These devices form the basis of the studies in chapter 5.

Since this groundbreaking study, the group has managed to raise the operational temperature to 140 K, primarily through improving modal confinement with the active region and reducing free-carrier losses in the metal contacts [22]. Electrically injected GeSn lasers have also been achieved for devices with different geometries by CEA, but the operational temperature remains below 100 K [139]. To push towards higher temperature operation in future devices, it is important to analyse the performance limitations in current designs. That is the focus of the following chapter.

4.2 **k**·**p** Theory

In order to analyse the carrier recombination processes limiting the operational temperature, it is fundamental to model the material bandstructure. In this body of work, a time-independent perturbation theory in the form of the **k**·**p** methodology is adopted.

The basis of **k**·**p** theory involves writing the single electron Schrödinger equation in terms of the cell-periodic components $u_{n\mathbf{k}}(\mathbf{r})$ of the Bloch states, $\psi(\mathbf{r})$. Such treatment allows the Hamiltonian to be written using terms which are **k**-independent, and terms with an explicit **k**-dependence. The **k**-dependent terms may then be treated using perturbation theory, allowing the Schrödinger equation to be solved approximately for a limited range of wave vectors in the proximity of a high-symmetry point. For simplicity and clarity, the basic principles of **k**·**p** theory are initially demonstrated through the derivation of a single-band model for the lowest conduction band, leading to the well-known effective mass model. Generally, the technique involves selecting a suitable set of basis vectors to construct the crystal eigenstates, from which the matrix form of the Hamiltonian is

obtained. The Hamiltonian is then diagonalised to yield the \mathbf{k} -dependent eigenvalues and eigenstates. These aspects of the theory are demonstrated through derivation of the Luttinger-Kohn model in section 4.2.2. Here, the three highest, spin-degenerate valence bands are explicitly considered, whilst the impact of spin-orbit coupling and energetically remote bands (not belonging to the basis set) are also included. This is extended in section 4.2.3 by including coupling with the lowest conduction band using the Kane model, before exploring the effects of strain using the Bir-Pikus Hamiltonian.

4.2.1 The Single Band Model

The derivation of all k-p models begin by considering the movement of a single-electron in a periodic potential, $V(\mathbf{r})$, which reflects the periodicity of the crystal lattice. If spin-orbit coupling effects are initially neglected for simplicity, the single-electron wavefunction, $\psi(\mathbf{r})$, will satisfy the Schrödinger equation,

$$\left[\frac{\mathbf{p}^2}{2m_0} + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (4.1)$$

where $\mathbf{p} = -i\hbar\nabla$ is the momentum operator, and m_0 and E are the electron rest mass and energy, respectively. Since the crystal potential is invariant under translation by a Bravais lattice vector \mathbf{R} , (i.e. $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{R}$), the electron wavefunction must also possess this translational invariance. As such, the general solution of equation 4.1 may be expressed as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \quad (4.2)$$

where $u_{n\mathbf{k}}(\mathbf{r})$ is a periodic function characterised by

$$u_{n\mathbf{k}}(\mathbf{r}) \equiv u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}). \quad (4.3)$$

Here n denotes the band index, and \mathbf{k} is a wavevector within the first Brillouin zone. This is known as the Bloch theorem.

The Schrödinger equation may then be written in terms of the cell periodic function $u_{n\mathbf{k}}(\mathbf{r})$ by substituting equation 4.2 into 4.1, giving

$$\left[-\frac{\hbar^2}{2m_0} \left(\nabla^2 + 2i\mathbf{k} \cdot \nabla + k^2 \right) + V(\mathbf{r}) \right] u_{n\mathbf{k}}(\mathbf{r}) = E_n(\mathbf{k}) u_{n\mathbf{k}}(\mathbf{r}), \quad (4.4)$$

which simplifies to

$$\left[\underbrace{\frac{\mathbf{p}^2}{2m_0} + V(r)}_{H_0} + \underbrace{\frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p} + \frac{\hbar^2 k^2}{2m_0}}_{H_1} \right] u_{n\mathbf{k}}(\mathbf{r}) = E_n(\mathbf{k}) u_{n\mathbf{k}}(\mathbf{r}) \quad (4.5)$$

by using $\nabla = -\mathbf{p}/i\hbar$. The Hamiltonian for the system is then the sum of the unperturbed Hamiltonian, denoted H_0 , and a term H_1 stemming from the interaction with the crystal potential.

To understand the fundamental concepts, the simple case of a single CB close to the Γ -point is considered, where n denotes the conduction band and m denotes the states belonging to energetically remote bands. Assuming that the CB state is known at Γ and H_1 may be treated as a perturbation to H_0 in the vicinity of $\mathbf{k} = 0$, the dispersion about $\mathbf{k} = 0$ may be approximated to second-order using non-degenerate, time-independent perturbation theory as [140]

$$E_n(\mathbf{k}) \approx E_n(0) + \frac{\hbar^2 k^2}{2m_0} + \underbrace{\frac{\hbar}{m_0} \mathbf{k} \cdot \langle u_{n0} | \mathbf{p} | u_{n0} \rangle}_{=0} + \frac{\hbar^2}{m_0^2} \sum_{m \neq n} \frac{|\langle u_{m0} | \mathbf{k} \cdot \mathbf{p} | u_{n0} \rangle|^2}{E_{n0} - E_{m0}}, \quad (4.6)$$

where

$$\mathbf{p}_{mn} \equiv \langle u_{m0} | \mathbf{p} | u_{n0} \rangle = -i\hbar \int_{\Omega} u_{m0}^*(\mathbf{r}) \nabla u_{n0}(\mathbf{r}) d(\mathbf{r}) \quad (4.7)$$

and the third term on the RHS is zero due to symmetry. Therefore, the band dispersion differs from the free electron dispersion only through the final term on the RHS, which stems from interactions with remote band states. By considering the components of \mathbf{k} explicitly, Eqn. 4.6 may be rewritten as

$$E_n(\mathbf{k}) \approx E_n(0) + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{\alpha, \beta} k_{\alpha} k_{\beta} \sum_{m \neq n} \frac{p_{nm}^{(\alpha)} p_{mn}^{(\beta)}}{E_n(0) - E_m(0)} \quad (4.8)$$

where $\alpha, \beta = x, y, z$ such that $p_{nm}^{(\alpha)}$ is defined as the component of the matrix element in α . By grouping the second and third terms in Eqn. 4.8, the energies for the perturbed system may be written in a form similar to the free electron model:

$$E_n(\mathbf{k}) \approx E_n(0) + \frac{\hbar^2}{2} \sum_{\alpha, \beta} k_{\alpha} k_{\beta} D_{\alpha, \beta}, \quad (4.9)$$

where $D_{\alpha,\beta}$ is

$$D_{\alpha,\beta} = \frac{1}{m_0} \left(\delta_{\alpha,\beta} + \frac{2}{m_0} \sum_{m \neq n} \frac{p_{nm}^{(\alpha)} p_{nm}^{(\beta)}}{E_n(0) - E_m(0)} \right). \quad (4.10)$$

As such, $D_{\alpha,\beta}$ can therefore be thought of as a tensor describing a direction dependent electron *effective mass*, m_n^* , through

$$D_{\alpha,\beta} \equiv \left[\frac{1}{m_n^*} \right]_{\alpha,\beta} \equiv \frac{\partial^2 E_n(\mathbf{k})}{\partial k_\alpha \partial k_\beta}. \quad (4.11)$$

The components of the tensor are then adjustable parameters in the model, which can be obtained through cyclotron resonance measurements or calculated from atomistic band-structure modelling techniques such as density functional theory (DFT) [141].

It should be noted that the effective mass tensor elements are solely determined by the momentum matrix elements $p_{nn'}^{\alpha,\beta}$ which are weighted by the proximity of band n' to band n . It can be shown that, in the vicinity of $\mathbf{k} = 0$, the CB closest to the band edge shows spherically symmetric S-like character, with states denoted $|S\rangle$, whilst the next highest CB and the VB show predominantly P-like character, with triply-degenerate states of the form $|X, Y, Z\rangle_c$ and $|X, Y, Z\rangle_v$, respectively. By considering only the influence of these bands, since they are closest, the only non-zero momentum matrix elements are then

$$\langle S | p_x | X \rangle_c \equiv \langle S | p_y | Y \rangle_c \equiv \langle S | p_z | Z \rangle_c, \quad (4.12)$$

$$\langle S | p_x | X \rangle_v \equiv \langle S | p_y | Y \rangle_v \equiv \langle S | p_z | Z \rangle_v, \quad (4.13)$$

such that the CB is isotropic, with effective mass [142]

$$\frac{1}{m_n^*} = \frac{1}{m_0} + \frac{2P^2}{\hbar^2 E_0} + \frac{2P'^2}{\hbar^2 E_0}. \quad (4.14)$$

The constants P^2 and P'^2 have been defined as $\hbar^2/m_0^2 |\langle S | p_x | X \rangle_v|^2$, and $\hbar^2/m_0^2 |\langle S | p_x | X \rangle_c|^2$, respectively. In general P' is small compared to P and often neglected from the analysis [142]. In the single band model, the dispersion of the CB closest to the band edge about Γ in a zinc blende structure is therefore

$$E_n(k) \approx E_n(0) + \frac{\hbar^2 k^2}{2m_n^*}. \quad (4.15)$$

A similar derivation can also be utilised for the highest VB states.

4.2.2 The 6-band Luttinger-Kohn Model

When modelling the VB, the energetic proximity of the LH, HH and SO bands mean it is no longer appropriate to consider interactions as perturbations. As such, to accurately model the VB structure, it is necessary to consider these interactions explicitly using a 6-band model. This method was first devised by Luttinger and Kohn in their 1955 paper [143], building upon the 3-band model of Dresselhaus, Kip and Kittel [144]. The key difference between these models was the inclusion of the SO interaction in the Hamiltonian, which allowed the degeneracy lifting of the SO band to be correctly reproduced.

In order to accurately reproduce the bandstructure, the effect of remote band interactions must be included in the model. This is nominally achieved through Löwdin's method of renormalisation [145]. The process involves dividing eigenstates of the system into two classes denoted A and B. Class A consists of states which interact strongly with each other and weakly with states in class B; these are the basis states. Class B states then refer to all other, energetically remote states. The interactions between class A and B states are then removed using an iterative process, resulting in renormalised interactions between the Class A states.

The Hamiltonian describing interactions between the Class A states are obtained to second order using perturbation theory for degenerate states as [76]

$$H_{nm}^{(A)} = \left[E_n(0) + \frac{\hbar^2 k^2}{2m_0} \right] \delta_{nm} + \underbrace{H'_{nm}(\mathbf{k}) + \sum_{j \in A \neq n} \frac{H'_{nj}(\mathbf{k})H'_{jm}(\mathbf{k})}{E_n(0) - E_j(0)}}_{=0}. \quad (4.16)$$

Here $H' = \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p}$ such that $H'_{nm} = \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p}_{nm}$. $E_n(0)$ is the energy of the basis state labelled $|u_{n0}\rangle$. Since class A states all have p-like symmetry, $H'_{nm} = 0$ for all n, m . By adding the perturbative corrections of the remote bands from Löwdin's theory, the renormalised Hamiltonian is [145]

$$\begin{aligned} H_{nm} &= H_{nm}^{(A)} + \sum_{j \in B} \frac{H'_{nj}(\mathbf{k})H'_{jm}(\mathbf{k})}{E_n(0) - E_j(0)} \\ &= \left[E_n(0) + \frac{\hbar^2 k^2}{2m_0} \right] \delta_{nm} + \frac{\hbar^2}{m_0^2} \sum_{\alpha, \beta} k_\alpha k_\beta \sum_{j \in B} \frac{p_{nj}^{(\alpha)} p_{jm}^{(\beta)}}{E_n(0) - E_j(0)} \end{aligned} \quad (4.17)$$

where j represents class B states. It can be seen that the only non-trivial contributions to the Hamiltonian are the interactions between class A and B states. It is found that

only three distinct types of interaction produce non-zero matrix elements, these are [146]

$$\begin{aligned}
A &= \frac{\hbar^2}{2m_0} \left(1 + \frac{2}{m_0} \sum_{j \in B} \frac{p_{\alpha j}^{(\alpha)} p_{j\alpha}^{(\alpha)}}{E_A - E_j(0)} \right), \\
B &= \frac{\hbar^2}{2m_0} \left(1 + \frac{2}{m_0} \sum_{j \in B} \frac{p_{\alpha j}^{(\beta)} p_{j\alpha}^{(\beta)}}{E_A - E_j(0)} \right), \quad \text{for } \alpha \neq \beta \\
C &= \frac{\hbar^2}{m_0^2} \sum_{j \in B} \frac{p_{\alpha j}^{(\alpha)} p_{j\beta}^{(\beta)} + p_{\alpha j}^{(\beta)} p_{j\beta}^{(\alpha)}}{E_A - E_j(0)}, \quad \text{for } \alpha \neq \beta.
\end{aligned} \tag{4.18}$$

The effective mass and non-parabolicity of the LH, HH and SO bands in the vicinity of the Γ -point are therefore determined solely by these interaction terms. Linear combinations of these terms are used to define the more commonly utilised Luttinger parameters

$$\begin{aligned}
\gamma_1 &= -\frac{2m_0}{3\hbar^2} (A + 2B), \\
\gamma_2 &= -\frac{m_0}{3\hbar^2} (A - B), \\
\gamma_3 &= -\frac{m_0}{3\hbar_0} C.
\end{aligned} \tag{4.19}$$

These parameters act as fitting parameters or can be obtained experimentally from cyclotron resonance measurements [147].

The SO interaction is treated as a perturbation to Eqn. 4.5, with the SO Hamiltonian is given by

$$H_{SO} = \frac{\hbar}{4m_0^2 c^2} [(\boldsymbol{\sigma} \times \nabla V) \cdot \mathbf{p} + \underbrace{(\boldsymbol{\sigma} \times \nabla V) \cdot \mathbf{k}}_{=0}] = \frac{\hbar}{4m_0^2 c^2} (\nabla V \times \mathbf{p}) \cdot \boldsymbol{\sigma}, \tag{4.20}$$

where the rightmost term is neglected from the analysis since the momentum of the electron in the orbital, where the spin-orbit interaction is strongest, is significantly greater than the crystal momentum, ($\hbar \mathbf{k} \ll \mathbf{p}$) [148]. Here $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli spin matrices [76]

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.21}$$

which act on the $|\uparrow\rangle, |\downarrow\rangle$ eigenspinors as

$$\sigma_x |\uparrow\rangle = |\downarrow\rangle, \quad \sigma_x |\downarrow\rangle = |\uparrow\rangle, \quad (4.22)$$

$$\sigma_y |\uparrow\rangle = i |\downarrow\rangle, \quad \sigma_y |\downarrow\rangle = -i |\uparrow\rangle, \quad (4.23)$$

$$\sigma_z |\uparrow\rangle = |\uparrow\rangle, \quad \sigma_z |\downarrow\rangle = -|\downarrow\rangle. \quad (4.24)$$

Knowing that the VB maxima is p-like in nature, a good choice of basis states are those reflecting the symmetries of the triply degenerate p-orbitals, $|X\rangle, |Y\rangle, |Z\rangle$. To account for the two distinct spin-projections, the full set in the ‘LS’ basis is then $|X \uparrow\rangle, |Y \uparrow\rangle, |Z \uparrow\rangle, |X \downarrow\rangle, |Y \downarrow\rangle, |Z \downarrow\rangle$, where \mathbf{L} is the orbital angular momentum as before, and \mathbf{S} is the spin. In this basis, the total Hamiltonian is not diagonal for $\mathbf{k}=0$, and therefore cannot be solved exactly, it is therefore customary to perform a transformation to the JM_J basis, for which exact solutions can be found at the zone-centre [142]. Here $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the total angular momentum, and M_J is the associated z-axis projection. The new basis set reads [146], [149]

$$|u_{10}\rangle = \left| \frac{3}{2}, \frac{3}{2} \right\rangle = -\frac{1}{\sqrt{2}} |(X + iY) \uparrow\rangle, \quad (4.25)$$

$$|u_{20}\rangle = \left| \frac{3}{2}, \frac{1}{2} \right\rangle = -\frac{1}{\sqrt{6}} |(X + iY) \downarrow\rangle + \sqrt{\frac{2}{3}} |Z \uparrow\rangle, \quad (4.26)$$

$$|u_{30}\rangle = \left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{6}} |(X - iY) \uparrow\rangle + \sqrt{\frac{2}{3}} |Z \downarrow\rangle, \quad (4.27)$$

$$|u_{40}\rangle = \left| \frac{3}{2}, -\frac{3}{2} \right\rangle = \frac{1}{\sqrt{2}} |(X - iY) \downarrow\rangle, \quad (4.28)$$

$$|u_{50}\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X + iY) \downarrow\rangle + \frac{1}{\sqrt{3}} |Z \uparrow\rangle, \quad (4.29)$$

$$|u_{60}\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X - iY) \uparrow\rangle - \frac{1}{\sqrt{3}} |Z \downarrow\rangle. \quad (4.30)$$

Due to the symmetry of p-like basis states, the only non-zero matrix elements occur between states of the same spin and are cyclical permutations of the form [142]

$$\Delta_{so} \equiv \frac{3i\hbar}{4m_0^2c^2} \langle X \uparrow | (\nabla V \times \mathbf{p})_y | Z \uparrow \rangle. \quad (4.31)$$

Utilising this fact, it can be shown that H_{SO} in JM_J basis is a diagonal Hamiltonian given by [142]

$$H_{SO} = \frac{\Delta_{so}}{3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{pmatrix}, \quad (4.32)$$

illustrating that the inclusion of the spin-orbit interaction splits the sixfold degeneracy at the Γ -point into a $|3/2, M_J\rangle$ quadruplet and a $|1/2, M_J\rangle$ doublet. The quadruplet forms the doubly spin-degenerate LH and HH bands which are separated from the spin-degenerate spin-orbit split off band by an energy Δ_{so} .

By adding H_{SO} to Eqn. 4.17 and inserting the Luttinger parameters in place of the interaction terms, the total 6-band Luttinger-Kohn Hamiltonian in the JM_J basis is [146]

$$H_{LK} = \begin{pmatrix} E_{HH}(\mathbf{k}) & \sqrt{2}S(\mathbf{k}) & -S(\mathbf{k}) & 0 & -R(\mathbf{k}) & -\sqrt{2}R(\mathbf{k}) \\ & E_{LH}(\mathbf{k}) & Q(\mathbf{k}) & R(\mathbf{k}) & 0 & \sqrt{3}S(\mathbf{k}) \\ & & E_{SO}(\mathbf{k}) & -\sqrt{2}R(\mathbf{k}) & -\sqrt{3}S^*(\mathbf{k}) & 0 \\ & & & E_{HH}(\mathbf{k}) & \sqrt{2}S^*(\mathbf{k}) & -S^*(\mathbf{k}) \\ & & & & E_{LH}(\mathbf{k}) & Q(\mathbf{k}) \\ & & & & & E_{SO}(\mathbf{k}) \end{pmatrix} \quad (4.33)$$

where the \mathbf{k} dependent functions are

$$E_{HH}(\mathbf{k}) = -\frac{\hbar^2}{2m_0} [(\gamma_1 + \gamma_2) k_{\parallel}^2 + (\gamma_1 - 2\gamma_2) k_z^2], \quad (4.34)$$

$$E_{LH}(\mathbf{k}) = -\frac{\hbar^2}{2m_0} [(\gamma_1 - \gamma_2) k_{\parallel}^2 + (\gamma_1 + 2\gamma_2) k_z^2], \quad (4.35)$$

$$E_{SO}(\mathbf{k}) = -\Delta_{so} - \frac{\hbar^2}{2m_0} \gamma_1 (k_{\parallel}^2 + k_z^2), \quad (4.36)$$

$$Q(\mathbf{k}) = -\frac{\hbar^2}{\sqrt{2}m_0} \gamma_2 k_{\parallel}^2 + \frac{\sqrt{2}\hbar^2}{m_0} \gamma_2 k_z^2, \quad (4.37)$$

$$R(\mathbf{k}) = \frac{\sqrt{3}}{2} \frac{\hbar^2}{m_0} [\gamma_{av} (k_x - ik_y)^2 - \mu (k_x + ik_y)^2], \quad (4.38)$$

$$S(\mathbf{k}) = \sqrt{\frac{3}{2}} \frac{\hbar^2}{m_0} \gamma_3 (k_x - ik_y) k_z. \quad (4.39)$$

Since the Hamiltonian is Hermitian, the lower elements are not explicitly depicted in Eqn.

4.33. Here $k_{\parallel}^2 = k_x^2 + k_y^2$, and γ_{av} and μ have been defined for concision, taking the values $\frac{1}{2}(\gamma_2 + \gamma_3)$ and $\frac{1}{2}(\gamma_3 - \gamma_2)$, respectively [149].

From inspection of the Hamiltonian, a number of deductions can be made about the resultant bandstructure, Firstly, by examining the 3×3 off-diagonal block, it is found that the elements on the diagonal are zero, illustrating that there is no coupling between states differing only in spin. It is also evident that anisotropy in the band non-parabolicity stems solely from the right-most term in $R(\mathbf{k})$. For typical III-V and group IV materials γ_2 and γ_3 take nominally similar values. As such, μ can be set to zero in the so-called ‘axial approximation’, which greatly reduces the region of the Brillouin zone that is to be calculated explicitly, whilst simultaneously providing a good approximation to the full bandstructure. This technique is implemented during any calculations involving QW based-devices in this study.

4.2.3 The 8-band Model

The 6-band model derived above is able to accurately model VB non-parabolicities in the vicinity zone centre by accounting for interactions with the lowest CB and more energetically remote states through the experimentally determined Luttinger parameters. This is particularly useful for investigating IVBA transitions. However, for modelling interband transitions it is necessary to accurately model the non-parabolicities of both the CB and VB. To do this, the basis set is expanded such that coupling between the CB and VB is included explicitly. The 8-band basis states are then [150]

$$|u_{10}\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle = |S \uparrow\rangle, \quad (4.40)$$

$$|u_{20}\rangle = \left| \frac{3}{2}, \frac{3}{2} \right\rangle = -\frac{1}{\sqrt{2}} |(X + iY) \uparrow\rangle, \quad (4.41)$$

$$|u_{30}\rangle = \left| \frac{3}{2}, \frac{1}{2} \right\rangle = -\frac{1}{\sqrt{6}} |(X + iY) \downarrow\rangle + \sqrt{\frac{2}{3}} |Z \uparrow\rangle, \quad (4.42)$$

$$|u_{40}\rangle = \left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{6}} |(X - iY) \uparrow\rangle + \sqrt{\frac{2}{3}} |Z \downarrow\rangle, \quad (4.43)$$

$$|u_{50}\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = -|S \downarrow\rangle, \quad (4.44)$$

$$|u_{60}\rangle = \left| \frac{3}{2}, -\frac{3}{2} \right\rangle = \frac{1}{\sqrt{2}} |(X - iY) \downarrow\rangle, \quad (4.45)$$

$$(4.46)$$

$$|u_{70}\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X + iY) \downarrow\rangle + \frac{1}{\sqrt{3}} |Z \uparrow\rangle, \quad (4.47)$$

$$|u_{80}\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X - iY) \uparrow\rangle - \frac{1}{\sqrt{3}} |Z \downarrow\rangle, \quad (4.48)$$

where $|u_{10}\rangle$ and $|u_{20}\rangle$ are the s-like states associated with the doubly spin-degenerate CB. By explicitly including these states, non-zero momentum matrix elements arise between s-like and p-like states with equivalent spin in class A, which are used to define the Kane momentum matrix element,

$$P \equiv -i \frac{\hbar}{m_0} \langle S | p_x | X \rangle \equiv -i \frac{\hbar}{m_0} \langle S | p_y | Y \rangle \equiv -i \frac{\hbar}{m_0} \langle S | p_z | Z \rangle. \quad (4.49)$$

Derivation for the renormalised Hamiltonian proceeds as per the 6-band model. However, as a result of the non-zero momentum matrix elements, the two terms in Eqn. 4.16 are now non-zero. The terms that are linear in $|\mathbf{k}|$ describe the CB-VB interband coupling, while the terms that are quadratic in $|\mathbf{k}|$ act as corrections to Eqn. 4.17. This occurs because the s-like states are now class A such that their effect on the bandstructure is calculated explicitly, and are thus removed from the term describing the effect of class B states. As such, the Luttinger parameters from the 6-band model, defined from the momentum matrix elements of interactions between class A and class B states, must be adjusted accordingly. In this work, the most commonly implemented correction is adopted, such that the parameters are transformed by [151]

$$\gamma_1 \rightarrow \gamma_1 - \frac{E_P}{3E_g}, \quad (4.50)$$

$$\gamma_2 \rightarrow \gamma_2 - \frac{E_P}{6E_g}, \quad (4.51)$$

$$\gamma_3 \rightarrow \gamma_3 - \frac{E_P}{6E_g}, \quad (4.52)$$

where $E_P = 2m_0|P|^2/\hbar^2$ is the Kane parameter, and E_g is the material bandgap. Similarly, the effective electron mass must also be adjusted. The resultant inverse effective mass is denoted s_C and defined as [152]

$$s_C = \frac{1}{m_c^*} - \frac{E_P}{3} \left(\frac{2}{E_g} + \frac{1}{E_g + \Delta_{so}} \right). \quad (4.53)$$

The resultant, fully-renormalised Hamiltonian is then given by [149], [151]

$$H(\mathbf{k}) = \begin{pmatrix} E_{CB}(\mathbf{k}) & -\sqrt{3}T_+(\mathbf{k}) & \sqrt{2}U(\mathbf{k}) & -U(\mathbf{k}) & 0 & 0 & -T_-(\mathbf{k}) & -\sqrt{2}T_-(\mathbf{k}) \\ & E_{HH}(\mathbf{k}) & \sqrt{2}S(\mathbf{k}) & -S(\mathbf{k}) & 0 & 0 & -R(\mathbf{k}) & -\sqrt{2}R(\mathbf{k}) \\ & & E_{LH}(\mathbf{k}) & Q(\mathbf{k}) & T_+(\mathbf{k}) & R(\mathbf{k}) & 0 & \sqrt{3}S(\mathbf{k}) \\ & & & E_{SO}(\mathbf{k}) & \sqrt{2}T_+(\mathbf{k}) & \sqrt{2}R(\mathbf{k}) & -\sqrt{3}S(\mathbf{k}) & 0 \\ & & & & E_{CB}(\mathbf{k}) & -\sqrt{3}T_-(\mathbf{k}) & \sqrt{2}U(\mathbf{k}) & -U(\mathbf{k}) \\ & & & & & E_{HH}(\mathbf{k}) & \sqrt{2}S^*(\mathbf{k}) & -S^*(\mathbf{k}) \\ & & & & & & E_{LH}(\mathbf{k}) & Q(\mathbf{k}) \\ & & & & & & & E_{SO}(\mathbf{k}) \end{pmatrix}, \quad (4.54)$$

where the matrix elements are

$$\begin{aligned} E_{CB}(\mathbf{k}) &= E_{CB}(0) + \frac{\hbar^2}{2m_0} s_C (k_{\parallel}^2 + k_z^2), \\ E_{HH}(\mathbf{k}) &= E_{HH}(0) - \frac{\hbar^2}{2m_0} \left[(\gamma_1 + \gamma_2) k_{\parallel}^2 + (\gamma_1 - 2\gamma_2) k_z^2 \right], \\ E_{LH}(\mathbf{k}) &= E_{LH}(0) - \frac{\hbar^2}{2m_0} \left[(\gamma_1 - \gamma_2) k_{\parallel}^2 + (\gamma_1 + 2\gamma_2) k_z^2 \right], \\ E_{SO}(\mathbf{k}) &= E_{SO}(0) - \frac{\hbar^2}{2m_0} \gamma_1 (k_{\parallel}^2 + k_z^2), \\ T_{\pm}(\mathbf{k}) &= \frac{1}{\sqrt{6}} P (k_x \pm ik_y), \\ U(\mathbf{k}) &= \frac{1}{\sqrt{3}} P k_z, \\ S(\mathbf{k}) &= \sqrt{\frac{3}{2}} \frac{\hbar^2}{m_0} \gamma_3 k_z (k_x - ik_y), \\ R(\mathbf{k}) &= \frac{\sqrt{3}}{2} \frac{\hbar^2}{m_0} \left[\gamma_{av} (k_x - ik_y)^2 - \mu (k_x - ik_y)^2 \right], \\ Q(\mathbf{k}) &= -\frac{1}{\sqrt{2}} \frac{\hbar^2}{m_0} \gamma_2 k_{\parallel}^2 + \sqrt{2} \frac{\hbar^2}{m_0} \gamma_2 k_z^2. \end{aligned} \quad (4.55)$$

It can be seen that the matrix elements $Q(\mathbf{k})$, $S(\mathbf{k})$, $R(\mathbf{k})$ describing coupling between the VB states are equivalent to those in the 6-band model, whilst terms $T_{\pm}(\mathbf{k})$ and $U(\mathbf{k})$ have been introduced to describe coupling between CB and VB states. As per the LK model, the diagonal entries of the upper right off-diagonal block illustrate that there is zero coupling between states differing only in spin. Two additional zero entries arise due to the orthogonality of the two spin states, which ensures that the momentum matrix elements between s-like and p-like states of opposing spins is zero. This results in weaker HH-CB coupling when compared to that of the LH-CB and SO-CB.

4.2.4 Pseudomorphic Strain Corrections

Even in bulk structures, lattice mismatch between epitaxial layers can result in levels of strain which can influence the bandstructure in an appreciable manner. As such, a correction must be made to the unstrained Hamiltonians derived previously.

Any strain can be decomposed into a ‘hydrostatic’ or ‘shear’ strain. Hydrostatic strains refer to those where the symmetries of the crystal lattice are preserved, and there is a change only in volume. This results in a shift in the band edge energies, without lifting the VB degeneracy. When the strain is tensile in nature, the bandgap narrows, whereas a compressive strain widens the bandgap. In contrast, a shear strain refers to those which act to break the lattice symmetries whilst preserving the volume. The symmetry breaking results in VB degeneracy lifting.

For the case of biaxially-strained layers, the strain is included in the model by way of a corrective perturbative Hamiltonian as described initially by Bir and Pikus [153], given for the 8-band model as [154]

$$H_{BP} = \begin{pmatrix} O_{BP} & -\sqrt{3}T_{BP} & \sqrt{2}U_{BP} & -U_{BP} & 0 & 0 & 0 & 0 \\ & P_{BP} + Q_{BP} & \sqrt{2}S_{BP} & -S_{BP} & 0 & 0 & 0 & 0 \\ & & P_{BP} - Q_{BP} & \sqrt{2}Q_{BP} & 0 & 0 & 0 & 0 \\ & & & P_{BP} & 0 & 0 & 0 & 0 \\ & & & & O_{BP} & -\sqrt{3}T_{BP}^* & \sqrt{2}U_{BP} & -U_{BP} \\ & & & & & P_{BP} + Q_{BP} & \sqrt{2}S_{BP}^* & -S_{BP}^* \\ & & & & & & P_{BP} - Q_{BP} & \sqrt{2}Q_{BP} \\ & & & & & & & P_{BP} \end{pmatrix}, \quad (4.56)$$

with

$$O_{BP} = a_c (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}), \quad (4.57)$$

$$P_{BP} = -a_v (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}), \quad (4.58)$$

$$Q_{BP} = \frac{b}{2} (\varepsilon_{xx} + \varepsilon_{yy} - 2\varepsilon_{zz}), \quad (4.59)$$

$$R_{BP} = -\frac{\sqrt{3}}{2b_v} (\varepsilon_{xx} - \varepsilon_{yy}) + id_v \varepsilon_{xy}, \quad (4.60)$$

$$S_{BP} = -\frac{d_v}{\sqrt{2}} (\varepsilon_{zx} - i\varepsilon_{yz}), \quad (4.61)$$

$$(4.62)$$

$$T_{BP} = \frac{P}{\sqrt{6}} [(\varepsilon_{xx} + i\varepsilon_{yx})k_x + (\varepsilon_{xy} + i\varepsilon_{yy})k_y + (\varepsilon_{xz} + i\varepsilon_{yz})k_z], \quad (4.63)$$

$$U_{BP} = \frac{Pk_z}{3} (\varepsilon_{zx}k_x + \varepsilon_{zy}k_y + \varepsilon_{zz}k_z), \quad (4.64)$$

where [146]

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a_0 - a}{a}, \quad (4.65)$$

$$\varepsilon_{zz} = -\frac{2C_{12}}{C_{11}}\varepsilon_{xx}, \quad (4.66)$$

$$\varepsilon_{xy(yx)} = \varepsilon_{xz(zx)} = \varepsilon_{yz(zy)} = 0. \quad (4.67)$$

Here a_0 , a and C_{12} , C_{11} are the unstrained lattice constant, strained in-plane lattice constant and elastic stiffness constants, respectively. O_{BP} and P_{BP} both describe hydrostatic strains, while Q_{BP} represents a shear strain.

T_{BP} and U_{BP} are historically taken to be zero, and hence most of the off diagonal terms are zero. The only surviving elements are O_{BP} , P_{BP} , and Q_{BP} . As such, the only corrections to be made to the unstrained Hamiltonian are given by [155]

$$E_{CB} \rightarrow E_{CB} + O_{BP}, \quad (4.68)$$

$$E_{HH} \rightarrow E_{HH} + P_{BP} + Q_{BP}, \quad (4.69)$$

$$E_{LH} \rightarrow E_{LH} + P_{BP} - Q_{BP}, \quad (4.70)$$

$$E_{SO} \rightarrow E_{SO} + P_{BP}, \quad (4.71)$$

$$Q \rightarrow Q + \sqrt{2}Q_{BP}. \quad (4.72)$$

It should be noted that Eqns. 4.69 - 4.72 are also valid for the 6-band model.

The bulk 6- and 8-band models, with strain corrections, used in this work were written in Python as functions to be used with the software *kpymod*, developed by Dr. Dominic Duffy. A link to the file containing said functions is provided in appendix A.

4.3 Parameterisation

Due to the semi-empirical nature of the k-p model, to calculate the pressure relations, bandstructure and bulk gain in the subsequent sections, numerous parameters are required as inputs. Since concerted focus on group-IV materials such as GeSn and SiGeSn

is relatively recent, a range of material parameters are given in current literature. Due to their importance in electronics, the pure elemental parameters for Si and Ge are generally well-established, however α -Sn has historically been of less interest, and hence many parameters are not available in literature. Greater uncertainty exists when modelling group-IV binary or ternary alloys. Due to a lack of experimental data, many of the parameters for these multinary alloys can only be approximated through linear interpolation. Such an approximation can lead to significant inaccuracies should the non-linearities be strong. Of the binary materials, GeSn is the most widely studied in terms of non-linear parameters, however there is comparatively little information for SiGe, and even less for SiSn. Thus, as GeSn, and SiGeSn to a lesser extent, with low Sn contents are investigated, it is likely that the models will be suitably accurate.

In this work, parameters from a range of sources are collated. As far as possible, the parameters most commonly used in literature are utilised. However, where minor differences are found in more recent publications, these values are adopted preferentially. The values implemented in models detailed in this chapter are given in table 4.1.

4.3.1 Linear Interpolation

A semiconductor alloy with low levels of ordering can be treated under the virtual crystal approximation (VCA). Here the crystal is modelled as homogeneous medium of atoms with properties that are weighted averages of the constituent elements. In the absence of bowing, material parameters can be estimated by linear interpolation using Vegard's law [170]. For a group-IV multinary alloy of $\text{Si}_x\text{Ge}_{1-x-y}\text{Sn}_y$, the equation reads,

$$P_{\text{SiGeSn}} = xP_{\text{Si}} + (1 - x - y)P_{\text{Ge}} + yP_{\text{Sn}}, \quad (4.73)$$

where P is a given parameter, and x and y are the fractions of Si and Sn in the alloy, respectively. In the case of binary GeSn, x is simply set to 0, and a similar approach can be made for SiSn and SiGe.

In this study, all of the parameters listed in table 4.1 are linearly interpolated, with the exception of the Γ - and L-valley bandgaps. Due to the lack of currently available literature, the level of accuracy of this approach is unknown. However, compositional dependence has been illustrated for parameters such as the deformation potentials and stiffness constants in III-V materials [172], [173], making it plausible that such a relation may exist in group

Parameter	Symbol (units)	Si	Ge	Sn
Bandgap (Γ)	$E_g^\Gamma @ 0 \text{ K (eV)}$	4.185 ^{a)}	0.8981 ^{a)}	-0.41 ^{a)}
Varshni Parameter (Γ)	$\alpha^\Gamma \text{ (eV/K)}$	3.91e-4 ^{b)}	5.82e-4 ^{b)}	0
Varshni Parameter (Γ)	$\beta^\Gamma \text{ (K)}$	125.0 ^{b)}	296.0 ^{b)}	0
Bandgap (L)	$E_g^L @ 0 \text{ K (eV)}$	2.176 ^{c)}	0.744 ^{b)}	0.092 ^{a)}
Varshni Parameter (L)	$\alpha^L \text{ (eV/K)}$	4.744e-4 ^{b)}	4.561e-4 ^{b)}	0
Varshni Parameter (L)	$\beta^L \text{ (K)}$	235.0 ^{b)}	210.0 ^{b)}	0
Δ Bandgap (Δ)	$E_g^\Delta @ 0 \text{ K (eV)}$	1.1695 ^{d)}	0.931 ^{b)}	-
Varshni Parameter (Δ)	$\alpha^\Delta \text{ (eV/K)}$	4.73e-4 ^{d)}	4.774e-4 ^{b)}	-
Varshni Parameter (Δ)	$\beta^\Delta \text{ (K)}$	636.0 ^{d)}	235.0 ^{b)}	-
Spin-Orbit Splitting Energy	$\Delta_{\text{SO}} \text{ (eV)}$	0.044 ^{a)}	0.297 ^{a)}	0.8 ^{a)}
Valence Band Offset	$\Delta E_{\text{VBO}} \text{ (eV)}$	-1.27 ^{e)}	-0.69 ^{e)}	0 ^{f)}
Elastic Constant	$c_{11} \text{ (GPa)}$	165.77 ^{a)}	128.53 ^{a)}	69.00 ^{a)}
Elastic Constant	$c_{12} \text{ (GPa)}$	63.93 ^{a)}	48.26 ^{a)}	29.30 ^{a)}
Elastic Constant	$c_{44} \text{ (GPa)}$	79.62 ^{a)}	66.30 ^{a)}	36.20 ^{a)}
Deformation Potential (Γ)	$a_c^\Gamma \text{ (eV)}$	-10.06 ^{g)}	-8.24 ^{h)}	-6.00 ^{g)}
Deformation Potential (L)	$a_c^L \text{ (eV)}$	-0.66 ^{h)}	-1.54 ^{h)}	-2.14 ⁱ⁾
Deformation Potential (Δ)	$a_c^\Delta \text{ (eV)}$	14.0 ^{j)}	5.75 ^{k)}	-
VB Deformation Potential	$a_v \text{ (eV)}$	2.46 ^{g)}	1.24 ^{g)}	1.58 ⁱ⁾
Shear Deformation Potential (Γ)	$b^\Gamma \text{ (eV)}$	-2.1 ^{g)}	-2.9 ^{g)}	-2.7 ^{l)}
Shear Deformation Potential (L)	$b^L \text{ (eV)}$	18.0 ^{k)}	15.9 ^{k)}	-
Shear Deformation Potential (Δ)	$b^\Delta \text{ (eV)}$	-	9.75 ^{k)}	-
Pressure Coefficient (Γ)	$dE_g^\Gamma/dP \text{ (meV/kbar)}$	12.7 ^{m)}	14.6 ^{m)}	16.6 ^{m)}
Pressure Coefficient (L)	$dE_g^L/dP \text{ (meV/kbar)}$	4.2 ^{m)}	5.8 ^{m)}	5.9 ^{m)}
Pressure Coefficient (Δ)	$dE_g^\Delta/dP \text{ (meV/kbar)}$	-1.4 ^{m)}	-0.7 ^{m)}	-1 ^{m)}
Luttinger Parameter (6-band)	$\gamma_1 \text{ (eV)}$	4.22 ^{a)}	13.38 ^{a)}	-12 ⁱ⁾
Luttinger Parameter (6-band)	$\gamma_2 \text{ (eV)}$	0.39 ^{a)}	4.24 ^{a)}	-8.45 ⁱ⁾
Luttinger Parameter (6-band)	$\gamma_3 \text{ (eV)}$	1.44 ^{a)}	5.69 ^{a)}	-6.84 ⁱ⁾
Kane Matrix Element	$E_P \text{ (eV)}$	21.6 ^{a)}	26.3 ^{a)}	23.8 ⁿ⁾

TABLE 4.1: ref. ^{a)} [156], ref. ^{b)} [157], ref. ^{c)} [158], ref. ^{d)} [159], ref. ^{e)} [160], ref. ^{f)} [161], ref. ^{g)} [162], ref. ^{h)} [163], ref. ⁱ⁾ [164], ref. ^{j)} [165], ref. ^{k)} [166], ref. ^{l)} [167], ref. ^{m)} [168], ref. ⁿ⁾ [169]

IV alloys. This is suggested for GeSn in a 2017 arXiv paper in which bowing terms for the deformation potentials are extracted from an empirical pseudopotential model (EPM) [171]. Since similar terms are not found elsewhere in peer reviewed literature, these are not adopted in the model utilised in this body of work. As this study deals only with bulk layers, the small residual strains are such that any bowing in the elastic constants or deformation potentials is not likely to significantly alter results.

A non-linear compositional dependence of the Luttinger parameters in GeSn has also been reported in a small number of studies [174], [175]. Here, parameters for an 8-band k-p model were again extracted from EPMS and quadratic formulas were fitted to the second order to extract the compositional dependence of the parameters. Figure 4.4 depicts

the Luttinger parameters calculated through linear interpolation (solid) and using the empirical model of [174]. The minimum Sn composition is chosen as 8% as alloys below this concentration are not considered in these studies, and the upper limit is capped at 15% due to the range over which the empirical formula is valid. Relatively good agreement between both models can be observed over this range. However, the empirical expression is unable to replicate the Luttinger parameters for Ge which are widely reported and well-defined in the literature. As such, we opt to use the linearly interpolated values.

