Improved Telecom Laser Photonics: Investigating Confinement Effects within $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaSb}/\text{GaAs}$ Quantum Ring-in-Well Structures

Christopher John Redman MSci

Physics Department

This thesis is submitted for the degree of
Doctor of Philosophy

September 2023

Improved Telecom Laser Photonics: Investigating Confinement Effects within $In_xGa_{1-x}As$/GaSb/GaAs Quantum Ring-in-Well Structures

Christopher John Redman

Abstract

This project aims to develop high performance III-V gain media used for silicon-integrated near-infra-red (NIR) lasers with efficient light coupling. Molecular beam epitaxy (MBE) was used to successfully create advanced GaSb quantum ring (QR) materials operating in the telecommunication O-band near room temperature (RT). These structures were the focus of an opto-thermal study and were complimented with simulations to probe and understand the underlying carrier recombination mechanisms. A series of GaSb/GaAs quantum ring-in-well (RWELL) structures were grown, with a change in width of the dilute $In_{0.06}Ga_{0.94}As$ quantum wells (QWs), using a novel QR formation technique involving in-situ annealing. The structures were characterised with atomic force microscopy (AFM) and photoluminescence (PL) measurements. Power sweeps were performed on each sample within a range of 0.07 mW to 260 mW. Temperature sweeps were performed from 4.2 K to almost RT. Both temperature and power varied linewidth, emission energy and intensity were studied as a function of QW width. The temperature-dependent analysis, simulations, and study of electron confinement are to the best of our knowledge, performed for the first time. The RWELL gain media were found to have higher activation energies, and therefore greater thermal efficiencies at low temperature (LT), than standard GaSb/GaAs QRs. The activation energies reduce with QW width to a value no less than standard GaSb/GaAs QRs. It was concluded that the wavefunction overlap of the carriers plays a big role in the quenching of PL at 4 K. Strong confinement of the electrons increases the Coulomb energy, oscillator strengths and binding energies, hence expected to provide stable PL at RT. Further optimisations of the growth parameters are still required for the RWELL structures to outperform the reference sample at high temperature (HT). The samples characteristically blueshift with excitation power, demonstrating their type-II band alignment. The rate of blueshift is found to have a dependence on the QW width, and therefore the electron localisation.

Contents

Journal Publications

Carrier Dynamics of Type-II $In_{0.06}Ga_{0.94}As$/GaSb/GaAs Quantum Ring-in-Well Structures (Prospective 2025).

Conference Presentations

Monolithic GaSb Quantum Rings (QR) on Silicon for Telecom Lasers via Aspect Ratio Trapping (ART) Technology. H. Lei, Z. M. Jin, <u>C. Redman</u>, S. Kafanov, P. D. Hodgson, M. Hayne, Y. Pashkin and Q. Zhuang. UK Semiconductors, Sheffield, UK (2018).

IOP "Lightning Talk" Competition, UK Semiconductors, Sheffield, UK (2018).

Poster Presentations

Optically Efficient Type-II GaSb/GaAs Quantum Ring-in-Well for Telecom Lasers. C. Redman, P. D. Hodgson, M. Hayne, Q. Zhuang. 11th International Conference on Quantum Dots, Munich, Germany (2020).

Remarkably, the last place I'd have expected to encounter Jesus is midway through a PhD in physics. Those who followed my academic progression will know that I was completely saved from self-destruction. After many compelling experiences, significant coincidences, and consistently answered prayers during a time of such adversity; I can say both wholeheartedly and undoubtedly that although I have very little understanding on how to define Him, **I absolutely cannot refute the existence of a real, sovereign, ever-loving, merciful, gracious, patient, and all-powerful God.** I take great pleasure in devoting this work with thanks and reverence, entirely to His wonderful Majesty in eternal Glory. All praise and adoration to the Father, Son, and Holy Spirit! This was His victory.

My heart further swells with gratitude to everyone sent to guide, encourage, and care for me along my path, even more so for the few who after everything, still stand by and support me. Mum, I wish there were words to adequately say how much I appreciate you. None of this would have been possible without all your kindness, strength, and generosity over the years. Dad, you are sorely missed and I cherish the time we shared, you will always be in my heart. I love you both so much and couldn't have wished for better parents. I can't forget how you constantly strived to grant me the opportunities you never had. Darling Simba and my family; brothers and sisters at Fountains, Life Church, Mission, Acts, St. Pio Friary, and Light Church; The Salvation Army team; my mentors at The Well, and ROC; the amazing volunteers at Inchurches, Homeless Not Hopeless, Immanuel Project, Grace Kitchen/The Well, Bradford Central Foodbank, JKN Outreach, Monday Munch, Penny Appeal, and the Langar Seva Society; my supervisors Prof Young, Dr Zhuang, and Prof Robinson; my colleagues, especially Peter, Manus, and Dr J. Llorens; Dr S. Birner; staff at the workshops and stores; along with the teams at Jobcentre Plus, Bevan Healthcare, and the Homeless Outreach Partnership. Thank you so dearly to you all.

Thank you Father for everything, for reconciling us to You through Your precious Son. In Jesus' mighty name, I pray that everyone reading this is blessed with Truth: that they are personally invited to come intimately know and experience You. That they witness how You delightfully listen and tirelessly act on our behalf. Please grant them eternal life and prosperity, along with an abundance of fulfilment and limitless evidential satisfaction in the Holy Spirit's presence. May they receive shalom and the realisation they were never alone, or forsaken, and most importantly; loved beyond comprehension. May they understand that the agonies of life are not brought upon us as a direct consequence of Your will, though circumstances are always used for eventual good. The single set of footprints we looked down to see during our darkest hour were not our own, but the Lord's as He carried us to safety. Hosanna in the highest, Amen.

I declare that the contents of this thesis are the result of my own studies. I have exercised reasonable care to ensure that this research is original, and to the best of my knowledge does not break any UK law or infringe any third party's copyright or other intellectual property right. Explicit references acknowledge where I have consulted the work of others. I confirm that this work has not been submitted in whole, or in part, for any other degree or qualification, at this university, or any other academic institution.

- Christopher John Redman (Sept. 2023).

Chapter 1

Introduction

The rapid growth of data centres, high-performance computing requirements, and cloud-based applications has driven network traffic and power consumption at an exponential rate; therefore, accelerating the need for efficient, high-capacity, and compact optical interconnects (Figure 1.1). Moreover, the UK government has committed [1] to the roll-out of fibre-to-the-premises architecture in order to maximise the security and resilience of the UK's networks. Clear targets have been set to establish gigabit-broadband at 85 % of all UK premises by 2025, and nationwide (> 99 %) by 2030. As part of this, upstream laser diodes (LDs) which meet international network standards (see Ref. [2]) must be implemented at every site.
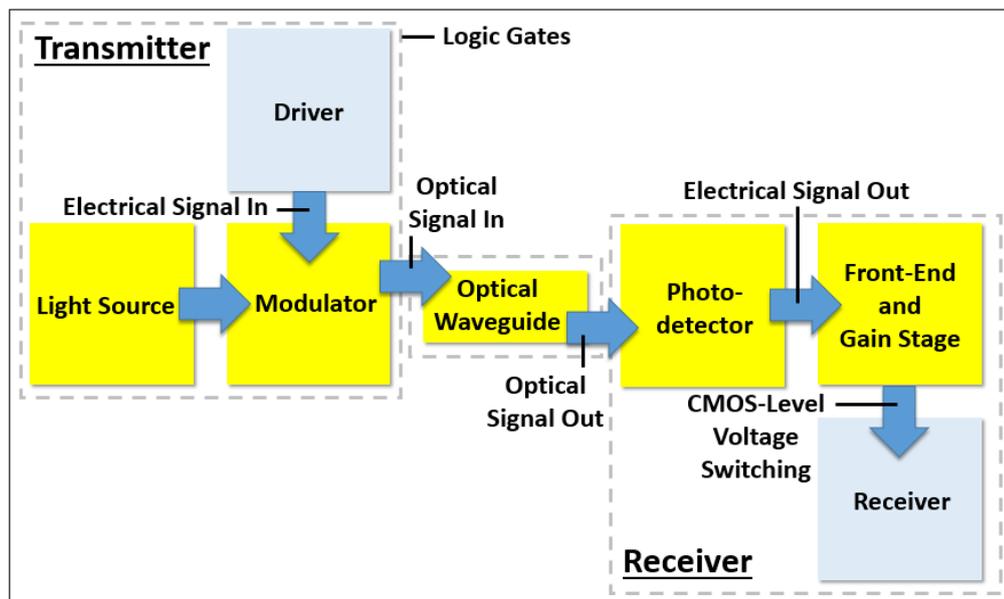
Figure 1.1: Schematic diagram of a general optical interconnect, i.e. a communication system that passes a binary signal from a transmitter to the receiver of an integrated circuit (IC). The connection is made by an optical waveguide which channels light akin to water inside a hollow pipe. The optical and electrical components are yellow and grey, respectively. Figure adapted with permission from [3].

The LD is essentially a LED placed inside an optical cavity, which uses reflectors to achieve stimulated emission, i.e. the interaction of an incident photon with an electron in an excited state. During this process, the electron de-excites across the bandgap (see Subsection 2.1.2) causing the emission of a second photon with the same frequency (ν), direction of travel, polarisation, and wave-phase (Θ) as the first. This is not to be confused with spontaneous emission (see Subsection 2.3.3).

Most telecommunication networks are hybrid integrated systems, whereby the use of an external LD is standard. A connection between the light source and waveguide is typically achieved by butt-coupling [4]. Not only is this approach known to reduce wall-plug efficiency, i.e. the ratio of the total optical output power (radiant flux [W]) to the input electrical power (consumer figure of merit for electrical-to-optical-power efficiency); it becomes increasingly difficult to scale down circuit size in accordance with Moore's law: i.e. the prediction that the number of transistors on an IC doubles approximately every two years.

Monolithic epitaxy is an attractive alternative to fabricate LDs, and is implemented at the Quantum Technology Centre, Lancaster University. Photonic ICs are cheaper to assemble, easier to mass-produce, and save energy by greatly reducing coupling losses between each component. The performance standards of opto-electronics grown by MBE and other techniques (e.g. metal-organic chemical vapour deposition (MOCVD) [5–9], pulsed laser deposition [10], etc.) have much improved, however they are still yet to satisfy the market requirements of this information age.

Many significant developments of the LD have been accomplished since the late 20th century; thus, shaping the telecommunications industry and influencing the way we communicate. Consequently, modern LDs outperform other technologies, e.g. gas lasers, and dye lasers etc. Vertical-cavity surface-emitting laser (VCSEL) structures [11–15] (see Appendix A) are a popular choice of LD for high-speed (gigabit) networking solutions, particularly in the context of short-range fibre-optic communication [12], because of the following coveted properties:

- Low manufacturing cost,

- Feasibility of on-chip testing and characterisation to ensure successful production during the stages of device processing (quality control),

- Compact structure designs which facilitates down-scaling towards dense integration of photonic devices (output power is proportional to the number of emitters in a 2D VCSEL array),

- High-quality, circular output beam that is compatible with optical fibres,

- Longevity of structures and reliable lasing with prolonged operation,

- High external quantum efficiency (η_{ext}),

- Low threshold current. This is the least amount of supplied current necessary before lasing can take place, i.e. the gradual change from spontaneous to stimulated emission,

- High-speed modulation,

- Wide modulation bandwidth

- Precision of wavelength, λ, is ideal for increasing bandwidth by wavelength-division multiplexing, i.e. the superposition of data streams, at different frequencies in a given band, for simultaneous transmission of data packets through a single optical fibre.

The modulation speed is limited by the damping of the intrinsic modulation response (inter-symbol interference), self-heating and parasitic impedance [14] for the reasons outlined below. Further information can be found in Ref. [15]; where the authors present a model to analyse the speed response of VCSELs, by taking self-heating into account for the first time.

VCSEL operating temperatures can often reach 85 °C with high current throughput, due to the dissipation of phonons. Effective cooling systems are required to maintain optimal continuous-wave (CW) lasing and prevent thermal degradation of photonic ICs. Operating expenses are therefore increased, along with the carbon (C) footprint of an organisation. Hence, the latest communication systems require LDs which are able to maintain their RT modulation bit rates at higher T, without modifying the operating parameters.

Even the most precise LD cannot produce perfect monochromatic emission. A light pulse (data packet) is quasi-monochromatic, i.e. a composition of slightly different λ-values. Chromatic dispersion refers to the broadening of these spectral components as they travel, at dissimilar phase velocities, through an optical fibre core. This effect is dependent on the fibre length and spectral width. Signal quality (amount of distortion), depends on the overlap of consecutive light pulses. Chromatic dispersion is thus a limiting factor on the optimum data transmission rate (bit rate) [bits/s], and the minimum number of optical repeaters needed for a long-haul telecommunications link. On the contrary, chromatic dispersion helps to reduce the impact of non-linearity in multi-mode fibres. This can lead to improved network performance with careful design and management.

Signals also attenuate in optical fibres, therefore the communications sector is in-

terested in light sources which emit at λ-values where transmission losses, i.e. the absorption of data packets, are minimised. Table 1.1 lists these windows in ascending order.

Telecommunication Band	Wavelength Range (nm)
850 nm	$810 - 890$
Original (O)	$1260 - 1360$
Extended (E)	$1360 - 1460$
Short (S)	$1460 - 1530$
Conventional (C)	$1530 - 1565$
Long (L)	$1565 - 1625$
Ultra-Long (U)	$1625 - 1675$

Table 1.1: Wavelength bands for optical communication.
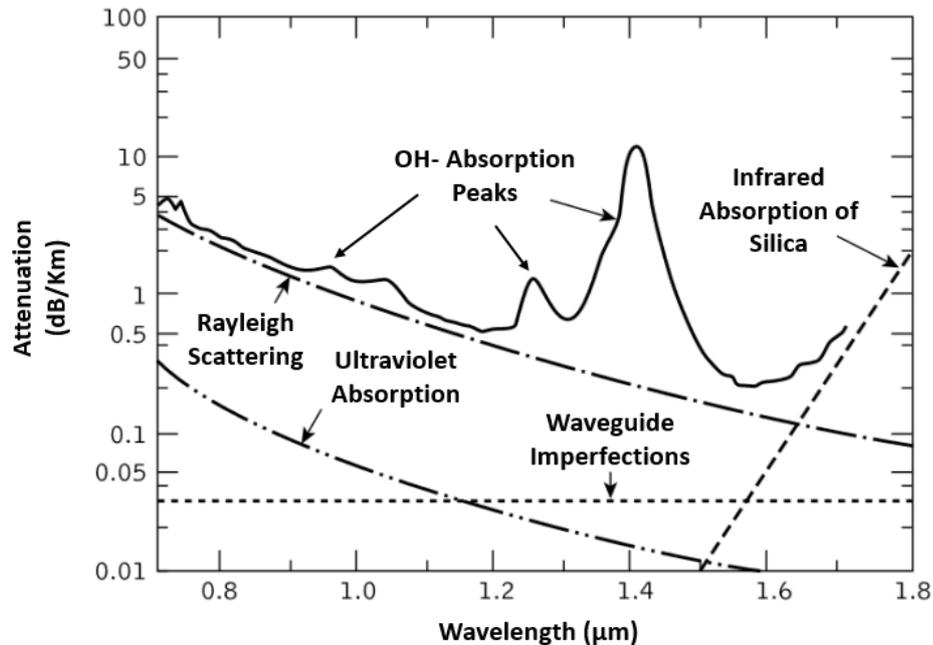
Figure 1.2: The measured attenuation inside a low loss single mode silica-based optical fibre, is plotted as a function of wavelength (solid line). The dashed and dotted lines represent the calculated attenuation for some of the loss mechanisms contributing to the net fibre loss are included. Figure adapted with permission from [16].

In typical fibres, the dispersion (Figure 1.3) increases between the O- and C- bands, whereas the attenuation per km is highest (lowest) in the O-band (C-band) (Figure 1.2). Due to this, LDs operating near 1310 nm are more suited to passive-optical networks and Ethernet architectures (local, metropolitan, or wide area networks) etc. Alternatively, LDs operating in the C-band are desirable for reliable communication across ultra-long distances (e.g. submarine cables). Although, that is not to disregard their potential effectiveness as next-generation, high-capacity, and short-reach
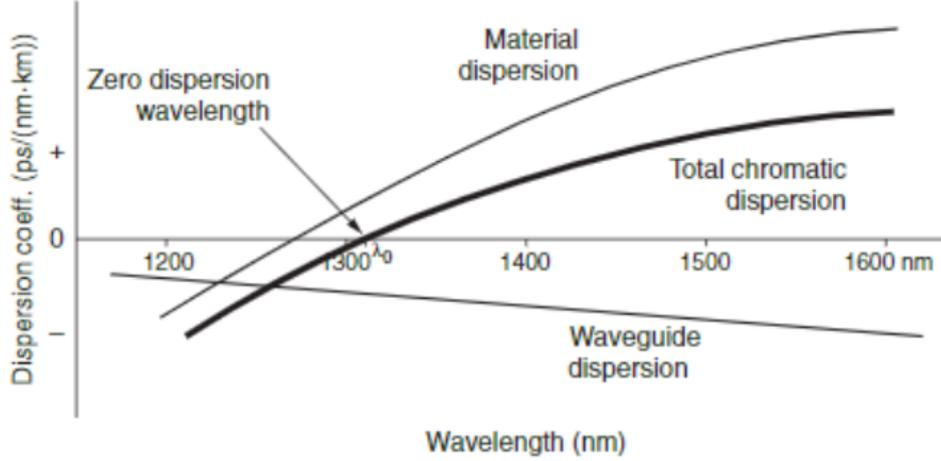
Figure 1.3: Plot of the total dispersion coefficient, per km, as a function of λ for a standard optical fibre. The total dispersion is a combination of the material and waveguide dispersions. Here, λ_0 is the zero dispersion wavelength, i.e. the point at which the material and waveguide dispersions are counterbalanced. Figure reproduced with permission from [17].

optical interconnects. In fact, upstream and downstream data can be sent through a single fibre simultaneously (full-duplex communication), with the implementation of two LDs; one operating in the O-band and the other in the C-band.

The O-band is the favoured optical channel in high-volume, state-of-the-art data centres and other cost-sensitive markets, because this λ-range eliminates operating expenses associated with the power consumption of dispersion management technologies. Addressing the global need for lower power consumption is a step towards the realisation of greener data centres. On this note, it is essential to improve the speed and efficiency of in-house connections, hence further motivating the gain media presented in this thesis.

Since silicon (Si) is a very cheap and abundant material, there is a general longing in the semiconductor industry to develop Si-based lasers for improved economy and streamlined fabrication. Thus, indirectly reducing the carbon footprint of any organisation that employs this technology. It would be ideal to fabricate a telecommunication LD using bulk Si, though doing so is very challenging due to its indirect bandgap. Instead, group III-V alloys are often selected for an active region, due to their: natural emission in the O-L bands, capability of fine-tuning a direct bandgap, and ease of epitaxial growth regardless of epi-structure intricacy.

One of the most common quantum structures found in commercial 1.3 μm LDs, are $In_xGa_{1-x}AsP$ QWs lattice-matched to InP substrates [18]. InP substrates are relatively fragile and restricted in wafer size availability; larger substrates can be manufactured with GaAs at a fraction of the fabrication cost. InP-based LDs also tend to suffer from poor thermal stability due to weak quantum and optical confine-

ments.

GaAs-based gain media typically emit at an unsuitably short λ for mid-to-NIR applications. Additionally, (In,Al)GaAs compounds are inherently susceptible to recombination-enhanced defect reactions [19–21]. Phonon dissipation (localised at defect sites) increases the temperature of the lattice, thereby increasing the mobility and density of defects; resulting in a reduced characteristic temperature (Equation 5.4). Hence, necessitating developments of GaAs-based materials to allow efficient operation in the telecommunication bands, by tailoring E_g along with the confinement of defects and carriers.

Type-I InAs/GaAs materials lead contemporary GaAs-based research results; the best of which currently only operate around 1.3 μm and therefore set a reasonable benchmark to compare our GaSb/GaAs-based samples against and gauge market competition. Type-II GaSb/GaAs active regions are less explored, but provide several key benefits over traditional type-I systems (limited by choice of application), due to an extremely deep hole localisation potential [22] and separation of charges. GaSb/GaAs LDs have better temperature stabilities, since holes find it difficult to escape the nanostructures at elevated T, and recombination rates are limited by their concentration. A reduction in binding energy, E_b (see Subsection 2.3.2) also causes type-II materials to emit at a longer λ than either of their separated compounds [23]. Furthermore, samples can exhibit Mott transitions, [24] and the Aharanov-Bohm effect [25]; both of which are absent in type-I nanostructures. Excitons have longer radiative lifetimes [26] and a greater population inversion, due to the spatial separation of carriers (weaker Coulomb interaction) [27]. Longer recombination lifetimes are not ideal for lasing, since they cause weaker PL emission, especially at HT. The strong electronic confinement reduces at short distances away from the GaSb interface; thus lowering E_b and the activation energy, E_a [28]. Further work on GaSb/GaAs QD- and QR- based architectures is required to enhance the rate of radiative recombination. It is a question of how to do this effectively.

A possible solution is to increase the electron density close to the nanostructures. The confinement far away from the QRs can be increased with the implementation of a thin cladding layer above and below the QR layers, using a material with a relatively large CB offset (e.g. $Al_xGa_{1-x}As$ [29] or AlSb [30]). This will blueshift emission due to an increase of the in-plane exciton reduced mass, resulting from quantum confinement and VB-mixing effects [31]. Improving type-II PL in this way is likely to deteriorate the interface quality, and introduce significant strain into the system if the materials are not lattice-matched. This is undesirable since there is already a substantial accumulation of strain during growth due to high lattice

6

mismatch between and GaAs and other common III-V's, e.g. InAs (7.2 %) and GaSb (7.8 %). Therefore making it easier for dislocations to form, yet harder to extend λ.

As for tackling dislocations, thermal annealing comes with the cost of enhanced diffusion of group-III atoms and variation of the morphology of buried nanostructures; adding deviations to the expected opto-electronic properties. Despite this, many techniques such as in situ thermal annealing [32, 33] have been well established to improve interface smoothness, and lessen the impact dislocations have on the structural quality of devices; consequently reducing the emission linewidth, i.e. the full width at half maximum (FWHM), Γ_{PL}.

Several approaches within epitaxial overgrowth [34] and aspect ratio trapping [35] have been proposed to reduce the density of dislocations and suppress upward threading, including; two-step growth [36]; selective area growth [37]; various types of buffer layers [6, 7, 38, 39]; strained-layer superlattice (SLS) filtration [40–45] and thermal annealing [32, 33, 46–48]. Despite improving threading dislocation (TD) density, these techniques significantly increase structural complexity and further inhibit the realisation of well-integrated Si-based opto-electronics.

Attempts to achieve GaAs-based lasing at $\lambda = 1.55$ μm have been made over the years using various material systems, e.g. dilute nitrides [49] and bismides [50, 51] etc. Significant Auger recombination in dilute nitrides hinders radiative recombination. Whereas bismides are still somewhat obscure in comparison to other well-established materials. Both dilute nitrides (e.g. InGaN/GaN) and bismides (e.g. GaAsBi) exhibit poor carrier lifetimes due to defect-related recombination [52].

Structural modifications can also extend λ, such as with bilayer 0D nanostructures [53, 54]. It has been demonstrated that coupling two layers of QRs enhances luminescence and lengthens recombination lifetimes [54]. Coupled nanostructures form tensile-strained pockets in the sandwiched GaAs matrix. Although, increasing λ in this manner is unlikely to be as effective as intentionally lowering the conduction band (CB) edge (e.g. alloying), without significant defect densities. Furthermore, thin separation layers in between coupled nanostructures render the growth surface prone to the formation of floating layers of antimony (Sb), and indium (In), than than their decoupled-counterparts during arsenic-antimony (As-Sb) exchange, or gallium-indium (Ga-In) exchange, respectively. Therefore, they have an inherently wider Γ_{PL}. There is also little scope to tune the density of electrons surrounding the nanostructures this way, such as with layers of graded composition [50] or SLSs [55]. Stacking QRs also leads to a larger Γ_{PL} due to inhomogeneous distribution of sizes in each subsequent layer [56–58], along with unintended vertical stacks of nanoscopic

islands.

So, a way of keeping electrons close to the nanostructures is required, along with a better way of controlling the change in nanostructure morphology during MBE growth. If λ can also be extended to 1.55 µm, then an opportunity is presented for such a gain media to launch the latest generation LD technologies into a wider market within the long-distance communications sector. Theoretically, all three issues can be simultaneously addressed, by putting the QRs in a container; i.e. their insertion into a QW [59], which acts as an electron reservoir above the nanostructures.

Generally speaking, the III-V/Si integration platform has gained traction amongst an enthusiastic research community, and often regarded as the most promising route towards commercial deployment of next-generation photonic devices. These systems are still in their infancy and remain at the frontier of nanotechnology research. Effort is primarily focused on tackling the fundamental challenges of growth, outlined in Subsection 2.2.1, which commonly lead to the failure of conventional Si-based QW lasers [60]. Specifically, large defect densities and the interfacial polar/non-polar behaviour of III-V/Si cause devices to degrade after long periods of constant current stress. Whilst remaining open to the possibility of other dissipative mechanisms; Liu et al. [21], maintain that the gradual degradation of laser performance [61] is predominantly correlated with slow, recombination-enhanced climb [19] of TDs. GaAs-based QW LDs grown on Si substrates therefore exhibit poor performance characteristics and short CW longevities of up to $\sim$ 200 h at 30 °C [8].

Notably, QDs/QRs have surpassed QWs in the realisation of higher-performance LDs [11, 20, 21, 40, 41, 53, 62–78], light-emitting diodes (LEDs) [79–82], detectors [56, 77, 83, 84], and charge-based memory devices [5, 85, 86]. Additionally, these 0D nanostructures have unlocked the possibility of quantum communication and computation [87], and with it, a quantum internet [88]; via single-photon sources [89–91]. QDs and QRs are less sensitive to non-radiative recombination centres than higher dimensional structures, owing to their small size [64]. They were also found to be less strained and yield fewer defects in their near vicinity, in comparison to QWs [92]. The plausibility of increasing QD density by over 50 % is suggested in the work of Ref. [73]. This is of importance, since the intensity of a LD can be improved by increasing the ratio of radiative 0D nanostructures (e.g. QDs and QRs) to defects [74]. Despite research efforts ceasing to form active regions entirely from bulk and 2D materials, in favour of exploring the more interesting lower-dimensional nanostructures; arrangements of multiple quantum wells (MQWs) and SLSs are still often implemented as crucial components of innovative designs, e.g. to increase modal gain [75].

This field has seen swift progression of pioneering work towards monolithic III-V/Si technology operating at 1.3 µm. Namely, the first pulsed- [62] and CW- [63] mode operation of a III-V QD laser on a Si(001) substrate, at RT. The first reliability study [21] on a III-V/Si QD LD, where the absence of any sudden failure over 2700 h of CW operation at 30 °C is reported. A record-breaking extrapolated longevity exceeding 22 years was achieved at 80 °C, with an InAs/GaAs narrow ridge-waveguide Fabry-Perot cavity laser [41]. This dramatically increases to over $\sim$ 1150 years, when aged at 35 °C, for a similar structure presented in the work of Ref. [20]. A record HT CW operation up to 165 °C was maintained with a ridge-waveguide Fabry-Perot laser containing an InAs/InGaAs/GaAs QD-in-QW stack active region [40]. RT CW threshold current densities have reduced to values near 50 A/cm^2 [40, 64], and are comparable to the RT characteristics of QD lasers on GaAs substrates. The lowest recorded CW threshold current density reported for a GaAs-based QD LD operating at RT, is 8.8 A/cm^2 [65], however the λ here is slightly less than ideal at 1.22 µm. Relatively high output powers of 176 mW (20 °C) [66], 105 mW (RT) [64], 63 mW (80 °C) [67], and 19 mW (125 °C) [68] have also been demonstrated. These achievements would not be possible without being able to manufacture QD lasers on Si with TD densities of the order $10^6 - 10^8$ cm^{-2} [69].

Importantly, the targets of this PhD project are:

- Design, grow and characterise GaSb QRs in an InGaAs MQW structure on a GaAs substrate; which exhibit bright and efficient RT PL at 1.31 µm.

- Develop and optimise this structure to achieve emission at longer λ; aim to show a shift regardless of intensity/efficiency up to 1.55 µm).

- Understand the dominant radiative and non-radiative recombination mechanisms at play.

The main scientific novelties of this thesis are the epitaxial design, simulation, growth, and optical properties of NIR gain media. Discussion focuses on the integrated intensity, Γ_{PL}, and reduced quenching at LT observed with this novel GaSb QRs-in-asymmetric-QW structure. A comprehensive account of fundamental concepts critical to the understanding of results is given in the next chapter. The focus of Chapter 3 is to introduce the experimental techniques and characterisation methods used to obtain and process the datasets. The considerations to maximise the accuracy of results are also included. Theoretical and experimental results are then presented and discussed in Chapters 4 and 5 respectively. Conclusions are summarised in Chapter 6. Finally, Chapter 7 covers prospective work had setbacks not caused a drastic reduction of available lab-time.

Chapter 2

Fundamental Concepts

Nanostructures confined in 3D have revolutionised semiconductor laser gain media. Due to their importance, this chapter is divided into three major topics to guide the reader from the conceptualisation of semiconductors and band engineering, to a way in which heterostructures with low-dimensional features can be manufactured, through to how they produce light and their optical properties. Specifically, this chapter explores the basics of semiconductor physics, photo-absorption and fluorescence; along with the monolithic growth of gain media using MBE, and the difficulties faced during nanofabrication. By the end of the chapter the reader will have an intuitive grasp of how semiconductor lasers work, along with the hypotheses outlined in Section 5.1.

The current section begins with a brief introduction to quantum theory in order to provide a foundation for explaining the electronic transitions inside the nanostructures presented within this work. For completeness, an initial derivation is included to show where fermions and bosons come from mathematically. This leads to an introduction of the Pauli exclusion principle, along with the singlet and triplet states. The simple Kronig-Penney model is employed to illustrate how bands of *allowed* states arise. The T-dependence of the bandgap is then discussed, followed by a common method to generalise the 1D Brillouin zone (BZ) to the 3D case. Moreover, both isotropic and anisotropic cases of the effective mass approximation are discussed to provide insight on how strain impacts the movement of electrons. The electronic density of states (DoS) is then introduced and given for each dimensional scenario, followed by an explanation of how states are filled with increasing T in order to derive carrier density. This leads to a discussion on free carriers and the equilibrium carrier concentration. This section closes with consideration of direct and indirect bandgaps and the various types of band structure to show how band alignments may differ.

2.1 Band Theory of Semiconductors

2.1.1 Brief Introduction to Quantum Mechanics

Quantum theory is the foundation of modern physics and attempts to explain the nature of matter and energy at the (sub)atomic level. The fundamental principles of quantum mechanics are:

Postulates of Quantum Mechanics

1. Everything that can be known of a state within a quantum system is completely specified by a wavefunction, $\psi(\mathbf{r}, t)$, where $\boldsymbol{r}$ and t are the position vector and time, respectively. All possible states of this system are encoded into a Hilbert space, i.e. a generalised form of Euclidean space.

2. Every physically measurable quantity (an observable) in classical mechanics must have a linear Hermitian operator associated to it in quantum mechanics, whose eigenvectors construct a complete basis (set of all possible system states). An operator, $\hat{A}$, is a mathematical construct which acts on a function to produce a new function, i.e. a mapping between spaces. A matrix, $\mathbf{A}_{ij}$, represents a function: it is a rectangular array of i rows and j columns, where each element is either a scalar or an expression. For $\mathbf{A}_{ij}$ to be Hermitian, it must be square and equivalent to its conjugate transpose, $\mathbf{A}_{ij}^{\dagger} = \mathbf{A}_{ji}^{*}$. This ensures that the eigenvalues associated with the possible measurement outcomes are real.

3. When the operator $\hat{A}$, associated with an observable acts on the wavefunction; it returns the corresponding measurement value of that observable multiplied by the wavefunction, i.e. only eigenvalues, Ξ, which satisfy $\hat{A}\psi = \Xi\psi$ are measured. The state does not necessarily have to be an original eigenstate of $\hat{A}$. An arbitrary state can be expanded into a complete set of eigenvectors of $\hat{A}$ ($\hat{A}\psi_i = \Xi_i\psi_i$) as $\psi = \sum_i \Xi_i\psi_i$, where the linear sum of independent functions may be infinite in principle. The probability of observing the eigenvalue Ξ_i is $|\Xi_i|^2 = \Xi_i^*\Xi_i$. Importantly, the state of a system is set to the eigenstate $|\Xi_i\rangle$ immediately after a measurement yields the eigenvalue Ξ_i.

4. The time-dependent Schrödinger equation below describes how the wavefunction evolves in time:

$$\hat{\mathcal{H}}\psi(\boldsymbol{r}, t) = i\hbar\frac{\mathrm{d}\psi}{\mathrm{d}t}. \qquad (2.1)$$

The Hamiltonian is denoted by $\hat{\mathcal{H}}$ and represents the sum of potential and kinetic energies of all particles within a given system. The imaginary unit

and the reduced Planck constant are denoted by i, and $\hbar = h/2\pi$, respectively; where h is the Planck constant. Additionally, a unitary operator $\hat{U}(t_0, t)$ transforms a state at time, t_0, to a future state at time, t. For an operator to be unitary, it must be bounded and linear in the Hilbert space and satisfy $\hat{U}^*\hat{U} = \hat{U}\hat{U}^* = \mathbb{1}$, where $\mathbb{1}$ is the identity operator.

The time-independent Schrödinger equation (Equation 2.2) is a linear partial differential equation that predicts the non-relativistic behaviour of an electronic wavefunction in a closed quantum system.

$$\left[-\frac{\hbar^2}{2m^*}\hat{\nabla}^2 + V(\boldsymbol{r}) \right] \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}). \tag{2.2}$$

Where m^* is the effective mass (see Section 2.1.2), $\hat{\nabla}$ represents the 3D vector differential operator with respect to each of the Cartesian coordinate axes. The potential is dependent on spatial coordinates, hence represented by $V(\boldsymbol{r})$. Energy eigenvalues are denoted by E. All other symbols have their usual meanings.

5. The probability, Pr, of finding a specific measurement outcome, Ξ_i, of an observable is given by the Born rule:

$$\Pr(\Xi_i) = |\langle \Xi_i | \psi \rangle|^2. \tag{2.3}$$

6. The expectation value, $\langle A \rangle$, of the observable corresponding to operator $\hat{A}$ is an average of many measurements. It is given by computing the expectation value integral with respect to the wavefunction:

$$\langle A \rangle = \frac{\int_{-\infty}^{\infty} \psi^* \hat{A} \psi}{\int_{-\infty}^{\infty} \psi^* \psi}. \tag{2.4}$$

Constraints on the Wavefunction

In order to represent a real, physically observable system, the following constraints on the wavefunction boundary conditions must be met:

- The wavefunction must be a solution to the Schrödinger equation.

- The wavefunction, and its derivative, must be continuous everywhere.

- The wavefunction (probability amplitude) returns a probability density $|\psi(\boldsymbol{r}, t)|^2$, i.e. the likelihood of finding a particle between x_{pos} and $(x_{\text{pos}} + \mathrm{d}x_{\text{pos}})$. In order for this value to be physically meaningful, the probability must be nor-

malised such that the total probability for all space is equal to unity.

- The total wavefunction must include electronic spin.

Fermions vs. Bosons

Electrons are indistinguishable particles. Thus the probability density of the electron cloud should not change upon particle exchange, so the new two-particle wavefunction must have the following property:

$$|\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)|^2 = |\psi(\boldsymbol{r}_2, \boldsymbol{r}_1)|^2. \tag{2.5}$$

This will only be true if the wavefunctions before and after permutation between different orbitals are related by a phase factor $e^{i\Theta}$:

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = e^{i\Theta}\psi(\boldsymbol{r}_2, \boldsymbol{r}_1). \tag{2.6}$$

There is an implicit requirement that a double permutation ought to reproduce the original wavefunction state, therefore:

$$e^{i\Theta}e^{i\Theta} = e^{2i\Theta} = 1. \tag{2.7}$$

This is only possible if Θ is a multiple of π, or 0. The exchange operator, $\hat{E}_{\leftrightarrows}$, is a convenient tool which takes this into account. Indeed the eigenvalues are limited to $+1$ and -1, causing the wavefunction to be either symmetric or antisymmetric respectively, upon exchange, as shown by Equation 2.8:

$$\hat{E}_{\leftrightarrows}|\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)\rangle = \pm|\psi(\boldsymbol{r}_2, \boldsymbol{r}_1)\rangle. \tag{2.8}$$

In the case where $\Theta = \pi$, the phase factor $e^{i\Theta} = -1$; so the wavefunction output changes sign upon interchanging coordinates. These particles are known as *fermions* and have half-integer *spin*. Here, "spin" refers to an intrinsic angular momentum proportional to $\hbar$.

Actually, any wavefunction used to describe multiple fermions must also be anti-symmetric with respect to their permutation. This includes those approximated as products of single fermionic functions, ϕ_i, where Coulomb interactions are ignored for simplicity:

$$|\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_i)\rangle \approx \phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2)\ldots\phi_i(\boldsymbol{r}_i). \tag{2.9}$$

In the context of this work, the wavefunction subscript distinguishes each of the electrons contained within the system. Notably, there is a condition that $\hat{E}_\leftrightarrows$ commutes with the Hamiltonian in order to provide common eigenfunctions. This is not satisfied by Equation 2.9, unless $\phi_1 = \phi_2$.

For completeness, when $\Theta = 0$, and $e^{i\Theta} = +1$: the wavefunction states before and after a single permutation are identical. These particles are known as *bosons* and they have integer spin, e.g. photons, gluons, deuterium nuclei etc.

It follows from a linear combination of all permutations, that for the antisymmetric (ψ^-) and symmetric (ψ^+) states:

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}|\phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2) - \phi_1(\boldsymbol{r}_2)\phi_2(\boldsymbol{r}_1)\rangle, \qquad (2.10\text{a})$$

$$|\psi^+\rangle = A_{12}|\phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2) + \phi_1(\boldsymbol{r}_2)\phi_2(\boldsymbol{r}_1)\rangle + A_{11}|\phi_1(\boldsymbol{r}_1)\phi_1(\boldsymbol{r}_2)\rangle + \ldots$$
$$\ldots A_{22}|\phi_2(\boldsymbol{r}_2)\phi_2(\boldsymbol{r}_1)\rangle. \quad (2.10\text{b})$$

Where the A_{11}, A_{22}, and A_{12} terms represent the combined expansion and normalisation constants.

Pauli Exclusion Principle

Clearly, the antisymmetric state collapses to zero when $\phi_1(\boldsymbol{r}_1) = \phi_2(\boldsymbol{r}_2)$; hence the Pauli exclusion principle: no two electrons can simultaneously occupy the same state (spin-orbital) within a quantum system, i.e. they cannot have the same quantum numbers. Specifically, there are four quantum numbers which describe the discrete characteristics of an electron cloud: principle (n), angular momentum (l), magnetic (m_l) and spin (m_s). The energy level at which an electron is situated is described by n. The magnitude of orbital angular momentum corresponds to l. These first two quantum numbers determine the size of an electron shell and the shape of the electron subshell, respectively. The energy sublevel splitting is described by m_l. Trivially, m_s represents the electronic spin. The last two quantum numbers project the angular momentum and the spin angular momentum along a specific axis respectively.

For a pair of these spin-1/2 fermions to exist in the same atomic or molecular orbital (same spatial coordinates), they must have opposite spin states: *up* and *down* with $m_s = 1/2$, and $m_s = -1/2$ respectively. More generally, the Pauli exclusion principle applies to any identical particles with half-integer spin.

The Singlet and Triplet States

Equations 2.10a and 2.10b lead to the following four cases represented by Figure 2.1. Two coupled fermions have total spins $S_t = 0$ and $S_t = 1$, in an antisymmetric (*singlet*) and symmetric configuration (*triplet*) respectively.
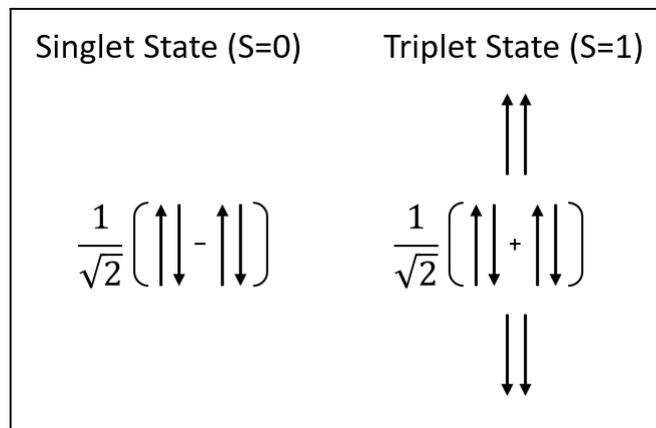
Figure 2.1: Illustration to show four combinations of spin pair states. Only one of the two-particle wavefunctions is anti-symmetric, the other three are symmetric. These are referred to as the singlet and triplet states, respectively. The up and down arrows signify the corresponding spin state. A normalisation factor must be taken into account for the mixed states.

All electron spins are paired anti-parallel in a singlet ground-state and remain this way after excitation. Whereas they are no longer paired for the triplet state, so the excited electron spin is oriented parallel to that of the ground-state.

A transition from a singlet to triplet state (vice versa) is "forbidden", because a post-absorption spin-flip without spin-orbit (SO) coupling violates the law of conservation of angular momentum. It is worth re-emphasising that the triplet state is not applicable to this project and a brief discussion on its origin has been included purely for completeness.

The probability for an electron to transition between bands depends on the shape of the initial and final eigenstate wavefunctions, and the interaction strength between photons and a specified eigenstate. Oscillator strengths are used to describe transition probabilities, whereas selection rules are used to predict if a transition is allowed.

Fermi's Golden Rule

Fermi's golden rule (oscillator strength) describes the probability of a transition to occur in a quantum system, per unit time, from an initial energy eigenstate to a continuous distribution of a final energy eigenstates. Thus it is applicable to electronic absorption and recombination [93]. The rule may be applied to discrete

final eigenstates even if there is a loss of coherence during the transition process, e.g. through carrier relaxation or perturbative noise. The general form of the rule is related to the inverse of mean lifetime and derived for semiconductors in Ref. [94]:

$$
\begin{aligned}
\Gamma_{\mathrm{i}\to\mathrm{f}} &= \frac{2\pi}{\hbar}\left|\langle\psi_{\mathrm{f}}|\hat{\mathcal{H}}'|\psi_{\mathrm{i}}\rangle\right|^2 \delta(E_{\mathrm{f}} - E_{\mathrm{i}} \pm \hbar\omega_{\mathrm{a}}), \\
&= \frac{2m_0}{\hbar^2}|M_{\mathrm{opt}}|^2(E_{\mathrm{f}} - E_{\mathrm{i}}).
\end{aligned}
\tag{2.11}
$$

Here, $\Gamma_{\mathrm{i}\to\mathrm{f}}$ is the transition probability per unit t, and the subscript labels "i" and "f" refer to the initial and final states, respectively. The transition rate is essentially constant regardless of whether the perturbing Hamiltonian describing the photon interaction, $\hat{\mathcal{H}}'$, is t-dependent or not [95]. Trivially, it is assumed that $\hat{\mathcal{H}}' \neq 0$, since electric dipole transitions are allowed. The angular frequency, ω_{a}, can be found from the dispersion relation of light: $\omega_{\mathrm{a}} = 2\pi\nu = c|\boldsymbol{k}|/n_{\mathrm{r}}$, where c is the speed of light, $\boldsymbol{k}$ is the wave-vector, n_{r} is the refractive index, and m_0 is the electron rest mass. All other symbols have their usual meanings.

Specifically, the matrix element, M_{opt}, for optical transitions is expressed as:

$$
M_{\mathrm{opt}} = \langle\psi_{\mathrm{f}}|x_{\mathrm{pos}}|\psi_{\mathrm{i}}\rangle = \int_{-\infty}^{\infty}\psi_f^*(x)x_{\mathrm{pos}}\psi_i(x)\mathrm{d}^3 x_{\mathrm{pos}} = M_\mu\Delta\psi,
\tag{2.12}
$$

Here, $|\psi\rangle$ is the envelope function of a given state. Importantly, M_{opt} can be separated into the product of two components; i.e. the intersubband dipole moment, M_μ, and the electron-hole wavefunction overlap, $\Delta\psi$, thus $M_{\mathrm{opt}} \propto \Delta\psi$ [96].

$$
M_\mu = \langle\psi_{\mathrm{e}_n}|\hat{\mu}|\psi_{\mathrm{h}_n}\rangle = \frac{1}{4\pi^3}\int_{\mathrm{BZ}}\psi_{\mathrm{e}_n}^*(\boldsymbol{k})(-e\boldsymbol{r}_{\mathrm{d}}\cdot\boldsymbol{E})\psi_{\mathrm{h}_n}(\boldsymbol{k})\mathrm{d}\boldsymbol{k},
\tag{2.13a}
$$

$$
\Delta\psi = \langle\psi_{\mathrm{e}_n}|\psi_{\mathrm{h}_n}\rangle = \int_{-\infty}^{\infty}\psi_{\mathrm{e}_n}^*(x_{\mathrm{pos}})\psi_{\mathrm{h}_n}(x_{\mathrm{pos}})\mathrm{d}x_{\mathrm{pos}}.
\tag{2.13b}
$$

Here, $\hat{\mu}$ is the dipole moment operator, e is the elementary charge constant, $\boldsymbol{r}_{\mathrm{d}}$ is the displacement vector, and $\boldsymbol{E}$ is the electric field. The subscript labels e_n and h_n denote the principle quantum number of the electron and hole respectively. All other symbols have their usual meanings.

2.1.2 Band Theory of Solids

The Band Theory of Solids attempts to predict carrier motion inside a material by describing the state of an electron interacting with charged atoms in the material lattice. It is used to distinguish between the different types of solid in terms of their

conductivities and can explain the electronic mean-free path in crystals.
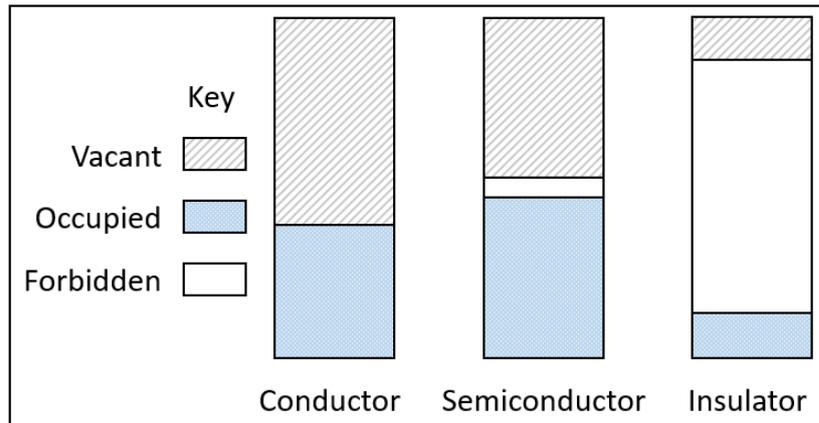
Figure 2.2: Schematic band diagrams to show the band edge positions structure for a metal, semiconductor and insulator. The CB and valence band (VB) are represented by the grey and blue bands respectively. The dotted line represents the Fermi level (see Subsection 2.1.3). A conductor has no bandgap, whereas an insulator has a bandgap that is too wide for electronic excitations to occur.

An insulator (e.g. rubber) has a very large bandgap, whereas the CB and VB overlaps for a metal (e.g. Cu). Semiconductors are an intermediary case. The relatively small gap between the bottom of their CBs and the top of their VBs allows for their conductivity to be influenced; hence they are commonly used in a wide range of applications that require a controlled flow of current, e.g. through an IC. These materials are distinct to semimetals in which the bands only just touch. Semiconductors are further distinguished by the level of doping within them. Intrinsic semiconductors are pure (e.g. Ge), whereas extrinsic semiconductors contain impurities. Extrinsic semiconductors are either n- or p- type; referring to whether the impurities they contain are pentavalent or trivalent, i.e. if they donate an electron or hole to the lattice, respectively.

Many fundamental concepts related to the electronic structure of semiconductor crystals have been obtained through the analysis of 1D atomic chains [97]. The Kronig-Penney model is renowned and will suffice to introduce the band theory required for this thesis i.e.: how solutions to Equation 2.2 yield allowed and forbidden energies in a uniform periodic potential.

Kronig-Penney Model

An ideal crystal has a periodic lattice, where each lattice site is substituted with an atom or molecule. Hence, a free electron moves through a periodic potential depicted in Figure 2.3 below. The Kronig-Penney model approximates the lattice potential by replacing constituent potentials with an alternating series of QWs and barriers, thus

allowing for the possibility of quantum tunnelling between sites. The weak periodic potential assumption is justified two-fold against the large electrostatic force close to an ion. The Pauli exclusion principle prevents valence electrons from bypassing the inner occupied orbitals. In addition, the core electrons screen the magnitude of nuclear charge experienced by conduction electrons. To further simplify the model, the unit cell of the periodic lattice is considered; comprising half of a barrier (region II) either side of the QW (region I).
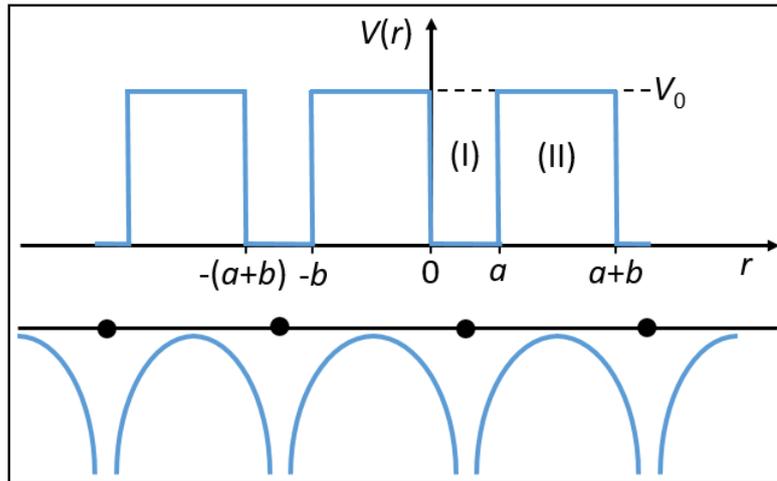
Figure 2.3: Upper: Potential energy as a function of position for the Kronig-Penney model. Notably, the lattice constant, a, is a geometrical parameter which is proportional to inter-atomic distance, hence quantifies the spacing between adjacent unit cells. The QWs (region I) and barriers (region II) are therefore assumed to have widths a and b, respectively. The barrier height, V_0, is finite. Lower: Sketch to show a more realistic potential arising from a lattice, where the black circles represent atoms. For simplicity, the example presented here is limited to a generic 1D solid. Both plots are aligned to reflect the attractive potential experienced by an electron.

The barrier height limits $V(\boldsymbol{r}) = 0$ and $V(\boldsymbol{r}) = \infty$ lead to the nearly-free electron (NFE) and tight-binding (TB) formalisms, respectively. The NFE formalism is very similar to the Drude-Sommerfeld model of electron gasses, by using Bloch theory to relax the assumption that ions do not interact with the valence electrons. Therefore carriers move almost freely throughout the crystal lattice. Similar to the free electron picture, the electron-electron interactions are still assumed negligible (the same is true for the TB formalism). Whereas the TB formalism treats electrons as if they are tightly bound to an atomic orbital at each lattice site. Therefore it takes solutions at each core and observes how they change when interacting with nearest neighbours through a hopping parameter, t_{hop}. Hopping is equally likely in all directions provided the system is isotropic; it occurs with probability t_{hop}^2/E_0^2, where E_0 represents the free electron energy (see Equation 2.18). In a way, the electrons may be thought of as moving through a discrete space.

Both models have different dispersion relations that are identical near the Γ-point,

i.e. the centre of the first BZ (at $|\boldsymbol{k}| = 0$), and they are in agreement on the forbidden states at the BZ edge. However, they differ insofar that the number of primitive unit cells, N_{cell}, is finite in the TB model, thus has a finite number of states; all of which are contained in the reduced zone scheme (all points in $\boldsymbol{k}$-space are separated by $2\pi/aN$). Whereas the NFE model has an infinite number of states, that extend well beyond $|\boldsymbol{k}| > \pi/a$. This leads to the main qualitative difference between the models, namely, that both approximations describe semiconductors from opposite extremes; i.e. the TB picture asserts that solids are effectively insulators, with small perturbations that make them metallic, and vice versa for the NFE picture. Other formalisms of the Kronig-Penney model exist, although they tackle various subtleties beyond the scope of this thesis.

Bloch showed that the solution to the wavefunction in a periodic potential must be of the exponential form according to Equation 2.14. Specifically, the eigenfunctions are products of a periodic function, $u_{n,\boldsymbol{k}}(\boldsymbol{r})$, with a plane wave. Importantly, $\boldsymbol{k}$ inside a periodic potential no longer represents a carrier momentum, but a crystal momentum ($\hbar\boldsymbol{k}$) instead, i.e. Bloch states are not eigenstates of the momentum operator, $\hat{p}$.

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = u_{n,\boldsymbol{k}}(\boldsymbol{r})\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{2.14}$$

The resultant potential field of the entire ionic ensemble $V(\boldsymbol{r})$ inherits the periodic nature of the lattice, hence it can be shown that $V(\boldsymbol{r}) = V(\boldsymbol{r}+\boldsymbol{R})$. By definition, this echoes through the character of the wavefunction. The lattice translation vector, $\boldsymbol{R} = y_1\boldsymbol{a_1} + y_2\boldsymbol{a_2} + y_3\boldsymbol{a_3}$, where $(y_1, y_2, y_3) \in \mathbb{Z}^+$. The primitive lattice vectors (shortest possible vectors) are represented by $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, and $\boldsymbol{a_3}$. It follows that:

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{R}) = \psi_{n,\boldsymbol{k}}(\boldsymbol{r})\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{R}}. \tag{2.15}$$

Equation 2.15 is slightly modified in the TB approximation. A complete wavefunction constructed from a linear combination of valence orbitals, $\phi_n(\boldsymbol{r})$, of all the atoms in a primitive cell (smallest possible unit cell) is known as a Wannier function. This is expressed as:

$$\psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N_{\mathrm{cell}}}} \sum_R \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{R}}\phi_n(\boldsymbol{r}-\boldsymbol{R}). \tag{2.16}$$

The atomic orbitals are eigenfunctions of the monatomic Hamiltonian, and their energy levels are identified by $n \in \mathbb{Z}^+$. The location of the yth atom is specified by $\boldsymbol{R}$ and the eigenfunction corresponding to its orbital is represented by $\phi_n(\boldsymbol{r}-\boldsymbol{R})$. Importantly, the factor of $N_{\mathrm{cell}}^{-1/2}$ ensures that the Bloch state is normalised for the number primitive unit cells. Equation 2.16 satisfies the periodicity requirement

(Equation 2.15).

The origin of bands is easiest to see in the TB picture: Since electrons must occupy different states and fill up the core energy levels of an atom due to the Pauli exclusion principle, every atom has a slightly different energy and is associated with a discrete energy level. When N atoms are brought close enough to each other, the energy levels of the outer-most orbit electrons are affected by neighbouring atoms. This causes an N-fold splitting of energy levels due to exchange energy, i.e. orbitals overlap and form a molecular orbit with N distinct energy levels. A solid typically consists of $\sim 10^{23}$ atoms; such a large number causes a continuum of energy states to emerge and leads to a band of energies, e.g. 1s energy band for the 1s energy level. The characteristic spacing between the energy levels is inversely proportional to the distance between neighbouring atoms.
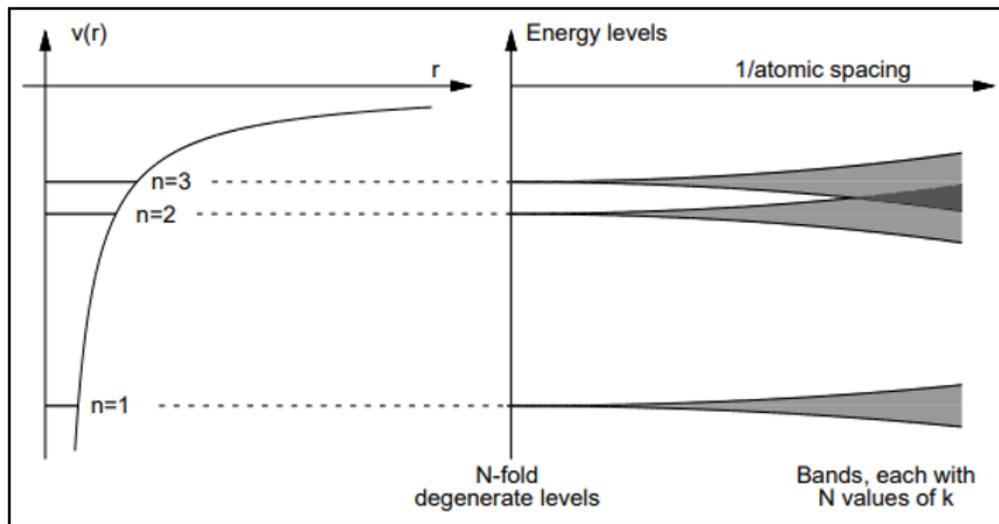
Figure 2.4: Schematic diagram to illustrate the birth of continuous bands in a periodic array of atoms; according to the TB formalism. The non-degenerate energy levels of a single atom are shown (left), where $V(r)$ refers to the atomic potential as a function of radial distance. These energy levels are sketched as a function of the (mean) average interatomic spacing for N atoms (right). Figure adapted with permission from [98].