4.3.2 Bowing

Deviations from linear scaling with composition is typically accounted for via the inclusion of a bowing parameter. This will often be determined from experimental measurements but are occasionally extracted from EPM or density functional theory models. With the inclusion of bowing, Eqn. 4.73 then reads

$$P_{SiGeSn} = xP_{Si} + (1 - x - y)P_{Ge} + yP_{Sn} + x(1 - x - y)b_{SiGe} + y(1 - x - y)b_{GeSn} + xyb_{SiSn}, \quad (4.74)$$

where b is the bowing parameter for the subscripted alloy.

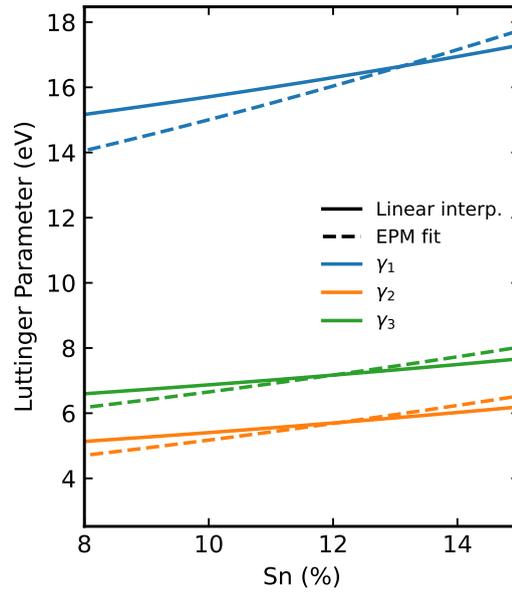


FIGURE 4.4: Luttinger parameters as a function of Sn, calculated using the empirical model devised in [171] (dashed), with standard linear interpolation utilised in this study (solid).

Since heightened temperature results in an increased lattice constant, the bandgap is inherently temperature dependent. The dependence can be approximated using numerous semi-empirical approaches such as the Päsler and Bose-Einstein models, but here the Varshni model is applied. This model was chosen because it is the most widely used, and thus has the most parameters available in literature. The equation is given by

$$E_g^i(T) = E_g^i(0) - \frac{\alpha T^2}{\beta + T}, \quad (4.75)$$

where i denotes the CB valley, and α and β are semi-empirical values typically determined through fitting of experimental data. This is used to calculate the temperature dependent bandgap of each group IV element, which are inserted into equation 4.74.

GeSn is the most widely researched group-IV binary and as such, numerous studies have attempted to quantify the bandgap bowing parameters. Recent literature values for b_{GeSn}^Γ lie between 1.94-2.92 eV [176]–[178], while b_{GeSn}^L is between 0.89-2.28 eV [177], [179]–[181], with most quoted at close to 1 eV. However, a number of groups have reported that the bowing parameter itself should possess both a temperature- and composition-dependence [177], [182], [183]. In this study, a temperature-dependent bowing term for GeSn is opted for, as this is heavily utilised in recent work. The terms for the Γ - and L -valley bandgaps are given by [184]

$$b_{GeSn}^\Gamma(T) = 2.55 - 4 \times 10^{-4}T \quad (4.76)$$

$$b_{GeSn}^L(T) = 0.89 - 7 \times 10^{-4}T. \quad (4.77)$$

These are in good agreement with other reported values whilst also including the temperature dependence.

For SiGe, the bowing parameters are taken to be $b_{SiGe}^\Gamma = 0.21$ eV and $b_{SiGe}^L = 0.335$ eV [185]. SiSn has been explored far less than the aforementioned binaries, and as such, the range of quoted bowing parameters has a large degree of uncertainty, spanning -21 [186] to 24 eV [187] for the Γ bandgap. Values of $b_{SiSn}^\Gamma = 2.124$ eV and $b_{SiSn}^L = 3.915$ eV from [185] are chosen, for consistency. Since the difference in electronegativity and atom size between Sn and Si is greater than that of Sn and Ge, it is expected that the bowing should be greater for SiSn than GeSn. This is reflected in the choice of parameters for the model. However, it should be noted that these large differences may qualify SiSn as a highly mismatched alloy (HMA) and the breakdown of the VCA [155]. As is the case

with other HMAs, such as dilute nitrides [188], this may mean a single bowing term may be inappropriate over the whole composition range. The bowing parameters for SiGe and SiSn are only utilised in modelling the SiGeSn cap layer. In this case, the low fractions of both Si and Sn mean that the relative contributions of these two bowing terms are low compared to the GeSn bowing, which further justifies the choice of parameters.

4.4 Evaluation of Model Accuracy

Multiple models of varying accuracy are available for calculating electronic bandstructure from which the material gain can be calculated. First principle calculations utilising DFT can be highly accurate but often require the use of computationally demanding hybrid-functionals to produce accurate bandgaps. In contrast, parabolic band models have a simple functional form which allows the bandstructure and gain spectra to be calculated with minimal computational cost. The k·p approach is semi-empirical in nature, striking a balance between both models in terms of accuracy and run-time. However, the inclusion of band interactions mean that gain must be calculated numerically by iterating over the Brillouin zone, rather than using the analytical solution afforded by parabolic bands. It is therefore essential to determine whether the 8-band k·p model outlined in Sec. 4.2.3 is most appropriate for calculation of bulk gain, given the required computational resources and run-time.

4.4.1 III-V Materials

The accuracy of the k·p bandstructure is assessed by way of comparison with a parabolic model and an EPM. GaAs is chosen for the comparison since it is a relatively wide gap material with well-known k·p parameters. This ensures that the elements of the Hamiltonian can be checked without complications stemming from large CB-VB interaction and uncertainty in material parameters.

Fig. 4.5 illustrates that all 3 models are able to accurately replicate the bandstructure for $|\mathbf{k}| < 0.5 \text{ nm}^{-1}$. Above this value, the parabolic model deviates significantly from the EPM model, which would result in distorted gain and IVBA spectra at higher levels of injection. In contrast, the k·p model is able to replicate the VB almost exactly below $|\mathbf{k}| = 1 \text{ nm}^{-1}$, whilst providing a better approximation of the CB.

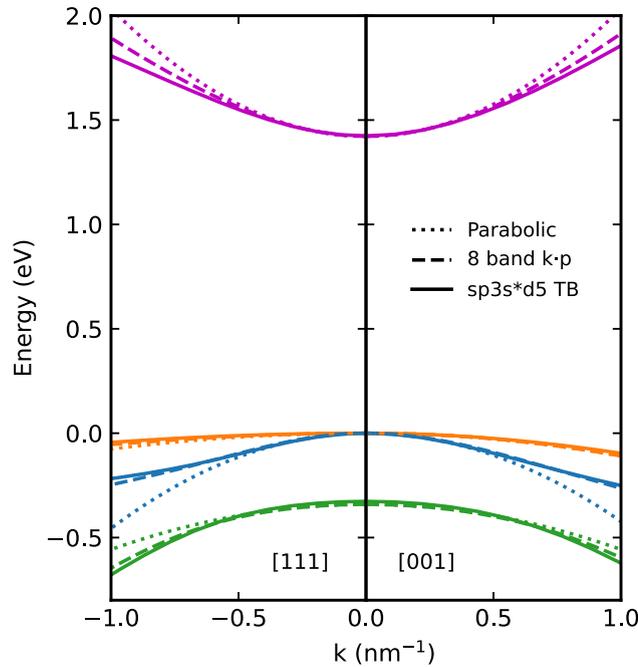


FIGURE 4.5: Bandstructure along the [001], and [111] directions in GaAs. The 8-band model is depicted by the dashed line, and compared to the parabolic model (dotted) and bandstructure calculated using an EPM, available on nanoHUB [118].

4.4.2 Group IV Materials

Fig 4.6 depicts the bandstructure of Ge using parabolic (dotted), 8-band k-p (dashed) and $sp^3s^*d^5$ tight-binding (solid) models. The tight-binding bandstructure was calculated using the *Band Structure Lab* online tool, available on nanoHUB [118] and is the most accurate model over extended ranges in k . It can be seen that the parabolic model is able to reproduce the bandstructure below $0.3\text{-}0.4\text{ nm}^{-1}$ along both directions, however significant deviations occur beyond this range. As the effect of band interactions are included, the k-p model is able to more accurately replicate the non-parabolicity at higher k . Since increased levels of carrier injection or thermal broadening will cause higher energy electron- (and lower energy hole-) states to be populated, the k-p method is most applicable in these circumstances.

4.5 Summary

In this chapter the (Si)GeSn material system was introduced and it's use in the development of truly direct-gap group IV lasers discussed, with an emphasis on electrically

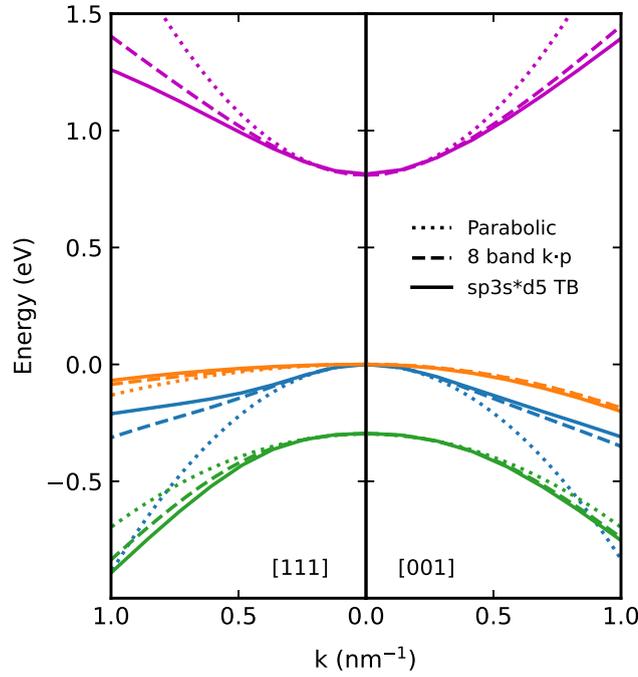


FIGURE 4.6: Bandstructure along the [001], and [111] directions in Ge. The 8-band model is depicted by the dashed line, and compared to the parabolic model (dotted) and bandstructure calculated using an EPM, available on nanoHUB [118].

injected devices. The theory behind the k-p method was introduced in a semi-conceptual manner, before outlining the strain-dependent 6- and 8-band models to be used in forthcoming chapters for gain and IVBA calculations. Parameters required for modelling group IV elements and multinary alloys are then collated through a comprehensive review of the available literature, before detailing how the obtained parameters are utilised within the bandstructure calculations. The accuracy of the model is then compared with more and less accurate approaches to test the robustness of the model.

Having established a number of appropriate models for calculating bulk bandstructure, we may now proceed with the analysis of carrier recombination processes in electrically injected, bulk group-IV laser heterostructures.

Chapter 5

Performance Limiting Mechanisms in Electrically Injected GeSn Lasers

5.1 Sample Information

The laser devices investigated during this body of work are FP bulk double GeSn/SiGeSn heterostructures, reported by the University of Arkansas in 2020 [21]. These structures consist of five layers, grown on 200 mm diameter, industry standard Si(100) wafers using chemical vapour deposition (CVD). The heterostructure design is depicted in Fig. 5.1, whilst table 5.1 gives detailed information on each layer, listed from bottom to top. These devices were reported in [21], and further growth information can be found in [189].

The wafer was processed into 100 μm , broad-ridge FP waveguide structures with cavity lengths of 1.7 mm, and topped with Cr/Au n and p contacts. Top (lateral) contacting is often deployed in laser devices grown on Si such that injected carriers can bypass thick layers of highly defective buffer material. Lasing was demonstrated up to a maximum of 100 K, with a threshold close to 3 kA/cm². The primary aim of this study is to understand why lasing is limited to such low temperatures in electrically injected devices, such that future designs can be adapted to improve performance. This is achieved by analysing the carrier recombination mechanisms using a combination of high hydrostatic pressure characterisation and k·p bandstructure modelling.

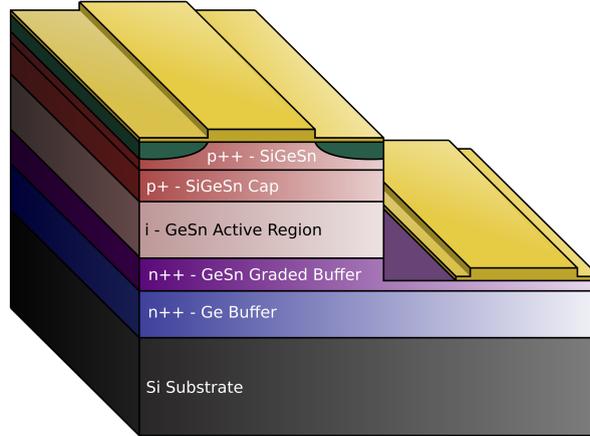


FIGURE 5.1: Schematic diagram on the top (lateral)-contacted FP ridge waveguide structures discussed in this chapter.

Layer	Material	Thickness (nm)	Doping (cm^{-3})	Strain (%)
Buffer	Ge	500	1×10^{19} (n-type)	0.1
Graded Buffer	$\text{GeSn}_{0.08-0.105}$	700	1×10^{19} (n-type)	-0.14
Active Region	$\text{GeSn}_{0.105}$	1000	-	-0.27
Cap	$\text{Si}_{0.03}\text{GeSn}_{0.08}$	170	1×10^{18} (p-type)	-
Ohmic Contact	$\text{Si}_{0.03}\text{GeSn}_{0.08}$	70	1×10^{19} (p-type)	-

TABLE 5.1

5.2 High Hydrostatic Pressure Measurements

As the maximum operating temperature for electrically pumped GeSn lasers is $\sim 100 - 140$ K, it is vital to determine the efficiency limiting processes such that future device design can be optimised to mitigate them. Analysing the threshold current as a function of hydrostatic pressure can provide key insight into the recombination characteristics, since each recombination pathway has a distinct pressure dependence that is tied intrinsically to the bandstructure. In this section these relations are used to quantify the contributions of each process, and stipulate the impact on higher temperature operation.

The devices were cleaved from an individual laser bar, with initial testing on the probe-station indicating two working devices, hereby referred to as samples A and B. Other devices were rendered defective, likely due to the quality of the Cr/Au contacts which showed visible peeling. Working devices were subsequently wire bonded and mounted onto individual laser headers. For each measurement run, a single device was loaded into the CuBe pressure cell and connected to the standard laser circuit depicted in section 3.3. As lasing was limited to temperatures below 100 K, the pressure cell was housed

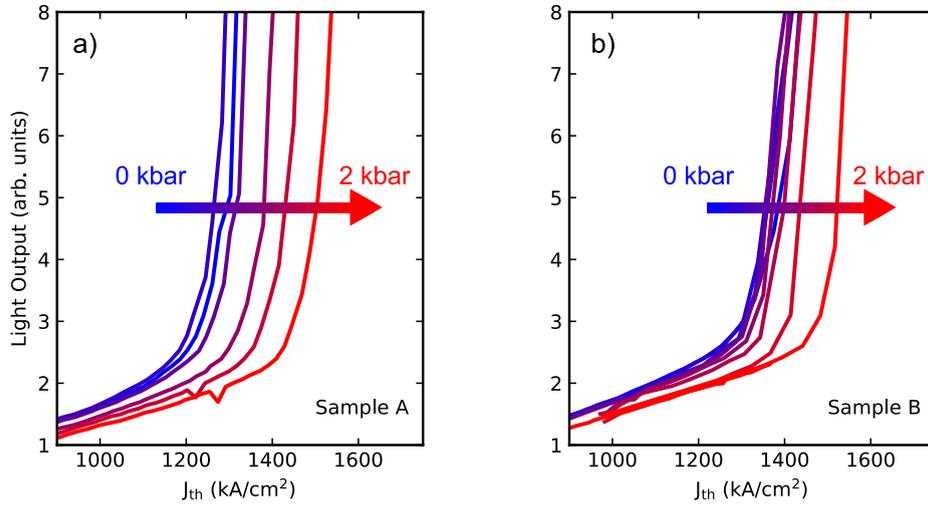


FIGURE 5.2: LI curves depicting the change in threshold current density with increasing pressure between 0-2.0 kbar at 85 K for samples a) A and b) B, respectively.

within a closed-cycle cryostat and pumped to a minimum achievable temperature of 85 K. The set-up was left for 1 hour to ensure the sample had equilibrated and the system temperature was stable. In order to minimise self-heating effects, square pulses of 500 ns were supplied at a frequency of 1 kHz. The pressure was then altered between 0-2 kbar. Since the cell utilised in this set-up is older than for RT measurements, 2 kbar was the highest achievable pressure, likely due to wear degrading the seal. Measurements of the light output as a function of current were then taken whilst the pressure was increased to, and decreased from, 2 kbar. This was done to ensure there was no irreversible device degradation as a result of the applied pressure. Low levels of light output paired with weakened spectrometer sensitivity at the predicted emission wavelength meant spectral measurements were not possible.

The LI curves obtained for a single device are illustrated in Fig. 5.2. Threshold was obtained using the method outlined in section 3.3. At atmospheric pressure, the devices illustrate threshold currents around 1.15 kA/cm^2 , in good agreement with those reported at 77 K in the original paper. Below threshold, the devices illustrate notable levels of spontaneous emission, evidenced by the gradient of the LI curve sub-threshold, likely stemming from large levels of carrier injection and the bulk nature of the devices. The threshold is seen to increase by $\sim 1.2\times$ between 0 and 2 kbar.

To understand the underlying cause of this shift, the expression for the threshold current density as a function of pressure is considered. This is dependent on the injection efficiency and the rate of carrier depletion from the Γ -valley. The threshold current as a function

of pressure can therefore be written in terms of the recombination processes as

$$J_{th}(P) = J_{th}^{def}(P) + J_{th}^{rad}(P) + J_{th}^{lk}(P), \quad (5.1)$$

where J_{th}^{def} , J_{th}^{rad} and J_{th}^{lk} are the pressure-dependent current densities of carriers recombining via defect-related, radiative, and leakage processes at threshold, respectively. At low temperatures, the thermally-activated Auger-Meitner processes are expected to be small and are therefore neglected from the analysis. The relations for radiative and defect-related recombination in bulk devices are [190]

$$J^{def}(P) \approx const. , \quad (5.2)$$

$$J^{rad}(P) \propto [E_g^\Gamma(P)]^{5/2} . \quad (5.3)$$

Here, leakage refers to any process by which injected carriers occupy states other than those in the active region Γ -valley, for example physical loss of carriers to states in neighbouring layers or electrons occupying indirect valley states in the active region. The latter is of utmost importance in these devices due to the proximity of the Γ - and L-band edges. Using a simple model, the pressure dependence of the leakage current density can be given by

$$J^{lk}(P) = J^{lk}(0) \exp\left(-\frac{dE_a}{dP} \frac{P}{k_B T}\right), \quad (5.4)$$

where the activation energy can be approximated as $E_a = E_L - E_{fc} \approx E_L - E_\Gamma$ [191]. Here, $E_{L/\Gamma}$ are the band minima of the respective valleys and E_{fc} is the conduction band quasi-Fermi energy, which is assumed to be close to the direct-valley minima at threshold. The movement of the different conduction band valley minima with pressure are given for bulk Si, Ge and Sn by the pressure coefficients dE_g^Γ/dP , dE_g^L/dP , and dE_g^Δ/dP listed in Tab. 4.1. Parameters for binary GeSn and ternary SiGeSn are then obtained through linear interpolation of the constituent elements. The resulting flat band edge diagram for the active region and cap material as a function of pressure is illustrated in Fig. 5.3. This shows that the pressure coefficients in the bulk GeSn and SiGeSn cap layer are similar for a given CB valley, which ensures that carrier leakage to the direct minima in the cap is negligible. Leakage to the L-valley in the cap is also expected to be minimal, since the large DoS in the lower energy L-valley of the active material will ensure those states are preferentially filled. This, paired with low levels of thermal carrier distribution broadening at 85 K, results in carriers being well confined to the active region.

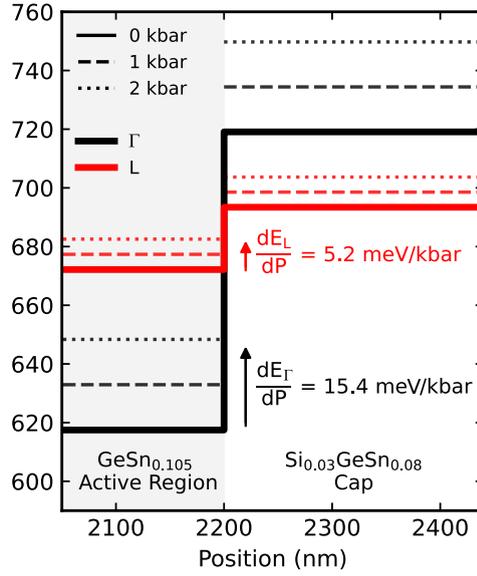


FIGURE 5.3: Movement of the flat band edges in the active region and cap layers as a function of increasing pressure (0 - 2 kbar). Black lines represent the band edge at the Γ -point, while red lines denote the L-point minima.

The threshold current densities for devices A (squares) and B (circles) are normalised at 0 kbar and plotted as a function of pressure in Fig 5.4. The green, orange and blue dashed lines depict the pressure dependence of the defect, radiative and leakage terms, given by Eqn. 5.2 - 5.4, respectively. Whilst leakage and defect related terms are independent of the choice of bowing parameters, the radiative term is inherently dependent on b_{GeSn}^Γ . The orange line utilises the bowing parameter given in Eqn. 4.76, whilst the filled area depicts the uncertainty in the values due to the range of b_{GeSn}^Γ cited in literature.

To extract the relative contributions, Eqn. 5.1 is normalised to give

$$\frac{J_{th}(P)}{J_{th}(0)} = \underbrace{\frac{J_{th}^{def}(0)}{J_{th}(0)}}_{\alpha} + \underbrace{\frac{J_{th}^{rad}(0)}{J_{th}(0)}}_{\beta} [E_g^\Gamma(P)]^{5/2} + \underbrace{\frac{J_{th}^{lk}(0)}{J_{th}(0)}}_{\gamma} \exp\left(-\frac{dE_a(P)}{dP} \frac{P}{k_B T}\right). \quad (5.5)$$

α , β and γ are then used as fitting parameters to determine the fraction of each process at threshold and atmospheric pressure. The carrier density required to fulfil the Bernard-Duraffourg condition is calculated and used to determine $J_{th}^{rad}(0)$. This provides a minimum value for $J_{th}^{rad}(0)$ and is used to reduce the fitting boundaries for β , which is set to vary between 0.05 - 1. The methodology for this calculation will be discussed in section 5.3. As such, γ is allowed to vary between 0 - 0.95. Since $\alpha + \beta + \gamma \equiv 1$ by definition, the expression can be fitted using only these two variables, improving confidence in

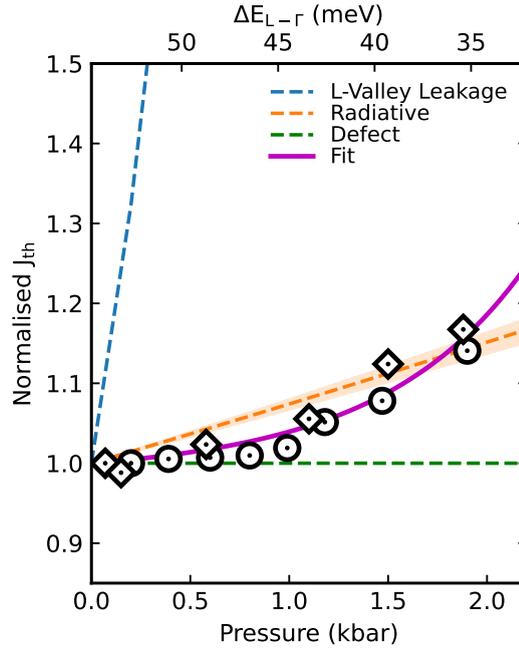


FIGURE 5.4: The change in threshold current density with pressure for samples A (diamonds) and B (circles), at 85 K, normalised to atmospheric pressure. The trends for each recombination pathway are denoted by dotted lines, with shading of equivalent colour representing the associated uncertainty. The fit is attained through a least squares minimisation of both data sets and is given by the solid magenta line.

the obtained fit. A standard, non-linear least-squares method is utilised for fitting.

As seen from the magenta line in Fig. 5.4, a good fit to the data is achieved with values of $\gamma = 1.15 \pm 0.02\%$, $\beta = 6 \pm 4\%$, $\alpha = 93 \pm 4\%$. The uncertainty bounds for the fit are not plotted since the result is too small to be visible. These results indicate that the threshold current density at 85 K is dominated by defect related recombination, accounting for around 90% of the total threshold. Transmission electron microscopy images highlight a markedly lower dislocation density in the active region of the sample when compared to the graded buffer layer, which is shown to be considerable [189]. It is therefore likely that an overwhelming fraction of carriers are not reaching the active region, due to significant non-radiative recombination localised at crystalline defects in the buffer layers, resulting in a low injection efficiency.

In this model, a constant Sn content and strain was assumed across the width of the active layer, as per Table 5.1. Since a graded buffer was used, it is expected that the Sn content is indeed fixed. However, the strain calculated from XRD will be an average over $1 \mu\text{m}$ of bulk material. It is therefore likely that there is a slight gradient in the strain profile. This solely impacts the radiative trend through a minor position-dependent shift in the

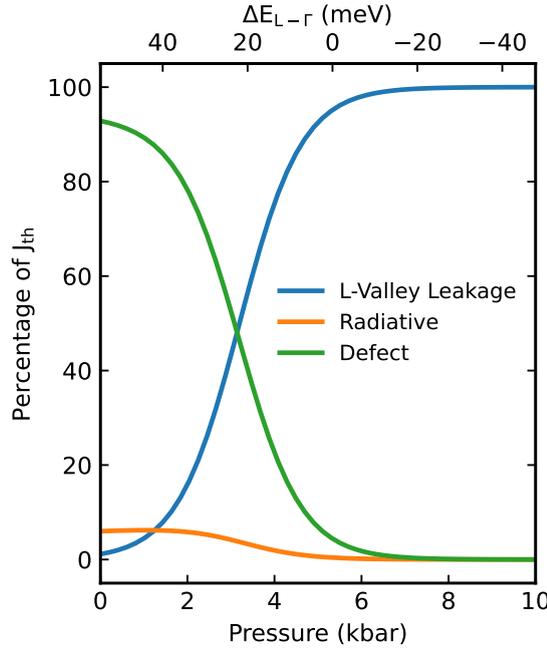


FIGURE 5.5: The percentage contribution of each process to the total threshold as a function of increasing pressure, determined by extrapolating the fit obtained in Fig. 5.4. The corresponding level of directness at each pressure is highlighted on the top axis.

bandgap energy, however the magnitude of strains involved and low radiative contribution render this effect negligible.

Whilst growth quality will likely be able to greatly reduce the effect of defect-related recombination as the field matures, this experiment provides strong evidence that carriers are occupying indirect-valley states, even at temperatures as low as 85 K. By extrapolating the fit to higher pressures and dividing by the total threshold current, it can be seen in Fig. 5.5 that an increase of only 3.15 kbar, equivalent to a decrease in the directness $\Delta E_{L-\Gamma} = 32$ meV, is required for the L-valley leakage current to become equal to the defect current at threshold (~ 1 kA/cm²). An increase of a further 10 meV results in a leakage current $4\times$ greater than the defect current. This illustrates that even a small to moderate decrease in the material directness results in a significant increase in carriers occupying indirect states. Since carriers occupying indirect states are unable to recombine radiatively in an efficient manner, they can therefore act as a source of optical loss through free carrier absorption (FCA). The subsequent relaxation through non-radiative recombination acts to further increase device temperature, broadening the carrier distribution and further populating the L-valley, resulting in a thermal runaway process [102].

In the above analysis, the threshold current density is shown to increase due to enhanced

L-valley occupation arising from tuning the offset between the Γ - and L-valleys using pressure. Similarly, the L-valley occupation may also be raised by increasing the thermal broadening of the carrier distribution. Since small to moderate decreases in the material directness elicited large degradation in device performance, it is important to probe whether the increasing indirect state occupation may be responsible for limiting the maximum operational temperature in electrically pumped, bulk devices.

5.3 Temperature Dependent Modelling

The feasibility of a given material or device heterostructure can be well assessed through modelling of the gain spectra. The material gain is a measure of the intrinsic ability of a semiconductor material to amplify light, and is entirely bandstructure dependent. Above a certain carrier density, the Bernard-Duraffourg condition will be met and a positive value for material gain will be produced. However, additional carrier injection will be required to compensate for optical loss in the device and achieve net gain. The threshold gain, g_{th} , is therefore defined as the amount of material gain needed to overcome said losses, and is given as

$$g_{th} = \frac{1}{\Gamma} (\alpha_i + \alpha_m) = \frac{1}{\Gamma} \left[\alpha_i + \frac{1}{2L} \ln \left(\frac{1}{R_1 R_2} \right) \right], \quad (5.6)$$

where Γ is the optical confinement factor, given by the overlap between the optical mode and the active region. α_i is the intrinsic optical loss per unit length due to processes such as inter-valence band absorption (IVBA), carrier scattering and FCA. α_m is the mirror loss per unit length arising from non-perfect reflectivity at the cavity facets. Here, L is the cavity length and $R_{1,2}$ are the facet reflectivities. In the case of uncoated facets, the mirror loss may be rewritten as $\alpha_m = 1/L \cdot \ln(1/R_f)$ where $R_f = R_1 = R_2$.

An ideal method for assessing feasibility of a material or device heterostructure at higher temperatures would be to model the temperature-dependent gain. In doing so, the carrier densities required to reach lasing threshold could be determined, which could be used to model the threshold current density using the standard ABC model discussed in section 2.4.

In practice, this requires accurate calculations of temperature-dependent IVBA, scattering and FCA losses. Whilst IVBA losses may be accounted for by including inter-valence band

transitions when calculating material gain from the bandstructure (to follow), FCA and scattering losses are far more difficult to quantify. This stems from the fact that the intraband transitions must be mediated by phonons to ensure momentum conservation. As such, first principle calculations require knowledge of materials parameters such as acoustic and optical phonon deformation potentials that are not documented for such novel materials [103]. Calculations of these parameters using first principle methods such as DFT are beyond the scope of this work. Other groups have attempted to quantify FCA in GeSn alloys using analytical models such as those derived from classical Drude theory, which requires knowledge of temperature- and carrier density-dependent carrier mobilities [192], or from empirical equations derived from measurements of pure Ge [193]. Using such methods, Hong et al. claim that FCA becomes so strong that net positive gain is only achievable for bulk alloys with <18% Sn for any level of injection (unstrained) [192], a claim that is disproved by the observation of lasing in optically pumped devices with in excess of 20% Sn two years prior [138].

Since a physically accurate and temperature-dependent model for FCA and scattering is unavailable, the scope of the modelling is limited to the material gain, specifically the minimum carrier density required to satisfy the Bernard-Duraffourg condition.. By neglecting optical loss processes, a lower boundary for the threshold carrier density, known as the transparency carrier density, is obtained. This allows for assessment of the suitability of GeSn as a bulk active region material from a fundamental bandstructure perspective.

5.3.1 Bulk Gain Modelling

With knowledge of the bulk active region bandstructure, calculated using the 8-band strain-corrected Kane model of Sec. 4.2.3, the material gain for a given photon energy can be determined by integrating over a truncated range of the Brillouin zone and summing over transitions between the conduction and valence bands as

$$G(\hbar\omega) = \frac{\pi q^2}{n_r c \varepsilon_0 m_0^2 \omega} \sum_{c,v} \iiint \frac{dk_x dk_y dk_z}{(2\pi)^3} |\vec{e} \cdot P_{cv}|^2 \times \delta(E_c - E_v - \hbar\omega)(f_c - f_v). \quad (5.7)$$

Here, q is the electron charge, n_r is the wavelength dependent refractive index, ε_0 is the permittivity of free space, c is the speed of light, m_0 is the electron mass, ω is photon angular frequency. $|\vec{e} \cdot P_{cv}|^2$ is the momentum matrix elements which are determined from the eigenvectors of the 8-band Hamiltonian.

Eqn. 5.7 assumes each state can be modelled using a perfect delta function. In practice, both carrier-carrier and carrier-phonon interactions broaden the state, resulting in a finite state lifetime which is characterised by the intraband relaxation time, τ . Thus, the delta function is replaced by a broadened distribution of states which are able to contribute to the gain at a specific photon energy. To fully incorporate the effect on the calculated material gain, τ should be calculated for each scattering interaction, ensuring the correct temperature, carrier density and \mathbf{k} -space dependencies. In practice, this again requires knowledge of material properties, such as phonon deformation potentials, that have not been parameterised for these materials. A standard hyperbolic secant lineshape given by

$$S(\hbar\omega) = \frac{1}{\delta\pi} \operatorname{sech}\left(\frac{E_c - E_v - \hbar\omega}{\delta}\right), \quad (5.8)$$

is therefore adopted. Here, $\delta = \hbar/\tau$ is the linewidth (full-width at half-maximum). A lack of literature is available for the form of the linewidth in bulk devices. As such, a simple temperature-dependent carrier lifetime, $\tau = 100 \text{ fs} \times 300 \text{ K/T}$ is adopted based on the reporting of quantum well based devices [194]. The choice of linewidth model has negligible effect on the qualitative trends observed in this work, but is included in the model for future work on materials requiring more quantitative gain calculations.

$f_{c(v)}$ is the Fermi-Dirac distribution describing the thermal distribution of electrons at quasi-equilibrium, defined as

$$f_{c(v)} = \left[1 + \exp\left(\frac{E_{c(v)}(k_x, k_y, k_z) - E_{f,c(v)}}{k_B T}\right) \right]^{-1} \quad (5.9)$$

where $E_{f,c(v)}$ is the conduction (valence) quasi-Fermi level. The quasi-Fermi levels are determined for a given injection by iterating an initial trial value until convergence is reached for the chosen carrier density. The carrier densities for the conduction and valence bands in bulk materials are given by

$$n_{3D} = \sum_{i \in c} \iiint \frac{dk_x dk_y dk_z}{(2\pi)^3} \left[1 + \exp\left(\frac{E_i^\Gamma(k_x, k_y, k_z) - E_{f,c}}{k_B T}\right) \right]^{-1} + 8 \left(\frac{m_L^{\text{dos}} k_B T}{2\pi \hbar^2}\right)^{3/2} F_{1/2}\left(\frac{E_c^L - E_{f,c}}{k_B T}\right) \quad (5.10)$$

and

$$p_{3D} = \sum_{i \in v} \iiint \frac{dk_x dk_y dk_z}{(2\pi)^3} \left[1 + \exp\left(\frac{E_{f,v} - E_i^v(k_x, k_y, k_z)}{k_B T}\right) \right]^{-1}, \quad (5.11)$$

respectively. The left-most term in Eqn. 5.10 denotes the carrier density of the conduction band in the Γ -valley and is determined directly from the 8-band $k \cdot p$ bandstructure. The second term accounts for carriers in the L-valley by using the semi-classical Thomas-Fermi density of states for parabolic bands. m_L^{dos} is the density of states effective mass and $F_{1/2}$ is short-hand for the Fermi-Dirac integral. Here, $F_{1/2}$ is approximated using the treatment of Fukushima in order to minimise computational cost, whilst ensuring high accuracy is maintained across the full range of carrier injections used in this study. The factor of 8 arises from a $2 \times$ spin- and $4 \times$ valley-degenerate L-valley. Here, the $k \cdot p$ model is not utilised to model the L-valley dispersion relation since the eigenstates of the system cannot easily be obtained from symmetry arguments, and reliable empirical parameters for such a new material system are not currently available.

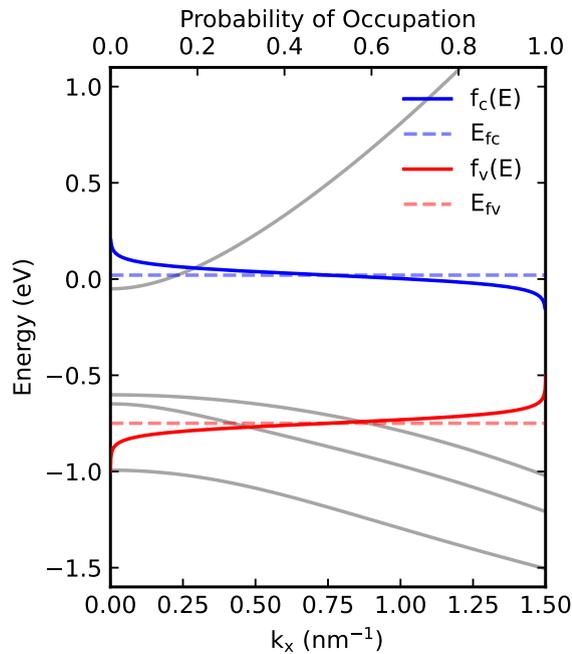


FIGURE 5.6: The Energetic positions of the VB (red, dashed) and CB (blue, dashed) quasi-Fermi levels and the associated thermally broadened Fermi-functions (solid) for 0.5% bi-axially strained $\text{GeSn}_{0.08}$ under an injection of $1 \times 10^{20} \text{ cm}^{-3}$. The probability of electron (blue) or hole (red) occupation is given by the top axis, whilst the bandstructure is illustrated in grey.

When computing the bandstructure, the chosen range in $|\mathbf{k}|$ must ensure sufficient states are calculated to avoid saturation at high temperatures and injections. Fig 5.6 illustrates the bandstructure of bulk $\text{GeSn}_{0.08}$ with a compressive strain of 0.5% at 300 K. These variables were selected to produce the least favourable configuration for gain within the range of values considered in this study, thus requiring the highest injection carrier densities. The solid blue and red lines represent the conduction and valence Fermi distributions under a very high injection of $1 \times 10^{20} \text{ cm}^{-3}$, respectively. At these levels of injection, a

$|\mathbf{k}|$ -range of $0\text{-}1.5\text{ nm}^{-1}$ is the minimum necessary to ensure an adequate number of states in the LH band. Whilst carriers at higher $|\mathbf{k}|$ will not contribute to gain, calculation of the states is required to achieve an accurate estimate of the valence band quasi-Fermi level, which is key in determining transparency.

5.3.2 Temperature Dependent Gain in Bulk GeSn

In order to determine the feasibility of bulk GeSn as an active medium, the gain spectra were calculated as a function of injected carrier density and temperature. The transparency carrier density was then extracted for each temperature and plotted.

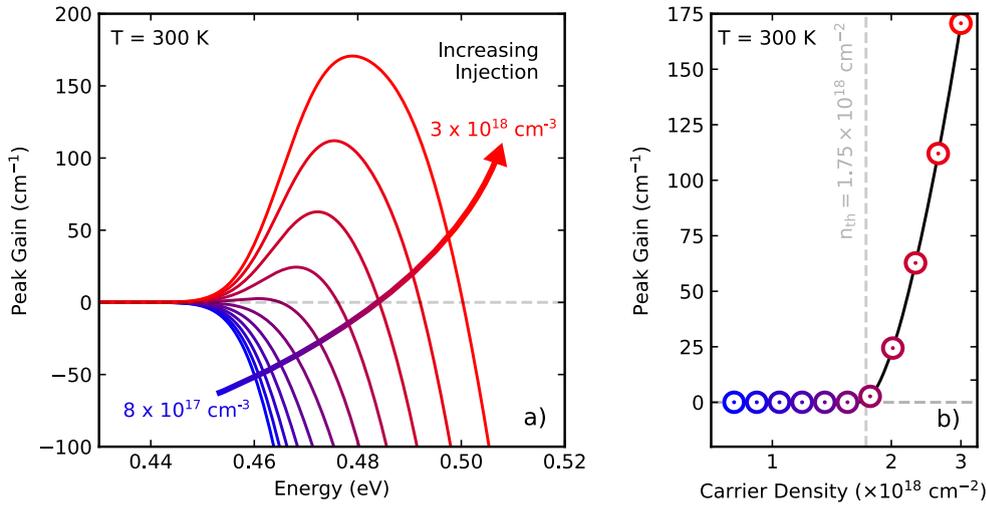


FIGURE 5.7: a) Gain spectra calculated as a function of increasing injection carrier density between $0.8\text{-}3 \times 10^{18}\text{ cm}^{-3}$ for GeSn_{0.105}, with a bi-axial strain of -0.27% at RT. b) illustrates the peak gain as a function of the carrier density, showing how the transparency carrier density is obtained through bisection of the data.

Initially, the active region of the devices in section 5.2, with a Sn composition of 10.5% and a residual compressive strain of -0.27% , were modelled. For these calculations, a set of 21 carrier densities were modelled per temperature point. A subsection of the RT gain spectra as a function of increasing carrier density are illustrated in Fig. 5.7 a), whilst Fig. 5.7 b) illustrates the threshold, determined from an interpolated fit of the gain peaks as a function of injection using bisection. Since the transparency carrier density was determined through interpolation, the accuracy of the root was optimised by decreasing the range of carrier densities used. To do this, the transparency point was initially estimated using the analytical equations of the parabolic band approximation. Bounds for the carrier density sweep were then taken to be plus or minus an integer

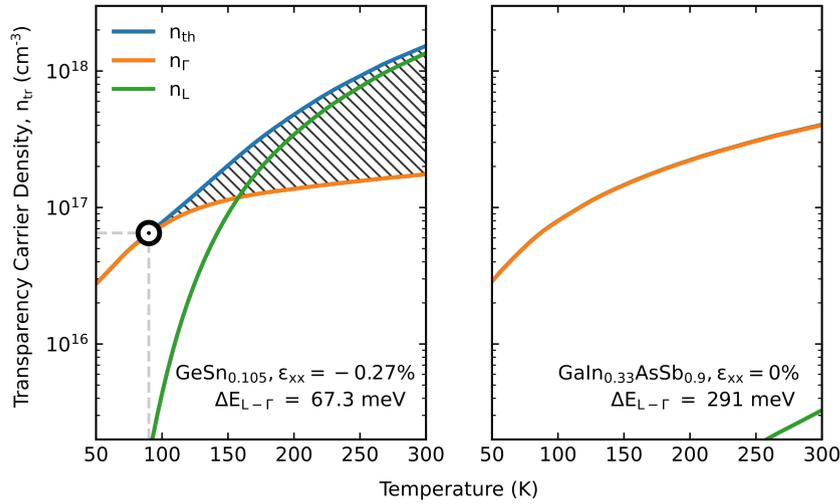


FIGURE 5.8: a) Transparency carrier density with temperature (blue), and the distribution of those carriers between the Γ (orange) and L-valleys (green). The white circular marker indicates the temperature at which over 2% of carriers are occupying the indirect valley. b) illustrates the same relations for GaIn_{0.33}AsSb_{0.9} to highlight the difference when using material with high levels of directness.

multiple of the calculated value. This process was repeated for 11 temperatures between 0 and 300 K.

Fig. 5.8 a) shows the transparency carrier density as a function of temperature for the GeSn devices, indicated by the solid blue line. n_{tr} was subsequently used to calculate the quasi-Fermi energies, from which Eqn. 5.10 was decomposed to calculate the Γ and L-valley carrier concentrations. These are represented by the orange and green lines, respectively. At low temperatures, the material has transparency densities akin to standard III-Vs at $\sim 3 \times 10^{16}$ cm⁻³. This is illustrated by way of comparison with the III-V-Sb material system to be discussed in the following chapter, shown in Fig. 5.8 b). Below 90 K, the Γ -valley carrier density accounts for more than 98% of the threshold carrier density, indicating that almost all injected carriers occupy direct states. This is in good agreement with the experimental results of the previous section where only $\sim 1\%$ of the threshold current was attributed to L-valley leakage at 85 K.

Above 90 K, the density of indirect carriers at transparency exceeds 2%, marking a significant change in the observed trend. This is therefore denoted as the break-point temperature. As the temperature is increased further, the fraction of injected carriers occupying indirect states increases. As such, a higher density of total carriers is required to achieve the necessary direct valley density for transparency. At only 150 K the density of carriers in the Γ - and L-valleys are equivalent. At RT, the transparency carrier density is

$1.75 \times 10^{18} \text{ cm}^{-3}$, $\sim 50\times$ greater than at 50 K. By comparison GaIn_{0.33}AsSb_{0.90} elicits a RT threshold only $13\times$ greater than at 50 K. This disparity stems primarily from the difference in directness. At RT n_L accounts for 88.5% of the transparency carrier density. By comparison, the large levels of directness in GaIn_{0.33}AsSb_{0.9} results in a RT L-valley percentage just 0.81% of n_{tr} .

Since this metric represents the minimum carrier density in the L-valley needed to produce gain, real devices with this composition would exceed 50% indirect occupancy at 150 K. Thus, this is likely limiting the operational temperature of current generation devices and suggests high temperature lasing through electronic pumping is not practicable using this composition. It is therefore necessary to investigate the parameter space of bulk growth, which encompasses the Sn composition and some small residual levels of strain, to determine whether RT operation may be feasible for alternative compositions.

5.3.2.1 Residual Strain Optimisation

The active region of the investigated devices has a quoted strain of -0.27% , determined through XRD reciprocal space mapping measurements conducted by the grower [189]. As discussed previously, increasing the compressive strain increases the energy of the Γ valley minima at a faster rate than the L-valley, resulting in a reduced material directness. Therefore, it is necessary to assess how much this small residual compressive strain affects

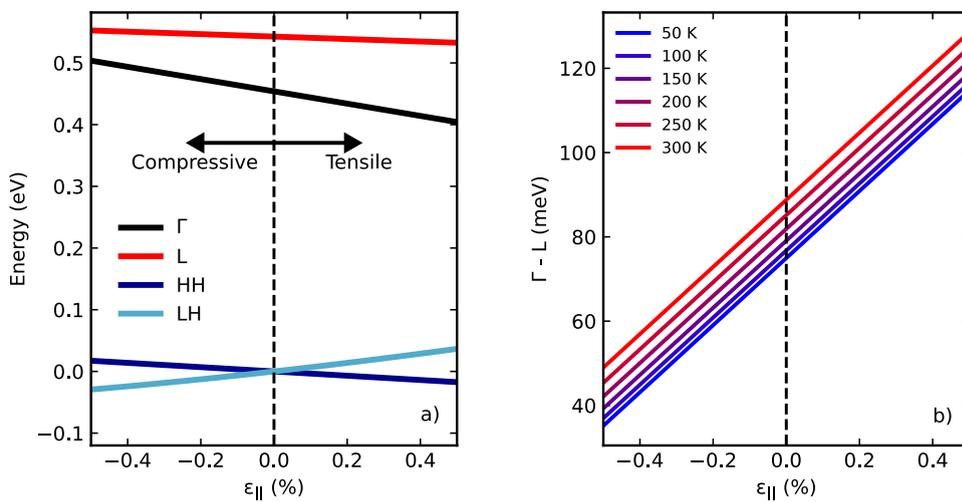


FIGURE 5.9: a) Energetic positions of key band extrema in GeSn_{0.105}, as a function of strain, with respect to the unstrained VB edge. b) Highlights how the Γ -L splitting changes with strain and increasing temperature, showing marginally increased directness at higher temperatures, and rapidly increasing directness with increasing tensile strain.

the temperature dependence of the transparency carrier density reported in the previous section. Additionally, evaluating whether the reduction of compressive strain or the inclusion of small tensile strains could yield large performance improvements is essential for informing next-generation device design.

The bounds for the residual strain values are taken to be -0.5 to $+0.5\%$, representing a rough upper limit for achievable values in bulk devices. Fig 5.9 a) illustrates the effect of strain on the flat band edges of the device at RT, highlighting the lifting of degeneracy in the VB and the change in offset between the Γ and L minima. The material directness as a function of strain is further highlighted in Fig 5.9 b) for temperatures between 50-300 K. This illustrates that eliminating residual compressive strain would result in an increase in directness in excess of 20 meV, while introducing a tensile strain of the same magnitude would result in a ~ 50 meV increase. Since a decrease in directness of just 30 meV was shown to have a significant deleterious effect on the experimentally observed device performance, it can be postulated that even small residual strains may play an important role in determining thresholds and maximum operational temperatures.

The process outlined in section 5.3.2 is again utilised for a 10.5% Sn alloy, and repeated for -0.5% , 0.5% and unstrained material. Fig. 5.10 a) depicts the compressively strained case. At 50 K, a marginally lower threshold of $1.5 \times 10^{16} \text{ cm}^{-3}$ is observed compared to the unstrained material, shown in Fig. 5.10 b). However, the L-valley occupation at transparency is significantly higher for all temperatures, stemming from a decrease in directness of 20 meV. As such, at transparency, 50% of carriers occupy the L-valley at

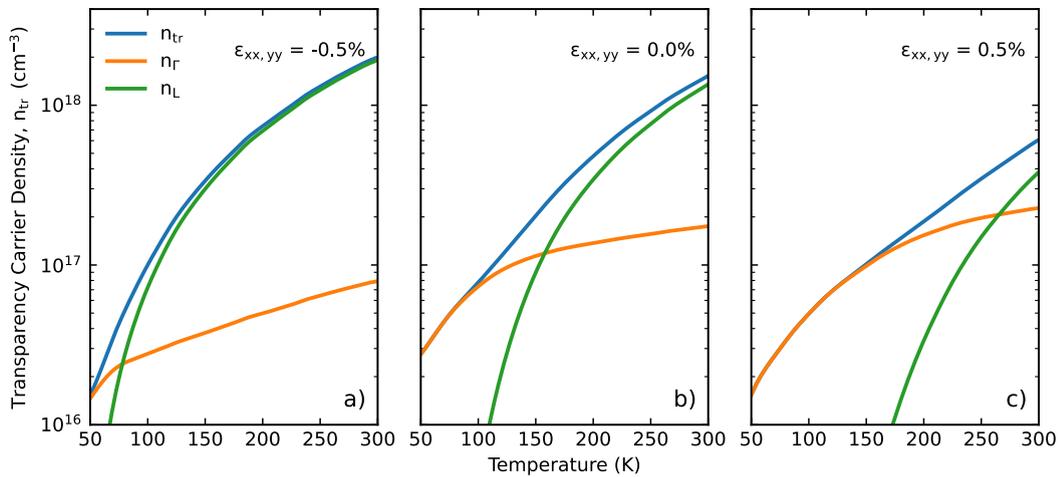


FIGURE 5.10: Transparency carrier density as a function of temperature (blue), and the distribution of those carriers between the Γ (orange) and L-valleys (green) for a) -0.5% , b) 0% and c) 0.5% strained GeSn_{0.105}, respectively.

just 75 K, markedly less than the 150 K observed in unstrained material. At RT indirect carriers account for 96% of n_{tr} . Despite this, the transparency carrier density is only $1.3\times$ higher for the compressively strained material when compared to unstrained GeSn_{0.105}.

In contrast, the inclusion of tensile strain significantly improves the VB and CB band-structure from a gain perspective, and therefore lowers the transparency carrier density. At 50 K, the required carrier density is equal to that of the compressively strained material. However, at high temperature, threshold is over $3\times$ lower than for the compressively strained material and $2.5\times$ lower than unstrained material. Carriers are equally divided between the L- and Γ -valleys at 250 K, an increase of 100 K over unstrained material. In the interest of applications, the inclusion of 0.5% tensile strain drastically alters the emission wavelength, pushing it to $3.4\ \mu\text{m}$ at RT compared to $2.55\ \mu\text{m}$ for a compressive strain of the same magnitude.

To better understand the origins of the observed differences, the behaviour of the quasi-Fermi levels are analysed as a function of carrier density. Fig. 5.11 a) represents the difference between the VB quasi-Fermi energy and the VB edge at 50 K. Both the compressive (red) and tensile (blue) materials show VB Fermi levels closer to, or further into the VB for any given carrier injection. This physically stems from a decrease in the VB DoS due to strain induced degeneracy lifting. As depicted in Fig. 5.11 b) for tensile material, the LH band is ‘light’ in the growth direction, but ‘heavy’ in-plane and vice-versa for

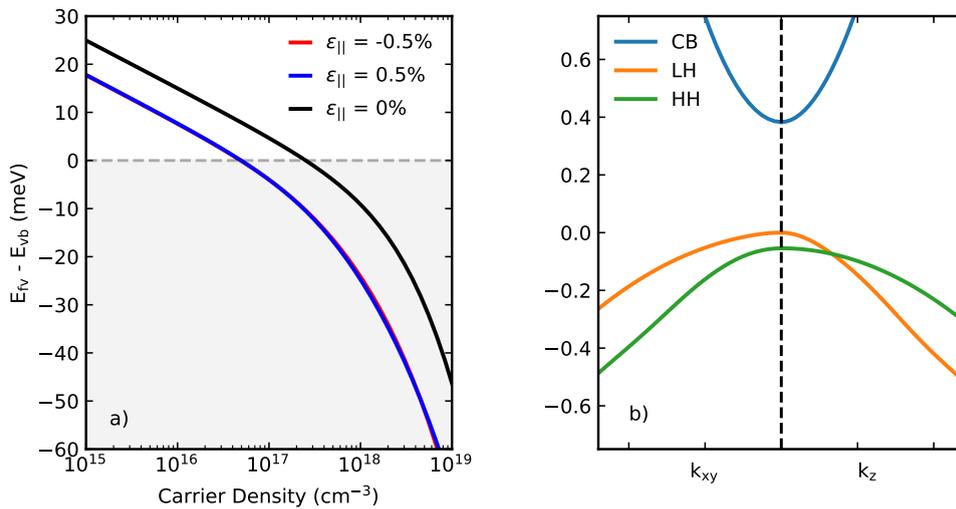


FIGURE 5.11: a) The relative position of the VB quasi-Fermi level with respect to the valence band edge as a function of increasing injection for -0.5% (red), 0% (black) and 0.5% (blue) strained GeSn_{0.105} at 50K. b) Effect of strain on the VB structure in the k_z and k_{xy} directions.

the HH band. The result is that the DoS is comparable for compressive or tensile strains of equal magnitude [84]. As such, the rate of movement of the VB quasi-Fermi energy with carrier density are equivalent for the strained bulk materials in Fig. 5.11 a).

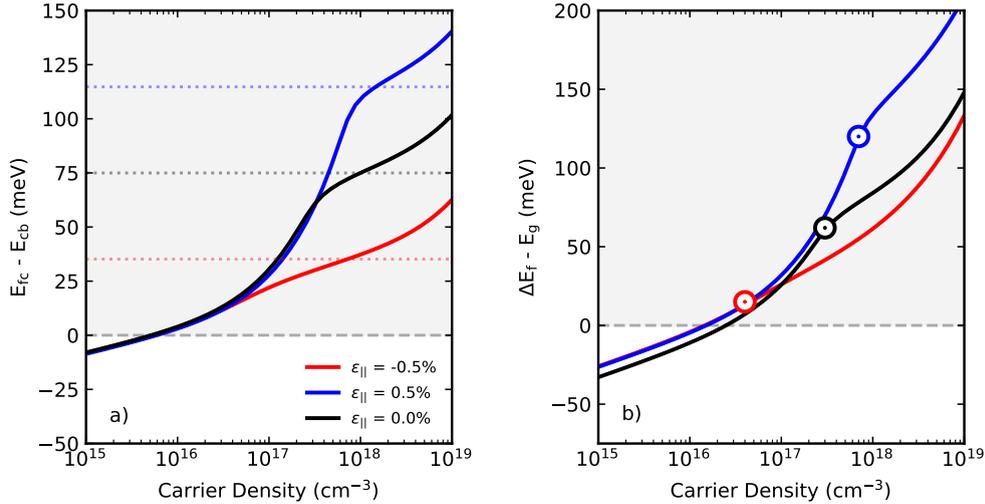


FIGURE 5.12: a) The relative position of the CB quasi-Fermi level with respect to the valence band edge as a function of increasing injection for -0.5% (red), 0% (black) and 0.5% (blue) strained GeSn_{0.105} at 50K. Dotted lines of comparable colour show the relative energy of the L-valley minima with respect to the Γ point energy, highlighting the change in directness with strain. b) The difference between the quasi-Fermi splitting and the bandgap for each strain state. A value of 0 indicates the carrier density for which the Bernard-Durrafourg condition is satisfied.

Fig. 5.12 a) depicts the difference between the CB quasi-Fermi energy and the CB edge at 50 K, with dotted lines of equivalent colour denoting the energies of the L-valley above the CB edge for each strain state. Below $4 \times 10^{16} \text{ cm}^{-3}$, the CB quasi-Fermi level penetrates into the CB equally for a given carrier density independent of the material strain. This illustrates that, at low temperatures, the thermal broadening is low enough that the carriers are confined to the Γ -valley. As the carrier density increases, the effects of directness become apparent. For the compressively strained material, the decreased directness results in a markedly reduced rate of penetration into the CB with carrier density above $4 \times 10^{16} \text{ cm}^{-3}$. The effect of the L-valley on the movement of the CB quasi-Fermi level becomes apparent at $3 \times 10^{17} \text{ cm}^{-3}$ and $7 \times 10^{17} \text{ cm}^{-3}$ for unstrained and -0.5% tensile strained material, respectively.

These two effects combine in Fig. 5.12 b) which illustrates the difference between the quasi-Fermi level splitting and the bandgap. It is observed that the transparency condition is satisfied at $1.5 \times 10^{16} \text{ cm}^{-3}$ for strained material and $2.5 \times 10^{16} \text{ cm}^{-3}$ for unstrained material. At these low carrier densities, the movement of the CB quasi-Fermi level is

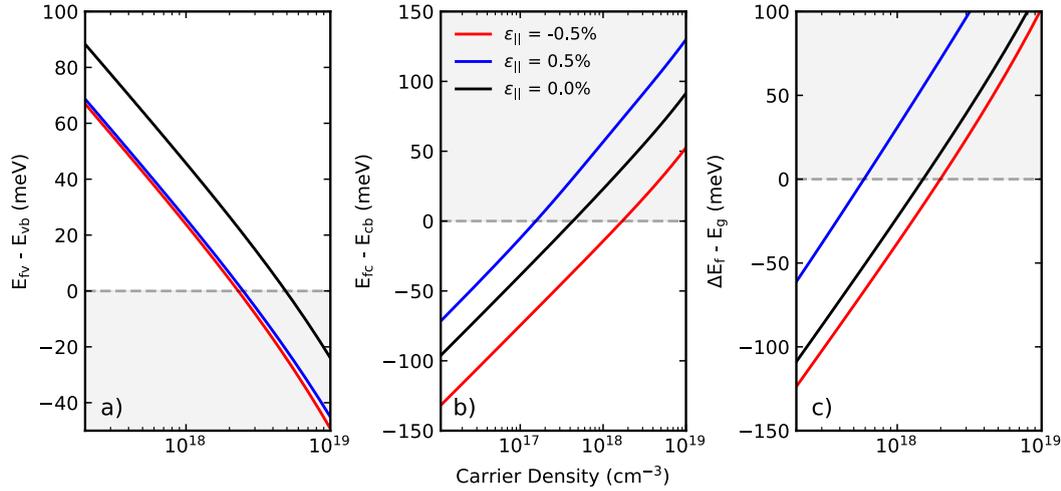


FIGURE 5.13: The separation between the a) VB quasi-Fermi level and the VB edge, b) CB quasi-Fermi level and CB edge, and c) quasi-Fermi separation and bandgap, as a function of carrier density at RT, for -0.5% (red), 0% (black) and 0.5% (blue) strained GeSn_{0.105}, respectively.

equal for all modelled strain states, thus the improvement in the transparency carrier density at low temperature can be solely attributed to the the improved VB DoS as a result of degeneracy lifting.

The same calculations were subsequently repeated at 300 K in Fig. 5.13 a), b), and c). The effect of thermal broadening is to smooth out the trends observed at 50 K. An improvement in the VB DoS is again observed in a) for both strained cases, however the penetration of the CB quasi-Fermi level into the CB shown in b) is distinct for each strain state at 300 K, even at low injection. A carrier density $\sim 10\times$ higher is required for the CB Fermi-level to reach the CB band edge for the compressively strained material compared to tensile. Thus, for residual compressive strains, the benefits of lower VB DoS is offset by the low levels of directness, resulting in transparency being achieved at a carrier density marginally higher than unstrained material. Contrastingly, for tensile GeSn_{0.105}, improvements to the DoS is the combination of both improved VB DoS and increased directness.

Despite a $2.5\times$ lower transparency carrier density than unstrained material, the fraction of indirect carriers at RT in 0.5% tensile GeSn_{0.105} remains in excess of 60%. It can be concluded that the small residual strains available to bulk devices are unable to provide sufficient improvement in carrier confinement to make RT lasing feasible when using 10.5% Sn devices. Additionally, in bulk devices where layers are thick, such levels of strain would mean the layer exceeds the critical thickness, resulting in high defect densities that would

greatly hinder performance. As such, other metrics should be considered for tuning the material directness.

5.3.2.2 Bulk Composition Optimisation

Since tuning of the strain achievable in bulk structures with 10.5% Sn is unable to provide the necessary levels of k-space carrier confinement, the effect of varying Sn content is investigated for unstrained alloys. Fig. 5.14 depicts the L- Γ splitting as a function of Sn% and temperature. It can be seen that an increase of 2% Sn constitutes a ~ 40 meV increase in material directness, with in excess of 20% Sn required to achieve levels comparable with standard III-V materials.

The calculations of transparency carrier density as a function of temperature are repeated for Sn content between 6-22%, constituting the range of compositions attained for optically pumped bulk devices reported in the literature. Larger Sn-content alloys are excluded from the analysis since the validity of the virtual crystal approximations becomes uncertain due to potential ordering effects, and the k-p model is known to produce spurious solutions for very low bandgap materials [195].

Fig. 5.15 again illustrates the transparency carrier density and the distribution of the carriers between the Γ and L-valleys as a function of temperature. Each subplot displays the Sn composition and the RT material directness. As expected, n_{tr} decreases with increasing Sn content for all temperatures. In the low Sn content alloys (6-10%), the direct

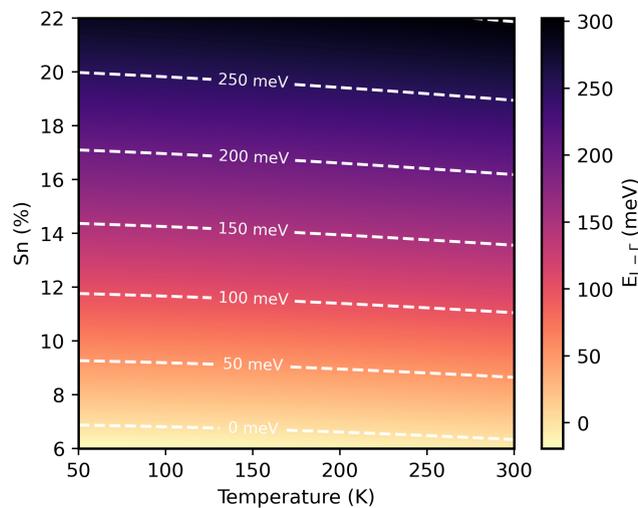


FIGURE 5.14: Directness as a function of temperature and Sn% in unstrained alloys.

valley carrier density at transparency is low. This reflects that the Bernard-Durraforg condition is achieved despite preferential population L-valley states. Since the L-valley DoS is significantly larger than the Γ -valley, the change in CB quasi-Fermi level with

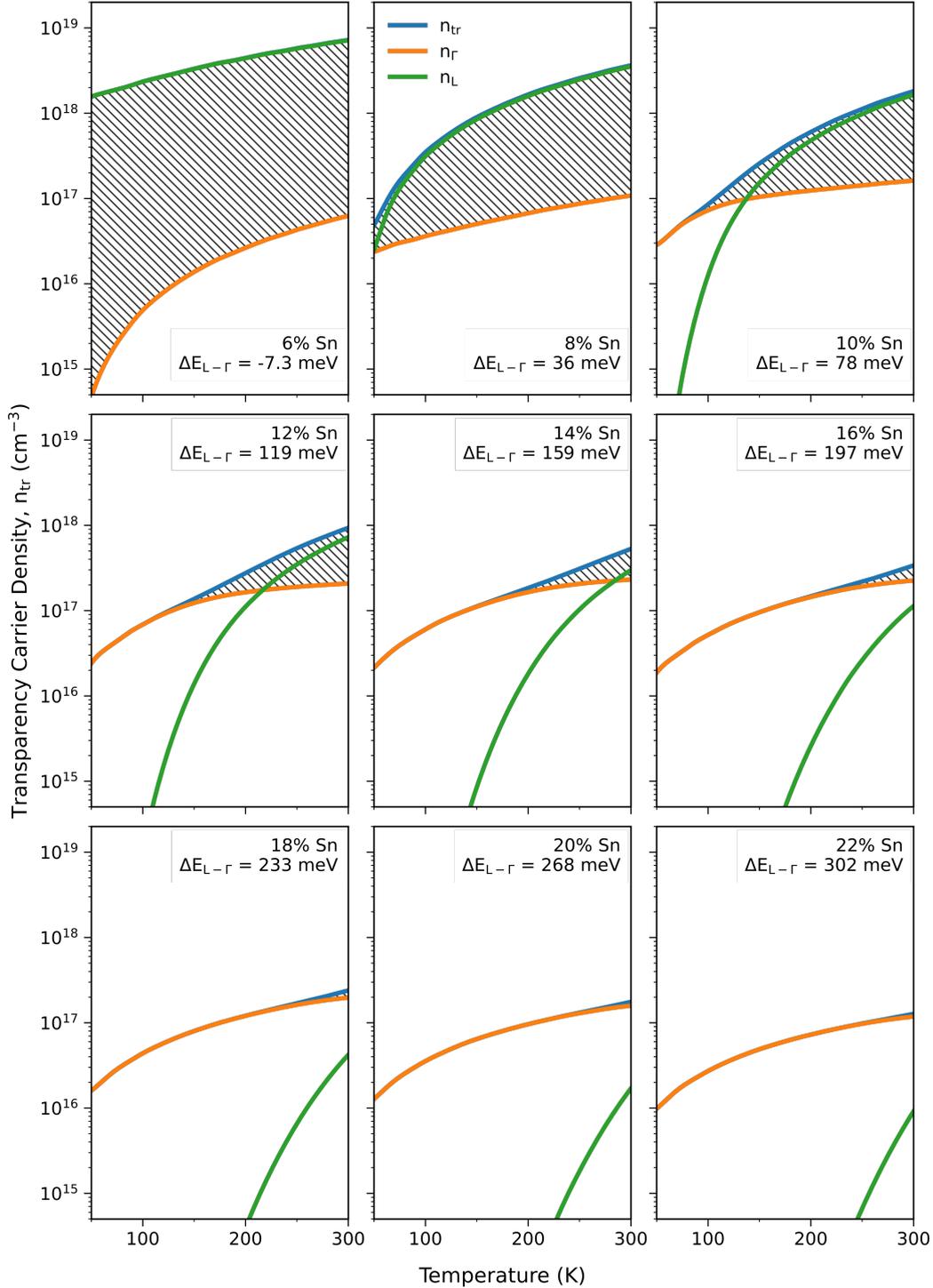


FIGURE 5.15: Transparency carrier density (blue) as a function of temperature, and the distribution of those carriers between the Γ (orange) and L-valleys (green), for increasing Sn content between 6-22%. Each subplot illustrates the level of directness at RT for the given alloy.

injection, dE_{fc}/dn , is notably lower. Transparency is therefore achieved primarily through greater penetration of the VB quasi-Fermi level into the VB, requiring a significantly higher carrier density.

Between 6-10% Sn, n_{Γ} increases since the CB edge has greater Γ character, raising dE_{fc}/dn . Above 10%, at low temperature, a marginal decrease in n_{tr} is observed with increasing Sn content. This is the culmination of two distinct processes. Firstly, the directness becomes large enough that carriers are well confined to the Γ valley at the levels of injection required to reach transparency. Secondly, there is a notable decrease in the charge carrier effective masses with Sn content. Fig 5.16 a) shows the effective masses of the Γ electrons (black), and heavy (dark blue) and light (light blue) holes, respectively, normalised to their values at 6% Sn. The decrease in effective mass for all carriers results in marginally lower CB and VB density of states, raising dE_{fc}/dn and dE_{fv}/dp such that n_{tr} is reduced. These two basic premises can be used to explain all aspects of the observed trends.

To better depict the effect of Sn content on L-valley occupation, the transparency carrier density is calculated as a function of temperature for Sn contents between 6-22% in 1% increments. Fig. 5.16 b) illustrates the variation of n_L/n_{tr} with temperature and Sn composition. Contour lines represent indirect occupations of 40%, 20%, 10%, 5% and 2%, from blue to red, respectively. From this, it is observed that it is possible to limit

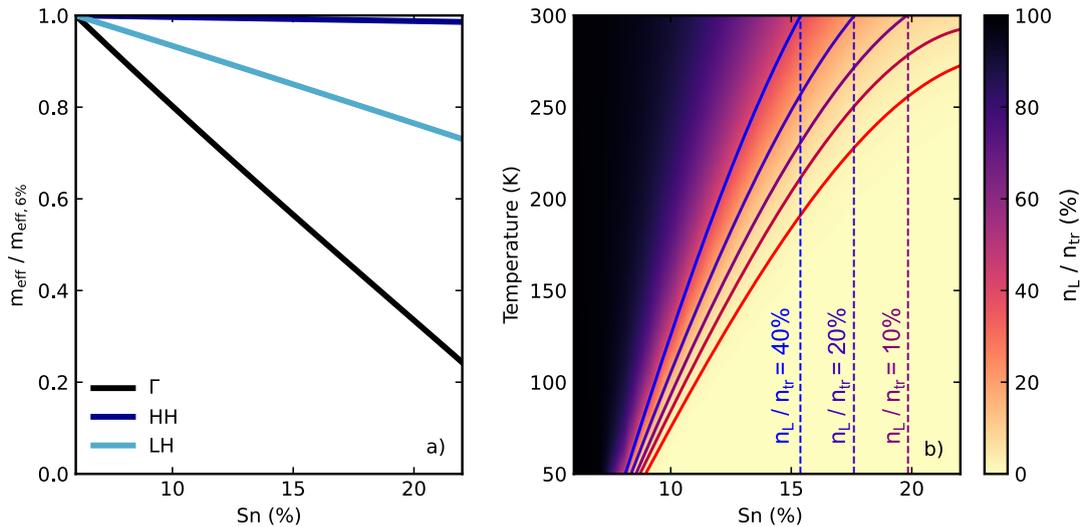


FIGURE 5.16: a) the carrier effective masses of the CB, and the LH and HH bands, as a function of Sn content, normalised to the value at 6%. b) The ratio of carriers in the L-valley at transparency to the total transparency carrier density as a function of Sn content and Temperature. Coloured contours depict 2%, 5%, 10%, 20% and 40% from red to blue, respectively.

the fraction of indirect carriers to less than 2% for temperatures below 200 K by using commonly utilised GeSn_{0.15}, as per [22]. This suggests that there may indeed be scope for higher temperature operation than current devices should defect formation be suitably suppressed. However, at RT, n_L/n_{tr} is in excess of 40% for alloys with 15% Sn, and even 20% Sn material still has 10% of carriers residing in L-valley states.

As such, RT lasing may only be attainable in GeSn alloys with in excess of 20% Sn. Near RT operation was experimentally demonstrated in devices with such compositions, requiring significant pump power densities. However, deployment as an active region in electrically pumped bulk devices comes with notably more challenges in terms of design. These include, but are not limited to:

1. Inherent difficulty with incorporating Sn content above 15% due to low solubility.
2. The need for higher Sn content buffers and barriers to reduce compressive strain in the active layer, introducing larger defect densities through heightened lattice mismatch with Ge and/or Si.

5.4 Intervalence Band Absorption

The intervalence band absorption is calculated as per Eqn. 5.7, with the conduction-valence band transition matrix replaced by the intervalence band transition matrices, denoted $|M|^2$. In the following section, the derivation of the transition matrix elements for IVBA processes, as described in [104], are reiterated for clarity. The \mathbf{k} -dependent matrices can be determined by noting that the valence band wavefunctions can be written to first order using perturbation theory as

$$|J, m_J, \mathbf{k}\rangle = e^{i\mathbf{k}\cdot\mathbf{r}} \left(|J, m_J\rangle + \frac{\hbar}{m_0} \sum_l \frac{\langle u_l | \mathbf{k} \cdot \mathbf{p} | J, m_J \rangle}{E_0 - E_l} |u_l\rangle \right), \quad (5.12)$$

where $|J, m_J\rangle$ is the valence band state at $\mathbf{k} = 0$ with total angular momentum J and z-projection M_J . Summation over the states $|u_l\rangle$ is to account for mixing with states other than the valence bands included explicitly in the basis set.

By definition, the transition matrix elements are then given as

$$M = \langle i, \mathbf{k} | \hat{\mathbf{e}} \cdot \mathbf{p} | j, \mathbf{k} \rangle \quad (5.13)$$

$$= \sum_{\alpha, \beta} e_{\alpha} \hbar k_{\beta} \left(\delta_{i,j} \delta_{\alpha, \beta} + \frac{2}{m_0} \sum_l \frac{\langle i, 0 | p_{\alpha} | u_l \rangle \langle u_l | p_{\beta} | j, 0 \rangle}{E_0 - E_l} \right), \quad (5.14)$$

where $\alpha, \beta = x, y, z$ and i, j are states belonging to the basis set.

Recalling that the Hamiltonian takes the form

$$H_{ij} = \left[E_j(0) + \frac{\hbar^2 k^2}{2m_0} \right] \delta_{ij} + \frac{\hbar^2}{m_0^2} \sum_{\alpha, \beta} k_{\alpha, \beta} \sum_l \frac{\langle i, 0 | p_{\alpha} | u_l \rangle \langle u_l | p_{\beta} | j, 0 \rangle}{E_0 - E_l}, \quad (5.15)$$

which, within the context of the 6-band model, can be expressed in terms of the standard Luttinger-parameters as

$$H = \begin{pmatrix} P+Q & -S & R & 0 & -S/\sqrt{2} & \sqrt{2}R \\ & P-Q & 0 & R & -\sqrt{2}Q & \sqrt{3/2}S \\ & & P-Q & S & \sqrt{3/2}S^+ & \sqrt{2}Q \\ & H.C. & & P+Q & -\sqrt{2}R^+ & -S^+/\sqrt{2} \\ & & & & P+\Delta_{SO} & 0 \\ & & & & & P+\Delta_{SO} \end{pmatrix} \quad (5.16)$$

where H.C. denotes the hermitian conjugate, the transition matrix elements for a given transition can be determined through simple comparison. The term $\sum_l \langle i, 0 | p_{\alpha} | u_l \rangle \langle u_l | p_{\beta} | j, 0 \rangle / (E_0 - E_l)$ in Eqn. 5.12 can be found by equating Eqns. 5.15 and 5.16. For example, considering transitions between the HH and SO bands, the mixing terms are given by:

$$\sum_l \frac{\langle 1 | p_{\alpha} | u_l \rangle \langle u_l | p_{\beta} | 6 \rangle}{E_0 - E_l} = \frac{m_0}{\sqrt{2}} \left[\sqrt{3}\gamma_2 (\delta_{\alpha x} \delta_{\beta x} - \delta_{\alpha y} \delta_{\beta y}) - i\sqrt{3}\gamma_3 (\delta_{\alpha x} \delta_{\beta y} - \delta_{\alpha y} \delta_{\beta x}) \right], \quad (5.17)$$

and

$$\sum_l \frac{\langle 4 | p_{\alpha} | u_l \rangle \langle u_l | p_{\beta} | 6 \rangle}{E_0 - E_l} = \frac{m_0}{2\sqrt{2}} \left[\sqrt{3}\gamma_3 (\delta_{\alpha x} \delta_{\beta z} + \delta_{\alpha z} \delta_{\beta x} + i\delta_{\alpha y} \delta_{\beta z} + i\delta_{\alpha z} \delta_{\beta y}) \right]. \quad (5.18)$$

The polarisation averaged transition matrix elements can then be calculated using Eqn. 5.13, where e_{α}^2 is taken to be 1/3 in all directions and the resultant matrix elements are

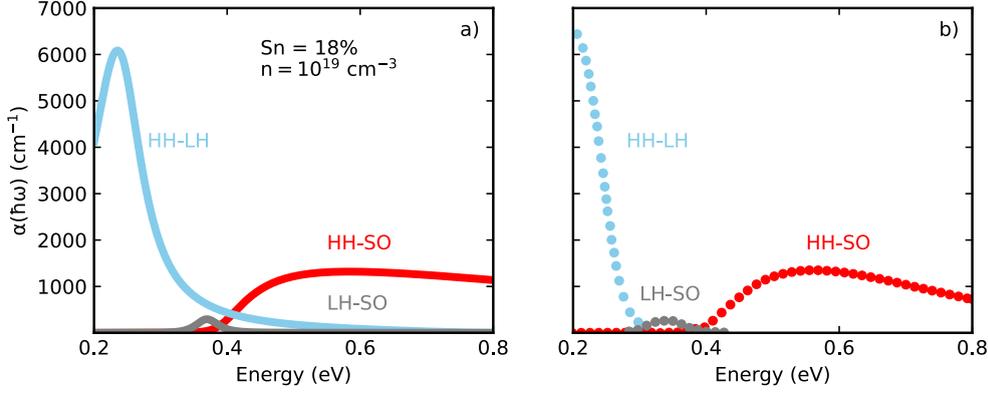


FIGURE 5.17: IVBA spectra for HH-LH (blue), LH-SO (grey) and HH-SO (red) transitions in $\text{GeSn}_{0.18}$ under an injection of $1 \times 10^{19} \text{ cm}^{-3}$ a) calculated from the 6-band k-p method and b) taken from the EPM model of [196].

summed to account for the double spin-degeneracy of each band. The elements for HH to SO and LH to SO transitions are therefore

$$\sum |\langle i | \hat{\mathbf{e}} \cdot \mathbf{p} | j \rangle |_{HH \rightarrow SO}^2 = 2\hbar^2(\gamma_2^2 + \gamma_3^2)(k_x^2 + k_y^2) + \frac{\hbar^2}{2}\gamma_3^2(k_z^2 + k^2), \quad (5.19)$$

$$\sum |\langle i | \hat{\mathbf{e}} \cdot \mathbf{p} | j \rangle |_{LH \rightarrow SO}^2 = \frac{2}{3}\hbar^2\gamma_2^2(3k_z^2 + k^2) + \frac{3\hbar^2}{2}\gamma_3^2(k_z^2 + k^2), \quad (5.20)$$

where $k^2 = k_x^2 + k_y^2 + k_z^2$. Since the bandstructure is calculated for positive values of $k_{x,y,z}$, the resultant IVBA profile is multiplied by 8 to account for the remaining octants of the Brillouin zone.

Fig. 5.17 a) and b) illustrate the IVBA as determined by the 6-band Hamiltonian used in this work, and using EPM calculated bandstructure, taken from [196]. The spectra show the contributions from transitions between each pair of VBs for 18% GeSn under an injection of $n = 1 \times 10^{19} \text{ cm}^{-3}$. HH-LH and HH-SO transitions show elongated tail absorption features which likely stem from differences in the calculated bandstructure at high- \mathbf{k} where the k-p method becomes less accurate. However, good agreement is achieved for the peak absorption magnitude of each interband transition. This reaffirms that Eqns. 5.19 and 5.20 are suitably accurate models for the interband matrix elements. The previous sections have illustrated that high Sn content and/or tensile strain are required to reduce transparency carrier density. Here, the effect of varying these parameters on the IVBA loss is analysed and used to reassess the optimal composition for a bulk GeSn laser. Fig. 5.18 illustrates the RT IVBA, interband absorption and net absorption as a function of injected carrier density for GeSn with 10%, 16% and 20% Sn, respectively. At 10% Sn, the bandgap is such that the emission peak lies close to the HH-SO IVBA maxima.

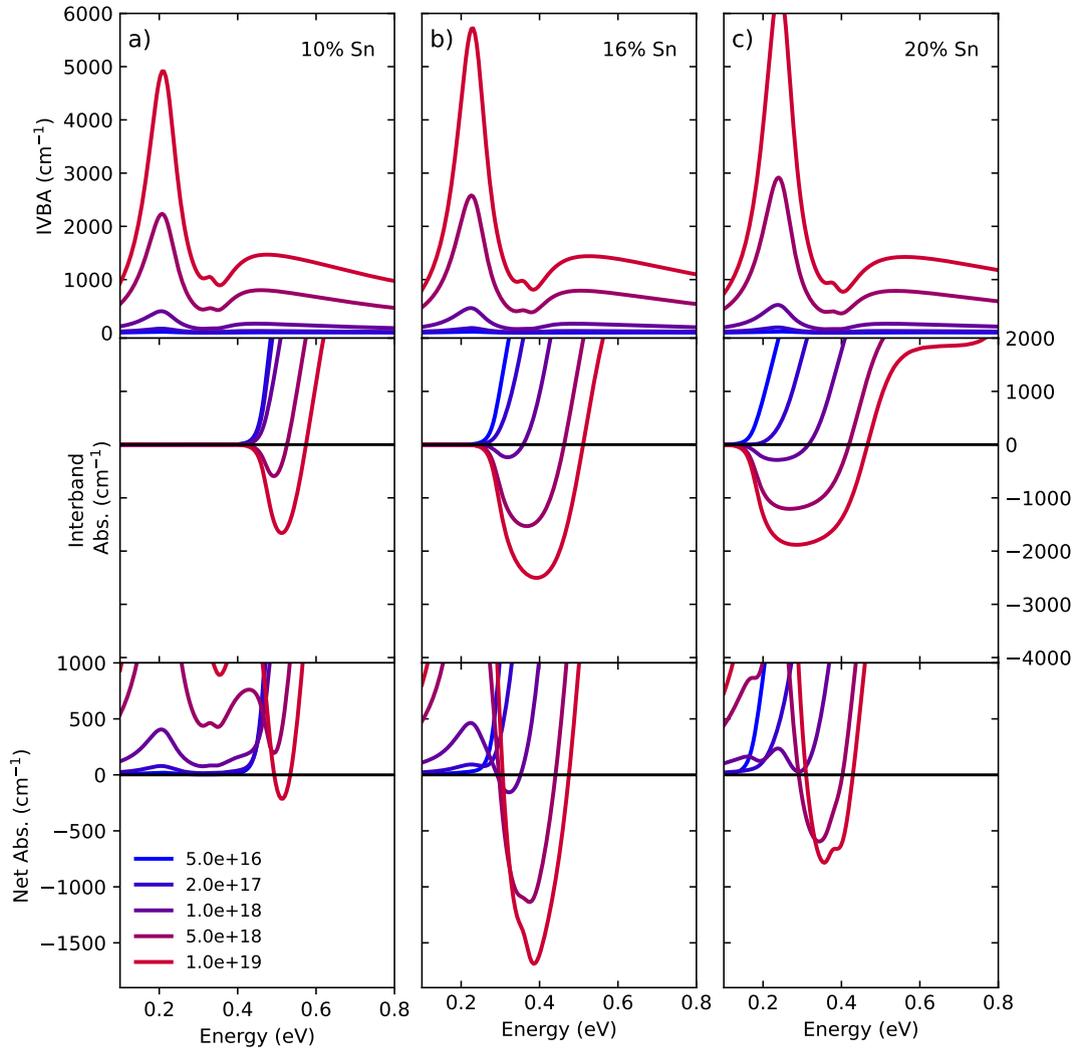


FIGURE 5.18: The IVBA, interband absorption (material gain) and net absorption as a function of carrier density for a) 10%, b) 16% and c) 20% GeSn alloys. Positive values denote absorption while negative values represent gain.

Since large carrier densities are required to achieve gain, stemming from high L-valley occupation, there is comparatively low gain compared to IVBA, resulting in small levels of net gain achievable only for injections in excess of $\sim 10^{19} \text{ cm}^{-3}$.

For 16% Sn, depicted in b), the peak gain for a given injection is heightened as a result of increased levels of directness. The bandgap reduction shifts the emission spectra to lower energies, such that the peak gain coincides with a minima in the IVBA spectra

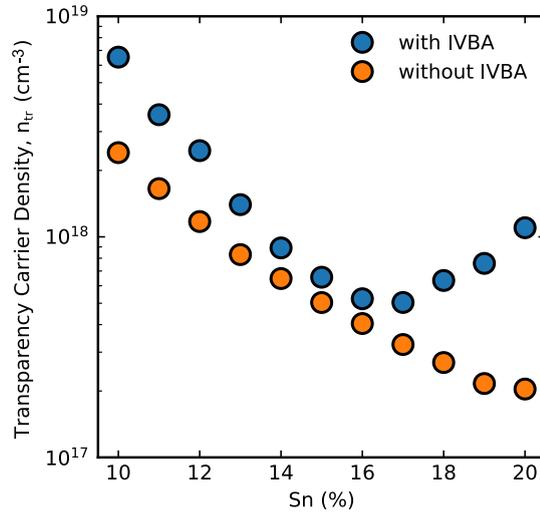


FIGURE 5.19: The RT transparency carrier density calculated with IVBA (blue) and without IVBA (orange), indicating that increasing Sn% can begin to degrade performance due to heightened IVBA loss from LH-HH transitions.

corresponding to LH-SO transitions. Despite improved direct-valley occupation, increasing the Sn content further does not yield greater net gain for a given transition. This is illustrated in Fig. 5.18 c). At this composition, the peak gain is reduced and the spectra is broadened for higher injections since the CB DoS decreases with increasing Sn content. The emission peak also shifts towards lower energies, where HH-LH IVBA is strong. These effects combine such that net gain is only achievable for energies close to the LH-SO peak, while strong absorption inhibits carriers close to the band edge contributing to gain.