The origin of the bandgap is easier to see in the NFE picture: Electrons are reflected at Bragg planes (see Section 3.5), i.e. planes in $\boldsymbol{k}$-space that perpendicularly bisect a reciprocal lattice vector (from the origin) [99], as a natural consequence of the periodic lattice. This occurs due to the phase velocity changing direction upon crossing a zone boundary. When the wave-vectors of both wavefunctions are 180 ° out of phase, the group velocity, $\boldsymbol{V_g}$ (Equation 2.17) of these opposing waves is approximately zero, hence they are described as standing waves.

$$\boldsymbol{V_g} = \frac{\mathrm{d}\omega_a}{\mathrm{d}\boldsymbol{k}} = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}\boldsymbol{k}}. \tag{2.17}$$
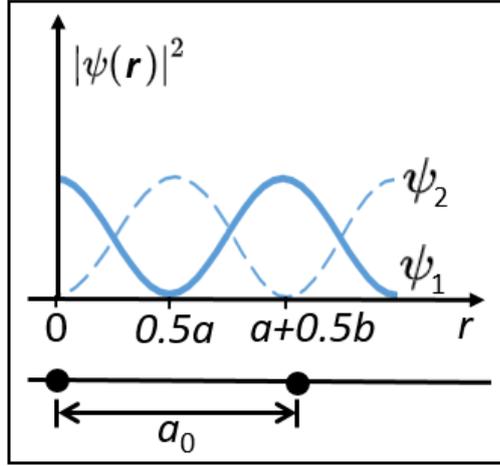
Figure 2.5: Sketch to show the probability amplitude of two opposing electronic wavefunctions in a 1D potential. The atoms are represented by black circles. Note the sinusoidal curves are not to scale.

As shown in Figure 2.5, the electronic wavefunction ψ_1 peaks at the lattice sites, whereas the second, ψ_2, peaks in the middle of the QW. There is a jump in energy between wavefunctions following the requirement that it, and its derivative, must be continuous everywhere. Thus, leading to energies which an electron in the solid cannot possess, hence the term forbidden states. These gaps occur when $\boldsymbol{k} \cdot a = \pm n\pi$ and cause the continuum of energy levels to separate into bands (see Figure 2.6).

In comparison, the TB formalism leads to an electronic wavefunction similar to an atomic orbital of the atom, as if it were monatomic. Due to limited interaction with neighbouring atoms, its energy will be close to its ionization energy. Eigenvalues in the (isotropic) NFE dispersion relation are given in terms of $\boldsymbol{k}$:

$$E(\boldsymbol{k}) \; = \; \Delta E_{\mathrm{c}} \; + \; E_0(\boldsymbol{k}) \; = \; \Delta E_{\mathrm{c}} \; + \; \frac{\hbar^2 |\boldsymbol{k}|^2}{2m_{\mathrm{e}}^*}. \tag{2.18}$$

Where $E(\boldsymbol{k})$ represents the energies of an electron with effective mass, m_{e}^*, at $|\boldsymbol{k}| = n\pi/a$, and ΔE_{c} is the CB offset. Notably, this is the same form as for the free electron. The only difference here is that the electron mass is replaced by an effective mass due to a weak interaction with the lattice. Generally, the energy of a particle may be interpreted as entirely kinetic provided that $E(\boldsymbol{k})$ conforms to this parabolicity. As shown in Figure 2.6, an electric bandgap opens up where the Drude-Sommerfeld (free-electron model) used to cross. This is not to be confused with the optical bandgap later discussed in Subsection 2.3.2.

There are no confined states within the bandgap of a perfect, infinite, and periodic
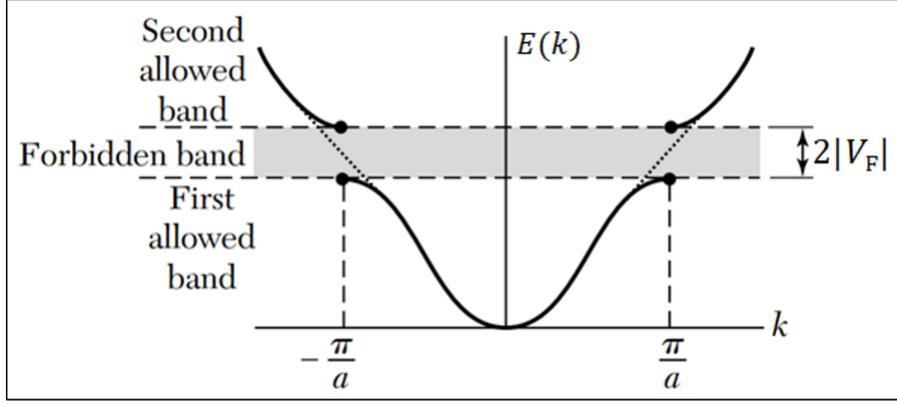
Figure 2.6: Plot of the modified electron dispersion solution in 1D contrasting both the free electron model and the NFE formulation of the Kronig-Penney model for a monotomic linear lattice. The dotted line represents the free electron dispersion. The magnitude of the energy gap is equal to twice the Fourier component of the crystal potential, $|V_F|$, since the normalised wavefunctions at the BZ boundary are $\sqrt{2}\cos(\pi x_{\mathrm{pos}}/a)$ and $\sqrt{2}\sin(\pi x_{\mathrm{pos}}/a)$. Figure adapted from [97].

lattice. Defects interfere with the local potential and create extra states that trap carriers. These states are different in energy to the band states and typically fall within the bandgap.

The range of k-values corresponding to the allowed states are BZs. The number of allowed k-values in a BZ is equal to the number of unit cells in the crystal. The first BZ is the primitive cell for the reciprocal lattice, i.e. the analogue of the Wigner-Seitz cell in real space. For a cubic lattice, the first BZ exists between $-\pi/a < k < \pi/a$.

The real band structure plot of bulk GaAs is presented in Figure 2.7 to exemplify the relative band positions and their general curvature. The curvature of each band is extremely important and influences how the (quasi-)particles behave (see Subsection section 2.1.2). An infinitely large lattice of homogeneous material is assumed in Figure 2.7. For confined heterostuctures however, the band structure is no longer continuous; data is plotted in real space, e.g. along the growth direction. Band diagrams of the gain media investigated in this work are illustrated in Figure 5.11.

The VB and CB appear closest to the Fermi level (mid-bandgap). The VB refers to the band of valence electron energies, where electrons are loosely bound to the atomic nucleus. The CB is the lowest range of vacant electronic states, hence the energy levels pertaining to free electrons. The CB sits at a higher range of energies above the VB. For semiconductors, both bands are separated by a small gap (between Γ_{6c} and Γ_{8v}), therefore it requires an external input of energy to push valence electrons beyond it into the CB; thus governing the electrical conductivity of a solid. The carriers at the band edges are the ones which participate in generation
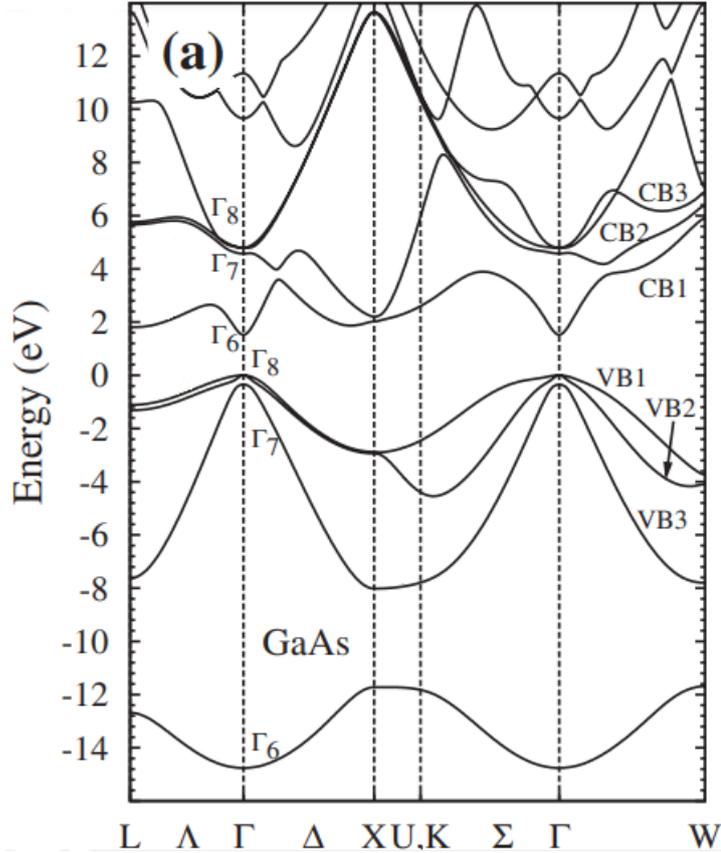
Figure 2.7: Plot to show the band structure of bulk GaAs. The highest three VBs are labelled VB1, VB2, and VB3; these are the heavy hole (HH) (Γ_{8v}), light hole (LH) (Γ_{8v}), and SO bands (Γ_{7v}), respectively. The lowest three CBs are labelled CB1 (Γ_{6c}), CB2 (Γ_{7c}), and CB3 (Γ_{8c}). The highest three VBs, CB2, and CB3, have p-like character at Γ, whereas Γ_{6c} is s-like. The symbols marked along the x-axis are the points of high symmetry shown in Figure 2.8. Figure reproduced with permission from [100].

and recombination processes (see Subsection 2.3.3). Consequently, their properties in these states are of interest to the semiconductor community, so only the dispersion at the band extrema is typically considered.

The 3-D Brillouin Zone

Discussion so far has focused on the simple cubic lattice, however, many semiconductors have a diamond or zinc-blende (ZB) structure, owing to a face-centred cubic lattice with either a one- or two- atom basis, respectively. Here, the face-centred cubic structure is used to describe the first BZ in 3D (Figure 2.8). The dimensions are relatively complicated in this case. For instance, the distances from its centre to the L- and X- planes are $\sqrt{3}\pi/a$ and $2\pi/a$, respectively.

Energy bands in 3D are harder to visualise. To solve for a real solid, the 1D $u_{n,\boldsymbol{k}}(\boldsymbol{r})$
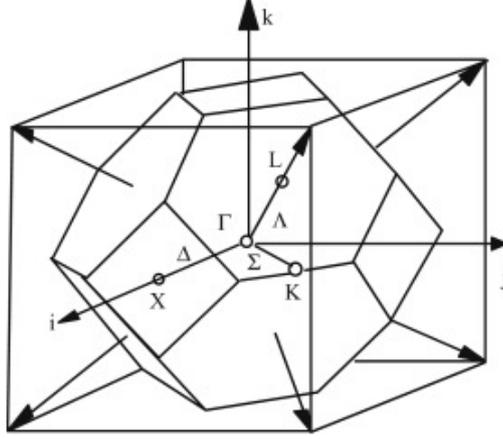
Figure 2.8: Schematic diagram to show the first BZ in 3D for ZB and diamond structure crystals. The labels refer to points of high symmetry. The dashed enclosure is a guide for the eye and extends between $-2\pi/a < \boldsymbol{k} < 2\pi/a$, so it is flush with the X-planes.

must be expanded in a Fourier series to obtain its 3D generalisation:

$$u_{n,\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} f_{\boldsymbol{G}}^n(\boldsymbol{k})\mathrm{e}^{i\boldsymbol{G}\cdot\boldsymbol{r}}. \tag{2.19}$$

Where $f_{\boldsymbol{G}}^n$ represents the Fourier coefficients. The reciprocal lattice vector, $\boldsymbol{G}$, is defined by:

$$\mathrm{e}^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1. \tag{2.20}$$

Where $\boldsymbol{G} = y_1'\boldsymbol{b_1} + y_2'\boldsymbol{b_2} + y_3'\boldsymbol{b_3}$, and $y_1', y_2', y_3' \in \mathbb{Z}^+$. The three primitive reciprocal lattice vectors $\boldsymbol{b_1}$, $\boldsymbol{b_2}$, and $\boldsymbol{b_3}$, are expressed in terms of the primitive real lattice vectors in Equations 2.21a, 2.21b, and 2.21c. Importantly, $\boldsymbol{b_1}$, $\boldsymbol{b_2}$, and $\boldsymbol{b_3}$, are normalised by the volume of a primitive unit cell, $V_c = |\boldsymbol{a_1} \cdot (\boldsymbol{a_2} \times \boldsymbol{a_3})|$.

$$\boldsymbol{b_1} = \frac{2\pi}{V_c}\boldsymbol{a_2} \times \boldsymbol{a_3}, \tag{2.21a}$$

$$\boldsymbol{b_2} = \frac{2\pi}{V_c}\boldsymbol{a_3} \times \boldsymbol{a_1}, \tag{2.21b}$$

$$\boldsymbol{b_3} = \frac{2\pi}{V_c}\boldsymbol{a_1} \times \boldsymbol{a_2}. \tag{2.21c}$$

Critical paths between points of high symmetry in reciprocal space need to be inspected in a similar way to the 1D case. These contours are then used to create surfaces of constant energy for specific bands (see Figure 2.9).

Effective Mass Approximation

Electrons are unable to move as freely in a semiconductor as opposed to in a vacuum, due to all the atomic interactions. Hence, the electron is treated with an effective mass. This is the apparent mass the electron has in response to all the forces acting upon it. Hence the band structure retains some simplicity by approximating the behaviour of carriers as identical to that of particles in the free electron gas model, only with different masses.

$$(m^*)^{-1} = \frac{1}{\hbar^2} \frac{\delta^2 E}{\delta \boldsymbol{k}^2} \bigg|_{|\boldsymbol{k}|=k_0}. \tag{2.22}$$

Notably, the inverse of $(m^*)^{-1}$ is no longer a constant like m_0. It depends on the curvature of the bands at their extrema. This can be found by assuming $E = \hat{p}^2/2m^*$, then expanding $E(\boldsymbol{k})$ into a Taylor series about $|\boldsymbol{k}| = k_0$, and finding the second derivative with respect to $\boldsymbol{k}$ at that point. Here, k_0 refers to the wave-vector offset that characterises the position of the CB minima.

A convex (concave) parabola corresponds to a positive electron (negative hole) effective mass. By convention, values for the hole effective masses are positive. Despite this, it is possible for an effective mass to be negative even with convex bands. This causes electrons to gain velocity in the opposite direction in response to electric and magnetic forces. Despite being negatively charged, they move as if they have positive charge and mass. Furthermore, weakly (strongly) curved parabola give large (small) m^*-values, which leads to distinct hole character that authors refer to as either *heavy* or *light*.

An external force, $\boldsymbol{F}$, on an electron inside a crystal is:

$$\boldsymbol{F} = \hbar \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t}. \tag{2.23}$$

Therefore carriers experience different magnitudes of force in different directions, thus accelerate to different degrees depending on the lines of action. As a result, constant-energy surfaces change; they are no longer spherical, but ellipsoidal, as shown in a) and b) of Figure 2.9. This produces anisotropy in $E(\boldsymbol{k})$.

Most semiconductors have an anisotropic $E(\boldsymbol{k})$. For that reason, carriers are no longer comparable to the free electron model and cannot be described accurately with a parabola (albeit Equation 2.22 does remain a decent approximation very near
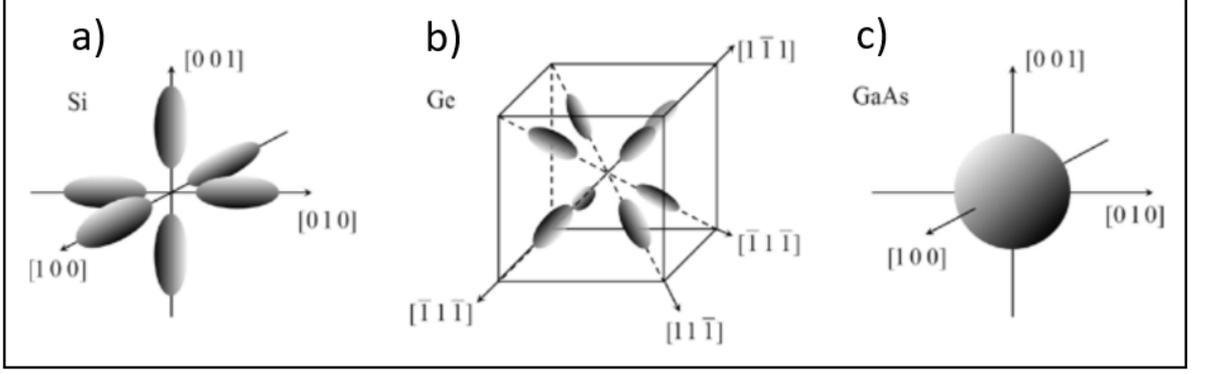
Figure 2.9: Sketches of the CB constant-energy surfaces for a) Si, b) Ge, and c) GaAs. The degeneracies of the Γ-, X-, and L- valleys are 1, 6, and 8 respectively. Notably, the X- and L-valleys are shared with the adjacent first BZ, thus only half the degeneracy is considered.

the Γ-point). The CB minimum in this case can be approximated by:

$$E(\boldsymbol{k}) = E_0 + \frac{\hbar^2}{2m_{\mathrm{x}}^*}(k_{\mathrm{x}} - k_{0,\mathrm{x}})^2 + \frac{\hbar^2}{2m_{\mathrm{y}}^*}(k_{\mathrm{y}} - k_{0,\mathrm{y}})^2 + \frac{\hbar^2}{2m_{\mathrm{z}}^*}(k_{\mathrm{z}} - k_{0,\mathrm{z}})^2. \qquad (2.24)$$

Where the subscript labels x, y, and z refer to the Cartesian axes aligned to the principle axes of the ellipsoids. Non-zero k_0-values imply that the CB minima is no longer positioned at the centre of the BZ.

Generally, the reciprocal effective mass is represented a tensor, $\boldsymbol{M}_{ij}^{-1}$, whose elements are typically functions of $\boldsymbol{k}$ and correspond to the effective masses of the anisotropic case above. The parabolic approximation is recovered when all of its diagonal elements are constant and all off-diagonal elements are zero. For the special case of isotropic bands, all diagonal, and all off-diagonal elements must be equal.

$$\boldsymbol{M}_{ij}^{-1} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial \boldsymbol{k}_i \partial \boldsymbol{k}_j}. \qquad (2.25)$$

Temperature Dependence of the Band Structure

Semiconductor materials exhibit a strong T-dependence; i.e. they behave like insulators at 0 K and their resistance decreases as they are heated. The bandgap is usually small enough to be bridged with an input of thermal energy from a phonon interaction. In addition, the crystal lattice expands with T, i.e. heating the electrons causes their vibrations about the ions to intensify. The magnitude of electron displacement is proportional to the size of the unit cell (considering the Heisenberg uncertainty principle). It therefore requires less energy to excite an electron into the CB, so the bandgap shrinks. The empirical Varshni relation [101] is used to model

the change of bandgap energy, E_g, with T:

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{\beta + T}. \tag{2.26}$$

Where $E_g(0)$ is the bandgap at absolute zero. The coefficients α and β are fitting parameters related to the thermal expansion of the lattice, and the Debye temperature respectively.

2.1.3   Carriers at Equilibrium

Density of States

The DoS describes the number of available states at each energy level. It is required to determine the distribution of carriers in a solid and to calculate absorption and emission rates.
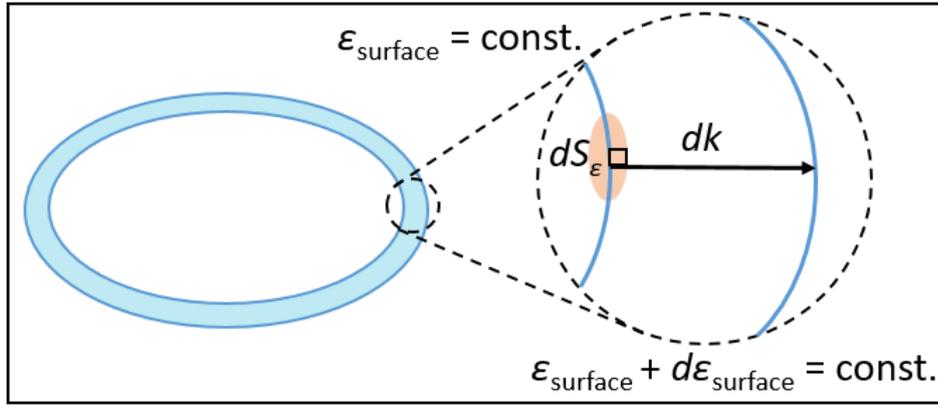
Figure 2.10: Schematic diagram to illustrate a Bloch band energy shell of allowed quantum states, with radius k and infinitesimal thickness δk. The number of allowed states, per unit volume, is contained in the energy range $(\epsilon, \epsilon + \delta\epsilon)$.

The volume of the shell in Figure 2.10, V_{shell}, can be calculated by multiplying the spherical surface area, δS_ϵ, by δk:

$$V_{\text{shell}} = 4\pi k^2 \cdot \delta k \bigg|_{k=0}, \tag{2.27}$$

In reciprocal space, the number of allowed k-values per unit volume ($V = 8\pi^3/L^3$) is:

$$\frac{N_k}{V} = \frac{1}{(2\pi)^3}, \tag{2.28}$$

The number of states in the shell may be then found by multiplying the volume of

the spherical shell with the number of allowed $\boldsymbol{k}$-states per unit volume ($\boldsymbol{k}$-space):

$$N_{\mathrm{k}} = \left. \frac{4\pi k^2}{(2\pi)^3} \delta k \right|_{k=0}. \tag{2.29}$$

An expression for k can be found by rearranging the parabolic dispersion relation (Equation 2.18), whilst noting the CB and VB energies $E_{\mathrm{c,v}} = E_{\mathrm{g}} - E_{\mathrm{i}}$. This is then differentiated with respect to energy and both expressions are substituted into N_{k}. The DoS for bulk materials, $g(E)_{\mathrm{3D}}$, is then obtained once the two electronic spin states are accounted for:

$$g(E)_{\mathrm{3D}} = \frac{\sqrt{E_{\mathrm{g}} - E_i}}{2\pi^2} \left(\frac{2m^*}{\hbar^2} \right)^{3/2}. \tag{2.30}$$

The shells no longer retain their spherical symmetry as the energy increases. At sufficiently high energy, the shell intersects the BZ boundary, hence the shell volume decreases and the number of allowed states with it [102].

The number of available states is also dependent on the system dimensionality. The low-dimensional cases are outlined here, where $g(E)_{\mathrm{1D}}$ is included for completeness. Notably, 2D films technically have a depth since they are atomically thin crystalline solids (without dangling bonds), however their aspect ratio allows for them to be considered negligibly thin. They bind to other epi-layers via the relatively weak Van der Waals forces.

Provided that one considers either an areal element $(k_{\mathrm{x}}, k_{\mathrm{y}})$, or a linear element k_{x}, the above approach may be used to find the DoS in 2D and 1D, respectively; since the free motion of carriers in a 2D structure is confined in one $\boldsymbol{k}$-direction, and in two $\boldsymbol{k}$-directions for the 1D case. Whereas the QD is confined in all three $\boldsymbol{k}$-directions; i.e. all of its available states exist as discrete energies described by $\delta(E - E_i)$, which is similar to the Dirac-delta function. In real QDs however, inhomogeneous morphology leads to a broadening of $\delta(E - E_i)$. The DoS for a QW and a QD are displayed in Figure 2.11 as a function of energy, and are expressed by Equations Equation 2.31a and Equation 2.31b:

$$g(E)_{\mathrm{2D}} = \frac{m^*}{\pi \hbar^2} \sum_i \mathrm{H}(E_{\mathrm{g}} - E_i), \tag{2.31a}$$

$$g(E)_{\mathrm{0D}} = \sum_i 2\delta(E_{\mathrm{g}} - E_i). \tag{2.31b}$$

Where $\mathrm{H}(E - E_i)$ represents the Heaviside step function and the summation appears

by taking all energy levels into account.
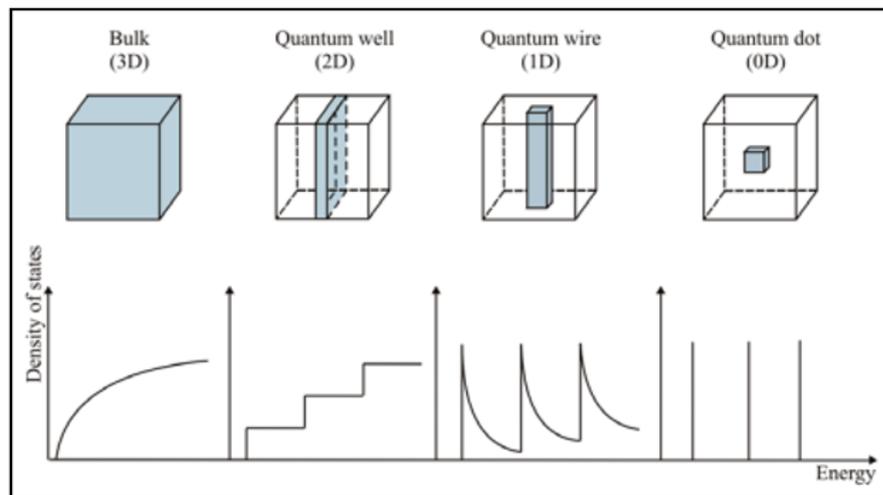
Figure 2.11: Schematic to show the energy dependence of the DoS for the 3D, 2D, and 0D cases. Each plot is accompanied by a sketch to illustrate what the corresponding structure looks like in real space.

Fermi-Dirac Distribution

Statistical thermodynamics is used to observe how states are filled as a function of T. In short, states are filled from the bottom-up and T causes disorder in the arrangement of electrons.
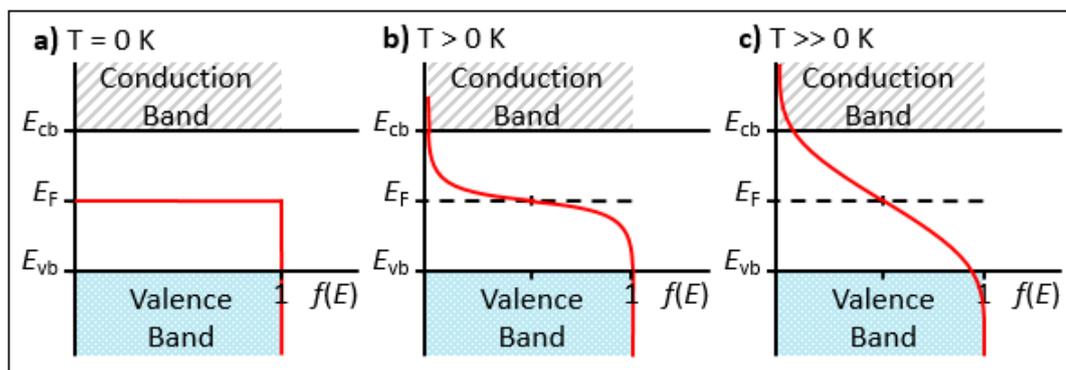
Figure 2.12: Schematic bandgap diagrams overlaid with the inverse of $f(E)$, to show how T changes the probability of a state being occupied in each part of the band structure. As T increases, more electrons have energy greater than the Fermi level (mid-gap).

At 0 K, all electrons settle into the lowest possible energy states and build a "Fermi sea" of electrons. The Fermi level is analogous to the sea surface, i.e. it separates the occupied states from the vacant. The Fermi surface is the 3D extension of the Fermi level. The Fermi energy (E_F) is therefore the highest energy that an electron may occupy at absolute zero. By definition, E_F changes with T since thermal energy affects the level occupation of carriers. The Fermi level position also changes with

the addition of impurities. Thus doping provides an extra control degree of freedom to manipulate the opto-thermal properties of a material.

Indeed, E_{F} is a key parameter of the Fermi-Dirac distribution, $f(E)$ [K], which determines the probability of electron occupancy at different energy levels. In order for electrical conductivity to occur, $f(E) > 0$ inside the CB. This means that $f(E)$ is also positive inside the bandgap, although it is worth reminding that $g(E) = 0$ for this region. Increasing T smears $f(E)$, as shown in Figure 2.12. At equilibrium, $f(E)$ is:

$$f(E) = \frac{1}{1 + \mathrm{e}^{(E-E_{\mathrm{F}})/k_{\mathrm{B}}T}}. \tag{2.32}$$
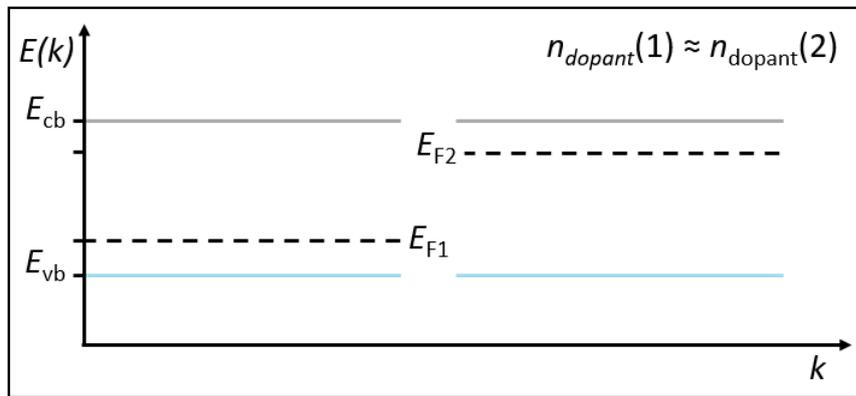
Figure 2.13: Schematic energy band diagram to show where the Fermi energy is located at 4 K in a layer of p-type (left) and n-type (right) semiconductor. The bandgap is small enough that low concentrations of dopants can dramatically change conductivity. Here, the amount of doping is assumed similar in material 1 and 2.

Figure 2.13 illustrates how incorporating n-type dopants into an intrinsic semiconductor alloy causes the Fermi level to gain energy and rise due to state-filling; the closer that the Fermi level is to the CB, the more that the $f(E)$ tail penetrates the band and the easier it is to excite electrons across the bandgap. The opposite is true for p-type dopants due to carrier recombination.

Carrier Density

The carrier density at a particular energy level is calculated by the product of the DoS with the corresponding occupation probability. For the electron density within the entire CB, this is:

$$n(E) = \int_{E_{\mathrm{c}}}^{\infty} g(E)f(E)dE. \tag{2.33}$$

For the hole density across the entire VB:

$$p(E) = \int_{-\infty}^{E_{\mathrm{v}}} g(E)[1 - f(E)]dE. \tag{2.34}$$

The probability of finding an electron at $E = E_{\mathrm{F}}$ is 50% for an intrinsic semiconductor, hence the Fermi level resides mid-bandgap as shown in Figure 2.12. The bands are close enough to the Fermi level for a small number of electrons to thermally populate the CB. The resulting carrier density for both carrier types is symmetrically distributed about E_{F}. Conversely, extrinsic semiconductors have an asymmetrical carrier distribution, whereby the concentration of electrons (holes) in the CB (VB) increases with the insertion of n-type (p-type) atoms.

Free Carriers

The number of free carriers, in either of the bands, can be increased through doping or thermal excitation. These carriers are able to move throughout the crystal lattice. The intrinsic carrier concentration, n_{i}, is a constant which refers to the number of free carriers without doping. It depends on the material bandgap (as well as T), and is defined by the Law of Mass Action [82]:

$$n_0 p_0 = n_{\mathrm{i}}^2. \tag{2.35}$$

Here n_0 and p_0 represent the concentration of majority and minority (quasi)particles in the semiconductor at equilibrium; depending on which of the electron and hole populations are most abundant. It follows from this relation that the doping level is inversely proportional to the number of minority carriers.

The equilibrium carrier concentration n_{eq}, is defined as the total number of carriers in the CB and VB at constant T, when no external bias or stimulus (e.g. light) is applied to a semiconductor. For majority carriers, this is equal to:

$$n_{\mathrm{eq}} = n_{\mathrm{i}} + n_{\mathrm{x}}. \tag{2.36}$$

Where n_{x} denotes the concentration of excess carriers due to doping.

Photo-absorption (see Subsection 2.3.2) is one of multiple processes that drives a state of non-equilibrium inside a semiconductor. The number of excess carriers increases, changing the total carrier concentrations according to Equations 2.37a and 2.37b [82]:

$$n_{\mathrm{e}} = n_0 + \Delta n, \tag{2.37a}$$

31

$$n_{\text{h}} = p_0 + \Delta p. \tag{2.37b}$$

Where Δn and Δp are the additionally-generated electrons and holes respectively.

Optically-Induced Charged Depletion

Carbon atoms occupy substitutional sites of As and act as shallow acceptors in As-based III-V materials. Significant unintentional doping of electrically-active carbon during growth causes samples to exhibit optically-induced charge depletion (OICD) [22] at low T and P. For C-doped samples, it is energetically-favourable for acceptor holes to exist in the wetting layer (WL) and QDs/QRs, when the laser source is switched off. Although the hole occupancy is relatively high inside the QRs and WL at low P, the recombination rate is low and limited by the small number of injected electrons. Increasing P increases the number of electrons in the CB which then recombine quickly and reduce the hole occupancy. Additionally, the acceptor states trap holes, thus impeding recombination with GaSb hole states as they are filled. With enough P, the relative photo-generated carrier and acceptor densities become similar. Further increasing P saturates the acceptor states. At this point, the GaSb hole density is minimal. When the intensity of the laser source, I_{ls}, is increased even further, the excess photo-generated holes start occupying the WL and QRs due to band-filling and intra-band recombination processes. Increasing P beyond this point results in blueshifted emission energy from the intended recombination channel, i.e., higher energy levels in the CB are accessed (high inversion population). Since the higher order electron states are filled, high energy electrons can recombine with holes at even lower energies in the VB; thus skewing the resultant emission. The photo-generated carrier density is much greater than the acceptor density at high P, hence acceptor states are no longer discernible in the PL spectra [22].

2.1.4 Band Alignment

Most electronic devices require careful implementation of at least two different types of semiconductor material (even if the difference is due to doping). Resulting heterostructures can therefore be (mis)aligned in either real- or momentum- space; leading to further classifications of semiconductor materials. Here discussion will first focus on band alignment in $\boldsymbol{k}$-space, then on epitaxial layers in real- space.

Direct/Indirect Bandgaps

The minimal- (maximal-) energy state in the CB (VB) is characterised by a crystal momentum in the BZ. Thus a material distinction is made depending on whether there is a change in kinetic energy, E_{kin}, during the band transition.
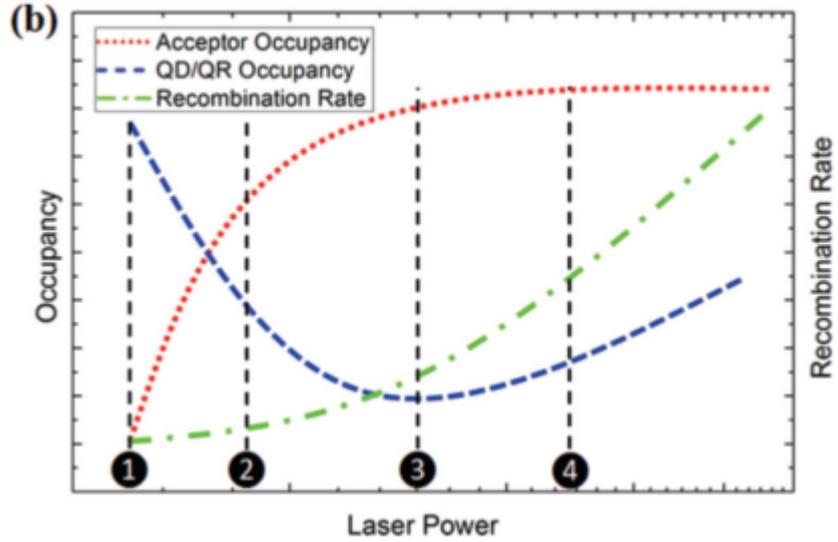
Figure 2.14: Figure to show the P-dependence of the acceptor occupancy, QD/QR occupancy and recombination rate. Figure from [22].
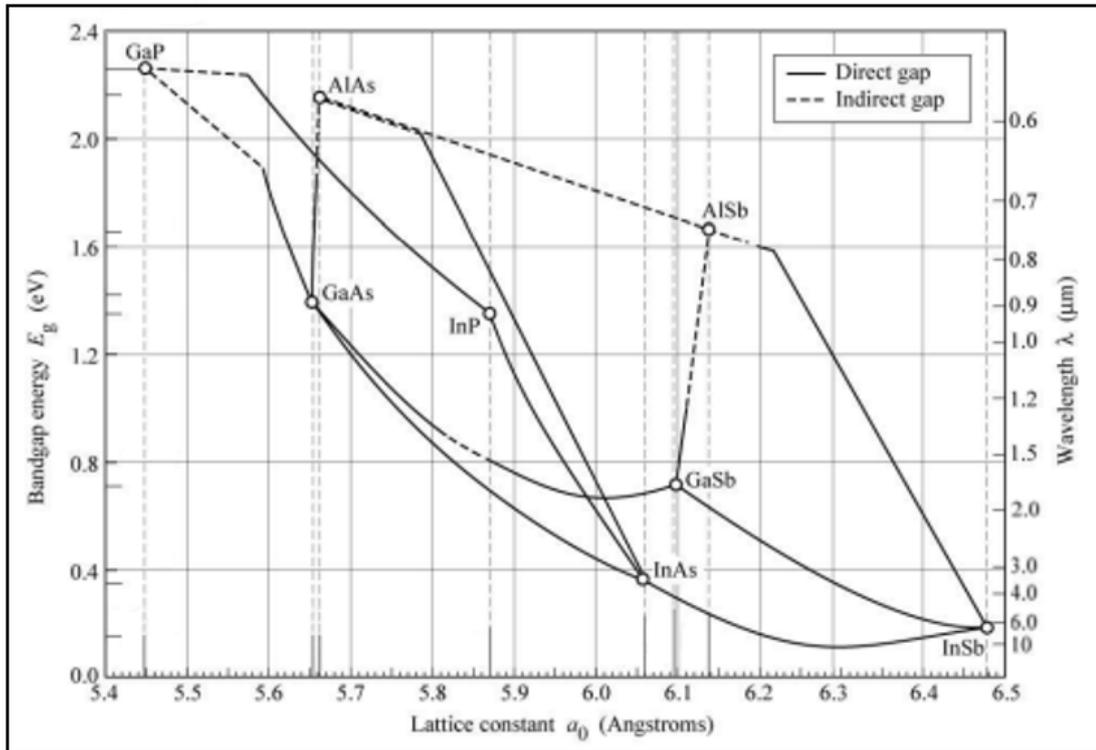
Figure 2.15: Bandgap energies are plotted as a function of bulk lattice constant, a_0, for common III-V binaries (see Subsection 2.2.2) at RT. The corresponding λ-values that are typically achieved with combinations of these materials are marked on the right axis. Figure adapted from [82].

The bandgap is direct if the crystal momentum, $\boldsymbol{k}$, of the electrons and the holes are the same in both bands; i.e. there is no change in momentum during the transition, hence an electron can directly emit a photon equal to E_g. Direct bandgap semiconductors tend to have an S-type symmetry for their CB minima and display a small amount of p-type behaviour away from the zone edge.
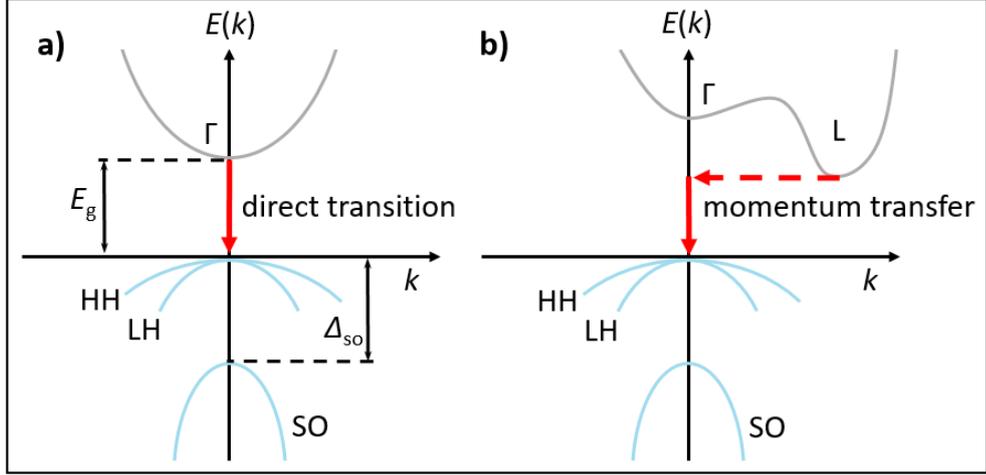
Figure 2.16: Schematic band structure to show how the CBs and VBs align in $\boldsymbol{k}$-space for a) direct, and b) indirect, bulk semiconductors. Here Γ and L refer to the corresponding conduction valleys. The SO (split-off) energy is represented by Δ_{so}. The CB minima and VB maxima are aligned for a direct bandgap, whereas for an indirect bandgap, the lowest energy valley is no longer at the Γ-point, i.e. the extrema exist at different values of $|\boldsymbol{k}|$.

Whereas for an indirect gap, the wave-vectors are different. A photon cannot be emitted because the electron must pass through an intermediate state; transferring momentum to the crystal lattice, via the emission and absorption of phonons, in order to account for the difference in momentum between the initial and final states. Therefore light is usually difficult to generate with indirect materials.

Type-I, Type-II and Broken Bandgap Structures

Electron affinity, χ_i, reflects the ability of an atom to accept an electron and is defined as the energy difference between the CB and vacuum level. This will vary for different materials (in contact) and determines the resultant energy difference between CB edges:

$$\Delta E_{\mathrm{c}} = \chi_1 - \chi_2. \tag{2.38}$$

When placed in contact, semiconductors which differ in electron affinity and Fermi level cause the appearance of energy discontinuities at the abrupt interface between their lattices; i.e. create a difference in (real-space) band alignments, as shown in Figure 2.17. Due to the dependence on growth conditions, the approach to aligning the bandgaps of differing semiconductors is not unanimously agreed upon by the semiconductor community. Conventions differ as whether to align the midgaps, VB edges, or as discussed here, using the electron affinity to align the CB against the vacuum reference level.

The difference in bandgap between both materials may be estimated in terms of the
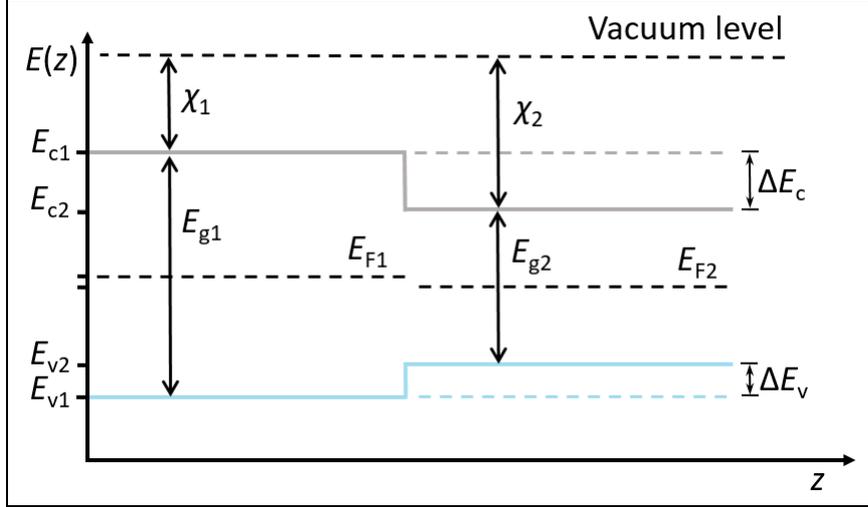
Figure 2.17: A schematic energy band diagram of two intrinsic semiconductor layers in contact. Here, intrinsic materials are considered for simplicity. The electron affinities, Fermi levels, bandgaps, CB and VB edges of each material are labelled as χ_i, $E_{\mathrm{F}i}$, $E_{\mathrm{g}i}$, $E_{\mathrm{c}i}$, and $E_{\mathrm{v}i}$ respectively, where $i = 1, 2$ and refers to either of the two layers. Although both materials here are intrinsic (E_F lies at the midgap), $E_{\mathrm{F}1} \neq E_{\mathrm{F}2}$ due different bandgap magnitudes and band offsets ($\Delta E_{\mathrm{c,v}}$). The vacuum level refers to the free electron energy.

CB and VB band offsets ΔE_c and ΔE_v:

$$\Delta E_\mathrm{g} = E_{\mathrm{g}2} - E_{\mathrm{g}1} = \Delta E_\mathrm{c} + \Delta E_\mathrm{v}. \tag{2.39}$$

Where the sign of each offset depends on the respective band alignments. Moreover, the right-hand side of Equation 2.39 distinguishes whether the net band structure is type- I, II, or III (broken), i.e. a distinction is made between type- I and II based on whether Equation 2.11 is zero, as illustrated in Figure 2.18. Each type exhibits different carrier recombination behaviour and they are discussed below.

In the type-I band alignment, one of the band offsets is positive and the other is negative. Figure 2.18a shows a layer of material with a relatively small bandgap between two layers of a material with a larger bandgap. Both carriers are therefore confined to (and recombine in) the smaller bandgap material, due to the lowering and raising of its CB and VB edges, respectively. InGaAs/GaAs is a classic example of a type-I heterostructure and has a 60:40 (CB:VB) band offset ratio [103].

For type-II band alignments, ΔE_c and ΔE_v are both either positive (type-IIa) or negative (type-IIb). Notably, electrons and holes are spatially separated due to the confinement of carriers on opposite sides of the heterojunction interface. Electrons (holes) are concentrated where the CB (VB) edge is lowered (raised), whereas the recombination of carriers occurs at the interface. Although the transition is spatially-indirect, it may still be either direct or indirect in **k**-space. Well-known examples of
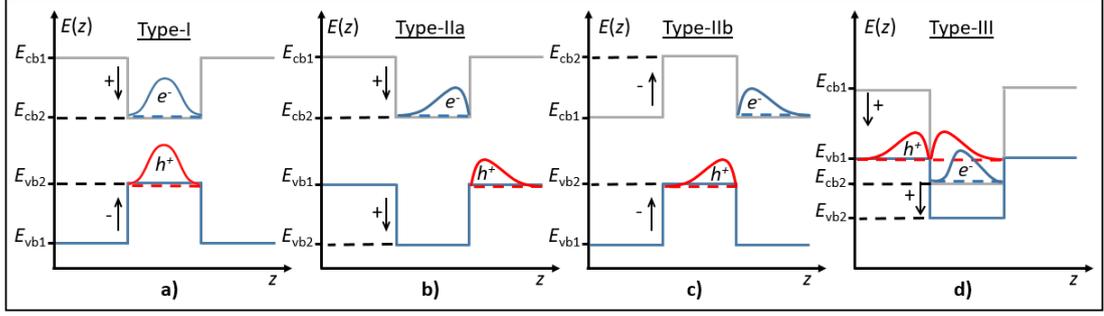
Figure 2.18: Schematic band diagrams to show the different possible types of band alignment for a QW, i.e. a semiconductor sandwich of two different materials. Carrier wavefunctions are labelled e^- for the electrons and h^+ for holes. The arrow direction indicates whether a band edge is raised or lowered, and its sign indicates whether the change in offset is positive or negative. Spatially separating the carriers creates a dipole, thereby producing an electric field and causing the carriers to collect near the heterojunction; for simplicity it is assumed that neither carriers are significantly confined. The symmetry in a) is due to carriers spreading across the entire layer. This is the same for holes in d), however the electron wavefunction is still confined by the VB edge of the adjacent material, i.e. less perturbed by holes in the lower layer. Note the figure is not to scale and band edge effects have been assumed negligible for simplicity.

type-IIa include GaSb/AlSb and InAs/AlSb. These offer a small lattice mismatch, at the cost of mixing direct and indirect bandgaps. Alternatively, the GaSb/(In)GaAs is an example of the type-IIb material system.

Type-III (broken bandgap) structures occur when one of the band offsets is larger than the bandgap of one of the materials, e.g. InAs/GaSb. In this case, the InAs CB edge exists at a lower energy than the GaSb VB edge.

2.2 III-V Epitaxy

This section focuses on heterogeneous parameter mismatch and alloying of semiconductors, the coherent growth of strained thin-films, and the subsequent onset of the relaxed-layer regime necessary to create a wide variety of self-assembled type-II nanostructure active regions. Lastly, crystallographic defects are discussed and how they affect device quality.

Epitaxial growth, i.e. monolithic integration, refers to the formation of an entire epi-structure from a single crystal. It is a popular approach for creating gain media, due to: low cost, dense integration, large wafer compatibility [104], and production simplicity. Successful epitaxy can be achieved either homogeneously or heterogeneously. Homoepitaxy refers to the epitaxial growth of films onto substrates formed of the same material. Direct atomic interactions between the film and substrate produce extremely high quality structures with an almost negligible defect density; at the cost of severely limited material choice and suitable applications, along with difficulty in separating epi-layers from a substrate since the lattice lacks any obvious

discontinuity. In contrast, heteroepitaxy refers to the deposition of a film onto a dissimilar substrate. Therefore, the composition in a heterostructure varies along its growth axis. Material combinations widen the scope for potentially advantageous characteristics and interesting phenomena. Heterogeneous integration includes other fabrication methods such as flip-chip and wafer bonding.

Advancement of epitaxial growth technologies along with understanding of thermo-dynamics and growth kinetics behind thin-film formation (see Subsection 2.2.3) has paved the way for the fine-tuning of a device's electronic properties to meet market demands e.g. low threshold current for high-speed applications [76, 81], Precise control of parameters is crucial for practical gain media designs that yield desirable opto-electronic devices suited to one of many potential applications.

The prior section discussed how layering different semiconductor materials in a heterostructure is extremely useful for creating new devices. While these materials have different electronic properties, they have different physical properties too; the most obvious being a. The rate at which a material expands during heating also tends to change and is quantitatively described by the coefficient of linear thermal expansion (CTE), α_T. These physical properties are tabulated below for several compounds commonly investigated in this field of research.

Material	a (nm)	α_T (10^{-6}/°C)
Si	0.5431[a]	2.59[b]
InP	0.5870[c]	4.56[d]
GaP	0.5451[c]	5.91[d]
GaAs	0.5653[c]	6.40[d]
InAs	0.6058[c]	5.16[d]
GaSb	0.6096[c]	7.17[e]

Table 2.1: RT Lattice constant and CTE data, for Si and several common $A^{III}B^{V}$ binary compounds. It is assumed that the α_T-values are T-independent. See Equation 2.41 for estimating ternary values. Values with exponents a-e are extracted from Refs. [105–109] respectively.

2.2.1 Parameter Mismatch

The monolithic fabrication of high-performance devices is fundamentally difficult. Dissimilar material parameters have an associated elastic strain energy that can drastically alter opto-thermal properties of the material and lead to various defects outlined in Subsection 2.2.4. Consequently, parameter mismatch is a limiting factor to successful growth and must be addressed in order to yield commercially competitive devices. Hence, pseudomorphic growth techniques remain under extensive research.

Layer deposition of a material with lattice constant, a, onto a substrate of lattice constant, a_0, will cause a lattice mismatch, $f(a) = \Delta a / a_0$, where $\Delta a = a - a_0$. Specifically, this is the relative difference between the in-plane lattice constants, $a_\parallel$, at the interface of two materials. The issues related to this type of mismatch are discussed in further detail in Subsection 2.2.3.

The corresponding expression for CTE mismatch is obtained by substituting a for α_T for each respective layer in f. Differences in CTE do not usually pose an issue until it is time to cool back down from growth temperature to RT. During MBE growth, the substrate often reaches 600 °C depending on the choice of materials. The material with a greater CTE value will contract more than the other layer whilst approaching RT; yielding a residual stress of a few percent in layers approximately several μm thick. This residual stress acts as an additional driving force for the generation of dislocations.

Furthermore, since severe CTE mismatch leads to exaggerated differences in rates of thermal expansion/contraction during a temperature cycle. This type of mismatch may cause delamination or crack formation, in order to relieve localised stress and decrease the overall energy of the material system, thereby jeopardising the I_{PL} yield and the number of successfully produced devices. Growth conditions and pre-existent nucleation sites (e.g. growth defects and surface contaminants) strongly affect the density of cracks and how they are distributed throughout the film. The cracking threshold for GaAs/Si has been experimentally observed at $6 - 7$ μm [35], therefore consideration of the critical thickness, h_c (see Subsection 3.1.1) places the target QW thickness well-under CTE-imposed limits.

The use of GaAs substrates looks promising for integration with the Si platform, since GaAs/Si has a reasonably small lattice mismatch of 4.1 % and a CTE ~ 2.5 times larger than that of Si. An excellent review on the heteroepitaxial growth of III-V semiconductors on Si is given in Ref. [104]. Table 2.2 displays the generally-accepted CTE and lattice parameter mismatch values for the common III-V binary materials in Table 2.1 with respect to Si.

Material	$f(a)$	$f(\alpha_T)$
InP	8.08	76.06
GaP	0.37	128.19
GaAs	4.09	147.10
InAs	11.54	99.23
GaSb	12.24	176.83

Table 2.2: Percentage mismatch values [%] for the lattice constant, $f(a)$, and CTE, $f(\alpha_T)$, of several common binary $A^{III}B^V$ compounds, relative to Si.

There is an additional major challenge to overcome in the context of integrating III-V/Si semiconductor technologies. A third type of mismatch known as an antiphase boundary (APB) [9, 110] occurs due to growing a polar binary on a non-polar monatomic lattice; where a polar crystal is defined as a material containing a non-zero dipole moment. APBs exist around regions where the sublattices are misaligned, such as across single atomic steps. Further information is included in Appendix B.

2.2.2 Semiconductor Alloys

Alloying is a powerful way to further broaden the range of semiconductor applications by tailoring bandgap, band offsets and lattice parameters. Elements from two or more different groups in the periodic table are commonly mixed to form compound semiconductors with unique properties, e.g. high electron mobility. These materials can be classified as oxide semiconductors, or in terms of cation-anion reactions between periodic table groups. In particular, A III-V compound semiconductor may contain: boron (B), aluminium (Al), gallium (Ga), or In from group 13; and nitrogen (N), phosphorus (P), arsenic (As), Sb, or bismuth (Bi) from group 15. Ternary compositions are composed of three elements and allow for bandgap adjustment, whereas quaternaries (4 elements) and higher order alloys allow for the simultaneous adjustment of bandgap and lattice constant. Moreover, a subset of materials known as dilute nitrides exist, whereby small amounts of nitrogen are added, to say, the group-V flux used in this work.

Semiconductor alloys can be either intrinsic or extrinsic depending on their level of doping, i.e. intrinsic materials are (relatively) pure. Importantly, the combination of direct and indirect bandgap materials may limit practicality for opto-electronics, due to the range of compositions which result in an overall indirect character.

Vegard realised that the lattice parameter of a binary material may be approximated with a weighted mean of the lattice parameters of the two constituent materials at the same T [111]. This empirical rule (Equation 2.40) is only valid under specific conditions, i.e. ideal systems where the lattice parameters of the bulk binary materials differ by less than 5 % [112].

$$a_{A_{(1-x)}B_x} = (1-x)a_A + xa_B. \tag{2.40}$$

It is widely accepted that a is linear with alloy composition, x. It does not necessarily follow that any other parameter which is approximately a linear function of the lattice parameter may also be linearly interpolated; as is the case for bandgap

energies. For this case a second term, characterised by the bowing parameter, Ω, is included as a quadratic correction to account for strain-induced band-edge curvature as a function of composition. Estimations for ternary alloy parameters may be obtained by:

$$Q(x) = xQ(0) + (1 - x)Q(1) - \Omega x(1 - x). \tag{2.41}$$

Where $Q(x)$, $Q(0)$, and $Q(1)$ are non-specific parameters corresponding to $In_xGa_{1-x}As$, InAs, and GaAs, respectively. The bowing is assumed to be upward (downward) when Ω is negative (positive).

The discrepancy between Equation 2.40 and Equation 2.41 originates from the assumption that an alloy can be ideally treated as a binary compound of known symmetry with zero disorder, thus ignoring inter- and intra- band coupling [113]. Disorder arises in real ternaries from the bimodal distribution of bond lengths, and asymmetric charge distribution. The latter is caused by variation of internal electronic structure between two different atoms sharing sites of mixed sublattice, hence the electronic wavefunction is no longer adequately described by Bloch functions [113]. Subsection 3.1.2 outlines how the 8-band $\boldsymbol{k} \cdot \boldsymbol{p}$ parameters are treated with bowing.

2.2.3 Strained Layers

Assuming growth is restricted to the [001] crystal axis, there are two ways in which the lattice mismatch is naturally accommodated for. The one that is energetically favoured depends on layer thickness and the elastic properties of the materials. The strained layer regime is favourable for thin films, and typically occurs for lattice mismatches above $\sim 2\ \%$ [117], whereas thick layers tend to relax by misfit dislocations (MDs). Strain and quantization effects influence the effective carrier masses, and heavily limit layer thickness, hence the realisable structure, e.g. single, multiple or coupled layers etc.

Pseudomorphic Layers

Coherent atomic bonds at the interface between different materials are formed in the pseudomorphic regime. During this type of growth, the lattice of the epi-layer becomes elastically distorted due to the tendency for adatoms to mimic the substrate lattice below, despite the mismatch of a. Such a perturbation to the lattice creates a lateral and homogeneous strain field throughout the epi-layer, thus driving deformation in the growth- (z-) direction via the Poisson effect [114]. The biaxial strain parallel to the sample surface, $\varepsilon_{\parallel}$, is expressed by Equation 2.42a. Whereas, the uniaxial strain perpendicular to the sample surface, $\varepsilon_{\perp}$, is expressed by Equation 2.42b. Surface undulations, impurities and conglomerations of In compound with

lattice mismatch to cause anisotropy in the strain distribution which propagates and intensifies along the growth axis.

$$\varepsilon_\parallel = \frac{a_0}{a} - 1, \tag{2.42a}$$

$$\varepsilon_\perp = -D\varepsilon_\parallel. \tag{2.42b}$$

The biaxial Poisson ratio, D, depends on the interface orientation and is a combination of the elastic constants, C_{ij}; these are components of the elasticity tensor, $\boldsymbol{E}_{ijkl}$, i.e. a fourth order tensor that relates the second-order stress ($\boldsymbol{\sigma}_{ij}$) and strain ($\boldsymbol{\varepsilon}_{kl}$) tensors by Hooke's law (Equation 2.43):

$$\boldsymbol{\sigma}_{ij} = \sum_{kl} \boldsymbol{E}_{ijkl}\boldsymbol{\varepsilon}_{kl}. \tag{2.43}$$

Conventionally, Voigt notation is used for its simplicity. Where $11 \rightarrow 1, 22 \rightarrow 2, 23 \rightarrow 4, 32 \rightarrow 4, 13 \rightarrow 5, 31 \rightarrow 5, 12 \rightarrow 6$, and $21 \rightarrow 6$. There are only three unique, non-zero elastic constants for cubic ZB crystals, i.e. C_{11}, C_{12}, and C_{44}. In the growth direction of a cubic lattice, $D[001] = 2C_{12}/C_{11}$ [119].