To analyse the effect of IVBA losses on potential device performance, the transparency carrier densities reported in section 5.3.2.2 are recalculated for 8-20% Sn content alloys at RT with IVBA included. Fig. 5.19 illustrates that increasing Sn content to improve the levels of directness no longer necessitates improved transparency carrier densities when IVBA is accounted for. Instead, n_{tr} is minimised for Sn contents around 17% in unstrained material. This occurs since the bandgap energy aligns with the spectral minimum between the LH-SO and HH-SO IVBA peaks. In practice, higher current densities will be required to overcome other optical losses such as mirror loss and carrier scattering processes. Since the peak blue shifts with increasing injection, a nominally higher Sn fraction would be needed to realign the gain peak with the IVBA minimum.

It should be noted that gain and IVBA are inherently linked, since IVBA will deplete holes from the valence band which could otherwise have contributed to gain. Full modelling of the phenomena would therefore require knowledge of the inter-band and intervalence

band transition rates. The calculation illustrated in this work is typical of historical presentations of the phenomena in literature, and accurately represents the IVBA loss and interband gain assuming fixed Fermi-levels.

5.5 Auger-Meitner Resonances

To raise the maximum operational temperature in devices, primary emphasis has currently been placed on increasing material directness by either incorporating tensile strain through various means, or by pushing for higher Sn content layers. However, the impact of doing so on other aspects of the bandstructure is often neglected, resulting in device heterostructures that are more challenging to fabricate yet do not yield significant performance improvements [22].

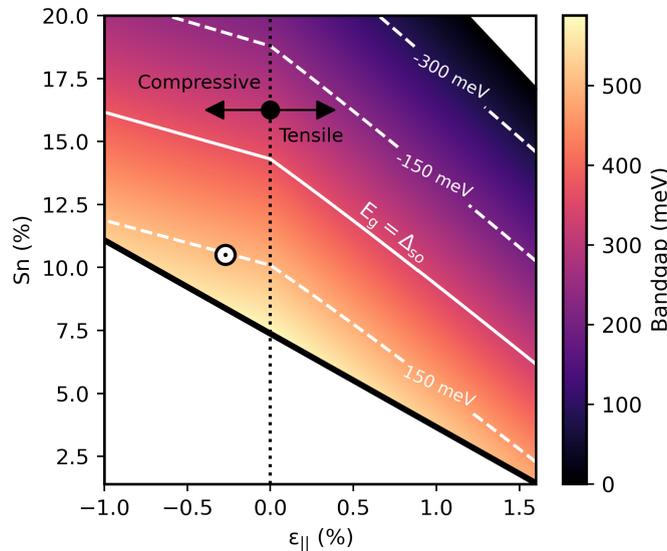


FIGURE 5.20: The bandgap as a function of Sn content and bi-axial strain. Compositions producing negative or indirect bandgaps are neglected and shown in white. White contours depicts the difference between the bandgap and spin-orbit splitting energies ($E_g - \Delta_{SO}$). The white marker indicates the experimentally investigated devices.

In previous sections, the effects of strain and Sn content on the VB degeneracy and carrier effective masses have been discussed with regards to indirect valley occupation, and therefore their qualitative effects on performance. However, the effects on the spin-orbit splitting energy, Δ_{SO} , relative to the bandgap energy has not been explored. This metric is of paramount importance to device performance since the SO band is energetically removed from the band edge, resulting in near 100% electron occupation under injection. As such, transitions between the SO band and the LH and HH bands can result in large optical and/or carrier losses through CHSH Auger-Meitner recombination.

Fig. 5.20 depicts the bandgap of GeSn as a function of Sn content and strain at RT. The coloured area denotes compositions where a direct, positive bandgap is obtained. White contour lines illustrate the SO splitting energy with respect to the bandgap, $E_g - \Delta_{SO}$, and the investigated device is represented by the white, circular marker. It is evident that the increase in Sn content and/or tensile strain required to improve material directness results in a shift towards a CHSH resonance where E_g and Δ_{SO} become equal. However, for higher Sn content and/or tensile strain, the resonance can be avoided. This comes at the cost of a greatly reduced bandgap which is known to promote CHCC Auger-Meitner recombination [12].

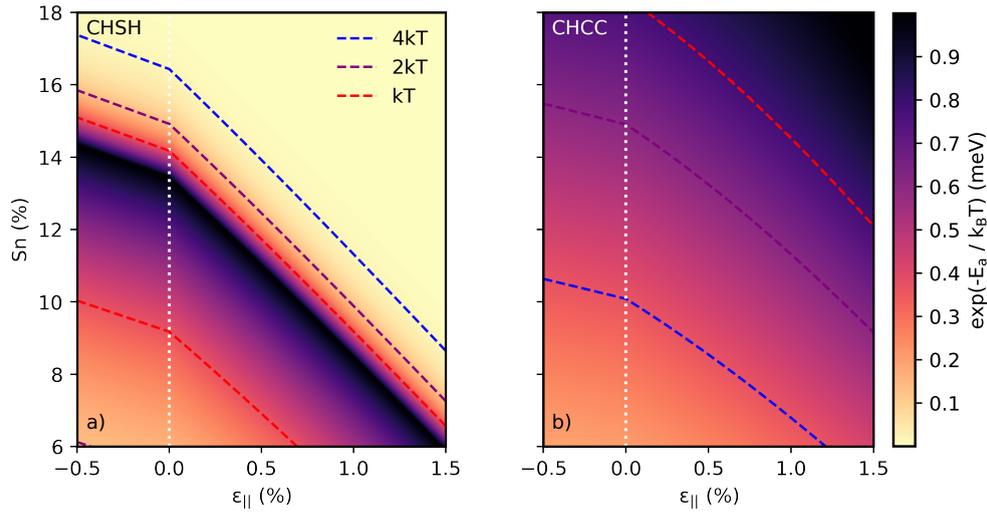


FIGURE 5.21: The exponent scaling the cubic-like carrier densities in Eqns. 5.21 and 5.22 as a function of strain and Sn content at RT for the a) CHSH and b) CHCC processes. Dashed lines depict contours where the Auger-Meitner activation energy is kT, 2kT or 4kT, respectively.

To evaluate the trade off between CHSH and CHCC recombination, the activation energies for both processes are calculated as a function of strain and Sn composition using the simple models outlined in section 2.4.3. Since the Auger-Meitner current densities are given as

$$J_{chsh} = C_{0,chsh} p^2 n \exp(-E_{a,chsh}/k_B T), \quad (5.21)$$

$$J_{chcc} = C_{0,chcc} p n^2 \exp(-E_{a,chcc}/k_B T), \quad (5.22)$$

the activation energies are used to determine $\exp(-E_a/k_B T)$ for both processes at RT, illustrated in Fig. 5.21 a) and b), respectively. Dark regions denote compositions where the right most term in Eqns. 5.21 and 5.22 are high, while light regions highlight compositions where this term is low, resulting in reduced Auger-Meitner current densities. For

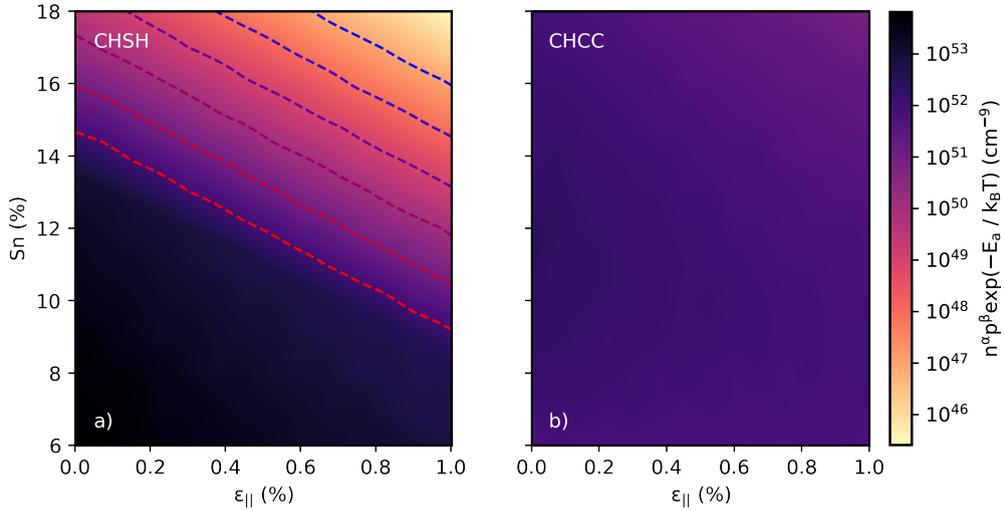


FIGURE 5.22: Eqns. 5.21 and 5.22 as a function of Sn content and increasing bi-axial tensile strain for a) CHSH and b) CHCC processes, respectively. Each dashed contour line equates to a $10\times$ reduction in the total factor (from red to blue).

the CHSH process in Fig. 5.21 a), a resonance is observed for unstrained material with $\sim 14\%$ Sn, occurring at lower Sn fractions with increasing tensile strain. Dotted contour lines represent the compositions for which the activation energies are $1\times$, $2\times$ and $4\times$ the thermal energy at RT, corresponding to values $2.7\times$, $7.4\times$ and $54.6\times$ lower than at the resonance points. However, the reduction in CHSH comes at the cost of a reduced bandgap, resulting in an increase in the CHCC exponential factor. It is therefore prudent to see whether the decreased L-valley occupation at these compositions is able to sufficiently offset the effect of activation energy reduction. To do so, the exponentials displayed in Fig. 5.21 are premultiplied by np^2 and pn^2 for CHSH and CHCC, respectively. Here, n and p are the electron and hole carrier densities calculated at transparency. In this analysis, only direct Auger-Meitner transitions are considered, and hence $n = n_\Gamma$ such that $n \neq p$. The results of the premultiplication for CHSH and CHCC are displayed in Fig. 5.22 a) and b), respectively.

It is observed that the rapidly increasing activation energy for the regime in which $E_g < \Delta_{SO}$ combines with decreased transparency carrier densities offering a pathway to significant CHSH reduction through heightened Sn content and tensile strain. Each dotted contour line in Fig. 5.22 a) depicts a $10\times$ reduction in $np^2 \exp(-E_{a,chsh}/k_B T)$. Conversely, the effect of lower transparency carrier densities acts to offset the lower activation energies for the CHCC process. The result is that $pn^2 \exp(-E_{a,chcc}/k_B T)$ is comparatively insensitive to composition, with marginal increases $\ll 10\times$ observed at the highest Sn

contents and tensile strains. As such, it can be expected that the overall Auger-Meitner currents will be reduced in devices utilising high Sn fractions and/or tensile strains.

The above analysis is dependent on a number of assumptions. Firstly, the theoretical model of Beattie and Landsburg is adopted, which assumes parabolic bands and nominally includes only direct-transitions, neglecting phonon-mediated Auger-Meitner processes [197]. Since the VBs exhibit fairly strong non-parabolicity, it is unclear the extent to which these approximations are valid. However, greater accuracy will be attained for higher Sn content and tensile strains since the required injection for transparency is lower, resulting in lower occupation of high- k states where non-parabolicity is higher. Secondly, it is assumed that we are able to operate within the framework of the Boltzmann approximation. Lastly, the values of the C_0 coefficients are not discussed during this analysis since they are dependent on the matrix elements of the given material. The remarks therefore assume that the value is less sensitive to composition than the carrier densities and activation energies.

Under these assumptions, it is expected that the employed model will underestimate the activation energies, however the general trends should be considered when designing devices. In future, full calculations of the Auger-Meitner coefficients could be conducted utilising EPM modelled bandstructure using Takeshima's Green's Function based approach [198], [199]. In doing so, the impact of IVBA and Auger-Meitner related losses could be compared such that the heterostructure design could be optimised accordingly.

5.6 Summary & Conclusions

It has been shown experimentally that defect related recombination is the dominant mechanism for raising the threshold current density in first generation, electrically injected bulk GeSn lasers. The use of high hydrostatic pressure measurements at low temperature also uncovered an appreciable leakage current pathway in which electrons occupy indirect L-valley states. Temperature dependent modelling of bulk gain was used to determine the lower boundary of L-valley leakage from a purely bandstructure perspective through analysis of the transparency carrier density. Increasing Sn content results in heightened levels of directness which decreases the required carrier density. However it is found that the fraction of indirect carriers at RT remains in excess of 10% for material with the highest Sn content (largest levels of directness), which would greatly hinder key device performance

metrics. Incorporation of small tensile strains may provide a pathway for comparably minor improvements, but at the cost of significantly more complicated heterostructure designs for typical FP lasers.

By including the effects of IVBA, it is found that increasing Sn content does not lead to improved transparency carrier densities due to strong LH-HH absorption in small bandgap materials. It should be noted that a balance must be struck between large Auger-Meitner currents and strong IVBA when choosing an optimal active region Sn content for a bulk device. Current estimates show Sn fractions around 16-17% may be optimal due to decreased IVBA, however, knowledge of typical loss values and the Auger-Meitner coefficients would allow this value to be more accurately determined. Optical loss values could be probed experimentally using gain measurements [200], or by measuring the differential quantum efficiency as a function of cavity length [201]. Auger-Meitner coefficients can be probed theoretically [198], [199] in conjunction with standard lifetime measurements using layers with high crystalline quality.

Chapter 6

Recombination Processes in Epitaxially Grown GaInAsSb QW Lasers on Si

6.1 Introduction

Type I based QW devices are currently able to access the 2.0 - 3.5 μm spectral range under continuous wave (CW) operation at RT [8]. Such devices are based on the quaternary AlGaAsSb and GaInAsSb alloys which are wide and narrow gap, respectively, thus forming good barrier/cladding and well materials. Typically, the whole structure, other than the wells, are grown to be lattice matched to GaSb, with only alloy compositions and layer thicknesses varied [5].

By utilising this methodology, high powered type I lasers have been reported for emission up to 2.5 μm , boasting 1.8 W output under CW pumping at RT with a 28% wall plug efficiency [6], with hundreds of mW output achieved for emission above 3 μm [7]. At this time, CW operation has been realised up to 3.44 μm [8], whilst pulsed devices can reach 3.7 μm [9]. The low threshold current densities afforded by ICLs and their apparent insensitivity to defects now makes them preferable for applications requiring emission above $\sim 3 \mu\text{m}$ [202]. However, type I QW devices operating between 2.05-2.65 μm have been reported with threshold current densities of less than 100 A/cm², putting them among the lowest of any FP edge emitting lasers [203]–[205].

Despite this, the threshold current density of these devices remain highly temperature sensitive. A literature review of devices by Eales et al. illustrates a strong increase in threshold current density and reduction in characteristic temperature T_0 (associated with heightened temperature sensitivity) for devices with wavelengths between 2 and 3.75 μm [91]. The degradation in performance with wavelength is most often attributed to 3 major contributions:

1. A reduction in the bandgap energy requires the growth of more challenging GaInAsSb and AlGaAsSb layers due to the presence of large miscibility gaps which can elicit defective material [206]. It is also widely appreciated that in doing so, the refractive index contrast is reduced leading to lower optical confinement factors, and thus less modal gain and heightened optical loss from increased overlap with the doped cladding layers [207].
2. Assuming a fixed barrier composition, strain and QW thickness, the VB offset is reduced with increasing wavelength until type II band alignment is achieved for emission above $\sim 3 \mu\text{m}$ [4]. This can lead to large hole leakage currents which are exacerbated when thermal broadening of the carrier distribution is high. Simultaneously, the CB confinement is heightened which can lead to carrier transport issues in devices with multiple QWs.
3. Strong Auger-Meitner recombination processes close to RT. It has been shown that devices emitting below 2 μm are dominated by strong CHSH recombination [208]. The Auger-Meitner recombination coefficient is minimised for emission at around 2.1 μm , producing devices with the lowest RT thresholds. Above this, CHCC and CHLH processes become dominant and act to drastically increase the Auger-Meitner coefficient further between $\sim 2.25 - 3.25 \mu\text{m}$ [91].

Epitaxial growth of such devices on Si would enable PICs and OEICs for environmental and healthcare sensing applications to be produced in a cheaper and more scalable fashion than current flip-chip bonding techniques. However, numerous growth-related issues can lead to degraded performance. These include, but are not limited to, lattice, polarity and thermal expansion coefficient mismatch, interfacial chemical reactions, and dopant diffusion issues [209], as discussed in Chapter 1. As such, it is vital to characterise the performance differences between devices grown on native III-V substrates and those on

Si, placing emphasis on any changes in the dominant carrier recombination pathways to help inform future growth.

In this study, temperature and pressure dependent characterisation techniques are employed to analyse devices operating $\sim 2.3 \mu\text{m}$ on Si and GaSb substrates. ‘2+6’ band k-p modelling of the bandstructure and carrier distribution is then utilised to support experimental findings, which suggest Shockley-Read-Hall (SRH) defect recombination, spatial leakage of holes and heightened carrier occupation of L-valley states may result in the heightened thresholds observed at RT in the devices on Si. In comparison, devices on GaSb are strongly dominated by CHCC Auger-Meitner recombination.

6.1.1 Sample Information

Devices on Si substrates were provided by the group of Professor Eric Tournié at the Université de Montpellier. The epitaxial stack is illustrated schematically in Fig. 6.1. The active region contains two 9 nm $\text{Ga}_{0.67}\text{In}_{0.33}\text{As}_{0.12}\text{Sb}_{0.88}$ QWs, separated by a 25 nm $\text{Al}_{0.25}\text{Ga}_{0.25}\text{As}_{0.02}\text{Sb}_{0.98}$ barrier. The active region is then sandwiched between 380 nm thick $\text{Al}_{0.25}\text{Ga}_{0.25}\text{As}_{0.02}\text{Sb}_{0.98}$ waveguiding layers, which is clad using 1.3 μm of $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}_{0.07}\text{Sb}_{0.93}$. The $\text{InAs}_{0.92}\text{Sb}_{0.08}$ buffer layer is employed to act as a marker during processing and is the layer through which electrons are injected [210].

A key merit of these devices is that they are grown on on-axis, (001) oriented Si with only a 0.5° residual miscut, which makes it compatible with standard CMOS fabrication processes [211]. By comparison, previous GaSb based QW lasers on Si were grown on substrates with miscuts of around 5° since this acted to reduce the density of antiphase domain boundaries [212]. In 2019 Delli et al. reported devices with defect densities of $3 \times 10^7 \text{ cm}^{-2}$ using strained AlSb/GaSb superlattices as a dislocation filter layer (DFL) on 4° miscut (001) oriented Si [213]. By contrast the upper bound for threading dislocation density in the devices reported in this body of work is estimated at only $8.5 \times 10^7 \text{ cm}^{-2}$, despite the absence of a DFL and growth on on-axis Si. Further information on the growth of these structures can be found in the initial paper by the growers [210] and the accompanying supplementary information document [214].

While nominal compositional data was provided by the growers and reported in [210], to ensure internal self-consistency of the analysis, we also opt to determine the epitaxial

structure by fitting digitised high resolution x-ray diffraction (HRXRD) data using kpy-mod. Fitting was performed by Dr Dominic Duffy due to his extensive expertise with the XRD technique. From this fitting we can make a number of statements. Firstly, the dotted lines in the inset of Fig. 6.2 denote the calculated peak position for GaSb (blue) and Si (red), assuming both are fully relaxed. It can be seen that these exactly match the experimentally observed peaks, which confirms zero residual strain in the GaSb buffers. Secondly, the occurrence of a singular peak at the GaSb diffraction angle reflects that the AlGaAsSb waveguide, cladding and barrier layers are indeed all lattice-matched to GaSb. Finally, Fig. 6.2 a) illustrates the fit using the composition data provided by the grower. It can be seen that the active region interference fringes are shifted to smaller angles, representing a smaller than expected in-plane strain (-1.2%). This potentially stems from differences in lattice constant, elastic constants or quaternary interpolation approach [77] between the programs used for fitting.

Fig. 6.2 b), shows the fit to the data using kpy-mod. In this simulation, the In content and layer thickness of the QW were allowed to vary, as well as the barrier thickness. In reality, altering the Sb and In content by the same amount would have similar effects on the lattice

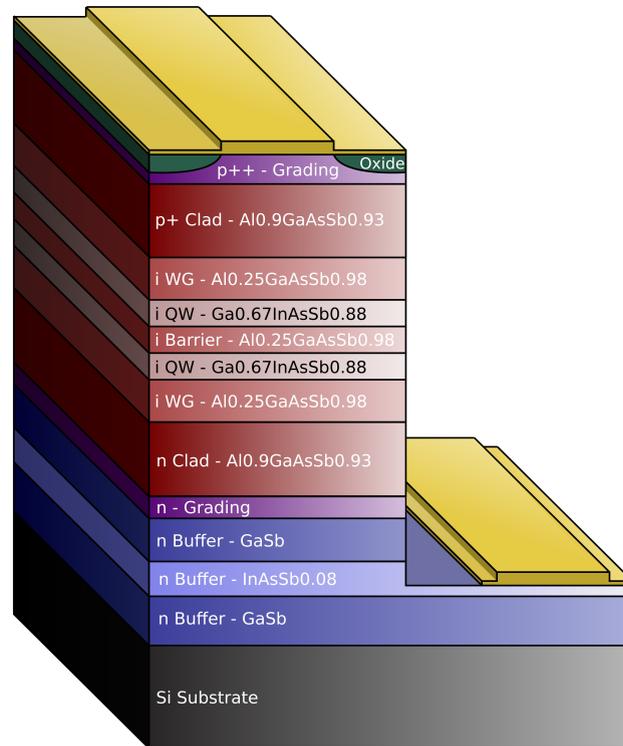


FIGURE 6.1: Schematic diagram illustrating the device layer structure and content (as reported by the grower [210]) and lateral contacting scheme, showing electron injection through the n-doped InAsSb buffer. Relative layer thicknesses are not to scale.

	QW (pre-fit)	QW (post-fit)	Barrier (pre-fit)	Barrier (post-fit)
Thickness (nm)	9	8.1	25	21.5
In (%)	33.0	38.9	-	-

TABLE 6.1

constant, and therefore the observed HRXRD spectra. Thus, due to the quaternary make-up of these layers, it is not possible to distinguish this difference using this method alone. To limit the number of fitting variables, only the In content was varied. The updated layer thicknesses and compositions used for modelling during the rest of this work are given in table 6.1. Using these compositions, the QW in-plane strain is determined to be -1.57% and excellent agreement with the experimental data is achieved.

The samples provided came in the form of two laser bars, denoted A and B, each containing a number of 100, 80, 10 and 8 μm ridge devices, separated by 80 μm n-contact pads, as illustrated in Fig. 6.3 a). The 8 μm ridges could not be tested in this work since the wire on the clip used to contact the devices was too wide. RT LI curves were collected using a probe station to identify working devices. A number of the remaining devices on both bars were not used as the bars had not cleaved cleanly and there was visible evidence of

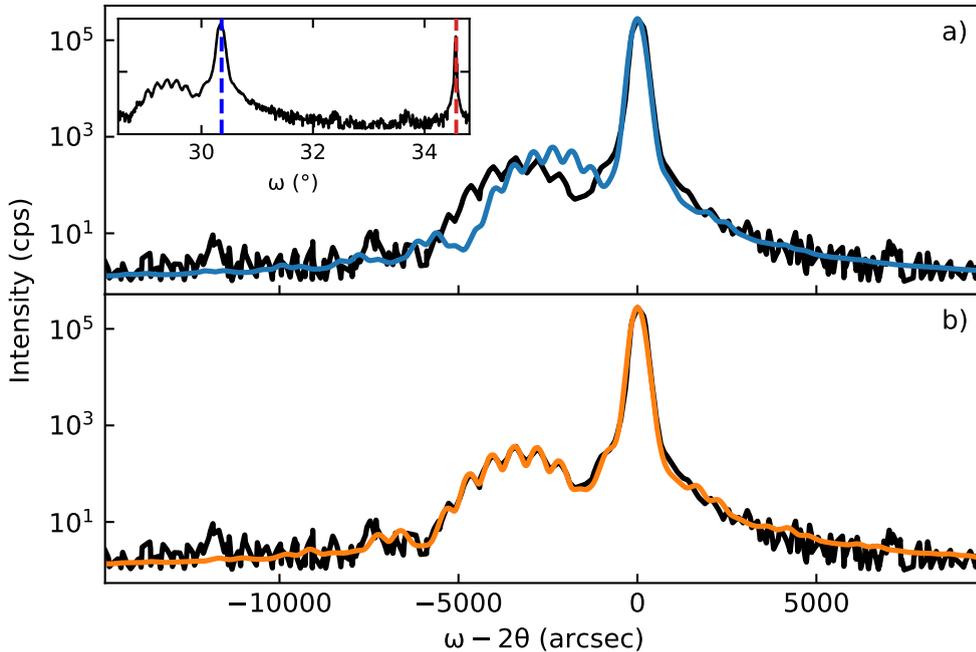


FIGURE 6.2: Digitised HRXRD data, obtained from [210], illustrated by the black lines. a) The simulated spectra using the values reported in the paper is illustrated in blue. b) The orange line represents the fitted data using the internal database.

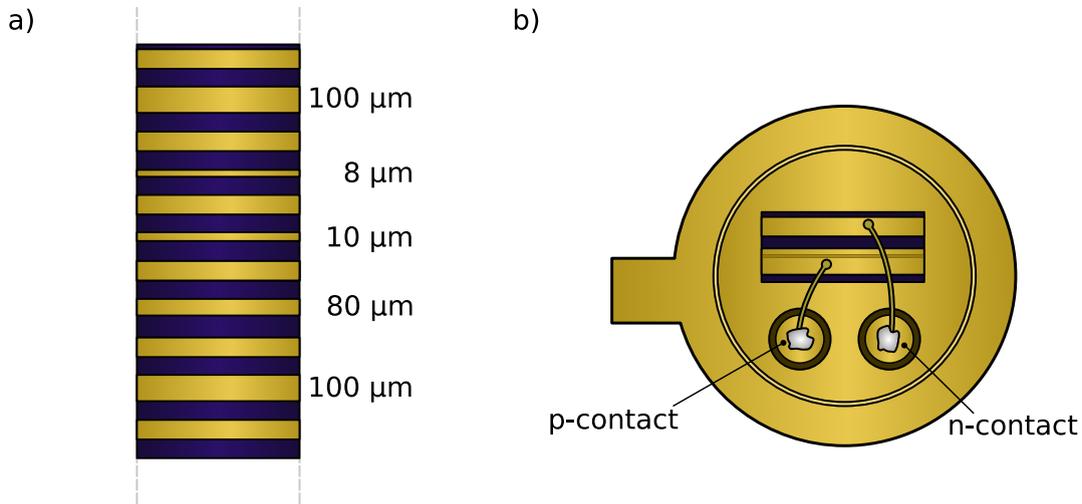


FIGURE 6.3: Schematic depictions of a) The laser bars, highlighting the layout of the p- and n-contact layers, and b) the devices pre-mounted on TO-headers.

crystallographic plane switching, resulting in irregular facets. Lasing was demonstrated in one 100 μm and one 10 μm device per bar. These are hereafter referred to as BR100A, BR10A, BR100B and BR10B, respectively. ‘BR’ denotes that the device is from a laser bar, ‘100’ and ‘10’ refer to the ridge width of the given device, whilst ‘A’ and ‘B’ highlight which bar they belong to.

Four TO header mounted devices were also obtained; two 100 μm and two 10 μm ridges. These are denoted PM100A, PM100B, PM10A and PM10B, where ‘PM’ reflects the fact that these devices are pre-mounted, ‘100’ and ‘10’ are the ridge widths and ‘A’ and ‘B’ are used to highlight the specific device of a given width.

The performance of these devices are compared to a broad area device grown on native GaSb, operating at 2.35 μm , fabricated by the group of Professor Markus Amann at Technische Universität Munchen. This device was previously reported on by former group member Dr. Eales [95]. Any data from this device was obtained by Dr. Eales in 2018 and is reproduced here with permission.

6.2 Temperature Dependent Measurements

6.2.1 Threshold Current Density

The laser performance for devices grown epitaxially on Si was initially investigated through analysis of the threshold current density as a function of temperature. LI curves were collected for three separate devices on Si and compared to the device on GaSb detailed in the previous section. Two of the devices were selected from laser bar A: a broad area $100\ \mu\text{m}$ ridge, BR100A, and a narrow $10\ \mu\text{m}$ ridge, BR10A. LI curves from these devices were measured by placing them in the Leybold closed cycle He cryostat. Here, the clip fixed the position of the laser whilst providing the p-contact. Since the devices are top contacted, the n-contact was made using silver epoxy paint between the contact pad and the metal base plate. This configuration enabled measurements between 20-300 K. In contrast, the pre-mounted, narrow $10\ \mu\text{m}$ ridge device (PM10A) was measured using the Oxford Instruments Optistat DN2 cryostat, permitting temperatures between 78-350 K. For both set-ups, emission from a single facet was focused into a bifurcated fibre using a calcium fluoride lens. The outputted signal was then split between an OceanOptics NIR 512-2.5 USB spectrometer for spectral measurements, and an Infrared Associates InSb detector for obtaining LI curves. The general set-up was as described in section 3.1.

Fig. 6.4 illustrates the raw LI curves with increasing temperature for a) PM10A, b) BR10A and c) BR100A, respectively. Whilst collecting data for PM10A, 20 K temperature steps were used between 100-300 K, with 10 K steps then used between 300-350 K. For BR10A,

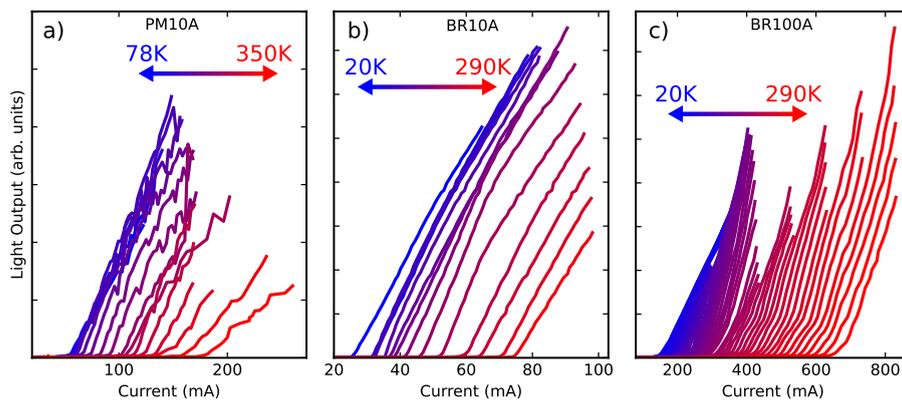


FIGURE 6.4: LI curves obtained as a function of increasing temperature from a) PM10A, b) BR10A, and c) BR100A. Measurements of PM10A were conducted using the Oxford instruments DN2 cryostat, whereas BR10A and BR100A were measured using the Leybold closed cycle He cryostat.

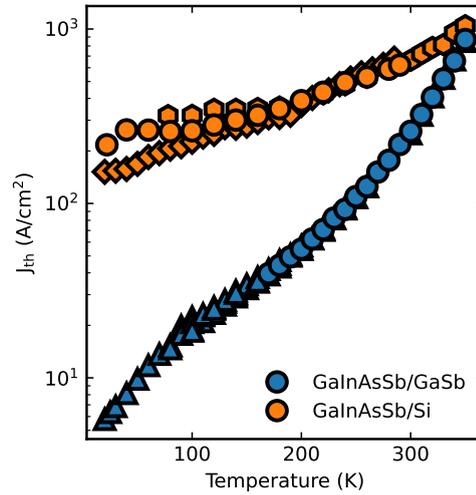


FIGURE 6.5: Threshold current density as a function of temperature for broad area devices on native GaSb (blue), two $10\ \mu\text{m}$ ridge width devices on Si (orange circles and hexagons) and a $100\ \mu\text{m}$ broad area device (orange diamonds).

a 20 K step size was used between 20-280 K, with a single data set collected at 290 K. These were typically chosen to reduce the time taken to measure each device, however for the broad area device, a step size of 5 K was utilised for the entire temperature range. The threshold currents were subsequently extracted as a function of temperature using the methodology outlined in section 3.4. These values were then converted to threshold current densities by dividing by the device area, given by the product of the cavity length and ridge width. This assumes negligible current spreading.

Fig. 6.5 shows the threshold current density, J_{th} , as a function of temperature for the three devices, along with the device on native GaSb. As initially reported by Eales, the observed threshold current in the device on GaSb is dominated by radiative recombination, which was confirmed using spontaneous emission measurements [95]. This results in devices that are relatively temperature insensitive below ~ 150 K. The resultant threshold current density is therefore very low, taking values of 5-30 A/cm², reflecting that highly temperature dependent loss processes such as Auger-Meitner and carrier leakage are well suppressed, whilst also indicating good material quality through low defect-related currents. The measured devices are broad area, with $100\ \mu\text{m}$ stripes, thereby minimising the impact of current spreading on the estimated threshold current. Above 150 K there is a so-called ‘break-point’, as defined by Sweeney et al. [215]. This indicates the temperature at which the thermally activated process provides an appreciable contribution to threshold, thereby changing the observed trend. Above this temperature, there is a rapid

increase in J_{th} , which Eales attributes to CHCC Auger-Meitner recombination [91].

As per the devices on native substrates, the threshold currents in devices on Si are relatively insensitive to temperature changes below 150 K. However, contrastingly, the threshold current density at 20 K is between 28 and 36× higher in the devices on Si. Since thermally activated processes such as Auger-Meitner recombination and leakage are significantly reduced at low temperatures, the difference in threshold can likely be attributed to strong defect-related recombination due to large defect densities in the buffer and active layers, or potentially poor carrier injection efficiency stemming from the use of lateral injection. Above 150 K, the devices on Si show lower temperature sensitivity compared to those on GaSb such that the threshold current densities become comparable at 350 K, despite the large difference observed at low temperature. Both broad area and narrow ridge devices on GaSb elicit comparable threshold current densities. If current spreading was appreciable in these devices, the 10 μm ridges would overestimate J_{th} , causing disparity with the 100 μm stripe lasers. As such, throughout this work it is assumed that lateral current spreading is negligible. The focus of this chapter is then to fully characterise the difference in behaviour between the devices on Si and those on native GaSb.

6.2.2 Spectra

Temperature dependent spectral measurements of BR10A and PM10A are illustrated in Fig. 6.6 a) and Fig. 6.7 a), respectively. Spectra were captured at different current values above threshold which, paired with the standard temperature-dependent change in external quantum efficiency, explains the observed difference in peak intensities. The peak energies were subsequently extracted as a function of temperature and plotted in Fig. 6.6 b) and Fig. 6.7 b).

The Varshni relation presented in Eqn. 4.75 was subsequently fitted to both data sets using a least-squares approach. Since the emission is multi-modal, all peaks were plotted at each temperature. As such, the fit represents an effective average emission energy at each temperature. For PM10A, the fitted parameters are given by: $E_g(0) = 0.613 \pm 0.001$ eV, $\alpha = (3.6 \pm 0.2) \times 10^{-4}$ eV/K and $\beta = 180 \pm 30$ K. For BR10A, the parameters are: $E_g(0) = 0.612 \pm 0.001$ eV, $\alpha = (4.2 \pm 0.5) \times 10^{-4}$ eV/K and $\beta = 240 \pm 60$ K. Thus, the fitting parameters for both devices agree within the fit uncertainty.

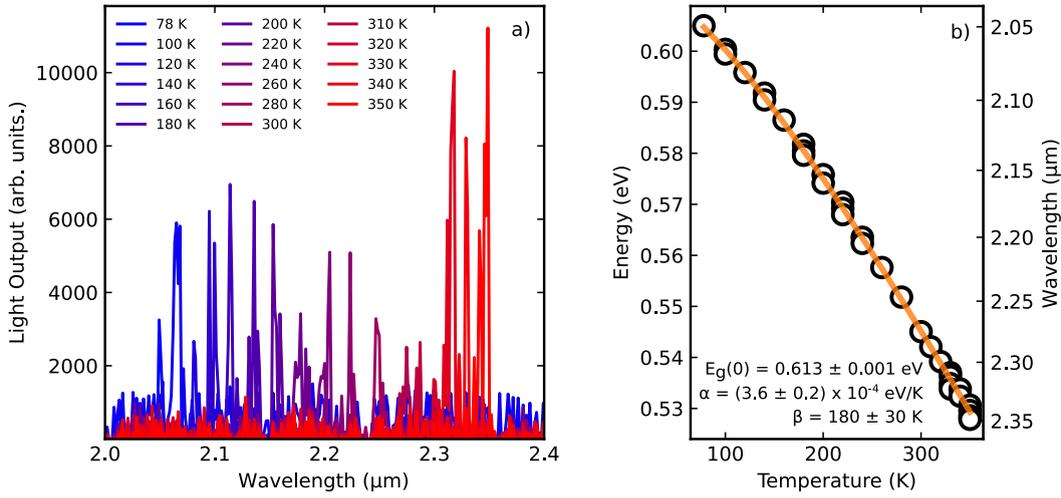


FIGURE 6.6: a) Spectral measurements of BR10A taken above threshold as a function of increasing temperature (red to blue). b) The Varshni curve and coefficients obtained from the peak positions as a function of temperature.

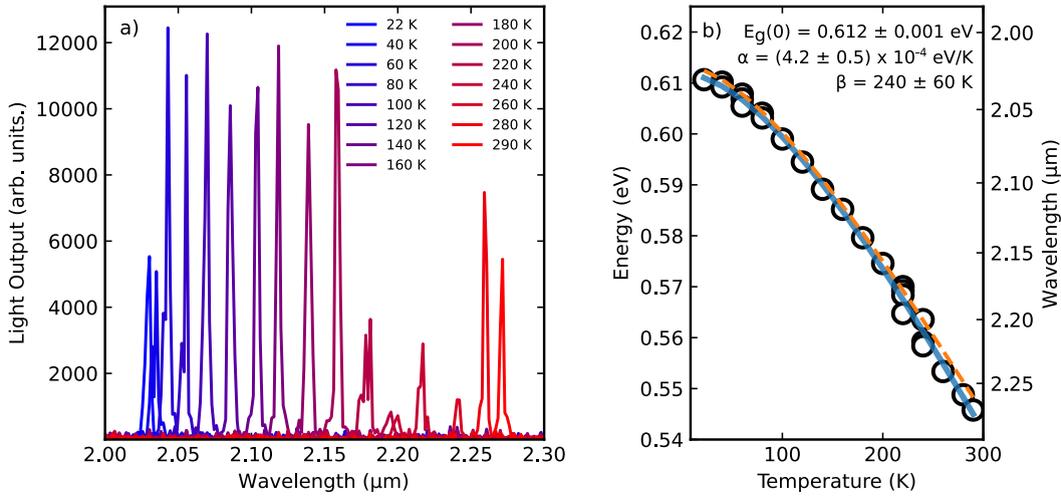


FIGURE 6.7: a) Spectral measurements of PM10A taken above threshold as a function of increasing temperature (red to blue). b) The Varshni curve (blue) and coefficients obtained from the peak positions as a function of temperature. The orange dashed line shows the Varshni fit from BR10A.

Fig 6.8 illustrates the Varshni fit to the data for PM10A which is plotted with the experimentally modelled transition energies using the layer parameters provided by the grower (red) and those determined from the refitted XRD data (green). It can be seen that the modelled Varshni trend, obtained using the ‘2+6’ band k-p model, accurately replicates the temperature dependence of the emission energies, but shifted to energies ~ 10 meV higher. The minor difference is likely due to a combination of blue-shift from band-filling effects and red-shift from band-gap renormalisation [216], uncertainty in reported ternary

and quaternary bowing parameters, or could stem from inaccuracies in the reported active region compositions. It can be seen that using the compositions obtained from the XRD fit conducted in this work produces a better estimation of the emission energy. This again reinforces that such values may be more representative of the growth. These small alterations in composition will have the greatest impact on the observed emission energy but comparatively minor impact of the temperature and pressure dependent modelling in this chapter. As such, the devices are modelled using the values reported in Table 6.1.

6.2.3 Characteristic Temperature

The near-exponential increase in threshold current density with temperature can be simply approximated as [217]

$$J_{th}(T) = J_0 \exp\left(\frac{\Delta T}{T_0}\right) \quad (6.1)$$

where J_0 is the reference current density, ΔT is the temperature difference between J_{th} and J_0 . T_0 is the so-called characteristic temperature which by definition is the temperature change required to raise J_{th} by a factor e . Therefore, this quantity provides a measure of the threshold current density's sensitivity to temperature, with a higher T_0 correlating to improved thermal stability. Although the characteristic temperature itself is inherently

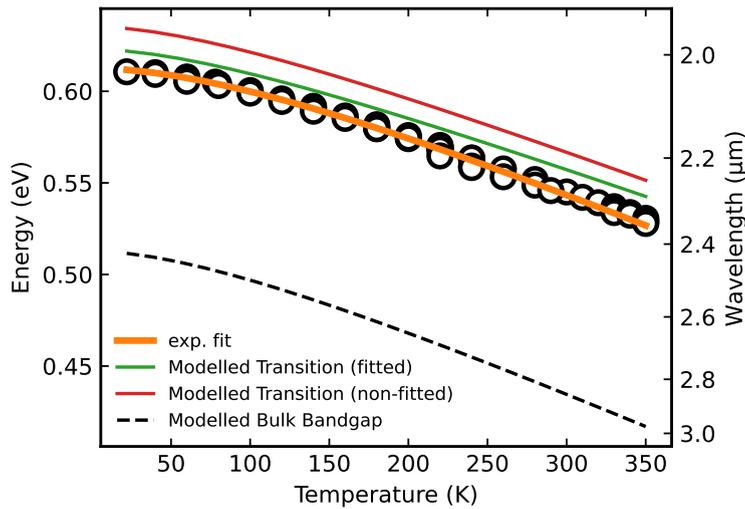


FIGURE 6.8: The Experimental Varshni fit (orange), compared to the modelled transition energy obtained from a k-p model. The red line shows the expected trend when using the heterostructure composition provided by the grower, whilst green is obtained using the adjusted values subsequent to XRD fitting.

dependent on temperature, it can be expressed over a small range as [218]

$$\frac{1}{T_0} = \frac{1}{J_{th}} \frac{dJ_{th}}{dT} = \frac{d \ln(J_{th})}{dT}. \quad (6.2)$$

Invoking the standard ABC model outlined in section 2.4, it has been previously shown that the characteristic temperature of defect-related, radiative and Auger-Meitner recombination in an ideal quantum well can be given by [219]

$$T_{0, rad}(T) = T, \quad (6.3)$$

$$T_{0, def}(T) = \frac{2T}{3}, \quad (6.4)$$

$$T_{0, aug}(T) = \frac{T}{3 + E_a/k_B T}, \quad (6.5)$$

respectively, where E_a is the Auger-Meitner activation energy. This model assumes that loss and gain parameters are not temperature dependent. It should also be noted that spatial and $|\mathbf{k}|$ -space leakage currents are neglected from this analysis. Whilst an expression for spatial leakage currents have been derived in literature [12], the resulting equation is a high-order polynomial requiring a number of assumptions. Estimating values of constants in this expression would likely lead to a non-physical model for the process which would greatly hinder interpretation. However, it is known that all leakage processes are highly temperature-dependent, and as such, a low characteristic temperature akin to that of Auger-Meitner dominated currents may be postulated. Leakage currents will be quantitatively explored during the interpretation of pressure dependent measurements and temperature- and pressure-dependent modelling later in the chapter.