The in-plane, and out-of-plane ($a_\perp$) lattice constants change with strain according to Equations 2.44a and 2.44b, respectively. If a layer has completely relaxed, then it has retained its original lattice constant both parallel and perpendicular to the surface plane. All subsequent unstrained layers then assume the new lattice constant. However, if a layer is completely unrelaxed, then it has the same $a_\parallel$ as the layer beneath and $a_\perp$ changes to accommodate the strain.

$$a_\parallel = a_0, \tag{2.44a}$$

$$a_\perp = a \left[1 - D \left(\frac{a_0}{a} - 1 \right) \right]. \tag{2.44b}$$

As illustrated in Figure 2.19: tensile strain occurs when the epi-layer lattice constant is smaller than that of the substrate ($a < a_0$), hence the epi-layer shortens in the growth direction and elongates in the direction perpendicular to it. The converse is true for compressive strain ($a > a_0$) shown in Figure 2.20.

Strain Dependence of the Band Structure

Strain modifies the effective mass and shifts the band edges. The energy levels change due to two strain contributions: hydrostatic and shear. Hydrostatic strain
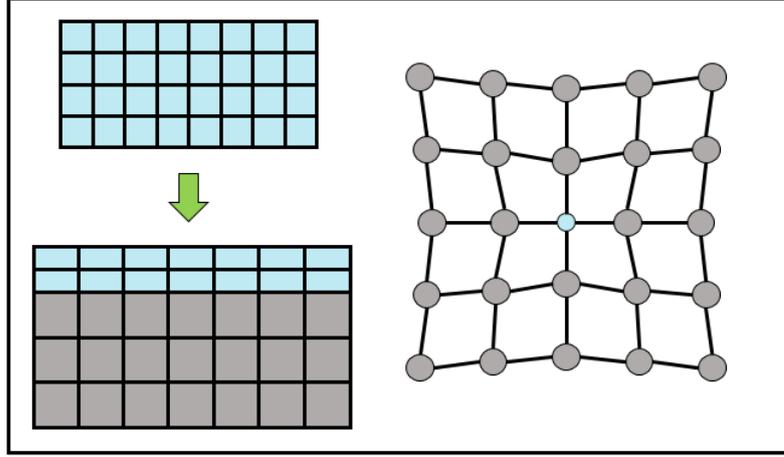
Figure 2.19: Diagram to show how a semiconductor of relatively small lattice constant adapts, under tensile strain, to the substrate of larger lattice constant. To the right is a cross-sectional view of the squashed lattice structure around a small dopant atom.
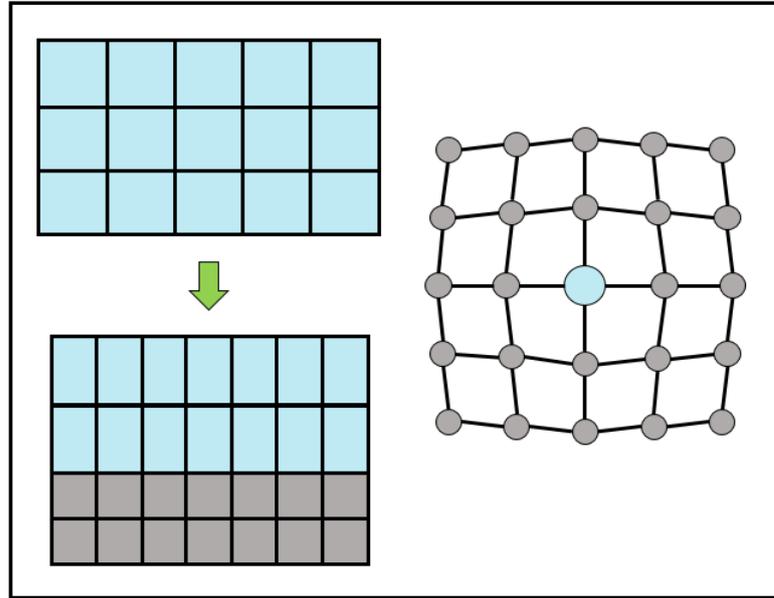
Figure 2.20: Diagram to show how a semiconductor of relatively large lattice constant adapts, under compressive strain, to the substrate of smaller lattice constant. To the right is a cross-sectional view of the inflated lattice structure around a large dopant atom.

shifts the band edges by $\Delta E_{\mathrm{c,v}}^{\mathrm{hy}}$ in the Van de Walle model [119] by:

$$\Delta E_{\mathrm{c,v}}^{\mathrm{hy}} = a_{\mathrm{c,v}} \left(\frac{\Delta V}{V} \right) = a_{\mathrm{c,v}}(2\varepsilon_\parallel + \varepsilon_\perp), \tag{2.45}$$

Where a_{c} and a_{v} are the hydrostatic CB and VB deformation potentials, respectively, and the fractional volume change $\Delta V/V = \mathrm{Tr}(\varepsilon) = \varepsilon_{\mathrm{xx}} + \varepsilon_{\mathrm{yy}} + \varepsilon_{\mathrm{zz}}$. Importantly, $\Delta E_{\mathrm{v}}^{\mathrm{hy}}$ shifts the average energy of the VBs, $E_{\mathrm{v,av}} = E_{\mathrm{hh}} + E_{\mathrm{lh}} + E_{\mathrm{so}})/3)$, where E_{hh}, E_{lh}, and E_{so} refer to the HH, LH, and SO band energies, respectively. Shear strain enters the Nextnano^{++} VB Hamiltonian through the matrix elements l_ε, m_ε, and n_ε

of a strain-dependent interaction operator:

$$l_\varepsilon = a_v + 2b \qquad (2.46a)$$

$$m_\varepsilon = a_v - b \qquad (2.46b)$$

$$n_\varepsilon = \sqrt{3}d \qquad (2.46c)$$

Where, b and d, represent the tetragonal and rhombohedral VB shear deformation potentials, respectively. Only the hydrostatic strain component affects $E_{v,av}$. This is the same for the CB edge in direct semiconductors due to being at the centre of the BZ. The L- and X- bands however, are affected by shear strain. Isotropic dispersion is a consequence of the degeneracy at the top of the VB. Shear strain lifts this degeneracy and splits the HH and LH bands apart in proportion to the magnitude of strain, thus adding to the SO splitting in the absence of strain. The combined shear- and SO- induced energy shift for the HH (ΔE_{hh}^{sh}), LH (ΔE_{lh}^{sh}) and SO (ΔE_{so}^{sh}) bands below are with respect to $E_{v,av}$ [119].

$$\Delta E_{hh}^{sh} = \frac{\Delta_{so}}{3} - \frac{1}{2}\delta E_{sh}, \qquad (2.47a)$$

$$\Delta E_{lh}^{sh} = -\frac{\Delta_{so}}{6} + \frac{\delta E_{sh}}{4} + \frac{1}{2}\sqrt{\Delta_{so}^2 + \Delta_{so}\delta E_{sh} + \frac{9}{4}\delta E_{sh}^2}, \qquad (2.47b)$$

$$\Delta E_{so}^{sh} = -\frac{\Delta_{so}}{6} + \frac{\delta E_{sh}}{4} - \frac{1}{2}\sqrt{\Delta_{so}^2 + \Delta_{so}\delta E_{sh} + \frac{9}{4}\delta E_{sh}^2}. \qquad (2.47c)$$

Where Δ_{so} is the split-off energy, i.e. the energy difference between the SO band and the other two VBs at $|\boldsymbol{k}| = 0$. The shear strain contribution to the energy shift, δE_{sh}, also depends on the interface orientation; for the growth direction, this is:

$$\delta E_{sh}[001] = 2b(\varepsilon_\perp - \varepsilon_\parallel). \qquad (2.48)$$

Chemical Potential

Thermodynamic criteria for the classification of crystal growth mechanisms are based on the equilibrium shape of a cubic crystal deposited onto a substrate. This may be derived in terms of the overall change in specific surface energies of the substrate ζ_{sub}, epi-layer ζ_{epi}, and interface ζ_{inter} ($\Delta\zeta = \zeta_{epi} + \zeta_{inter} - \zeta_{sub}$) [120]. Surface energy is a measure of the disruption of inter-molecular bonds (excess energy) when a surface is created. Surfaces are less energetically favourable than bulk material. Hence work is required to either build up a surface area or to cleave bulk material. Alternatively, the interrelation of binding energies between two deposited atoms

(cohesion), and between that of an atom in the substrate and one in the epi-layer (adhesion) may be considered [121]. Both expressions are identical [122].

The Gibbs free energy, G, is the energy readily available to do useful work (at constant T):

$$\Delta G = \Delta H - T\Delta S. \tag{2.49}$$

Where enthalpy (ΔH) is the amount of heat energy transferred in a chemical process under constant pressure, and entropy (ΔS) is the measure of molecular disorder of a closed thermodynamic system.

The chemical potential, μ can be expressed as either the rate of change of internal energy of a system, U, or G, due to the change in number of atoms of a given substance in the over-layer, whilst keeping the number of atoms for all other types of particles, i, constant [123].

$$\mu_i = \left(\frac{\partial U}{\partial N_i}\right)_{S,V,N_{i\neq j}} = \left(\frac{\partial G}{\partial N_i}\right)_{P,T,N_{i\neq j}}. \tag{2.50}$$

The chemical potential of a thin film depends on film thickness, since the attraction between atoms within the top monolayer (ML) and the substrate decreases, as the distance from the interface increases. Furthermore, introduction of MDs due to either alloying or pseudomorphic growth can increase the hydrostatic stress, which in turn increases G. At 0 K, μ is the Fermi energy. Markov proposed a model [122] for the chemical potential for the first few MLs of deposited material:

$$\begin{aligned}
\mu(N) &= \mu_{\text{3D}} + [\epsilon_{\text{wl}} - \epsilon_{\text{sub}}(N) + \varepsilon_{\text{d}}(N) + \varepsilon_{\text{h}}(N)], \\
&= \mu_{\text{3D}} + a_0^2 \Delta\zeta + \varepsilon_{\text{h}}(f).
\end{aligned} \tag{2.51}$$

Where $f = f(a)$, and N refers to the lattice misfit and the number of atoms in the over-layer, respectively. The bulk chemical potential of the deposited material is denoted by μ_{3D}; ϵ_{wl} and ϵ_{sub} are the desorption energies required to release an adsorbed atom from the WL and the substrate, respectively; ε_{d} and ε_{h} are the MD energy per atom and the homogeneous strain energy per atom contained in a single ML of deposit, respectively (assuming the first few MLs of deposited material are equally strained).

Two regions of different chemical potential (e.g. across a heterojunction) form a gradient that acts as a thermodynamic force driving mass transport ($\Delta\mu = \mu(N) - \mu_{\text{3D}}$), i.e. the migration of an atomic species from higher to lower chemical potential; in order to reduce G. Peierls et al. [123] showed in the limit of small strains, that

the mechanism behind the mode of film growth is closely associated with how the chemical potential changes with respect to the number of atoms in a deposited film.

Wetting
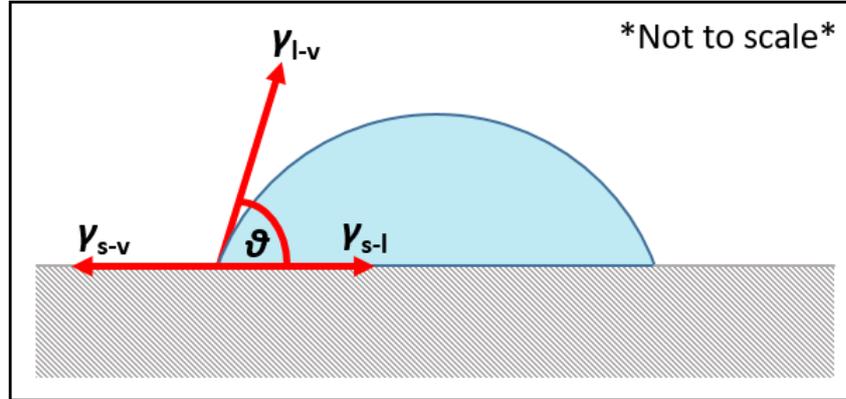
Figure 2.21: Schematic to show equilibrium shape of a liquid droplet on a smooth substrate. The interfacial tensions are represented by the red arrows.

Wetting is a measure of the contact area between a liquid and a solid surface. It depends on inter-molecular interactions, hence fundamental in understanding the bonding of two substances. The degree of wetting, coined *wettability*, is particularly useful in droplet epitaxy. It is conveniently conceptualised (see Figure 2.21) by the Young-Dupré equation below [124]:

$$\gamma^{\mathrm{sv}} = \gamma^{\mathrm{sl}} + \gamma^{\mathrm{lv}} \cos\left(\theta_{\mathrm{eq}}\right). \tag{2.52}$$

Where γ^{sv}, γ^{sl}, γ^{lv} and θ_{eq} are the solid-vapour, solid-liquid and liquid-vapour interfacial tensions, and equilibrium contact angle, respectively. The contact angle is an inverse measure of wettability [125] and is formed between the liquid-vapour interface and the substrate surface.

A contact angle $< 90\,°$ ($> 90\,°$) is considered low (high) and indicates that wetting is favourable (unfavourable). Liquids are strongly stressed by a hydrophilic surface and droplets tend to flatten out since the contact angle is small. Conversely, large contact angles correspond to hydrophobic surfaces, upon which compact liquid droplets are formed. Equation 2.52 is only applicable to flat, rigid, chemically homogeneous (ideal) solid surfaces, since surface roughness affects the contact angle [126]. Non-ideal surfaces result in contact angle hysteresis, a phenomenon defined where the advancing (θa) and receding (θr) contact angles are unequal [127].

Wettability was found [128] to be consistent with nucleation of 3D clusters, i.e. the Bauer criteria well-describe the transition from planar to 3D growth of islands

despite the trade-off between the cost of additional surface energy and energetic gain due to strain relaxation relative to the WL [129].

Growth Modes

For the epitaxy of thin films onto crystal surfaces, there are three primary growth mechanisms (see Figure 2.22) determined by the interaction strength between adatoms and the surface they collect on [128]: Frank-van-der-Merwe (FM), Volmer-Weber (VW), and Stranski-Krastanov (SK). The latter, being of importance to this work, is an intermediary case between the other two. The following discussion focuses on the distinctions between them.
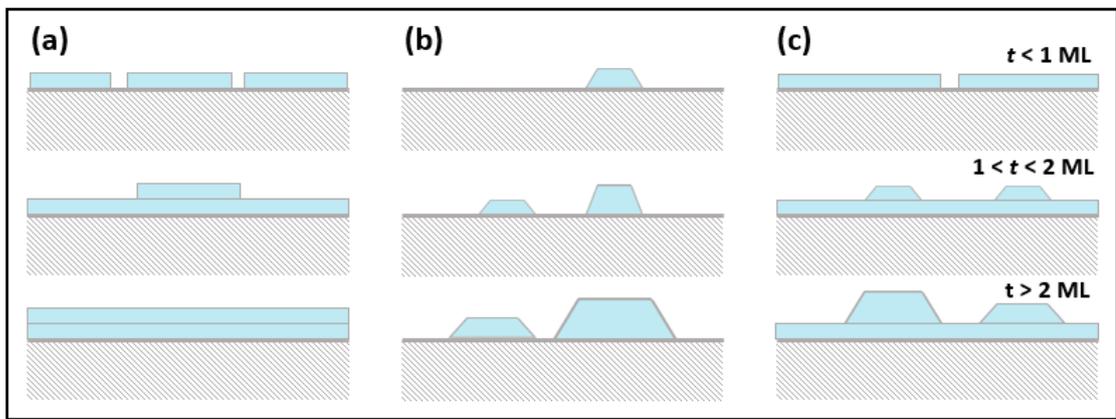
Figure 2.22: Schematic representation of the three main growth modes at three various stages of surface coverage.

The FM mechanism refers to 2D layer-by-layer growth, whereby adatoms preferentially stick to vacant surface sites and complete a full, atomically smooth layer before subsequent layers form. This growth mode is characterised by complete wetting of the substrate and negligible misfit ($\mu_{3D} \gg \varepsilon_{d,h}$), due to relatively strong adhesion ($\epsilon_{sub} > \epsilon_{wl}$), which corresponds to $\Delta\zeta < 0$ and a positive μ ($\partial\mu/\partial N_i > 0$).

FM growth is used to create basic planar structures. Bulk material can be simply deposited on a substrate. Secondly, a QW is formed by a very thin layer of semiconductor sandwiched in between a material of wider bandgap. A MQW is a periodic QW structure and features alternating layers of materials with different bandgaps; notably, carriers are confined to the z-direction and unable to interact with those in other layers due to a sufficiently thick barrier layer. Furthermore, a superlattice structure is formed when the thickness of barrier layers in a MQW are reduced enough for the carrier wavefunctions to overlap.

Conversely, VW growth refers to 3D island formation without the existence of a WL, whereby adatoms are much more attracted to each other than they are to the surface;

continued growth and coarsening results in the formation of rough multi-layer films. This growth is characterised by incomplete wetting due to relatively strong cohesion ($\epsilon_{sub} < \epsilon_{wl}$). This corresponds to $\Delta\zeta > 0$ and a negative μ ($\partial\mu/\partial N_i < 0$).

The VM mechanism is desirable for droplet epitaxy. This growth technique is effective at growing nanowires and nanodrilling a substrate in order to provide a patterned template for selective growth of a range of nanostructures [130].

SK growth is a two-step process characterized by the transition of 2D adsorbate layers, up to several MLs thick, into the nucleation and coalescence of 3D islands. Lattice misfit must be non-zero and $\Delta\zeta < 0$ for the formation of 3D islands above a WL. Importantly, stacking faults can be generated at the coalescence of these grown islands [131]. The WL and island are in different phases with respect to G, hence they cannot be in equilibrium with each other [128].

Three distinct morphologies are typically formed using SK growth, i.e. QDs, QRs, and clusters. InAs and GaSb QDs typically have the approximate shape of a truncated pyramid [132], whereas QRs appear as two lobes of a cleaved torus. A "cluster" is an umbrella term representing a wide variety of undesirable nanostructures due to being highly disordered. The Sb concentration is strongly inhomogeneous in the case of GaSb/GaAs, hence it is possible to observe clusters that contain GaAs-linked chains of pure GaSb regions, as shown in Figure 2.23(c).
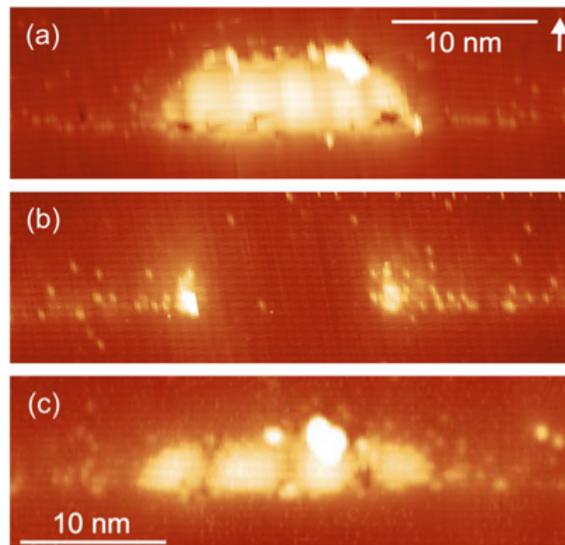
Figure 2.23: cross-sectional scanning tunnelling microscopy (XSTM) images of quintessential GaSb/GaAs nanostructures; revealing a) a QD, b) a QR cleaved centrally, and c) a segmented cluster. The arrow indicates the growth direction and the bright points are filled electron states. The scale is the same for a) and b). Figure adapted from [132].

A critical lattice mismatch f_c exists, which represents the boundary between SK and FM growth. It depends on material parameters such as chemical bond strength

(material stiffness) and degree of anharmonicity (i.e. a Lennard-Jones potential with steeper repulsion and weaker attraction branches) [128]. The growth of coherent 3D nano-islands is favourable when $|f(a)| > f_c$; below f_c, the strain energy is eventually accommodated by MDs (see Subsection 2.2.4). Moreover, f_c limits the thickness of a stable WL to the range of action of the inter-atomic bonding [128]. Taking $In_xGa_{1-x}As/GaAs(001)$ into consideration, Walther et al. found that an In concentration of ~ 25 % must be exceeded in order for nucleation to occur [133].

Nucleation

The transition between strained and relaxed regimes occurs when elastic stress caused by lattice mismatch, during the first few MLs of pseudomorphic growth, accumulates until the film reaches h_c. This parameter depends on $f(a)$ and chemical potential differences. The planar film is therefore stable up this point. At h_c, the total strain energy begins to exceed the energy required for a biaxially-strained film to undergo a structural transformation. Thus it becomes energetically favourable for the film to relax irreversibly; i.e. it returns to its natural bulk lattice parameter and forms as many atomic bonds as possible with the nearest neighbours underneath. Here, strain is compensated for by dangling bonds and other crystallographic defects [134]; some of which may form TDs [135]. In the latter case the deformation energy is largely dependent on surface quality rather than layer thickness.

Growth of stable lattice-matched layers thicker than h_c is possible, due to the energy required to form a new dislocation being greater than the energy to sustain one [136]. Though this depends on growth temperature, growth rate and interface quality, i.e. the process must not drastically increase surface roughness or introduce a significant density of defects. Although undisclosed, this actually happened during the growth of GaInAlAs/InP MQWs in the work of Ref. [137]. Interestingly, the materials would relax during the device processing stages; leading to the presumption here, that etching and cleaving increased the free surface energy, hence triggered the nucleation (and glide) of dislocations. This process is analogous to the supercooling of water.

The nucleation concept is based on the instability of planar growth against clustering. The N-dependence of G of the growing film has an inflection point, where the curvature changes from positive to negative. This point is observed to occur at a slightly greater thickness than h_c [123, 128]. QDs nucleate because ε_h is lower in the nano-island than it is in the WL. The onset of relaxation of misfit strain at the island edges overcompensates for the surface energy of the 3D crystallite side facets [128]. To elaborate: Anharmonic chemical bonding influences the adhesion of 2D islands to the WL; thus leading to vertical displacements of atoms close to

the 2D island step-edges, away from the minima of their potential valleys. Atoms then detach from the edges of the initial ML-tall island and attach to the double step of the bi-layer island growing above. Nucleation of the second layer typically occurs at the edges as corners since the atomic separations here closely resemble the lattice parameter of the bulk crystal [138, 139]. This yields a tendency towards VW growth, i.e. a positive wetting parameter $\Phi = 1 - \epsilon_{\mathrm{sub}}/\epsilon_{\mathrm{wl}}$. Larger displacements of the edge atoms results in weaker average wetting of the substrate, therefore Φ increases with additional island height [128].

Monolayer islands may turn into multilayer islands by the consecutive layer-by-layer nucleation and growth of single MLs, or by the nucleation and lateral growth of multilayer islands. These multilayer islands remain constant in height. Layer-by-layer (lateral) island growth is favourable when compressed, stiff (tensile, soft) films are deposited.

2.2.4 Crystalline Defects

It is highly unlikely for a real crystal to exist with a perfect structure. At the very least, it is currently impossible with cutting-edge technology to produce a 100 % pure semiconductor in bulk quantity; even the purest materials contain $\sim 10^{12}$ impurities cm^3 [82], i.e. it is increasingly difficult to fabricate materials with impurity concentrations lower than parts per billion. Although seemingly negligible compared to the total density of atoms for typical bulk samples (GaAs: $\sim 4 \times 10^{22}$ atoms/cm^3 [140]); these sites govern some of the most important properties of crystals such as material strength and carrier mobility.

Impurities can appear in different forms, e.g. as small distinct crystals (precipitates) or as clusters, however more often than not they are dispersed as single atoms throughout the crystal [141]. It was previously accepted that impurities are distributed randomly in a semiconductor. However, it is currently understood that Coulomb correlation effects result in significantly ordered and non-Poissonian distributions, which may alter the material transport properties [142, 143].

The maximum impurity concentration achievable in a semiconductor is limited by segregation effects. Particles redistribute themselves during epitaxial growth in order to reduce their concentration as implied by Equation 2.53 [144]:

$$\nu_{\mathrm{d}} = \frac{eC_{\mathrm{D}}}{k_{\mathrm{B}}T} \sqrt{\frac{2exV_0^{\mathrm{ps}}}{\varepsilon_{\mathrm{x}}}}. \tag{2.53}$$

Where x is the concentration of an incorporated substance, V_0^{ps} is the barrier height

of the pinned surface, C_D is the diffusion coefficient of the impurity, and ε_x is the absolute permittivity of a given substance. Here, ε_x refers specifically to the impurity. All other symbols have their usual meanings.

The repulsive interaction between impurities is due to the total incorporation energy, E_{inc}, being increased, i.e. it is easier for a diffusion hop to occur. This is because impurities are ionic, therefore electrostatic interactions during growth causes them to either cluster to form a bulk defect, i.e. a small region of a different phase; or drift towards the surface since free carriers are depleted near it, due to surface states pinning the Fermi level. Furthermore, a higher concentration increases the Fermi energy due to state-filling (electric field at the surface increased), which in turn increases E_{inc} [144].

Defects often cause the occurrence of deep levels in the bandgap. A crystalline defect is defined as an interruption of the regular periodicity of atoms in a crystalline solid. Sites within a specific region may be identically occupied, but not all are perfectly positioned [141], e.g. dislocations. Alternatively, sites within a given region may be perfectly positioned, but not all are occupied with identical atoms, e.g. dopants (excluding surface contaminants). Types of defect include: point, line, planar, and bulk. Elastic lattice deformations are not considered as defects; they are responses to an external force. A true imperfection is one that persists through any elastic distortion, akin to a hole in a rubber band [141].

The following discussion focuses on point and linear defects, owing to their significance to this work. Although planar and bulk defects will cause little interference on the measured PL spectra in Section 5.4, that is not to disregard the possibility of their existence entirely.

Point defects

Point defects are those which occur only at or around a single lattice point. They are not considered to have any spatial dimension. These defects typically involve at most a few extra or missing atoms. Many types of point defect exist: e.g. interstitial, vacancy, substitutional, antisite etc. Moreover, complexes of various defect species can occur, e.g. a vacancy binding to an impurity, or shared atomic sites (split interstitials) [145]. Weatherly et al. are the first authors to spatially resolve and image individual buried point defects at the nanoscale, by using high-resolution cathodoluminescence to spatially resolve, image and analyse InGaN/GaN QWs [146].

The simplest cases to consider are the vacancy and interstitial defects. The former is an empty lattice site whereas the latter is an atom of the host material embedded

within an interstitial site [141]. Usually, an interstitial solid solution is formed if the defect is relatively small compared to the host atoms, i.e. occupy sites between nearest neighbours. Atoms such as hydrogen [147] can occupy interstices at low energies. When defects are similar in magnitude to the host atoms, then they usually occupy the main crystal lattice sites and form a substitutional solid solution [141].

Given that it requires an energy, E_{pd}, to form a specific point defect inside a crystal lattice; the Boltzmann distribution may be used to estimate the probability of occurrence, thus predict the equilibrium concentration of defects, n_{pd}:

$$n_{\mathrm{pd}} = \frac{N_{\mathrm{pd}}}{N_{\mathrm{ds}}} = \mathrm{e}^{-E_{\mathrm{pd}}/k_{\mathrm{B}}T}. \tag{2.54}$$

Where N_{pd} is the number of point defects and N_{ds} is the number of potential defect sites.

Line Defects

Line defects are confined to 1D, whereby atoms in the crystal lattice are misaligned as shown below. The two fundamental types of dislocation are edge dislocations and screw dislocations. Dislocations are often referred to as *mixed*, since they usually have a combination of edge and screw qualities.
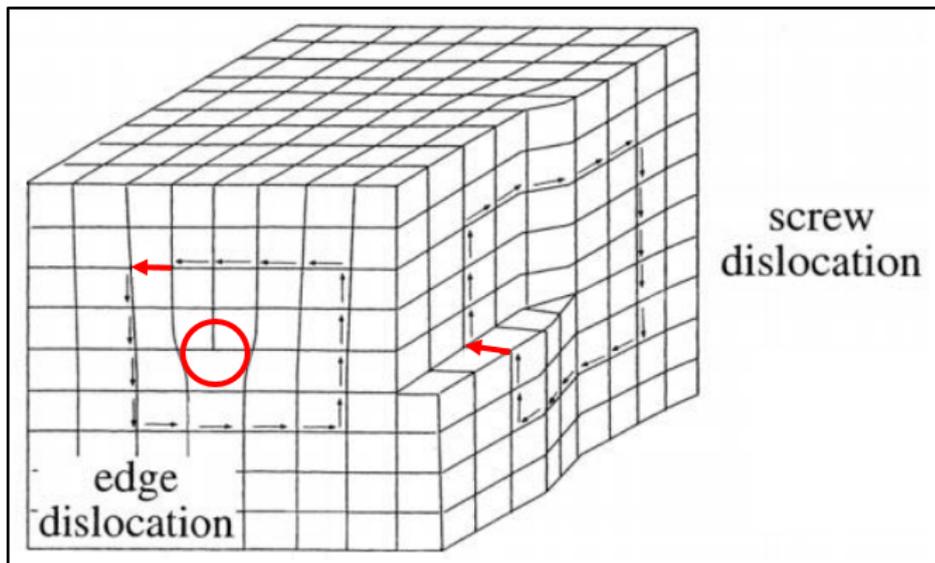
Figure 2.24: Schematic to show two types of dislocation which may manifest at a heterojunction; provided that $\Delta a/a$ is significant. The red arrows represent the Burgers vector, **b**, of the edge and screw dislocations, whereas the red circle highlights a dangling bond site. Figure adapted from [148].

Edge dislocations due to the termination of an atomic plane in the middle of a crystal. The adjacent planes bend to compensate for the breaking of periodic conditions.

It is helpful to think of these defects as an extra half plane of atoms. Screw dislocations are defects where atomic planes follow a helical path around a dislocation line. Their presence causes a distortion of the crystal lattice, i.e. strain, expressed in terms of b. For a screw (edge) dislocation, b is parallel (perpendicular) to the dislocation line.
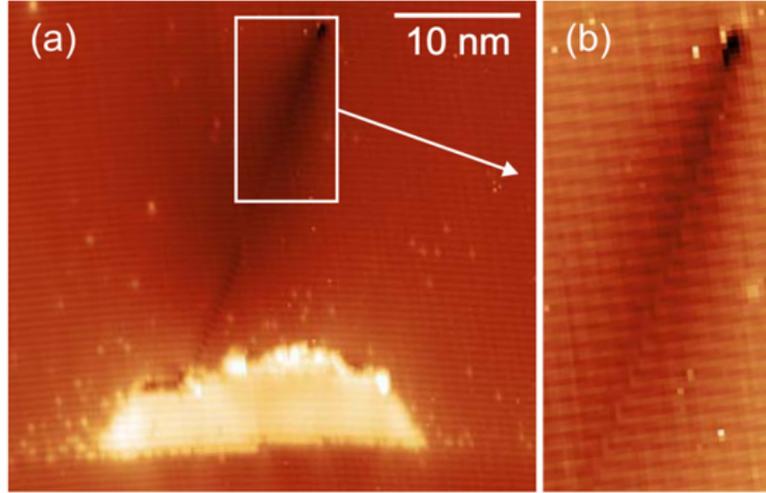
Figure 2.25: Filled-state XSTM image of a screw location above a GaSb/GaAs QD. To the right is a magnification of the defect core; showing the phase mismatch of the atomic rows at the dislocation boundary. The dislocation terminates ~ 15 nm above the surface at the cold cap interface. Figure adapted from [132].

Dislocations are able to move due to the breaking, and re-bonding, of atoms in the surrounding planes with atoms at the terminated region. Due to conservation of energy the dislocation cannot terminate within a crystal. They propagate from a strained layer surface and must glide to the edge of the wafer or turn up towards the growth interface to form TDs [149]. It may penetrate the substrate or bend at an interface into a MD [150]. For the GaSb/GaAs material system, dislocation glide enhances diffusion of Sb atoms, leaving an Sb-rich region trailing behind the movement of a dislocation core richer in Sb [33].

High densities of TDs can develop during growth to relax significant amounts of mismatch-induced strain. It is energetically favourable for TDs to form when the distance from dislocation to the sample edge is much greater than the distance between the dislocation and the epi-surface, hence why TDs are abundant in the laser structures discussed here. In fact, the majority of research devoted to the III-V/Si epitaxial platform so far has been based around the suppression of TD formation during growth for the following reason. TDs have associated trap states which act as non-radiative recombination centres and tend to accrue further point defects in their vicinity. This promotes further non-radiative recombination and gradual device degradation through enhanced dislocation climb [21, 151], therefore
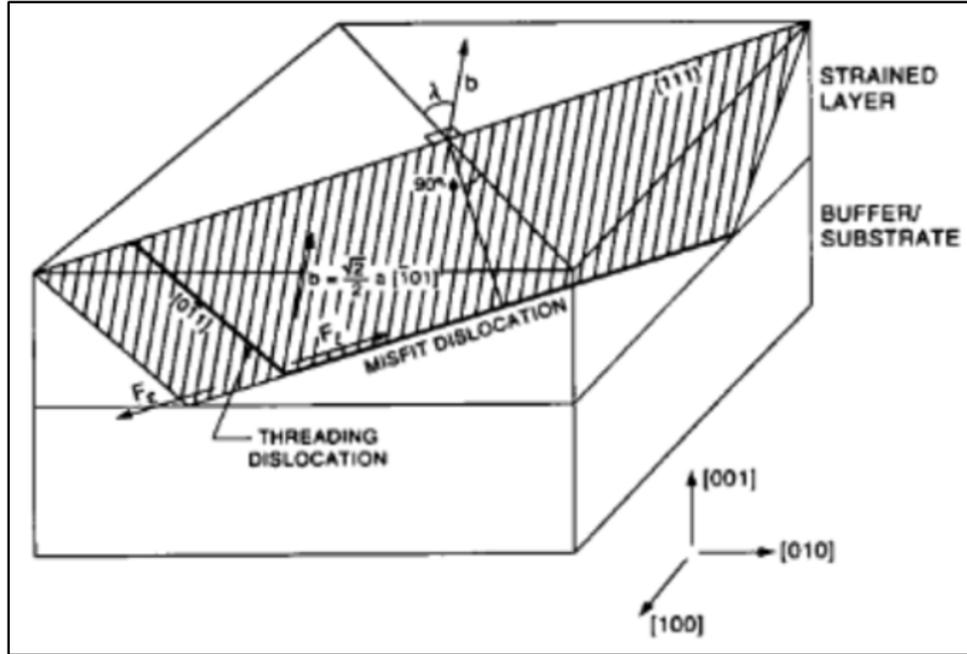
Figure 2.26: Schematic to show the dislocation flow process. The hatched region is a (111) plane in a strained-layer. Figure taken from [115].

yielding lower quantum efficiencies and shorter lifetimes in III-V/Si photonic devices. Comprehensive discussions on MD interactions and dynamics can be found in Ref. [152] and [153], respectively.

2.3   Light Production in Semiconductors

Here, light classification is discussed to clear up any potential confusion typically associated with optical processes. This section then explains how semiconductors typically absorb and emit light; supplemented with detailed illustrations of radiative and non-radiative recombination mechanisms. Here, the bulk case is shown for simplicity and to maintain transparency of the bands in $\boldsymbol{k}$-space, however Figure 2.31 depicts the 0D case relevant to these structures. Finally, the internal and external quantum efficiencies of these materials are introduced.

2.3.1   Classification of Light

Light can be classified as either "incandescent" or "luminescent". Incandescence is light emitted from a hot body due to an increase in T, e.g. a tungsten filament. Incandescence has a continuous spectrum and is not to be confused with thermoluminescence. The latter referring to ionizing irradiation inducing a displacement of electrons; heating the material then triggers de-excitation and produces light. Conversely, luminescence is an umbrella term for a wide range of phenomena, where a

luminescent material otherwise known as a "phosphor" emits light whilst relatively cold.

Importantly, PL refers to any quantum mechanical, light-emitting process induced by photon absorption. The category is further subdivided into "fluorescence" and "phosphorescence" depending on how long PL occurs for after photoexcitation. More specifically, whether the radiative transition requires a change in spin multiplicity. Fluorescence occurs very shortly after photoexcitation and decays rapidly after the excitation stops; whereas phosphorescence is long-lived PL that continues long after photoexcitation. Most photoluminescent events are fluorescent.

A photon is essentially a piece of light, i.e. a quantum region of space with a propagating and rapidly fluctuating electromagnetic field. The electric field oscillates at ν, and has an associated energy E:

$$E = h\nu = \frac{hc}{\lambda}. \tag{2.55}$$

Where λ is the photon wavelength inside a vacuum, and the other constants have their usual meanings.

2.3.2   Photo-Absorption

There are different types of absorption event which occur under a range of conditions. The way in which an atom can absorb light depends on its state. Extending this principle to an ensemble of atoms, the bulk electronic properties of a lattice determine the way in which a semiconductor absorbs light. To elaborate, discussion here begins with a focus on general inter-subband transitions. The main transitions relevant to this work are then summarised. Notably, the vibrational and rotational atomic modes are not applicable to this study.

When a photon interacts with a carrier: there is a probability that the carrier will absorb it, and use its energy to "jump" from its current subband (with energy E_1) to a vacant higher-energy subband (of energy E_2), subject to selection rules; leaving a hole behind. The ν of the photon involved in the absorption process corresponds to the exact difference in energy levels, as expressed by Equation 2.56:

$$\nu = \frac{E_2 - E_1}{h}. \tag{2.56}$$

If ν does not match any of the differences in subband energies, then the light cannot be absorbed and it continues along its path. Notably, the simultaneous absorption

of multiple photons is possible provided the combined energies match the energy required to make the transition.

Inter-Band Absorption
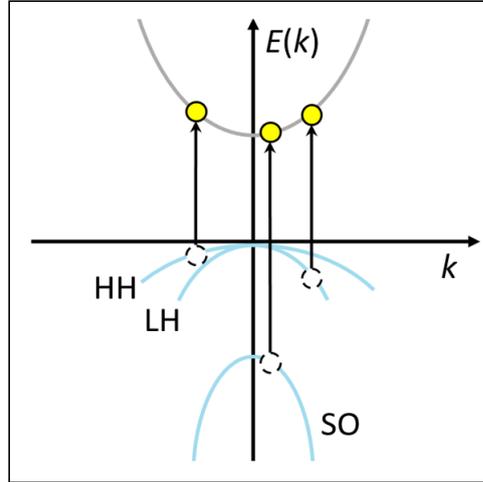
Figure 2.27: Energy schematic to show possible direct electronic transitions during an inter-band absorption event. The yellow and empty circles represent the conduction electrons and the valence holes, respectively. The arrows represent the energetic direction of the transitions. Furthermore, a continuum of state exists above and below the CB and VB, respectively.

The inter-band absorption process happens across the bandgap, as illustrated in Figure 2.27. A pair of free carriers are generated when the incident photon has an energy greater than the bandgap. These hot (quasi-)particles can move independently through the lattice.

Carriers are unstable post-excitation. Non-equilibrium states thermalise by vibrational relaxation, i.e. phonon interactions cause the hot carriers to quickly lose energy in the form of heat to the crystal lattice. The carriers relax rapidly towards each other due to the opposite signs of band curvature. Energy and momentum are conserved as the electron moves to a lower vibrational energy level in the same electronic state. The carriers reach their minimum energy state over very short timescales relative to the radiative lifetime, hence the relaxation process happens at a much faster rate than that at which the carriers radiatively recombine. Thus, carriers tend to collect at the band extrema.

The number of generated electron-hole pairs depends on the laser power density. Subband states start to become more populated with an increase in P and this leads to the band-filling effect. Assuming negligible reflections, the rate of photon absorption is determined by the material thickness and absorption coefficient, $\alpha(\lambda)$ [cm^{-1}]. The intensity of light inside a material, $I(z)$, decays exponentially with depth, z, and can be calculated with Equation 2.57:

$$I = I_{\mathrm{s}} \mathrm{e}^{-\alpha(\lambda)z}. \tag{2.57}$$

Where I_{s} is defined as: the excitation power, P, per unit (cross-sectional) area, A, at the surface. The inverse of $\alpha(\lambda)$, i.e. the penetration depth, determines the distance a photon can penetrate a material before it is absorbed by it; hence dependent on the material itself and the incident photon λ. For a direct transition [154]:

$$\alpha(\lambda) \propto \sqrt{h\nu - E_{\mathrm{g}}}. \tag{2.58}$$

Here, E_{g} refers to the energy of the optical bandgap (see Figure 2.28). A thin material appears transparent to a photon if it has a small $\alpha(\lambda)$-value, such as inside the active region of the gain media of this work.

Assuming a one-to-one relationship between a single photo-absorption event and the creation of a single electron-hole pair; then differentiating Equation 2.57 with respect to z will return the change in light intensity across a thin horizontal cross-section, i.e. the generation rate, G, at a specified depth inside the material:

$$G = \alpha(\lambda)N_0 \mathrm{e}^{-\alpha(\lambda)z}. \tag{2.59}$$

Where N_0 is photon flux at the surface, in units of photons/unit area/second.

Exciton Formation

It is possible for an electron at the CB minima to bind with a hole at the VB maxima. This forms an electrically neutral quasi-particle known as an *exciton*. A positronium-like atom is the simplest way to visualise this, where an electron and hole orbit around their common centre of mass. Excitons can be observed at low P during PL measurements.

Since MBE was used, instead of MOCVD, to grow the samples presented in this thesis; large amounts of unintentional doping is unlikely [158]. However, it is even less likely that the background C-doping is zero, due to the age of the system, and given the deep confining potential of GaSb/GaAs nanostructures; the possibility of these nanostructures containing holes prior to optical pumping cannot be excluded. It is therefore deemed appropriate to regard the term "exciton" in a more general sense [47], as to include charged excitons and other exciton complexes [159] (e.g. biexcitons [160]). Due to the spatial separation of electrons and holes: any excitons in this work will be fundamentally different from those found in typical type-I

systems. Hence, here they are analogous to atoms with electrons in stable orbits around a positively charged nucleus [161, 162].

The CB electron is weakly attracted to the hole by the Coulomb potential $U(\boldsymbol{r})$:

$$U(\boldsymbol{r}) = -\frac{e^2}{4\pi\varepsilon_0\varepsilon_r a_X}, \tag{2.60}$$

Where a_X denotes the exciton Bohr radius (Equation 2.61), i.e. the separation between the carriers, ε_0 is the vacuum permittivity, and $\varepsilon_r = \langle\varepsilon\rangle/\varepsilon_0$ is the relative permittivity (dielectric constant); here, $\langle\varepsilon\rangle$ is the average dielectric constant of both carrier-confining layers. The reduced mass of the electron-hole system, $\mu^* = m_e^* m_h^*/(m_e^* + m_h^*)$, and the Bohr radius, $a_B \approx 0.053$ nm, can be used in Equation 2.61 to calculate a_X [94]:

$$a_X = \frac{m_0 \varepsilon_r a_B}{\mu^*}. \tag{2.61}$$

Excitonic stability is only achieved when $U(\boldsymbol{r})$ is strong enough to prevent phonon interactions from breaking the carriers apart, i.e. when $E_b > k_B T$ [163, 164]. Specifically, this is the energy required to ionize the exciton, or more appropriately for the context of this work; remove the electron from the region it is confined to. The electrostatic force arising from $U(\boldsymbol{r})$ is screened by a net repulsion from the density of electrons surrounding the hole. Therefore sufficiently increasing temperature, or the carrier density, ionizes the excitons; which releases free carriers that further screen the Coulomb interaction, resulting in the acceleration of excitonic ionization [165]. These competing forces cause the exciton to have less energy than the free carriers, hence the exciton state is represented as a forbidden state within the bandgap (see Figures 2.31 and 2.28).

There are two main types of excitons depending on the characteristics of the semiconductor material containing them e.g. dielectric constant, doping, crystalline quality, and architecture. Materials with a large dielectric constant (e.g. semiconductors) can store relatively more electrical energy. The internal electric field is enough to significantly reduce the Coulomb interaction between the constituent carriers. As a result, the exciton has a larger radius than the lattice spacing. These are *Wannier-Mott*, or *free* excitons, owing to their ability to move throughout the lattice with an associated kinetic energy, E_{kin}. Notably, a small electron effective mass also favours the formation of this type of exciton.

In contrast, materials with a small dielectric constant typically allow for a stronger Coulomb interaction between an electron and hole. Hence the exciton radius tends

to be relatively small. Usually, radii are of the same order of magnitude as the size of the unit cell. These are known as *Frenkel* excitons; they are localised states of electron-hole pairs residing at the same lattice cell, whether due to confinement effects or the binding to an (artificial) atom or defect. Hence their movements are limited to hopping mechanisms (or scattering processes). This type of exciton is relevant to the gain medium presented in this thesis, due to the dimensionality of the QW, i.e. the electrons are effectively pinned to the heterojunction, whilst the holes are contained within regions of GaSb (predominantly QRs).

Two definitions of bandgap emerge for semiconductors with a significant E_b, as shown in Figure 2.28. The optical bandgap, $E_g^{optical}$, is the threshold for photon absorption, i.e. the energy required to form one bound exciton (Figure 2.28). The electrical bandgap, $E_g^{electric}$, is the difference between CB and VB edges, i.e. the energy required to form one free electron-hole pair, thus: $E_g^{electric} > E_g^{optical}$. The electrical bandgap is useful for LDs and LEDs, whereas optical bandgap measurements are important for solar cells.
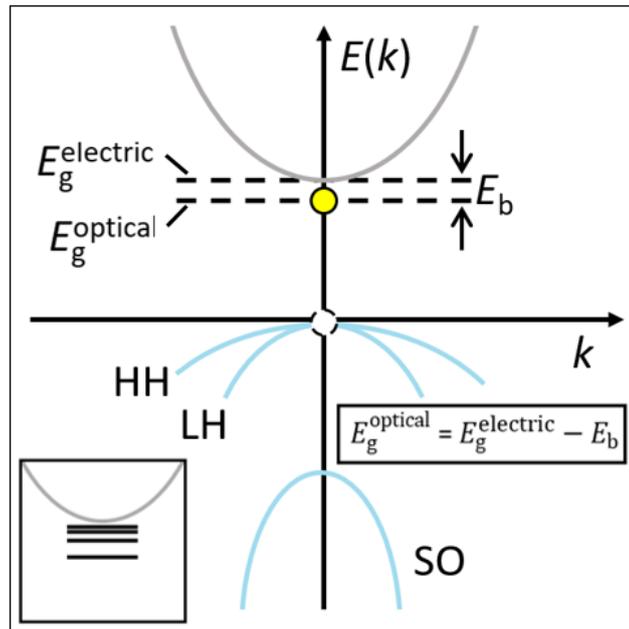
Figure 2.28: Bulk band structure schematic to show the location of carriers upon the formation of a bound exciton. Incident photons often have just enough energy to create the exciton, but not enough energy to separate the carriers. The transition arrow is omitted to emphasise the creation of an exciton. Inset: Close up of the energy levels corresponding to excitonic states within the bandgap.

Excitation under resonant conditions [90] yields a hydrogen-like absorption spectra for a direct bandgap semiconductor; specifically, a parabolic spectrum of exciton energies with a large density of excited states. Bound excitons are created at the ground state, $k_X = |0\rangle$, without any change of momentum ($E_{kin} = 0$). For an absorption transition with no net change in momentum: the carrier group velocity,

$\mathbf{V_g}$ must equal zero. This can only be true where the gradient of the dispersion relation is zero, i.e., at the CB and VB extrema.

Non-vertical excitation cannot occur under resonant conditions; without the involvement of a phonon, in a higher-order process, which has a tiny probability of occurring [163]. This is because $\lambda \gg a$. Therefore, the value of $|\mathbf{k}|$ associated with the photon is negligible in comparison with the dimension of the Brillouin zone $(2\pi/a)$. As a result, the quasi-momentum of the absorbed photon cannot sufficiently change $\mathbf{k_X}$ for the required non-vertical transition.

It is not to say that a free exciton ($E_{kin} \neq 0$) cannot easily be created in a direct bandgap semiconductor. Incident photons with an energy significantly greater than the bandgap generate free electron-hole pairs, which bind to excitons and then thermalize by emission of phonons into the crystal lattice. The duration of this process is very short [163]. The crystal lattice must be of high purity (and at a LT) for a free exciton to exist; since impurities are very efficient at localising the exciton wavefunction, thereby nullifying E_{kin}.

Intra-Band Absorption
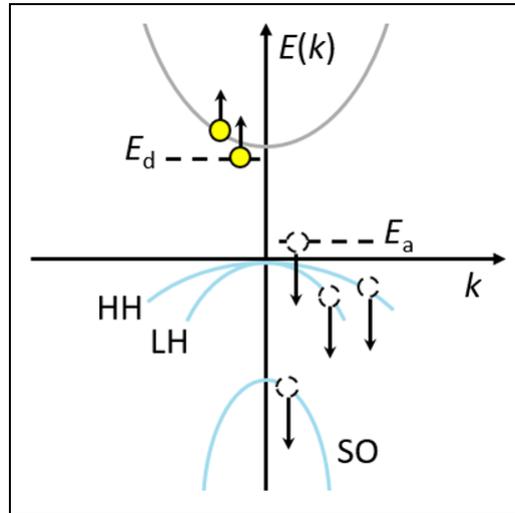
Figure 2.29: Energy schematic to show possible direct carrier transitions during an intra-band absorption event. The yellow and empty circles represent the conduction electrons and the valence holes, respectively. The arrows represent the energetic direction of the subband transitions. The upper and lower forbidden states are the donor and acceptor states at energies E_{donor} and $E_{acceptor}$ respectively. Here, higher order transitions are neglected.

The intra-band absorption process is illustrated in Figure 2.29, where a photon interacts with an electron (hole) in the CB (VB). The carrier is then raised (lowered) to an energy level with greater (reduced) energy. This process generally includes interactions with carriers in bound states at energies near the band edges, and transitions from higher order levels.

Intra-band absorption poses a greater threat to the performance of mid-infrared devices in comparison to those operating at shorter λ (e.g. NIR, visible). Moreover, it generally starts to dominate other absorption mechanisms when the incident photon energy is less than the T-dependent bandgap energy. A quantum mechanical treatment [155], where the photon energy was set comparable to the bandgap, revealed that this process is insignificant in semiconductor lasers. For the above reasons, this process is neglected in the analysis of Chapter 5.

Inter-Valence-Band Absorption
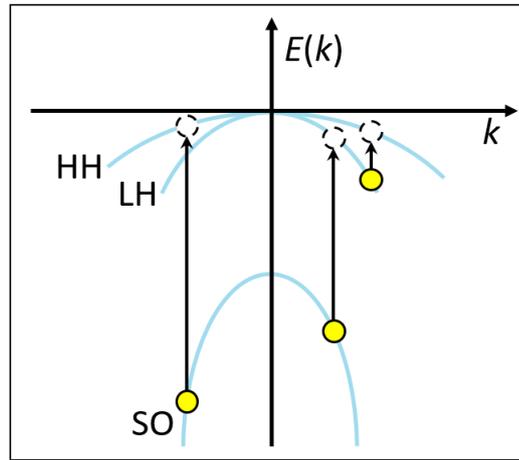
Figure 2.30: Energy schematic to show possible direct carrier transitions during an inter-valence-band absorption event. It should theoretically be possible to excite an electron from the LH band to the HH band; however in reality, materials rarely have a bandgap energy approximately equal to the energy difference between the SO and HH/LH bands.

Inter-valence-band absorption is the process illustrated in Figure 2.30, whereby a photon resulting from a radiative recombination event is re-absorbed by an electron in the split-off band, and excites the electron to a vacant state in the HH or LH bands. This process is significant when the bandgap is approximately equal to the difference in energy between the SO and HH/LH bands, i.e. $E_{\mathrm{g}}(T) \approx E_{\mathrm{so}} - E_{\mathrm{hh}}$. For the GaSb/GaAs material system, the bandgap is greater than this difference in VB energies; except at large values of $\boldsymbol{k}$. It is assumed that the majority of recombination happens at (or very close to) the Γ-point, hence inter-valence-band absorption is also ruled out as a significant process in the reference sample. The same cannot be justified for the RWELL samples because E_{g} is expected to reduce and may become comparable to the $E_{\mathrm{so}} - E_{\mathrm{hh}}$ energy difference as the CB edge is lowered.

2.3.3 Fluorescence

A transition across the bandgap can be made by releasing a photon with a longer λ

than that of the incident light, in one of several ways illustrated in Figure 2.31. Both parts of the fluorescence processes take enough time for the emission of scattered light to be considered isotropic.

In this work, PL emission is very likely dominated by fluorescence caused by band-to-band and excitonic recombination. The probabilities of donor-to-acceptor, band-to-acceptor, and donor-to-band are low. It is unlikely for many donor and acceptor sites to be located within proximity of each other and $\Delta\psi$ is reduced for these RWELL structures.

Indeed, it is also possible for the electron to immediately scatter elastically back down to E_v. This process emits a photon, at the same λ as the incident photon, with a direction of scatter that conserves momentum. Elastic scattering will have a negligible effect on the PL results presented in Section 5.4; since the photon will either: escape, be reabsorbed elsewhere, or be indistinguishable from the PL peak of the frequency-doubled laser source (see Subsection 3.3.1).
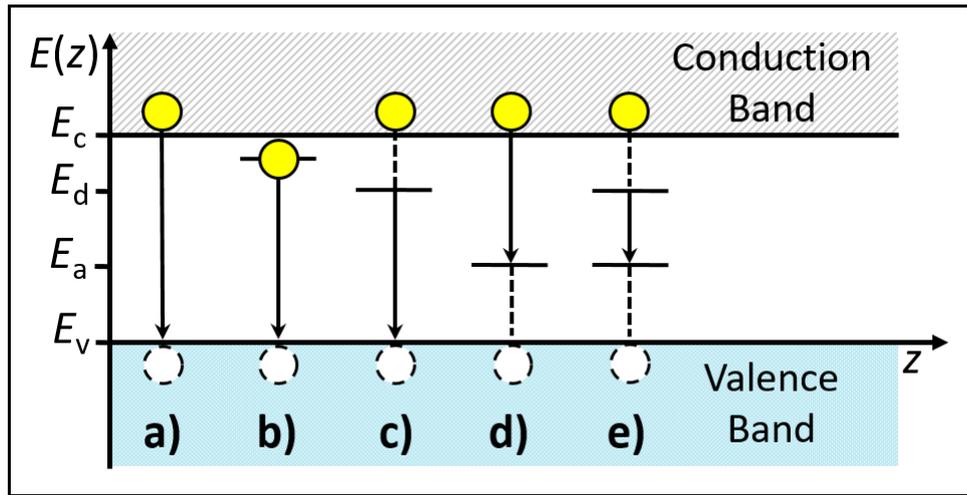
Figure 2.31: Schematic band diagram of the following radiative recombination mechanisms: a) band-to-band, b) excitonic, c) donor-to-valence-band, d) conduction-band-to-acceptor, and e) donor-to-acceptor recombination. Phonon emission is represented by dashed lines. Although deep-level transitions can sometimes have a radiative nature, as observed in n-type GaN [156]; here they will be discussed in the context of Shockley-Read-Hall (SRH) recombination (see Subsection 2.3.3) and are omitted from this diagram accordingly.

Band-to-Band Recombination

Band-to-band (BtB) recombination refers to the common process where an unbound electron in the CB recombines with an unbound hole in the VB, as shown in Figure 2.31. A photon is formed with average energy:

$$h\nu = E_\mathrm{g} + \frac{1}{2}k_\mathrm{B}T. \tag{2.62}$$

The first term denotes the bandgap energy at T, whereas the second term is the thermal contribution to the photon energy, taking equipartition of energy into account whilst assuming one degree of freedom. The Boltzmann constant is referred to as k_B. The emitted photon has a similar energy to the bandgap, thus weakly absorbed by the semiconductor. This is because electrons and holes tend to collect close to the bottom of the CB minima and the top of the VB maxima, respectively.

The probability for a radiative electron-hole recombination event to occur is proportional to the concentrations of both carriers. The radiative recombination rate, per unit volume, can be approximated as [82]:

$$\tau_r^{-1} = -\frac{dn_h}{dt} = -\frac{dn_e}{dt} = Bn_h n_e. \tag{2.63}$$

Where B is the bimolecular recombination coefficient. Typical values are of the order $\sim 10^{-10}$ cm^{-3}s^{-1} ($\sim 10^{-13}$ cm^{-3}s^{-1}) for direct (indirect) bandgap bulk III-V semiconductors [157]. This difference is to be expected since a phonon is needed to conserve momentum and energy for indirect gap materials (see Subsection 2.1.4).

A significantly smaller B factor for Si indicates that BtB recombination is much less probable; thus highlighting the necessity for quantum efficiency improvements (see Subsection 2.3.4) should materials compete with market-leading gain media.

Excitonic recombination

Normally, excitons can only recombine at the band minima due to the rate of carrier relaxation and opposing signs of the CB and VB curvatures where $k \neq 0$. The recombination of a bound exciton produces a photon with energy according to Equation 2.64.

$$h\nu = E_g + E_b. \tag{2.64}$$

PL can be used to experimentally define E_b (Equation 2.65a) as the difference between the electron-hole emission peak, $E_{e_1-hh_1}$, and the exciton emission peak, E_X. If E_X is unresolved in the PL spectra, a theoretical estimate of E_b may be obtained (Equation 2.65b) for the exciton ground-state (n=1), by subtracting the theoretical exciton energy from the combined electron and hole subband energies [166].

$$E_b = E_{e_1-hh_1} - E_X, \tag{2.65a}$$

$$= E_{eh} - \frac{4\mu^* R_y}{m_0 \varepsilon_r^2}. \tag{2.65b}$$

Where $R_y = 2.18 \times 10^{-18}$ J, represents the Rydberg energy, and all other symbols have their usual meanings. The factor of $(n-1/2)^2 = 1/4$ arises in the denominator of the exciton eigenenergy term, due to the dimensionality of the QW [167]. The carrier confinement energy, $E_{eh} = E_h + E_e$, is the addition of hole confinement energy from the nanostructures E_h, and electron confinement energy from the QW, E_e. If a carrier is trapped, e.g. by a defect, or strain-induced potential minima, then E_{eh} is augmented by the energy required for the carrier to overcome the barrier caused by the trap potential at that site. The hole confinement energy, E_h, inside a QD can be estimated by Equation 2.66:

$$E_h = \frac{\pi^2 \hbar^2}{2m_h^* \varnothing^2}. \tag{2.66}$$

Where $\varnothing$ represents the QD diameter, and not the radius [168]. The electron confinement energy, E_e, inside a QW with potential barrier height V_0, can be estimated using the finite QW model (Equation 2.67), of which, tunnelling effects are taken into account:

$$E_e = \frac{\hbar^2 k_n^2}{2m_e^*}. \tag{2.67}$$

Here, m_e^* is the effective mass of the electron inside the QW, and the constant k_n is referred to (with caution) as a wavenumber. Importantly, k_n is not related to momentum; rather, it is a count of the number of crests belonging to a wavefunction inside the QW. Inside the QW, $k_n = \sqrt{2m_e^* E}/\hbar$; whereas within the barrier, $k_n = \sqrt{2m_{barrier}^*(V_0 - E)}/\hbar$ and $m_{barrier}^*$ is the electron effective mass inside this region.