The blue markers in Fig. 6.9 illustrate the temperature dependence of T_0 for devices on native GaSb, as reported by Eales [95]. At low temperatures, it can be seen that T_0 follows a trend indicating that the threshold current is dominated by a mix of radiative and defect-related recombination. At ~ 140 K there is a break-point, illustrating the onset of a highly-temperature dependent process which acts to drastically reduce T_0 . This was attributed to Auger-Meitner recombination by the author. However, the trend continues well below the expected relation for Auger-Meitner dominated currents, indicated by the solid red line. This could be the result of an optical loss process with appreciable temperature dependence, which is neglected during this analysis, but would act to further reduce the expected T_0 relation for each process [95]. Else it may indicate that another mechanism with high temperature sensitivity, such as carrier leakage, may be contributing.

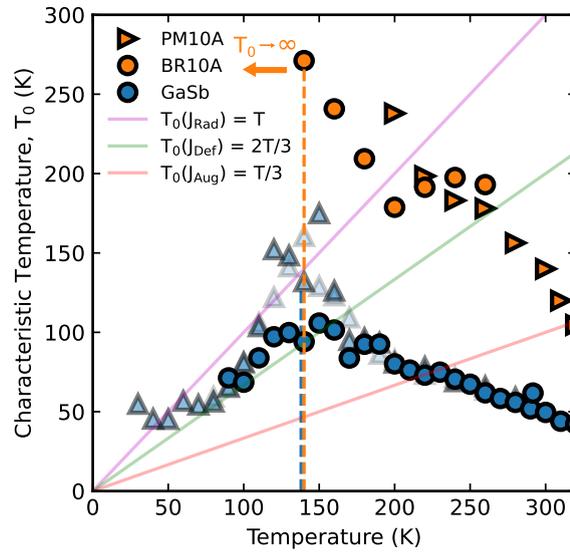


FIGURE 6.9: Variation in T_0 with temperature for devices on GaSb (blue) and Si (orange). Pink, green and red solid lines represent the expected behaviour when the current is dominated by radiative, defect or Auger-Meitner recombination, respectively. Dotted lines denote break-point behaviours where the observed trend changes in each device type.

The orange markers in Fig. 6.9 illustrate the T_0 values obtained using a 5-point moving average of the threshold current density with temperature data from PM10A. Interpretation of T_0 is complicated by the large threshold currents at low temperature. Whilst the obtained values of dJ_{th}/dT are comparable to the devices on GaSb, the low temperature thresholds result in artificially inflated T_0 values in excess of 500 K, which can be misinterpreted as heightened stability. A break-point is again observed in these devices at ~ 140 K, falling in-line with the expected trend for Auger-Meitner dominated currents at 320 K. However, since data is not obtained for higher temperatures it is not possible to ascertain whether it then follows this trend or reduces further, as per the devices on GaSb.

As such, the difference in low temperature threshold behaviour for devices on Si and GaSb makes direct comparison of T_0 difficult outside of general qualitative statements. A different technique is therefore required to provide quantitative analysis of recombination mechanisms in these devices.

6.2.4 Spontaneous Emission

One method for obtaining further information about the radiative recombination component and the dominant recombination pathways is to measure the spontaneous emission (SE) as a function of current and temperature. To ensure the effects of gain and loss are filtered from the collected emission, measurements of SE are typically conducted by collecting light along the growth direction. Since Si is nominally non-absorbing for wavelengths above $\sim 1.2 \mu\text{m}$ and the devices are top contacted, light could be collected directly through the substrate without any extra processing steps.

Due to issues with alignment to the $10 \mu\text{m}$ stripes and the necessity of using a non-mounted device, SE was only collected for the broad area device BR100A. Here, the ridge was aligned with a hole in the copper base plate of the laser mount, which housed a thin fibre. In this measurement, the fibre was placed onto the window of an Infrared Associates InSb detector.

Fig. 6.10 a) depicts the LI curves for SE collected as a function of increasing temperature, with white markers denoting the light output at threshold. A super-linear trend is observed below threshold at lower temperatures, which is again indicative of defect dominated currents. Fig. 6.10 b) shows the normalised light output as a function of temperature for fixed currents of 25 and 50 mA, respectively. In both cases, the spontaneous

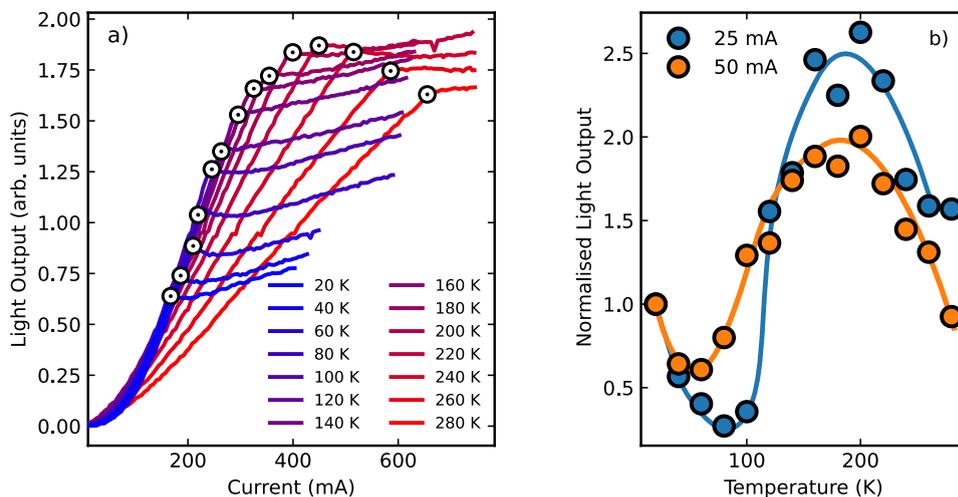


FIGURE 6.10: SE curves as a function of increasing temperature (blue to red) for BR100A. White markers are used to highlight the threshold current at each temperature. b) Normalised light output as a function of temperature at 25 mA (blue) and 50 mA (orange). Lines are a guide to the eye.

emission initially decreases with temperature, since the defect current dominates. The radiative current is known to increase linearly with temperature, resulting in increased light output between $\sim 80 - 180$ K, before the SE decreases once more as a thermally activated process such as Auger-Meitner or leakage begins to dominate above 200 K.

The curves show strong pinning of the SE above threshold. This suggests that the quasi-Fermi levels are fixed above threshold. Typically this is attributed to the rate of carrier injection matching the stimulated emission rate above threshold, leading to a fixed carrier density. However, it should be noted that the CB quasi-Fermi level may appear fixed despite an increasing carrier density if excess carriers are occupying L-valley states, due to the large DoS. These indirect carriers would not recombine radiatively, therefore leaving the SE pinned. Similarly, if the VB DoS in the well is low, significant hole leakage may occur at higher temperatures resulting in the VB quasi-Fermi level pinning close to the barrier VB edge where a near continuum of unconfined states exist.

6.2.4.1 Z-Analysis

Since the spontaneous emission L_{spont} is proportional to the radiative recombination rate, Bn^2 , the ABC model may be written in terms of the SE as

$$I \propto a' L_{spont}^{1/2} + b' L_{spont} + c' L_{spont}^{3/2} \propto \left(\sqrt{L_{spont}} \right)^Z, \quad (6.6)$$

where a' , b' and c' are unknown parameters. Here, the primes are used to differentiate the parameters from the standard constants of the ABC model. As such, if defect, radiative or Auger processes are dominant, Z will take values of 1, 2 and 3, respectively. By rearranging for Z , this becomes

$$Z = \frac{d \ln(I)}{d \ln(\sqrt{L_{spont}})}, \quad (6.7)$$

illustrating that Z can be obtained from the SE data. Fig. 6.11 a) illustrates an example plot of $\ln(I)$ vs $\ln(\sqrt{L_{spont}})$ at 120 K. The Z value at threshold is denoted Z_{th} and is obtained from the gradient using the first 10 data points below threshold. This process is repeated for the remaining data sets and the resulting Z -values are plotted as a function of temperature in Fig. 6.11 b).

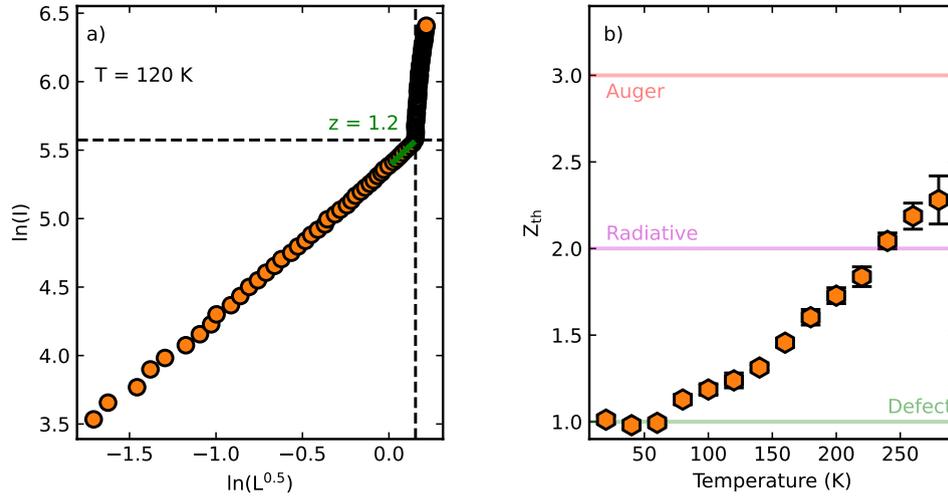


FIGURE 6.11: a) plot of $\ln(I)$ vs $\ln(L_{spon}^{1/2})$ for data obtained from BR100A at 120 K illustrating how Z_{th} is obtained. b) The temperature dependence of the obtained Z_{th} values, reinforcing defect dominated current at low temperature.

At the lowest temperatures, $Z_{th} = 1$ reflecting that the threshold current is dominated by defect related recombination. This supports the postulates made from the temperature dependence of J_{th} in section 6.2. Above 60 K, Z_{th} takes non-integer values and slowly rises with increasing temperature. At such low temperatures any thermally activated processes such as Auger-Meitner or L-valley leakage are expected to be suppressed. As such, the small rise is likely the result of an increasing radiative current component with temperature which is an expected trend [220].

At RT, Z_{th} increases to 2.25. Interpretation of this result is complicated since such a value could indicate a radiatively dominated device with a smaller contribution from higher order processes, or it may be that there remains a significant defect related current at RT and another thermally activated pathway with a large dependence on carrier density becomes prominent. Since L_{spon} at threshold is seen to decrease with temperature above 200 K, the radiative current must be dropping whilst the overall threshold current is increasing. As such, the fractional contribution from radiative recombination is reducing, and thus, the latter interpretation is most plausible.

Due to the simplistic nature of the ABC model invoked in this analysis, spatial and k-space leakage currents are not included explicitly. However, it is possible that the raised Z_{th} value at high temperature is the result of either Auger-Meitner recombination, spatial hole leakage or L-valley leakage, since all are expected to be highly dependent on the

carrier density and temperature. Using the temperature dependent techniques discussed here, it is not possible to distinguish between these effects.

6.3 Pressure Dependent Measurements

Whilst temperature dependent characterisation techniques were able to highlight the effect of large defect densities on the device performance at low temperatures, it was not able to distinguish the dominant recombination process at high temperature. Consequently, a second technique was required to separate the effects of L-valley leakage, hole leakage and Auger-Meitner recombination processes. In an attempt to do this, RT high-hydrostatic pressure measurements were conducted.

The general set-up of the pressure cell, pulse generator, oscilloscope and lock-in amplifier for RT measurements are as described previously in section 3.3. Emission through the cell window was collected into a bifurcated optical fibre with one output focused onto an Infrared Associates GaSb detector and the other connected to an OceanOptics NIR 512-2.5 USB spectrometer. LI curves and spectra were then obtained for steps between 0-7 kbar, this constitutes the range over which the cell could hold pressure before minor leaks began to occur. Measurements were obtained whilst increasing and then subsequently decreasing the pressure to assess repeatability.

6.3.1 Spectra

Fig. 6.12 a) and Fig. 6.13 a) illustrate the spectra of PM10A during two separate measurement runs. It should be noted that, as per the temperature-dependent measurements, the current was set to a nominal value above threshold to capture the spectral data. This, paired with changes in radiative efficiency, explains the variation in peak intensity. However, in this case, only isolation of the peak energies is required. Spectral data was obtained while increasing and decreasing the pressure to assess whether the process was repeatable.

Fig. 6.12 b) and Fig. 6.13 b) depict the extracted peak energies as a function of increasing pressure. For both runs, the change in emission energy with pressure and the emitted wavelength for a given pressure remain the same within the quoted uncertainty. This reinforces that the process is fully reversible in tuning the bandstructure, and affirms

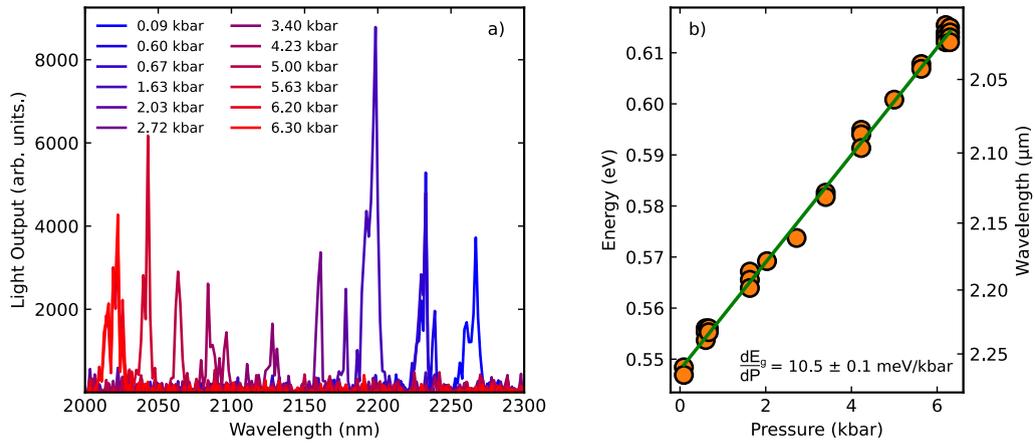


FIGURE 6.12: a) Spectral measurements of PM10A as a function of increasing pressure (blue to red) during an initial run. b) The emission energy as a function of pressure, allowing the pressure coefficient to be obtained from the gradient.

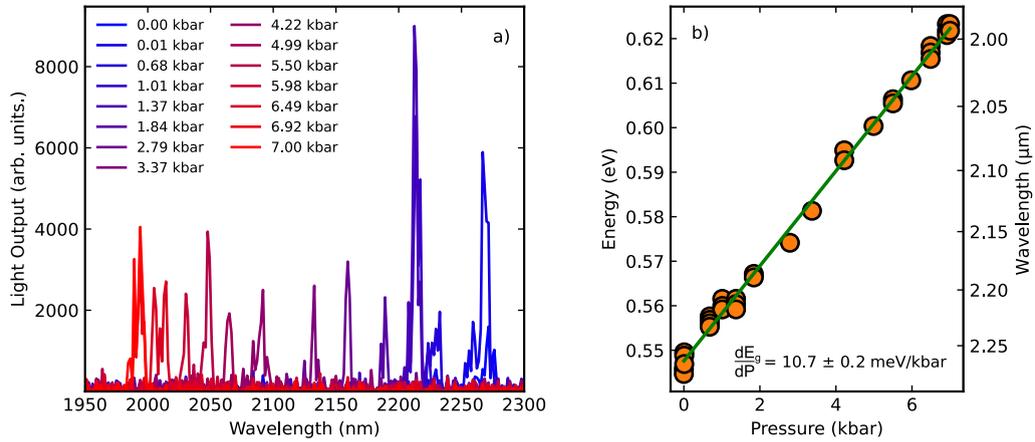


FIGURE 6.13: a) Spectral measurements of PM10A as a function of increasing pressure (blue to red) during a second run. b) The emission energy as a function of pressure, from which the pressure coefficient is obtained.

that any discrepancies in repeated pressure measurements are not due to variations in the pressure itself. The averaged pressure coefficient of $10.6 \pm 0.1 \text{ meV/kbar}$ is in line with the general empirical rule which states that the Γ -minima in III-V material shifts at a rate of 10-15 meV/kbar [110].

6.3.2 Threshold Current Density

Pressure dependent LI curves were obtained twice for PM10A and once for PM10B. The threshold current density at each pressure was extracted from the plot using the methodology outlined in section 3.4. Fig. 6.14 shows the variation in threshold current density

with pressure for a) PM10A (run 1), b) PM10A (run 2), c) PM10B and d) the device on GaSb, normalised to their respective values at atmospheric pressure. Solid markers represent measurements taken whilst the pressure was being increased, while faded markers indicate values obtained with decreasing pressure.

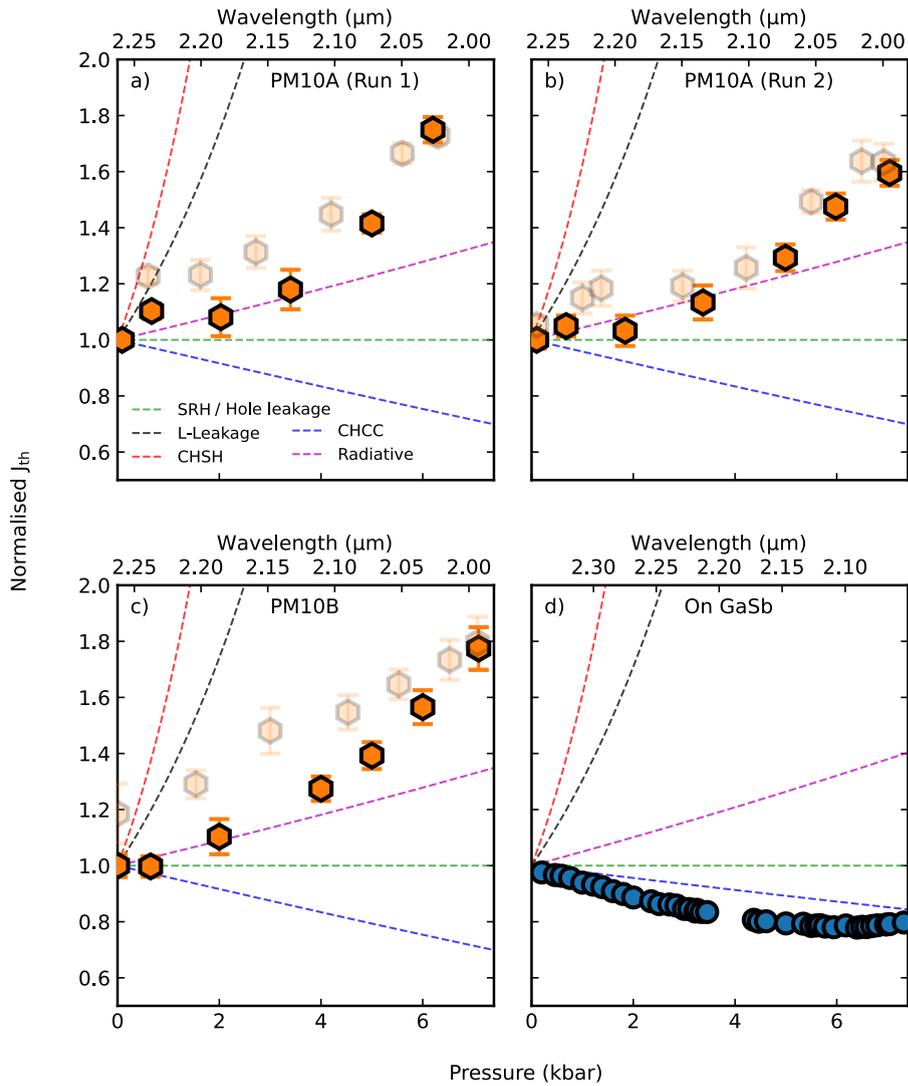


FIGURE 6.14: The threshold current density as a function of pressure for a) PM10A (first run), b) PM10A (second run), c) PM10B and d) the device on GaSb. Data are normalised to their respective values at atmospheric pressure. Red, black, pink, green and blue dashed lines show the trends for devices dominated by CHSH, L-leakage, radiative, defect/hole leakage and CHCC currents, respectively.

It is evident for PM10A (run 1) (a) and PM10B (c), that the thresholds are higher than expected when the pressure is lowered, such that the current at atmospheric pressure does not return to its pre-measurement value. This discrepancy may be the result of a number of factors such as changes to the contacts, facets or even the introduction of defects owing to differences in the lattice constants of III-V and group IV materials with

pressure. However, it can be seen in Fig. 6.14 b) that the effect is greatly reduced on subsequent runs over the same pressure range. This suggests that the degradation is predominantly introduced during the first run, and subsequent application of pressures within the same range will have comparatively little effect. As such, this data set is used for any subsequent fitting.

With this in mind, we return to the analysis of trends in Fig. 6.14. For devices on native GaSb, illustrated in Fig. 6.14 d), the threshold current density initially decreases with pressure. Such behaviour is characteristic of CHCC Auger-Meitner recombination. A turning point is observed at $\sim 2.12 \mu\text{m}$ indicating the onset of a secondary process, previously attributed to the onset of appreciable CHSH Auger-Meitner recombination [95].

In contrast, the threshold current density is shown to increase constantly with pressure for devices on Si. Such behaviour is indicative of either CHSH Auger-Meitner recombination or increasing L-valley occupation. Eales showed that regardless of strain and heterostructure design, CHSH was dominant in devices operating below $2.0 \mu\text{m}$ whilst CHCC was dominant above $\sim 2.1 \mu\text{m}$ [91]. It is therefore less likely that the observed increase is purely related to CHSH Auger-Meitner processes in these devices. If the fractional contribution of this is assumed to be negligible at atmospheric pressure, a fit to the data may be obtained using

$$\frac{J_{th}(P)}{J_{th}^0} = \frac{J_{def}(P)}{J_{def}^0} F_{def}^0 + \frac{J_{rad}(P)}{J_{rad}^0} F_{rad}^0 + \frac{J_{hole}(P)}{J_{hole}^0} F_{hole}^0 + \frac{J_L(P)}{J_L^0} F_L^0 + \frac{J_{chcc}(P)}{J_{chcc}^0} F_{chcc}^0, \quad (6.8)$$

where subscripts *def*, *rad*, *hole*, *L* and *chcc* denote SRH defect recombination, radiative recombination, hole leakage, L-valley leakage and CHCC Auger-Meitner recombination, respectively. A superscripted '0' marks values at atmospheric pressure. The defect related current is expected to be pressure independent, whilst the radiative current for a QW based device is proportional to E_g^2 [190]. Since the pressure coefficients of the barrier and well materials are nominally the same, the VB offset is maintained such that the hole leakage term is also independent of pressure. As per the previous chapter, the L-valley leakage current is proportional to $\exp(-dE_a^L/dP \cdot P/k_bT)$ where the activation energy is given by the different between the L and Γ minima, $E_a^L = E_L - E_\Gamma$. CHCC current has the same pressure dependence, however the activation energy is given by $E_a^{chcc} = E_g \cdot m_c / (m_c + m_h)$, where m_c and m_h are the in-plane effective masses of the conduction and heavy hole bands, respectively. Invoking these relations, Eqn. 6.8 can be

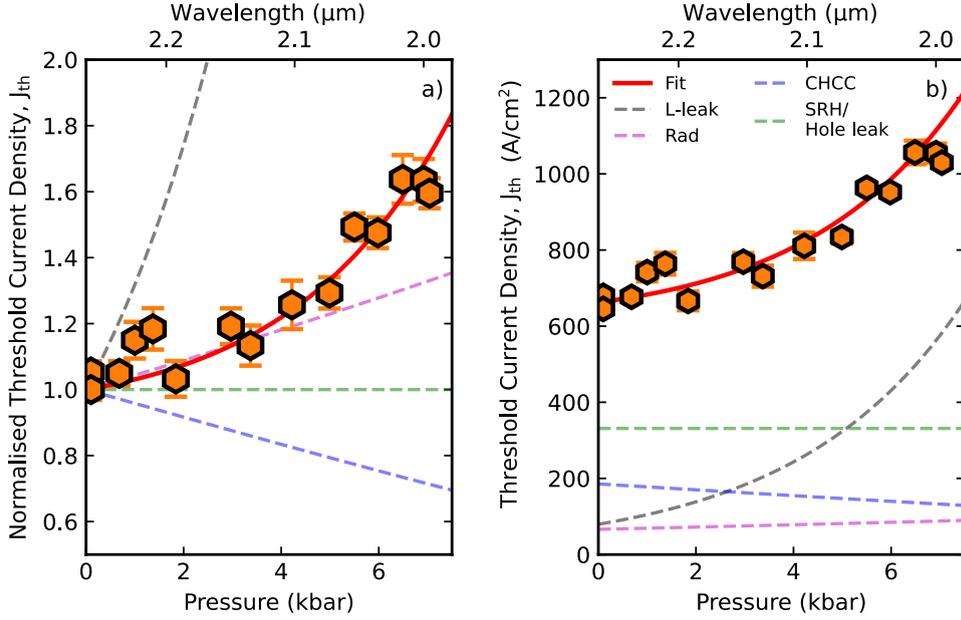


FIGURE 6.15: a) The threshold current density as a function of pressure for PM10A (run 2). Dashed black, pink, green and blue lines represent the pressure dependencies of L-leakage, radiative, SRH defect/hole leakage and CHCC currents, respectively. The solid red line shows the fit to the data. b) Non-normalised data, with dashed lines showing the weightings of each process as a function of pressure.

simplified to

$$\frac{J_{th}(P)}{J_{th}^0} = F_{ind}^0 + \frac{E_g(P)^{3/2}}{E_g(0)^{3/2}} F_{rad}^0 + \exp\left(-\frac{dE_a^L(P)}{dP} \frac{P}{k_B T}\right) F_L^0 + \exp\left(-\frac{dE_a^{chcc}(P)}{dP} \frac{P}{k_B T}\right) F_{chcc}^0 \quad (6.9)$$

where F_{ind}^0 is the cumulative fraction of threshold current attributed to the pressure independent defect and hole leakage processes.

Boundaries for the fitting parameters were implemented to reduce the number of possible solutions. For example, F_{rad}^0 was allowed to vary between 0.05 and 0.25 based on simulations of the radiative current density assuming an optical loss ranging from zero to 10 cm^{-1} . This represents a conservative estimate based on the experimentally determined values of 2-6 cm^{-1} reported by the growers [214]. The minimum bound of F_{ind}^0 was based on the current density at 20 K, which was shown to be dominated by defects. It is not expected that this contribution would lessen with temperature, and thus the minimum bound is determined by $J_{th}(20 \text{ K})/J_{th}(300 \text{ K}) \approx 0.2$. Since F_{ind}^0 also encapsulates contributions from hole leakage, which would worsen with increasing thermal broadening, it is possible that F_{ind}^0 may be considerably higher than 0.2 at RT, should hole leakage be significant. As such, the upper limit was left at 1. As CHCC recombination was shown

to be dominant in the devices on native substrates at a similar wavelength, it can be expected that it should be appreciable for the devices on Si. However, the magnitude of this is unknown so F_{chcc}^0 was allowed to vary between 0 and 1. The fit can then be conducted using only these three variables since $F_L^0 = 1 - F_{ind}^0 - F_{rad}^0 - F_{chcc}^0$, by definition, which was implemented as a condition for the least-squares fit.

The obtained fractional contributions to threshold at atmospheric pressure are $F_{ind}^0 = 0.50$, $F_{rad}^0 = 0.10$, $F_{chcc}^0 = 0.28$, and $F_L^0 = 0.12$, with the fit illustrated in Fig. 6.15 a). It can be seen that a reasonable fit to the data is attained. The relative contributions of each process as a function of pressure are more easily seen in Fig. 6.15 b). It should be noted that there is large uncertainty in the values of F_{chcc}^0 and F_{ind}^0 due to the similarities of their pressure dependencies and the scatter in the data. However, F_{chcc}^0 cannot be greater than F_{ind}^0 as no decrease in J_{th} is observed over the first 2 kbar. In this situation, the threshold current at RT and atmospheric pressure for devices on Si is dominated by a combination of large defect and/or hole leakage currents, followed by significant CHCC recombination. While, in contrast to the devices on GaSb, leakage of carriers to L-valley states could be non-negligible, representing $\sim 12\%$ of threshold.

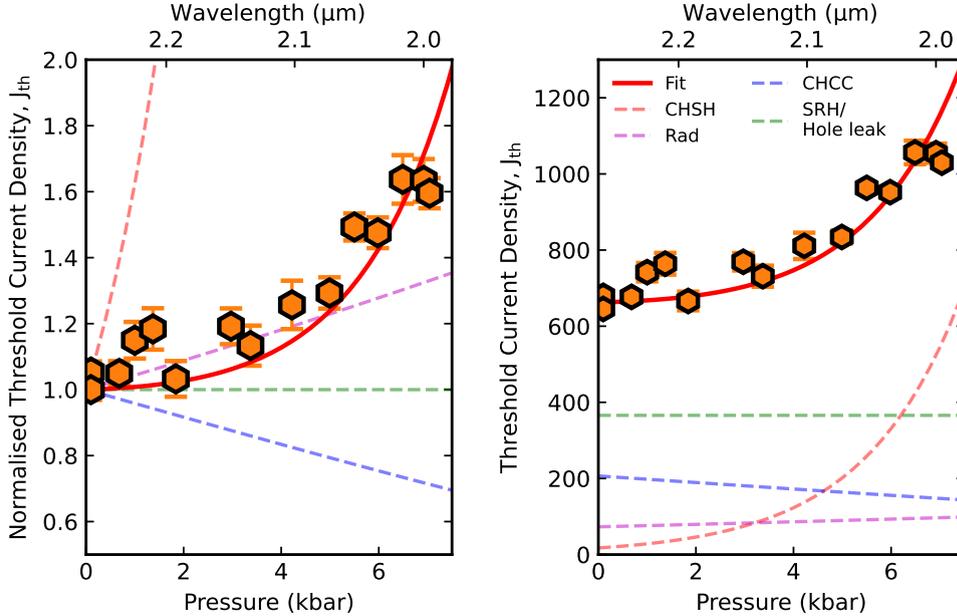


FIGURE 6.16: a) The threshold current density as a function of pressure for PM10A (run 2). Dashed red, pink, green and blue lines represent the pressure dependencies of CHSH, radiative, SRH defect/hole leakage and CHCC currents, respectively. The solid red line shows the fit to the data. b) Non-normalised data, with dashed lines showing the weightings of each process as a function of pressure.

The same process is repeated assuming L-valley leakage is negligible and CHSH Auger-Meitner recombination is present instead. In this instance, the fitting parameters are obtained as $F_{ind}^0 = 0.53$, $F_{chcc}^0 = 0.29$, $F_{rad}^0 = 0.15$, $F_{chsh}^0 = 0.03$. The fit to the data is displayed in Fig. 6.16 a). Fig. 6.16 b) again highlights the fractional current contributions as a function of pressure. Similar fractions of defect/hole leakage and CHCC currents are obtained at atmospheric pressure when compared to fitting with L-valley leakage. However, much lower F_{chsh}^0 is required to reproduce the observed trend.

From this analysis, some general observations can be made. For example, it may be postulated that CHCC recombination remains important in these devices. However, growth on Si promotes increased defect recombination currents, and differences in heterostructure design make the devices on Si more prone to hole leakage, which account for a greater proportion of the total current density when combined. The latter process shall be explored in greater detail in section 6.4.2. Despite this, the general trend of increasing J_{th} with pressure can be well replicated by invoking either L-leakage or CHSH Auger-Meitner processes.

6.3.3 Relative Quantum Efficiency

More information can be obtained by analysing the relative gradient of the LI curves above threshold. The gradient is a measure of the external differential quantum efficiency, given by

$$\frac{1}{n_{ext}^d} = \left[\frac{\alpha_i}{\ln(R^{-1})} L_c + 1 \right] \frac{1}{n_{int}^d}, \quad (6.10)$$

where α_i , R and L_c are the internal optical loss, mirror reflectivity and cavity length, respectively. n_{int}^d is the internal differential quantum efficiency, which is dependent on three separate contributions [221]. Firstly, when injecting current, there is a lateral spreading. As such, the current density is only large enough to overcome optical losses and produce gain over a certain region below the contact. This phenomenon gives rise to the differential current spreading efficiency, given by [67]

$$\eta_s^d = \left[\frac{\Delta J_{las}}{\Delta J} \right]_{J > J_{th}} \quad (6.11)$$

where ΔJ is the change in total injected current density, and ΔJ_{las} is the change in injected current density that can contribute to lasing above threshold.

The second contribution to the internal differential quantum efficiency is the result of spatial leakage of carriers into layers other than the active material. For example, through carrier recombination in the barriers before reaching the wells, or through leakage over the hetero barrier into surrounding layers. The resultant term is therefore given by [67]

$$\eta_{inj}^d = \left[\frac{\Delta J_{qw}}{\Delta J_{las}} \right]_{J > J_{th}}, \quad (6.12)$$

where ΔJ_{qw} is the change in the injected current that ends up in the quantum wells. If the quasi-Fermi levels in the surrounding heterostructure are pinned above threshold, then $\eta_{inj}^d = 1$.

The third term corresponds to the fraction of carriers recombining in the QW that do so via stimulated emission, denoted η_{rad}^d . Above threshold, if the quasi-Fermi levels pin, then the spontaneous emission and non-radiative current remains fixed and $\eta_{rad} = 1$, reflecting that any increase in current into the QW results in stimulated emission. If the quasi-Fermi levels do not pin in the QW above threshold, a fraction $1 - \eta_{rad}^d > 0$ of the current in the QW recombines via spontaneous emission and/or non-radiative processes.

With this understanding, the variation in quantum efficiency with increasing pressure is assessed using the gradient of the LI curves above threshold. Fig. 6.17 illustrates the differential quantum efficiency of PM10A and PM10B in a) and b), respectively. The data was subsequently normalised to their respective values at atmospheric pressure to assess the relative change only. Both devices show very strong decreases in η_{ext}^d with increasing pressure. Previous analysis by the growers suggest internal optical losses of only $2-6 \text{ cm}^{-1}$ at atmospheric temperature and pressure [214]. With such low levels of loss and pressure-dependent bandstructure simulations suggesting strong IVBA suppression, it is unlikely that a pressure dependent increase in optical loss is responsible for the observed trend.

As such, it can be inferred that the decrease in η_{ext}^d likely stems from a decrease in the internal differential quantum efficiency, η_{int}^d . Since the heterostructure offsets are maintained during pressure measurements, any amount of spatial hole leakage should be constant such that η_{inj}^d is invariant. Thus, the large change in efficiency can plausibly be attributed to a change in η_{rad}^d . In making this statement, it is once more assumed that the lateral current spreading is negligible and independent of pressure. In reality the effective carrier masses increase with pressure, predominantly in the LH and CB, which

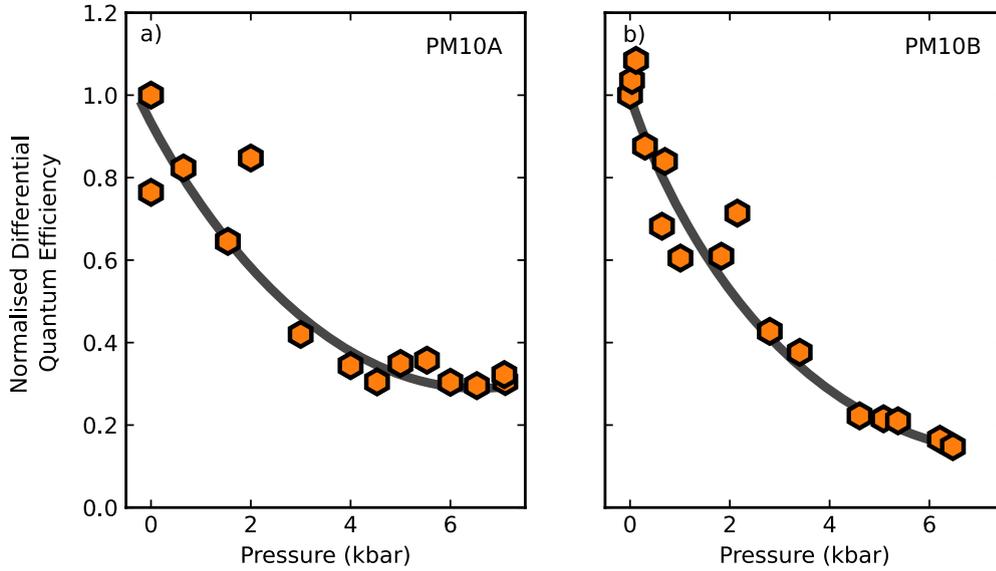


FIGURE 6.17: The differential quantum efficiency as a function of pressure, obtained from the slope of the LI curves above threshold for a) PM10A and b) PM10B. Values are normalised to atmospheric pressure. Grey curves are intended as a guide to the eye.

would act therefore decrease mobility and likely improve lateral carrier confinement. This is in contrast to the decrease in emission efficiency observed here.

Measurements of SE from PM100A illustrated that the quasi-Fermi levels pin when lasing, highlighted by a constant SE with increasing current above threshold for all temperatures in Fig. 6.10. This is particularly important as it insinuates that direct valley Auger-Meitner recombination currents are fixed above threshold and cannot be responsible for the observed decrease in η_{ext}^d . While SE measurements could not be conducted using the existing pressure set-up to verify pinning, if it is assumed that pinning persists as a function of pressure as per the temperature dependent measurements, then the drop in efficiency may then be explained using L-valley leakage. If such indirect states become occupied, the shift in CB quasi-Fermi energy with injection is massively suppressed due to the large density of indirect valley states. As such, the SE appears pinned above threshold since the Γ -valley occupation is nominally the same, whilst η_{rad}^d is greatly reduced due to an increased total current into the QW which cannot recombine via stimulated emission.

6.4 Modelling

The temperature dependent measurements of stimulated and spontaneous emission in section 6.2 illustrated the presence of a large defect-recombination current pathway which

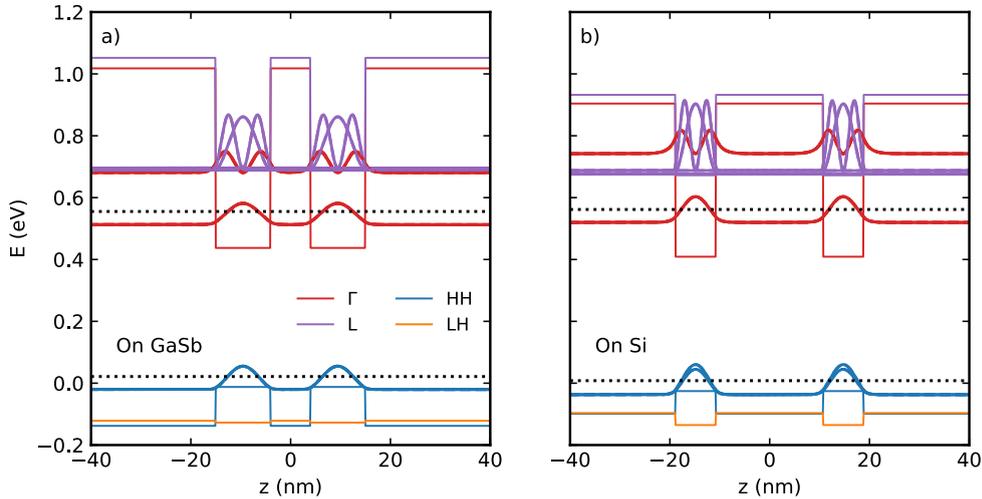


FIGURE 6.18: Flat band diagrams showing the bulk band edges, confined Γ and L states in the CB and the first confined state in the VB for a) devices on GaSb and b) those on Si.

greatly impacted the total threshold current density of devices on Si. It also demonstrated the existence of a thermally activated, highly temperature dependent process that raised threshold significantly above 180 K. In an attempt to determine the contributing process, pressure dependent measurements of threshold were conducted. The distinct behaviour of J_{th} with increasing pressure for devices on Si and those on GaSb hinted at differences in the bandstructure. Through analysis of the differential quantum efficiency with increasing pressure, a decrease, which may be indicative of leakage to L-valley states, is observed. However, this assumes that the Fermi-levels do indeed pin well in the device, which cannot be verified through pressure-dependent measurements of the SE. Thus, to attempt to resolve whether CHSH Auger-Meitner or L-leakage are responsible for the increase in J_{th} with pressure, the full bandstructure of these seemingly similar active regions must be modelled.