Many particle interactions, caused by increasing I_{ls}, will change E_{PL}. This energy shift of excitonic recombination ΔE_X is separable into three distinct contributions [159]:

$$\Delta E_X = \Delta E_{e-e} + \Delta E_{h-h} + \Delta E_{e-h}. \tag{2.68}$$

Where ΔE_{e-e}, ΔE_{h-h}, and ΔE_{e-h} are contributions from the electron-electron, hole-hole and electron-hole interactions, respectively. In type-I structures, all three contributions are approximately equal, owing to their similar electron and hole wave functions [169]. However for a type-II structure, ΔE_{e-h} is relatively smaller due to the spatial separation of charges. The RWELL structure explored in this work is expected to have a greater energy shift ΔE_{e-e} than standard GaSb/GaAs QRs [47, 81, 91, 170], due to an increase of electron interactions with additional QW confinement.

Donor-to-Band Recombination

Donor-to-band recombination is a radiative process that usually occurs at LT, where an excess electron is trapped in a state just below the CB edge, i.e. at a *donor* level. The bound electron recombines with a free hole inside the VB and emits a photon with an energy according to Equation 2.69. The vacant donor site remains positively charged until another electron is trapped.

$$hv = E_{\mathrm{g}} - E_{\mathrm{ion}} + \frac{1}{2}k_{\mathrm{B}}T. \tag{2.69}$$

Importantly, an ionization energy E_{ion}, must be spent through scattering and/or tunnelling processes to free the electron. Hence, the luminescence has a longer λ than the incident photon.

Band-to-Acceptor Recombination

This type of radiative recombination is similar to the above process, except for bound holes. An excess hole is trapped in a state just above the VB edge, referred to as the *acceptor* level. A free electron inside the CB then recombines with the bound hole and creates a photon with energy according to Equation 2.69. The filled acceptor site will have a net negative charge until another hole is trapped.

Both types of free-to-bound recombination processes exhibit a faster decay, with increasing T, than that of BtB, due to thermal excitations out of the bound states. In other words, the donor (acceptor) states begin to merge with CB (VB) states at sufficiently HT.

Donor-to-Acceptor Recombination

A further case considers the possibility of radiative recombination in materials with significant concentrations of impurities. At LT, trapped electrons recombine with trapped holes and release photons with energy according to Equation 2.70.

$$hv = E_{\mathrm{g}} - (E_{\mathrm{donor}} + E_{\mathrm{acceptor}}) + \frac{e^2}{4\pi\varepsilon_0\varepsilon_r r_{\mathrm{d}\to\mathrm{a}}}. \tag{2.70}$$

Where E_{donor} and E_{acceptor} refer to the donor and acceptor ionization energies, respectively. Coulomb attraction between both types of carrier is taken into account by the last term. Here $r_{\mathrm{d}\to\mathrm{a}}$ refers to the donor-acceptor separation and is not to be confused with a_{X}. All other symbols have their usual meanings.

There are T-dependent electron-hole recombination processes which only generate

phonons instead of photons. It is worth highlighting that the non-radiative recombination channels discussed in the following Subsections can never be completely avoided, since any semiconductor will have at least some native defects. Increasing T causes a drop in I_{PL}, thus degrading the performance of LDs. Typically, non-radiative mechanisms begin to dominate beyond $T \approx 100$ K [172].

Surface Recombination

Dangling bonds (see Figure 2.24) are partially filled electron orbitals which can exist on a semiconductor surface and between layers (at interfaces). These are caused by a discontinuity of the crystal lattice periodicity, i.e. where neighbouring atoms of the same type are missing.

Since valence orbitals are unable to form a chemical bond, they can either reconstruct themselves to form bonds between neighbouring atoms within the same plane or bind to impurities during growth (e.g. carbon). The former creates a new atomic structure with different energy states to that of bulk material, whereas the latter injects new energy states; both of which act as non-radiative recombination centres. The "forbidden" states can either be donor-like, or acceptor-like, and the surface recombination rate depends on the velocity of minority carriers migrating to the surface ($\sim 10^6$ cm·s^{-1} in GaAs [82]).

As a rule of thumb, to minimise non-radiative losses from surface impurities: the distance between the surface and a p-n junction should at minimum be three multiples of the magnitude of the free carrier diffusion length [157]. Moreover, post growth techniques such as surface passivation, e.g. the deposition of an ultrathin InP layer [173], aim to eliminate dangling bonds and reduce impurity concentrations within the material.

Auger Recombination Processes

An Auger recombination process describes one of four types of strongly T-dependent, non-radiative, mechanisms. When an electron and a hole recombine in a BtB transition; the energy is absorbed by a third carrier, rather than emitted as a photon. This excites the electron (hole) higher (deeper) into the CB (VB). Subsequently, the carrier relaxes towards the band extrema by emitting phonons. Each of the processes are shown in Figure 2.32. Notably, it is unlikely for Auger recombination to occur in semiconductors with a large bandgap, due to typically large electron effective masses and therefore large transfers of quasi-momentum. Thus for the Auger event to occur in this case, holes must exist at an unrealistically-deep energy state.

The CHCC process involves a collision between two electrons in the CB. One of them
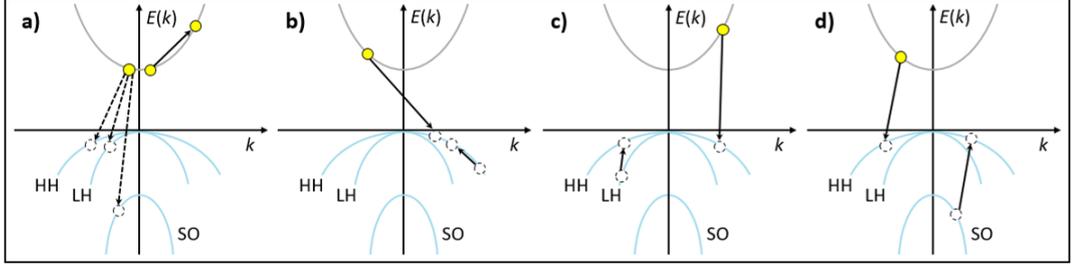
Figure 2.32: Schematic to show the four main non-radiative Auger recombination processes: a) conduction-heavy conduction-conduction (CHCC), b) conduction-heavy heavy-heavy (CHHH), c) conduction-heavy light-heavy (CHLH), and d) conduction-heavy spin-orbit-heavy (CHSH).

recombines with a hole in the VB and the excess energy excites the other electron higher into the CB. However the remaining processes involve one electron and two holes. An electron-hole recombination event either: promotes a HH to a valence state of higher energy (CHHH); excites a hole from the LH band to the HH band (CHLH); or excites a hole in the SO band to the HH band (CHSH). The CHCC process occurs predominantly in semiconductors with an abundance of electrons, while the other processes typically require an abundance of holes.

Limiting the density of final states available to the third party carrier will suppress Auger processes and mitigate inter-valence-band resonances [79, 174]. The CHCC Auger recombination process is most likely to be of significance to this work; since varying the carrier overlap with changes to electron confinement, will change the number of final vacant Auger states that the electrons are able to scatter into. Additionally, CHHH, CHLH and CHSH Auger processes are less likely to have an observable impact, since all three have a p-type semiconductor requirement; moreover, the HH and LH bands in GaSb/GaAs are rather wide, and degenerate close to the band maxima.

The Auger non-radiative recombination rate, R_{Auger}, for the CHCC process is given by:

$$R_{\text{Auger}} = C_n n_e^2 n_h. \tag{2.71}$$

Whereas for the remaining Auger processes:

$$R_{\text{Auger}} = C_p n_e n_h^2. \tag{2.72}$$

The Auger coefficients, $C_{n,p}$ are usually different due to variation between the CB and VB structures. In the high excitation regime, the concentration of non-equilibrium carriers surpass that of the equilibrium carriers; therefore the total Auger

recombination rate is:

$$R_{\text{Auger}} = (C_{\text{p}} + C_{\text{n}})n_{\text{e}}^3 = C_{\text{Auger}}n_{\text{e}}^3. \tag{2.73}$$

Clearly, the luminescence efficiency is only significantly impacted by Auger recombination at large excitation intensities (or carrier injection currents), as indicated by the cubic carrier concentration dependence. Moreover, the Auger recombination rate is strongly T-dependent:

$$R_{\text{Auger}} \propto \text{e}^{(-E_{\text{a}}/k_{\text{B}}T)}T^3. \tag{2.74}$$

Where the exponential term dominates and E_{a} corresponds to a specific Auger process. The activation energy of the CHCC process, $E_{\text{a}}^{\text{CHCC}}$, can be calculated according to Equation 2.75, with E_{g}, Δ_{so}, and the effective masses of the electrons (m_e^*), HHs (m_{hh}^*), and SO holes (m_{so}^*).

$$E_{\text{a}}^{\text{CHCC}} = \frac{m_e^* E_{\text{g}}}{m_e^* + m_{\text{hh}}^*}. \tag{2.75}$$

Shockley-Read-Hall Recombination

There is debate whether SRH or Auger recombination is the main cause of performance degradation of luminescent devices (see [82] vs. [176]), which is further fuelled by uncertainty that arises upon consideration of the exact respective weights of defect-related scattering mechanisms in both types of process.

SRH recombination refers to non-radiative processes due to trapping of free carriers native defects within an active region, e.g. vacancies, interstitials, impurities, and dislocations etc. Specifically, it refers to recombination mid-bandgap via deep trap levels. Although that is not to disregard the capability defects have to act as scattering centres and also cause indirect Auger processes; as described in Ref. [175], where observations of a single InGaN/GaN QW were used to derive a linear dependence between the SRH and bimolecular defect-assisted Auger coefficients.

The SRH recombination rate is defined as [177]:

$$R_{\text{SRH}} = \frac{p_0\Delta n + n_0\Delta p + \Delta n\Delta p}{(n_{\text{T}}\nu_{\text{h}}\sigma_{\text{h}})^{-1}(n_0 + n_1 + \Delta n) + (n_{\text{T}}\nu_{\text{e}}\sigma_{\text{e}})^{-1}(p_0 + p_1 + \Delta p)}. \tag{2.76}$$

Where the trap concentration is denoted by n_{T}; the electron and hole thermal velocities by $\nu_{\text{e,h}}$; the carrier capture cross sections of the traps for the electron and holes by $\sigma_{\text{e,h}}$; and it is assumed that $\Delta n = \Delta p$. The electron and hole concentrations

when the trap level has an energy equal to the Fermi energy, are represented by n_1 and p_1, respectively. They are expressed in terms of the energy difference between the deep trap level (E_T) and the Fermi level of an intrinsic semiconductor (E_F^i):

$$n_1 = n_i e^{\left(E_T - E_F^i\right)/k_B T}, \tag{2.77a}$$

$$p_1 = p_i e^{\left(E_T - E_F^i\right)/k_B T}. \tag{2.77b}$$

Substituting $R_{SRH} = \Delta n, p/\tau$ returns the non-radiative lifetime of either excess carrier. This expression is simplified upon distinguishing what the majority carriers are, and assuming a small deviation away from equilibrium. Consequently, the following relations hold:

	n-type:			p-type:	
$n_0 \gg p_0$,		(2.78a)	$p_0 \gg n_0$,		(2.79a)
$n_1 \gg p_1$,		(2.78b)	$p_1 \gg n_1$,		(2.79b)
$n_0 \gg n_1$,		(2.78c)	$p_0 \gg p_1$,		(2.79c)
$\Delta p \ll n_0$.		(2.78d)	$\Delta n \ll p_0$.		(2.79d)

The minority non-radiative recombination rates for electrons $(\tau_{n_0})^{-1}$ and holes $(\tau_{p_0})^{-1}$ are then:

$$\frac{1}{\tau_{n_0}} = n_T \nu_e \sigma_e, \tag{2.80a}$$

$$\frac{1}{\tau_{p_0}} = n_T \nu_h \sigma_h. \tag{2.80b}$$

Thus proving that the rate of minority carrier capture does indeed limit the SRH recombination rate.

2.3.4 Total Recombination Rate

The amount of light produced as excess carriers return to a state of equilibrium depends on the relative contributions of radiative and non-radiative processes. The total recombination rate τ^{-1} is given by the sum of the inverse radiative (τ_r) and non-radiative (τ_{nr}) lifetimes:

$$\tau^{-1} = \tau_r^{-1} + \tau_{nr}^{-1}. \tag{2.81}$$

Furthermore, the internal quantum efficiency, η_{int}, can be defined as:

$$\eta_{int} = \frac{\tau_r^{-1}}{\tau_r^{-1} + \tau_{nr}^{-1}}. \tag{2.82}$$

The RT internal luminescence efficiencies of the first III-V structures were extremely low (< 1 %). Currently, high-quality bulk and QW semiconductors can have η_{int}-values in excess of 90 %. This improvement is attributed to improved crystal quality, and reduced defect and impurity concentrations [82].

Not all emitted photons are able to escape from within the material due to reabsorption effects and total internal reflections [178]. For optoelectronic devices, η_{ext} is closely related to η_{int} and is expressed by Equation 2.83 as the ratio of the number of emitted photons (N_{ph}) to the number of photons absorbed by the system (N'_{ph}) [179].

$$\eta_{\text{ext}} = \frac{N_{\text{ph}}}{N'_{\text{ph}}}. \tag{2.83}$$

Hence, an external quantum efficiency (EQE) value close to unity (in the absence of an optical cavity) indicates that the light generation is extremely efficient.

Chapter 3

Research Methods

Chapter 3 discusses the theoretical and experimental methods employed in this work. Specifically, the simulations, epitaxial growth, and characterisation of the reference and RWELL samples. A discussion on the h_c calculation is first presented. Following a brief overview of the computational methods used to calculate band structure, the $\boldsymbol{k} \cdot \boldsymbol{p}$ envelope function theory is introduced, along with its implementation in 3D with Nextnano^{++}. This Chapter then focuses on the important ZB parameters (Tables 3.1 and 3.2) required for the 8-band model [211], along with the assumptions made. Bowing is considered and the procedure used to interpolate the parameters is outlined in section 3.1.2. In Subsection 3.2, the MBE system is described in detail, along with the growth techniques and procedure used to form the QRs. Next, the PL characterisation technique is discussed in Subsection 3.3, along with configuration of the implemented system. Uncertainties corresponding to the optical measurements and their analysis are considered, which leads to discussion on how the data was processed. This chapter concludes with details on the physics of AFM and XRD. Although the simulation chapter in this thesis is presented before the experimental chapter, the calculations were actually performed post-experiment due to an unforeseen change of project during Covid-19 lockdown. The new research aim validated the adopted fitting procedure and generated further understanding for the interpretation of results.

During this project, I became a competent and independent operator of the MBE system. I am able to grow a range of monolithic semiconductor structures; including bulk heterostructures, thin films (MQWs and digital alloys), and 0D nanostructures. A relatively new technique to form QRs without the need for a hot upper layer was applied for the first time to InGaAs-capped GaSb/GaAs QRs. I was extremely fortunate to assist with MBE repair and maintenance tasks including: bake-out, calibration, leak detection, crucible, shutter, and ion gauge filament replacement,

mass spectrometer instalment, and part-salvage from an older system. Experimental work beyond the scope of this thesis included forming GaAs templates for selective area growth by nanodrilling, and "V"-trench patterning of Si substrates for III-V/Si integration. Supplementary information on this technology employed at the Lancaster Quantum Technology Centre is included for completeness in Appendix B. Furthermore, I gained a lot of experience in PL spectroscopy and started to build a new µ-PL system (including a: 405 nm laser source, Oxford Instruments He Microstat, iHR320 monochromator, SR830 DSP lock-in amplifier, InSb detector etc.). I became proficient with AFM and XRD techniques and performed all of the measurements and analysis in Chapter 5. I wrote a Python script to estimate h_c for an $In_xGa_{1-x}As$ QW as a function of x. Dr J. Llorens assisted me in simulating the structures of Chapter 4, in a collaboration with the Institute of Micro and Nanotechnology, Madrid. I adapted a Python script that was kindly provided to me in order to estimate the 8-band $In_xGa_{1-x}As$ parameters. These were implemented into Nextnano^{++}, to create 3D simulations of the strain, band structure, wavefunction probabilities, and transition energies of a typical nanostructure (excluding the WL) found in the grown samples of Chapter 5. These results were then compared with experimental data to supplement the conclusions made from a limited data-set.

3.1   Simulations

3.1.1   Critical Thickness Calculation

Equation 3.1 desribes the h_c model used in this work [114,115]. It is deemed a reasonable rule-of-thumb for placing an upper limit on the choice of a stable $In_{0.06}Ga_{0.94}As$ layer thickness for the following reasons. Experimental designs tend to stay well clear of theoretical h_c-values. Maximising the QW thickness is not as important in this work as observing the effects by changing it, hence deeper theoretical consideration of h_c does not aid the objectives of this thesis. Moreover, the isotropic approximation significantly underestimates h_c in the InGaAs system [116], thus providing further assurance of high quality growth.

$$h_c = \frac{|\boldsymbol{b}|(1 - D\cos^2(\alpha_B))(ln(h_c/|\boldsymbol{b}|) + 1)}{4\pi f(a)(1 + D)\cos(\lambda_B)}. \tag{3.1}$$

Where $|\boldsymbol{b}|$ is the slip distance, D is assumed to be along the [001] direction, λ_B is the angle between the Burgers vector and the construction line perpendicular to the glide-surface plane intersection, and α_B is the angle between the Burgers vector and the dislocation line (see Figure 2.26). Low-misfit systems tend to be predominantly

populated by 60 ° dislocations which glide along the $\langle 110 \rangle$ and $\langle 110 \rangle$ directions of the (111) plane. The few edge dislocations that form at the interface are due to reactions of complementary 60 ° dislocations [115], thus $\cos(\lambda_B)$ and $\cos^2(\alpha_B)$ in Equation 3.1 are assumed to be 1/2 and 1/4, respectively. All other symbols have their usual meanings.

A comprehensive review on the historical development of the models used to calculate h_c is presented in Ref. [117]. The concept developed into the most commonly-used approach presented by Matthews and Blakeslee [114]. Various implicit expressions of h_c emerged from this theory, necessitating clarification on the correct form and associated parameter values in the work of Ref. [115]. It is interesting to note that an explicit solution for h_c has since been identified [118]; whereby the equation proposed by Matthews and Blakeslee is reformulated in terms of the Lambert W-function.

3.1.2   Quantum Ring Modelling

A number of methods, ranging from ab initio to the empirical, exist to calculate the electronic structure of crystals. Ab intio methods are calculations from first principles, e.g. density functional theory, or better yet, quasi-particle calculations for the ground-state of materials whilst accounting for electron correlations; and the non-equilibrium Green's function, which describes transport phenomena in nanoscale devices. Conversely, empirical methods use a group of adjustable and experimentally derived parameters to fit known features of the bulk band structure. Literature tends to focus on pseudopotential- or TB- based approaches. Solving by first principles is the most accurate. Nevertheless, improved accuracy comes at the price of poorer efficiency due to requiring extended computation time and resources. Large N-body systems severely impact the algorithmic efficiency, hence a supercomputer (cluster architecture) is often required to reduce run-times to within viable limits. Quantum computers are expected to solve this problem entirely; hence it is highly likely that their establishment will lead to new predictions of subtle physics currently unknown with the most realistic models.

$k \cdot p$ Perturbation Theory

This method is a semi-empirical approximation to the real band structure near the Γ-point. Eigenfunctions are described as linear combinations of multi-band envelope functions multiplied by the corresponding periodic parts of the Bloch functions. It offers a pretty good cost balance with respect to finding the band dispersion and effective masses around the centre of the BZ, hence is favoured in this work; given

that only ground-state recombination needs to be simulated, thus solutions far away from Γ are not of interest.

Notably, the fundamental theoretical contributions to this method are as follows: Development of $\boldsymbol{k} \cdot \boldsymbol{p}$ theory is attributed to the work of Luttinger and Kohn [183], and applied to heterostructures by Bastard [184]. The coupling of the CB and VB was derived by Kane [185]. The inclusion of strain within the Hamiltonian was considered by Bir and Pikus [186, 187]. Moreover, the problem of operator ordering is attributed to Burt [188] and the treatment of spurious modes by Foreman [189].

There are various models which differentiate by the amount of bands considered and how they are coupled. The simplest of these, which encapsulates more than one VB at the Γ-point, is the 4-band model. Bloch functions are assumed to be spin degenerate in this model, i.e. only the CB, HH and LH bands are considered. The SO band is included by taking the SO interaction into account. This leads to the 6-band model, where there are two Bloch functions for each VB corresponding to spin-up and spin-down states. So far treatment has ignored the coupling between the CB and VB; taking this into account leads to the 8-band model. To provide the CB-VB coupling, the Kane interaction term is considered by inserting $P_0 k_{1,2,3}$ in the conduction-valence blocks, $\hat{\mathcal{H}}_{\mathrm{cv,vc}}(\boldsymbol{k})$. Where the momentum matrix element at the band extrema, P_0, is related to the optical momentum matrix element (Kane parameter), E_{p}, by:

$$ P_0 = \sqrt{\frac{\hbar^2}{2m_0} E_{\mathrm{p}}} \tag{3.2} $$

At Γ, the same effective masses are always obtained using either of the 2-, 4-, 6- or 8- band models. Increasing the number of bands improves the description of the band structure at large k-values. Indeed, a simpler model is preferable with respect to efficiency optimisation; however, the type-II nature of the samples under study necessitated the additional complexity incurred by the 8- band model. Moreover, the single band approximation is not applicable to this work, since the effective masses for each band are left uncorrected by strain.

Nextnano[++]

Nextnano[++] software [180] was used to model the electronic properties of the grown samples in 3D. The code (developed by Andlauer [181, 182]), is based upon the multi-band $\boldsymbol{k} \cdot \boldsymbol{p}$ envelope-function method. This software provides an intuitive environment to solve the Schrödinger, Poisson and current equations self-consistently for any arbitrary structure, provided that the required fundamental shapes exist

within its library. The parameters within the Nextnano^{++} database are drafted mostly from Vurgaftman [107] and can be overwritten easily. Calculations are solved in real-space (Cartesian basis), utilising a discrete grid of N nodes. Notably, layers are assumed to have abrupt interfaces.

A number of calculations were made in two separate investigations. The width of an In$_{0.06}$Ga$_{0.94}$As QW was first varied up to 25 nm, in increments of 5 nm. These results were compared to the GaSb/GaAs QR reference calculation. Next, x was varied between $10 \leq x \leq 30$ % for a 10 nm QW, in order to check the feasibility of λ extension using this structure. The calculation sequence implemented here, begins by calculating the strain minimisation and piezoelectric charge density [213]. These charges arise where piezoelectric constants change abruptly, i.e. in an inhomogeneous strain field, or at a heterojunction; due to the shear strain-induced displacement of carriers within the unit cell. Large shear strains occur at the edges of the QR due to geometry nd lattice misfit. Since the structure has interfaces with edges that are much sharper than the grown samples, the total charge density may be over estimated in the simulations. A high strain solver residual accuracy ($< 10^{-12}$) is implemented, and the density of pyroelectric charges is assumed negligible for ZB structures in Nextnano^{++}. Next, the 8-band $\boldsymbol{k} \cdot \boldsymbol{p}$ envelope function is approximated to determine the bandgap and confinement energies of the carrier states. Finally, the program iterates through self-consistent calculations of the Schrödinger and Poisson equations to determine the carrier wavefunctions. Carrier densities are computed from their respective wavefunctions, which are generated with each quantum iteration. The previous value for carrier density is substituted into the Poisson equation, during the next calculation cycle, in order to update the electrostatic potential. The solver needs to compute a reasonable number of states to avoid truncating the quantum density; a problem which often leads to issues with convergence and direction of the electrostatic potential. A selection of 20 electron and 20 hole eigenvalues were enough to maintain the accuracy of the solver, whilst keeping runtimes and computational effort low.

Parameter Choice

The correct choice of parameters is crucial in order to obtain an accurate description of the band structure. The ZB 8×8 Hamiltonian parameters are tabulated in Table 3.1. Namely, the: effective CB mass ($m_e^*[001]$), modified Luttinger parameters (γ_1', γ_2', γ_3', and κ'), $E_g(0 \text{ K})$, offset of the three VB energy average (ΔE_v), Δ_{so}, inversion asymmetry parameter (B_{inv}), a_c, a_v, b and d.

Empirical methods rely critically on correct parameters. The Nextnano^{++} material database contains some errors, hence it was ignored in favour of directly importing

the values generated by the python script. The code interpolates an inventory of binary material data and creates an output file containing the corresponding ternary values. Some parameters were kept the same due to their origins (see Subsection 3.1.2). Importantly, the T-correction for the bandgap is computed with the Varshni parameters. Care must therefore be taken to avoid applying the thermal expansion twice, by using the RT lattice constant and the bandgap at 0 K.

Parameter	GaSb	GaAs	InAs	Ω	$In_{0.06}Ga_{0.94}As$
a [Å]	6.09590[a]	5.65325[a]	6.05830[a]	0[a,b]	5.67755
a_{exp} [µÅ]	47.2[a]	38.8[a]	27.4[a]	0[a,b]	38.1
ε_0	15.69	12.90	15.15	0	13.04
ε_∞	14.44	10.89	12.25	0	10.97
C_{11} [GPa]	88.42[*a]	122.10[*a]	83.29[*a]	15.50[b]	118.90
C_{12} [GPa]	40.26[*a]	56.60[*a]	45.26[*a]	2.20[b]	55.80
C_{44} [GPa]	43.22[*a]	60.0[*a]	39.59[*a]	17.90[b]	57.77
e_{14} [C/m^2]	-0.168[c]	-0.160[c]	-0.044[c]	0[a]	-0.153
$E_g(0\ K)$ [eV]	0.812[a]	1.519[a]	0.417[a]	0.477[a]	1.426
α [meV/K]	0.4170[a]	0.5405[a]	0.2760[a]	0[a]	0.5246
β [K]	140[a]	204[a]	93[a]	0[a]	197
a_c [eV]	-7.50[a]	-7.17[a,b]	-5.08[a,b]	-1.47[b]	-6.96
a_v [eV]	0.80[a]	1.16[a,b]	1.00[a,b]	0[a,b]	1.15
b [eV]	-2.00[a]	-2.00[a,b]	-1.80[a,b]	-0.20[b]	-1.98
d [eV]	-4.70[a]	-4.80[a,b]	-3.60[a,b]	-0.43[b]	-4.60
ΔE_v [eV]	1.777	1.346	1.390	-0.380[a]	1.370
Δ_{so} [eV]	0.760[a]	0.341[a]	0.390[a]	0.150[a]	0.335
γ_1	13.40[a]	6.98[a]	20.00[a]	8.67	7.27
γ_2	4.70[a]	2.06[a]	8.50[a]	3.75	2.23
γ_3	6.00[a]	2.93[a]	9.20[a]	-	3.07
$\Delta\gamma$	1.300	0.870	0.700	0.481[a]	0.833
κ	3.18	1.20	7.68	0[a]	1.59
γ_1'	2.32	0.66	2.81	2.02	0.68
γ_2'	-0.84	-1.10	-0.09	0.43	-1.06
γ_3'	0.46	-0.23	0.61	-	-0.04
$\Delta\gamma'$	1.300	0.870	0.700	-2.844	1.020
κ'	-2.36	-1.96	-0.91	0	-1.90
S [m_0]	-2.2517	-2.8758	-4.7916	1.2840	-3.0632
E_p [eV]	27.00[a]	28.80[a]	21.50[a]	-1.48[a]	28.45

Table 3.1: 8-band $\mathbf{k} \cdot \mathbf{p}$ input parameters and their corresponding bowing values for the $In_{0.06}Ga_{0.94}As$ ternary, at the Γ-point where applicable. The Landé factor, g, and B_{inv} are set to zero. Values with an asterix (*) denote erroneous elastic constants tabulated in Ref. [107], and have been corrected by a factor of 10. T-dependent constants are specified at RT unless otherwise stated. The temperature correction of E_g and a is computed in Nextnano^{++} using the Varshni and lattice expansion parameters. The γ-values are adimensional as Nextnano^{++} input parameters; they effectively behave as effective masses in the Hamiltonian. The 5th column corresponds to the InGaAs Bowing constant (see Subsection 3.1.2). Values with exponents a-c are extracted from Refs. [107, 214, 215] respectively; the rest are either calculated from other parameters (see Table 3.2), or obtained from the Nextnano^{++} database.

Although the B_{inv} parameters are non-zero, they are assumed negligible. In fact, there is somewhat more numerical error corresponding to the uncertainty of this empirical parameter, than the amount of correction it introduces to the Nextnano^{++} output. Furthermore, B_{inv} is much less important for an ensemble due to its tiny energy scale. It is usually relevant when studying single nanostructures, e.g. fine structure splitting of excitons.

The $a_{\mathrm{c,v}}$ variables are based upon measurements of the bandgap with hydrostatic strain. The deformation potential is split into a sum of CB and VB contributions, $a = a_{\mathrm{c}} + a_{\mathrm{v}}$, hence the values tabulated in Ref. [107] are positive. Conventionally [119], Nextnano^{++}, and many authors find it intuitive to reflect the bandgap as a transition energy, hence take the difference of deformation potentials ($a = a_{\mathrm{c}} - a_{\mathrm{v}}$) and list values with the opposite sign. Large $a_{\mathrm{c,v}}$-values correspond to band structures which are strongly affected by strain. The sign of $a_{\mathrm{c,v}}$ indicates how the CB edge shifts under tensile ($\varepsilon > 0$) and compressive ($\varepsilon < 0$) strains.

The isolated CB curvature at the Γ-point is given by the inverse of m_{e}^*. The interaction with all other remote bands is approximated implicitly. In the 8-band model, however, S is not an effective mass, but a fictitious value in the CB elements, $\hat{\mathcal{H}}_{\mathrm{cc}}(\boldsymbol{k}, \hat{\varepsilon}) = E_{\mathrm{c}} + \{\hbar^2/2m_0\}Sk^2 + \mathrm{Tr}(\hat{\varepsilon})$. It provides the correct band curvature from the empirical E_{p}, E_{g}, and Δ_{so} parameters, and is derived by removing the explicit VB contributions expressed in the 8-band Hamiltonian matrix. The 6-band effective mass is removed by the following procedure: the Kane interaction term is introduced and then $1/m_{\mathrm{e}}^*$ is replaced with S, the full Hamiltonian is diagonalised, and the highest eigenvalue is differentiated twice with respect to $\boldsymbol{k}$, finally $|\boldsymbol{k}| = 0$ is substituted to arrive at Equation 3.3.

$$
\begin{aligned}
S &= \frac{m_0}{m_{\mathrm{e}}^*} - \frac{E_{\mathrm{p}}}{3}\left(\frac{2}{E_{\mathrm{g}}} + \frac{1}{E_{\mathrm{g}} + \Delta_{\mathrm{so}}}\right), \\
&= \frac{m_0}{m_{\mathrm{e}}^*} - E_{\mathrm{p}}\frac{E_{\mathrm{g}} + \frac{2}{3}\Delta_{\mathrm{so}}}{E_{\mathrm{g}}(E_{\mathrm{g}} + \Delta_{\mathrm{so}})}.
\end{aligned}
\tag{3.3}
$$

Specifically, the contribution from the HH and LH bands is $2E_{\mathrm{g}}^{-1}$, and from the SO band: $(E_{\mathrm{g}} + \Delta_{\mathrm{so}})^{-1}$. By construction, this curvature is equal to the effective mass of the CB since nothing has been added, rather the matrix representation of the Hamiltonian has been renormalised.

Despite its consistency, the validity of the 8-band curvature makes it quite common for spurious solutions to appear. Specifically, the implementation of inaccurate $\boldsymbol{k} \cdot \boldsymbol{p}$ parameters is one known cause. The spurious wavefunctions are not physically sensible and may strongly oscillate or spike at heterojunctions. Furthermore, their CB

eigenvalues may also reside in the bandgap. There is freedom in fixing the S-value, hence it is possible to remove the spurious solutions in a few ways, however none are a universal fix. Nextnano^{++} grants the option of completely neglecting remote bands, by setting $S = 1$ and then re-scaling all relevant parameters accordingly; starting with E_p:

$$E_\mathrm{p} = \left(\frac{m_0}{m_\mathrm{e}^*} - 1\right) \frac{E_\mathrm{g}(E_\mathrm{g} + \Delta_\mathrm{so})}{E_\mathrm{g} + \frac{2}{3}\Delta_\mathrm{so}} \tag{3.4}$$

Bowing

Since m_e^*, E_p, E_g, and Δ_so all change with composition, S inherits their combined bowing. Some authors assume Δ_so follows Vegard's Law for the InGaAs QW since it was found to cause little difference ($4\,\% - 14\,\%$ depending on T, x, and whether the SO interaction is included) to the calculation accuracy with respect to experimental data [214]. This is because Δ_so at the Γ-point ranges from $0.32 - 0.36$ eV, and $0.37 - 0.41$ eV for GaAs and InAs, respectively [107]; therefore expected to weakly impact the compositional bowing of band energies [214]. It was decided however to use the bowing value of 0.15 stated in Ref. [107] to maintain the highest possible accuracy in aid of the interpretation of experimental results in Chapter 5. Importantly, the interpolation of $\Delta\gamma = \gamma_3 - \gamma_2$ was used to consider the warping of the VB:

- Establish the HH and LH masses (presented in Table 3.2) for each binary in terms of γ_1 and γ_2:

$$\frac{m_0}{m_\mathrm{hh}^*[001]} = \gamma_1 - 2\gamma_2, \tag{3.5a}$$

$$\frac{m_0}{m_\mathrm{lh}^*[001]} = \gamma_1 + 2\gamma_2. \tag{3.5b}$$

- Then compute bowing on both the HH and LH effective masses. The bowing values were taken from [107].

- The correct γ_1 and γ_2-values are recalculated from the updated $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}$ hole masses by rearranging Equations 3.5a and 3.5b.

- Finally, to calculate γ_3: $\Delta\gamma$ is used, along with its corresponding bowing parameter [107] and the updated γ_2-value.

The 8-band γ' variables are derived using the 6-band variables by Equations 3.6a-d. Therefore, they inherit any bowing applied to the 6-band variables. However, care should be taken to ensure they are interpolated correctly. Directly modifying the InGaAs 6-band parameters erroneously calculates the binary γ'-values with E_p and E_g corresponding to $\mathrm{In}_{0.94}\mathrm{Ga}_{0.06}\mathrm{As}$, hence the resulting γ'-values lay outside of the binary ranges. To avoid this, the effective bowing constants were found by

Effective Mass	GaSb	GaAs	InAs	Ω	In$_{0.06}$Ga$_{0.94}$As
$m_{\mathrm{e}}^*[001]$	0.0390[a]	0.0670[a]	0.0260[a]	0.0091[a]	0.0640
$m_{\mathrm{hh}}^*[001]$	0.250	0.3497	0.3333	-0.1450[a]	0.3568
$m_{\mathrm{lh}}^*[001]$	0.044	0.0901	0.0270	0.0202[a]	0.0852
$m_{\mathrm{so}}^*[001]$	0.120	0.172	0.140	0.000[a]	0.170

Table 3.2: Carrier effective masses at the Γ-point for the In$_{0.06}$Ga$_{0.94}$As ternary, in units of $[m_0]$.

rearranging Equation 2.41 for Ω, whilst using the strain-corrected effective masses. These were then modified to quadratically interpolate the binary γ_1' and γ_2'-values. The γ_3'-value was found by modifying the bowing constant of $\Delta\gamma'$ and implementing the approach outlined above.

$$\gamma_1' = \gamma_1 - \frac{E_{\mathrm{p}}}{3E_{\mathrm{g}}}, \tag{3.6a}$$

$$\gamma_2' = \gamma_2 - \frac{E_{\mathrm{p}}}{6E_{\mathrm{g}}}, \tag{3.6b}$$

$$\gamma_3' = \gamma_3 - \frac{E_{\mathrm{p}}}{6E_{\mathrm{g}}}, \tag{3.6c}$$

$$\kappa' = \kappa - \frac{E_{\mathrm{p}}}{6E_{\mathrm{g}}}. \tag{3.6d}$$

Bowing is typically assumed null for the hydrostatic and shear deformation potentials. However, Khomyakov et al. [214] predicted InGaAs ternaries to exhibit a non-negligible upward compositional bowing (a_{v} is still assumed to follow Vegard's law). Elastic constants were also found to deviate from Vergard's law [214], thus affecting the perpendicular strain through D. Accurate band edge energy calculations therefore require quadratic interpolation of these constants; particularly in the case of In-rich InGaAs ternaries. It was deemed necessary to encapsulate such bowing in this work despite growing samples with a weak ternary alloy, since the alloy concentration was increased up to 30 % in further calculations.

Simulated Structure

Two 3D structures are created to model the reference and RWELL samples. These are presented in Figure 3.1. For the reference, a single, flat ring-shaped GaSb nanostructure is encased in a GaAs matrix which extends 50 nm in space from the top and bottom surfaces of the QR, and 50 nm radially from either either edge of the QR. The RWELL samples are modelled similarly to the reference, with the exception that the QR is capped with an In$_x$Ga$_{1-x}$As QW of varied x and width.
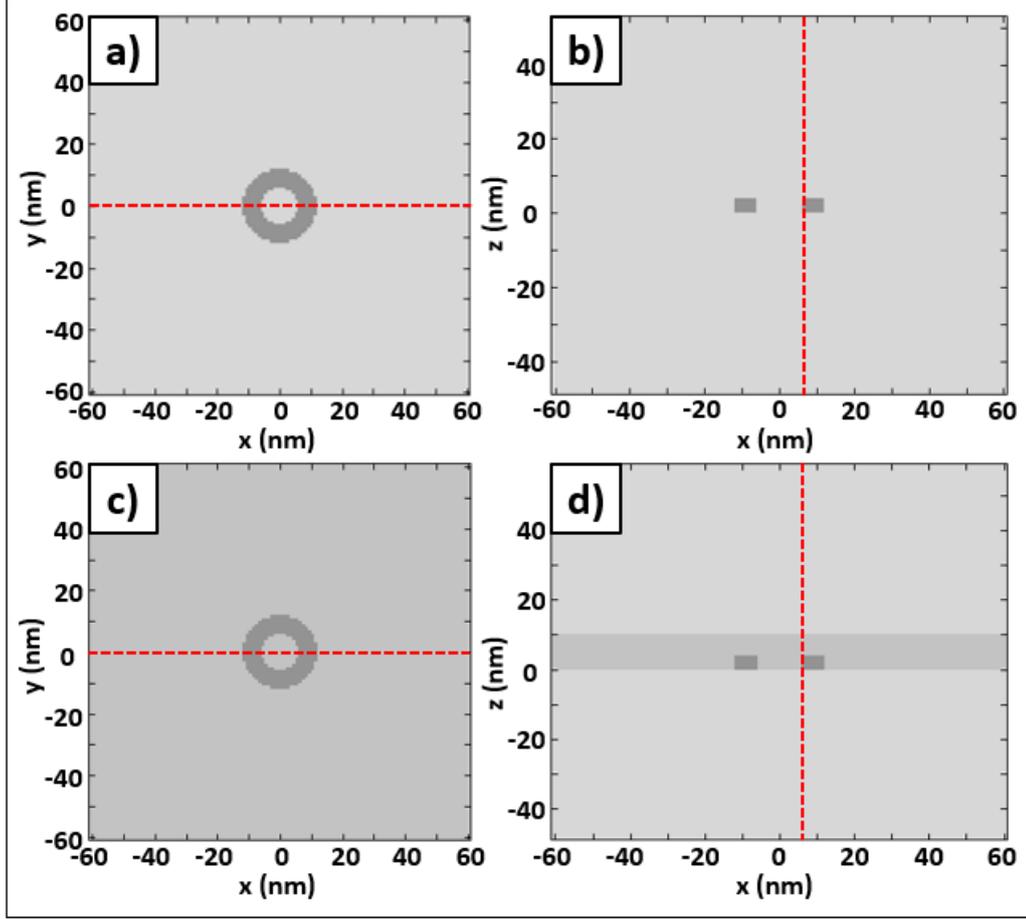
Figure 3.1: 2D cross section plots to show the a) top view, b) side view, of the simulated GaSb/GaAs QR structure; along with the c) top view, and d) side view of the simulated 10 nm $In_xGa_{1-x}As$-capped GaSb/GaAs QR structure. Here, greyscale is used to distinguish between GaAs, $In_xGa_{1-x}As$ and GaSb, where the shading reflects a. The red dotted lines are guides for the eye to show where the 1D-slices are taken.

For these simulations, the GaAs matrix extends 50 nm in the z-direction from the top and bottom surfaces of the QW. The QR is assumed to be 4 nm tall, thus the dissolution effect has been ignored in this work. Neglecting the likely formation of a weak Sb-incorporated quarternary coating around the QR simplifies the QR model. A reasonable QR thickness was assumed from Ref. [47], where 0.5 is suggested for the ratio of thickness to the overall size of the QR. An AFM measurement (Figures 5.9 and Figure 5.8) informed the estimate of the average diameter for the QRs ensemble, thus the inner and outer radii are set at 6 and 12 nm, respectively.

A non-homogeneous grid was applied to balance the computational effort of a larger structure with decent resolution. The line separation was set to 2 nm between the grid boundary and 10 nm above, below, and either side of the nanostructure. The line separation then approaches 1 nm to obtain a higher resolution in the region of interest. In these simulations strain is assumed to have zero relaxation through blurred interfaces. The substrate (not shown) is assumed to be completely relaxed

by using Dirichlet boundary conditions.

3.2 Molecular Beam Epitaxy

Monolithic (bottom-up) growth circumvents the need for transfer and bonding mentioned in Subsection 2.2.3, hence it remains attractive for its simplicity in light of convergence towards cheaper, efficient and complementary metal-oxide semiconductor (CMOS) compatible production. Notably, CMOS compatibility is a widely-debated topic with regards to definition and plausibility. Top-down heterogeneous approaches to fabrication, such as the selective removal of bulk material to create a free standing structure, tend to be costly and the necessary long wet-etch times are likely to damage the epitaxial film.

MBE growth is advantageous over other monolithic approaches to semiconductor fabrication, since it allows for very precise control of composition and thickness. It is also easy to produce abrupt interfaces.

3.2.1 Equipment Setup I
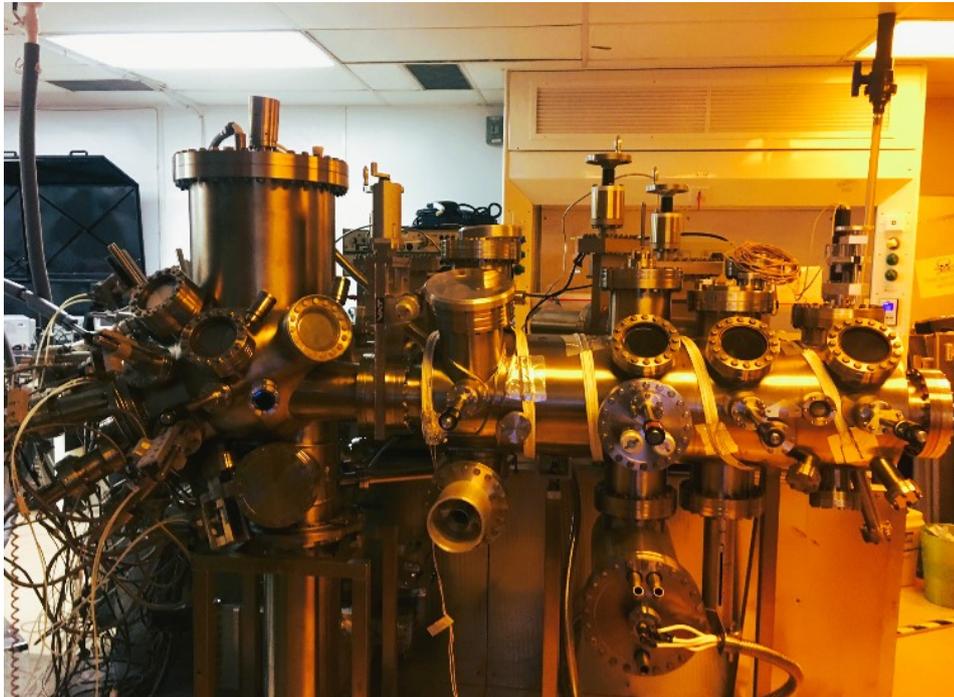
Figure 3.2: Photograph of the MBE system used to grow the samples of this work.

All of the GaSb/GaAs QR samples in this work were grown via solid-source MBE; using the vintage VG-V80H MBE reactor shown in Figure 3.2. Essentially, an MBE system comprises a vacuum subsystem, a deposition subsystem, a computer (and software), and bakeout accessories. The vacuum subsystem incorporates valves, an

air compressor, router, turbo, two ion pumps, and a cryo-pump. During growth, the current vacuum system can maintain pressures down to 10^{-11} mbar through the series of pumping technologies and the use of liquid nitrogen (LN_2). The deposition sub-system can be thought of as having three main sections: a fast entry lock, a preparation chamber and a growth chamber (see Figure 3.3). LN_2 is supplied to the header tank (growth chamber) from a dewar. The substrate is held in a rotating manipulator inside the growth chamber under ultra-high vacuum. There is also an area specifically dedicated to sample outgassing that is furthest away from the growth chamber. The bakeout accessories include heater lamps and stainless steel panels which enclose the MBE system in an insulating box.
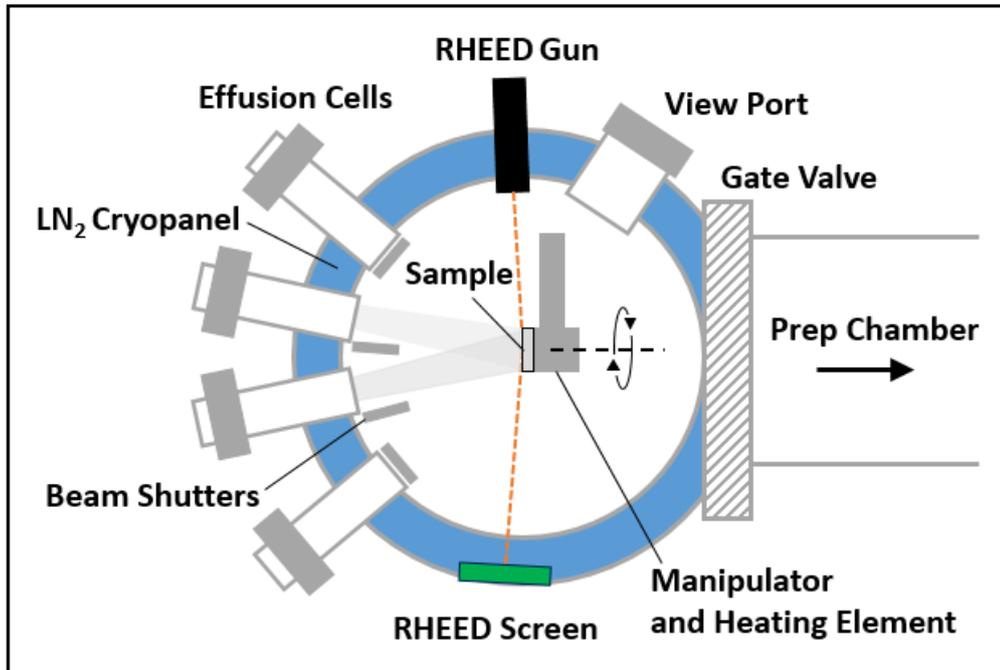
Figure 3.3: Schematic diagram of the simplified MBE growth chamber (not to scale). Here, the Ga and As cell shutters are both opened to grow GaAs. Growth is then terminated by closing the Ga shutter.

The growth chamber is loaded with Knudsen cells for each available group-III source (In, Ga etc.). Cracker cells are also equipped for the As and Sb sources; these are a special type of cell with two different T-zones. The lower-T zone produces a tetrameric As_4 or Sb_4 vapour. The higher-T cracking zone is then used to break the tetramers into As_2 or Sb_2 dimers. These dimers react more aggressively than their corresponding tetramers,, and therefore increase reaction rates.

The cells are directed towards the sample surface as shown in Figure 3.3. Each cell contains a pyrolytic boron nitride crucible to store the source material. The crucibles are heated by a tungsten filament to induce sublimation (vaporisation in the case of Ga), and build up a T- and material- dependent pressure. Once the

shutter is opened, a homogeneous beam of gaseous molecules ejects from the cell nozzle and reaches the surface of the substrate. Thus layers of semiconductor are deposited over time, via adsorption.

We are able to automate our growths by remotely actuating the shutters with a pre-written recipe and the use of pneumatically driven controllers (stepper motors for the cracker cells). The cell temperatures can be ramped and are monitored by thermocouples. A heater allows fine control of the substrate temperature, T_{sub}, inside the manipulator. Ion gauges are fitted to the preparation and growth chambers to monitor the background pressure and the beam equivalent pressure (BEP), respectively.

The BEP is an approximate measure of the local vapour flux, that would strike the sample surface if the manipulator was set to the growth position. BEP measurements establish a working range for the cell and gather quantitative data for estimating deposition rate. The measurements should start and finish with the group-III and group-V cells, respectively. This is to minimise the time it takes to stabilise the BEP reading; since group-III beams have a smaller cross-section than group-V beams, which tend to disperse as high pressure vapors. The reading may take up to a couple of minutes to settle. This depends on the size of the system, pumping capacity, and how efficient the cryopanel is at cooling the system and condensing contaminants on the chamber walls. The Al cell should always be done last, to prevent heating neighbouring cells, as it requires the hottest temperatures in excess of 1100 °C.

During system calibration, the manipulator is rotated to the BEP position to protect it, then the shutters (and valve if present) are opened. The monitoring ion gauge filament is in direct contact with the beam and is bombarded with particles. After recording the base pressures, the shutters are closed again, allowing the vacuum to recover. The background BEP is then noted down and subtracted from the open-shutter BEP to obtain the net flux for that specific temperature.

An MBE system is often equipped with other in situ monitoring equipment such as a mass spectrometer. Arguably, reflection high-energy electron diffraction (RHEED) is one of the most important tools during MBE growth as it allows the operator to observe layer growth in real-time.

Reflection High Energy Electron Diffraction

The RHEED equipment comprises of an electron gun as the electron beam source, and a phosphor screen to detect a visible diffraction pattern. Both components are positioned at a shallow angle at either side of the sample surface.

The amount by which the electron beam diffracts clearly depends on surface rough-ness. The spot intensity changes periodically as each layer is completed; leading to oscillations which are used to measure growth rate. A complete oscillation is observed for the deposition of a full ML. The RHEED will usually display an intense streaky pattern for a smooth layer, whereas the surface will be roughest around 0.5 MLs and provide a weaker spotty pattern due to scattering. Amorphous materials (e.g. oxide layers) prevent the observation of any diffraction pattern. To determine the growth rate in [ML/s], the number of oscillations are counted and then averaged over the time taken for them to dampen completely. The overall oscillation intensity drops with increasing number of MLs because the electron beam was focused on the original surface. Approximately 30 oscillations is typically considered good. This depends however, on the V:III flux ratio, and smoothness of the initial surface. Notably, the period of subsequent oscillations may be unequal; differences indicate perturbations to growth rate caused by shutter transients, which significantly impacts the nominal thickness of thin layers.

The RHEED pattern is an indication of the surface configuration and growth kinetics. The spacing between the lines is proportional to the in-plane reciprocal lattice constant. The number of lines in the pattern and their separation changes during lattice reconstructions. A reconstruction is analogous to a unit cell insofar that it recreates the entire surface when repeated. GaAs(001) is used as an example below to describe the reconstruction process. The reason why III-V surfaces reconstruct is usually due to lattice termination at a surface or the volatile nature of the group-V species when varying temperature conditions. Above the As-terminated surface layer, there are no Ga atoms for the As atoms to bind to. Incomplete electronic shells cause the surface As atoms to bind to each other and dimerise, thus double the space is occupied at each lattice point. An atomic species always dimerises in the same direction, since motion is restricted by the other type of atom as the surface atoms try move together, i.e. $\langle\bar{1}10\rangle$ and $\langle110\rangle$ for As and Ga, respectively [190]. However it is not energetically favourable for them all to stay aligned in the family of $\langle110\rangle$ directions due to coulomb repulsion, hence the (2×1) reconstruction does not form. Instead, they tend to form periodic pairs and leave behind double-dimer spaced gaps in a (2×4) reconstruction, where the $\langle110\rangle$ and $\langle\bar{1}10\rangle$ azimuths are 2x and 4x, respectively. Each azimuth corresponds to the movement of atoms in the perpendicular plane due to the orientation of the sample with respect to the electron beam.

Any amorphous As that condensed on the GaAs surface begins to sublime when heated to $\sim$ 300 °C which results in a c(4×4) reconstruction. The As-As dimers begin to sublime around $\sim$ 400 °C, i.e. the added heat (with no added flux) frees

the back-bonded As dimers from the As lattice sites, which results with the (2×4) reconstruction. Further heating with no added flux causes As-Ga bonds to break which leaves behind a considerably rough surface, until enough heat is supplied for the upper surface to fully thermally decompose into Ga and As droplets. Thus it is important to supply an abundance of As flux. This overpressure must compensate the loss of As from the surface, especially beyond 400 °C [191].

3.2.2 Growth Techniques

Discussion below focuses on how the apparatus was used to create the samples. There are different techniques available that lead to QR formation; e.g. precise control of the flux ratio between Sb and background-As during post-growth Sb soaking [81], and annealing during a growth interruption post-deposition of a thin cap [47, 192, 193]. The latter technique is employed in this work, based upon the cold cap technique [194] previously developed by Kamarudin et al. [195]. To the best of our knowledge, this has not been used to form InGaAs-capped GaSb/GaAs QRs before.

During capping, the local strain inside the GaSb QDs is initially enhanced as the nanostructures are surrounded with material(s) of reduced lattice constant. The goal is to then trigger dissolution and intermixing effects by annealing, in order to form QRs. This technique allowed us to take full advantage of the positive and "negative" effects of annealing, especially since the extent of blueshift resulting from it (see Subsection 5.1.1) was found to be negligible.

Traditionally, the evolution of GaSb QDs to QRs happens with very little intermixing, when nanostructures are cold-capped with GaAs, and a hotter GaAs cap layer is deposited directly above [57, 194]. Deposition of the second *hot* cap layer is necessary to give sufficient energy to allow the relaxation of accrued strain from the initial cap layer, via mass transport, i.e. lateral Sb diffusion, As/Sb exchange, and Sb segregation [194]. Diffusion of Sb occurs within regions of large strain, the magnitude of which is maximal at the QD's peak and centre [86, 132], so the Sb will be redistributed toward the GaAs/GaSb boundary, i.e. they erupt spewing Sb into the local GaAs matrix via As-Sb exchange, analogous to a nanoscopic volcano. This leads to a QR-like morphology, where the final capping (or annealing) temperature influences their inner and outer radii. Smaller uncapped QDs tend to either remain as QDs, often decorated with TDs starting from their top due relaxation of local strain; or form clusters during the capping process [132]. It is reasonable to presume that the same formation process occurs for GaSb/GaAs QRs capped with InGaAs, especially since the In concentration used here is low.

In order to avoid the destruction of InGaAs surface and protect the nanostructures whilst ramping to HTs, the InGaAs QW was capped by an ultra-thin layer of GaAs with a $\sim$ 2 ML nominal thickness. Surface states are not expected to interfere strongly with the characteristics of the dominant recombination channel, due to the relative band offsets. It may however, either incorporate some segregated In and extend the effective QW width, or provide new radiative channels of weak, if not negligible, I_{PL}.

Furthermore, to preserve the physical and optical properties of nanostructures whilst exposed to atmospheric conditions; a layer of semiconductor with a relatively larger bulk bandgap is often passivated on top of the final layer. This coating is selected taking bandgap alignments into consideration.

3.2.3   Growth Method

QW thicknesses were varied according to Table 3.3. The growth recipes followed in this work are described by the growth protocols and tabulated growth parameters presented in Appendix C. A reference sample was grown under similar conditions to the other RWELL samples. For this structure, the $In_{0.06}Ga_{0.94}As$ layer was replaced by a 5 nm GaAs cap at the same growth temperature used for the QWs. Post-growth, the samples showed a mirror-like surface which indicated a high crystalline quality.

| Sample ID	InGaAs QW Thickness (nm)	Cold Cap Thickness (nm)	Cold Cap Temperature (°C)	Annealing Temperature (°C)
A1447	0 (Reference)	5	475	670
A1458	5	5	475	670
A1459	10	5	475	670
A1464	15	5	475	670

Table 3.3: Layer thicknesses and temperatures used.

Importantly, the T_{sub}-values quoted throughout this thesis are the setpoint temperatures. Discrepancies of T_{sub} between samples are difficult to comment on, due to losing the temperature conversion chart amongst other important data from the lab PC. However, the quoted setpoint temperatures in the $620 - 700$ °C range will be approximately (100 ± 10) °C greater than the real T_{sub}-values for three reasons; the first being that the relationship between the reading and the actual T_{sub} is linear in this region [196], the second being that the samples started to deoxidise around this temperature, the third is that the intended temperature during GaAs growth was 580 °C and high quality growth was achieved. In earlier symmetric RWELL growths

(see Tables C.5 and C.6), T_{sub} was increased incrementally by 20 °C. The optimum growth temperature was inferred from the smallest Γ_{PL} and brightest I_{PL}, since the dislocation density will be minimised. Poor growth was observed at both limits of the investigated T-range, hence a temperature of (570 ± 10) °C corresponds to the optimum InGaAs growth window. Therefore the setpoint T-range of $500 - 600$ °C will be approximately (50 ± 10) °C greater than T_{sub}. For temperatures between $450 - 495$ °C there is an estimated (40 ± 20) °C difference as the intended cap growth will have been at, or very close to 430 °C [57].

The specific QW temperature and growth interruption time during annealing was based on the Γ_{PL} and I_{PL} results of an 8 nm $In_{0.2}Ga_{0.8}As$/GaAs MQW [197], where the authors concluded that the narrow T-range of $520 - 540$ °C was optimum. Furthermore, high-quality GaAs growth is typically achieved at 580 °C [45, 198]. Temperatures lower than this can result in poor quality material, which causes an increase in background carrier concentration, i.e. from C-doping [198]. The annealing temperature was based on the work of Ref. [33].
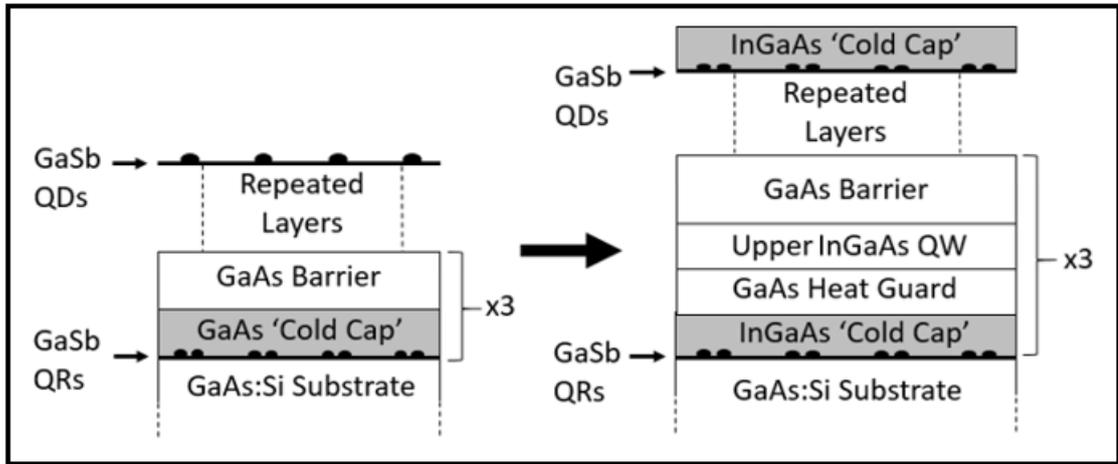
Figure 3.4: Schematic to show the epi-layer design of the samples grown in this study. There are 3 repeated stacks of GaSb QRs embedded at the bottom of an $In_{0.06}Ga_{0.94}As$ QW. Each stack was separated by a 250 nm-thick GaAs barrier grown under the same conditions as the buffer layer. A final layer of half-capped QDs was then grown to allow the typical QR morphology for each sample to be characterised. The dimensions are not to scale.

Quarters were first cleaved from a 2 " n-type GaAs(001) wafer and mounted on separate sample holders (free from In-bonding) and then stored in the epi-rack. A substrate was then selected and transferred to the manipulator. The substrate was annealed in situ at a temperature around 700 °C to remove any native oxide. After deoxidation, a GaAs buffer layer was grown at a temperature of 670 °C and growth rate of 1.0 ML/s, followed by three repeated stacks of GaSb QRs embedded at the bottom of an $In_{0.06}Ga_{0.94}As$ QW. Each stack was separated by a 250 nm-thick GaAs barrier grown under the same conditions as the buffer to prevent any

coupling effects. The reference sample was finished with a layer of uncapped QDs to probe their morphology, whereas the surface QDs were buried under 5 nm of $In_{0.06}Ga_{0.94}As$ in the other samples to observe the effects of annealing part-way through QW deposition.

The active region was grown by the following procedure: First, the growth temperature was reduced to a temperature of 540 °C under a flux of As_2. The As-valve was then shut (over 1 minute) for a 5 minute growth interruption to purge the remaining background. The GaAs layer was then soaked with Sb for 30 s. Next, QDs were formed by depositing 2.1 MLs of GaSb, at a growth rate of 0.30 ML/s. The QDs were cold capped with an $In_{0.06}Ga_{0.94}As$ QW at a temperature of 475 °C, growth rate of 0.32 ML/s, and V/III BEP ratio close to 5 (As-flux $\approx 3 \times 10^{-7}$ mbar) for structural preservation of QRs whilst maintaining high quality (In)GaAs growth [57]. The substrate was then heated to 670 °C under As_2 flux and left for 30 s to induce Sb-As exchange and form the QRs. The As-valve was partially opened, whereas the Sb-valve was opened fully. It is important to protect the InGaAs at hotter T, to avoid evaporating the In away, so a thin layer of GaAs with nominal thickness of ~ 2 ML was used to cap the QW. After annealing, the temperature was reduced back to 540 °C in order to grow the remainder of the QW.

3.3 Photoluminescence Spectroscopy

PL spectroscopy is a characterisation technique that probes the optical properties of materials. It utilises a laser to optically pump a sample by inducing photo-absorption, and a spectrometer-detector combination to record the response. It can be thought of as a way to determine the emission *fingerprint* of a material; since the resulting spectra are a means of characterising energy level differences, impurity/defect levels, Γ_{PL}, non-radiative mechanisms and material quality. Importantly, the incident photon energy of the laser source must be greater than the semiconductor bandgap to excite the electron non-resonantly out of the VB, and well above the bandgap into the continuum of states within the CB. PL is typically observed at (or near) the band edges.

This technique is theoretically non-destructive and allows for an arbitrary number of relatively quick measurements. With the configuration used here however, there is the pragmatic caveat that appropriate care must be taken to avoid accidentally scratching the sample surface with the tip of the optical fibres. The spectrometer grating and bandgap of the detector material must also be appropriately selected according to the expected PL signal.

3.3.1 Equipment Setup II
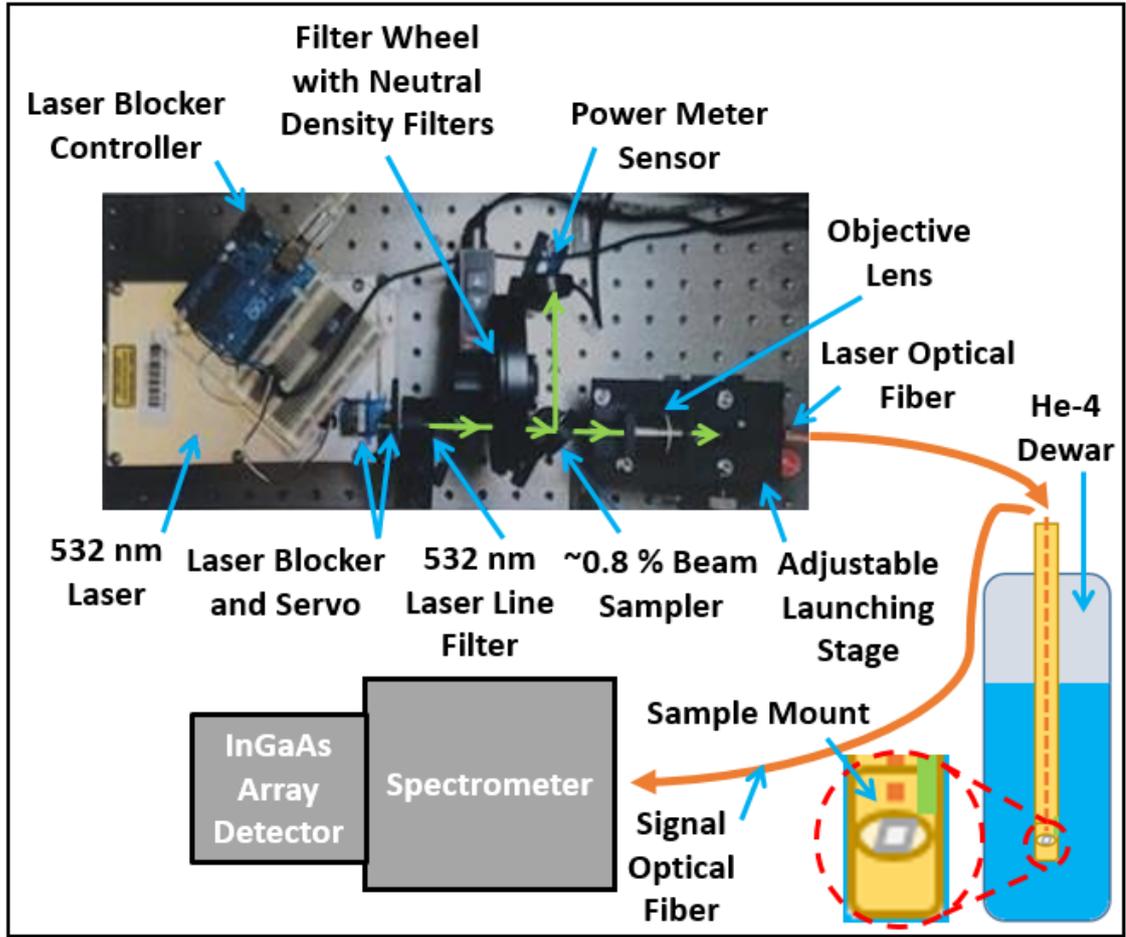
Figure 3.5: Schematic diagram of the PL system used to measure the samples of this work, complete with a photograph of the optical bench configuration. The green lines with arrows represent the lasing path (in air). The magnification(red dashed circle) shows a cross-section of the copper screw mount at the probe terminal, sample (grey rhombus) and the position of the temperature sensor (green bar). Figure adapted with permission from [199].

PL measurements were performed at varied T and P; using stabilised, CW emission from a diode-pumped, frequency-doubled, solid-state laser operating at 532 nm (visible green). This laser is capable of exciting electrons well beyond the (In)GaAs bandgap.

A 200 µm-thick silica core multi-mode optical fibre was used to transmit light from the laser source to the sample. When the laser is switched on and the shutter is open, a ~ 2 mm^2 spot illuminates the sample, corresponding to a surface intensity of $I_s = 7.96$ W/cm^{-2}. The infrared PL (and some of the internally-reflected laser beam) then exits the dewar via the 550 µm-thick silica core optical fibre. Both optical fibres are threaded in parallel through the centre of a long, hollow, copper rod with alignment normal to the copper screw mount located at the probe terminal, near the Cernox temperature sensor (LakeShore: CX-1050-SD-HT).

The sample emission was diffracted using an Acton SpectraPro-2300i spectrometer with a 30 cm focal-length, 150 gr/mm grating, and a 1.2 μm blaze wavelength. Finally, a Peltier-cooled iDus InGaAs DU492A 1.7 charge-coupled device (CCD) was cooled to −60 °C and used to detect the PL.

The rest of the PL hardware was controlled by a Labview script running on the connected PC. This program controls the current supplied to the laser and thereby modulates P and enables automated sweeps. A motorised filter wheel is installed to provide further functionality, however the zero-filter setting was selected for all measurements and P was attenuated by decreasing the supply of current to the laser. In order to measure the laser power at the sample surface, ~ 0.8 % of the laser beam is redirected by a beam sampler to a power meter [199].

InGaAs Charge-Coupled-Device Detector

InGaAs detectors are sensitive within the range of 800 nm to 1700 nm. The focal-plane array of the area scan CCD used here is composed of a dense 2D array of tiny photocells (pixels) at the focal plane of the lens. Specifically, an InGaAs absorption layer is deposited onto an InP substrate and capped with an ultra-thin InP layer. An anti-reflective coating applied to the top surface maximises photon throughput. This material is then In-bump-bonded (flip-chip) to a readout IC, which is connected to the detector preamplifier [200].

The focal-plane array senses incident photons and collects the generated charges. Charge intensity at a particular pixel corresponds to the amount of light captured by it. The integration time is the length of time that the shutter is open to expose the photocells and count incoming photons. Once the shutter is closed, the electrons are shifted to an output node by the readout IC, where they are amplified and converted to voltage. Notably, the charge accumulation for a CCD sensor is different to a CMOS sensor. Whereby the analogue-to-digital converter (ADC) is situated external to the focal-plane array, as opposed to being incorporated within each pixel of a typical CMOS image sensor. The function of an ADC is to convert time and amplitude signals from analogue (continuous) to discrete (digital). Transmission of digital signals occur via data lines, where the number of parallel lines per pixel is referred to as *pixel depth*. Data lines are regulated by the pixel clock, i.e. a high frequency pulse train able to discern valid information.

The Labview program can automatically control the integration time by setting an appropriate PL magnitude scaling factor. It is important to carefully select the integration time to obtain a high quality measurement. Photocells can be thought of as an array of closely packed buckets, which fill up with electrons. The length of

time that the buckets are filled is analogous to the integration time. Overexposure will flood the pixels and cause them to leak, akin to how buckets overflow into their nearest neighbours. This saturates the brightest wavelengths at peak maxima, where the most photocurrent is generated. If the integration time is set too low, then the spectra will be noisy.

Spectrometer

A key component of the spectrometer is the monochromator. This is essentially a special diffraction grating that is optimised to the λ of incident light. The light is split into constituent λ-components and then directed to the detector. Dispersion, i.e. the magnitude of λ splitting, is mainly dependent on the groove spacing d_{gr}. However, angle of incidence, θ_i, and the order of diffraction, n_d, also affect it. The λ of reflected light is related to the angle of diffraction, θ_d by Equation 3.7. Both angles are with respect to the axis perpendicular to the grating.

$$n_d\lambda = d_{gr}(\sin\theta_i + \sin\theta_d). \tag{3.7}$$

The grating has a saw-tooth profile to maximise the grating efficiency for a desired n_d-value and minimise the other orders. The grating efficiency is a measure of the diffracted light intensity and is heavily dependent on the blaze angle, θ_{bl}. This is the angle between the plane parallel to the grating surface and the sloped edges. Moreover, the diffraction power of a ruled blazed grating depends on the polarisation of the incident photons. TM (S-plane), and TE (P-plane) polarised light are orientated perpendicular, and parallel, to the grating grooves, respectively.

3.3.2 Photoluminescence Method

A 1×1 cm^2 piece of sample is cleaved and fixed to the screw mount with electrodag. The probe is then manually lowered (raised) into (out of) a dewar containing liquid helium-4 (LHe-4), to allow for T-measurements from ~ 4 K to RT. Helium boil-off escapes through the dewar exhaust valve and is routed to the external liquefaction room for recycling.

Precise temperature control is mostly dependent on patience whilst adjusting the height of the probe above the LHe-4 surface; it is important to do so very slowly (at least 30 minutes per sweep as a rule of thumb). Not only does this grant better measurement selectivity for data analysis, but it avoids excess LHe-4 boil-off (saving money), and ensures better stability of the temperature reading as thermal equilibrium is reached.

3.3.3   Random Uncertainty Sources

Sample Heating

When the laser is operating in CW mode and illuminates the sample for a significant length of time, it starts to heat the sample surface. The rate of heating is increased if the sample is not completely submerged in LHe-4. The resulting lattice expansion will result in a slightly different bandgap to what the specified temperature corresponds to (Subsection 2.1.2).

Rate of Submersion

The rate that the probe was lowered will be different for each sweep. When the probe is lowered too quickly, there is a small thermal lag between the real and recorded temperatures as thermal equilibrium is reached. This further invalidates the temperature readings (Subsection 3.3.4). Hence rushed measurements make it slightly more difficult to compare T-sweeps between samples.

Detector

The level of noise at the output of the InGaAs detector is limited by an aggregate of several noise sources (Equation 3.10). The read noise, $\sigma_{\rm r}$, is associated with generating electronic signals at the detector output whilst the shutter is closed (zero integration time); i.e. it comprises both random (Gaussian-distributed) fluctuations of the analogue signals within the pixels, and (uniformly-distributed) rounding errors after the signals have been quantized by the ADC. Since $\sigma_{\rm r}$ specifies the minimum number of detectable signal electrons [electrons RMS], it is proportional to the image contrast resolution. Thus $\sigma_{\rm r}$ limits discernible changes in signal amplitude, hence a low value is necessary to detect weakly emitting nanostructures. Notably, $\sigma_{\rm r}$ does not depend on either signal level or sensor temperature. However, it does depend on the pixel clock rate. This is the speed of pixel transmission necessary to refresh a full frame of pixels within a single cycle.

Fixed pattern noise, $\sigma_{\rm f}$, is caused by spatial non-uniformity of pixels. Defects occur due to fabrication difficulties associated with intricate sensor designs. This type of noise is independent of both T and signal level, and may be assumed negligible as is customary in the application of scientific-grade cameras.

Shot noise originates from the discrete nature of electrons and can be distinguished into two primary sources. Photon shot noise is caused by the sporadic arrival of photons at each pixel, whereas dark shot noise arises due to spontaneous thermal excitations of valence electrons into the CB. This type of noise depends on power

(electron density) and obeys Poisson statistics.

Photon shot noise, σ_s, depends on the signal level and is T-independent. Since the standard deviation associated with a Poisson distribution is proportional to the square root of its mean value, σ_s is given by:

$$\sigma_s = \sqrt{\eta_d N t}. \tag{3.8}$$

Where N is the average number of independent, random absorption events per second, and t is the duration of exposure. The quantum efficiency, η_d, is redefined here in the context of solar cells; as the ratio of the number of electrons generated inside a pixel, to the number of incident photons.

Conversely, thermal shot noise can manifest in the absence of light, hence is independent of the signal level. The accumulation of charge in a pixel whilst the shutter remains closed is referred to as *dark current* (I_d). The variation in the number of dark electrons during acquisition is the dark shot noise, σ_d expressed by Equation 3.9. Dark current depends trivially on T, integration time, and sensor quality. Hence, thermal shot noise dominates σ_r at elevated temperatures.

$$\sigma_d = \sqrt{I_d t}. \tag{3.9}$$

Dark current in the CCD can be reduced significantly with thermoelectric cooling. This is important for InGaAs arrays, since the small bandgap produces a relatively large thermally-generated signal which weakens the signal-to-noise ratio. Cooling also reduces the effects of *hot pixels*. This noise source is accounted for in the Labview code by automatically subtracting the dark acquisition from the bright signal during each measurement cycle, i.e. when the shutter is closed and reopened.

The total effective noise per pixel σ_{eff} can be estimated by Equation 3.10, as the quadrature sum of individual noise sources.

$$\sigma_{eff} = \sqrt{\sigma_r^2 + (\sigma_f^2 \sim 0) + \sigma_s^2 + \sigma_d^2}. \tag{3.10}$$

A system is limited by shot noise when capturing intense light under bright conditions; whereas it is light-limited when saturation has not been reached, yet the sensor exposure cannot be increased any further.

Analytical Error

Some random error related to the determination of the PL peak energy, E_{PL}, occurs

due to an irregular profile (that is likely influenced by background noise), and low resolution of data on some measurements. This will have a subsequent impact on finding the FWHM.

Background intensity changes with every slight adjustment to the fibre orientation. Moreover, there are many contributing emitters, $2 - 3$ of which appear dominant depending on the angle of surface illumination. Therefore the only semi-reliable information obtainable from the background sweeps are the peak positions, and all must be kept in mind during analysis. The QR emission is assumed to dominate background below 300 K in the λ-interval that spans the fitted Gaussian baseline. This is because it is only likely to significantly affect I_{PL}, and the ratio of intensities is used. All the samples are affected by background in a similar way and by similar amounts.

3.3.4   Systematic Uncertainty Sources

Dewar Contamination

When the rod is lowered too quickly, the inrush of air causes moisture to freeze on the probe and sample. This contaminates the dewar with ice crystals (see Subsection 3.3.4). Furthermore, accidentally pushing the optical fibres against the sample surface causes the accumulation of optically-active contaminants from other experiments inside the dewar, e.g. loose nanowires. The likelihood of having a large enough concentration of them to make a difference is extremely small.

Temperature Sensor

The temperature sensor is not in direct contact with the sample. It is fitted to the copper probe, as shown in Figure 3.5. This results in a slight temperature gradient which prevents the true temperature from being recorded. Taking the difference between the known temperature of LHe-4 and one of the most inaccurate measurements into consideration, the maximum uncertainty in a temperature reading is estimated to be ± 0.5 K. However these temperature differences do not significantly impact the analysis of data. Specifically, the I_{PL} results are normalised; the thermal corrections of the peak energies are very small; and trivially, the P sweeps are all at 4 K. This sensor has minimal field-induced temperature errors, hence this particular source of uncertainty is assumed negligible.

Moisture

Glass optical fibres are often unintentionally doped with metallic ions and water molecules; the latter being present in the air contained within the laser cavity too.

These impurities directly influence the PL spectra by absorbing the collected light. A clear dip was observed in all measured PL spectra at ~ 1.09 eV, which made peak decomposition and fitting difficult due to the unfortunate overlap with E_{PL} for the QR peak of the samples in this work. The OH− ion has an associated absorption peak energy in remarkable agreement at E_{PL} = 1.097 eV [201]. Therefore, the dominant cause of signal loss is attributed to the trapping of water vapour inside the optical fibre.

Detector and Spectrometer Efficiencies

Detectors are unable to convert photons to current at 100 % efficiency. Gratings have the potential to be very efficient; however they are unable to ideally reflect light due to their coating and the manufacturing process which produces slightly malformed grooves, along with irregularities in the spacing between them. Both of these components are λ-dependent and need to be compensated for.

There are two further sources of systematic error related to spectrometer calibration. The slit alignment is often adjusted between different experiments. Changing the angle of the incident light ray entering the spectrometer shifts the measured spectral position by a small amount. The second uncertainty is due to the grating being one of three that are mounted on a turret. The turret automatically rotates to change groove density. There is an uncertainty associated with the capability of the spectrometer to realign it back to the exact same angle.

The spectral response of the coated mirrors inside the spectrometer is likely to be almost flat for the λ-domain of interest at least 95 %, so any optical losses from the mirrors may be assumed negligible. This assumption relies on a diligently built system, where the apparatus is configured to suit experimental needs. To elaborate, the mirror coating must have been chosen with a suitable reflectivity window in mind that corresponds to the λ-range of interest. The mirrors and grating must also be well-positioned so that light is reflected onto the grating at an angle of θ_{bl} away from the grating normal, i.e. perpendicular to the surface.

3.3.5 Data Processing

Ignoring the detector and grating efficiencies has a significant impact on the measured I_{PL}. In some cases, this may even lead to a spectral lineshape containing artefacts, which influence conclusions on the sample structure; particularly when decomposing ambiguous peaks. Below is a description of the method used to improve the accuracy of the measured dataset. Surprisingly, despite how this important issue significantly affects the reliability of results, it is nearly always absent from the

analytical discussion of typical PL measurements in literature; prompting the question of how often it goes unaddressed, rather than simply considered superfluous detail for publication.

In order to process and analyse the PL spectrum, it is crucial to convert the raw data into a workable format. All X-domain data points were first converted from λ [nm] to E [eV]. Similarly, the Y-domain data points were converted from counts, to counts per second, by using the integration time recorded on each measurement output.

PL Normalisation

An online data reader tool was used to obtain estimates of published efficiency ratings for the detector [202, 203] and grating [204]. A mean average of the -25 °C and -90 °C curves is taken to approximate the detector efficiency at -60 °C. The grating efficiency is at RT. It is inferred from the use of an optical fibre that the light is unpolarised, hence a mean average of the S- and P- plane grating efficiencies is used [204, 205].
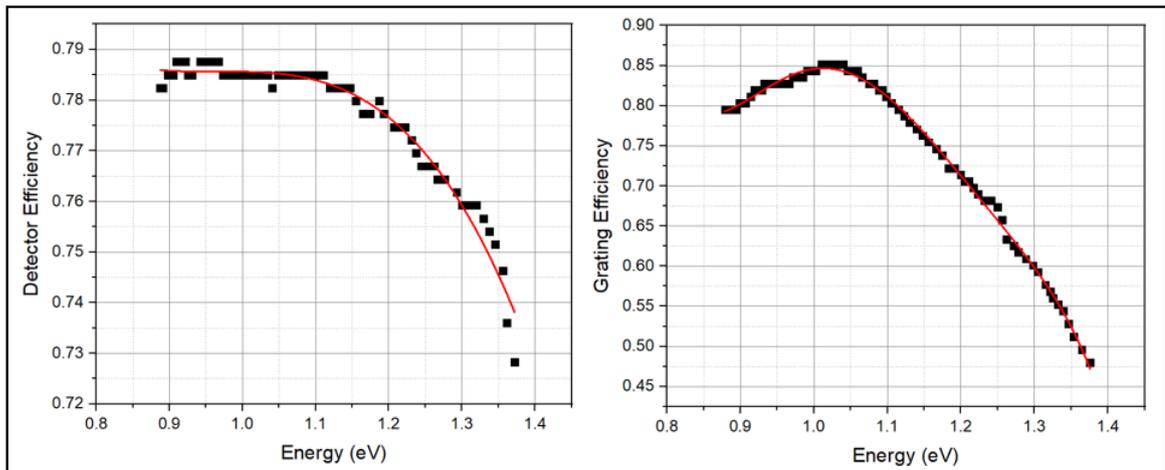
Figure 3.6: Plots to show the polynomial fittings applied to the raw digitised data for the detector efficiency (left) and grating efficiency (right).

A sensible high-energy curve cut-off was selected near the GaAs peak, to simplify analysis with an established λ-range of interest. Importantly, the cut-off position was kept constant for all curves of a given measurement sweep. It was important not to trim the low energy QR tail as this needs to be assessed during sample comparison. Next, a polynomial fit was made to obtain a smooth approximate curve that was easier to work with. The detector and grating efficiency curves were then linearly interpolated, to estimate their respective efficiencies at common E-axis data points. Each new curve must have the same initial E-coordinate (E_{lim} = low energy limit)

and increment (dE). The latter is automatically calculated by the software after setting a sampling interval (1000 points). Both arrays now have an equal number of uniformly-spaced data points along the E-axis, and can be multiplied together to create a net efficiency curve (Figure 3.7) that can be used to normalise the measurements. First, the linear interpolation must be repeated for each of the T- and P- sweeps; using the same parameters as before. Each interpolated measurement is then divided by the product of the efficiency curves.
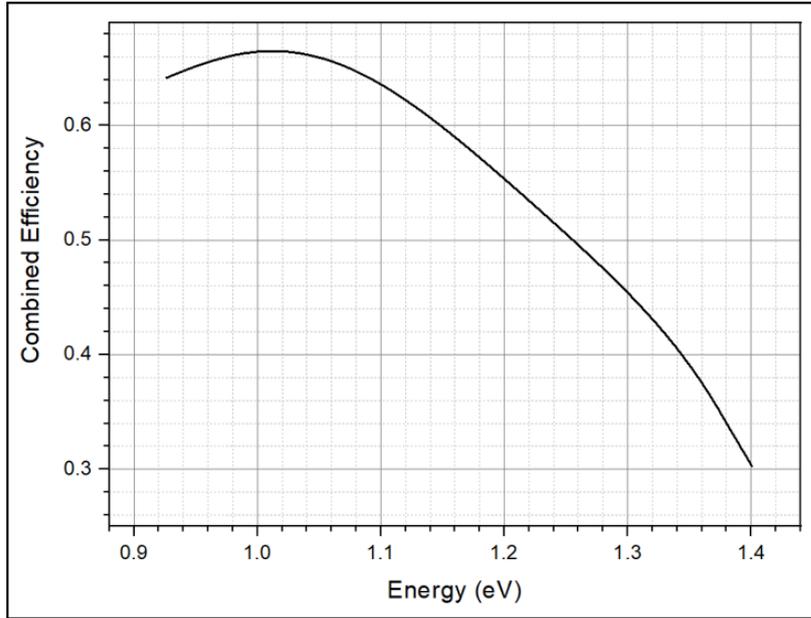

Figure 3.7: Plot of the resultant measurement efficiency used to normalise the data.

Absorption Dip Correction

The following approach mitigated the impact of the absorption dip. The lowest T-curve at full P was used to determine the absorption dip, since the PL should be at its brightest. Therefore the impact the absorption dip has on the spectra will be minimal and easiest to correct for by eye. Notably, the lineshape roughness caused by multiple sources of loss made it difficult to achieve a completely smooth correction with a single mini-Gaussian fit.

A Gaussian function was simulated with a magnitude expressed in terms of relative peak brightness, i.e. the percentage by which the peak dips by. For one of the T-sweeps, this was found to be approximately 2.965 % and is similar across all samples. This agrees well with ~ 3 % quoted in the Thorlabs specifications sheet of Ref. [206]. The baseline therefore was set to 1 counts/s, whilst the peak amplitude was set close to 1.03 counts/s, so the peak only corrected the data at a region specified by the FWHM. It was difficult to control the amplitude of the peak maxima at precisely 1.02965 counts/s, due to working with area and FWHM in the Origin

fitting dialogue. The area underneath a Gaussian curve, $A = h_{\mathrm{G}}\sigma(2\pi)^{1/2}$, where h_{G} and σ are the Gaussian height and standard deviation, respectively. The position of the peak maxima was set to 1.09 eV according to Refs. [201, 206]. Multiple attempts of trial and error were necessary to best guess the Gaussian height needed to compensate for the absorption dip. The array of Y-values were then linearly interpolated for the same range of E-values used in the data. In a similar way to above, the new array of Y-values taken from the simulated Gaussian plot was multiplied with the normalised measurements.
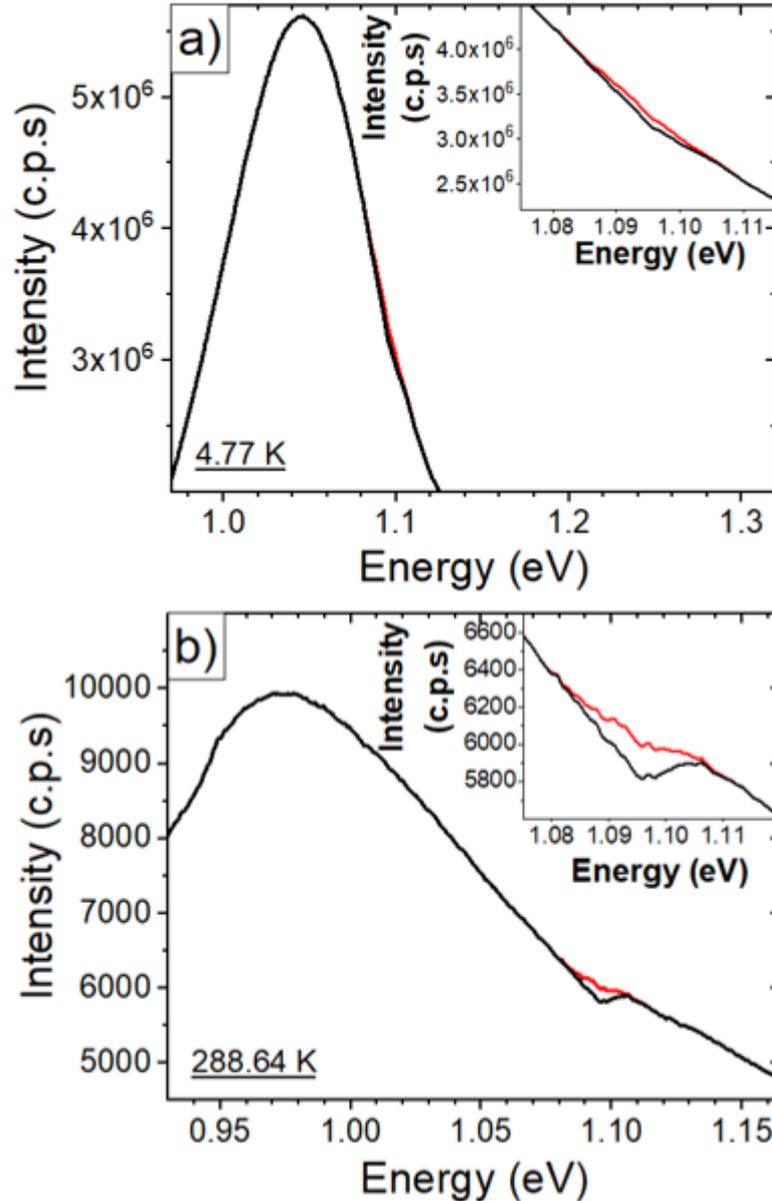
Figure 3.8: Plots to show the absorption dip correction for the reference sample at a) 4 K and b) near RT. The red line represents the corrected line. Magnifications of the dip are included within the inset axes.

After a convincing correction was achieved at LT, it was important to check the dip

had a negligible T-dependence. The same Gaussian was multiplied by the highest T-dataset at full power and this also resulted in a good correction to the lineshape as shown in Figure 3.8. It is reasonable to assume negligible T-dependence on the absorption dip magnitude for the entire sweep, since the highest and lowest T-curves seem reasonably corrected. Little T-dependence is to be expected since most of the fibre is at RT. Even though some of the fibre is contained within the probe and reached 4 K as it was lowered further into the dewar, the rod was submerged less than (1.0 ± 0.5) m which is much smaller than the units of fibre loss [dB/km]. Indeed, the optical fibre was exposed to moisture in the air, however humidity variation inside the lab was minimal due to the air conditioning unit.

Spectral Shift Correction

The last processing step dealt with a consistency issue between measurement sweeps. The WL peak was found to be ever so slightly misaligned on both 4 K measurements of the P- and T- sweeps. The issue is likely to be caused by the spectrometer slit alignment.

The laser line is actually a reliable feature that can be used to check confidence in the spectral output. It was compared to its expected value for each HT measurement. The LT measurements cannot be used as the inhomogeneously-broadened laser line is dominated by QR and WL emission. A Gaussian was fitted to the highest T laser line and the maxima was used to inform the appropriate spectral shift. The entire spectrum of each T-curve was shifted accordingly. This shift was very small. There will be negligible error associated with the discernment of the peak maxima.

The laser peak is barely visible on the LT P-sweep spectra for the 5 nm sample. The peaks are not expected to change across sweeps at the same T and P. Additionally, the WL peak should also remain at the same energy with varied P. Thus, the difference in WL peaks was then used to correct the P-varied spectra for sweeps where the laser line is not visible. This should be fine to correct for the slit alignment since the sweeps were performed consecutively for each sample.

3.4   Atomic Force Microscopy

AFM is a technique which can be used to map surface morphology with extremely high resolution (of the order of several nm). Measurements were performed using a Digital Instruments multimode scanning probe microscope driven by a Nanoscope IIIa controller (see Figure 3.9). They were then analysed using NanoScope Analysis software.
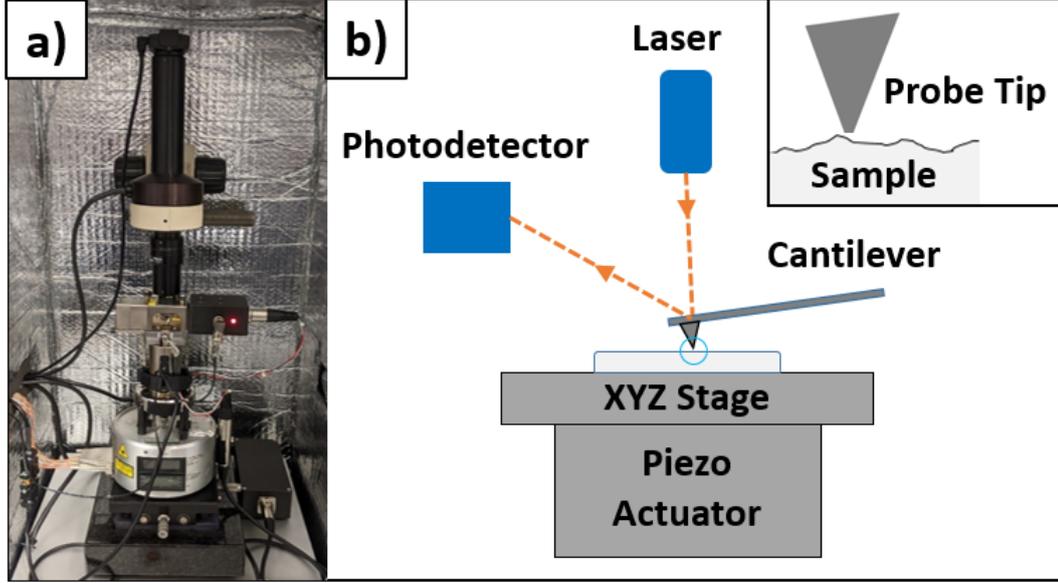
Figure 3.9: a) Photograph of the AFM apparatus used in this work, along side b) a schematic diagram to elucidate the operation of the main components. The orange dashed lines with arrows represent the laser beam path. Inset: Close up of the probe tip and sample surface.

The probe has a very fine crystalline silicon-tipped cantilever, with a force constant of approximately 40 N/m; which is scanned closely along the sample surface, using a piezoelectric motor for precise actuation. A laser is delicately focused onto the back of the cantilever and detects the deflection, by reflecting the beam back to a photodetector and measuring the exaggerated laser spot movement.

The cantilever acts as a spring with a resonant frequency, $f_0 \approx 300$ KHz. Assuming it is rectangular, homogeneous and isotropic, along with simple loading conditions, f_0 in the z-direction can be expressed as [207]:

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{k_{\text{spring}}}{M + m_{\text{eff}}}}. \tag{3.11}$$

Here, M is the load mass, and $m_{\text{eff}} \approx 0.24 x_{\text{w}} x_{\text{t}} x_{\ell} \rho$ [207] is the effective mass of the cantilever; where x_{w}, x_{t}, x_{ℓ}, and ρ are the cantilever width, thickness, length, and density, respectively. A small effective mass and clean probe tip are necessary to achieve an optimal high frequency and enable fast scan rates in tapping mode. The spring constant, k_{spring}, is defined in the z-direction by [208]:

$$k_{\text{spring}} = \frac{Y_{\text{m}} x_{\text{w}} x_{\text{t}}^3}{4(x_{\ell} - \Delta x_{\ell})^3}. \tag{3.12}$$

Here, the distance from the loading point (probe tip) to the end of the cantilever is denoted by Δx_{ℓ}, and Y_{m} is the Young's modulus. The more flexible the cantilever

is, the better it can sense deflections.

As the cantilever is brought into close contact with the sample, atomic forces act on the tip and cause it to bend normal to the surface. At very small tip to sample distances, a very strong repulsive force arises due to exchange interactions from the overlap of electronic orbitals at ranges of a few angstroms. The tip is in *contact mode* when under such force. Van der Waals forces arise when the tip is pulled back, up to a critical cut off distance when it is too far away. Instantaneous polarisation of an atom induces a polarisation in nearby atoms which creates this attractive force. The system is said to be in the *non-contact* and *tapping* modes under these conditions.

The controller can be operated in a number of modes. For this project, the constant height (contact) and tapping modes were used. Constant height mode is best used for taking micrographs of relatively flat surfaces, e.g. QWs; it was used in the preliminary AFM measurements of this project, where high-speed scans were required. Tapping mode was used in lieu of the constant force mode for the high resolution measurements, due to the combined benefits of contact and non-contact AFM.

In contact mode, either height or force can be kept constant. The probe tip is effectively dragged across the sample surface during the scan. This method is strongly affected by frictional and adhesive forces, which cause data discrepancies and damage to both the sample and the probe tip. The scanner moves at a constant height above the surface in constant height mode. Cantilever deflections are measured and used to directly map surface contours. For completeness, constant force mode produces images by using the cantilever deflection as an input feedback signal, to keep it at a fixed value and maintain a constant total applied force to the sample. According to Hooke's law, any changes in deflection will cause a change in height. These changes are measured and used to generate an image. In this method, the scanning speed is limited by the feedback circuit's response time.

Tapping mode minimises frictional forces, along with adhesion of contaminants to the tip due to a meniscus formed by capillary condensation [209]. The damage done to the surface and tip is thus reduced in comparison to contact mode. The cantilever oscillates up and down, as it is driven near resonance with a piezo crystal in the cantilever holder. The frequency and amplitude of the driving signal remains constant. The system is calibrated to provide the optimal oscillation frequency and the force set point is kept as low as possible. When the tip passes over an obstruction (depression), it has less (more) room to oscillate and therefore the oscillation amplitude decreases (increases). The tip amplitude is fed into a feedback circuit, which controls the height of the cantilever above the sample surface. The height is then adjusted accordingly and a constant oscillation amplitude and force on the sample

is maintained throughout the scan. The surface topology is imaged by mapping the correction in tip-sample height; at each intermittent contact.

3.5   X-Ray Diffraction Spectroscopy

Bragg's law (Equation 3.13) states that a superposition of wavefronts will scatter from a surface at an angle equal to the angle of incidence, and that the path difference must equal a whole number of wavelengths for constructive interference to occur. It is used to explain how X-Rays interact with crystalline facets. Notably, Bragg's law is a special case of the more general Laue diffraction.

$$y\lambda = 2d_\perp \sin(\theta_\mathrm{B}). \tag{3.13}$$

Where λ is the wavelength of the incident X-Ray beam, $d_\perp$ is the perpendicular distance between two planes, θ_B is the Bragg angle, and $y \in \mathbb{Z}^+$.

Diffraction occurs because the X-Ray λ is comparable to that of interatomic distances, as shown in Figure 3.10. When photons interact with an atom, they cause a perturbation to its electron cloud. Forcing the charges to move causes them to re-radiate electromagnetic waves at the same frequency, i.e. the incident X-Ray beam can be considered to have scattered elastically (Rayleigh scattering) according to the atomic positions in the lattice. Some of the scattered rays constructively interfere at specific angles and remain coherent, some do not and cancel each other out. This is an accumulative effect which intensifies with each lattice plane. However, real materials have many planes that cause destructive interference. Hence, measurements generally result with very sharp peaks that dominate noise. These peaks occur where constructive interference is maximised.

XRD scans can be used to study the perfection of a crystalline lattice and to characterise a lattice-matched thin film. This measurement technique can quickly determine attributes such as layer thickness, composition, relaxation etc. An $\omega - 2\theta$ measurement investigates a small region of reciprocal space by fixing the detector at twice the Bragg angle and then tilting the sample back and forth (ω-axis). The Bragg plane is therefore aligned with a number of planes that are divergent from the sample surface; yielding a distribution of planes as a function of tilt [210]. In an ideal crystal, the peak FWHM is caused by the spectral width of the X-Ray source and beam geometry. Poor epitaxial quality (mosaicity) and imperfections (dislocations, defects, and film curvature) in a real crystal lattice cause the FWHM to broaden.
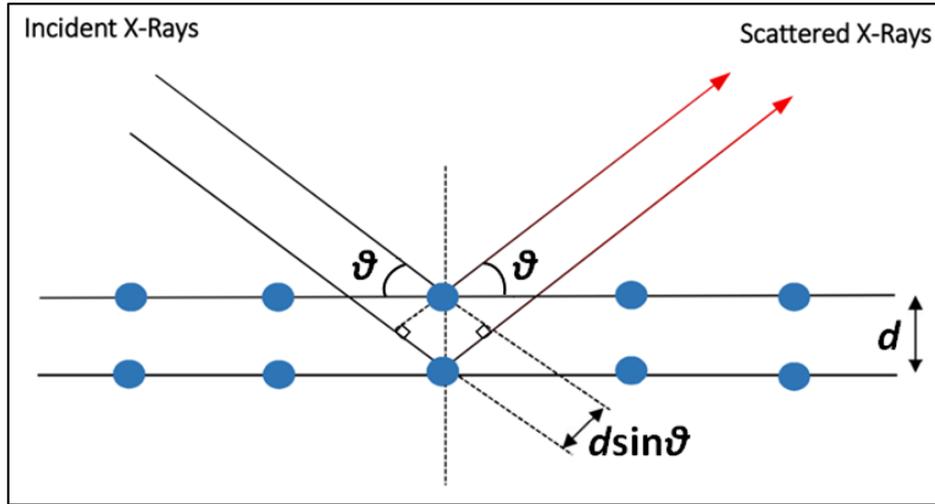
Figure 3.10: Diagram to show Bragg diffraction. The incident X-Rays are scattered and reflected in a specular manner. The path difference, between $y = 1$ and $y = 2$, is $2d_\perp \sin(\theta_B)$.
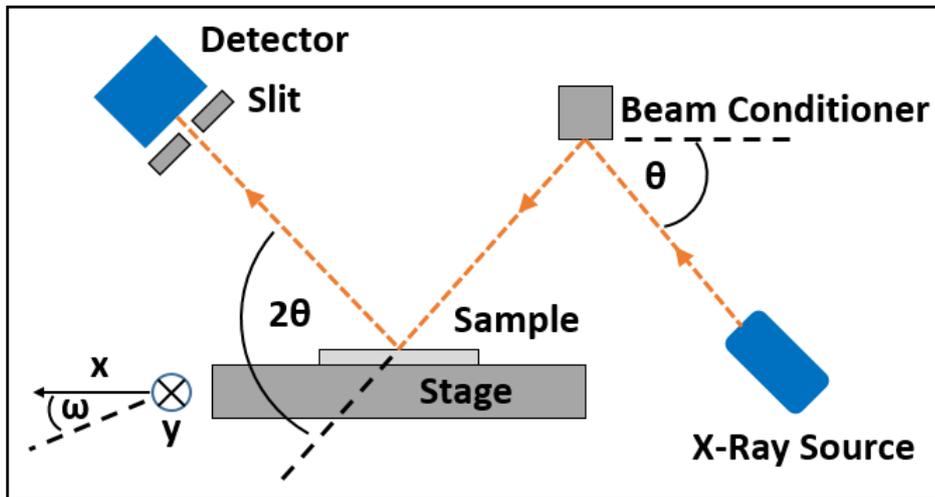
Figure 3.11: Schematic diagram of the XRD diffractometer.

An indication of the quality of epitaxial growth was obtained using a Bede QC200 diffractometer (see Figure 3.11). High resolution $\omega - 2\theta$ scans were produced and then Mercury RADS software was used to simulate the theoretical XRD result according to dynamical scattering theory. This was fitted to the measured data to determine the thickness of layers and the In composition of the QWs.

In this experiment, a GaAs beam conditioner (high-quality crystal with a well-defined lattice constant) is used to select the correct Bragg angle ($\theta_B = 33.025$ °) and find the InGaAs-004 diffraction peak. The incident radiation λ was 1.54056 Å. The lower and upper scan range limits were set to -5000 arsec and 3000 arcsec, respectively; with an increment of 4 arcsec for the 10 nm $In_{0.15}Ga_{0.85}As$/GaAs MQW. For the 15 nm and 20 nm samples, the lower and upper scan range limits were set to -4500 arcsec and 4500 arcsec, respectively; with an increment of 2 arcsec.

The MQW structures were replicated in the fitting software with a simple model. Each of the GaAs layers were assumed to be completely relaxed, and the relaxation of the InGaAs QW layers was assumed zero, however for the 10 nm MQW, the relaxation was set to 0.2 % to optimise the fit. This makes extremely little difference to the analysis. Fitting the theoretical peaks to the measured data allowed the software to automatically extract the thickness of each layer and the In composition of the QW layers.

Chapter 4

Theoretical Results

Simulations of a single QR layer were performed to validate the assumption of size and shape, further understand the emission spectra of this structure, and check feasibility of wavelength extension. For the model to produce meaningful results that well describe the experimentally observed PL; the confinement needs to be taken into account, along with strain-induced band deformation, the electron-hole coupling, and the piezoelectric charge density. Furthermore the resolution of the grid spacing and number of eigenvalues affect the quality of calculated data.

As far as we are aware, there have been no band structure simulations performed for InGaAs-capped GaSb/GaAs QRs before, therefore the electronic properties of these nanostructures have not been studied in detail. Furthermore, the majority of simulated GaSb/GaAs QRs are based on truncated pyramids, however, due to thermochemical edge reconfiguration it is likely that QRs are more circular in shape. Circular QRs are frequently observed in AFM, e.g. in the work of Ref. [54, 81, 193]. Notably, a small population do appear to have more of a rectilinear shape with rounded vertices.

This chapter begins with the presentation of h_c results. Calculation results are then presented for the reference structure and RWELL samples, where the strain profiles, band diagrams (with confinement energies and exciton E_b-values), transition energies, and wavefunction probabilities are presented.

4.1 Critical Thickness

The values of h_c for a single $In_xGa_{1-x}As$/GaAs layer, are plotted as a function of In concentration, x, in Figure 4.1. This calculation was used to inform the decision on the appropriate range of QW thicknesses to investigate. Notably, the dislocation core

dimensions become comparable to h_c, for In compositions greater than 40% [116]. Experimental results are therefore expected to deviate from linear elastic theory beyond this concentration, thus calculated values for this interval are increasingly unreliable. The data calculated using the People and Bean model appears to over estimate h_c at low In x and then starts to drop off at 60 %, highlighting the necessity for further improvement, e.g. that which is proposed in the Dodson–Tsao kinetic model [222]. Although significantly lower, the real value of h_c is most likely closer to the Fitzgerald curve considering the 15 nm RWELL sample results displayed in Figure 5.12.
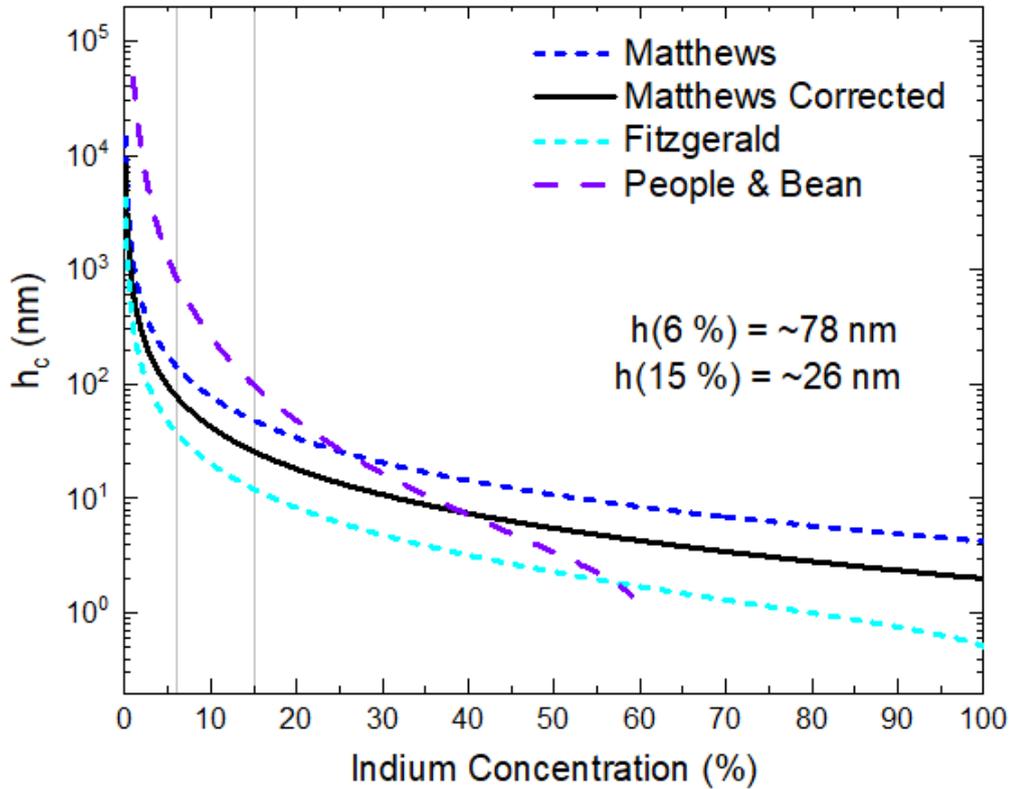
Figure 4.1: Plot to show h_c as a function of x, for a single $In_x Ga_{1-x} As$/GaAs layer. The composition ranges between $0 \leq x \leq 100$ %. The model proposed by Matthews and Blakeslee is plotted with the original parameters and those proposed by Fitzgerald, and then compared against that of People and Bean. The solid curve was used to discern the appropriate QW thickness.

## 4.2 Electronic Properties

4.2.1 Strain Distributions

The numerical strain minimisation calculation is a critical feature of the simulations presented in this Chapter, whereby the total elastic energy in a continuum elasticity model is minimised. Strain not only directly modifies the band-edge energies through the linear deformation potentials, but produces polarisation charges that augment

the electrostatic potential. The position-dependent total charge density is therefore calculated included in the Poisson equation.

To assess the dominant strain tensor components contributions to the hydrostatic strain acting on the QW and the QR, 2D cross-sections of the lateral (xx)- and transverse (zz)- strain tensor components are plotted at $(x, 00002, z)$ and $(x, y, 0.0002)$, respectively. The results for the RWELL structure with the most strain (30 % alloy concentration), are presented in Figure 4.2. All other calculations displayed a similar QR-induced strain field of varying magnitude.
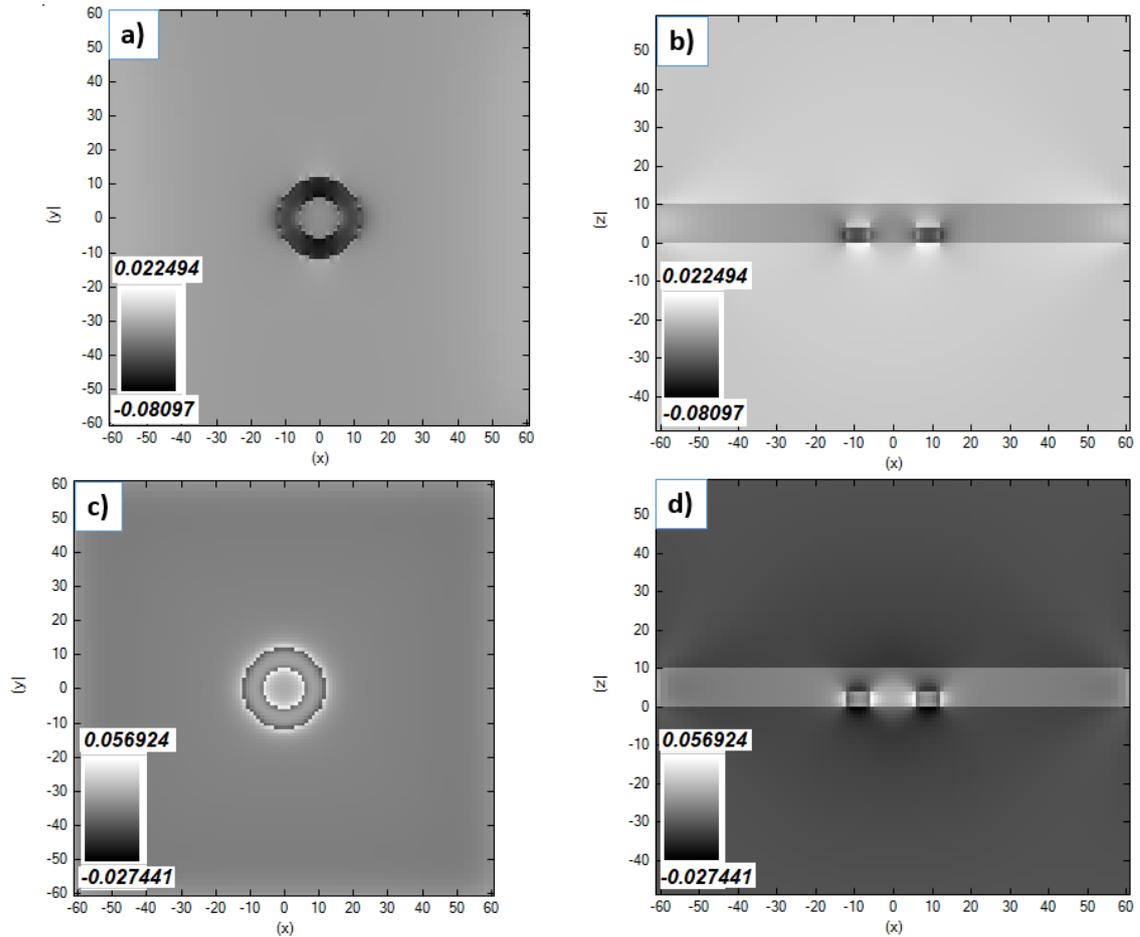
Figure 4.2: 2D strain plots of the $x = 30$ % RWELL structure, to show the top-view (left) and side-view (right) of the position-dependent magnitudes of the a,b) lateral and c,d) transverse tensor components. Bright regions correspond to where the lattice is stretched and dark regions correspond to where it is compressed.

A non-negligible relative change in lattice parameter is observed almost everywhere throughout the simulated structure. The Poisson effect is clearly visible by the opposite signs of the perpendicular strain components (bright and dark). Here, the lateral strain tensor components in the x- and y- directions are assumed equal by symmetry arguments. Above and below the QR, the GaSb lattice stretches the InGaAs QW unit cell laterally, which causes a positive (tensile) strain component in the x-direction. The reverse is true for the compressed InGaAs either side of the

QR. The out-of-plane lattice constant reduces to accommodate the change in lateral strain component and corresponds to a negative (compressive) strain component in the z-direction.

The magnitude of lattice distortion reduces with h from the nanostructure, since the unit cells of each subsequent layer experience less force than at the interface due to elastic neighbouring bonds, thus retain more of their bulk character. In other words, there is a balance of strain acting on the local (In)GaAs lattice by the GaSb lattice, and vice versa. Similarly, the lateral component of strain inside the QW reduces with distance from the QRs.

Notably, the lateral strain component inside the GaAs barriers and the QW is relatively uniform away from grid boundary. Although there is significant lateral tensile strain above the QR, it is much less than that in the GaAs directly below the QR due to the differing a-values. The transverse strain component is maximal in the QW at the InGaAs/GaSb interface and minimal in the QW layers above the QR surface. There is a significant contribution to the hydrostatic strain inside the QR core by the transverse strain component, which is minimal at the origin. It can be seen that the GaSb lobes have the most compressive lateral strain, which reduces towards the outer QR diameter. This is reasonably explained by the relative volume difference between the lobe regions and the surrounding material, in addition to the counter-force acting on the GaSb unit cell by the $In_{0.3}Ga_{0.7}As$ QW. The WL has not been included in these simulations, therefore in reality, it is speculated that the magnitude of the lateral strain component, both here and at the bottom of the QW will reduce, since the (In)GaAsSb layer will have an intermediate lattice parameter.

For each investigation, the hydrostatic strain of the most and least strained RWELL structures was analysed and plotted in Figure 4.3, in order to understand how QW width and alloy composition affects the CB edge. A vertical 1D slice of the hydrostatic strain is taken through $(0.0002, 60002, z)$. A horizontal 1D slice of the hydrostatic strain is also taken along the x- direction at $(x, 0.0002, 0.0002)$. .