Fig. 6.18 a) and b) show the calculated band alignment for the device on GaSb and those on Si, respectively. These are calculated using a ‘2+6’ band model available in kpymod. For clarity, the effects of band bending are ignored initially such that typical band offsets can be observed. A full Poisson-Schrödinger model is implemented during the optimisation stage. The dotted lines depict the CB and VB quasi-Fermi levels at transparency, achieved at an injected carrier density of $1.2 \times 10^{12} \text{ cm}^{-2}$ for the device on GaSb and $1.3 \times 10^{12} \text{ cm}^{-2}$ for devices on Si. As expected, it can be seen that the bulk CB edges are comparable in the well for both devices, owing to similar active region

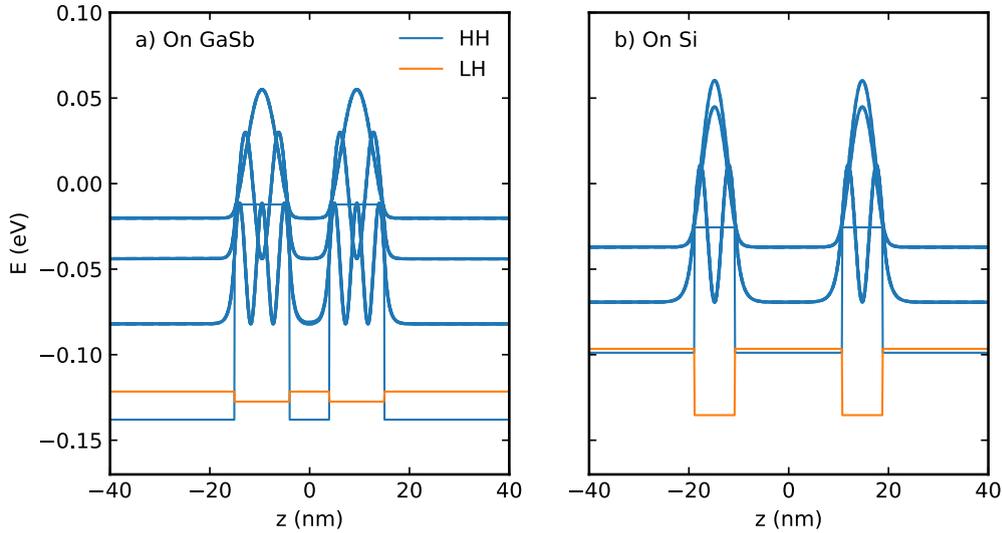


FIGURE 6.19: The VB structure (flat band), showing the confined states for devices on a) GaSb and b) Si.

compositions and lasing energies. In both instances, the lowest L-valley state is situated ~ 160 meV above the lowest Γ -state. However, the second Γ state is of significantly higher energy in the device on Si, which may make L-valley occupation higher for a given injection above transparency.

Additionally, the VB offset for the devices on GaSb is close to 100 meV, compared to 60 meV for the devices on Si, as illustrated in Fig. 6.19. In literature, an offset of 150 meV is often quoted as being adequate for suppression of hole leakage [222]. As such, both devices may be prone to increased leakage currents at higher temperatures, but more so in devices on Si where hole confinement is greatly reduced. It is therefore necessary to further probe how these minor differences in bandstructure may be impacting the observed experimental trends.

6.4.1 Conduction Band Carrier Distribution

To assess the potential of L-valley leakage in these devices it is prudent to analyse how carriers distribute themselves in k-space for injection levels above transparency and observe how this varies with temperature and pressure. Since increasing carrier density results in heightened splitting of the quasi-Fermi levels, the approach of Blood et al. is adopted, using the so called thermally normalised injection level, $F = (\Delta E_f - E_0)/k_B T$ [67]. Here a value of $F = 0$ is obtained when $\Delta E_f = E_0$ i.e. at transparency. Values of $F = 1, 2, 3$

etc. then denote quasi-Fermi level separations that are integer multiples of the thermal energy above the minimum transition energy.

As per the previous chapter, the bandstructure is calculated for a given temperature and pressure and the quasi Fermi levels are calculated for a set of input carrier densities. The quasi Fermi levels are then utilised to determine the population of the different CB valleys at a given injection. Fig. 6.20 illustrates the fraction of carriers in Γ , L and X as a function of the thermally normalised injection at RT and atmospheric pressure. At transparency, it can be seen that 16.6% of injected carriers occupy L-valley states. This illustrates that the bandstructure of the device is such that a large fraction of the injected current will be lost to indirect states where stimulated emission cannot occur. The requirement for higher carrier densities to compensate for this process will likely raise the total threshold current density due to increased parasitic losses such as Auger-Meitner recombination and spatial hole leakage.

The above calculation is repeated as a function of temperature, ranging from 50 - 300 K. Fig. 6.21 a) plots the fraction of carriers in the L-valley as a function of thermally normalised injection. It can be seen that L-valley occupation becomes appreciable between 150 - 200 K, and rises most with increasing inversion for the highest temperatures. Fig. 6.21 b) illustrates the fraction of L-valley carriers at transparency as a function of temperature, as calculated for the active regions on Si (orange) and those on GaSb (blue).

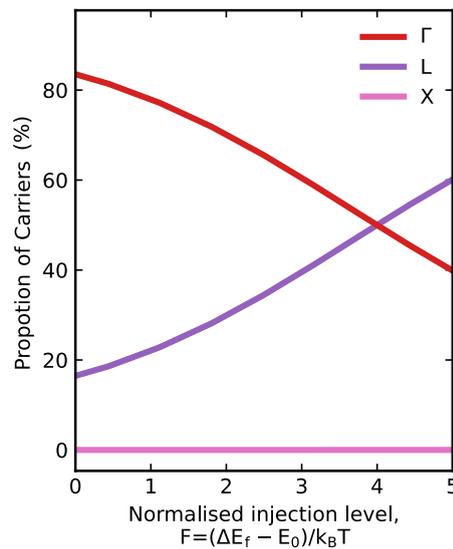


FIGURE 6.20: The fractional distribution of carriers between the Γ , L and X valleys as a function of increasing thermally normalised injection.

In both sets of devices, negligible indirect valley occupation occurs below 175 K, however this fraction rises exponentially with temperature above this value. It can be seen that L-valley fraction at transparency rises faster with temperature for the devices on Si when compared to those on GaSb. At RT, 10% of injected carriers are indirect for the device on GaSb, whilst over 17% are indirect in the device on Si. This discrepancy can be attributed to the energetic difference between the second Γ state in the two heterostructure designs, as discussed previously. As such, due to a lower concentration of L-valley carriers, it is possible that the devices on GaSb substrates could be dominated by CHCC Auger-Meitner recombination, whilst L-valley leakage could be appreciable in the devices on Si, despite seemingly similar active region designs. It should be noted that the break point temperature of ~ 150 K observed in this model is in good agreement with the onset of non-radiative processes in the experimental measurements of threshold with temperature.

The temperature is then fixed to 300 K and the same calculations are conducted as a function of increasing pressure. In Fig. 6.22 a), the fraction of indirect carriers is seen to increase with pressure. Fig. 6.22 b) depicts the percentage of L-valley carriers at transparency with increasing pressure for devices on Si (orange) and on GaSb (blue). The fractional increase with pressure is similar in both devices, since the pressure coefficients are comparable. This ensures that $n_{L,tr}/n_{tr}$ remains $> 8\%$ higher in the devices on Si

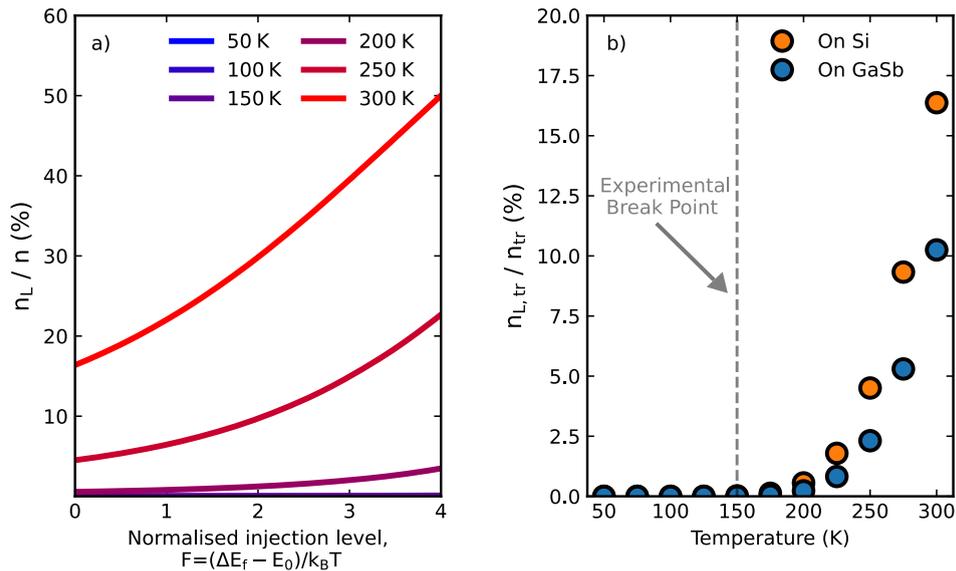


FIGURE 6.21: a) The fraction of carriers occupying L states as a function of increasing inversion and temperature (blue to red) for the active region on Si. b) The fraction of carriers in the L-valley at the transparency carrier density as a function of temperature for devices on Si (orange) and GaSb (blue).

across the whole pressure range. The heightened proportion of carriers in the L-valley of the devices on Si may then explain the discrepancy in the pressure trends observed experimentally.

6.4.2 Valence Band Carrier Distribution

In order to probe the potential influence of hole leakage, a Poisson-Schrödinger solver was used to determine the transparency carrier density for devices on Si and GaSb at RT. The DoS were then utilised to assess the fraction of carriers above the VB barrier offset at transparency.

For simulations of the device on Si, the active region consisted of two 8.1 nm quantum wells separated by a 21.5 nm barrier. This was sandwiched between two 25 nm waveguides. This produced a total simulation domain of 87.7 nm, which was to be matched when simulating the active region of the devices on Si.

A heatmap illustrating the position and energy dependence of the carrier density at transparency is given in Fig. 6.23. Here, grey lines depict the band edges, whilst the red dashed lines show the quasi-Fermi levels. The spatial dependence of the electron and hole carrier densities are further elucidated in the topmost figure. Whilst electrons are well confined

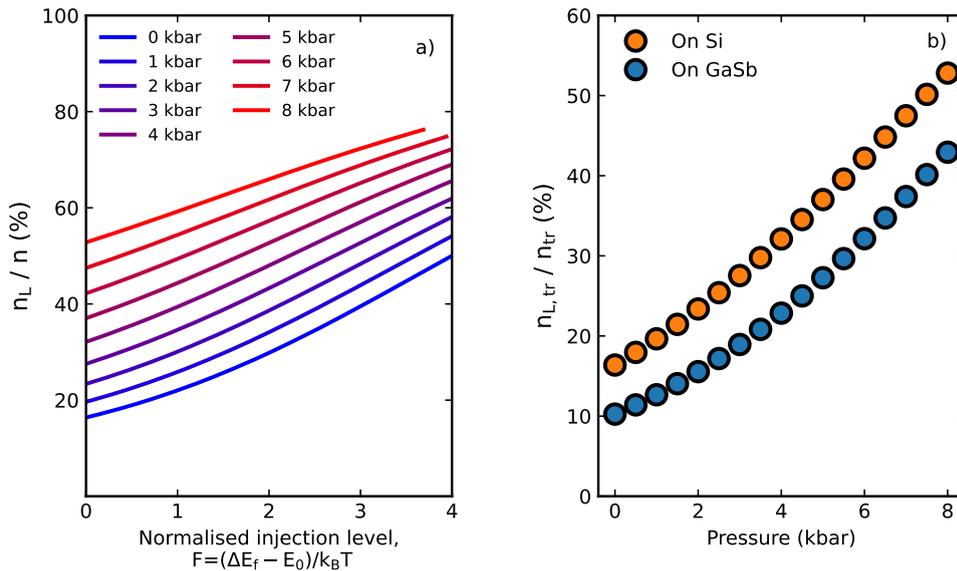


FIGURE 6.22: a) The fraction of carriers occupying L states as a function of increasing inversion and pressure (blue to red) for the active region on Si. b) The fraction of carriers in the L-valley at the transparency carrier density as a function of pressure for devices on Si (orange) and GaSb (blue).

to the QW regions, the hole density in the barriers is appreciable. This is further highlighted by the rightmost plot, illustrating the 2D DoS, weighted by the Fermi function. The orange highlighted area represents the sheet-carrier density at energies greater than the VB barrier offset. By comparing to the total carrier density, it is found that 51.6% of holes at transparency are existing in states above the barrier energy.

The calculations were repeated for the device on native GaSb. In these devices, two 11 nm wells are placed between three 11 nm barriers. To ensure the same simulation domain size, 16.3 nm of the waveguide layers are included on either side of the outer barriers. The results are given in Fig. 6.24. From the heatmap and 2D Fermi weighted DoS plots, it can be seen that there is a higher DoS close to the VB edge as a result of having wider wells. When combined with the larger VB barrier, this yields far fewer carriers occupying

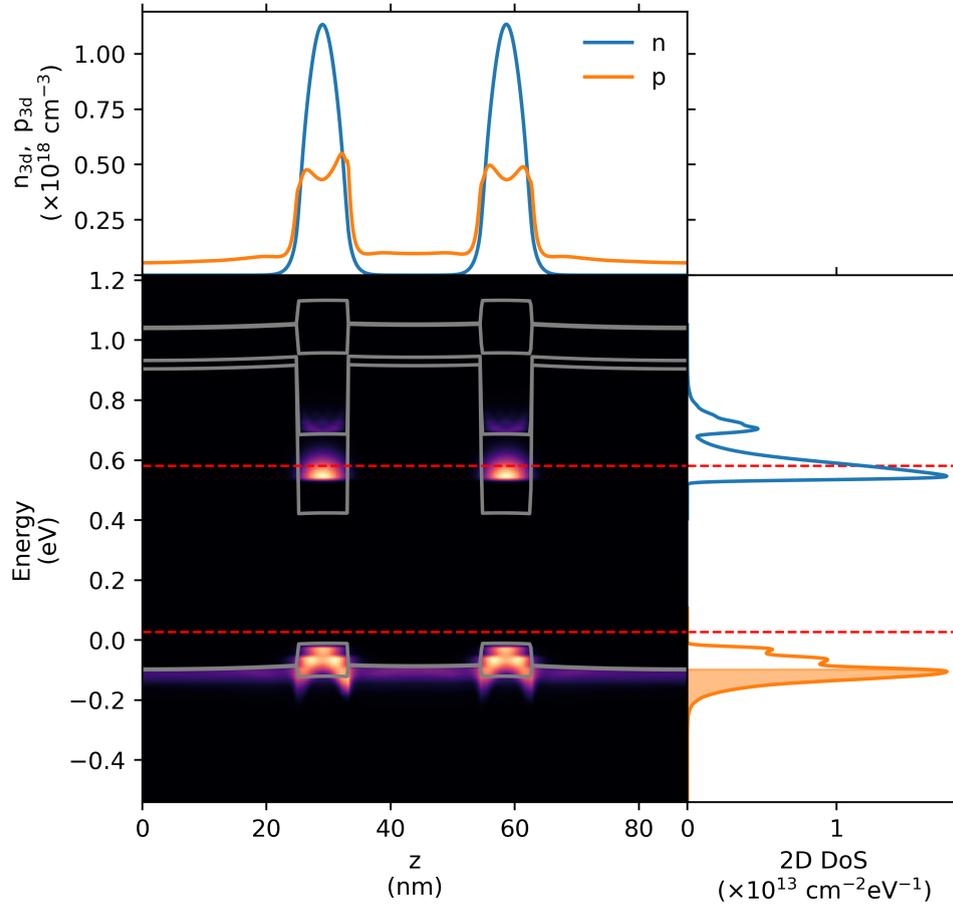


FIGURE 6.23: Heatmap illustrating the spatial and energetic variation in carrier density at the transparency condition for the active region on Si. The topmost plot shows the spatial variation of the 3D carrier density, whilst the rightmost figure shows the 2D DoS weighted by the Fermi function. The filled orange region denotes the portion of the sheet carrier density occupying states above the barrier in the VB.

energies above the barrier edge. At RT, 22.4% exist in this energy range at transparency. Even larger electron confinement is illustrated in these devices, meaning electron leakage in both types of devices are likely to be negligible.

Fig. 6.25 shows the fraction of carriers above the VB barrier edge as a function of temperature, calculated using the methodology outlined above. The low barrier offset and reduced DoS in the devices on Si result in a much sharper rise with increasing temperature compared to those on GaSb. For all temperatures the leakage carrier density is in excess of $2\times$ higher for the active region on Si, and $> 10\times$ higher below 175 K.

It should be noted that these fractions are strongly dependent on the size of simulation domain, since a larger domain would introduce a greater number of total states, and thus

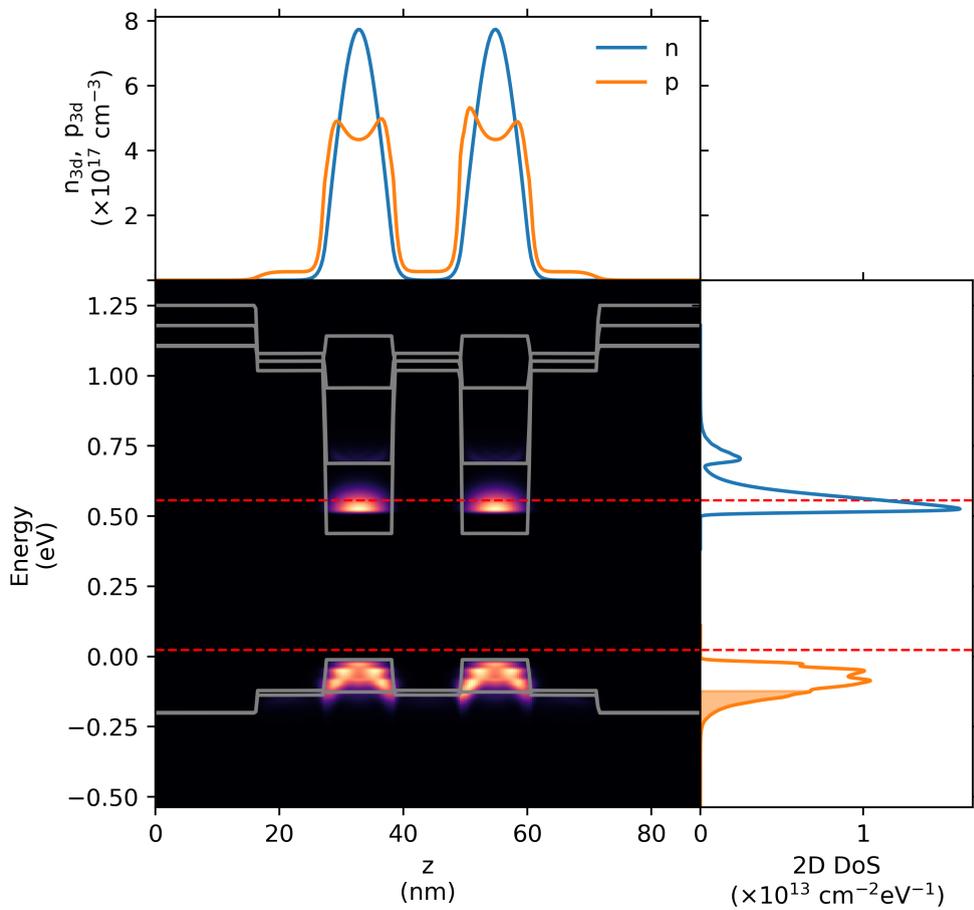


FIGURE 6.24: Heatmap illustrating the spatial and energetic variation in carrier density at the transparency condition for the active region on GaSb. The topmost plot shows the spatial variation of the 3D carrier density, whilst the rightmost plot shows the 2D DoS weighted by the Fermi function. The filled orange region denotes the portion of the sheet carrier density occupying states above the barrier in the VB.

artificially raise the leakage fraction. Additionally, hole mobility is known to be very low compared to that of the electron, and nominally decreases with increasing temperature, which may go some way to offsetting the effects of thermal carrier distribution broadening.

As such, it can be postulated that hole leakage may be exacerbated for the active region on Si compared to the design on GaSb. However, full drift-diffusion modelling of the entire heterostructure is likely required to make quantitative, temperature-dependent estimations of leakage currents within the devices.

6.4.3 Layer Optimisation

It is prudent to check whether minor changes to the active region design may be able to reduce the L-valley occupation for temperatures above 200 K, whilst also decreasing the fraction of holes occupying non-confined states.

6.4.3.1 Quantum Well Optimisation

The primary goal is to reduce the energy of the second Γ state such that it is below the lowest L state. One way to attain this is to increase the thickness of the well, this reduces the confinement energy, and therefore lowers the energy of the Γ states. Since the L states have a large effective mass, their energy is comparatively unaffected by minor

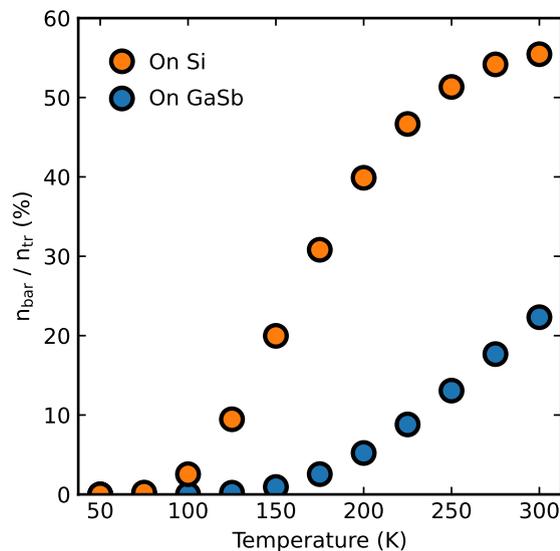


FIGURE 6.25: The fraction of carriers occupying states higher than the barrier energy as a function of temperature for the active regions on Si (orange) and GaSb (blue).

changes in well thickness. As the energy of the ground state transition is reduced when using a thicker well, the Sb or In fractions must be increased in order to attain the same emission wavelength. Additionally, increasing the well thickness can also be beneficial in decreasing hole leakage since the number of confined hole states can be raised.

Fig. 6.26 a) shows the emission wavelength as a function of Sb content and QW thickness. It should be noted that if the layer is wider than some critical thickness, it will no longer be perfectly strained to the previous layer. Instead, it becomes energetically favourable for relaxation to occur in the plane, resulting in the formation of defects, which degrades performance. Here, the critical thickness, h_c , is calculated using the Matthews-Blakeslee model as [68]

$$h_c = \frac{b}{2\pi\varepsilon_{\parallel}} \frac{1 - \nu\cos^2\alpha}{(1 + \nu)\cos\lambda} \left[1 + \ln\left(\frac{h_c}{b}\right) \right], \quad (6.13)$$

where ν is the Poisson ratio for a biaxially strained cubic lattice, given by $\nu = 2C_{12}/C_{11}$. The most common dislocation type for cubic crystals is with a propagation direction of $\alpha = \lambda = 60^\circ$ and a Burgers vector of $a/2\langle 110 \rangle$, giving a magnitude of $b = a\sqrt{2}/2$. Thus, the critical thickness is calculated based on the energy required to form this type of dislocation. Whilst Eqn. 6.13 can be solve iteratively, for example using a Newton-Raphson method, a direct-solve method based on the Lambert W_{-1} function is implemented in kpymod [223].

The greyed region in Fig. 6.26 represents compositions where this critical thickness is

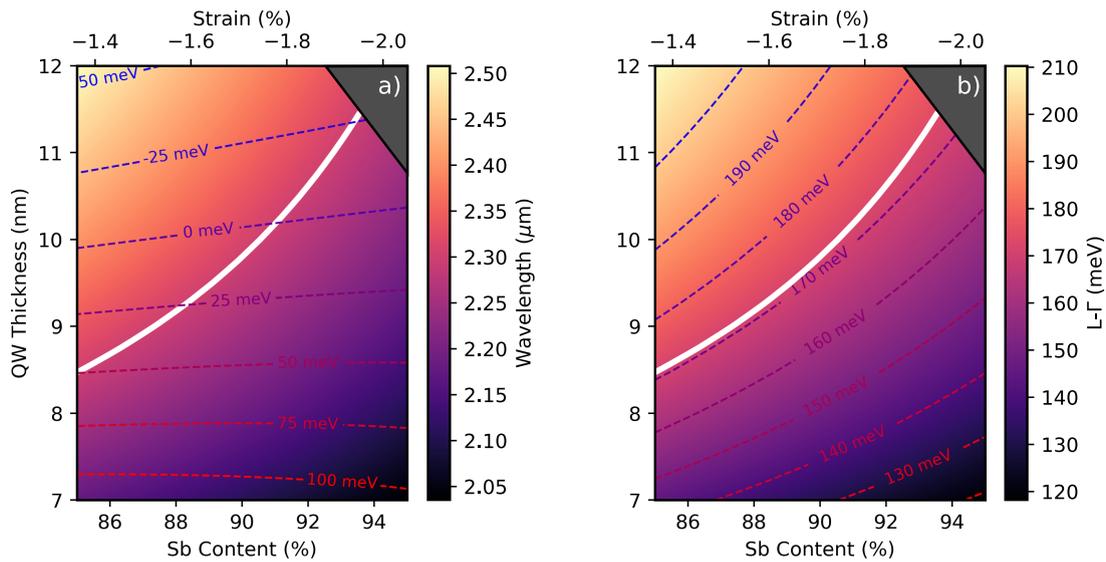


FIGURE 6.26: a) Heatmap of emission wavelength as a function of QW thickness and Sb content. Contours represent Γ_2 -L. b) Heatmap and contours of L- Γ_1 as a function of QW thickness and Sb content. In both plots, greyed regions denote compositions in excess of the critical thickness while white lines show pairs of values producing emission at 2.3 μm .

exceeded. The solid white line denotes pairs of thickness and Sb content that produce emission at $2.3 \mu\text{m}$. In Fig. 6.26 a), dotted contours represent the energetic value of Γ_2 -L in meV, where Γ_2 is the second highest Γ state in the QW. Increasing the well thickness does indeed lower the energy of Γ_2 , however the critical thickness of the layer excludes high Sb compositions with thick wells which would otherwise produce the greatest reduction in Γ_2 -L.

In order for the improvements to the Γ_2 energy to yield reduced L-valley occupation, the splitting between the lowest L state and Γ_1 must be maintained or increased. Fig. 6.26 b) illustrates the energetic splitting between the lowest energy L and Γ states. It is found that increasing Sb content and well thickness produces very minor improvements to the directness for devices emitting at $2.3 \mu\text{m}$. Thus, it can be expected that the fraction of L-valley carriers should decrease.

The fraction of indirect carriers are once more calculated as a function of temperature and pressure for a device employing a thicker well with higher Sb content. Values of 10.9 nm and 92.7% are selected to avoid compositions that are too close to the critical thickness, whilst providing appreciable improvements in Γ_2 -L. Fig. 6.27 a) verifies that this change results in a $\sim 36\%$ reduction in the fraction of L-valley carriers, falling from 16.4% to 10.5%. This brings the devices on Si in line with those on native GaSb. The fraction of L-carriers as a function of pressure is illustrated in Fig. 6.27 b), which highlights that this metric can also be reduced in line with the devices on native GaSb.

The inset of Fig. 6.27 a) shows the effect of the minor change in QW thickness and composition on the fraction of carriers occupying states above the barrier. The results for the non-optimised well are given by faded markers, with post-optimised results showing that the fraction at RT is more than halved. Again, the waveguide thickness was adjusted such that the overall simulation domain was maintained across all calculations.

6.4.3.2 Barrier and Waveguide Optimisation

Further improvement to the VB can be achieved by varying the composition of barrier and waveguide layers. Type I AlGaAsSb/InGaAsSb devices are typically characterised by substantial imbalance between the electron and valence band offsets [222]. Since electrons have far higher mobility, barrier compositions are chosen to provide adequate CB offset, often resulting in insufficient hole confinement, as demonstrated in these devices.

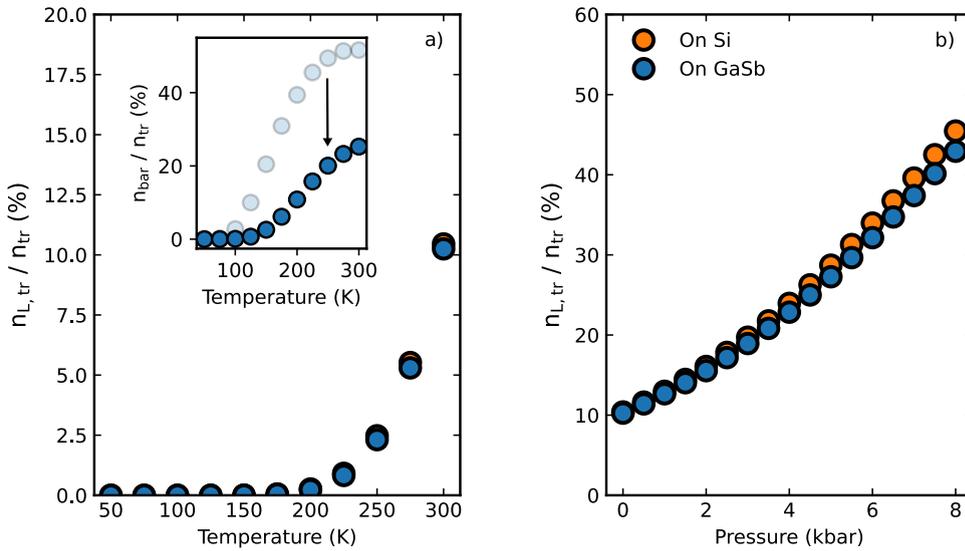


FIGURE 6.27: a) and b) show the fraction of carriers in the L-valley at transparency as a function of temperature and pressure, respectively, subsequent to increasing the well thickness and Sb content. The inset in a) depicts the decrease in carriers above the VB barrier edge at transparency after this alteration.

In the previous section the Sb content was raised in order to retain emission at $2.3 \mu\text{m}$ after the well was made thicker. This increases the compressive strain in the QW which raises the HH band edge energy, improving the VB offset from 60 meV to 99.4 meV. A secondary effect is that an additional confined HH state is obtained since the confinement energy is reduced by using a wider well. Improved device performance stemming from increased compressive strain has been experimentally demonstrated in multiple studies and is now widely appreciated [224]. However, further increases in compressive strain and layer thickness to enhance hole confinement are not possible due to the critical layer thickness. Thus, further refinements to the VB offset can only be achieved by altering the barrier and waveguiding layers.

In AlGaAsSb alloys lattice-matched to GaSb, simultaneous reduction of the Ga and Sb content is required to increase the bandgap whilst maintain lattice matching. However, a comparatively small change in Sb is required for a given change in Ga, as lattice matching can be achieved for any fraction of Ga, but Sb content must remain in excess of 0.9. As a result, the Ga-containing binary alloys have the greatest influence on the change in properties such as the bandgap and valence band offset. As the fractional contribution of GaAs and GaSb decreases, the bandgap increases due to the wider gaps of Al-based binaries. Similarly, since the natural VB offset is larger in AlAs, hole confinement is increased.

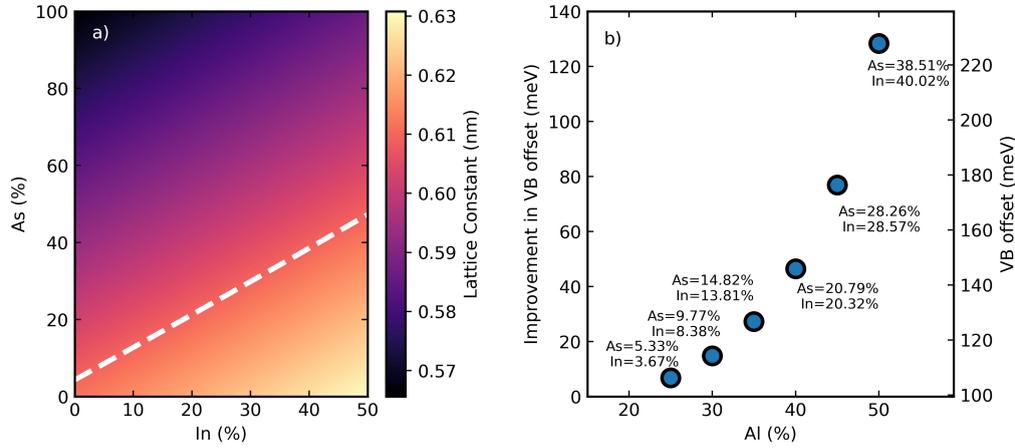


FIGURE 6.28: a) Heatmap of lattice constant as a function of In and As content for Al=0.5. The white dashed line shows compositions that are lattice matched to GaSb. b) The energetic improvement in VB confinement energy as the Al fraction is increased for compositions that are lattice matched to GaSb and reproduce the conduction band offset from the original barrier material.

Despite this, the bandgap increases much faster than the hole confinement energy, such that the predominant effect of the compositional change is an increase in the electron confinement energy. This can result in electron distribution inhomogeneity in multi-well devices. Thus, to attain appropriate levels of both electron and hole confinement it may be preferable to adopt quinary AlGaInAsSb layers for the barriers, which allows the CB and VB offsets to be tuned independently [225].

To test the extent to which the VB offset could be improved, the material was first defined as $\text{Al}_x\text{Ga}_{1-x-y}\text{In}_y\text{As}_z\text{Sb}_{1-z}$ where x and z can take values between 0 and 1, and y can vary between 0 and $1-x$. The Al fraction was fixed and the lattice constant calculated whilst the In and As contents were allowed to vary, producing a heatmap akin to that shown in Fig. 6.28 a). The dashed white line illustrates the pairs of As and In contents required to produce lattice matching to GaSb for the given Al fraction. The bulk HH and CB energies were subsequently calculated for the lattice matched compositions. From here, the composition required to reproduce the CB energy of the original barrier material was determined. The VB energy was extracted for the same composition.

The process was repeated for Al content between 0.25 - 0.50, and the difference between the VB offset for the quaternary and quinary barriers are plotted in Fig. 6.28 b). It is observed that Al= \sim 0.45, corresponding to In and As contents around 28% would be required to attain a 80 meV improvement in the VB confinement energy, producing a total confinement in excess of 175 meV. This situation is modelled in Fig. 6.29, producing

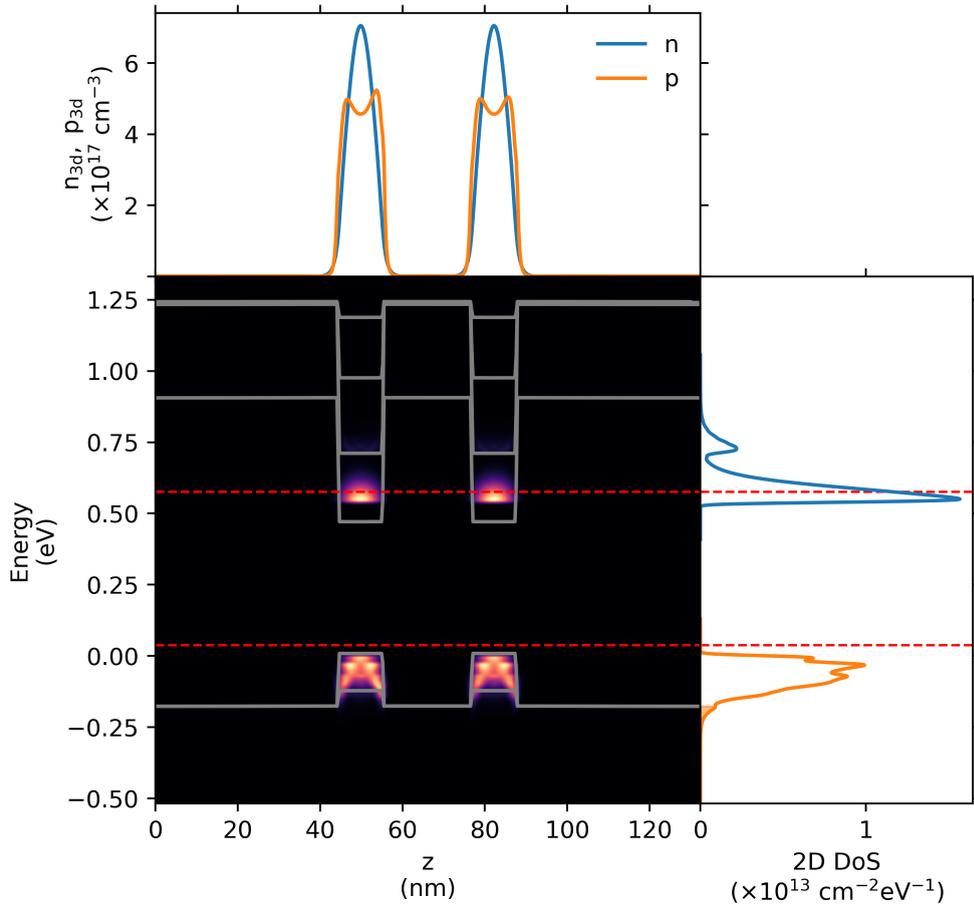


FIGURE 6.29: The spatial and energetic variation in carrier density at transparency when using thicker wells with higher Sb content and quinary barriers. The top and rightmost plots verify a greatly reduced hole density above the VB edge while strong CB confinement is retained.

a hole leakage carrier density of only 2.5%. In reality, a full drift-diffusion model is necessary to quantify the leakage current density such that the barrier offset can be adjusted appropriately through In incorporation. Since this increases the bandgap, the refractive index contrast between the waveguide and the cladding layers will be reduced, therefore decreasing the modal confinement and reducing modal gain for a given sheet carrier density. As such the In content in the barrier will require fine tuning to balance the effects of transparency carrier density and modal overlap to attain the lowest possible thresholds.

6.5 Summary & Conclusions

The performance of type I Sb-based QW devices grown epitaxially on Si substrates were compared to devices with a similar active region grown on native GaSb. Temperature dependent measurements of the threshold current density and Z-analysis from spontaneous emission data showed that the current was dominated by defect related recombination at low temperatures. Since the n-contact is placed on the InAsSb buffer layer, this is attributed to large threading dislocation densities in these lower buffer layers due to sizeable lattice mismatch with the Si substrate.

The large rise in threshold current density with temperature associated with Sb-based lasers was again observed in the devices on Si. Typically this is the result of Auger-Meitner recombination or increasing hole leakage. The mechanism causing this shift was probed using high-hydrostatic pressure at RT. The threshold in devices on GaSb decreased with pressure, indicative of CHCC dominated current, while devices on Si increased with pressure, suggesting a different dominant mechanism. Hole leakage is a pressure independent process, since band offsets are maintained with increasing pressure, and hence cannot be responsible for the observed trend. Both CHSH Auger-Meitner recombination and increasing L-valley occupation show similar trends, making quantitative fits to the data difficult. However, the drop in external quantum efficiency with pressure is better explained by increased L-valley occupation. From previous studies, regardless of strain or active region composition, CHSH Auger is only expected to only become important in devices emitting below 2 μm . As such it is proposed that enhanced L-occupation may explain the difference in pressure data and also contribute to raised thresholds at higher temperatures. This is not a process that is typically investigated in these devices and may warrant further investigation.

Temperature and pressure dependent Poisson-Schrödinger models were conducted to analyse the distribution of carriers in the conduction and valence bands at transparency. It was found that there was an increased fraction of indirect electrons in the devices on Si due to the use of thinner wells which increased the confinement energy and shifted the second lowest direct state to higher energies. This accounted for 16.6% of the injected electrons at transparency. The occupation of such states may increase non-radiative carrier recombination through Auger-Meitner or defect processes, and/or heighten free carrier absorption losses.

In addition, the VB offset in the devices on Si were around 40 meV lower than in the active region design for the devices on GaSb, at just 60 meV. This results in an increased fraction of carriers occupying bulk states above the barrier edge that may contribute to a hole leakage current.

In this work, it is suggested that the L-valley occupation can be reduced for a given emission wavelength by utilising thicker wells with higher Sb content. This results in enhanced compressive strain which also aids hole confinement. When modelling the Si devices using 10.9 nm thick wells with 92.7% Sb, so not to exceed the critical thickness, the reduced L-valley occupation seen in the devices on native GaSb is matched at transparency.

VB offsets can be further increased by implementing quinary barrier layers such that the VB and CB confinement can be tuned individually to ensure adequate carrier confinement. It is postulated that Al content of 45%, with In and As contents of 28% can be used to retain the electron confinement and emission wavelength of the original device, whilst increasing the hole confinement by ~ 80 meV to 180 meV total. This takes it beyond the 150 meV offset quoted to be adequate in literature. Whilst growth of such layers is highly complex, this technique has been adopted in previous studies and may provide a feasible route for enhanced high temperature performance.

Chapter 7

Conclusions & Future Work

In this chapter, the findings from the research conducted during the course of this project are summarised. Recommendations for future research directions are offered based on the insights gained from these investigations.

7.1 Electrically Pumped GeSn Lasers

The nature of the dominant recombination processes in electrically injected, first generation, bulk GeSn FP lasers was probed using high-hydrostatic pressure measurements at low temperature. The bulk active region consisted of 10.5% Sn material, with a residual compressive strain of -0.27%. A sizeable increase in J_{th} with pressure was observed at 85 K, which may only be attributed to increased L-valley occupation, stemming from reducing levels of directness. By performing a fit to the data, considering radiative, defect-related and L-leakage current contributions, it was estimated that $93 \pm 4\%$ of the 1.2 kA/cm² threshold current density was the result of defect related recombination. This may be expected due to the large defect density observed in SEM images of the graded buffer layer through which electrons are injected.

However, even at low temperature where thermal broadening of the carrier distribution is low, $\sim 1.15 \pm 0.02\%$ of the current at atmospheric pressure was attributed to L-valley carrier leakage. This indicated that the low maximum operational temperature could be in part due to carriers being lost to indirect valleys where radiative recombination lifetimes are long.

This prompted an investigation of the material system itself to see if bulk electrically injected GeSn lasers were plausible from a purely bandstructure perspective. To do this, the transparency carrier density, representing the minimum carrier density above which lasing can be achieved, was calculated as a function of temperature for GeSn alloys with varying Sn content. It was found that Sn fractions in excess of 15% were required to achieve at least 50% of carriers in direct states at RT, with over 20% needed to get the indirect fraction under 10%. The levels of strain available to bulk electrically injected lasers were unable to provide significant improvements to the threshold carrier density, but nonetheless, compressive strain should be minimised where possible.

By including IVBA losses, calculated directly from the bandstructure, the transparency carrier density ceases to fall with increasing Sn content. This stems from the gain peak overlapping with strong LH-HH absorption as the bandgap is reduced for high Sn content layers. As such, a turning point was found to exist at $\sim 16.5\%$ Sn, constituting the minimum possible transparency carrier density. Thus in bulk devices, improving directness may not yield better performance.

Alloys around 16% exhibit bandgap energies similar to their spin-orbit splitting energy which may exacerbate Auger-Meitner recombination losses. Basic modelling of the general Auger-Meitner current trends at transparency indicate that the CHCC current may be relatively independent of Sn content and strain, since the effect of the reduced bandgap is offset by that of increased directness. In contrast, the results suggest that the CHSH Auger-Meitner current may be the highest in unstrained alloys with less than 14% Sn, whilst order of magnitude reductions may be found by further increasing Sn and tensile strain. As such, if the primary focus is to improve the operational temperature of electrically pumped, bulk GeSn lasers at any wavelength, the interplay between Auger-Meitner and IVBA effects need to be accurately modelled to find an optimal composition.

7.1.1 Future Work

The above analysis suggests that electrically injected FP lasers with GeSn active regions may be fundamentally limited by large indirect carrier losses at high temperatures. As such, it is likely that changes to the material system and/or the heterostructure will be required to attain near RT operation. Details of both areas of interest, including potential studies, are discussed in the following sub-sections.

7.1.1.1 Alloying & Doping

One approach is to further explore the available parameter space. Alloying with other group-IV materials, such as C or Pb, is one route that has been tentatively explored in a handful of studies. As with III-V multinationals, exploring more group-IV alloys may allow for superior engineering of group-IV based heterostructures due to enhanced control of the bandstructure.

Initial reports of GePb suggested a composition of 3.5% Pb may be sufficient to produce a direct-bandgap material [226]. Broderick et al. show that a singlet state at $\sim 1\%$ Pb is produced but this has primarily indirect (L-like) character, so cannot be classed as truly direct material [227]. The authors find that 3-4% would indeed be required to produce truly direct material, with a resulting bandgap of only 0.3-0.4 eV. Above 6-7% it is postulated that the alloy would become semimetallic. The growth of GePb on Ge has been reported in a number of studies [228]–[230], including the 2023 study by Liu et al. where GePb QWs are grown on GePb [231]. The bandgaps of GePb_{0.035} and GePb_{0.072} were probed measuring the absorption edge, which were found to be 0.669 eV and 0.657 eV respectively, owing to strong confinement effects. Whilst this remains an active area of work, the toxicity, environmental impact and public opinion of Pb may limit any real commercial feasibility.

Carbon doping and alloying have also been explored for producing direct bandgap material. C has the smallest effective mass of the group-IV elements, and as such an increase in the bandgap energy and decrease in directness would be expected. However, Stephenson et al. suggest introducing dilute levels of C into a Ge matrix results in an impurity level just above the CB edge [232]. The interaction of this level with the Γ -valley minima results in a strong splitting of 170 meV per C% over the first 1%. Such values are in line with those observed for the first N% in dilute nitrides which are modelled using a band anti-crossing model (150-200 meV splitting) [233]. A decrease in bandgap was observed in QW samples with 0.2% C when compared to a Ge reference at 20 K using photoreflectance. Dey et al. state that the bandstructure of GeC produces a material that is far more favourable than GeSn for the production of QW devices [234]. They suggest that a larger Γ -valley effective mass results in direct material with a confined Γ -like state for widths as low as 5 nm. If true, this could be an important step for producing efficient group-IV QW laser

heterostructures, since SiGeSn and GeSn could then be used for SCH and cladding layers, further enhancing optical confinement.

However, in 2019 the validity of the model for dilute carbides was debated by Kirwan et al. who claimed the materials were in-fact only “quasi-direct” [235]. It is stated that this is because the impurity state arises primarily from interactions with the L-valley of the Ge matrix. This suggestion was based on DFT simulations which ascertained the pressure coefficient of the lowest CB state. A pressure coefficient akin to that of the L-valley in Ge suggests that the observed CB minima would be optically inactive. These findings were supported by Broderick et al. [236], however it should be noted that this was refuted by Gulyas et al. [237]. In contrast, this group used the inner product of the CB edge state with the eigenstates of the Ge model to determine the primary character of the carbon state.

Further work is required to determine the true nature of the CB edge and, therefore, whether it is optically active. Since atomistic modelling has produced conflicting conclusions, an experimental approach is desirable for reliable verification. However, these alloys are notoriously difficult to grow since C-C bonds and carbon clusters are energetically favourable compared to Ge-C bonds. Despite this, the growth of Sn free GeC was reported by Reza et al [238] in 2023. A simple pressure dependent experiment could therefore be designed to test the nature of the bandgap. This would involve measuring the absorption in a bulk layer as a function of wavelength at different pressures. The pressure coefficient of the absorption edge would indicate whether the state is primarily direct or indirect. The premise is similar to the study previously conducted to assess the Sn content required to produce the direct-indirect transition in GeSn [91].

7.1.1.2 Material Growth & Heterostructure Design

During the early stages of (Si)GeSn growth, the primary challenge was incorporating sufficient levels of Sn to product direct gap material, since the solid solubility of Sn in Ge is $< 1.1\%$ [239]. However, advances in far from equilibrium growth techniques involving low temperature and high deposition rate means high quality layers with $> 16\%$ Sn are now routinely achieved [240]–[242]. Despite this, as shown in chapter 5, defect related recombination still plagues electrically injected FP laser devices. This occurs because growing high Sn layers results in large compressive strains that are detrimental to gain.

Thus, the GeSn active region is grown beyond the critical thickness to encourage relaxation which results in the formation of large misfit dislocation arrays at the interface [17].

Introducing tensile strain into the active region of FP lasers is inherently difficult due to the requirement for wide bandgap waveguiding and cladding layers for both carrier and modal confinement. As such, whilst QW active regions could be implemented to reduce interfacial defect formation, the issues associated with compressive strain would persist and be further amplified by quantum confinement effects. The high Sn content and/or tensile strain required to attain high temperature lasing in electrically injected devices may therefore prove difficult using a FP cavity design.

A number of the design limitations have been avoided in optically pumped microdisk lasers grown on insulators where dielectrics are used as cladding layer. Here, the layers are covered with SiN which acts as a stressor layer, bringing the active region into a tensile strain. The under etching of the microdisk also acts to reduce the number of defects. Using this technique, RT lasing has been demonstrated, however, the production process is very involved [20], [243]. In particular, a transfer process to a host substrate is necessary. The requirement for this step makes the flip-chip bonding implementation of vastly mature III-V lasers far more appealing.

In order to have a more scalable implementation, electrically injected ridge waveguide lasers grown using aspect-ratio trapping (ART) may prove beneficial. ART is a selective area growth technique which acts to limit the propagation of $\langle 111 \rangle$ oriented threading dislocations by initialising the growth in a tall, narrow trench, etched into dielectric [244]. This results in dislocations terminating at the side walls of the trench without propagating into the active layers. However, issues with antiphase domains persist due to polarity mismatch when using III-V material. This scheme has been utilised by De Konnick et al to produce electrically injected InGaAs lasers on Si, with impressive performance metrics [245]. The heterostructure design is reproduced in Fig. 7.1 for clarity. The left structure is the n-contact, while the structure on the right illustrates the p-i-n laser diode and the p-contact.

Using this design, it can be imagined that GeSn with moderate Sn content could be used for the n- and intrinsically-doped material, with 10-12% Sn used for the QWs. It may be possible for a SiN layer to be placed between the device and the oxide, akin to the microdisk structures, thereby introducing tensile strain to enhance directness. Aspect ratio trapping using GeSn has been reported in a single 2024 study, in which Sn incorporation

again reflecting that the current was defect dominated. A steep rise in Z above 150 K was indicative of a highly temperature dependent mechanism, however the basic nature of the model meant it was not possible to deduce the responsible process.

In an attempt to recover the dominant recombination process at RT, J_{th} was measured as a function of pressure. Both defect and hole leakage processes are expected to be pressure independent, however J_{th} was seen to increase with pressure in devices on Si, in contrast to those on GaSb which decreased. This highlighted that CHSH or L-leakage processes must be present in the lasers on Si, however both trends were able to provide a good fit to the data. The decreasing differential efficiency with pressure was more in line with an L-leakage contribution, but further work was required to assess this possibility.

Calculating the flat bandstructure showed minor differences in the active region design between the two sets of devices, which made both L-leakage and hole-leakage more probable in the lasers on Si. It was established that broadening the wells by 2 nm and increasing the In content to $\sim 92 - 93\%$ could both decrease the L-valley occupation and limit the number of holes above the barrier edge at transparency. It was also demonstrated that further refinements to the barrier, such as using quinary AlGaInAsSb material would enable sufficient electron and hole confinement to be attained simultaneously.

As such, it was found that nominally small changes to the active region design can drastically alter the carrier recombination mechanisms present in GaSb based lasers. It was also confirmed that defect related recombination was significantly greater in the devices on Si. The following section outlines potential future work on similar devices, and alternative schemes for GaSb based mid-infrared emitters on Si operating around 2-3 μm .

7.2.1 Future Work

To better isolate the impact of the substrate on the resultant performance metrics, it may be necessary to grow the same heterostructure design on top of GaSb, GaAs and Si wafers simultaneously. The growers have previously implemented this technique when assessing antimonide-based ICLs [202]. In doing so, the effects of increased hole and L-valley leakages encountered during this study, which greatly hindered interpretation, would be negated.

This opens up the possibility for direct comparison of the active region by measuring the gain characteristics using a segmented-contact approach [200]. Using this technique

it is possible to deduce key information such as the total internal loss and transparency carrier density. Whilst relating gain to the current density can complicate active region comparison due to the inherent effect of non-radiative recombination, this issue can be bypassed by instead relating the gain to a level of population inversion, which can be obtained directly from the collected spectra [77].

Exploring alternative heterostructure designs may also be key to improving the performance of epitaxially grown GaSb based lasers on Si. For example, utilising the ART technique discussed in the previous sub-section to reduce the impact of threading dislocations. Another option may be to change tact and opt for ICL based devices on Si in order to cover the 2-3.5 μm range currently covered by type I QW devices. This stems from preliminary studies by Cerutti et al and Fagot et al. which suggest that ICLs may be more tolerant to defects [202], [247]. This is the result of two distinct features of the heterostructure design. Firstly, the structure must minimise non-radiative recombination occurring in the surrounding layers. In ICLs, this is achieved by having a hole injector (electron blocking) layer and an electron injector (hole blocking layer) sandwiching the active region. This stops electrons (holes) escaping to the p-side (n-side) of the structure [247]. Secondly, the bandstructure should be such that recombining via a defect state is less energetically preferable than radiative recombination. Assuming the defect states in these structures are mid gap, a type II (or III) band alignment routinely used in ICLs can be used to satisfy this requirement.

Bibliography

- [1] Fortune Business Insights. “Semiconductor Laser Market Size, Share & COVID-19 Impact Analysis, By Type, Application and Regional Forecasts, 2022-2029”. Available at <https://www.fortunebusinessinsights.com/semiconductor-laser-market-107170>.
- [2] D. M. Dourado, G. B. de Farias, R. H. Gounella, et al. “Challenges in silicon photonics modulators for data center interconnect applications”. *Optics & Laser Technology* vol. 144 (2021).
- [3] H. Shimizu. “Scientific breakthroughs and networks in the case of semiconductor laser technologies in the US and Japan, 1960s-2000s”. *Australian Economic History Review* vol. 51.1 (2011).
- [4] L. Cerutti, A. Vicet, and E. Tournié. “Interband mid-infrared lasers”. *Mid-Infrared Optoelectronics*. Elsevier, 2020, pp. 91–130.
- [5] E. Tournie and A. N. Baranov. “Mid-infrared semiconductor lasers: a review”. *Semiconductors and Semimetals* vol. 86 (2012).
- [6] Y. Chen, C. Yang, T. Wang, et al. “High-Power, High-Efficiency GaSb-Based Laser with Compositionally Linearly Graded AlGaAsSb Layer”. *Applied Sciences* vol. 13.9 (2023).
- [7] T. Hosoda, G. Kipshidze, G. Tsviid, et al. “Type-I GaSb-based laser diodes operating in 3.1 to 3.3 μm wavelength range”. *IEEE Photonics Technology Letters* vol. 22.10 (2010).
- [8] T. Hosoda, G. Kipshidze, L. Shterengas, et al. “Diode lasers emitting near 3.44 μm in continuous-wave regime at 300K”. *Electronics letters* vol. 46.21 (2010).

-
- [9] K. Vizbaras and M.-C. Amann. “Room-temperature 3.73 μm GaSb-based type-I quantum-well lasers with quaternary barriers”. *Semiconductor Science and Technology* vol. 27.3 (2012).
- [10] Y. Jiang, L. Li, R. Q. Yang, et al. “Type-I interband cascade lasers near 3.2 μm ”. *Applied Physics Letters* vol. 106.4 (2015).
- [11] J. R. Meyer, W. W. Bewley, C. L. Canedy, et al. “The interband cascade laser”. *Photonics*. Vol. 7. 3. MDPI. 2020, p. 75.
- [12] S. J. Sweeney, T. D. Eales, and I. P. Marko. “The physics of mid-infrared semiconductor materials and heterostructures”. *Mid-infrared Optoelectronics* (2020).
- [13] M. S. Vitiello, G. Scalari, B. Williams, et al. “Quantum cascade lasers: 20 years of challenges”. *Optics express* vol. 23.4 (2015).
- [14] M. Razeghi, N. Bandyopadhyay, Y. Bai, et al. “Recent advances in mid infrared (3-5 μm) Quantum Cascade Lasers”. *Optical Materials Express* vol. 3.11 (2013).
- [15] Y. Yao, A. J. Hoffman, and C. F. Gmachl. “Mid-infrared quantum cascade lasers”. *Nature Photonics* vol. 6.7 (2012).
- [16] J. De Jesus, T. A. Garcia, S. Dhomkar, et al. “Characterization of the three-well active region of a quantum cascade laser using contactless electroreflectance”. *Journal of Vacuum Science & Technology B* vol. 31.3 (2013).
- [17] V. Reboud, O. Concepción, W. Du, et al. “Advances in GeSn alloys for MIR applications”. *Photonics and Nanostructures-Fundamentals and Applications* (2024).
- [18] T. D. Eales, I. P. Marko, S. Schulz, et al. “Ge_{1-x}Sn_x alloys: Consequences of band mixing effects for the evolution of the band gap Γ -character with Sn concentration”. *Scientific reports* vol. 9.1 (2019).
- [19] S. Wirths, R. Geiger, N. Von Den Driesch, et al. “Lasing in direct-bandgap GeSn alloy grown on Si”. *Nature photonics* vol. 9.2 (2015).
- [20] D. Buca, A. Bjelajac, D. Spirito, et al. “Room temperature lasing in GeSn microdisks enabled by strain engineering”. *Advanced Optical Materials* vol. 10.22 (2022).
- [21] Y. Zhou, Y. Miao, S. Ojo, et al. “Electrically injected GeSn lasers on Si operating up to 100 K”. *Optica* vol. 7.8 (2020).
-

- [22] S. Acharya, H. Stanchu, R. Kumar, et al. “Electrically Injected mid-infrared GeSn laser on Si operating at 140 K”. *IEEE Journal of Selected Topics in Quantum Electronics* (2024).
- [23] B. Ghyselen, C. Navone, M. Martinez, et al. “Large-diameter IIIV on Si substrates by the smart cut process: The 200 mm InP film on Si substrate example”. *physica status solidi (a)* vol. 219.4 (2022).
- [24] Y. Wan, C. Shang, J. Norman, et al. “Low threshold quantum dot lasers directly grown on unpatterned quasi-nominal (001) Si”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 26.2 (2020).
- [25] X. Luo, Y. Cao, J. Song, et al. “High-throughput multiple dies-to-wafer bonding technology and III/V-on-Si hybrid lasers for heterogeneous integration of optoelectronic integrated circuits”. *Frontiers in Materials* vol. 2 (2015).
- [26] C. Xiang, W. Jin, D. Huang, et al. “High-performance silicon photonics using heterogeneous integration”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 28.3 (2022).
- [27] J.-S. Park, M. Tang, S. Chen, et al. “Heteroepitaxial growth of III-V semiconductors on silicon”. *Crystals* vol. 10.12 (2020).
- [28] A. Y. Liu and J. Bowers. “Photonic integration with epitaxial III–V on silicon”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 24.6 (2018).
- [29] MODTRAN. “Atmospheric transmission data”. Available at http://modtran.spectral.com/modtran_home.
- [30] E. Pannier and C. O. Laux. “RADIS: A nonequilibrium line-by-line radiative code for CO₂ and HITRAN-like database species”. *Journal of Quantitative Spectroscopy and Radiative Transfer* vol. 222-223 (2019).
- [31] J.-M. Fedeli and S. Nicoletti. “Mid-infrared (Mid-IR) silicon-based photonics”. *Proceedings of the IEEE* vol. 106.12 (2018).
- [32] R. Soref. “Group IV photonics for the mid infrared”. *Silicon Photonics VIII*. Ed. by J. Kubby and G. T. Reed. Vol. 8629. International Society for Optics and Photonics. SPIE, 2013, p. 862902.
- [33] M. Montesinos-Ballester, C. Lafforgue, J. Frigerio, et al. “On-chip mid-infrared supercontinuum generation from 3 to 13 μm wavelength”. *ACS photonics* vol. 7.12 (2020).
-

-
- [34] J. Hodgkinson and R. P. Tatam. “Optical gas sensing: a review”. *Measurement Science and Technology* vol. 24.1 (2012).
- [35] H. Zhang, Z. Wang, Q. Wang, et al. “Parts-per-billion-level detection of hydrogen sulfide based on doubly resonant photoacoustic spectroscopy with line-locking”. *Photoacoustics* vol. 29 (2023).
- [36] S. Lin, J. Chang, J. Sun, et al. “Improvement of the detection sensitivity for tunable diode laser absorption spectroscopy: a review”. *Frontiers in Physics* vol. 10 (2022).
- [37] L. Tombez, E. J. Zhang, J. S. Orcutt, et al. “Methane absorption spectroscopy on a silicon photonic chip”. *Optica* vol. 4.11 (2017).
- [38] H. D. Yallem, M. Vlk, A. Datta, et al. “Sub-ppm methane detection with mid-infrared slot waveguides”. *ACS Photonics* vol. 10.12 (2023).
- [39] F. Ottonello-Briano, C. Errando-Herranz, H. Rödjegård, et al. “Carbon dioxide absorption spectroscopy with a mid-infrared silicon photonic waveguide”. *Optics Letters* vol. 45.1 (2020).
- [40] M. Vlk, A. Datta, S. Alberti, et al. “Extraordinary evanescent field confinement waveguide sensor for mid-infrared trace gas spectroscopy”. *Light: Science & Applications* vol. 10.1 (2021).
- [41] R. Chandrasekar, Z. J. Lapin, A. Nichols, et al. “Photonic integrated circuits for Department of Defense-relevant chemical and biological sensing applications: state-of-the-art and future outlooks”. *Optical Engineering* vol. 58.2 (2019).
- [42] S. A. Holmstrom, T. H. Stievater, D. A. Kozak, et al. “Trace gas Raman spectroscopy using functionalized waveguides”. *Optica* vol. 3.8 (2016).
- [43] N. F. Tyndall, T. H. Stievater, D. A. Kozak, et al. “Waveguide-enhanced Raman spectroscopy of trace chemical warfare agent simulants”. *Optics Letters* vol. 43.19 (2018).
- [44] J. M. Morales, P. Cho, J. R. Bickford, et al. “Development of army relevant wearable Photonic Integrated Circuit (PIC) biosensors”. *Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XXII*. Ed. by J. A. Guicheteau and C. R. Howle. Vol. 11749. International Society for Optics and Photonics. SPIE, 2021, 117490G.
-

-
- [45] T. Greenhalgh, M. Knight, M. Inada-Kim, et al. “Remote management of covid-19 using home pulse oximetry and virtual ward support”. *British medical journal* vol. 372 (2021).
- [46] R. Baets. “Silicon-photonics-based spectroscopic sensing for environmental monitoring and health care”. *2021 Optical Fiber Communications Conference and Exhibition (OFC)*. 2021, pp. 1–42.
- [47] A. Amann, B. de Lacy Costello, W. Miekisch, et al. “The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva”. *Journal of Breath Research* vol. 8.3 (2014).
- [48] L. Pauling, A. B. Robinson, R. Teranishi, et al. “Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography”. *Proceedings of the National Academy of Sciences* vol. 68.10 (1971).
- [49] A. Amann, M. Ligor, T. Ligor, et al. “Analysis of exhaled breath for screening of lung cancer patients”. *memo-Magazine of European Medical Oncology* vol. 3 (2010).
- [50] H. Haick, Y. Y. Broza, P. Mochalski, et al. “Assessment, origin, and implementation of breath volatile cancer markers”. *Chemical Society Reviews* vol. 43.5 (2014).
- [51] P. J. Mazzone. “Exhaled breath volatile organic compound biomarkers in lung cancer”. *Journal of Breath Research* vol. 6.2 (2012).
- [52] C. Wang and P. Sahay. “Breath analysis using laser spectroscopic techniques: Breath biomarkers, spectral fingerprints, and detection limits”. *Sensors* vol. 9.10 (2009).
- [53] L. Ciaffoni, G. Hancock, J. J. Harrison, et al. “Demonstration of a Mid-Infrared Cavity Enhanced Absorption Spectrometer for Breath Acetone Detection”. *Analytical Chemistry* vol. 85.2 (2013).
- [54] C. Shi, D. Wang, Z. Wang, et al. “A mid-infrared fiber-coupled QEPAS nitric oxide sensor for real-time engine exhaust monitoring”. *IEEE Sensors Journal* vol. 17.22 (2017).
- [55] G. Dudzik, K. Abramski, and G. Wysocki. “Mid-IR parts-per-billion level nitric oxide detection using solid-state laser intracavity photothermal sensor (SLIPS)”. *Sensors and Actuators B: Chemical* vol. 380 (2023).
- [56] Y. L. Pham and J. Beauchamp. “Breath Biomarkers in Diagnostic Applications”. *Molecules* vol. 26.18 (2021).
-

-
- [57] S. Pal, K. B. Ozanyan, and H. McCann. “A spectroscopic study for detection of carbon-monoxide using mid-infrared techniques for single-pass measurement”. *Journal of Physics: Conference Series* vol. 85.1 (2007).
- [58] P. Fuchs, C. Loeseken, J. K. Schubert, et al. “Breath gas aldehydes as biomarkers of lung cancer”. *International Journal of Cancer* vol. 126.11 (2010).
- [59] J. Li, Y. Peng, and Y. Duan. “Diagnosis of breast cancer based on breath analysis: An emerging method”. *Critical Reviews in Oncology/Hematology* vol. 87.1 (2013).
- [60] M. Winkowski and T. Stacewicz. “Optical detection of formaldehyde in air in the 3.6 μm range”. *Biomedical Optics Express* vol. 11.12 (2020).
- [61] A. Sharma, R. Kumar, and P. Varadwaj. “Smelling the disease: Diagnostic potential of breath analysis”. *Molecular Diagnosis & Therapy* vol. 27.3 (2023).
- [62] A. E. Klingbeil, J. B. Jeffries, and R. K. Hanson. “Temperature-dependent mid-IR absorption spectra of gaseous hydrocarbons”. *Journal of Quantitative Spectroscopy and Radiative Transfer* vol. 107.3 (2007).
- [63] S. S. Sehnert, L. Jiang, J. F. Burdick, et al. “Breath biomarkers for detection of human liver diseases: preliminary study”. *Biomarkers* vol. 7.2 (2002).
- [64] D. Halmer, G. von Basum, P. Hering, et al. “Mid-infrared cavity leak-out spectroscopy for ultrasensitive detection of carbonyl sulfide”. *Optics Letters* vol. 30.17 (2005).
- [65] I. Vurgaftman, M. Lumb, and J. Meyer. “Bands and photons in III-V semiconductor quantum structures”. Vol. 25. Oxford University Press, 2020.
- [66] G. P. Agrawal and N. K. Dutta. “Semiconductor lasers”. Springer Science & Business Media, 2013.
- [67] P. Blood. “Quantum Confined Laser Devices: Optical Gain and Recombination in Semiconductors”. New York: Oxford University Press, 2015.
- [68] P. Harrison and A. Valavanis. “Quantum wells, wires and dots: theoretical and computational physics of semiconductor nanostructures”. John Wiley & Sons, 2016.
- [69] F. Bloch. “Über die quantenmechanik der elektronen in kristallgittern”. *Zeitschrift für physik* vol. 52.7 (1929).
- [70] K.-Y. Cheng. “III–V compound semiconductors and devices”. Springer, 2020.
-

-
- [71] P. Vogl, H. P. Hjalmarson, and J. D. Dow. “A semi-empirical tight-binding theory of the electronic structure of semiconductors”. *Journal of physics and chemistry of solids* vol. 44.5 (1983).
- [72] L. A. Coldren, S. W. Corzine, and M. L. Mashanovitch. “Diode lasers and photonic integrated circuits”. Vol. 218. John Wiley & Sons, 2012, pp. 532–534.
- [73] M. G. Bernard and G. Duraffourg. “Laser conditions in semiconductors”. *physica status solidi (b)* vol. 1.7 (1961).
- [74] S. Sweeney and A. Adams. “Optoelectronic devices and materials”. *Springer Handbook of Electronic and Photonic Materials* (2007).
- [75] J. Piprek. “Semiconductor optoelectronic devices: introduction to physics and simulation”. Elsevier, 2003.
- [76] D. J. Griffiths and D. F. Schroeter. “Introduction to quantum mechanics”. Cambridge university press, 2018.
- [77] D. A. Duffy. “Development of novel semiconductor materials and type-II heterostructures for infrared applications”. PhD thesis. University of Surrey, 2022.
- [78] A. Adams. “Band-structure engineering for low-threshold high-efficiency semiconductor lasers”. *Electronics Letters* vol. 22.5 (1986).
- [79] E. Yablonovitch and E. Kane. “Reduction of lasing threshold current density by the lowering of valence band effective mass”. *Journal of Lightwave Technology* vol. 4.5 (1986).
- [80] S. J. Sweeney, T. D. Eales, and A. R. Adams. “The impact of strained layers on current and emerging semiconductor laser systems”. *Journal of Applied Physics* vol. 125.8 (2019).
- [81] E. P. O’Reilly. “Valence band engineering in strained-layer structures”. *Semiconductor Science and Technology* vol. 4.3 (1989).
- [82] S. Tiwari. “Stress and strain effects”. *Semiconductor Physics: Principles, Theory and Nanoscale*. Oxford University Press, 2020.
- [83] S. J. Sweeney and J. Mukherjee. “Optoelectronic devices and materials”. *Springer handbook of electronic and photonic materials* (2017).
- [84] A. R. Adams. “Strained-layer quantum-well lasers”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 17.5 (2011).
-

-
- [85] S. J. Sweeney. “Novel experimental techniques for semiconductor laser characterisation and optimisation”. *Physica Scripta* vol. 2004.**T114** (2004).
- [86] W. Shockley and W. Read Jr. “Statistics of the recombinations of holes and electrons”. *Physical review* vol. 87.**5** (1952).
- [87] R. N. Hall. “Electron-hole recombination in germanium”. *Physical review* vol. 87.**2** (1952).
- [88] F. Zhao, M. E. Turiansky, A. Alkauskas, et al. “Trap-assisted Auger-Meitner recombination from first principles”. *Physical Review Letters* vol. 131.**5** (2023).
- [89] A. Ghiti, M. Silver, and E. O'Reilly. “Low threshold current and high differential gain in ideal tensile-and compressive-strained quantum-well lasers”. *Journal of applied physics* vol. 71.**9** (1992).
- [90] H. Choi and G. Turner. “Antimonide-based strained quantum-well diode lasers”. *Physica Scripta* vol. 1997.**T69** (1997).
- [91] T. D. Eales, I. P. Marko, A. R. Adams, et al. “Quantifying Auger recombination coefficients in type-I mid-infrared InGaAsSb quantum well lasers”. *Journal of Physics D: Applied Physics* vol. 54.**5** (2020).
- [92] G. G. Samsonidze and G. G. Zegrya. “Auger recombination in semiconductor quantum wells in a magnetic field”. *Physical Review B* vol. 63.**7** (2001).
- [93] G. Zegrya and V. Kharchenko. “New mechanism of Auger recombination of nonequilibrium current carriers in semiconductor heterostructures”. *Soviet physics, JETP* vol. 74.**1** (1992).
- [94] M. Dyakonov and V. Y. Kachorovskii. “Nonthreshold Auger recombination in quantum wells”. *Physical Review B* vol. 49.**24** (1994).
- [95] T. Eales. “Influence of electronic band structure on the efficiency of mid-infrared photonic devices.” eng. PhD thesis. University of Surrey, 2020.
- [96] A. S. Polkovnikov and G. G. Zegrya. “Auger recombination in semiconductor quantum wells”. *Physical review B* vol. 58.**7** (1998).
- [97] Y. Ben, F. Liang, D. Zhao, et al. “The investigation of carrier leakage mechanism based on ABC-models in InGaN/GaN MQW and its effect on internal quantum efficiency under optical excitation”. *Crystals* vol. 12.**2** (2022).
-

- [98] M. Grau, C. Lin, O. Dier, et al. “Room-temperature operation of 3.26 μm GaSb-based type-I lasers with quinary AlGaInAsSb barriers”. *Applied Physics Letters* vol. 87.24 (2005).
- [99] S. J. Sweeney, G. Knowles, and T. E. Sale. “Evaluating the continuous-wave performance of AlGaInP-based red (667 nm) vertical-cavity surface-emitting lasers using low-temperature and high-pressure techniques”. *Applied Physics Letters* vol. 78.7 (2001).
- [100] L. Pavesi. “Routes toward silicon-based lasers”. *Materials Today* vol. 8.1 (2005).
- [101] L. Hultdt. “Band-to-band Auger recombination in indirect gap semiconductors”. *physica status solidi (a)* vol. 8.1 (1971).
- [102] O. Moutanabbir, S. Assali, X. Gong, et al. “Monolithic infrared silicon photonics: The rise of (Si)GeSn semiconductors”. *Applied Physics Letters* vol. 118.11 (2021).
- [103] C.-Y. Tsai, C.-Y. Tsai, C.-H. Chen, et al. “Theoretical model for intravalley and intervalley free-carrier absorption in semiconductor lasers: beyond the classical Drude model”. *IEEE Journal of Quantum Electronics* vol. 34.3 (1998).
- [104] B. Marzban, D. Stange, D. Rainko, et al. “Modeling of a SiGeSn quantum well laser”. *Photonics Research* vol. 9.7 (2021).
- [105] S. J. Sweeney. “Radiative and non-radiative recombination in 1.3 μm and 1.5 μm semiconductor diode lasers”. PhD thesis. University of Surrey, 1999.
- [106] Thorlabs. “N-KB7 lens transmission data”. Available at https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=112.
- [107] Thorlabs. “CaF₂ lens transmission data”. Available at https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=4354.
- [108] Thorlabs. “UV fused silica lens transmission data”. Available at https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=123.
- [109] I. P. Marko and S. J. Sweeney. “Optical and electronic processes in semiconductor materials for device applications”. *Excitonic and Photonic Processes in Materials* (2015).
- [110] W. Paul. “Chapter 1 High pressure in semiconductor physics: A historical overview”. Ed. by T. Suski and W. Paul. Vol. 54. Semiconductors and Semimetals. Elsevier, 1998, pp. 1–48.
-

-
- [111] T. Hertsens. “An overview of laser diode characteristics: measuring laser diode characteristics”. *ILX Lightwave white paper* (2005).
- [112] A. F. Phillips, S. J. Sweeney, A. R. Adams, et al. “The temperature dependence of 1.3- and 1.5- μm compressively strained InGaAs (P) MQW semiconductor lasers”. *IEEE Journal of selected topics in quantum electronics* vol. 5.3 (1999).
- [113] A. Marinins, S. Hänsch, H. Sar, et al. “Wafer-scale hybrid integration of InP DFB lasers on Si photonics by flip-chip bonding with sub-300 nm alignment precision”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 29.3 (2022).
- [114] M. Theurer, M. Moehrle, A. Sigmund, et al. “Flip-chip integration of InP and SiN”. *IEEE Photonics Technology Letters* vol. 31.3 (2019).
- [115] M. Tang, J.-S. Park, Z. Wang, et al. “Integration of III-V lasers on Si for Si photonics”. *Progress in Quantum Electronics* vol. 66 (2019).
- [116] Z. Yan and Q. Li. “Recent progress in epitaxial growth of dislocation tolerant and dislocation free III-V lasers on silicon”. *Journal of Physics D: Applied Physics* vol. 57.21 (2024).
- [117] R. Soref. “Mid-infrared photonics in silicon and germanium”. *Nature photonics* vol. 4.8 (2010).
- [118] S. Mukherjee, K. Miao, A. Paul, et al. “Band Structure Lab”. 2006.
- [119] R. A. Soref and L. Friedman. “Direct-gap Ge/GeSn/Si and GeSn/Ge/Si heterostructures”. *Superlattices and Microstructures* vol. 14.2 (1993).
- [120] S. Kufner, J. Furthmüller, L. Matthes, et al. “Structural and electronic properties of α -tin nanocrystals from first principles”. *Physical Review B Condensed Matter and Materials Physics* vol. 87.23 (2013).
- [121] J. Zheng, Z. Liu, C. Xue, et al. “Recent progress in GeSn growth and GeSn-based photonic devices”. *Journal of Semiconductors* vol. 39.6 (2018).
- [122] Y. Miao, G. Wang, Z. Kong, et al. “Review of Si-Based GeSn CVD growth and optoelectronic applications”. *Nanomaterials* vol. 11.10 (2021).
- [123] S. Al-Kabi, S. A. Ghetmiri, J. Margetis, et al. “Study of high-quality GeSn alloys grown by chemical vapor deposition towards mid-infrared applications”. *Journal of Electronic Materials* vol. 45.12 (2016).
- [124] J. Liu, X. Sun, R. Camacho-Aguilera, et al. “Ge-on-Si laser operating at room temperature”. *Optics letters* vol. 35.5 (2010).
-

-
- [125] R. E. Camacho-Aguilera, Y. Cai, N. Patel, et al. “An electrically pumped germanium laser”. *Optics express* vol. 20.10 (2012).
- [126] L. Carroll, P. Friedli, S. Neuenschwander, et al. “Direct-gap gain and optical absorption in germanium correlated to the density of photoexcited carriers, doping, and strain”. *Physical review letters* vol. 109.5 (2012).
- [127] S. A. Srinivasan, C. Porret, M. Pantouvaki, et al. “Carrier scattering induced linewidth broadening in in situ P-doped Ge layers on Si”. *Applied Physics Letters* vol. 113.16 (2018).
- [128] F. Armand Pilon, A. Lyasota, Y.-M. Niquet, et al. “Lasing in strained germanium microbridges”. *Nature communications* vol. 10.1 (2019).
- [129] J. Ke, L. Chrostowski, and G. Xia. “Stress engineering with silicon nitride stressors for Ge-on-Si lasers”. *IEEE Photonics Journal* vol. 9.2 (2017).
- [130] M. G. Masteghin, V. Tong, E. B. Schneider, et al. “Stress-strain engineering of single-crystalline silicon membranes by ion implantation: Towards direct-gap group-IV semiconductors”. *Physical Review Materials* vol. 5.12 (2021).
- [131] D. Stange, S. Wirths, R. Geiger, et al. “Optically pumped GeSn microdisk lasers on Si”. *ACS photonics* vol. 3.7 (2016).
- [132] V. Reboud, A. Gassenq, N. Pauc, et al. “Optically pumped GeSn micro-disks with 16% Sn lasing at 3.1 μm up to 180 K”. *Applied Physics Letters* vol. 111.9 (2017).
- [133] J. Margetis, S. Al-Kabi, W. Du, et al. “Si-based GeSn lasers with wavelength coverage of 2–3 μm and operating temperatures up to 180 K”. *ACS Photonics* vol. 5.3 (2017).
- [134] F. Gencarelli, B. Vincent, L. Souriau, et al. “Low-temperature Ge and GeSn chemical vapor deposition using Ge_2H_6 ”. *Thin Solid Films* vol. 520.8 (2012).
- [135] D. Stange, N. von den Driesch, T. Zabel, et al. “GeSn/SiGeSn heterostructure and multi quantum well lasers”. *ACS Photonics* vol. 5.11 (2018).
- [136] N. von den Driesch, D. Stange, D. Rainko, et al. “Advanced GeSn/SiGeSn group IV heterostructure lasers”. *Advanced Science* vol. 5.6 (2018).
- [137] Q. M. Thai, N. Pauc, J. Aubin, et al. “GeSn heterostructure micro-disk laser operating at 230 K”. *Optics Express* vol. 26.25 (2018).
- [138] Y. Zhou, W. Dou, W. Du, et al. “Optically pumped GeSn lasers operating at 270 K with broad waveguide structures on Si”. *ACS Photonics* vol. 6.6 (2019).
-

-
- [139] B. Marzban, L. Seidel, T. Liu, et al. “Strain engineered electrically pumped SiGeSn microring lasers on Si”. *ACS Photonics* vol. 10.1 (2022).
- [140] E. Kane. “The k-p Method”. *Semiconductors and semimetals*. Vol. 1. Elsevier, 1966, pp. 75–100.
- [141] J. Laflamme Janssen, Y. Gillet, S. Poncé, et al. “Precise effective masses from density functional perturbation theory”. *Physical Review B* vol. 93.20 (2016).
- [142] M. Willatzen and L. C. L. Y. Voon. “The kp method: electronic properties of semiconductors”. Vol. 1. Springer, 2009.
- [143] J. M. Luttinger and W. Kohn. “Motion of electrons and holes in perturbed periodic fields”. *Physical Review* vol. 97.4 (1955).
- [144] G. Dresselhaus, A. Kip, and C. Kittel. “Cyclotron resonance of electrons and holes in silicon and germanium crystals”. *Physical Review* vol. 98.2 (1955).
- [145] P.-O. Löwdin. “A note on the quantum-mechanical perturbation theory”. *The Journal of Chemical Physics* vol. 19.11 (1951).
- [146] S. L. Chuang. “Physics of photonic devices”. John Wiley & Sons, 2012.
- [147] R. R. Goodman. “Cyclotron Resonance in Germanium”. *Physical Review* vol. 122.2 (1961).
- [148] I. M. Tsidilkovski. “Band structure of semiconductors: International series on the science of the solid state”. Vol. 19. Elsevier, 2016.
- [149] A. Meney, B. Gonul, and E. O'Reilly. “Evaluation of various approximations used in the envelope-function method”. *Physical Review B* vol. 50.15 (1994).
- [150] A. M. Cohen and G. E. Marques. “Electronic structure of zinc-blende-structure semiconductor heterostructures”. *Physical Review B* vol. 41.15 (1990).
- [151] M. Gladysiewicz, R. Kudrawiec, and M. Wartak. “8-band and 14-band kp modeling of electronic band structure and material gain in Ga (In) AsBi quantum wells grown on GaAs and InP substrates”. *Journal of Applied Physics* vol. 118.5 (2015).
- [152] M. Gladysiewicz, R. Kudrawiec, and M. Wartak. “Electronic band structure and material gain of III-V-Bi quantum wells grown on GaSb substrate and dedicated for mid-infrared spectral range”. *Journal of Applied Physics* vol. 119.7 (2016).
- [153] G. Bir and G. Pikus. “Symmetry and Strain-induced Effects in Semiconductors”. A Halsted Press book. Wiley, 1974.
-

-
- [154] A. Luque, A. Panchak, A. Vlasov, et al. “Four-band Hamiltonian for fast calculations in intermediate-band solar cells”. *Physica E: Low-dimensional Systems and Nanostructures* vol. 76 (2016).
- [155] H. S. Mczko, R. Kudrawiec, and M. Gladysiewicz. “Strain engineering of transverse electric and transverse magnetic mode of material gain in GeSn/SiGeSn quantum wells”. *Scientific Reports* vol. 9.1 (2019).
- [156] O. Madelung. “Semiconductors : group IV elements and III-V compounds”. 1991.
- [157] M. Levinshstein, S. Rumyantsev, and M. S. Shur. “Handbook Series on Semiconductor Parameters”. 1996.
- [158] S. Krishnamurthy, A. Sher, and A.-B. Chen. “Generalized Brooks formula and the electron mobility in $\text{Si}_x\text{Ge}_{1-x}$ alloys”. *Applied Physics Letters* vol. 47.2 (1985).
- [159] J. Singh. “Physics of Semiconductors and Their Heterostructures”. McGraw-Hill, 1993.
- [160] C. G. Van de Walle and R. M. Martin. “Theoretical calculations of heterojunction discontinuities in the Si/Ge system”. *Physical Review B* vol. 34.8 (1986).
- [161] G.-E. Chang, S.-W. Chang, and S. L. Chuang. “Strain-balanced $\text{Ge}_z\text{Sn}_{1-z}/\text{Si}_x\text{Ge}_y\text{Sn}_{1-x-y}$ multiple-quantum-well lasers”. *IEEE Journal of Quantum Electronics* vol. 46.12 (2010).
- [162] Y.-H. Li, X. G. Gong, and S.-H. Wei. “Ab initio all-electron calculation of absolute volume deformation potentials of IV-IV, III-V, and II-VI semiconductors: The chemical trends”. *Physical Review B* vol. 73.24 (2006).
- [163] C. G. Van de Walle. “Band lineups and deformation potentials in the model-solid theory”. *Physical Review B* vol. 39.3 (1989).
- [164] T. Brudevoll, D. S. Citrin, M. Cardona, et al. “Electronic structure of α -Sn and its dependence on hydrostatic strain”. *Physical Review B* vol. 48.12 (1993).
- [165] K. Ksova, P. Hapala, J. Valenta, et al. “Direct bandgap silicon: tensile-strained silicon nanocrystals”. *Advanced Materials Interfaces* vol. 1.2 (2014).
- [166] S. Adachi. “Properties of semiconductor alloys: group-IV, III-V and II-VI semiconductors”. John Wiley & Sons, 2009.
- [167] M. Willatzen, L. C. Lew Yan Voon, P. V. Santos, et al. “Theoretical study of band-edge states in Sn_1Ge_n strained-layer superlattices”. *Physical Review B* vol. 52.7 (1995).
-