Increasing the QW thickness and composition increases the compressive hydrostatic strain in the GaSb core and the surrounding material. Varying the indium content has a stronger effect (~ 2 % shift) on the hydrostatic strain in comparison to changing the QW width.

The transverse compressive strain inside the QR is maximal at the middle of its lobe. Whereas for the QW it is maximal above the QR. The lateral strain bows upwards towards the grid boundary for $12 < x_{pos} < 62$. This behaviour increases profoundly for the highest In % calculation and yields a triangular potential at the grid boundary. It is thus concluded that as strain increases, the model poorly
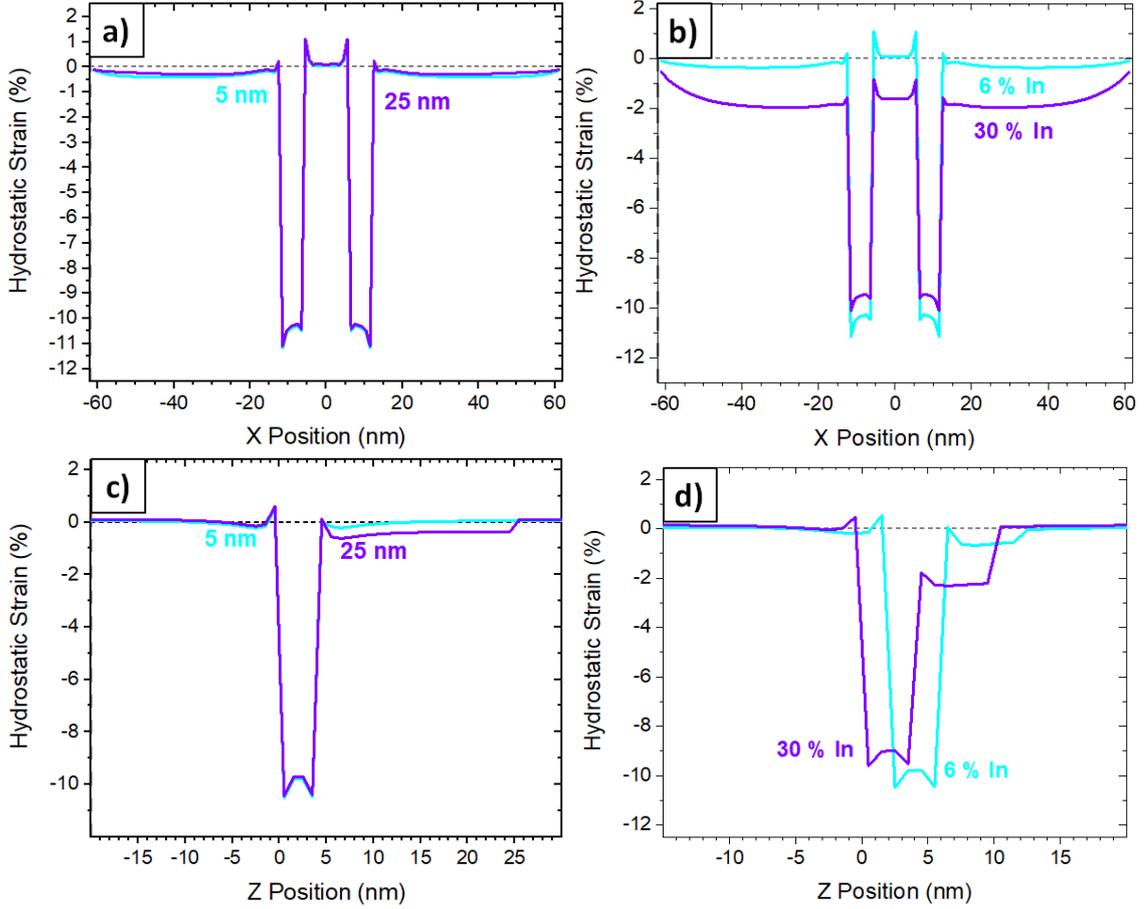
Figure 4.3: 1D strain profile plots of the hydrostatic strain for the simulations of most and least strain.

predicts the strain profile inside the QW, due to assuming zero lateral confinement. The lateral hydrostatic strain inside the QW region contained within the inner circumference of the QR core is positive and increases towards the InGaAs/GaSb interface.

4.2.2 Wavefunction Probabilities

The electron wavefunction probabilities are presented in Figures 4.4 and 4.5 for the QW width-varied and x-varied calculations respectively. It is concerning that no evidence has been observed to suggest that the electron sits in the QR centre. A wider diameter and thinner QR may reduce the strain and allow for this to happen, should the lateral confinement remain neglected.

The electron wavefunctions are well-confined to the QW along the z-axis. The electrons. Counter-intuitively, the electron wavefunction the electron wavefunction probability cloud appears to spread laterally outwards towards the edges of the simulation grid, with increasing strain despite this distance being much longer than the anticipated a_X. This is a direct result of the strain profile, which implies an
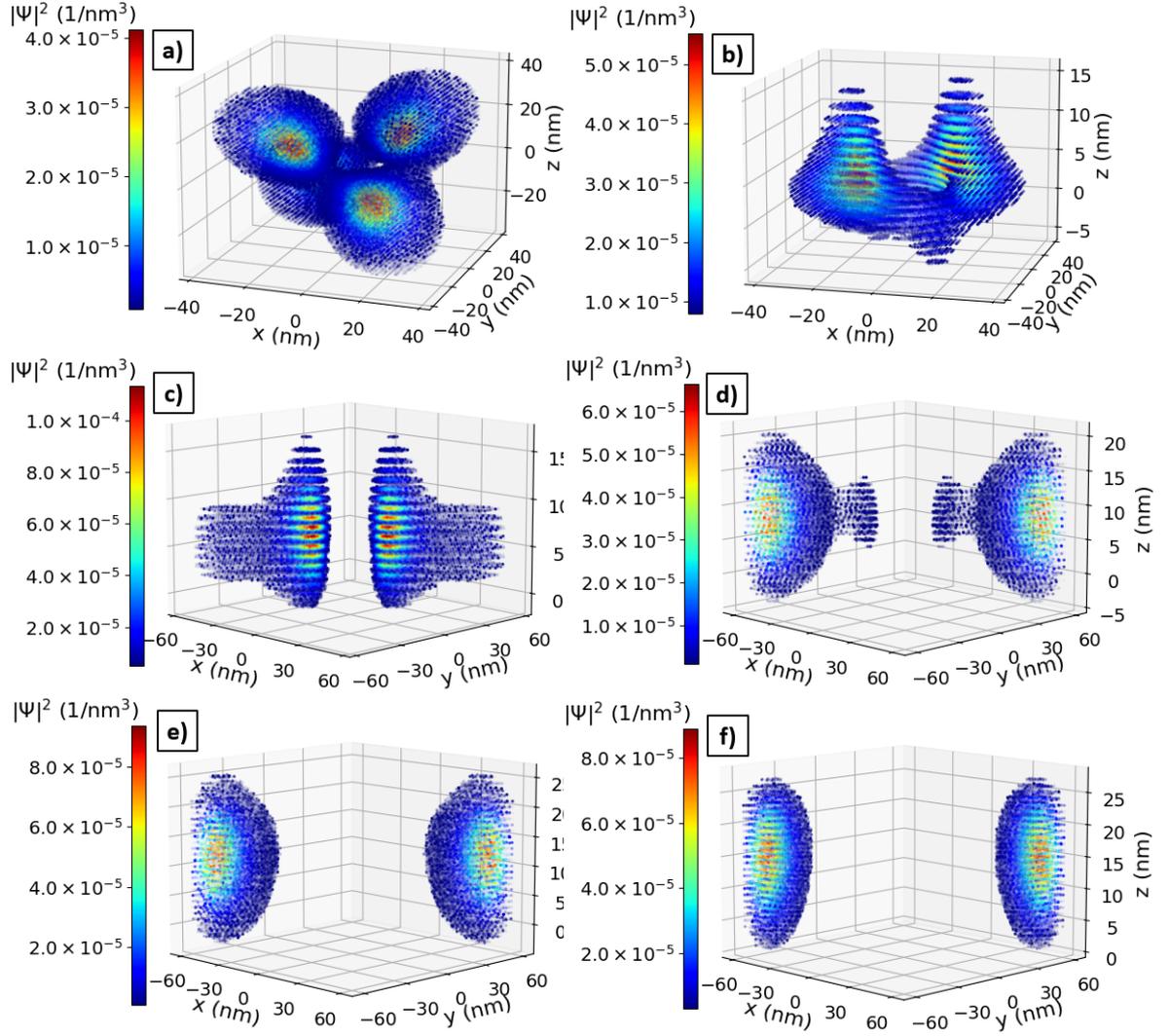
Figure 4.4: Calculated electron groundstate wavefunction probabilities in real space for a single a) GaSb/GaAs QR reference, and In$_{0.06}$Ga$_{0.94}$As-capped GaSb/GaAs QR at QW widths of b) 5 nm, c) 10 nm, d) 15 nm, e) 20 nm, and f) 25 nm.

antibonding scenario i.e. it is energetically favourable for the electrons to sit as far away from the QR as possible, since there is no lateral confinement potential from surrounding nanostructures at play in the model. The prediction of the electron's lateral position should therefore be taken lightly. This current model cannot give a precise lateral placement, however this work is a refreshing demonstration of the critical sensitivity of the carrier wavefunction overlap to the internal strain field inside these materials. It adds a slight deviation to the calculated transition energy. In the limit of weak strain, the model gives a reasonable approximation to the lateral exciton Bohr radius since the electrons are pinned to the outside of the GaSb QR as expected
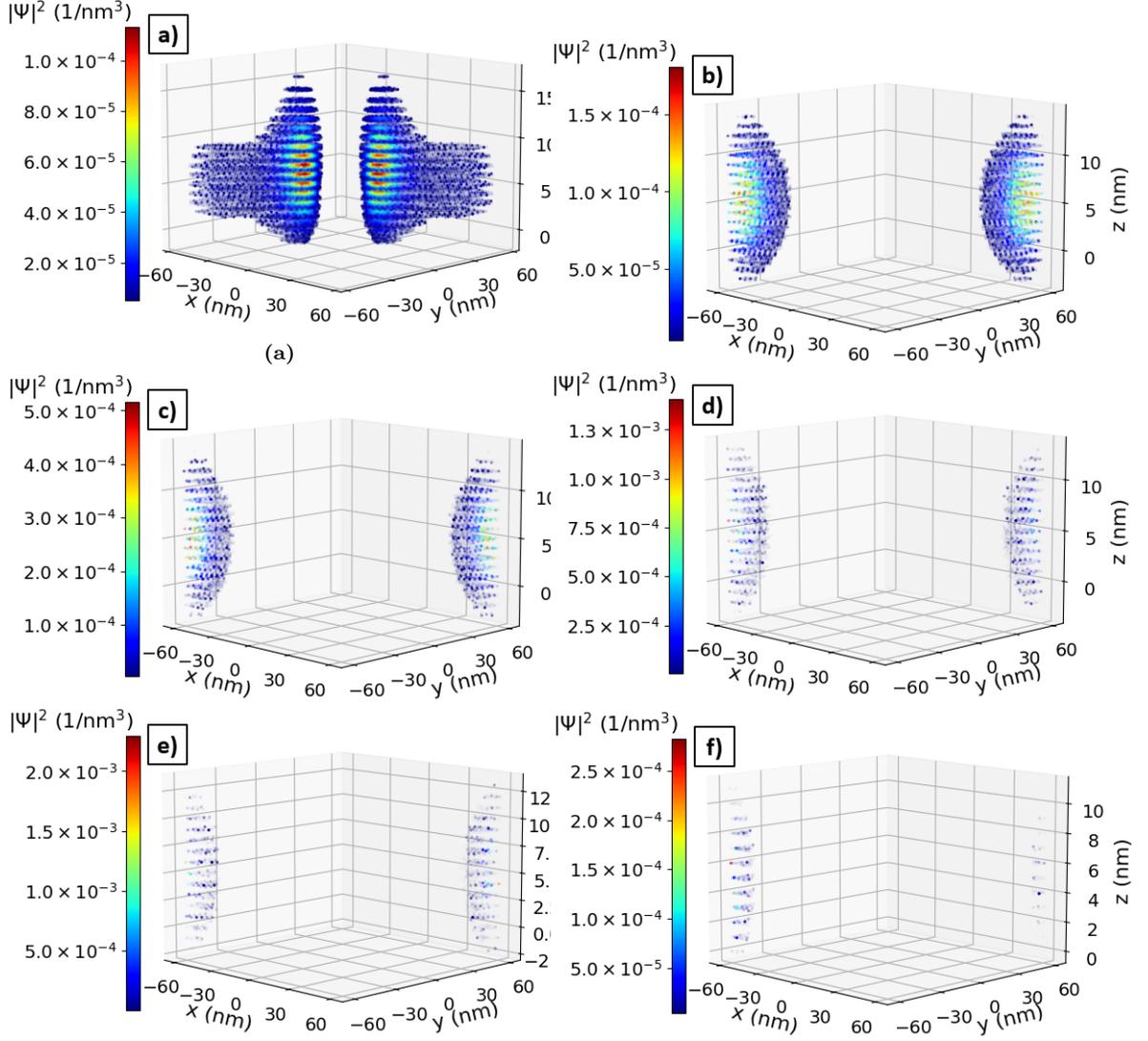
Figure 4.5: Calculated electron groundstate wavefunction probabilities in real space for a single In$_x$Ga$_{1-x}$As-capped GaSb/GaAs QR at alloy concentrations of b) 10 %, c) 15 %, d) 20 %, e) 25 %, and f) 30 %. The previous 6 %, 10 nm sample is included in a) for comparison.

4.2.3 Confinement Energies

Different aspect ratios aren't expected to significantly affect the transition energy, however changes to the confinement in the z-direction will. Transverse band edge diagrams of the RWELL samples are presented in Figures 4.6 and 4.7 to show how the CB edge changes with increasing indium content and QW width. The CB edge energy is significantly reduced near the grid boundary. Notably, the CB edge energy of the simulated results is spatially-dependent. Electron states underneath the QRs (GaAs) are counter-intuitively lower in energy than the lowest energy states inside the QW at coordinates $(0.3002, 0.5998)$. This effect is likely due to strain and the piezo-electric charge density, since the resolution of the "circle" is poor.. Thus, these simulations imply that it is energetically favourable for GaAs electrons to tunnel through the WL and recombine with the GaSb holes, where the electric polarisation

110

is large, i.e. at the sharp edges of the QR. The rough edges may reflect the behaviour of real disorder in the QRs. The bands are tilted at the heterojunctions, this effect cannot represent CC since there is only one electron and 1 hole in the simulated system. It must reflect the electrostatic potential and hydrostatic strain profile. The triangular potential well will be bigger in the real samples and this will likely affect the ground state energy of the electron and the localisation of its wavefunction. IT is most likely one of the main factors for the discrepancy in experimental and theoretical data points. The triangular potential wells look square due to the low 1×1 nm grid resolution.

The benefit of the subsequent reduction of the electron subband energy with increasing thickness becomes outweighed by the cost of increased strain and likelihood of defect-induced device failure. Thus, 10 nm is selected for the QW thickness in the investigation of In concentration. We expect better electron confinement with a deeper QW, hence better thermal stability.

It is expected that the wavelength should increase and the CB drop wih In % off of the assumption that the electrons recombine with the QRs vertically from above. This ought to lead to a reduction in CB edge. Actually these simulations show that the average position of the electrons is sensitive to the lateral strain component since the CB edge is lower further away from the QRs. These results agree with the observation that QRs provide enough confinement as a sub ten nm QW. Above 15 nm, the electrons prefer to sit on the outskirts of the QW. Notably there is little point in investigating the excited state with the current model, being able to rule out this potential emission by aligning the peak energies will allow for a better discernment of the source of the double peak in Chapter 5.

The transition energies follow a parabolic distribution as expected with the energy scaling of inverse width. Even though the grown samples did not exhibit the expected emission wavelengths, the conditions explored in this Chapter should give rise to them. Remarkably, these simulations suggest an optimistic feasibility for the significant extension of wavelength across the entire telecom bands. Thus, potentially allow for these structures to be used as ground-breaking, highly λ-tunable single-source emitters. Indeed the lateral confinement must first be investigated to improve the accuracy of the predicted electron subband energies. Solutions are already under development to realise this, e.g. selective epitaxial overgrowth, hybrid nanostructures etc.
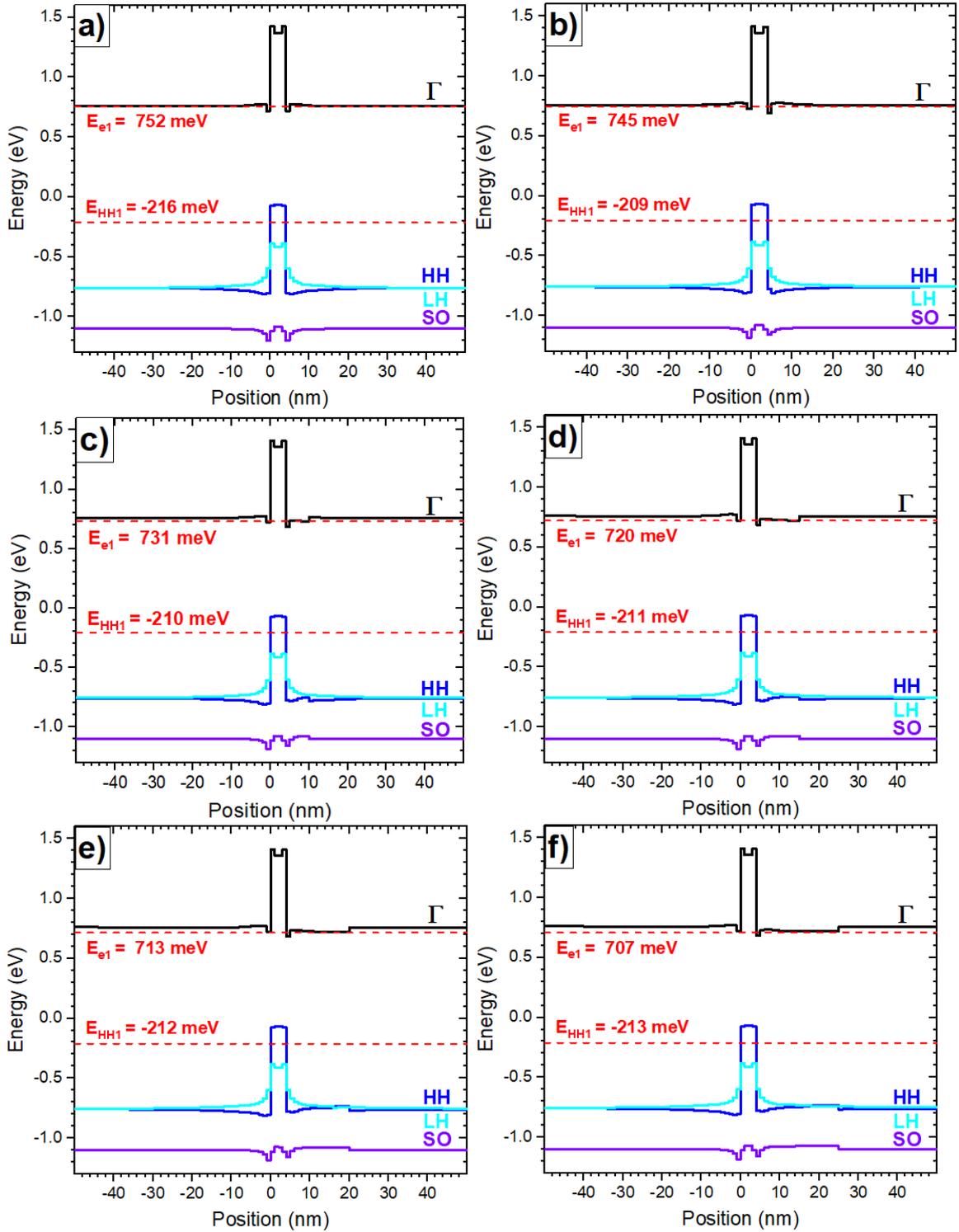
Figure 4.6: Calculated transverse band diagrams of a single a) GaSb/GaAs QR reference, and an $In_{0.06}Ga_{0.94}As$-capped GaSb/GaAs QR at QW widths of b) 5 nm, c) 10 nm, d) 15 nm, e) 20 nm, and f) 25 nm. The 1D-slice is taken along the z-axis at grid coordinates (0.002, 6.0002, z).
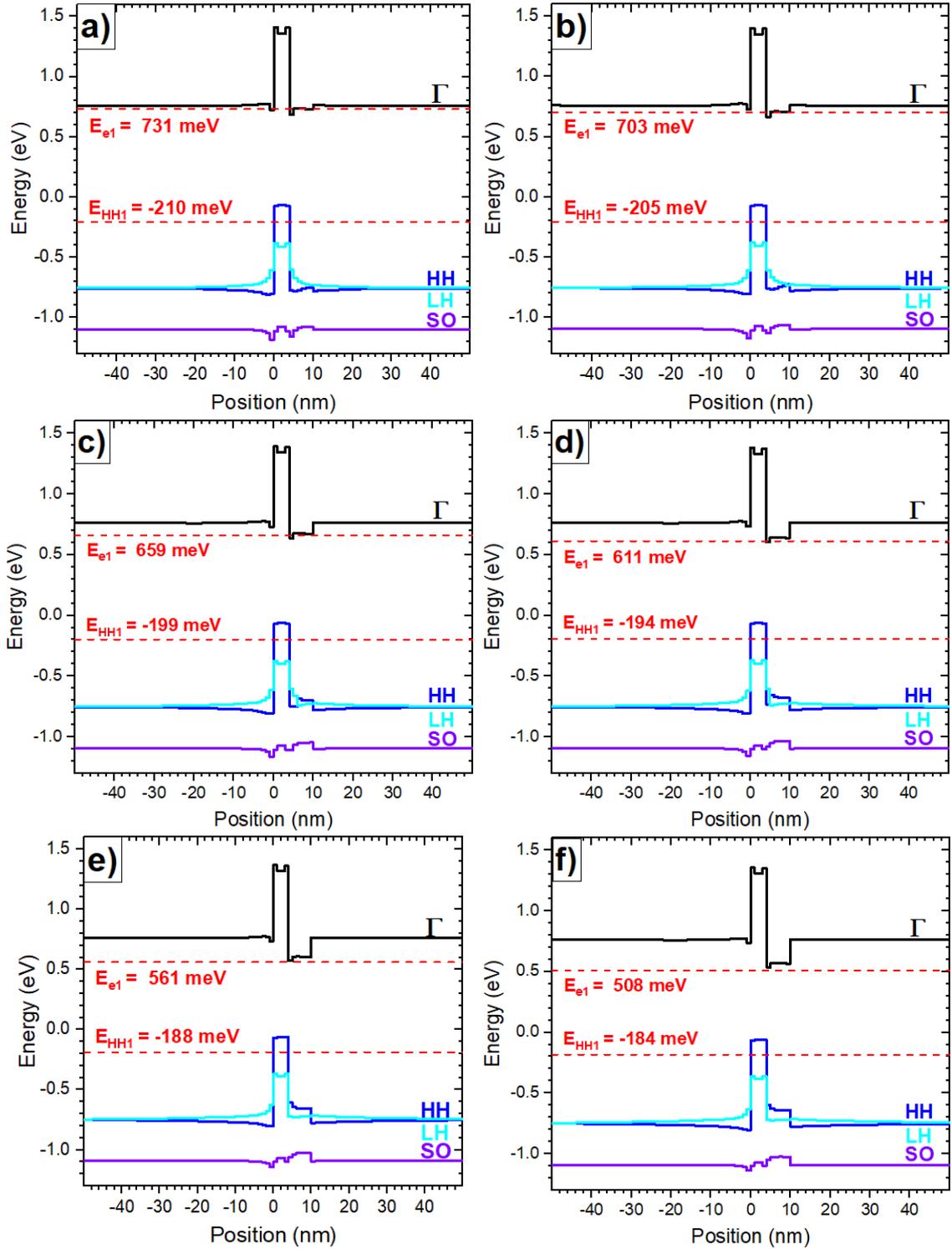
Figure 4.7: Calculated transverse band diagrams of a single In$_x$Ga$_{1-x}$As-capped GaSb/GaAs QR at alloy concentrations of b) 10 %, c) 15 %, d) 20 %, e) 25 %, and f) 30 %. The 1D slice is taken along the z-axis at grid coordinates (0.002, 6.0002, z).
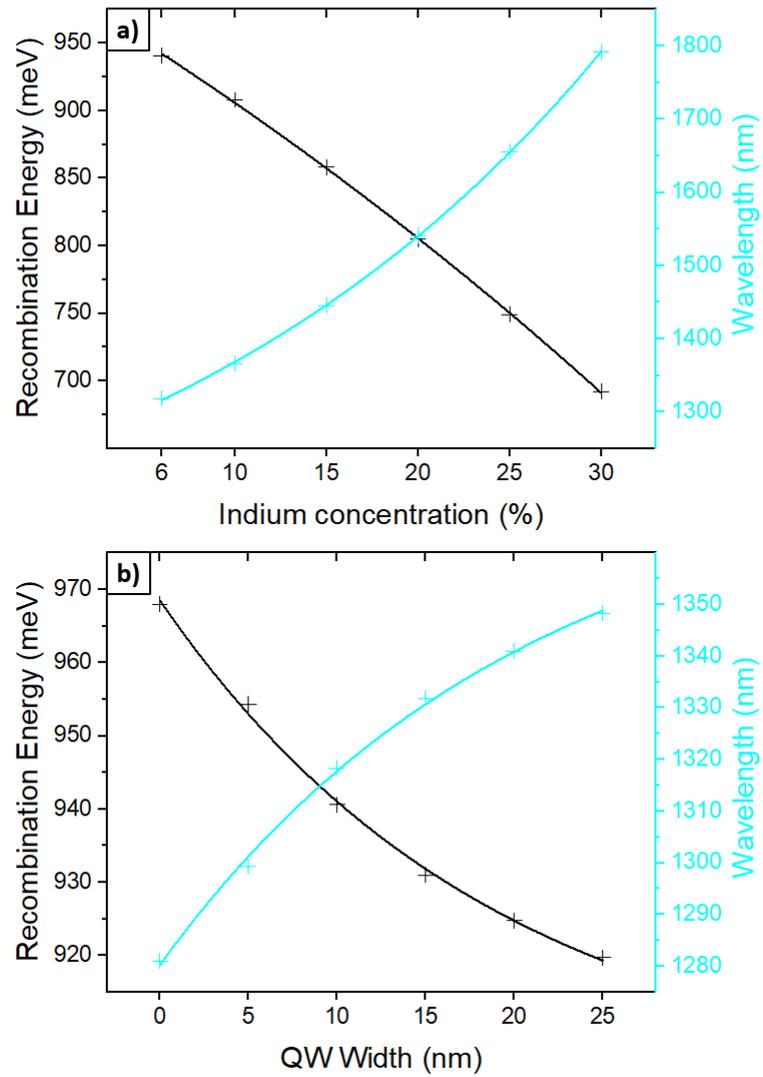
Figure 4.8: Simulated groundstate transition energies, at the Γ-point, as a function of a) QW width, and b) alloy concentration.

Chapter 5

Experimental Results

Here we developed the opto-thermal properties of the active region used for the GaSb/GaAs VCSEL [11]. We wanted to explore the potential use of this structure for telecom LDs, by expanding on the work in Ref. [59] with the inclusion of an annealing step and a triple stacked MQW structure; i.e. we deposited an asymmetric $In_{0.06}Ga_{0.94}As$ QW directly above the QRs in repeated layers. This was to use $\Delta\psi$ as an extra degree of freedom in tuning λ, reduce the variation of opto-thermal properties across samples, and utilise the increased internal strain of the asymmetric QW; by forcing a greater percentage of QD-like nanostructures to change into QRs via a better understood and predictable way. This work led to a nanostructure ensemble with improved height uniformity as inferred from a reduction in Γ_{PL} and less sub-linearity of I_{PL}, as a function of P. The features observed in the PL measurements were further understood, by investigating the QW width as an independent variable to probe confinement effects. In these structures, Γ_{PL} is poorly understood; therefore, it is important to investigate the effect of wavefunction penetration into the barrier layer, using a MQW to keep the defect density (e.g. dopants etc.) and disorder constant.

This chapter introduces considerations made with regards to the grown structures. The PL data is presented, analysed and discussed. The critical thickness was calculated for $In_xGa_{1-x}As/GaAs$ and presented in Figure 4.1.

5.1 Hypotheses

The temperature stability of QR PL can be adjusted by varying electronic confinement. The spatial separation of the electron and the hole in type-II band structures leads to faster thermal quenching of luminescence with increasing T. Conversely, stronger confinement effects in narrower QWs are expected to increase Coulomb in-

teraction and $\Delta\psi$ for the electrons and holes in type-II excitons, potentially leading to a larger thermal E_a, due to increased E_b of the exciton. This in turn should result in better thermal stability of PL.

5.1.1 Growth Technique Outcomes

Layer Stacking

Stacking layers of QDs/QRs [46, 57] increases complexity due to several design aspects of growth that accumulate strain; e.g. number of layers, their composition, the cap growth parameters, annealing temperature and duration, the separation thickness between layers, and the type of buffer used. Composition is affected by T_{sub}, atomic flux, growth rates, interruption times, and substrate rotational speed. These were carefully considered to achieve maximum growth quality. Although there is no minimum thickness for the hot cap; making it too thin can allow the lower MQW layers to influence the strain profile of the upper layers. This may lead to relaxation through the spontaneous formation of defects and vertical stacking etc. A MQW structure can have an inhomogeneous QR distribution, i.e. the average size and density of the QRs in each subsequent layer increases, due to the floating Sb [56, 57] (In [58]) layer during deposition; hence Γ_{PL} values, pertaining to the net QR emission of the entire stack, will be wider than that of single layer structures. This cost is outweighed by enhanced I_{PL}.

Here, fine-tuning of the bandgap is mostly achieved by the thickness and x of the QW, although the the intended size distribution of QRs will also play a key role. An account for the decisions made during experimental planning is given here.

Growth Rate

Reduced Γ_{PL} can be achieved with lower rates of growth, with the additional benefit of longer emitted λ [216, 217]. A low growth rate was chosen to form the QRs since the atoms are able to migrate longer distances across the sample surface. Thus, high crystalline quality is achieved and larger nanostructures grow, since it is energetically favourable for the atoms to assemble at existing QDs rather than form new ones. Self-limiting growth leads to better size homogeneity across the QD ensemble [218]. Growth interruptions are important because they give the atoms, hence QDs, more time to rearrange themselves; thereby indicating that overgrowth is a non-equilibrium process [192].

Cold-Capping

The volume of the capped nanostructures is expected to be slightly reduced compared to the surface nanostructures [219]. Therefore, to minimise this dissolution effect, the cap is grown with a colder T_{sub} than recommended for optimum growth quality, and the As/Ga flux ratio is set to ~ 5. The cold-cap effectively locks the buried atoms in place, so the layers above should not affect the QR geometry, regardless of how thick the hot cap or barrier layer is. A thin cap layer is often desirable as it enhances indirect excitonic radiative recombination at elevated T. However, states at the interface can easily capture carriers and weaken PL if it is made too thin [28].

The yield of resulting QRs, QDs and clusters is directly affected by thermal and strain energy; hence influenced by the thickness of the first cap layer [132], and the temperature of the second hotter cap during growth [57]. Smakman et al. [132] found that varying the thickness of the first GaAs cap layer in GaSb/GaAs structures strongly influences the outcome of the growth process. QDs were preserved using an 8 nm thick layer, whereas less than 10 % of the QDs remained when a 4 nm cap layer was applied. This is because large lattice mismatch gives rise to a Ga-Sb and In-As pair energy of 1.87 eV and 1.40 eV, respectively [194, 220]. GaSb QDs embedded in GaAs are therefore less stable than their InAs counterparts and self-destruct relatively easily [86]. A combination of this bond energy and large local strain fields [221] enables substantial exchange of Sb for As during capping, thus GaSb nanostructures preferentially form QRs over QDs.

The annealing step should have supplied enough thermal energy to drive As-Sb exchange, hence a large QR density is expected. However, the lattice mismatch and pair energy will be slightly reduced with an InGaAs cap. Whilst there is the possibility for an amount of intermixing (if at all significant), the successful formation of QRs with an InGaAs cap is more challenging than with the GaSb/GaAs material system. Due to this, both QW growth temperature and thickness require careful consideration.

Importantly, the nanostructures undergo partial relaxation, i.e. it is incorrect to assume zero residual strain. Band structure calculations in Ref. [162] show both strained and relaxed GaSb/GaAs offsets. The relaxed calculation indicates a type-I structure and clearly this is not the case when taking blueshift results into account both in this work and in literature (e.g. [170]).

The Dodson–Tsao kinetic model was used in Ref. [223] to explain the dependence of residual strain and crystalline quality on growth temperature. QWs with a low In concentration (12 %) were investigated at T_{sub}-values between 420 °C and 477 °C.

For film thicknesses in the range of QW widths studied here, it was found that the generation of dislocations is larger at lower growth temperatures and could possibly be a compensation for slow glide velocity. Therefore the cold cap technique requires an in situ annealing treatment to form the RWELL samples with high crystalline quality.

Thermal Annealing

Thermal treatments are commonly employed across the semiconductor industry to homogenise the distribution of group-III atoms, smooth interfaces (helpful for buffer continuity and graded layers), and lessen the density of non-radiative line defects. Consequently, quantum efficiencies and lifetimes are maximised, thus I_{PL} is increased, and Γ_{PL} is reduced. The associated cost is a widened GaSb bandgap [170], enhanced diffusion of group-V atoms [33] and slight variation of buried nanostructure morphology. The latter will offset the Γ_{PL} reduction in a trade-off. Samples may be annealed during growth or afterwards.

Rapid thermal annealing (RTA) can be used as a post-processing step to reduce strain and improve crystalline quality in fabricated devices [46, 224] at the risk of causing undesirable changes to nominal parameters such as E_a and λ [46]. Blurring the (In)GaAs/GaSb interface smears out the confinement potential [33, 46], whereas reducing nanostructure height and intermixing-induced VB edge lowering causes blueshift. It was found that annealed QDs lose (gain) uniformity in the lateral (out-of-plane) directions [225]. The structural quality of a layer may degrade if the RTA temperature is set too high. These techniques usually involve temperatures around $600 - 800$ °C and affect the entire sample.

Alternatively, in situ annealing techniques [32, 33, 47, 48] are more subtle and applied during epitaxy around 580 °C. Applying this annealing process after cold-capping is shown to further relax the strain in close proximity to the nanostructures, by enhancing intermixing; as inferred from larger QR openings with lower Sb content in the cross-sectional transmission electron microscopy (XTEM) images taken under dark-field (002) conditions in the work of Ref. [47]. The same is expected for our samples, thus slight changes to the spatial configurations of the exciton complexes are expected. A reduction in the yield of undesirable nanostructures is also expected, since MDs are known to act as points of nucleation for clusters.

Annealing the samples is therefore expected to increase the wavefunction overlap and lead to more type-I behaviour [46, 171]. Sb-As intermixing allows ψ_h to leak out of the QD, whilst ψ_e leaks in, as shown in Figure 5.1b. An increased optical oscillator strength is therefore to be expected [46], which ought to improve the

intensity of excitonic PL emission. However, this measurement is beyond the scope of the current investigation. Furthermore, an increased E_b results in an extended indirect exciton lifetime (inverse of the decay rate) relative to excitons observed in bulk and traditional GaSb/GaAs QD/QR materials; hence the PL emission should be observable at comparably higher T and P.
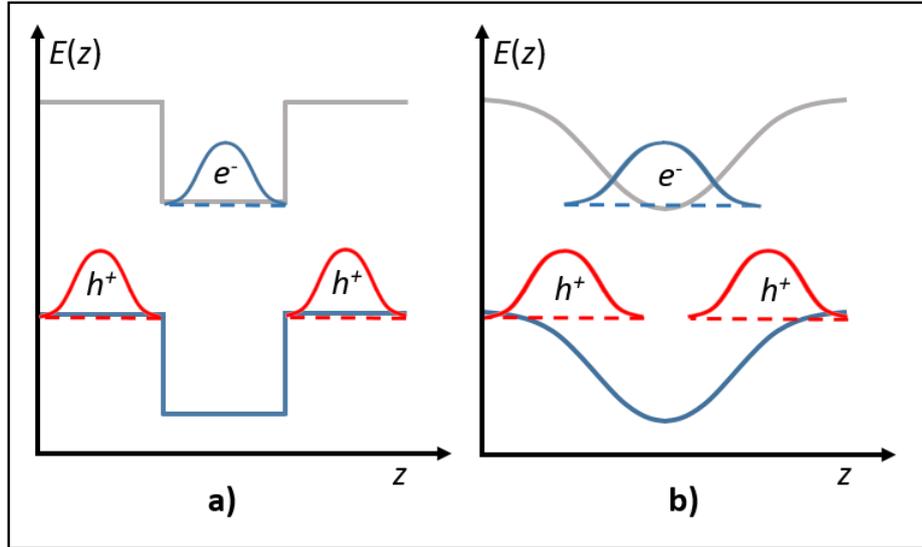
Figure 5.1: Bandgap diagram to show the $\Delta\psi$ at sufficiently HT and confinement in a type-IIa structure. The carrier distribution is assumed to be unaffected by charge separation for clarity in discerning the broadening effects due to increased confinement and intermixing. The upper and lower bands are the CB and VB, respectively. The sloped band edges in b) reflect how the lattice interface loses abruptness with significant interdiffusion.

5.1.2 Structural Properties

Layer Thickness

It was found in Ref. [226] that the In migrated approximately 60 ML, or 20 nm (2 significant figures) into the barrier. Our stacked QR layers are isolated by a 250 nm-thick barrier. It is reasonable to suggest that this layer ought to be thick enough for all of the floating layer of In to be fully desorbed despite an increase of segregation length at 670 °C, hence it is confidently presumed that each QR layer is insulated from In and Sb atoms migrating above from the lower layers. Notably, this issue will need to be addressed when coupling nanostructure layers, thus indicating how an RWELL structure may be advantageous for improved thermal stability when using this material system for lasers.

Furthermore, the periodic layer separation thickness influences the rate of carrier capture in QWs [227]. When the barrier thickness is less than ~ 20 nm, the carrier capture in a QW is a purely quantum mechanical and rapid process (thermalisation). In contrast, when the barrier thickness is larger than approximately ~ 50 nm, then

capture is dominated by diffusion of carriers from the barriers into the QWs. Therefore the latter is expected to be the predominant capture mechanism inside the QWs. Importantly, QW width had no observable dependence on the capture process in the work of Ref. [227].

Layer Composition

The intermixing between As and Sb is significantly less in SSMBE than for alternative techniques such as MOCVD, and its In-Ga analogue observed in InAs/GaAs QDs [228]. However, the application of an in situ annealing treatment to GaSb/GaAs QDs was not only found to reduce the height of the QDs, but to cause some intermixing at the interfaces due to diffusion of Sb [33]. This raises a question on the extent to which these RWELL samples resemble the intended nanostructure morphology and structural composition. A hypothesis is presented here, although ideally, further characterisation measurements would resolve any ambiguity (see chapter 7).

Intermixing of As in the QDs/QRs may be a contributing factor to any PL discrepancies. Hodgson et al. deposited $In_{0.8}Ga_{0.2}Sb$ QDs onto GaAs [229]. These nominal concentrations resulted with a highly intermixed $In_{0.08}Ga_{0.92}As_{0.95}Sb_{0.05}$ outer disc, surrounding an $In_{0.8}Ga_{0.2}As_{0.19}Sb_{0.81}$ core. The resulting recombination energies are increased, hence intermixing was primarily of group-V elements, not group-III. Intermixing will be significantly reduced in the samples presented here, since they have a much lower In %. Non-negligible amounts of As are expected in the nanostructures after annealing, however this should not affect results much as the amount of exchange will be similar in the reference sample.

A floating layer of Sb naturally forms during GaAs capping, acting as a surfactant, since the bond between Ga and Sb is weaker than that between Ga and As. A small amount of Sb atoms are incorporated into the GaAs matrix, whereas a significant number of them remain at the growth front and become incorporated into the layer above. Strain exacerbates this effect. Segregation of Sb increases the size and density of the QRs as the number of repeated layers increases, thus shifting PL emission to longer λ [56]. In attempt to combat Sb loss through group-V segregation, the GaAs is soaked with Sb prior to GaSb deposition.

It is widely accepted that strain also acts to reduce the In incorporation during growth; the reduction being greater at higher T_{sub}. Greater diffusion lengths of group-III atoms tend to broaden the emission FWHM. Notably, tensile stress is less effective at suppressing In incorporation than compressive stress [230]. The possibility of a strain-induced In/Ga exchange reaction [231,232] is worthy of deeper consideration here. The extent to which the In atoms migrate in an InGaAs/GaSb/GaAs

triple-stacked RWELL structure is yet to be investigated. However, for the reasons outlined below, it can be inferred with confidence that the In % inside the QRs really is negligible and has little effect on the PL response.

In segregation in the QW is expected to slightly compensate band bending (BB) effects in our samples, i.e. modify the triangular QW potential profile; since the CB edge is expected to be lower than usual towards the cap layer due. This might reduce the rate of blueshift with increasing P; and add to the average spatial separation of carriers, especially if a "W"-shaped potential is formed. Muraki et al. proposed a phenomenological model [233] to predict the variation of In composition across an InGaAs QW and GaAs barrier. The model suggests that as In atoms are incorporated into the alloy; a fraction of the incident flux, denoted by the segregation coefficient R, always segregates to the surface? layer during the deposition of each ML. A larger population of In is expected to increase surface roughening. Anticipating this, a low In concentration was employed so that the population of In at the surface would remain low [234]. The electrons are concentrated at the bottom of the QW due to the QRs, and the segregated In floats to the top of the QW, so FWHM broadening in a single RWELL layer would be insignificant. However the roughness is likely amplified through the z-axis in a triple-stacked MQW, and smoothed significantly upon annealing. In addition, lower nominal In concentrations are expected to yield smaller In-rich clusters, which can behave like QDs [224].

The In concentration in the yth layer may be estimated by:

$$x(n) = \begin{cases} rx_0(1 - R^Y), & \text{for } y \leq Y; \\ rx_0 R^{(y-Y)}(1 - R^Y), & \text{for } y > Y. \end{cases} \tag{5.1}$$

Where x_0 is the nominal In concentration and the sticking coefficient, denoted by r, is a ratio of adsorbed to impinged (including desorption effects) atoms. The segregation coefficient dominates the sticking coefficient within the QW by an order of magnitude Y, therefore the model predicts that the segregation rate is larger than the desorption rate. The QW is shallower than expected for lower values of r. The QW and subsequent barrier are represented by $y \leq Y$ and $y > Y$, respectively. The same In segregation and desorption characteristics in arsenides are also observed in antimonides [235] hence applicable to all sample regions in this work.

Fujimoto et al. investigated [226] the In distribution profile of a 30 nm InGaAs/GaAs QW using XRD and secondary-ion mass spectroscopy. The QW is deposited with similar growth conditions to the samples here ($x_0 = 6$ % and $T_{\text{sub}} = 540$ °C) except their QW is much thicker, therefore the tensile strain will be increased. The segregation rate R is assumed constant, and the In concentrations in a ML, and in the

next, are assumed to be in equilibrium. Their derived r-value fell to zero at 540 °C, i.e. the QW stopped growing. The sticking rate is not zero for the less-strained QWs presented here. A zero sticking coefficient would imply that all the QWs are limited to 5 nm. This is not the case considering the XRD measurements of Section 5.3 and the PL measurements of Subsection 5.4, where a clear trend with increasing QW width is revealed. However, it is possible here that part of the QWs grown at a hotter T_{sub} are shorter than anticipated. Notably, the model [236] employed above assumes that the desorption of In atoms within the upper GaAs barrier occurs at the same rate as the sticking. This assumption is likely invalid, hence the result was ignored with caution, in favour of the optimal $In_{0.2}Ga_{0.8}As$ growth temperature window of (530 ± 10) °C, to find a starting point for QW optimisation. For In desorption and segregation effects to become negligible, it takes a relatively cold T_{sub} of 450 °C [233]. The sticking rate should be increased for the lower T_{sub}-values here, thus the In concentration should be closer to the nominal value.

Realistically, we expect the WL of our samples to be strongly intermixed along with the QW at short lengths from the QRs, due to dilution through group-V exchange. The authors of Ref. [57] found that the Sb content of the WL can vary between $30 - 40$ % using a similar As/Ga flux ratio. This is presumably due to differences in strain distribution within the QRs and WL [221]. The QDs/QRs are likely to be almost pure [57,132] due to the instability described in section 2.2.3, which arises at the position where the equilibrium of the interfacial tensions between the cap, QD and under-layer are disturbed [86]. For the reference sample, the GaAs region in the QR centre should be almost entirely free of Sb, unlike the inner radius of InAs/GaAs QRs [220] which contain significant amounts of In. In atoms may conglomerate in the centre of our QRs due to strain. It is also possible that our structures are encased within a thin, dilute $In_xGa_{1-x}As_ySb_{1-y}$ top layer. This quaternary has been shown [237] to suppress group-V intermixing at relatively low In concentrations (~ 15 %) [237] hence may limit any annealing-related differences in composition.

5.1.3 Optical Properties

Quantum Confinement

Quantum confinement occurs when the QW thickness is comparable to the electronic de Broglie wavelength. Electrons are able to move freely in bulk materials, however their restricted movement in 3D-confined structures and this leads to discrete energy levels, therefore the CB and VB are split into subbands. Quantum confinement typically increases E_g in magnitude with the reduction of structure dimensions.

The lowered CB edge along with the barrier provided by the upper GaAs/InGaAs interface ought to enhance electronic confinement within the InGaAs layer. The dimensions of a 10 nm QW capping layer are comparable to a_X. Therefore choosing a QW size smaller than this in our work, will cause the structure to operate in the strong confinement regime.

The additional confinement from the QW forces the individual electron (ψ_e) and hole (ψ_h) wavefunctions to overlap more spatially, i.e. increase $\Delta\psi$ along the z-axis; thus increase the electronic density near the nanostructures.
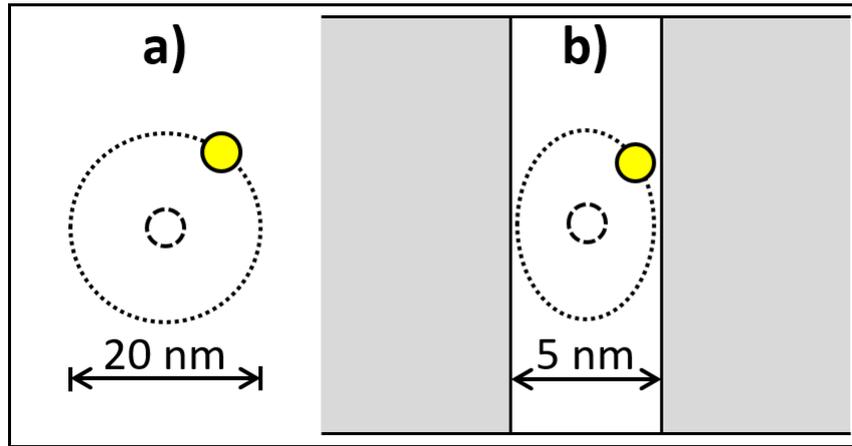
Figure 5.2: Sketches to show a) the typical value of a_X in bulk GaAs [170], and b) the compression of a_X inside a 5 nm-thick QW. The dotted line represents an electronic orbit in the classical picture.

Recombination rates are dependent on the energy difference in states and $\Delta\psi$. A greater percentage of carriers should therefore contribute to photon production, since RWELL systems are known to effectively suppress edge recombination [84]. Additionally, fewer defect-formed, non-radiative recombination channels are accessible within the restrictive active region.

Extension of Wavelength

Using an InGaAs QW should also theoretically lower the CB edge of the active region and further extend λ beyond typical 1.3 µm RT emission. Further gains in λ can be reasonably expected, since the lattice constant of InAs is closer to GaSb; thus an InGaAs cap is expected to slightly reduce the compressive strain inside the QRs, and therefore reduce the hole confinement energy.

5.2 Optimisation of GaSb Deposition

The first experiment of this project sought optimum GaSb growth conditions using this MBE system. Specifically, the duration that the GaAs was soaked with Sb,

and how long the Sb shutter was open for during GaSb deposition, was investigated. Simple GaSb/GaAs QR structures were fabricated. Notably, the growth method for these structures is based on a cold-cap followed by an immediate hot GaAs layer, rather than annealing mid-way. This method is simpler to achieve QRs, hence likely to yield reliable parameters and provide a strong foundation to explore more complex growth techniques. QDs were grown under identical conditions to the buried QRs for characterisation. Importantly, PL from the QDs is typically negligible due to surface states, therefore they are often disregarded as emitters unless formed in close proximity to the buried QD layers; in this case, the buried nanostructures function as a carrier reservoir [238].

Modifying the shutter time trivially adjusts the amount of material present on the GaAs substrate, thus influences the size and density of nanostructures. Suitable growth conditions for GaSb QRs have been well established. They are presumably based on satisfactory growth quality and optimum QR outer diameter in the work of Timm et al. [194]. Authors often use GaSb growth rates and nominal layer thicknesses of 0.3 ML/s, and 2.1 ML, respectively [46, 47, 56, 57, 194, 239]. Strain limits the total amount of GaSb in a QR layer, post capping, to ~ 2 ML [221]; the rest of the Sb behaves undesirably as a surfactant. Thus standard values were adopted for two samples and a third sample was grown at 1.8 ML.

Soaking the growth with Sb before the GaSb layer replaces the Ga–As bond with the Ga–Sb bond [240] and improves the interface quality. Two of the samples were soaked with Sb for 30 s whereas one sample was soaked for 50 s. Bremner et al. demonstrated that a 30 s soak minimises the corrugations at the GaAs–GaAsSb interface [241], i.e. it should reduce roughness and improve layer abruptness of our buffer-WL interface. The samples here are grown at different temperatures even though it would seem to only be a 40 °C difference than in their work. The temperature of the substrate is measured by a PID with the system employed here, and RHEED transitions are used to calibrate the measured T_{sub}, i.e. using surface reconstructions at the deoxidation temperature. There are further sources of error such as sample holder geometry and differences in thermal heat capacity due to the coating they have built up. Moreover, there is often a thermal lag. A rough estimate of the total uncertainty to the temperature measurement is ~ 20 °C. Additionally, a different Sb-flux was applied on top of differences in MBE systems; hence it was appropriate to replicate their experiment and determine a new optimum range, that encompasses the formation of a complete layer. Due to the T-dependence of the exchange reaction, a slightly more aggressive As-Sb reaction is expected, so the upper soak time limit was reduced by 10 s. There is more interest in exploring the upper soak time limit, instead of the lower limit, due to the greater I_{PL}-value

observed with a 60 s soak in the work of Ref. [241], resulting from a more triangular potential well. This result indicates a more abrupt interface.

Substrates were prepared in the same manner as subsection 3.2.3. The buffer layers for each sample were grown at 650 °C, with a growth rate of 1.0 ML/s. The growth temperature was then reduced to a temperature of 540 °C under an As_2 flux for all 3 samples. After, the As-valve was closed similar to the method in subsection 3.2.3 and the surface was soaked with Sb prior to GaSb deposition. A GaSb-rich is indicated by the change in RHEED pattern from a 2×4 to a 1×3 reconstruction [57]. The formation of the QRs was indicated by the change in RHEED pattern from streaky to spotty. The QDs were cold-capped with 5 nm of GaAs, at a temperature of 475 °C and growth rate of 0.3 ML/s. The QR layer was separated from the top layer of uncapped QDs by a 200 nm-thick GaAs barrier grown under identical conditions to the buffer.
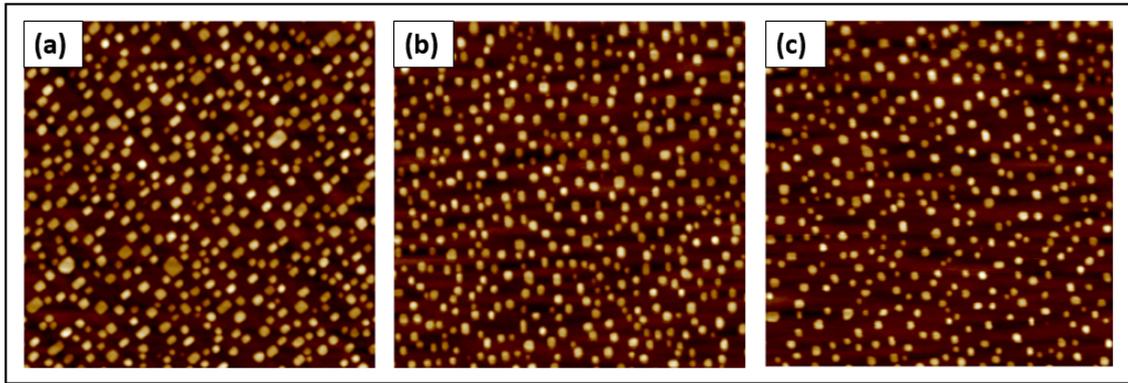
Figure 5.3: 1×1 μm^2 micrograph of the GaSb surface QDs at 540 °C for each deposition time (left to right: $(50 + 7)$s, $(30 + 7)$s, $(30 + 6)$s).

The next batch of growths investigated T_{sub} during GaSb deposition. As a starting point, T_{sub} was set to the estimated value for optimum InGaAs growth. This was to see how the nanostructures might form in high quality $In_xGa_{1-x}As$ QWs. In other words, to gauge their size and emission characteristics within a QW layer of minimal alloy disorder and roughness.

5.3 Structural Characterisation

Unfortunately, there is currently no XRD data for the asymmetric RWELL samples. However, previous $In_{0.15}Ga_{0.85}As$/GaAs MQW samples grown at 10 nm, 15 nm, and 20 nm, were characterised with XRD; along with a symmetric version of the structures analysed in this Chapter, i.e. the QR layer exists in the middle of the QW. These structures can be used to obtain an, albeit limited, understanding on the quality of QW growth for all of the capped-RWELL structures. The AFM and PL
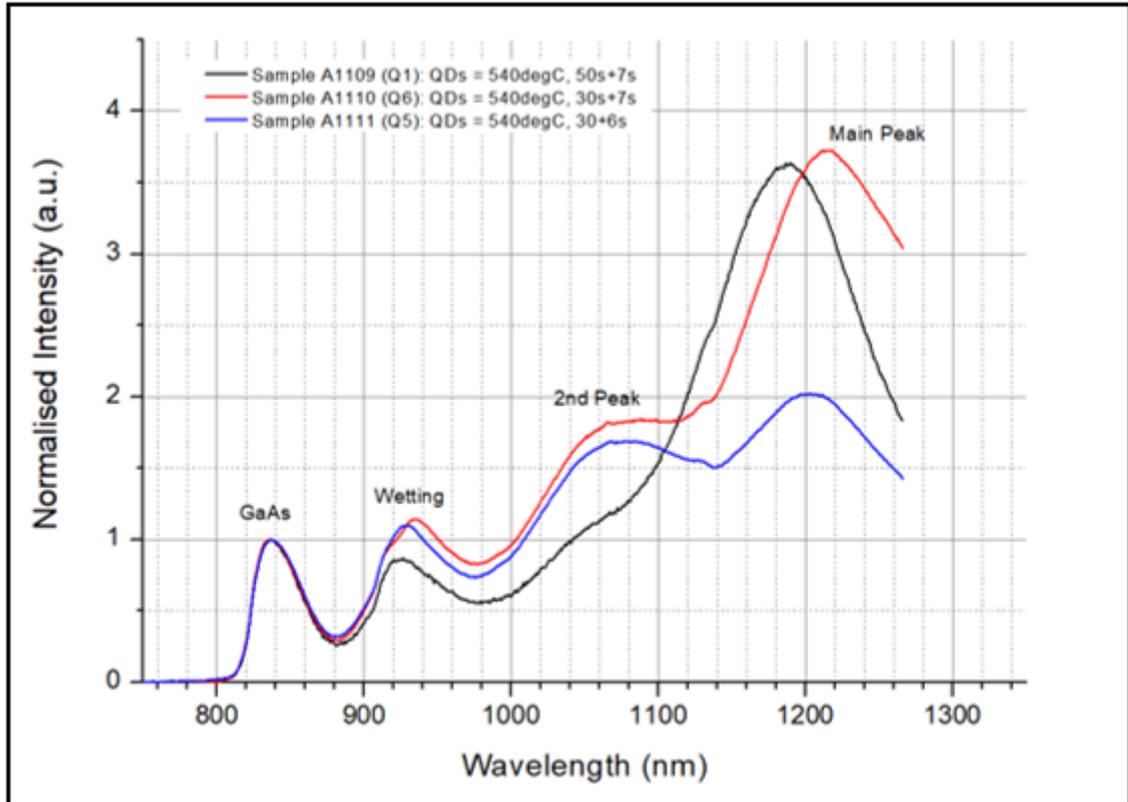
Figure 5.4: Comparison of PL emission for 3 GaSb/GaAs samples normalised to the GaAs peak at 4.2 K and 300 mW for each of the growth deposition times
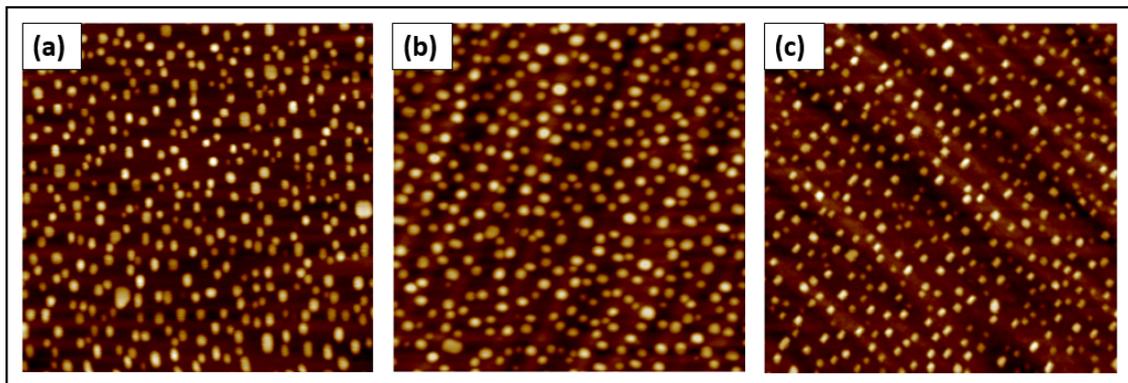
Figure 5.5: 1×1 µm^2 micrograph of the GaSb surface QDs for each sample, at different QR growth temperatures (left to right: 540 °C, 550 °C, 530 °C).

of these samples are beyond the scope of this current investigation, hence discussion is omitted. they are included purely for comparative reasons to assess the case for assuming high crystalline quality. The structure of the symmetric RWELL samples (Appendix C) is similar to that of the asymmetric RWELL samples,however the QWs are richer in In and separated by a thinner barrier (100 nm). The ultra-thin GaAs protective layer is also missing. Furthermore, there are differences in growth conditions compared to the asymmetric RWELL samples, namely: the InGaAs layer was grown at a faster rate of 1.0 ML/s; the GaAs layers were grown at a slower rate
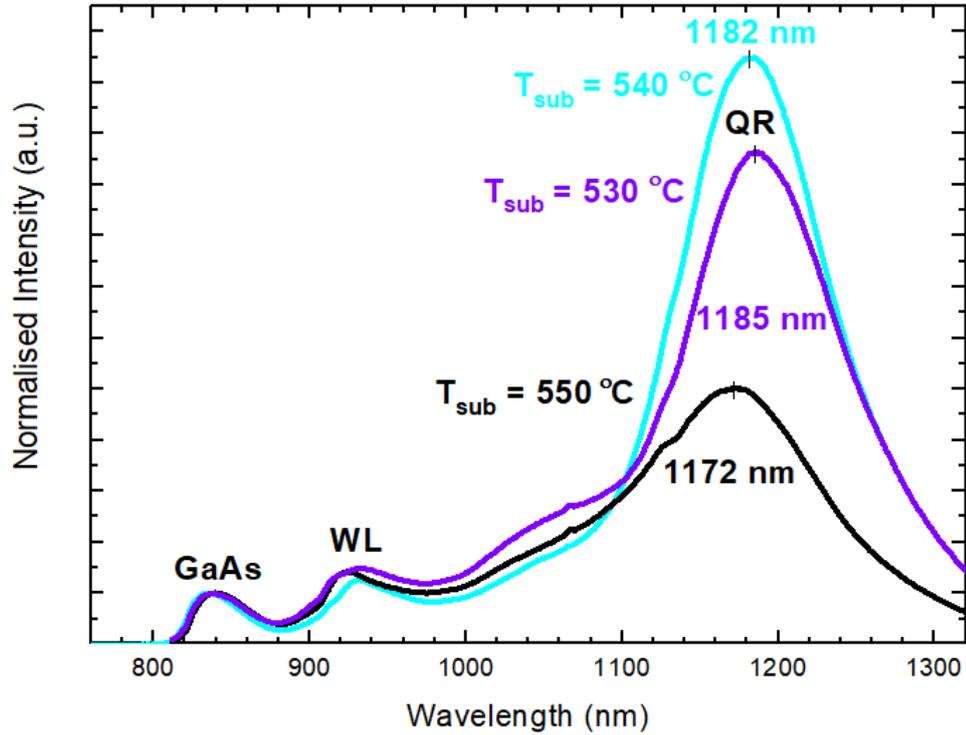
Figure 5.6: Comparison of 4.2 K PL emission for 3 GaSb/GaAs QDs/QRs samples grown on GaAs:Si substrates, using a 300 mW laser source. An optimum T_{sub}-value was found by varying T_{sub} during GaSb deposition, and observing how the corresponding change in density and size of nanostructures affects the PL emission. The data has been normalised to the GaAs peak intensity.

of 0.85 ML/s; a substrate setpoint temperature of 640 °C was used to grow the InGaAs layer and final GaAs cap layer; and the GaAs barrier layers were grown with a set point of 680 °C.

The hotter T-setpoint for the InGaAs QW is an optimum growth T_{sub} value (within uncertainty). Therefore the XRD results provide an upper bound on the crystalline growth quality, i.e. the FWHM for the asymmetric RWELL samples will be slightly wider than the data fit, due to the nucleation of dislocations. However, QW composition should be slightly more accurate in comparison to the nominal values for the asymmetric RWELL samples because of the lower T_{sub} and In content. The discrepancy of thickness is expected to be similar for both structures since the uncertainty depends mainly on shutter transients during growth.
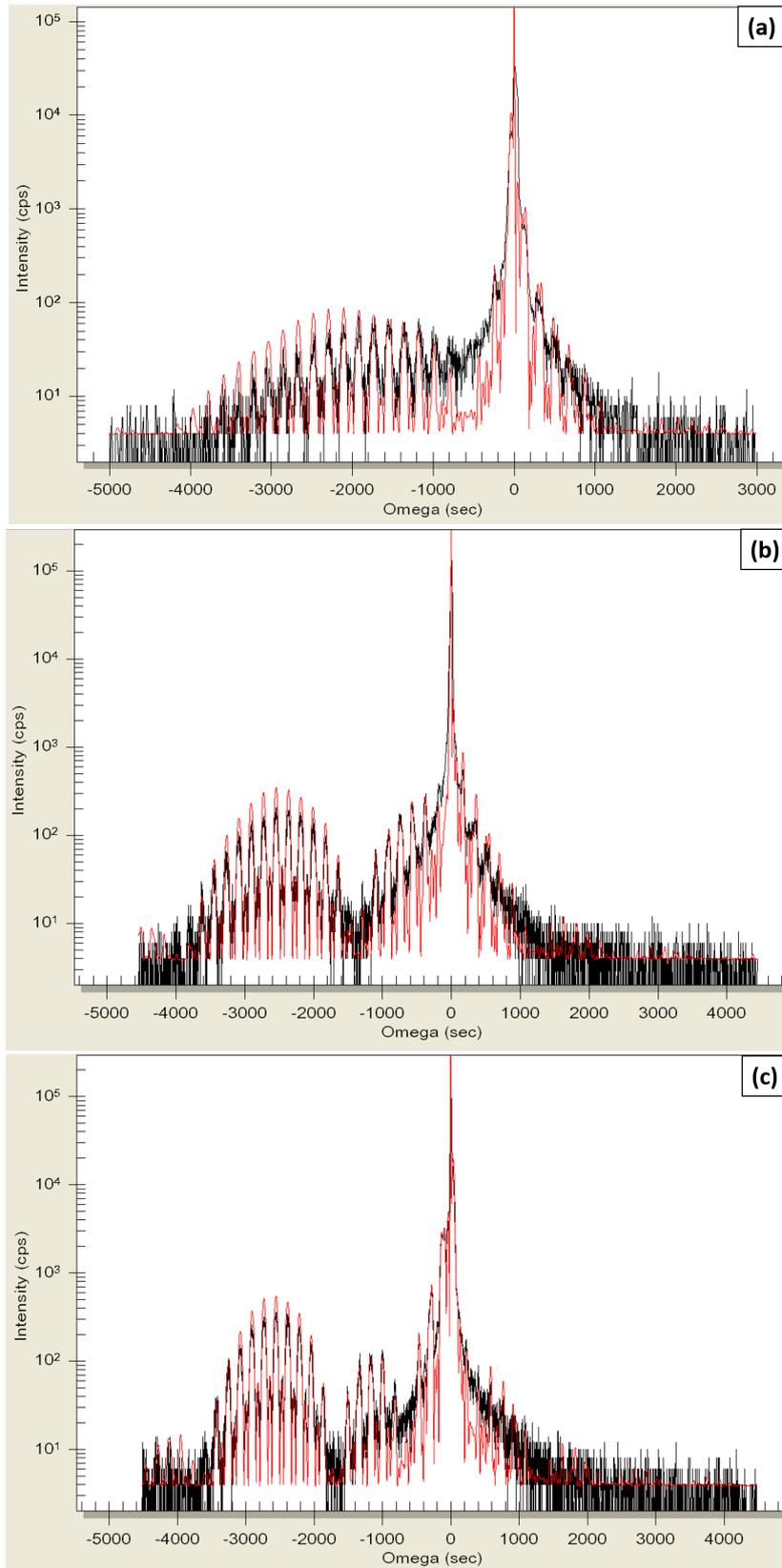
Figure 5.7: Rocking curve measurements are plotted (black) and overlaid with simulated data (red) for a triple-stack $In_{0.15}Ga_{0.85}As$/GaAs MQW of a) 10 nm, b) 15 nm, and c) 20 nm.

The rocking curve measurements of three symmetric $In_{0.15}Ga_{0.85}As$/GaAs MQW samples are presented in Figure 5.7. The fitting parameters corresponding to the

rocking curve measurements are extracted and tabulated in Table 5.1. Clearly, the model is in strong agreement with the measured data. High growth quality of the QW layers is inferred from the remarkable overlap of the fine, experimental and theoretical satellite peaks. It is worth emphasising that the model assumes a homogeneous distribution of In content throughout the QW. The negligible imperfection density is a result of MBE technology, and selection of layer thickness below h_c for an $In_{0.15}Ga_{0.85}As$ QW (Figure 4.1). Furthermore, the intensity of the satellite peaks increases with increasing QW thickness because the number of Bragg planes increases with thicker layers of material.

Sample ID	Nominal QW Thickness (Å)	Actual QW Thickness (Å)	Nominal x (%)	Actual x (%)
A1336	100.00	96.57	15.00	14.00
A1396	150.00	147.20	15.00	14.69
A1397	200.00	189.64	15.00	14.79

Table 5.1: Layer thicknesses and temperatures used.

A low amount of In segregation is evident for each sample since the measured In content is close to the intended values (Table 5.1). At most, the In concentration and thickness deficits were 1 % (10 nm sample) and 10.4 Å (20 nm sample), respectively. It is highly likely that these discrepancies are not due to segregation of In; rather, slight imprecision of cell temperature, V:III flux and growth times, along with the relaxation parameter used in the model. Considering the difficulty of epitaxy, the growth of these layers are considered successful and further recipes can be designed based upon these results.
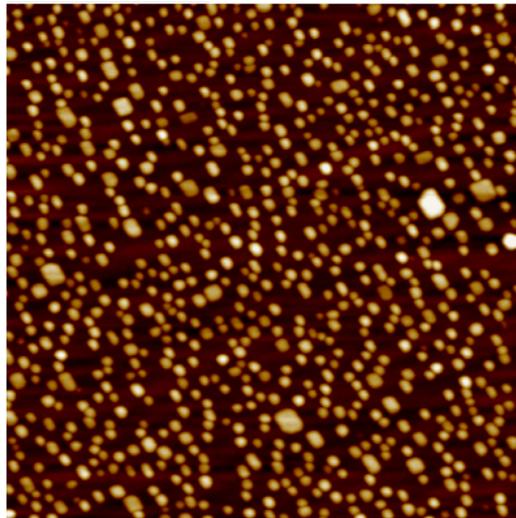
Figure 5.8: 1 μm × 1 μm micrograph of the reference sample GaSb surface QDs. The colour brightness reflects the relative height of the nanostructure.

QD morphology statistics were analysed and the results are presented in Figure 5.9. It is clear that the QD ensemble has an asymmetric normal distribution of diameters and heights. The histograms were therefore fitted with a BiGaussian, i.e. a combination of two half-Gaussians, to assess the mode average and skewness, ζ_S.
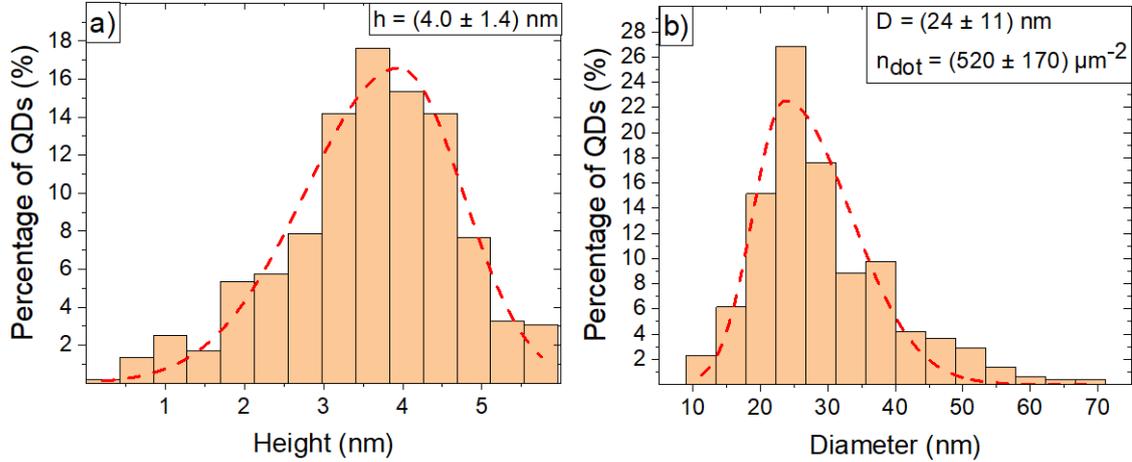
Figure 5.9: Histogram plots to show the a) height and b) diameter distributions of the ~ 522 sample size. a Bigaussian is used to fit the binned data (red dotted-line).

The dot density, $n_{dot} = (520 \pm 170)$ μm². The mode averages of the height and diameter are $h = (4.0 \pm 1.4)$ nm and $D = (24 \pm 11)$ nm respectively. The standard deviations of h and D, were calculated by the quadrature sum of the half-Gaussian widths above and below the mode average. The height distribution has a greater non-uniformity than the distribution of diameter. Specifically, the ζ_S-values of the resulting curve fits of h and D are 0.28782 and 0.77435 respectively. The buried QRs are expected to have an even greater non-uniformity of h, due to the dissolution effect.

The height distribution has a tail to the shorter QD side. The majority of QDs have a height of 3-5 nm, and $h < 3$nm for almost a quarter of their remainder. For the diameter distribution, the tail is towards the wider QDs. There are only several dots with $D > 50$ nm on the micrograph. The remaining counts are likely due to sampling errors in the Nextnano Analysis software, which are accounted for in the uncertainty of n_{dot}. Therefore D can be assumed to be uniformly-distributed (in approximation) around 24 nm. The variation of h and D will affect the in-plane and out-of-plane lattice constants of the local QW, respectively, along with the confinement energy of carriers. Thus the distribution tail of h is expected to have a greater impact on the transition energy than that of D.

Previous high resolution transmission electron microscopy (HRTEM) [47] and XSTM studies [132], show double-lobe features that are indicative of QR formation. No QRs were identified during the AFM measurements of this project. However, surface QRs

were not observed in the work of Refs. [47, 242] either, despite the clear formation of buried QRs. The buried structures are most likely similar to those in Ref. [195]. They are expected to be approximately circular [57, 81] and highly disordered [243]. HRTEM and XSTM seemed beyond the scope of this work since we can reasonably assume our nanostructures are consistent with previous studies by looking at the reference sample AFM. It is worth reminding that growth conditions are similar, well developed and found to yield similar morphology under varied conditions. Moreover, the λ-values here are very similar to in their work.
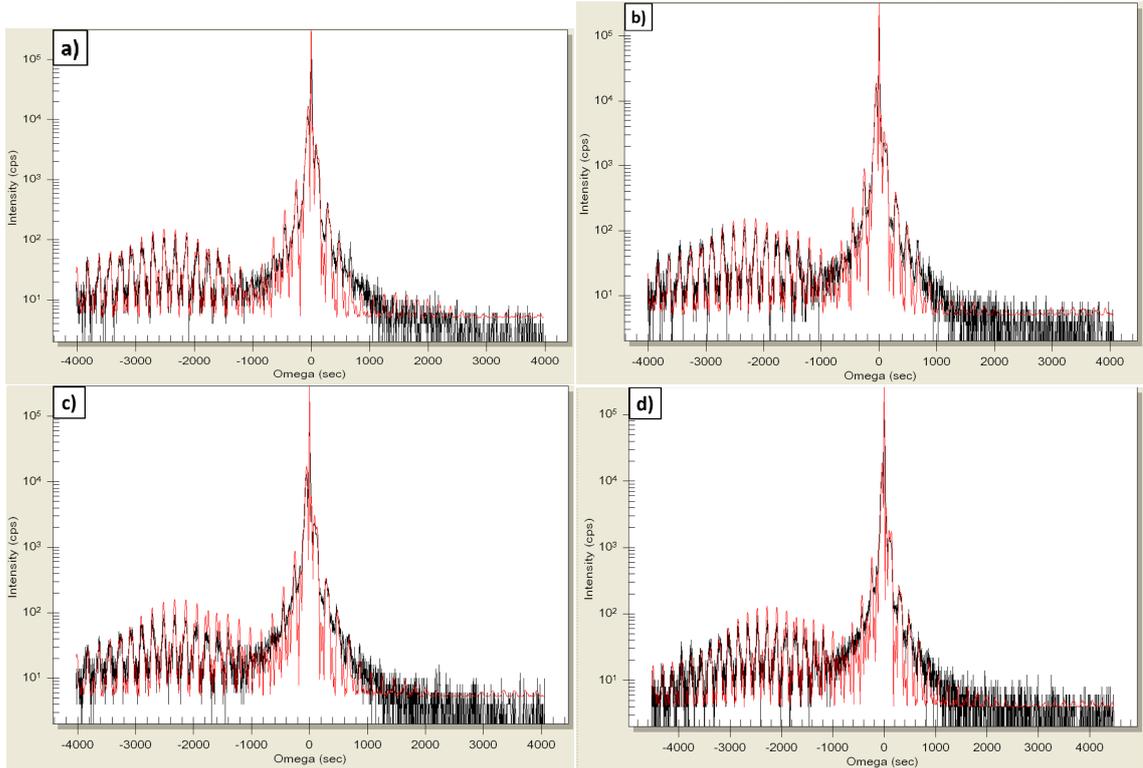
Figure 5.10: Rocking curve measurements are plotted (black) and overlaid with simulated data (red) for a triple-stack (symmetric) $In_{0.15}Ga_{0.85}As$/GaAs MQW of 10 nm width, grown at a T_{sub} setpoint of a) 520 °C, b) 560 °C, c) 600 °C, and d) 620 °C.

5.4 Asymmetric Ring-in-Well Photoluminescence

5.4.1 Peak Discernment
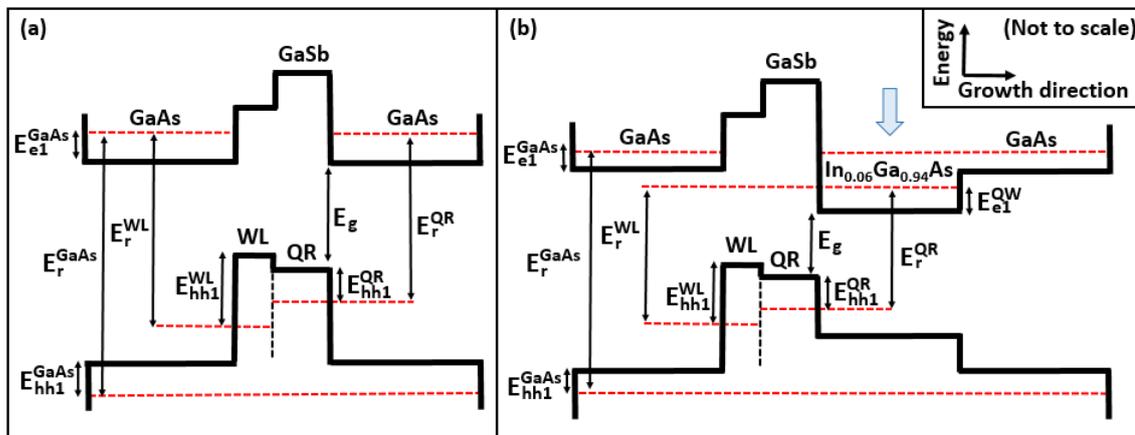
Figure 5.11: Bandgap diagrams of the previous (a) and developed (b) structures to show how the main radiative recombination channel has changed. The red dotted lines represent the ground-states, the blue arrow indicates the key difference between both structures; all other notation has its usual meaning. The relative GaSb VB offsets have been displayed taking into consideration the work of Pryor and Pistol [244]. The WL is assumed to be 100 % relaxed.

PL spectra were first measured at 4 K following the procedure outlined in Subsection 3.3.2, since the prominence of observed peaks are maximised in these conditions. Figure 5.11 shows the corresponding band diagram of the spectra shown in Figure 5.12. Flat band-edges are assumed for simplicity and highlight the pre-eminence of three distinct transition energies. The VB edge of the WL is raised above the QR band edge [244], and the QW layer is in contact with the WL in regions between the nanostructures and within their inner radii.

Figure 5.12 compares all four samples at a large excitation power of $P \approx 260$ mW. Here, the data has been normalised to the GaAs peak to better compare the performance of each sample. Notably, the spectra has not been corrected according to Subsection 3.3.5, so the reader can assess the small differences between original and modified spectra.

Three dominant resolved bands of emission are apparent for each sample, in agreement with Figure 5.11. The peaks close to 1.35 eV and 1.5 eV are attributed to emission from the QDs/QRs; WL and the GaAs matrix, respectively. The dominant peak near 1.05 eV is yet to be discerned with confidence although it is most likely to be emission from either QDs/QRs or CB to C-acceptors. Emission from the QW BtB transition is unobservable due to both the Coulomb attraction from the QRs and their large hole confinement energy (~ 600 meV for pure GaSb/GaAs QRs [29, 80, 86].
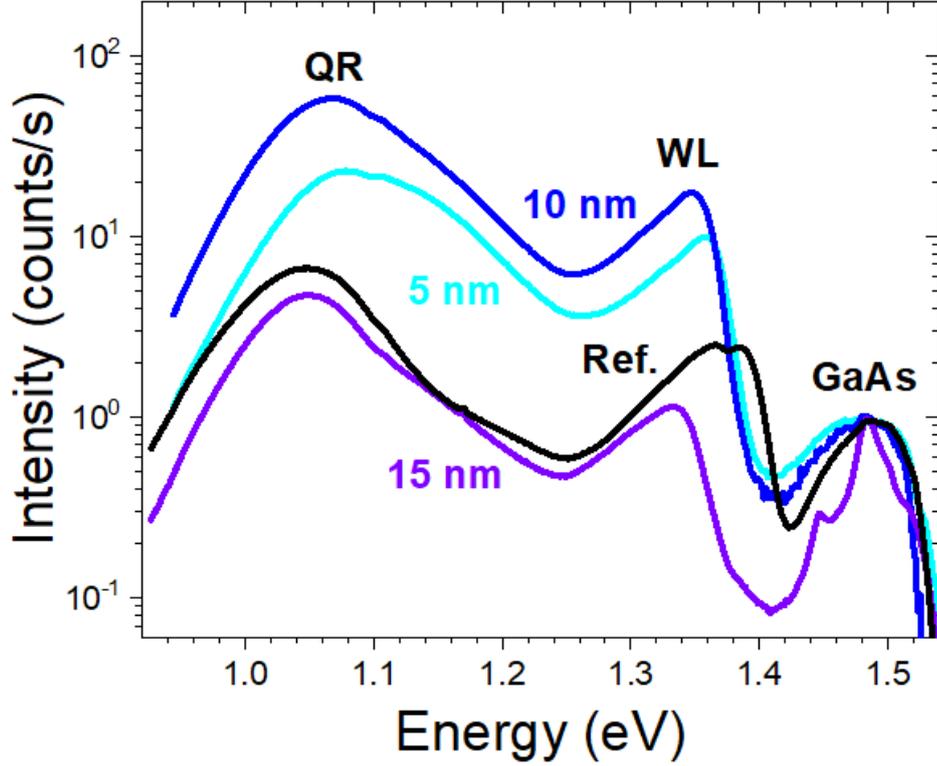
Figure 5.12: 4 K PL spectra comparison of the reference and RWELL samples, at full P.

The reported bandgap value of GaAs is 1.519 eV at 0 K and corresponds exactly to the barely resolved shoulder of the 15 nm GaAs sample, therefore this shoulder is attributed to the BtB transition of bulk GaAs. The ground-state E_b (free-to-bound transition) of C-acceptors in GaAs is 26 meV [245]. Therefore the majority of emission from the observed GaAs peak coincides well with C-acceptor states above the GaAs VB edge, despite the 15 nm sample showing emission at ~ 1.85 eV (8 meV difference may be due to excitonic binding and systematic uncertainty). The fact that acceptor states dominate emission at these λ-values implies a significant amount of accidental doping near the nanostructures where the electrons are concentrated, i.e. in the (In)GaAs cold-cap and the ultra-thin GaAs protective layer. Unintentional doping is surprising since MBE was used, and the characteristic "U"-shape dependence of E_{PL} with P [22, 158] was unobserved for the WL.

The small peak on the 15 nm sample at ~ 1.45 eV is rather interesting and provides a clue towards discernment of E_a in Subsection 5.4.3, since it is ~ 68 meV, and ~ 37 meV, lower in energy than the GaAs peak and C-doped GaAs peak respectively. It is likely caused by a QW BtB transition; or a type-II transition between electrons in the QW ground-state, and acceptor states in the GaAs barrier or protective layer. It is more likely to be a transition between the QW and the GaAs barrier, since the peak almost disappears at low P. The smaller density of hole states in the ultra-thin

protective GaAs layer would be most discernible at low P and become saturated as P increases.

This data (and the data in Figure 5.14), shows that as the width of the $In_xGa_{1-x}As$ QW increases from 5 nm to 15 nm, the PL emission redshifts by ~ 32 meV and the intensity (peak amplitude) decreases by nearly two orders of magnitude. The WL also redshifts by ~ 24 meV. The WL E_{PL} of each sample changes with confinement suggesting one of two possibilities. Either the WL composition is similar across the RWELL samples and tensile strain is only partially relaxed and lowered the CB edge; or strain-driven intermixing of In with increasing QW thickness has changed the layer composition. The former is most likely.

The shoulder near (~ 1.2 eV) on all QR peaks may indicate a significant background PL contribution since it is more prominent at low P. The background emission has a large Γ_{PL} near this energy, at (1.15 ± 0.2) eV. However, it is very unlikely that background emission is visible at full P and LT. Emission of ~ 1.15 eV from GaSb/GaAs QDs has been observed at 10 K [28], therefore depending on whether QRs were grown successfully, the dominant peak can be explained by 1 of 4 structure-related possibilities described below.

The first possibility is the transition from excited states of electrons at higher energy states in the (In)GaAs layer, to holes deeper in the QR VB ($e_2 - HH_2$). Secondly, although the AFM data shown in Figure 5.9 lacks strong evidence to consider a bimodal distribution, the active region still appears to contain a significant asymmetric distribution of nanostructure morphology. The variation of local QW strain must be non-negligible, since there is a slight dependence of the shoulder peak energy on QW-width, however, it is conjectured that shoulder for this case is more likely caused by band-filling and the subsequent optical activation of states within the smaller nanostructures (with increased carrier confinement). Wide QDs typically contain defects which cause non-radiative losses. However strikingly, an asymmetrical distribution of QDs was found in Ref. [246] to be optically-active even though the mode averages were found to be $D = 60$ nm and $h = 6$ nm respectively. These nanostructures have similar dimensions to the large nanostructures of the reference sample of this work. The third explanation is that tensile strain and confinement inside the GaAs regions of the inner QRs may change the electronic energy levels of states in regions outside of the QRs relative to the bound states within their centres. This effect is expected to be insignificant due to the rapid filling of a small DoS inside the QR centre. Alternatively, emission may be dominated by holes bound to C-acceptors ($e_1 - A^0$), with the higher energy peak corresponding to the ($e_1 - HH_1$) transition. If this was the case, one would expect the PL sub-peaks to be separated

by ~ 26 meV and this is not the case even with a generous approximation for the sub-peak energies.

There is a high-energy shoulder on the RWELL samples that is most pronounced on the 15 nm sample. The strain and band offset analysis in Chapter 4 also indicates the existence of a radiative dual-channel, i.e., the observation of an average of two recombination channels emitting at similar energies. This potentially occurs in one of two ways. A strain-related splitting of the intended (In)GaAs-to-GaSb channel transition energy could be caused by the recombination of electrons from InGaAs regions inside and outside of the QRs; or large P may lead to radiative recombination from the combined QW and GaAs barrier CB ground-states. If a radiative dual-channel is operational, it can be argued that the magnitude of relative intensity is a result of the differing surface areas at the inner and outer QR regions; or in the previous case, the transition rates between regions of differing material. An accurate value for the difference in CB edges between electrons situated inside the QR centre and outside of them cannot be calculated yet, due to the lack of lateral confinement with the current model and is left for future simulations. Alternatively, the filling of GaAs states and lowest energy states inside the QW layer may be simultaneous at sufficiently large P due to their small relative CB offset. The increase of the shoulder intensity relative to the QR peak corresponds to a greater proportion of electrons recombining from the QW. Other less plausible explanations may still account for the observed phenomenon in the RWELL samples. This may provide evidence that a thin $In_xGa_xAs_ySb_{1-y}$ dilute quaternary top-coating exists around the RWELL nanostructures in the QW due to QR formation. Increased strain would drive a greater amount of inter-diffusion. Although the asymmetry on the high energy slope is often caused by energy transfer between QRs of different sizes, this explanation cannot be justified without first assessing the relative intensities against P. The same applies to the possible occupation of excited states. In the case of the former, ground-state recombination dominates and the increased hole-hole Coulomb charging energy corresponds to the sub-peak separation. If the occupation of excited states is true, then the higher energy peak would exhibit an increasingly super-linear P dependence caused by the rapid occupation of higher energy states. Excited states would not be significant enough to be resolved since it would require the nanostructures to trap unlikely large densities of holes. A Mott transition may also occur at higher P; free carriers recombine at higher energy than when they are bound.

All WL peaks also have an asymmetric lineshape (low energy feature). It is unusual for a band edge to be abrupt. Long, low-energy and T-dependent tails are referred to as Urbach tails [247]. They often appear near the optical band edge because

135

of fluctuations in the electrostatic potential, with a dependence on disorder and impurities. It reflects the exponential part of the absorption coefficient spectrum as it falls to zero; i.e. local states are shunted below the CB edge, hence recombination decays exponentially into the bandgap. The dilution of the WL along the growth direction due to As-Sb exchange [57] may further account for the observed asymmetries. The reference sample should have slightly better structural quality than the In-containing RWELL samples, however similar tails are observed across all samples, thus indicate similar structural quality near the QRs in each of them.

The resolved double-peak of the reference is unlikely to be caused by ML fluctuations of WL thickness, i.e. hole localisation at the WL interface from islands/terraces formed by $1-2$ ML fluctuations of the WL thickness. All samples would exhibit the same phenomenon if this were the case. Clearly, any sub-peaks are restricted by the resultant WL profile and must be intrinsically different in shape. The peak cannot be due to potential C-acceptors either. With regards to this, the lower energy peak would be attributed to the localised states, and the higher energy peak would be the GaAsSb states. The higher energy peak increases in intensity, and the lower energy peak is visible at low P. At low P the typical acceptor occupancy is negligible. At full P the density of photo-generated carriers is greater than density of acceptor states, thus emission from them is unobservable due to their saturation.

Charged excitons, namely X^- trions, are characterised by the convolution of a hyperbolic secant profile, and an exponential tail function. The former function has heavier tails than a Gaussian and is related to phonon scattering, the latter accounts for recoil electrons left behind after recombination [248]. Stronger electron confinement increases the electron density near the WL in the RWELL samples, thus increases the probability for charged excitons to exist. A similar low energy tail is actually observed across all samples and observation of excitonic emission without a μ-PL measurement setup is dubious. That is not to completely disregard the possibility of observing the broad PL band with LT macro-PL as a composition of sharp lines [249]. Since electronic confinement is poorer in the reference sample and the measurement of PL is at high P; excitons are unlikely responsible for the WL tail.

An unresolved LO-phonon replica provides an alternative explanation to the asymmetry of the low energy slope. These replica peaks are caused by quantum confinement and piezoelectric effects in low-symmetry nanostructures. An emission line attributed to them was found to occur at ~ 31 meV below the WL in GaSb/GaAs QDs [28]. This result is supported by the work of Ref. [250], where Raman spectroscopy was used to obtain a phonon energy of ~ 29 meV in GaSb QDs. A close

inspection of all the raw sample data reveals a slight shoulder on the low-energy side of the broadened transition line. An approximate energy difference of $\sim (30 \pm 5)$ meV is obtained for a possible sub-peak fit in strong agreement with reported values. Phonon-assisted transition phenomena are most identifiable in the reference sample.

The ground-state emission from the RWELL structures studied here have a smaller E_{PL} than the sample in the work of Ref. [59]. The authors report a QR peak λ of ~ 1200 nm (1.035 eV) at 10 K with an excitation density of 0.5 W/cm^2. This is striking in comparison to the results presented here, considering the much lower nominal In concentration and large excitation power used. It is likely that the QRs in this work are bigger. Notably, the QR and WL peaks in this work are more clearly resolved, indicating improved growth conditions. The relative difference of intensities between these peaks is similar to in their work. Typical Γ_{PL}-values reported in literature for the ground-state of GaSb/GaAs QRs are typically below 100 nm (at low T and P) [54, 57, 81].

The remainder of this section focuses on the T- and P- dependencies of E_{PL}, Γ_{PL}, and I_{PL}.

5.4.2 Power-Dependent Measurements

PL from the InGaAs-capped GaSb/GaAs QRs was measured at 4 K whilst varying P. Figures 5.13a to 5.13d show the evolution of the two lowest energy peaks with P over a range of 0.07 mW to 260 mW. A single Gaussian fitting was used to extract E_{PL}, Γ_{PL}, and I_{PL}. Studying the evolution of integrated intensity rather than peak amplitude is preferable, due to less sensitivity to fibre orientation (illuminated surface area) and slight changes to the local nanostructure density across different regions. Indeed, a two-peak model would fit the data better for these samples. However, it was considered inappropriate for employment without a precise grasp of the Γ_{PL} and E_{PL}-values of the other emitters (and potentially background), despite a sensible evolution of the high energy feature with P. The uncertainty in the resulting fitted parameters is too large and accompanied by a risk of rationalising the data. Confident deconvolution and identification of these sub-peaks is further prevented by the small difference in energies. Instead, it is better to compare the average ground-state emission with a single Gaussian across samples as a figure of merit. Moreover, it is very likely that discrepancies between measured and simulated spectra are partly caused by band-tail states away from resonant peak energies.
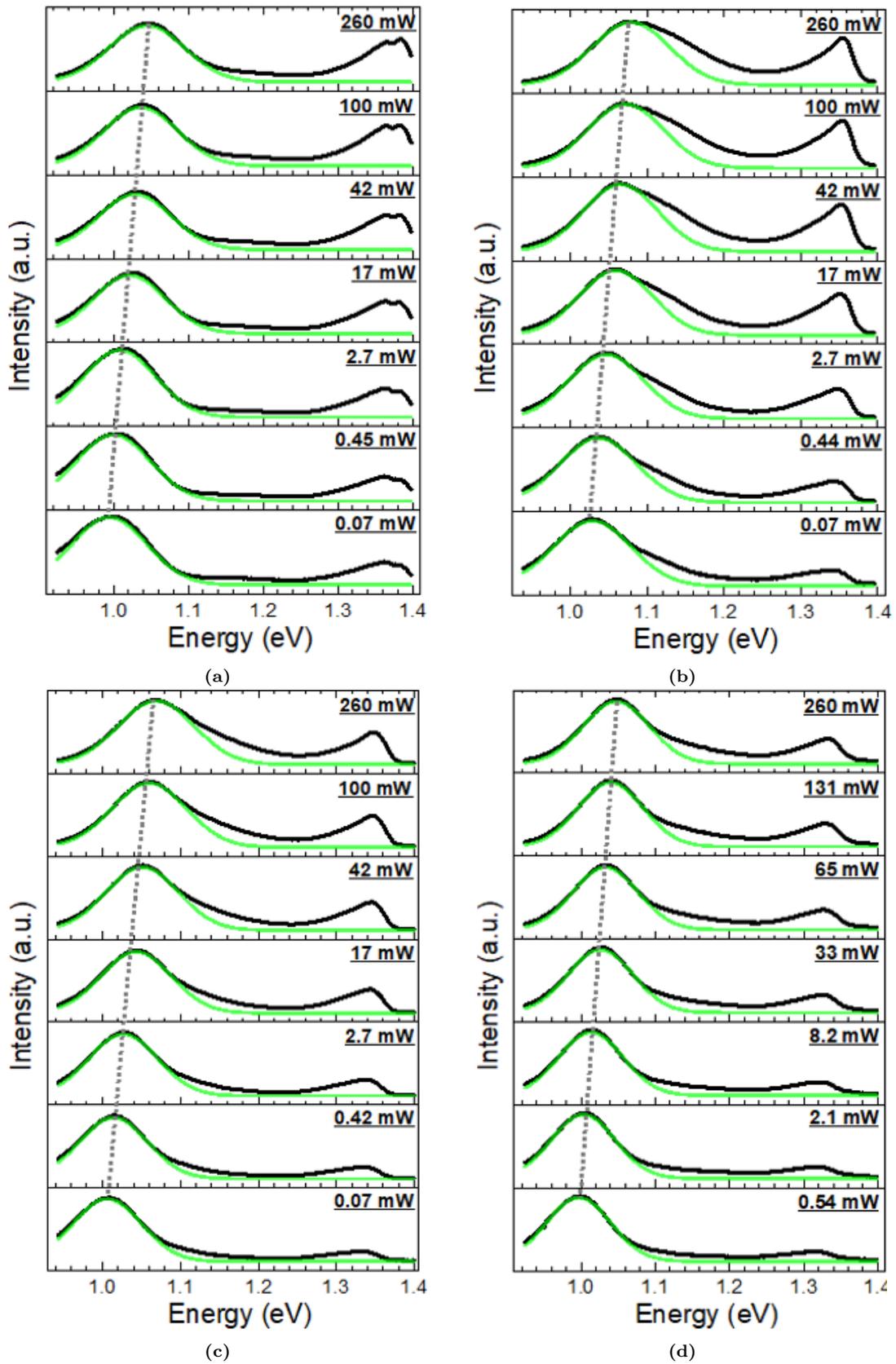
Figure 5.13: *P* sweeps: a) ref, b) 5 nm, c) 10 nm, d) 15 nm.

The single simulated Gaussian curve misfit of the high energy tail (QR emission peak) is the least discernible below ~ 1 mW and grows in intensity with P. The

discrepancy appears to be inversely proportional to QW width, thus related to the behaviour of electrons and their confinement, rather than the variation of hole confinement energy from an asymmetric distribution of QR shapes and sizes. Moreover, the strain used to form the nanostructures should be similar across all RWELL samples due to annealing after 5 nm of cold-cap during each growth. It is reasonable to suggest that an ensemble of many overlapping states arises due to fluctuations in local strain, hence why the peaks remain blurred together at high P. Recombination from biexcitons is ruled out since the approximate low energy sub-peak does not appear to reduce in I_{PL} with P, despite the potential higher energy sub-peak appearing to grow in intensity. As it stands, predominant recombination from e-h pairs cannot yet be ruled out. Broadening of the high energy tail and Γ_{PL} at high P for each of the RWELL samples in Figure 5.13 is a strong indication of excitonic resonance losses [24]. This is typically caused by many-body effects from a large electron density, where $\mathrm{e}^- - \mathrm{e}^-$ interactions, i.e. screening, dominate $\mathrm{e}^- - \mathrm{h}^+$ interactions. The possible overlap between neighbouring excitons occurs in samples where the average separation of randomly distributed QRs is small (approximately twice the Bohr radius). Excitonic ionisation can be increased by the additional screening from de-localised electrons [24]. The asymmetry cannot be from electrons confined within unintended In conglomerations either, since the emission peak would be at a longer λ.

The intensity of the WL peak increases relative to the QR peak with increasing P for all samples. This indicates an increasing rate of recombination in the WL. Large P-values fill the vacant CB states and a greater percentage of electrons transition to holes in the WL, as the QR states become saturated by the photo-generated carrier density. It is energetically unfavourable, in the RWELL samples at low P, for the electrons to recombine with holes in the WL due to strain inside the QW and the strong Coulomb potential of the QRs. A reduction of QW confinement reduces the relative WL peak intensity in the RWELL samples, which suggests a lower electron density near the nanostructures.

Figure 5.14 shows the P-dependence of E_{PL} for each sample. Recombination energy increases with increasing P in each sample. It is widely accepted that these strong blueshifts are a key PL characteristic from type-II structures. The mechanisms commonly invoked to explain this behaviour are capacitive charging (CC) [158, 161, 170, 239, 251], BB [76] and state-filling [252]. The primary cause of blueshift in GaSb/GaAs QR structures is CC and this was found to be an order of magnitude larger than the contribution from BB [170].
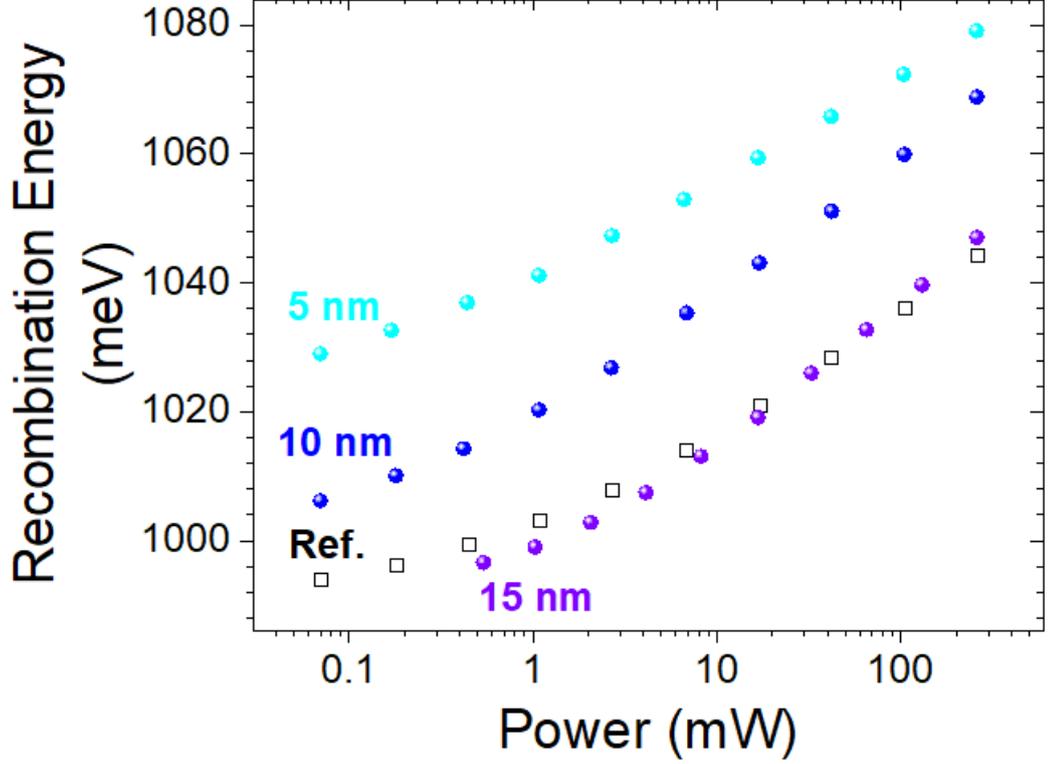
Figure 5.14: Recombination energy of the QR peak in each sample as a function of P. Each value has been thermally-corrected $(-k_{\mathrm{B}}T/2)$.

The internal electric field caused by the spatial separation of carriers creates a strain-dependent, triangular potential well at the interface between the nanostructures and the local matrix [253]. This causes the ground-state energy to shift proportionally to the electron density near the interface, $n_{2\mathrm{D}}$ [254]:

$$\Delta E_{\mathrm{BB}} = \left(\frac{\hbar^2}{2m_{\mathrm{e}}^*} \right)^{1/3} \left(\frac{18\pi^2 e^2}{8\varepsilon_0\varepsilon_{\mathrm{r}}} \right)^{2/3} n_{2\mathrm{D}}^{2/3}. \tag{5.2}$$

Importantly, the 0D nature of QDs and QRs must be taken into account. Charging effects are included by modelling the nanostructures as a nano-scale capacitor, where the addition of each subsequent charge results in an energy shift proportional to the number of holes trapped inside the nanostructure, n_{h} [158, 161, 170, 251]:

$$\Delta E_{\mathrm{CC}} = \frac{n_{\mathrm{h}}e^2}{C} = \frac{e^2 a_{\mathrm{X}}}{2A\varepsilon_0\varepsilon_{\mathrm{r}}} n_{\mathrm{h}}. \tag{5.3}$$

Where C is the total capacitance, A is the area of the nanostructure, and n_{h} is the number of holes trapped inside the nanostructure. All other symbols have their usual meanings. The electron and hole densities are often unknown, so they are related to the measured I_{PL}, via the bimolecular recombination equation previously discussed in Subsection 2.3.3. Charge neutrality is assumed $(n_{\mathrm{e}} = n_{\mathrm{h}})$ and leads to

140

$\Delta E_{\mathrm{BB}} \propto I_{\mathrm{PL}}{}^{1/3}$ and $\Delta E_{\mathrm{CC}} \propto I_{\mathrm{PL}}{}^{1/2}$ for BB and CC, respectively.

There is a ~ 36 meV difference in E_{PL} between the thickest and thinnest QW samples. Such an increase in energy with a reduction in QW width is consistent with quantum confinement and strain arguments. The energy of discrete hole (electron) states in the GaSb nanostructures ($\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}$ QW) are increased (lowered) by compressive (tensile) strain. Moreover, tensile strain tilts the bands, thus steepening the triangular potential well [28].

Counter-intuitively, the QR peaks have blueshifted away from the reference QR energy. Below ~ 30 meV, the 15 nm sample has a slightly longer λ than the reference sample, hence provides optimism for one of the main project objectives. However, the difference is only slight and at an impractically low P for industry. Further increases to the QW width may increase λ at the expense of poorer I_{PL} yield, due to inevitable dislocations if the critical thickness is exceeded. An important question is therefore raised as to why the QW has not significantly redshifted QR emission as expected. A few potential reasons are discussed below.

The most likely explanation is that confinement effects dominate the tensile strain-related reduction of the CB edge. Quantum confinement likely causes the majority of blueshift by increasing E_{g} and E_{b}. Electron interactions also gain significance, as the z- dimension is reduced, thus counteract $\mathrm{e}^- - \mathrm{h}^+$ interactions with increasing P; effectively increasing a_{X} [24] and contributing to blueshift. The simulation results of Chapter 4 (Subsection 4.2.3) suggest that it may be energetically favourable, at particular locations were the piezoelectric charge density is significant, for GaAs electrons to tunnel through the WL and recombine with the GaSb holes. Transition energies are inherently larger, since the GaAs states below the QRs are further spatially separated from the GaSb states in the RWELL samples, than GaAs states above the QRs in the reference. With regards to the VB, compressive strain inside the QDs/QRs is expected to push holes to lower energies. However, both confinement and strain energies inside the QRs will be similar for all samples because only the cap layer was changed between them, and even though InGaAs is closer in lattice constant to GaSb, than GaAs is, the dilute QW will reduce the compressive strain in the QRs by a negligible amount. In-situ thermal annealing is known to increase the GaSb bandgap [170], although conditions were similar across all samples, thus any related differences are negligible.

Different effective masses between the RWELL samples may also explain the observed blueshift dependence on QW-width. The tensile strain inside the QW reduces the effective mass [255] which spreads the electronic wavefunction (increases a_{X}) and compresses the electron dispersion, i.e. increases CB curvature far from the

Γ-point. This can be observed by investigating the parabolicity of the field dependence of the QR peak at low excitation [47]. Clearly any subtle increase in electron energy levels across the entire band does not adequately give the full picture at the Γ-point. Otherwise, by these arguments the 15 nm sample ought to emit the largest energy photons, since strain increases with increasing QW width. Besides, the tensile strain-induced reduction of effective mass is opposed by the increase of effective mass due to the penetration of the electronic wavefunction into the barrier (which increases with greater confinement), and non-parabolicity of the CB at large k-values (which reduces with confinement) [256]. Importantly, the carrier mobility depends on the net effective mass and impacts the significance of screening, such that small E-fields can easily capture electrons if there is an overall reduction of m_{e}^{*}; resulting in a deeper potential well at the heterojunction, i.e. more CC.

The slightly increased rate of blueshift for the RWELL samples above a critical excitation may possibly indicate the band-filling, and subsequent saturation, of a limited DoS inside the inner region of the QRs. However this is unlikely since the emission centre of mass (COM) hardly changes in Figure 5.13 and a small DoS of states will be eclipsed by recombination from regions in between adjacent QRs with a larger DoS. The strain-induced CB profile is a more likely explanation.

Increasing P will increase I_{PL}, due to greater QD/QR occupancy, by the relationship: $I_{\mathrm{PL}} = \eta P^{\aleph}$ (assuming $P \propto I_{\mathrm{ls}}$, where P is the nominal laser power). The exponent, $\aleph$, relates to the mechanism of recombination. Values of 1 and 2 typically indicate excitonic and free carrier recombination, respectively. A lowered $\aleph$-value is desirable for the objectives of this project, as it corresponds to increased recombination rates. Emission is therefore brighter at resonant λ-values, and the rate of blueshift is dampened (stabilisation of λ), i.e. higher energy states are kept unoccupied due to shorter carrier lifetimes. A more complicated carrier recombination mechanism is evident if $\aleph$ increases at higher T. The intercept, $\mathrm{Log}(\eta)$, of a linear fit of $\mathrm{Log}(I_{\mathrm{PL}})$ as a function of $\mathrm{Log}(P)$, is a comprehensive characteristic related to nanostructure efficiency. It includes effects due to geometry, carrier capture, absorption, and ionization. A lower value for η is associated with poorer recombination efficiency.

The P-varied integrated intensities of each sample at 4 K, along with their corresponding $\aleph$-values are presented in Figure 5.15. The propagated uncertainties taken from the least-squares fit are omitted due to the data points being much larger than the error bars. Moreover, the uncertainties introduced during analysis are much smaller than the uncertainties introduced during growth, particularly when discussing the resulting ensemble of nanostructures.
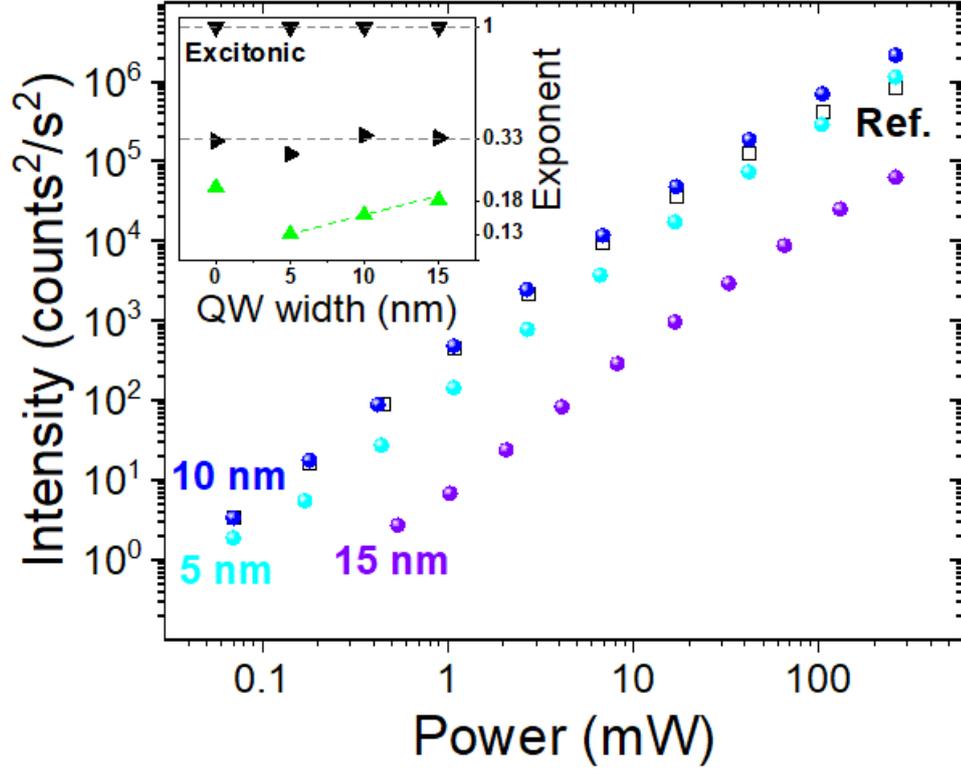
Figure 5.15: Integrated intensity of the QR peak in each sample as a function of P. Inset: the ℵ-values are plotted as a function of QW width, where the black and green triangles correspond to low and high P, respectively. Excitonic recombination dominates at low P and T.

The data in Figure 5.15 shows that $I_{\mathrm{PL}} \propto P$ and increases linearly below ~ 30 meV. Intuitively, a higher intensity of incident photons will create greater densities of electrons in the QW layer and populate the QRs with more holes, thus recombination rates will increase subject to nanostructure hole occupancy. The probability for a recombination event to occur will also increase with a reduced carrier separation.

The relatively low I_{PL} of the 15 nm sample indicates a greater amount of non-radiative recombination and consequently implies one of two things. A greater percentage of injected carriers is likely to populate the barrier layers and recombine more slowly there due to a relatively small carrier volume density.

Differential carrier-lifetime measurements in strained InGaAs/GaAs QWs confirm that recombination rates are in fact reduced in QWs of thinner QW width or smaller potential barriers [257]. Or this sample may have a greater defect density than the other samples, i.e. the number of band-tail states has increased as the QW approached the critical thickness. A consequent relaxation of the internal QW strain will facilitate carrier loss with increasing T through the thermal activation of dislocation trap states.

Discrete hole-charging can be inferred from increasingly sub-linear behaviour at higher excitation. Likewise, a sub-linear dependence can be explained by non-

radiative loss into the barriers [169]. The exact balance of mechanisms causing this behaviour is unclear as of yet. However, discrete hole charging is likely to account for most of this behaviour at LT.

Importantly, the inset of Figure 5.15 provides new evidence to confirm that the widely used approximation to the bimolecular rate equation cannot be used to discern between CC or BB for 0D confined type-II nanostructures. Moreover, it shows that the rate of blueshift depends on an average of separate contributions. At full P, there is a clear dependence of $\aleph$ on QW width, i.e. $\aleph$ increases for thicker QWs; therefore, the observed increase of I_{PL} saturation with QW width is related to the localisation of electrons. The weakened electron binding causes a reduction of the rate of radiative QW-to-QR transitions, which suggests that more electrons are de-exciting to states in unintended regions of the structure. As a result, the sub-linear behaviour of $I_{\mathrm{PL}}(P)$ intensifies.

Indeed, charging energy is a highly important factor and found [170] to dominate BB by an approximate factor of 10; though these QW-varied results prove that BB cannot be ignored completely. Besides, an upper limit to the amount of CC-related blueshift depends on the self-limiting number of holes that the average QR can contain [26]. Conceivably, the observed blueshift is a compound effect: the sum of contributions from CC, BB, changes in strain, and energy transfer. Nevertheless, it is concluded that the type-II nature of GaSb/GaAs QRs and InGaAs/GaSb/GaAs RWELLs truly leads to a non-linear rate of carrier transfer beyond a critical excitation density.

The generalised analysis outlined in Ref. [170] was employed here in attempt to distinguish the main mechanism responsible for recombination; without rationalising the data with prior assumptions of $\aleph$. First, the minimum transition energy was obtained by extrapolating a linear fit to the lowest excitation data points in a plot of E_{PL} as a function of I_{PL}. The minimum transition energy is the energy difference, at $P = 0$, between the ground-state and the CB edge. This value includes any accidental hole charging [158, 239]. The intercepts are then subtracted from the thermally-corrected (and P-varied) QR emission energies to obtain the net ground state energy shift, ΔE. The $\aleph$-value was then found by applying a linear fit to the trends in a plot of $\log(\Delta E)$ vs $\log(I_{\mathrm{PL}})$.

Evidently, $\aleph$ evolves with increasing P, hence a single value for $\aleph$ cannot be extracted. Three distinct $\aleph$-values are derived, due to the change of gradient between low P and high P. A gradient of 1 is observed at low laser power. As P increases, the exponents reduce to a value around 0.33, and then even further at full P, thus consistent with reduced carrier transfer into the QD/QRs. All $\aleph$-values are well
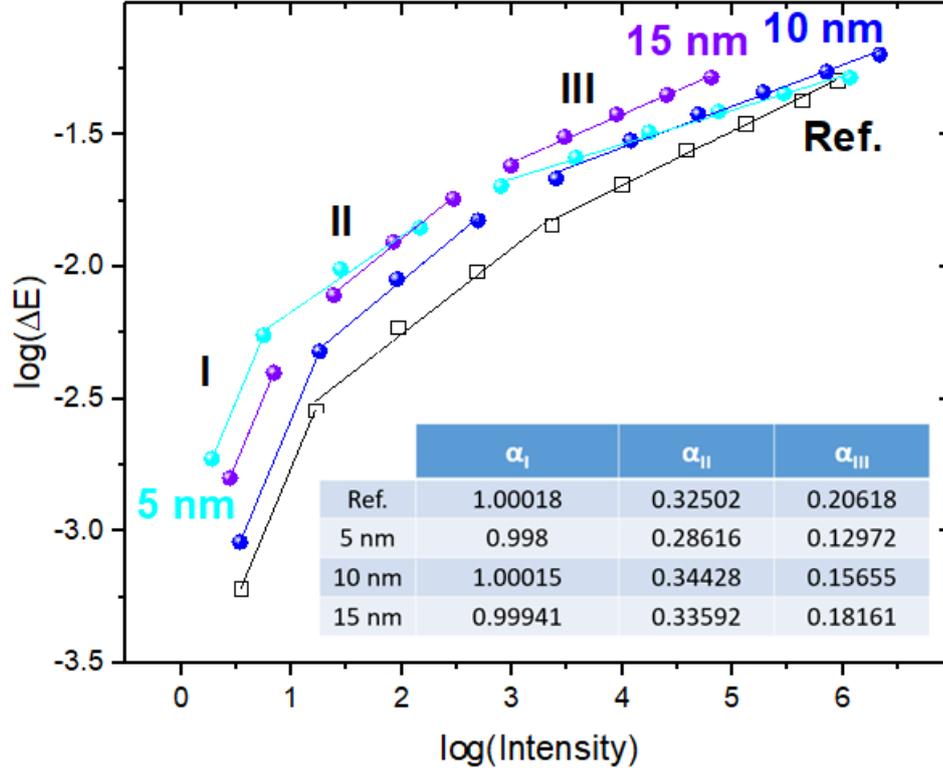
Figure 5.16: Plot of Log(ΔE) vs Log(I_{PL}) for the QR peaks of each sample. There appears to be three gradients corresponding to three distinct ranges of P. The inclusion of data points in each excitation range is not consistent for each sample due to poor data resolution, i.e. sparse measurements. Some gradients appear to meet at a common data point, where others do not. It does not affect the conclusions made since the gradients are used as a rough estimation for each range. As more measurement curves are taken (and accurately fitted), the corresponding exponents are likely to converge to a value in agreement with theory.

below 2, thus free carrier recombination is negligible in these samples at LT, and a Mott transition causing the higher energy shoulder can be ruled out.