- [168] S.-H. Wei and A. Zunger. “Predicted band-gap pressure coefficients of all diamond and zinc-blende semiconductors: Chemical trends”. *Physical Review B* vol. 60.8 (1999).
- [169] P. Lawaetz. “Valence-band parameters in cubic semiconductors”. *Physical Review B* vol. 4 (1971).
- [170] A. R. Denton and N. W. Ashcroft. “Vegard’s law”. *Physical Review A* vol. 43.6 (1991).
- [171] H. Lan and C. Liu. “Valence band structure calculations of strained $\text{Ge}_{1-x}\text{Sn}_x$ quantum well pFETs”. *arXiv preprint arXiv:1703.01812* (2017).
- [172] P. A. Khomyakov, M. Luisier, and A. Schenk. “Compositional bowing of band energies and their deformation potentials in strained InGaAs ternary alloys: A first-principles study”. *Applied Physics Letters* vol. 107.6 (2015).
- [173] N. Bouarissa. “Compositional dependence of the elastic constants and the Poisson ratio of $\text{Ga}_x\text{In}_{1-x}\text{Sb}$ ”. *Materials Science and Engineering: B* vol. 100.3 (2003).
- [174] S.-Q. Liu and S.-T. Yen. “Extraction of eight-band kp parameters from empirical pseudopotentials for GeSn”. *Journal of Applied Physics* vol. 125.24 (2019).
- [175] K. Lu Low, Y. Yang, G. Han, et al. “Electronic band structure and effective mass parameters of $\text{Ge}_{1-x}\text{Sn}_x$ alloys”. *Journal of Applied Physics* vol. 112.10 (2012).
- [176] V. R. Dcosta, C. S. Cook, A. G. Birdwell, et al. “Optical critical points of thin-film $\text{Ge}_{1-y}\text{Sn}_y$ alloys: a comparative $\text{Ge}_{1-y}\text{Sn}_y/\text{Ge}_{1-x}\text{Si}_x$ study”. *Physical Review B* vol. 73.12 (2006).
- [177] L. Jiang, J. Gallagher, C. L. Senaratne, et al. “Compositional dependence of the direct and indirect band gaps in $\text{Ge}_{1-y}\text{Sn}_y$ alloys from room temperature photoluminescence: implications for the indirect to direct gap crossover in intrinsic and n-type materials”. *Semiconductor Science and Technology* vol. 29.11 (2014).
- [178] H. Tran, W. Du, S. A. Ghetmiri, et al. “Systematic study of $\text{Ge}_{1-x}\text{Sn}_x$ absorption coefficient and refractive index for the device applications of Si-based optoelectronics”. *Journal of Applied Physics* vol. 119.10 (2016).
- [179] W.-J. Yin, X.-G. Gong, and S.-H. Wei. “Origin of the unusually large band-gap bowing and the breakdown of the band-edge distribution rule in the $\text{Sn}_x\text{Ge}_{1-x}$ alloys”. *Physical Review B* vol. 78.16 (2008).
-

- [180] S. Gupta, B. Magyari-Köpe, Y. Nishi, et al. “Achieving direct band gap in germanium through integration of Sn alloying and external strain”. *Journal of Applied Physics* vol. 113.7 (2013).
- [181] P. Moontragoon, Z. Ikoni, and P. Harrison. “Band structure calculations of Si–Ge–Sn alloys: achieving direct band gap materials”. *Semiconductor science and technology* vol. 22.7 (2007).
- [182] Y. Chibane and M. Ferhat. “Electronic structure of $\text{Sn}_x\text{Ge}_{1-x}$ alloys for small Sn compositions: Unusual structural and electronic properties”. *Journal of Applied Physics* vol. 107.5 (2010).
- [183] J. Gallagher, C. L. Senaratne, J. Kouvetakis, et al. “Compositional dependence of the bowing parameter for the direct and indirect band gaps in $\text{Ge}_{1-y}\text{Sn}_y$ alloys”. *Applied Physics Letters* vol. 105.14 (2014).
- [184] D. Rainko, Z. Ikonic, N. Vukmirovi, et al. “Investigation of carrier confinement in direct bandgap GeSn/SiGeSn 2D and 0D heterostructures”. *Scientific reports* vol. 8.1 (2018).
- [185] P. Moontragoon, R. Soref, and Z. Ikonic. “The direct and indirect bandgaps of unstrained $\text{Si}_x\text{Ge}_{1-x-y}\text{Sn}_y$ and their photonic device applications”. *Journal of Applied Physics* vol. 112.7 (2012).
- [186] H. Lin, R. Chen, W. Lu, et al. “Structural and optical characterization of $\text{Si}_x\text{Ge}_{1-x-y}\text{Sn}_y$ alloys grown by molecular beam epitaxy”. *Applied Physics Letters* vol. 100.14 (2012).
- [187] T. Wendav, I. A. Fischer, M. Montanari, et al. “Compositional dependence of the band-gap of $\text{Ge}_{1-x-y}\text{Si}_x\text{Sn}_y$ alloys”. *Applied Physics Letters* vol. 108.24 (2016).
- [188] S. Schulz, M. A. Caro, L.-T. Tan, et al. “Composition-dependent band gap and band-edge bowing in AlInN: A combined theoretical and experimental study”. *Applied Physics Express* vol. 6.12 (2013).
- [189] Y. Zhou, Y. Miao, S. Ojo, et al. “Supplementary document for Electrically injected GeSn lasers on Si operating up to 100 K ” (2020).
- [190] A. Adams, M. Silver, and J. Allam. “Semiconductor optoelectronic devices”. *Semiconductors and semimetals*. Vol. 55. Elsevier, 1998.
-

- [191] A. Aldukhayel, S. R. Jin, I. P. Marko, et al. “Investigations of carrier scattering into L-valley in $\lambda=3.5\mu\text{m}$ InGaAs/AlAs(Sb) quantum cascade lasers using high hydrostatic pressure”. *physica status solidi (b)* vol. 250.4 (2013).
- [192] H. Hong, L. Zhang, K. Qian, et al. “Limitation of bulk GeSn alloy in the application of a high-performance laser due to the high threshold”. *Optics Express* vol. 29.1 (2021).
- [193] Z. Chen, Z. Ikonic, D. Indjin, et al. “Design optimization of tensile-strained SiGeSn/GeSn quantum wells at room temperature”. *Journal of Applied Physics* vol. 129.12 (2021).
- [194] J. Lim, R. MacKenzie, S. Sujecki, et al. “Simulation of double quantum well GaInNAs laser diodes”. *IET optoelectronics* vol. 1.6 (2007).
- [195] D. Sytnyk and R. Melnik. “The Luttinger-Kohn theory for multiband Hamiltonians: A revision of ellipticity requirements”. *arXiv preprint arXiv:1808.06988* (2018).
- [196] Q. M. Thai, J. Chretien, M. Bertrand, et al. “GeSn optical gain and lasing characteristics modelling”. *Physical Review B* vol. 102.15 (2020).
- [197] P. Landsberg and A. Beattie. “Auger effect in semiconductors”. *Journal of Physics and Chemistry of Solids* vol. 8 (1959).
- [198] M. Takeshima. “Green’s-function formalism of band-to-band Auger recombination in semiconductors. Correlation effect”. *Physical Review B* vol. 26.2 (1982).
- [199] M. Takeshima. “Unified theory of the impurity and phonon scattering effects on Auger recombination in semiconductors”. *Physical Review B* vol. 25.8 (1982).
- [200] D. A. Duffy, I. P. Marko, C. Fuchs, et al. “The impact of band bending on the thermal behaviour of gain in type-II GaAs-based W-lasers”. *IEEE Journal of Selected Topics in Quantum Electronics* (2024).
- [201] S. Inada, M. Yoshita, M. Okano, et al. “Measurements of Cavity-Length-Dependent Internal Differential Quantum Efficiency and Internal Optical Loss in Laser Diodes”. Vol. 47.4R (2008).
- [202] M. Fagot, D. A. Díaz-Thomas, A. Gilbert, et al. “Interband cascade lasers grown simultaneously on GaSb, GaAs and Si substrates”. *Optics Express* vol. 32.7 (2024).
- [203] K. Kashani-Shirazi, K. Vizbaras, A. Bachmann, et al. “Low-Threshold Strained Quantum-Well GaSb-Based Lasers Emitting in the 2.5-to 2.7- μm Wavelength Range”. *IEEE Photonics Technology Letters* vol. 21.16 (2009).
-

- [204] R. Liang, J. Chen, G. Kipshidze, et al. “High-Power 2.2- μm Diode Lasers With Heavily Strained Active Region”. *IEEE Photonics Technology Letters* vol. 23.10 (2011).
- [205] A. Salhi, Y. Rouillard, J. Angellier, et al. “Very-low-threshold 2.4- μm GaInAsSb-AlGaAsSb laser diodes operating at room temperature in the continuous-wave regime”. *IEEE photonics technology letters* vol. 16.11 (2004).
- [206] V. Elyukhin, R. Peña-Sierra, B. R. Flores, et al. “The miscibility gap of $\text{In}_x\text{Ga}_{1-x}\text{Sb}_y\text{As}_{1-y}$ alloys”. *Journal of Physics: Condensed Matter* vol. 16.22 (2004).
- [207] M. Rattunde, J. Schmitz, C. Mermelstein, et al. “III-Sb-based type-I QW diode lasers”. *Mid-infrared semiconductor optoelectronics* (2006).
- [208] T. D. Eales, I. P. Marko, A. R. Adams, et al. “The nature of Auger recombination in type-I quantum well lasers operating in the mid-infrared”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 28.1 (2021).
- [209] E. Tournié, L. Monge Bartolome, M. Rio Calvo, et al. “Mid-infrared III-V semiconductor lasers epitaxially grown on Si substrates”. *Light: Science & Applications* vol. 11.1 (2022).
- [210] M. Rio Calvo, L. Monge Bartolomé, M. Bahriz, et al. “Mid-infrared laser diodes epitaxially grown on on-axis (001) silicon”. *Optica* vol. 7.4 (2020).
- [211] K. Volz, A. Beyer, W. Witte, et al. “GaP-nucleation on exact Si(001) substrates for III/V device integration”. *Journal of Crystal Growth* vol. 315.1 (2011).
- [212] J. Tatebayashi, A. Jallipalli, M. N. Kutty, et al. “Monolithically integrated III-Sb-based laser diodes grown on miscut Si substrates”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 15.3 (2009).
- [213] E. Delli, V. Letka, P. D. Hodgson, et al. “Mid-infrared InAs/InAsSb superlattice nBn photodetector monolithically integrated onto silicon”. *ACS Photonics* vol. 6.2 (2019).
- [214] M. R. Calvo, L. M. Bartolomé, M. Bahriz, et al. “Supplementary document for Mid-infrared laser diodes epitaxially grown on on-axis (001) silicon” (2020).
- [215] S. Sweeney, A. Phillips, A. Adams, et al. “The effect of temperature dependent processes on the performance of 1.5- μm compressively strained InGaAs(P) MQW semiconductor diode lasers”. *IEEE Photonics Technology Letters* vol. 10.8 (1998).
-

- [216] G. Tränkle, E. Lach, A. Forchel, et al. “General relation between band-gap renormalization and carrier density in two-dimensional electron-hole plasmas”. *Physical Review B* vol. 36.12 (1987).
- [217] N. Tansu, Y.-L. Chang, T. Takeuchi, et al. “Temperature analysis and characteristics of highly strained InGaAs-GaAsP-GaAs ($\lambda > 1.17 \mu\text{m}$) quantum-well lasers”. *IEEE Journal of Quantum Electronics* vol. 38.6 (2002).
- [218] C. R. Fitch, G. W. Read, I. P. Marko, et al. “Thermal performance of GaInSb quantum well lasers for silicon photonics applications”. *Applied Physics Letters* vol. 118.10 (2021).
- [219] E. P. O'Reilly and M. Silver. “Temperature sensitivity and high temperature operation of long wavelength semiconductor lasers”. *Applied Physics Letters* vol. 63.24 (1993).
- [220] D. A. Duffy, I. P. Marko, C. Fuchs, et al. “Performance characteristics of low threshold current $1.25 \mu\text{m}$ type-II GaInAs/GaAsSb W-lasers for optical communications”. *Journal of Physics D: Applied Physics* vol. 54.36 (2021).
- [221] P. Snowton and P. Blood. “The differential efficiency of quantum-well lasers”. *IEEE Journal of Selected Topics in Quantum Electronics* vol. 3.2 (1997).
- [222] G. Belenky, L. Shterengas, M. Kisin, et al. “Gallium antimonide (GaSb)-based type-I quantum well diode lasers: recent development and prospects”. *Semiconductor Lasers* (2013).
- [223] R. M. Corless, G. H. Gonnet, D. E. Hare, et al. “On the Lambert W function”. *Advances in Computational mathematics* vol. 5 (1996).
- [224] J. Chen, D. Donetsky, L. Shterengas, et al. “Effect of quantum well compressive strain above 1% on differential gain and threshold current density in type-I GaSb-based diode lasers”. *IEEE Journal of Quantum Electronics* vol. 44.12 (2008).
- [225] T. Lehnhardt, A. Herrmann, M. Kamp, et al. “Influence of GaSb and AlGaInAsSb as Barrier Material on $\sim 2.8\text{-}\mu\text{m}$ GaSb-Based Diode Laser Properties”. *IEEE Photonics Technology Letters* vol. 23.6 (2011).
- [226] W. Huang, B. Cheng, C. Xue, et al. “The band structure and optical gain of a new IV-group alloy GePb: a first-principles calculation”. *Journal of Alloys and Compounds* vol. 701 (2017).
-

- [227] C. A. Broderick, E. J. O'Halloran, and E. P. O'Reilly. "First principles analysis of electronic structure evolution and the indirect-to direct-gap transition in $\text{Ge}_{1-x}\text{Pb}_x$ group-IV alloys". *arXiv preprint arXiv:1911.05679* (2019).
- [228] T. T. McCarthy, A. M. McMinn, X. Liu, et al. "Molecular beam epitaxy growth and characterization of GePb alloys". *Journal of Vacuum Science & Technology B* vol. 42.3 (2024).
- [229] X. Liu, J. Zheng, Q. Huang, et al. "Investigation of temperature and H_2 on GePb/Ge multiple quantum well growth". *Journal of Physics D: Applied Physics* vol. 57.24 (2024).
- [230] H. Alahmad, A. Mosleh, M. Alher, et al. "GePb alloy growth using layer inversion method". *Journal of Electronic Materials* vol. 47 (2018).
- [231] X. Liu, J. Zheng, Q. Huang, et al. "Growth and characterization of GePb/Ge multiple quantum wells". *Journal of Alloys and Compounds* vol. 934 (2023).
- [232] C. A. Stephenson, W. A. O'Brien, M. Qi, et al. "Band anticrossing in dilute germanium carbides using hybrid density functionals". *Journal of Electronic Materials* vol. 45 (2016).
- [233] C. Stephenson, W. O'Brien, M. Penninger, et al. "Band structure of germanium carbides for direct bandgap silicon photonics". *Journal of Applied Physics* vol. 120.5 (2016).
- [234] T. Dey, A. W. Arbogast, Q. Meng, et al. "Why room temperature GeSn lasers need carbon". *2023 IEEE Silicon Photonics Conference (SiPhotonics)*. IEEE, 2023, pp. 1–2.
- [235] A. C. Kirwan, S. Schulz, and E. P. O'Reilly. "Nature of the band gap of Ge: C alloys: insights from hybrid functional density functional theory calculations". *Semiconductor Science and Technology* vol. 34.7 (2019).
- [236] C. A. Broderick, M. D. Dunne, D. S. Tanner, et al. "Electronic structure evolution in dilute carbide $\text{Ge}_{1-x}\text{C}_x$ alloys and implications for device applications". *Journal of Applied Physics* vol. 126.19 (2019).
- [237] I. A. Gulyas, C. A. Stephenson, Q. Meng, et al. "The carbon state in dilute germanium carbides". *Journal of Applied Physics* vol. 129.5 (2021).
-

- [238] M. S. Reza, T. Dey, A. W. Arbogast, et al. “Growth of tin-free germanium carbon alloys using carbon tetrabromide (CBr₄)”. *Journal of Applied Physics* vol. 134.18 (2023).
- [239] S. Wirths, D. Buca, and S. Mantl. “Si–Ge–Sn alloys: From growth to applications”. *Progress in crystal growth and characterization of materials* vol. 62.1 (2016).
- [240] J. Grant, G. Abernathy, O. Olorunsola, et al. “Growth of pseudomorphic GeSn at low pressure with Sn composition of 16.7%”. *Materials* vol. 14.24 (2021).
- [241] X. Liu, J. Zheng, C. Niu, et al. “Sn content gradient GeSn with strain controlled for high performance GeSn mid-infrared photodetectors”. *Photonics Research* vol. 10.7 (2022).
- [242] H. Zhao, G. Lin, Y. Zhang, et al. “Composition and strain effects on Raman vibrational modes of GeSn alloys with Sn contents up to 31% grown by low-temperature molecular beam epitaxy”. *Optical Materials* vol. 149 (2024).
- [243] J. Chrétien, Q. Thai, M. Frauenrath, et al. “Room temperature optically pumped GeSn microdisk lasers”. *Applied Physics Letters* vol. 120.5 (2022).
- [244] J. G. Fiorenza, J.-S. Park, J. Hydrick, et al. “Aspect ratio trapping: A unique technology for integrating Ge and III-Vs with silicon CMOS”. *ECS Transactions* vol. 33.6 (2010).
- [245] Y. De Koninck, C. Caer, D. Yulistira, et al. “GaAs nano-ridge laser diodes fully fabricated in a 300 mm CMOS pilot line”. *arXiv preprint arXiv:2309.04473* (2023).
- [246] H. Stanchu, G. Abernathy, J. Grant, et al. “Development of aspect ratio trapping growth of GeSn on Si for midwave infrared applications”. *Journal of Vacuum Science & Technology B* vol. 42.4 (2024).
- [247] L. Cerutti, D. A. Díaz Thomas, J.-B. Rodriguez, et al. “Quantum well interband semiconductor lasers highly tolerant to dislocations”. *Optica* vol. 8.11 (2021).
- [248] E. Kane. “Energy band structure in p-type germanium and silicon”. *Journal of Physics and Chemistry of Solids* vol. 1.1 (1956).
-

Appendix A

Further Notes on Bandstructure Calculations

Whilst researching the k - p theory, it was often difficult to decipher the origin of Hamiltonians presented in papers and textbooks. Different variants and notations were frequently adopted, which further complicated interpretation and made it difficult to ensure the correct matrix elements were being used. The aim of this Appendix is to further elucidate the theory outlined in the main text by reporting the derivation of the Dresselhaus-Kip-Kittel 3-band model, and how this leads to the 6-band Luttinger-Kohn model. Here, Löwdin renormalisation and formulation of the renormalised Hamiltonian matrix elements are stated explicitly in an attempt to provide a more accessible derivation than those provided in many textbooks. Google Drive links to the Python code used to model 6-band, 8-band bandstructure, gain and IVBA are then presented.

A.1 The 3-band Dresselhaus-Kip-Kittel Model

If relativistic effects are initially ignored, the basis states of the VB may be given by the three degenerate p -like ($l=1$) orbital states with x -, y - and z - symmetries, denoted as $|X\rangle, |Y\rangle, |Z\rangle$. As the basis states are by definition a complete set, the eigenfunctions at any \mathbf{k} may be given by a linear combination of the basis states, such that

$$|u_{n\mathbf{k}}\rangle = \sum_m c_m(\mathbf{k}) |u_{m0}\rangle, \quad (\text{A.1})$$

where $c_m(\mathbf{k}) = \langle u_{m0} | u_{n\mathbf{k}} \rangle$. Substituting into Eqn.4.5 and multiplying both sides by $\langle u_{n0} |$, the Schrödinger equation can be rewritten in matrix form as [140]

$$\sum_m \left[\left\{ E_n(0) + \frac{\hbar^2 k^2}{2m_0} \right\} \delta_{mn} + \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p}_{nm} \right] c_m(\mathbf{k}) = E_n(\mathbf{k}) c_n(\mathbf{k}), \quad (\text{A.2})$$

where $\mathbf{p}_{nm} = \langle u_{n0} | \mathbf{p} | u_{m0} \rangle$. Therefore in terms of the Hamiltonian matrix elements, \mathbf{H}_{nm} ,

$$\sum_m \mathbf{H}_{nm} c_m(\mathbf{k}) = E_n(\mathbf{k}) c_n(\mathbf{k}) \quad (\text{A.3})$$

$$\Rightarrow \sum_{m \neq n} \mathbf{H}_{nm} c_m(\mathbf{k}) = (E_n(\mathbf{k}) - \mathbf{H}_{nn}) c_n(\mathbf{k}). \quad (\text{A.4})$$

In order to accurately model the bandstructure, the effect of remote band interactions must be included in the model. This is nominally achieved through Löwdin's method of renormalisation [145]. The process involves dividing eigenstates of the system into two classes denoted A and B. Class A consists of states which interact strongly with each other and weakly with states in class B; these are the basis states. Class B states then refer to all other, energetically remote states. This is illustrated in Fig. A.1 a) for the case of a 3 or 6 state basis where the VB interactions are modelled explicitly. The interactions between class A and B states are then removed using an iterative process, resulting in renormalised interactions between the Class A states. To perform this renormalisation, the coefficients $c_n(\mathbf{k})$ are initially found by rearranging Eqn. A.4 and dividing between the two classes as [65]

$$c_n(\mathbf{k}) = \sum_{m \in A \neq n} \frac{\mathbf{H}_{nm}}{E_n(\mathbf{k}) - \mathbf{H}_{nn}} c_m(\mathbf{k}) + \sum_{j \in B \neq n} \frac{\mathbf{H}_{nj}}{E_n(\mathbf{k}) - \mathbf{H}_{nn}} c_j(\mathbf{k}). \quad (\text{A.5})$$

While Eqn. A.5 is valid for $n \in A$, the equation remains valid for $n \in B$, without the $n \neq m$ restriction. As such, Eqn. A.5 can be used to iteratively replace $c_j(\mathbf{k})$, giving [146]

$$c_n(\mathbf{k}) = \sum_{m \in A \neq n} \frac{\mathbf{H}_{nm}}{E_n(\mathbf{k}) - \mathbf{H}_{nn}} c_m(\mathbf{k}) + \sum_{j \in B} \frac{\mathbf{H}_{nj}}{E_n(\mathbf{k}) - \mathbf{H}_{nn}} \underbrace{\left(\sum_{m \in A} \frac{\mathbf{H}_{jm}}{E_j(\mathbf{k}) - \mathbf{H}_{jj}} c_m(\mathbf{k}) + \overbrace{\sum_{i \in B \neq j} \frac{\mathbf{H}_{ji}}{E_j(\mathbf{k}) - \mathbf{H}_{jj}} c_i(\mathbf{k})}^{= 0 \text{ (1st order)}} + \dots \right)}_{c_j(\mathbf{k})}. \quad (\text{A.6})$$

$$\Rightarrow \sum_{m \in A} \frac{c_m(\mathbf{k})}{E_n(\mathbf{k}) - \mathbf{H}_{nn}} \left(\underbrace{\mathbf{H}_{nm} + \sum_{j \in B} \frac{\mathbf{H}_{nj} \mathbf{H}_{jm}}{E_j(\mathbf{k}) - \mathbf{H}_{jj}}}_{\mathbf{U}_{nm}^A} - \mathbf{H}_{nm} \delta_{nm} \right) \quad (\text{A.7})$$

where \mathbf{U}_{mn}^A is the renormalised Hamiltonian. Rearranging Eqn. A.7, the renormalised eigenvalue problem to be solved then reads

$$\sum_{m \in A} (\mathbf{U}_{mn}^A - E_n \delta_{mn}) c_m(\mathbf{k}) = 0. \quad (\text{A.8})$$

\mathbf{H}_{nm} represents the Hamiltonian describing interactions between class A states and is determined using perturbation theory for degenerate states. To second order, the matrix elements are

$$\mathbf{H}_{nm} = \left[E_n(0) + \frac{\hbar^2 k^2}{2m_0} \right] \delta_{nm} + \underbrace{\frac{\hbar}{m_0} \sum_{\alpha} k_{\alpha} \sum_{m \in A \neq n} \mathbf{p}_{nm}^{(\alpha)} + \frac{\hbar^2}{m_0^2} \sum_{\alpha, \beta} k_{\alpha} k_{\beta} \sum_{j \in A \neq n} \frac{\mathbf{p}_{nj}^{(\alpha)} \mathbf{p}_{jm}^{(\beta)}}{E_0 - E_j}}_{= 0}. \quad (\text{A.9})$$

Since the basis states all have p-like symmetry, the momentum matrix elements of the form are zero for interactions between class A states. For the second term in \mathbf{U}_{nm}^A , $j \notin A$, the terms of the form \mathbf{H}_{nj} and \mathbf{H}_{jm} are therefore potentially non-zero, and can be written generally as

$$\mathbf{H}_{nj} = \frac{\hbar}{m_0} \sum_{\alpha} k_{\alpha} \sum_{\substack{j \in B \\ n \in A}} \mathbf{p}_{nj}^{(\alpha)}. \quad (\text{A.10})$$

We therefore obtain the renormalised Hamiltonian [146]

$$\mathbf{U}_{nm}^A = \left[E_n(0) + \frac{\hbar^2 k^2}{2m_0} \right] \delta_{nm} + \frac{\hbar^2}{m_0^2} \sum_{\alpha,\beta} k_\alpha k_\beta \sum_{\substack{j \in B \\ n, m \in A}} \frac{\mathbf{p}_{nj}^{(\alpha)} \mathbf{p}_{jm}^{(\beta)}}{E_0 - E_j} \quad (\text{A.11})$$

$$= E_n(0) \delta_{nm} + \sum_{\alpha,\beta} k_\alpha k_\beta D_{nm}^{\alpha,\beta} \quad (\text{A.12})$$

where we have defined

$$D_{nm}^{\alpha,\beta} = \frac{\hbar^2}{2m_0} \left[\delta_{nm} \delta_{\alpha,\beta} + \frac{2}{m_0} \sum_{j \in B \neq n, m} \frac{\mathbf{p}_{nj}^{(\alpha)} \mathbf{p}_{jm}^{(\beta)}}{E_0 - E_j} \right]. \quad (\text{A.13})$$

Since the valence band edge at Γ is taken to be the energetic zero, the renormalised Hamiltonian is given by the right-most term in Eqn. A.12, this is the form first devised by Dresselhaus, Kip and Kittel [144].

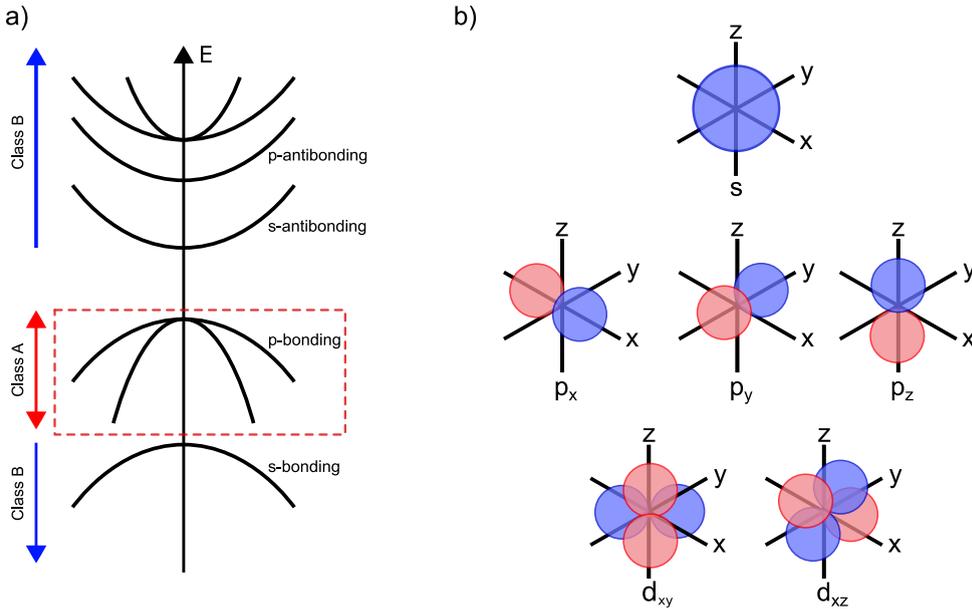


FIGURE A.1: a) schematic bandstructure using a 3-band model (with SO interactions neglected), illustrating the highest valence band as class A, and all other bands as class B, along with their typical orbital character at Γ . b) Symmetries of the s , p and d orbitals used during the calculations of the matrix elements in the Hamiltonian derivations presented in this work. Adapted from [65].

Whilst the interaction matrix elements remain unknown, we are now at a place where each matrix elements of the renormalised Hamiltonian may be derived as a function of k_x, k_y and k_z with a minimal number of fitting parameters, which may be obtained via experimental or atomistic modelling approaches. To do so, the non-zero coefficients $D_{nm}^{\alpha,\beta}$

must be determined through analysis of the state symmetries. For III-V and group IV materials, coupling between s, p and d states is considered, the orbitals of which are illustrated in Fig. A.1 b). The non-zero momentum matrix elements are given in Table A.1

We initially consider the case of the matrix elements D_{11} and D_{12} , where state 1 denotes $|X\rangle$ and 2 is $|Y\rangle$. Utilising Eqn. A.13 and the symmetry arguments in Table A.1, it can be shown that, for bulk material, the matrix elements are [65]

$$D_{11} = k_x^2 \langle X|p_x|S\rangle W'_s \langle S|p_x|X\rangle + k_y^2 \langle X|p_y|Z\rangle W'_p \langle Z|p_y|X\rangle \\ + k_z^2 \langle X|p_z|Y\rangle W'_p \langle Y|p_z|X\rangle + k_x^2 \langle X|p_x|d_a\rangle W'_d \langle d_a|p_x|X\rangle \\ + k_x^2 \langle X|p_x|d_b\rangle W'_d \langle d_b|p_x|X\rangle \quad (\text{A.14})$$

$$D_{12} = k_x k_y \left[\langle X|p_x|S\rangle W'_s \langle S|p_y|Y\rangle + \langle X|p_y|Z\rangle W'_p \langle Z|p_y|X\rangle \right. \\ \left. + \langle X|p_x|d_a\rangle W'_d \langle d_a|p_y|Y\rangle + \langle X|p_x|d_b\rangle W'_d \langle d_b|p_y|Y\rangle \right], \quad (\text{A.15})$$

where W'_s , W'_p and W'_d are fitting parameters characterising the strength of coupling for s-p, p-p and p-d interactions, respectively. By defining $W_s = W'_s \langle X|p_x|S\rangle^2$, $W_p = W'_p \langle X|p_y|Z\rangle^2$ and $W_d = \frac{2}{3} W'_d \langle X|p_x|d_a\rangle^2$, the matrix elements are then

$$D_{11} = k_x^2 (W_s + 2W_d) + (k_y^2 + k_z^2) W_p, \quad (\text{A.16})$$

$$D_{12} = k_x k_y (W_s - W_d + W_p). \quad (\text{A.17})$$

	Value
s-p coupling	$\langle X p_x S\rangle = \langle Y p_y S\rangle = \langle Z p_x S\rangle$
p-p coupling	$\langle X p_y Z\rangle = \langle Z p_x Y\rangle = \langle Y p_z X\rangle = \langle Z p_y X\rangle = \langle Y p_x Z\rangle = \langle X p_z Y\rangle$
p-d coupling	$\langle X p_x d_a\rangle = -\langle Y p_y d_a\rangle = -\sqrt{3} \langle X p_x d_b\rangle = -\sqrt{3} \langle Y p_y d_b\rangle$

TABLE A.1: The non-zero momentum matrix elements between s-, p- and d-states. Only two of the five total d orbitals are considered, denoted d_a and d_b . Taken from [65].

The other matrix elements form cyclic permutations of the above equations, resulting in the DKK Hamiltonian given by [144]

$$\mathbf{H}_{DKK} = \begin{pmatrix} Lk_x^2 + M(k_y^2 + k_z^2) & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & Lk_y^2 + M(k_z^2 + k_x^2) & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & Lk_z^2 + M(k_x^2 + k_y^2) \end{pmatrix} \quad (\text{A.18})$$

where $L \equiv -W_s - 2W_d$, $M \equiv -W_p$ and $N \equiv -W_s - W_p + W_d$, as per their notation.

A.2 The 6-band Luttinger-Kohn Model

In the previous model, the resulting bandstructure is triply degenerate at the zone centre, in stark contrast to experimental findings. The previous model was therefore subsequently adapted by Luttinger and Kohn to account for the spin-orbit interaction, which correctly reproduces the spin-orbit split off band energy [143]. The spin-orbit Hamiltonian is given by

$$H_{SO} = \frac{\hbar}{4m_0^2c^2} [(\boldsymbol{\sigma} \times \nabla V) \cdot \mathbf{p} + \underbrace{(\boldsymbol{\sigma} \times \nabla V) \cdot \mathbf{k}}_{=0}] = \frac{\hbar}{4m_0^2c^2} (\nabla V \times \mathbf{p}) \cdot \boldsymbol{\sigma}, \quad (\text{A.19})$$

where the rightmost term is neglected from the analysis since the momentum of the electron in the orbital, where the spin-orbit interaction is strongest, is significantly greater than the crystal momentum, ($\hbar\mathbf{k} \ll \mathbf{p}$) [148]. Here $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli spin matrices [76]

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{A.20})$$

which act on the $|\uparrow\rangle, |\downarrow\rangle$ eigenspinors as

$$\sigma_x |\uparrow\rangle = |\downarrow\rangle, \quad \sigma_x |\downarrow\rangle = |\uparrow\rangle, \quad (\text{A.21})$$

$$\sigma_y |\uparrow\rangle = i |\downarrow\rangle, \quad \sigma_y |\downarrow\rangle = -i |\uparrow\rangle, \quad (\text{A.22})$$

$$\sigma_z |\uparrow\rangle = |\uparrow\rangle, \quad \sigma_z |\downarrow\rangle = -|\downarrow\rangle. \quad (\text{A.23})$$

To account for the two distinct spin-projections, the basis set of the three band model is multiplied by the eigenspinors $|\uparrow\rangle, |\downarrow\rangle$ to form a six state basis. The new set is written in

the ‘LS’ basis as $|X \uparrow\rangle, |Y \uparrow\rangle, |Z \uparrow\rangle, |X \downarrow\rangle, |Y \downarrow\rangle, |Z \downarrow\rangle$ where \mathbf{L} is the orbital angular momentum as before, and \mathbf{S} is the spin. Due to the symmetry of p-like basis states, the non-zero matrix elements between states of the same spin are all of cyclical permutations of the form [142]

$$\Delta_{so} \equiv \frac{3i\hbar}{4m_0^2c^2} \langle X \uparrow | (\nabla V \times \mathbf{p})_y | Z \uparrow \rangle, \quad (\text{A.24})$$

such that [248]

$$H_{SO} = \frac{\Delta_{so}}{3} \begin{pmatrix} 0 & -i & 0 & 0 & 0 & 1 \\ i & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & -1 & 0 & i & 0 \\ 0 & 0 & -i & -i & 0 & 0 \\ 1 & i & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{A.25})$$

The full Luttinger-Kohn Hamiltonian is then obtained in the LS basis as

$$\mathbf{H}_{LK}^{LS}(\mathbf{k}) = \mathbf{D} \otimes I_{2 \times 2} + H_{SO}. \quad (\text{A.26})$$

It should be noted that Luttinger and Kohn redefine \mathbf{D} from that of the DKK model, such that

$$\mathbf{D} = \frac{\hbar^2 k^2}{2m_0} + \mathbf{D}_{DKK}, \quad (\text{A.27})$$

thus yielding [143]

$$\mathbf{D} = \begin{pmatrix} Ak_x^2 + B(k_y^2 + k_z^2) & Ck_xk_y & Ck_xk_z \\ Ck_xk_y & Ak_y^2 + B(k_z^2 + k_x^2) & Ck_yk_z \\ Ck_xk_z & Ck_yk_z & Ak_z^2 + B(k_x^2 + k_y^2) \end{pmatrix} \quad (\text{A.28})$$

where the coefficients A , B and C are linked to L , M and N as

$$A = \frac{\hbar^2}{2m_0} + L, \quad (\text{A.29})$$

$$B = \frac{\hbar^2}{2m_0} + M, \quad (\text{A.30})$$

$$C = N. \quad (\text{A.31})$$

Since Eqn. A.26 is not diagonal for $\mathbf{k} = 0$, and therefore cannot be solved exactly, it is customary to perform a transformation to the JM_J basis, for which exact solutions can

be found at the zone-centre [142]. Here $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the total angular momentum, and M_J is the associated z-axis projection. The new basis set reads [146], [149]

$$|u_{10}\rangle = \left| \frac{3}{2}, \frac{3}{2} \right\rangle = -\frac{1}{\sqrt{2}} |(X + iY) \uparrow\rangle, \quad (\text{A.32})$$

$$|u_{20}\rangle = \left| \frac{3}{2}, \frac{1}{2} \right\rangle = -\frac{1}{\sqrt{6}} |(X + iY) \downarrow\rangle + \sqrt{\frac{2}{3}} |Z \uparrow\rangle, \quad (\text{A.33})$$

$$|u_{30}\rangle = \left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{6}} |(X - iY) \uparrow\rangle + \sqrt{\frac{2}{3}} |Z \downarrow\rangle, \quad (\text{A.34})$$

$$|u_{40}\rangle = \left| \frac{3}{2}, -\frac{3}{2} \right\rangle = \frac{1}{\sqrt{2}} |(X - iY) \downarrow\rangle, \quad (\text{A.35})$$

$$|u_{50}\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X + iY) \downarrow\rangle + \frac{1}{\sqrt{3}} |Z \uparrow\rangle, \quad (\text{A.36})$$

$$|u_{60}\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X - iY) \uparrow\rangle - \frac{1}{\sqrt{3}} |Z \downarrow\rangle. \quad (\text{A.37})$$

The diagonalised spin-orbit Hamiltonian then takes the form [142]

$$H_{SO}^{JM_J} = \frac{\Delta_{so}}{3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{pmatrix}, \quad (\text{A.38})$$

illustrating that the inclusion of the spin-orbit interaction splits the sixfold degeneracy at the Γ -point into a $|3/2, M_J\rangle$ quadruplet and a $|1/2, M_J\rangle$ doublet. The quadruplet forms the doubly spin-degenerate LH and HH bands which are separated from the spin-degenerate spin-orbit split off band by an energy Δ_{so} .

From Eqns. A.26 and 4.32, the total Luttinger-Kohn Hamiltonian in the JM_J basis is given by [146]

$$H_{LK}^{JM_J} = \begin{pmatrix} E_{HH}(\mathbf{k}) & \sqrt{2}S(\mathbf{k}) & -S(\mathbf{k}) & 0 & -R(\mathbf{k}) & -\sqrt{2}R(\mathbf{k}) \\ & E_{LH}(\mathbf{k}) & Q(\mathbf{k}) & R(\mathbf{k}) & 0 & \sqrt{3}S(\mathbf{k}) \\ & & E_{SO}(\mathbf{k}) & -\sqrt{2}R(\mathbf{k}) & -\sqrt{3}S^*(\mathbf{k}) & 0 \\ & & & E_{HH}(\mathbf{k}) & \sqrt{2}S^*(\mathbf{k}) & -S^*(\mathbf{k}) \\ & & & & E_{LH}(\mathbf{k}) & Q(\mathbf{k}) \\ & & & & & E_{SO}(\mathbf{k}) \end{pmatrix} \quad (\text{A.39})$$

where the \mathbf{k} dependent functions are

$$E_{HH}(\mathbf{k}) = -\frac{\hbar^2}{2m_0} [(\gamma_1 + \gamma_2) k_{\parallel}^2 + (\gamma_1 - 2\gamma_2) k_z^2], \quad (\text{A.40})$$

$$E_{LH}(\mathbf{k}) = -\frac{\hbar^2}{2m_0} [(\gamma_1 - \gamma_2) k_{\parallel}^2 + (\gamma_1 + 2\gamma_2) k_z^2], \quad (\text{A.41})$$

$$E_{SO}(\mathbf{k}) = -\Delta_{so} - \frac{\hbar^2}{2m_0} \gamma_1 (k_{\parallel}^2 + k_z^2), \quad (\text{A.42})$$

$$Q(\mathbf{k}) = -\frac{\hbar^2}{\sqrt{2}m_0} \gamma_2 (k_x^2 + k_y^2) + \frac{\sqrt{2}\hbar^2}{m_0} \gamma_2 k_z^2, \quad (\text{A.43})$$

$$R(\mathbf{k}) = \frac{\sqrt{3}}{2} \frac{\hbar^2}{m_0} [\gamma_{av} (k_x - ik_y)^2 - \mu (k_x + ik_y)^2], \quad (\text{A.44})$$

$$S(\mathbf{k}) = \sqrt{\frac{3}{2}} \frac{\hbar^2}{m_0} \gamma_3 (k_x - ik_y) k_z. \quad (\text{A.45})$$

Since the Hamiltonian is Hermitian, the lower elements are not explicitly depicted in Eqn. A.39. γ_1 , γ_2 and γ_3 are the Luttinger parameters, which are fitting parameters related to the effective mass along given crystallographic directions, and are defined as linear combinations of A , B and C [146]:

$$\gamma_1 \equiv -\frac{2m_0}{3\hbar^2} (A + 2B), \quad (\text{A.46})$$

$$\gamma_2 \equiv -\frac{m_0}{3\hbar^2} (A - B), \quad (\text{A.47})$$

$$\gamma_3 \equiv -\frac{m_0}{3\hbar^2} C. \quad (\text{A.48})$$

Here $k_{\parallel}^2 = k_x^2 + k_y^2$, and γ_{av} and μ have been defined for concision, taking the values $\frac{1}{2}(\gamma_2 + \gamma_3)$ and $\frac{1}{2}(\gamma_3 - \gamma_2)$, respectively [149].

A.3 Python Code

The code developed over the course of this work for calculating bulk bandstructure using 6- and 8-band models as well as calculating material gain and IVBA are accessible using the following Google Drive [link](#)¹. It should be noted that these functions were written to be compatible with kpyomod version 0.15.1 and lower.

¹https://drive.google.com/file/d/1XCeeHwe29OoBA0lXH9tfei0-tXP4xd_I/view