At full P, $\aleph$ is greater for the reference than it is for the RWELL samples. This is most likely due to a_{X} and thickness of the cap layer. The ground-state of the RWELL samples is less prone to defects since the electronic wavefunction encounters a limited number of states in the thin QW region, in comparison with the vast number of states in bulk GaAs layers of the reference. Thus, the severity of sub-linear behaviour would be greater in the reference in agreement with observation. SRH transitions are typically characterised by $\aleph > 1$. Although this provides enough reason for an amount of scepticism in this explanation, SRH recombination cannot be ruled out completely since $\aleph$ is an average value corresponding to multiple simultaneous recombination processes across the entire QR ensemble. Alternatively, the reduced sub-linear behaviour with reduced QW width may be explained by suppression of CHCC recombination with confinement. Relative to thicker QWs; there are fewer vacant final Auger states, in the 5 nm QW, available for an electron to scatter into. In this way, radiative recombination rates would appear to increase as the QW width

is reduced. The estimated CHCC energies for each sample are tabulated below and do not match the exponents, hence this is ruled out.

It is argued that 0D confined type-II structures cannot be described by the current overly-simplistic recombination model, where contributions to blueshift are treated independently and often attributed as either one or the other. To clarify, greater electron confinement leads to enhanced electronic density near the QRs and attracts more holes into the QRs. This will increase the local inhomogeneous $\boldsymbol{E}$-field and assist the further capture of injected electrons, thus the steepening of the potential well is highly non-linear. A new model is required, which takes into account the total energy shift due to all mechanisms contributing to blueshift, along with strain. In the case of disordered samples, it may be necessary to include an extra small correction term related to state-filling.

The authors of Ref. [170] suspect that the charge neutrality condition does not hold; it is extremely likely that a new form of the equation is necessary to describe these structures accurately. Such a notion is supported by the suggestion made in Ref. [258], where it is posited that the coefficient B ought to include a carrier density dependence, i.e. $1/\tau = B(n)n^2$. Their results are for RT, where the carrier population follows a Fermi-Dirac distribution. The authors note that the carrier distribution cannot be accurately described by the Fermi function at LT, hence they advise to proceed with caution during the application of this dependence. Nevertheless, assuming a cubic dependence on carrier concentration as zeroth-order approximation seems to be in agreement with the mid-range excitation measurements here, thus CC is very likely dominant at RT in the majority of published articles which incorrectly attribute blueshift to BB. Therefore, correcting for the LT dependence of the carrier distribution is a good starting point for future theoretical development.

The charge neutrality assumption is invalid, and the resulting assumption that $I_{\mathrm{PL}} \propto n_{\mathrm{h}}^2$ is proven to inadequately describe the electronic behaviour for these samples for two reasons. The first being the observation of a clear dependence on $\Delta\psi$ at the heterojunction. This caused $\aleph$ to vary, at high P, well below the traditional BB threshold. A more complicated, or perhaps subtle, blueshift mechanism related to the behaviour of electrons must therefore exist at large excitations. Furthermore, the simulated structures of Chapter 4 reveal a consistent and significant reduction in transition energy than what was experimentally measured. Charge neutrality was assumed for the Poisson equation, and electron-electron interactions were ignored. Relaxation of these assumptions, along with additional consideration of multiply-charged QRs and slight changes to the QR dimensions is highly likely to place

the calculated transition energies in excellent agreement with measured PL. This work supports previous conclusions that the bimolecular recombination rate is often applied inappropriately to this type of nanostructure. The B coefficient is incorrect and does in fact need refinement to include a dependence on carrier density, although this is currently unknown at LT. Hence, the existing model can only be used to accurately describe QWs.

In Figure 5.17, Γ_{PL} is plotted as a function of P for each sample. A clear reduction of Γ_{PL} is revealed for the RWELL samples in comparison to the reference at an excitation of 260 mW.
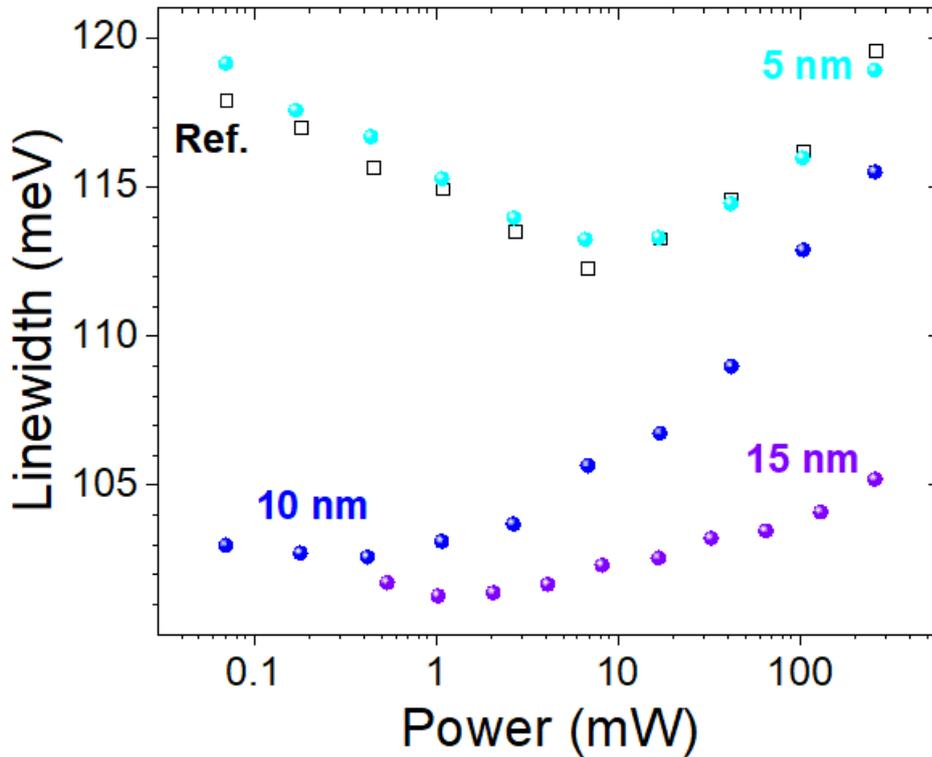
Figure 5.17: Γ_{PL} of the QR peak in each sample as a function of P.

It is surprising to observe a reduction in Γ_{PL} with increasing QW thickness. Although alloy disorder and scattering due to defects (impurities, dislocations etc.) are expected to play a key role, this dependence cannot be caused by them; since greater densities accumulate as layer thickness increases, thus result in the opposite FWHM dependence. Other mechanisms contributing to the inhomogeneous broadening of the FWHM must be considered to account for this peculiar behaviour. These are: phonon-induced homogeneous broadening, surface roughness (exacerbated by In segregation) and fluctuations of QW layer thickness due to deposition technique, penetration of the exciton wavefunction into the barrier layer, compositional differences between the QW and barrier, and strain-induced fluctuations of potential.

Furthermore, the number of indirect excitonic transitions in thinner QW samples will increase.

Electrons are susceptible to the potential fluctuations caused by atomic irregularities at the QW-barrier and QR/WL-QW interfaces, when the QW width is reduced such that a_X in the growth direction is larger than the layer thickness [256], as shown in Figure 5.18 and Figure 5.2. Increasing the confinement of electrons with a thinner QW increases the penetration of the electron wavefunction into the adjacent layers. Indeed, the 5 nm (15 nm) sample has the greatest (smallest) FWHM. Therefore, the QW width-dependent trend for the RWELL samples is primarily attributed to the reduction of QW dimensions which enhances tunnelling into the barriers, therefore increasing the significance of interfacial inhomogeneity.
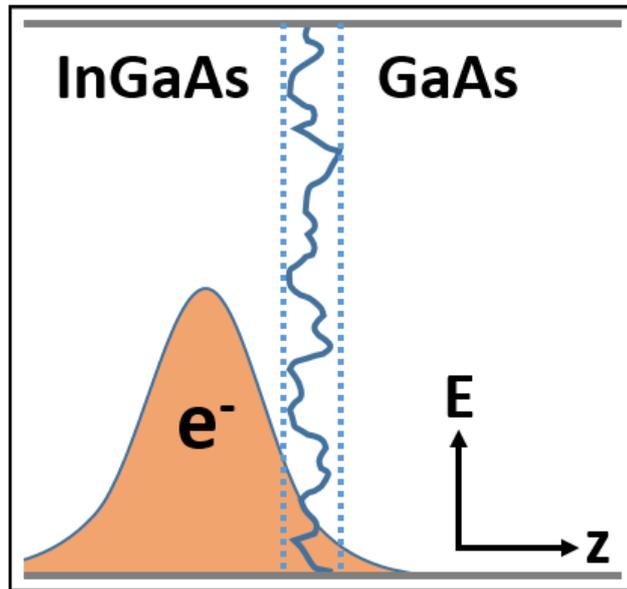
Figure 5.18: Sketch to show the penetration of the electron wavefunction into the barrier layer. Significant fluctuations of local QW width are caused by atomic irregularities illustrated here within the dashed region at the QW-barrier heterojunction.

Taking into consideration that the Γ_{PL}-values of the RWELL samples are less than the reference at full P, and the consistency of peak fitting; the majority of the contribution to FWHM broadening is unlikely to be a result of poor InGaAs growth. The possible mechanisms causing broadband emission that are mutually applicable to both types of structure are outlined here. Non-uniform nanostructure dimensions and shapes cause fluctuations in quantum confinement and local strain. Moreover, larger nanostructures have more hole states than smaller structures which changes the position of the energy levels. Extending the analysis applied to excitons in QWs [259] and QDs/QRs [86]; the hole confinement energy is sensitive to $1-2$ nm fluctuations in nanostructure height and quantisation in the x- and y- directions, i.e. the QR aspect ratio. The QRs will have a wide distribution of different thicknesses.

There will be additional quantisation in the x- and y- directions for electrons in between adjacent pieces of QR material, i.e. the distance between QRs, and within their inner radii. The above contributions to the lateral carrier confinement will reduce the energy required for exciton hopping. The sum of in-plane and out-of-plane confinement variation produces $e_1 - HH_1$ transitions with similar, but non-identical, recombination energies. Actually, the contribution of FWHM from variation in QR height and diameter should be very similar across all samples. This is because there were no differences to the annealing conditions across them. Interface roughness due to As-Sb exchange will also cause variation in the spatial separation of carriers and influence the local 3D strain profile, along with compositional effects on the local bandgap. Defect-related scattering of electrons is expected to dominate in the reference due to poor electron confinement, since all other contributions will be approximately similar to those in the RWELL samples, and the Γ_{PL} has increased beyond the 5 nm sample at low P.

Notably, a monotonic increase of Γ_{PL} is expected with increasing P. However, a $\sim$ 10 meV drop in Γ_{PL} is observed at $P < 10$ mW for the reference and 5 nm samples. This is very unusual to observe without an electric field applied to the gain media. An internal electric field is created across the QW by placing QRs at the bottom it, however screening of the quantum-confined Stark effect is rejected on the assumption that the in-built field will be too weak. It is even more curious that the samples with the most and least electronic confinement exhibit such a phenomenon, yet the 10 nm and 15 nm samples do not. The simplest explanation presents itself upon considering the peak fittings. The Γ_{PL} range-average is approximately half a minor tick in the spectra of Figure 5.13, which is comparable to the uncertainty of the fitted Γ_{PL}. This is easier to see with thinner plotted curves. The I_{PL} parameter is relatively insensitive to the accuracy of fitting due to its order of magnitude, whereas the peak position is confined to a very small P-domain of possible values. The reduction in Γ_{PL} seems to follow a consistent trend, however this could be a consequence of unintended bias rooted in masking a double-peak fit, and freedom of choosing the number of data points to include in the fit.

Alternatively, the Γ_{PL} reduction is adequately explained by a combined effect from OICD and band-tail states, i.e. the expansion of band edge states leading to a band-tail DoS with an energy that increases exponentially up to the free electron energy [260, 261]:

$$g(E) = \frac{N_\ell}{k_B T_0} e^{E_m / k_B T_0}. \tag{5.4}$$

Here, N_ℓ refers to the total number of localised states, and T_0 is a measure of the width of the exponential. Specifically, the characteristic temperature, T_0 quantifies

the T-sensitivity of a device. Large T_0-values represent better stability, i.e. slow increases of threshold current density and η_{ext} with T. The average trap state energy, with respect to the mobility edge, is represented by E_{m}. Since the holes are confined to the QRs and the electrons are free (subject to Coulomb attraction) in type-II structures, the excitonic properties are considered with respect to the confinement of electrons. Therefore, the mobility edge is redefined here as the energy which separates the localised and non-localised electron states instead of exciton states. The free exciton energy has also changed to reflect the energy required for an electron to overcome the average trapping potential.

Excitons are highly localised at low P and T. As carriers are injected into the system with small increases to the P, the lowest level(s) of the localised states become fully occupied at approximately ~ 7 meV. Some potential minima will have such a small energy barrier and small DoS, that they saturate rapidly with band-filling and cause additional electrons to spill over into deeper potential minima with vacant states. Band-filling opposes the broadening effect, since the vacant higher energy levels in these deeper potentials have a larger DoS. A reduction in Γ_{PL} therefore occurs in two ways. The average spatial variation of $\Delta\psi$ becomes more uniform with further filling of the larger/deeper potential minima. Additionally, the resulting emission of the slightly deeper potentials dominates the emission from weakly localising potentials and shifts the COM emission energy towards that of the deeply confined states.

OICD cannot be ignored due to the C-doped peak, so acceptor holes are likely present in the WL and QR at low P (and T). Assuming the QRs are charged at $P = 0$, then small increases to P will introduce more electrons into the system which recombine quickly with the holes and deplete them, thus the average hole occupancy reduces in the QRs, as acceptor states fill up. As previously discussed, a reduction in ΔE_{CC} will reduce E_{b}, and increase a_{X}, due to electronic screening [24]. Consequently, the electrons lose their sensitivity to strain-induced potential-fluctuations of (In)GaAs in close proximity to the QRs.

5.4.3 Temperature-Dependent Measurements
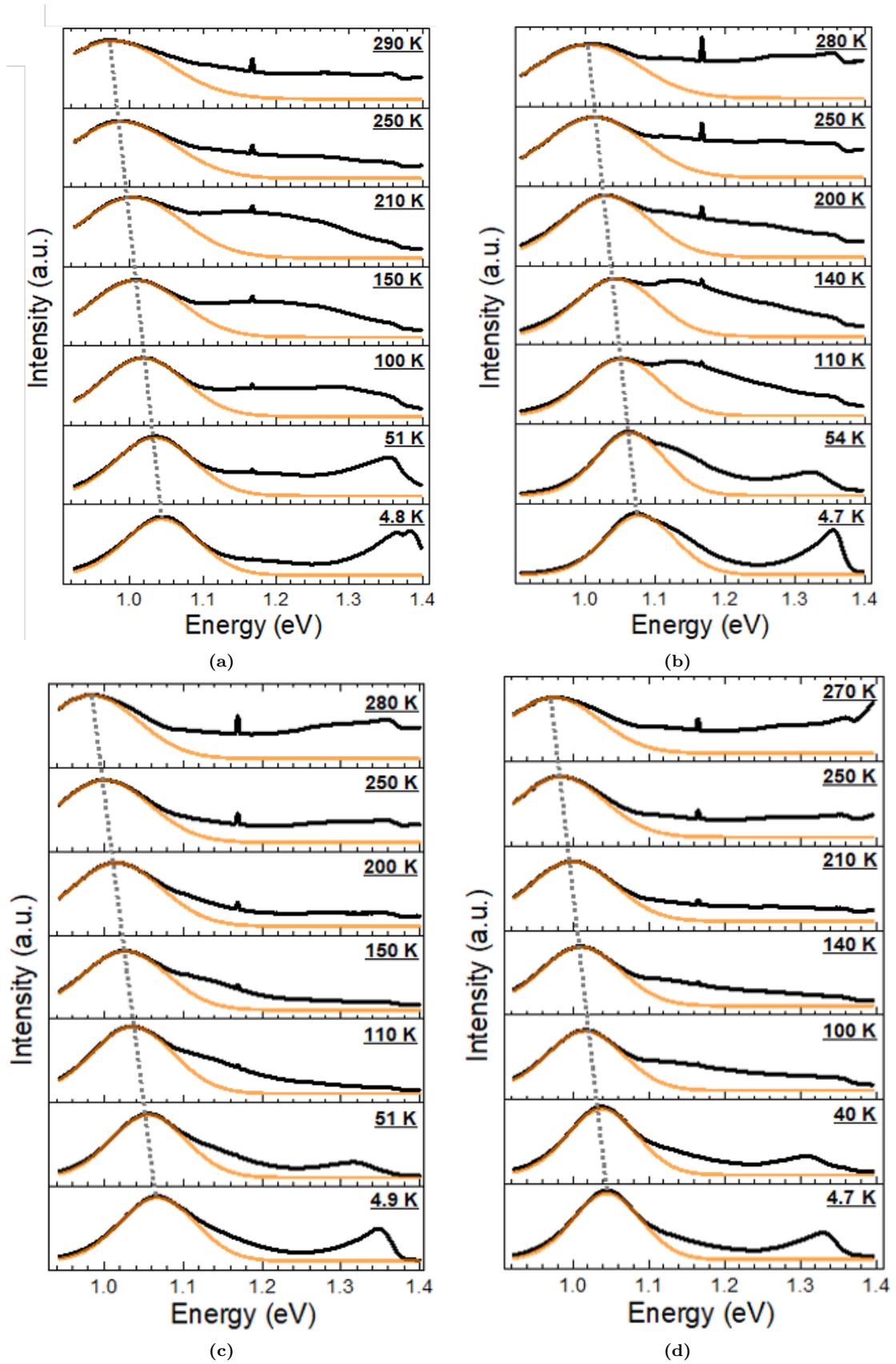
Figure 5.19: Temperature sweeps: a) ref, b) 5 nm, c) 10 nm, d) 15 nm.

Figures 5.19a-d show the evolution of the same two lowest energy peaks over a T-range of 4 K to $\sim$ 280 K. The background emission becomes significant towards RT.

The relative WL peak amplitudes at 4.2 K suggest that for thinner QW layers: the likelihood increases for a radiative recombination event to occur inside the WL, since the electron wavefunction is confined closer to the QW-WL interface. The WL PL decreases rapidly, with increasing T, relative to the QR peak; thus indicating a small WL E_a. Holes become thermally activated at sufficiently HT, and transfer from the WL to the QRs. Smaller T-increments than those presented in Figures 5.19a-5.19d are required for a precise estimation of the WL quenching temperature, hence the entire normalised dataset was consulted. For the reference sample, the WL PL fully quenches above (90 ± 5) K. At this T, the QR peak I_{PL} reduces to 3.0 % of its initial value at 4.2 K. The quenching of the WL appears to be slightly dependent on QW width. However, more T-precise measurements need to be taken to confirm this observation. The WL completely disappears at temperatures above (85 ± 5) K, (80 ± 5) K, and (70 ± 5) K for the 5 nm, 10 nm, and 15 nm samples, respectively. At these temperatures, the QR peak I_{PL}-values reduce to 1.1 %, 4.6 %, and 9.1 % of their initial 4.2 K values, respectively. Thus, it requires less thermal energy for holes to escape the WL in wider RWELL structures. This is adequately explained by the increased compressive strain inside the QRs (from a thicker QW).

In Figure 5.20, E_{PL} is plotted as a function of T for each sample. Notably, the longest λ RWELL sample emits around $\sim$ 1042 meV at 4 K and reduces by approximately 80 meV to $\sim$ 962 meV near RT.

An immediate observation is that the data slightly deviates away from the Varshni relation in the mid-temperature range. The observed trend is reminiscent of a weak characteristic "S"-shape dependence common in partially disordered semiconductors such as InGaN/GaN MQWs [260]. This behaviour is an indication of significant weak-trapping of carriers in band-tail states, caused by fluctuations of potential as illustrated in Figure 5.22. The T-dependence of the bandgap, $E_g^{BT}(T)$ (see Equation 2.26) is modified in the band-tail model [262] to account for carrier thermalisation (see Figure 5.22):

$$E_g^{BT}(T) = E_g(T) - \frac{\sigma_\ell^2}{k_B T} \tag{5.5}$$

Where σ_ℓ is an indication of the degree of localisation; a larger σ_ℓ-value corresponds to stronger localisation. Increasing P causes σ_ℓ to reduce until the strong "S"-dependence eventually converges to the anticipated Varshni bandgap dependence.
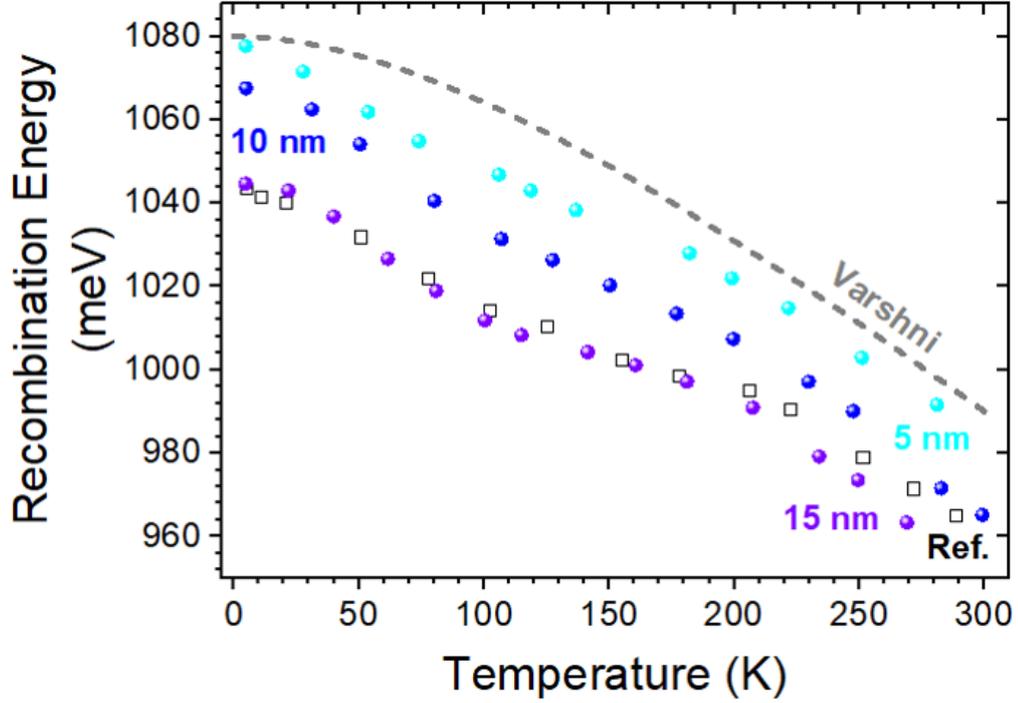
Figure 5.20: Recombination energy of the QR peak for each sample as a function of T.

The band-tail model explains a decreasing-increasing-decreasing dependence of E_{PL} at low P with increasing T. Emission first redshifts up to a maximum localisation energy, beyond which it then blueshifts until a temperature where the carriers are fully delocalised. For these samples, the electrons fully delocalise at an approximate temperature of ~ 150 K. Further heating causes the emission to redshift again [260]. Admittedly, a monotonic redshift manifests in the measured data here, however this is due to the power used ($P \approx 260$ mW). The amount of blueshift is overcompensated by the overall reduction in bandgap with increased T, hence slows the rate of redshift to plateau-like range of mid-temperatures. An observation of this localisation effect highlights the strong confinement of electrons by the QW.

The reference exhibits the same dependence as the RWELL samples, thus implies that In conglomerations are unlikely to have formed accidental QDs with confined states in the QW. The visibility of this dependence across all samples implies that the band-tail phenomenon almost certainly arises due to the types of QR-related disorder outlined in the P-varied Γ_{PL} analysis. The 15 nm and reference samples show the most pronounced localisation behaviour and the 5 nm sample shows the least. This agrees with expectations on the amounts of disorder introduced by increasing cap thickness. An increased rate of redshift for the 5 nm and 10 nm samples below ~ 40 K is a manifestation of their increased electron confinement.

Figure 5.21 displays the evolution of Γ_{PL} with T for the Gaussian fitting. It is used to help inform discussion on carrier motion. Notably, the localisation behaviour is

most visible from the 5 nm sample hence discussion on Γ_{PL} is in relation to this sample and each stage of the thermalisation process is assumed similar for the other RWELL samples.
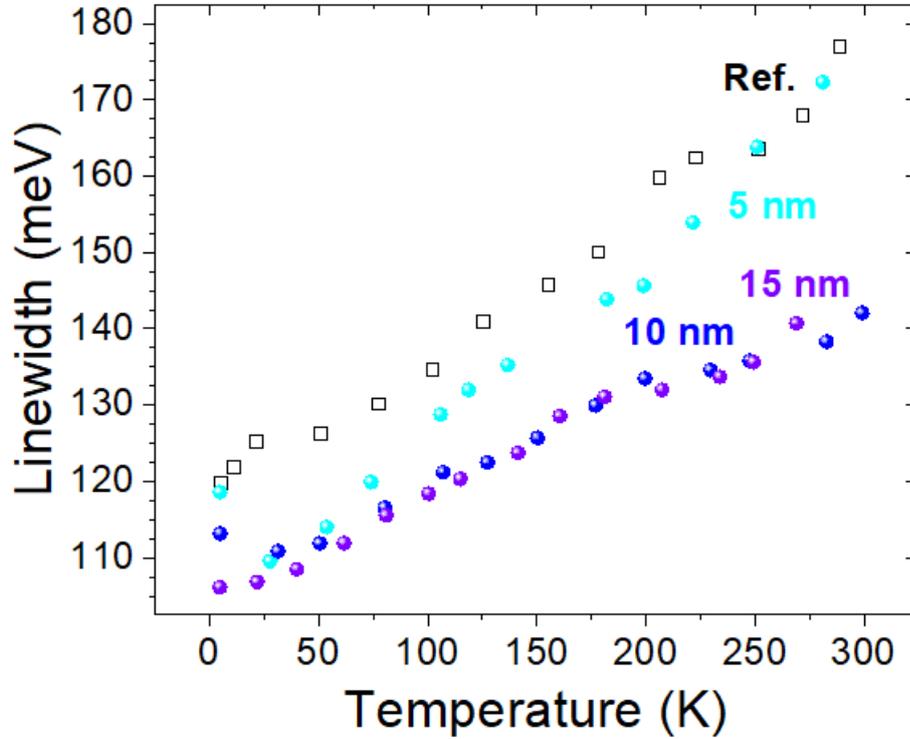
Figure 5.21: Γ_{PL} of the QR peak for each sample as a function of temperature.

Carriers are randomly distributed at 4 K, due to the variation of potential minima. Small increases to T thermally activate carriers that are weakly-localised in the shallowest traps, as shown by Figure 5.22a. An initial redshift and decrease in Γ_{PL} are indications of hopping to other localised states of lower energy, until a saturated redistribution is reached, as shown by Figure 5.22b. Usually at this temperature, E_{PL} is observed to be at a minimum when measuring with sufficiently low P. It can only be inferred from the change of gradient here, along with the Γ_{PL} minima, that saturation occurs at $T_{\mathrm{min}} \approx 40$ K. After ~ 40 K, Γ_{PL} starts to increase as the carrier distribution thermalises; electrons start to occupy higher energy states within the deeper confinement potentials, as shown by Figure 5.22c, thus causing the emission to blueshift up to a maximum at the free electron ground state. The rate of Γ_{PL} increase slows as temperatures approach 100 K, which corresponds to the T-range where electrons gain mobility in the deepest potential minima, as depicted in Figure 5.22d. The narrowing of the carrier distribution reduces the FWHM. This is countered by an increase of FWHM caused by the regular thermalisation of carriers with further heating of the sample towards the full-delocalisation temperature.
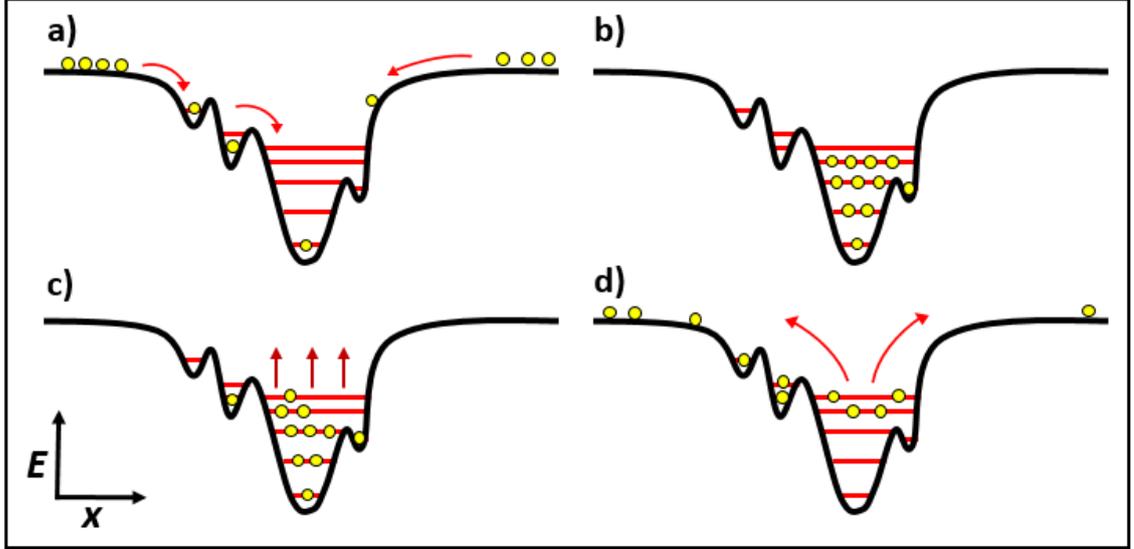
Figure 5.22: Sketches to show the electron distribution across several shallow potential minima, as T increases from a) 4 K to b) $T \approx 40$ K, c) $T > T_{\min}$, and d) $T \approx 100$ K.

Delocalised carriers are increasingly prone to phonon scattering with increasing T. Thus, the regular thermalisation of carriers (between $T_{\max} \approx 150$ K, and RT) broadens the FWHM at a faster rate and E_{PL} reduces. The steep gradient in Figure 5.20, at elevated T, indicates additional redshift. For the regular scattering of electrons, Γ_{PL} can be expressed as a function of T:

$$\Gamma_{\mathrm{PL}}(T) = \Gamma_0 + \Theta_{\mathrm{a}}T + \frac{2\Theta_{\mathrm{LO}}}{\mathrm{e}^{E_{\mathrm{LO}}/k_{\mathrm{B}}T} - 1}. \tag{5.6}$$

Here, E_{LO} is the energy of the LO phonons. The carrier-phonon interaction coefficients Θ_{LO} and Θ_{a} represent interactions with LO phonons and acoustic phonons, respectively.

For the 5 nm and ref samples, Γ_{PL} increases at an exaggerated rate above $T > 275$ K. This may be a fitting artefact caused by increasing background contributions to PL, whilst the peak broadens and redshifts towards the λ detection window cut-off ($E \approx 0.9$ eV).

The E_{a}-values of the predominant loss mechanisms for each sample were sought from the data presented in Figure 5.23, in order to assess the competing radiative and non-radiative recombination channels. The Arrhenius equation (Equation 5.7) was fitted to the normalised I_{PL} as a function of inverse T. The negative E_{a} implies a reduced probability of recombination at elevated T. There are two clear slopes with different gradients. This implies two quenching processes pertaining to the HT and LT limits.

$$\frac{I_{\mathrm{PL}}}{I_0} = \frac{1}{1 + A_1\mathrm{e}^{(-E_1/k_{\mathrm{B}}T)} + A_2\mathrm{e}^{(-E_2/k_{\mathrm{B}}T)}}. \tag{5.7}$$

Where $A_{1,2}$ and $E_{1,2}$ refer to the non-radiative recombination coefficients (corresponding to the rate of a specific Auger process present in the sample), and E_a at HT and LT, respectively. I_{PL}/I_0 is the ratio of integrated intensity to the extrapolated integrated intensity value at absolute zero. The I_{PL}-value at absolute zero and 4 K were assumed equal.
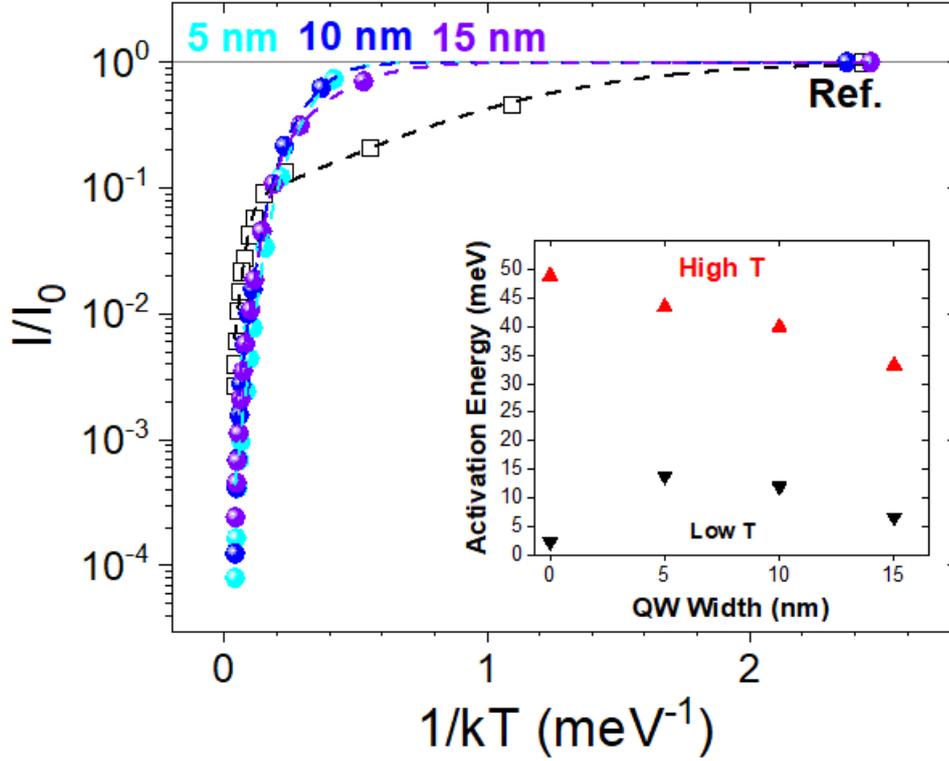
Figure 5.23: Ratio of integrated intensities, as a function of inverse T for the QR peak of each sample. The data is fitted by Equation 5.7. Inset: E_a is plotted as a function of QW width. The red and black triangles correspond to hotter and colder temperatures, respectively.

At LT, the RWELL sample data reveals a clear improvement to efficiency, in comparison to standard GaSb/GaAs QRs. From 4.2 K to $\sim$ 70 K, the PL of the RWELL samples quenches at a lesser rate, with increasing T, than the PL of the reference. This implies a larger E_a-value for the RWELL samples, than for the reference; which is only observed for the smaller E_a, even though non-radiative transitions at HT were expected to require less energy due to phonons. In the HT limit, where T increases from $\sim$ 70 K to RT, the converse is true, i.e. the E_a-value of the RWELL samples must be smaller than that of the reference. The smaller (larger) E_a-value is thus attributed to the loss mechanism at LT (HT).

Three assumptions are made in order to analyse the data further. First, the excitons thermally dissociate to free carriers with increasing T. Secondly, PL at LT is purely excitonic and the amount of thermal dissociation is negligible. Thirdly, excitons are at the ground-state energy; before they can recombine, they need to overcome an energy barrier (E_a) to reach the band edge.

156

The cold E_a-value of the reference sample is attributed to the average E_b, i.e. minimum amount of thermal energy required to break the excitons apart (on average), since electrons are only loosely bound to the GaSb/GaAs interface of the QRs and are easily ripped away by phonons with increasing T. The experimentally-obtained $E_a \approx 2$ meV of the reference reflects this, and strongly agrees with the calculated $E_b = 2.226$ meV reported for a GaSb/GaAs QR in Ref. [263].

Importantly, the E_a-values of the RWELL samples are all greater than that of the reference, and reduce with poorer electron confinement in the growth direction. The sensitivity of E_a to QW width therefore has to be either related to confinement or the density of defects. It is difficult to discern the LT mechanism of the RWELL samples with absolute certainty from the current datasets. These ambiguous mechanisms are elucidated here. Forcing the carriers closer together increases the Coulomb potential which acts like a stronger "glue" in the 5 sample, thus increasing E_a. This effect is amplified by greater amounts of CC caused by greater electron densities at the heterojunction. Moreover, the 5 nm sample will have the largest reduced mass according to Equation 5.8, due to its dependence on tunnelling of the electronic wavefunction into the barriers [256]; the larger $\Delta\psi$ increases E_b relative to the 15 nm and reference samples.

$$\frac{1}{m^*_{e,tot}} = \frac{|\psi_{QW}|^2}{m^*_{QW}} + \frac{|\psi_{barrier}|^2}{m^*_{barrier}} = \frac{|\psi_{QW}|^2}{m^*_{QW}} + \frac{(1 - |\psi_{QW}|^2)}{m^*_{barrier}}. \tag{5.8}$$

Where $m^*_{e,tot}$ refers to the combined increase of reduced effective mass from the QW and tunnelling effect, the electronic wavefunction probabilities inside the QW, and barrier, are $|\psi_{QW}|^2$, and $|\psi_{barrier}|^2$, respectively. The effect of non-parabolicity CB is negligible at the Γ-point and gains significance with increasing k.

Greater defect densities can increase E_a, by making it more probable for electrons to be trapped in deeper potential minima. When one of the carriers in a type-II exciton is contained inside a nanostructure, and the other is bound to a potential minima, or defect trap state, the associated E_b makes it more difficult for the exciton to dissociate. Greater accumulations of material increase the likelihood of dislocations, the density of In conglomerations, and concentration of carbon. In particular, the average confinement potential of accidental (Ga)InAs QDs will be deepen as more In atoms are incorporated and migrate.

On the other hand, the CB DoS will also increase with increasing QW thickness, thus at large P, it is more probable to excite an electron into a defect trap state. Therefore E_a might correspond to the minimum trapping energy of a shallow non-radiative recombination centre. The XRD data presented in Figures 5.7 and 5.10

implies a large density of non-radiative recombination centres is less likely to be causing the reduction of E_a. If E_a corresponded to C-doping, then these would likely be constant across all samples and this is not observed either. So it may be reasonably concluded that defects are unlikely to play a big role in these samples after all.

Regardless of whether the reduction in E_a is from a reduced $\Delta\psi$-value or a type of defect, the inset of Figure 5.23 shows clear evidence that reducing the QW thickness, i.e. increasing Coulomb interaction, and therefore increasing E_b results in better thermal stability. This is a remarkable result considering the dilution of the QW and is very promising on the route to stabilising bright O-band lasing at RT. An opportunity is presented here for further improvements to thermal stability by increasing In %. Further tweaks to the growth recipe are necessary to address issues of growth quality first. Emission can be slightly red-shifted to at least 1300 nm at RT with these samples by either growing a thicker QW or focusing on the QR dimensions. Consideration of strain is extremely important, as it increases with thicker QW layers and higher concentrations of In.

In the HT limit, the reference E_a-value is close to $E_a = 89$ meV of a triple-stack GaSb/GaAs QD structure in Ref. [28], where the authors attribute it to weak quantum confinement of electrons within the wide potential well of the CB. For the RWELL samples, the hot E_a is assigned to the thermal escape of free electrons out of the QW and into the GaAs layers. This claim is well-supported by the calculation results presented in Subsection 4.2.3, where the difference between the electron confinement energy and the CB offset at the QW-barrier interface closely match the extracted E_a-values.

The hot E_a-values decrease as the QW width increases from 5 nm to 15 nm, Indeed, the calculated CB edge offset at the GaAs/InGaAs interface reduces as the QW width increases. This phenomenon must be strain-related since strain tilts the CB edge. Therefore, the energy barrier at the upper interface of the dilute QW decreases; thus facilitating the excitation of electrons out of the QW into states within the GaAs layers.

The confinement inside the 250 nm GaAs barriers is expected to be negligible (89.9 μeV), thus electrons can be assumed to reside at the CB edge.

Assuming a 2 nm QW width, the confinement energy of the protective GaAs layer is approximately 1.4 eV (Equation 2.66), hence trap states from these protective layers are considered optically-inactive; i.e. both electrons and holes will be repelled by their respective confinement energy and $In_{0.06}Ga_{0.94}As$/GaAs band offset.

Chapter 6

Summary

The inclusion of a dilute InGaAs QW within the active region of GaSb/GaAs QR structure provides an extra degree of freedom to fine-tune the opto-electronic properties of a type-II structure, via $\Delta\psi$. The temperature stability of PL from the nanostructures can be easily modified by varying the electron confinement, and therefore $\Delta\psi$ [28]. Strong confinement yields an increased kinetic energy of indirect excitons, along with increases to Coulomb energy, oscillator strengths and binding energies, which ought to provide stable PL at RT. Notably, it offers more advantages than a simple increase of RT I_{PL}. We developed the asymmetric RWELL structure presented in Ref. [59], following on from the work of Ref. [47], with the use of isolated layer stacks in a MQW and a novel annealing technique to form the QRs. We further explored the potential use of this structure for lasing applications, by varying QW width to investigate confinement effects and the resultant optical properties. Decreasing the QW width has a large effect on the resulting LT transition energy; this is not observed in GaSb/GaAs QRs in $Al_xGa_{1-x}As$ QW structures [26] above 10 nm in the entire T-range, despite the prediction of it in their simulations that estimate a drop of ~ 200 meV . This is comparable to a similar prediction of ~ 200 meV of these structures by increasing the In concentration to values close to 30 %. Counter-intuitively, redshifted emission is not observed, other than at low excitation power and T, however the simulations suggest that wavelength extension is possible. It is reasonable to expect that a 5-10nm increase of the QW, along with the formation of bigger rings will shunt emission into the E-band. The simulations of Ref. [26] suggest that modifications to the CB edge by increasing x to 20% is the most likely way to achieve C band emission. This will require delicate refinement of the growth conditions to optimise the formation of QRs inside the QW, and achieve a homogenous layer of InGaAs above the nanostructures with smooth interfaces to fully exploit the advantages of using an InGaAs QW to cap the GaSb QRs; so that

they may serve other blueshift-insensitive applications. The potential mechanisms behind the blue-shift have been elaborated on, and it is boldly concluded here for the first time that extension of λ beyond the O-band using the relative band-offsets of an InGaAs/GaSb cap layer is probable. provided solutions can be developed to overcome the combined effect from electron correlation, strain and potentially, the average contribution of two very similar radiative recombination channels.

Improvements to Γ_{PL} and quenching, relative to standard GaSb/GaAs QR MQWs, are demonstrated, along with a larger blueshift across the same span of pumping power. Thermal quenching is found to be caused by two main non-radiative channels that activate in different temperature ranges. Hole densities govern rates of recombination in these samples and their thermal escape out of the QRs accounts for the HT E_{a}. The LT E_{a} is related to the thermal escape of electrons out of the QW. Both of these thermal activation mechanisms appear to be strain-sensitive. Annealing is presumed to have likely caused accidental, In-rich QDs in the QW and a band tail model was used to account for localised states in the QW near the CB edge due to them, and other effects which cause fluctuations in the potential minima. A strong dependence on QW width is clearly observed for Γ_{PL}, I_{PL}, E_{PL}, E_{a} and the rate of change of intensity with respect to P. For the 15 nm RWELL sample, I_{PL} reduces; an indication that the QW has approached the $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}$ critical thickness, despite being significantly under-estimated as a growth precaution. These findings reveal the potential of using QR-in-well structures for GaAs-based telecom lasers with improved performance at LT, however further modifications to the growth parameters are required to improve the performance at higher T. Surface roughness, alloy disorder and the annealing temperature are likely to impact the quality of QR formation and hence reduce the QRs confinement potential, thus leading to a lower E_{a} at RT.

Chapter 7

Future Work

These simulations and experimental data indicates a hopeful possibility of using this material system to create the first widely tunable single source emitter that operates across the entire telecom bands. The steps to realise it seem rather straight forward, it all depends on being able to drop the CB edge with successful growth of GaSb QRs on InGaAs layers, by increasing the In content to 20% (C-band) 30% (U-band). Otherwise the E band is likely achievable by refining growth conditions to reduce defect density and extend QW width beyond 15 nm along with the formation of slightly bigger rings. Might even be able to reach the 850 nm band making use of strained short QWs and even AlGaAs barriers (wf VB mixing). To increase wavelength further should the future student encounter any strain-related obstacles, they could perhaps look at coupling QR layers inside a QW to increase tensile strain inside the QW if the QW is sandwiched in between each layer. This would require a different approach to forming the rings than proposed here. Optimisation of the In concentration and QW growth temperature for higher structural quality is required for future work. Immediate work to be undertaken if there was more lab time, includes characterisation of the current samples to measure layer roughness and thickness and inform the growth parameters of the next samples. Varying T_{sub} whilst growing the QW will allow the optimisation of interfaces in the QW. Further to this, the annealing temperature should be varied to observe the effects of intermixing. Once a sample is able to emit very strong PL at RT, the In concentration in the QW can be properly investigated. Specifically, whether the drop in CB edge can be enough to induce long-wavelength emission without strain-related issues. This can be done by changing the In composition, and reintroducing the QRs back into the centre of the QW, along with using a graded QW.

To refine the current simulations, some cross-sectional AFM would be helpful to find out the QR geometry.High-resolution XRD, EDX for compositional measurements,

/X-SEM/HRTEM/XSTM etc to view the buried nanostructure morphology and gauge structural reasons for why the samples do not emit at the intended wavelength. The QWs may need to be grown at hotter T_{sub} to achieve optimal quality growth; the MQW sample XRD looks encouraging. A longer term goal: refine the model by taking density-functional theory into account, this will provide a small correction to the estimated binding energies. To check whether blueshift is because of weakened E_b, magneto-PL measurement should be undertaken to probe the exciton Bohr radii. An increased E_b results in an extended indirect exciton lifetime. The cryostat should be used to have a more precise control of T during PL measurements.

The lateral confinement results predicted by the current simulations needs further refinement by simulating different configurations of nanostructures, and their separations. It would also be advantageous to explore different avenues for the precise control of lateral confinement of the wavefunction. The accuracy of the carrier confinement energies will improve once an accurate wavefunction probability is obtained. From this, the binding energies can be estimated properly. We want the longest wavelength of emission possible, so need to find out if the tensile strain is too much as x increases, i.e. what is the lowest drop in CB edge we can observe whilst still forming high quality structures, and without shunting the QW states beyond the GaAs ground state. This may require a re-optimisation of the QW growth parameters. Along-side the inclusion of lateral confinement, it is worth including the WL to further improve the prediction accuracy

Once high quality growth conditions of GaSb/InGaAs QRs have been obtained, e.g. decent QR density, height and diameter, etc. Then try grow them in a symmetric MQW. This will isolate the effects of lateral confinement of electrons because the strain above and below the rings should be similar due to symmetry as implied by the strain calculations of Chapter 4. The surface roughening during growth makes it rather difficult. Should high quality growth be achieved for a dilute symmetric MQW, it is worth measuring the time-resolved PL for a QW with larger x to measure the radiative and Auger recombination rates. Additional capacitance-voltage measurements are required to determine the exact band-offset for the InGaAs/GaSb QRs. This technique simultaneously obtains the band edge discontinuities and the charge density at the interface, i.e. the in-built potential, by comparing bulk-doped samples with the band extrema away from the interface and the Fermi level distance, the VB and VB can be found.

Appendix A

VCSELs
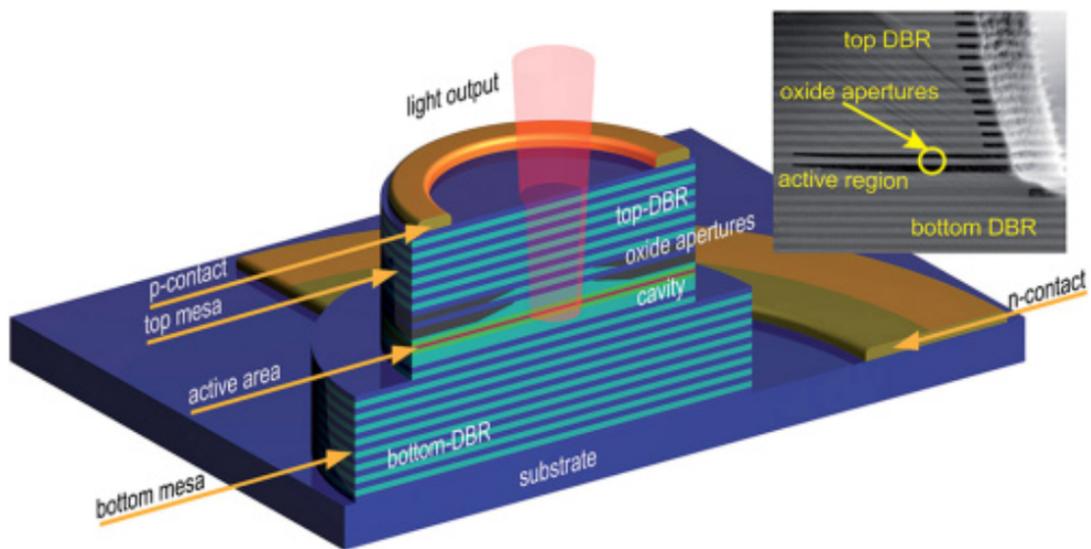
Figure A.1: Schematic of a top-emitting VCSEL. Selectively oxidised layers (dark grey) control the current path between the contacts (dark yellow). The profile of the beam is dependent on the diameter of this aperture. Inset: scanning electron microscopy (SEM) image of a cross-section of the structure after being cleaved. Figure reproduced with permission from [13].

A VCSEL laser is presented in Figure A.1. Injection of electrical current (via the top and bottom metal contacts) causes this type of monolithic LD to emit a laser beam from an aperture at its top surface. It uses two heavily-doped Bragg reflectors, above and below the active region, to form the optical cavity and induce resonance; thus amplifying the intensity of photons within the gain medium. The Bragg reflector is a stack of alternating thin films of high- and low- n_r, which reflects light at particular λ-values. The reflectivity of these stacks is usually $> 99\,\%$ (at the intended resonant λ), since the layer thickness is set to a quarter of the radiative recombination energy of carriers inside the active region, E_PL. One stack is designed with a larger number of periods to slightly increase its reflectivity and control which surface the generated

light emerges from. In basic designs, one reflector is p-type and the other is n-type. An electrically-insulating oxide layer, with high optical transmittance at the lasing λ (NIR for the applications of this work), is often placed between the active region and the Bragg reflector nearest the beam output. The oxidation aperture is used as a current window to efficiently pump the active region; reducing threshold current and optical loss.

Appendix B

Anti-Phase Domains

Growing on (001)-orientated Si used to be challenging due to the difference in number of valence electrons of the epi-layer material and that of the substrate; affecting the initial interface formation at the start of growth. In III-V/Si epitaxy, the group-V species preferentially nucleates on the top surface of the Si substrate, terminating at 1 ML. The surface is then chemically stabilised by forming As-As dimers, which makes it difficult to form bonds with group-III atoms. This forms a region on the Si(001) surface with ML steps covered with group-V atoms, i.e. an APB, as shown in Figure B.1. Due to these bonds, atoms in the region enclosed by the APB are flipped to the opposite order in which they would normally be configured in an ideal lattice. Each of these point defects are known as an antisite. Continued planar growth leads to a 3D growth mode with a high probability of nucleation at these antisites [131]. Thus TDs and stacking faults can form and further degrade device performance.

Recently, the epitaxial overgrowth scheme "ART+SAG" (aspect ratio trapping and selective area growth) has attracted much interest [264, 265] in comparison to other approaches that combat APBs for monolithic III-V/Si integration; e.g. wafer bonding [77], or $2-6$ ° substrate off-cuts towards the Si(110) planes [78]. The ART+SAG method offers a highly controllable and reproducible way to achieve monolithic III-V/Si integration, while merely employing an ultra-thin buffering layer.

Research at the Lancaster Quantum Technology Centre therefore proposes monolithic growth of our nanostructures on patterned "V"-groove Si [266], due to its alluring potential of defect suppression; access to the direct nature of the III-V bandgap; and ability to reap the many benefits of using the Si platform. Misfit strain is essentially relaxed via III-V nucleation twinned stacking faults aligned to the (111) facets of the V-grooves. Ridges exist at Si surface since the substrate is
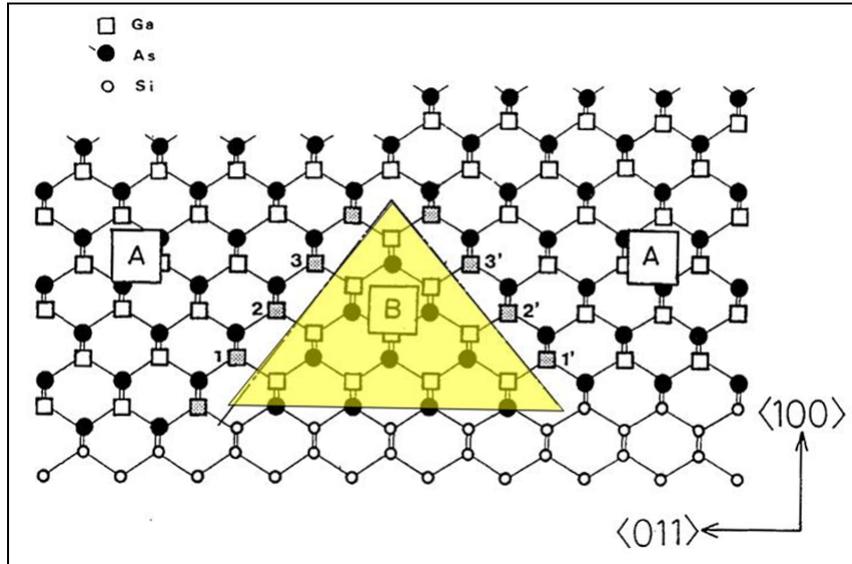
Figure B.1: Diagram to show an inversion domain. The yellow triangle highlights the APB formed during growth of GaAs/Si [110].

slightly undercut; these trap these defects and avoid the formation of APBs [72]. A high aspect ratio will allow an efficient defect trapping, i.e. the APB density decreases with reduction of trench width [267].

Appendix C

Growth Protocols

There was a system bake between the growth of the symmetric and asymmetric samples hence T_{Ga} is slightly different for these batches of growth.

Table C.1: Growth parameters used for the deposition-varied GaSb/GaAs QR samples. Here, R_{gr} and $\varnothing_{As,Sb}$ are the growth rate and valve position of the As and Sb cracker cells, t is the layer thickness.

Sample ID	Desorp. T (°C)	Buffer Layer		QR Layers		Cap Layers		Barrier Layers	
A1109	650	T_{sub}	650 °C	T_{sub}	540 °C	T_{sub}	475 °C	T_{sub}	650 °C
		T_{Ga}	955 °C	T_{Ga}	903 °C	T_{Ga}	903 °C	T_{Ga}	955 °C
		T_{As}	380 °C	T_{Sb}	510 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	210 mil	$\varnothing_{Sb}$	100 %	$\varnothing_{As}$	220 mil	$\varnothing_{As}$	210 mil
		R_{gr}	1.0 ML/s	R_{gr}	0.3 ML/s	R_{gr}	0.30 ML/s	R_{gr}	1.0 ML/s
		x_t	254 nm	x_t	2.1 ML	x_t	5 nm	x_t	204 nm
A1110	680	T_{sub}	650 °C	T_{sub}	540 °C	T_{sub}	475 °C	T_{sub}	650 °C
		T_{Ga}	955 °C	T_{Ga}	903 °C	T_{Ga}	903 °C	T_{Ga}	955 °C
		T_{As}	380 °C	T_{Sb}	510 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	210 mil	$\varnothing_{Sb}$	100 %	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	R_{gr}	0.3 ML/s	R_{gr}	0.30 ML/s	R_{gr}	1.0 ML/s
		x_t	204 nm	x_t	2.1 ML	x_t	5 nm	x_t	204 nm
A1111	650	T_{sub}	650 °C	T_{sub}	540 °C	T_{sub}	475 °C	T_{sub}	650 °C
		T_{Ga}	955 °C	T_{Ga}	903 °C	T_{Ga}	903 °C	T_{Ga}	955 °C
		T_{As}	380 °C	T_{Sb}	510 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	210 mil	$\varnothing_{Sb}$	100 %	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	R_{gr}	0.3 ML/s	R_{gr}	0.30 ML/s	R_{gr}	1.0 ML/s
		x_t	204 nm	x_t	1.8 ML	x_t	5 nm	x_t	204 nm

Table C.2: Growth parameters used for the T_{sub}-varied GaSb/GaAs QR samples.

Sample ID	Desorp. T (°C)	Buffer Layer		QR Layers		Cap Layers		Barrier Layers	
A1126	680	T_{sub}	680 °C	T_{sub}	540 °C	T_{sub}	475 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{Ga}	903 °C	T_{Ga}	903 °C	T_{Ga}	955 °C
		T_{As}	380 °C	T_{Sb}	510 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	220 mil	$\varnothing_{Sb}$	100 %	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	R_{gr}	0.3 ML/s	R_{gr}	0.30 ML/s	R_{gr}	1.0 ML/s
		x_t	339 nm	x_t	2.1 ML	x_t	5 nm	x_t	204 nm
A1127	680	T_{sub}	680 °C	T_{sub}	550 °C	T_{sub}	485 °C	T_{sub}	660 °C
		T_{Ga}	955 °C	T_{Ga}	903 °C	T_{Ga}	903 °C	T_{Ga}	955 °C
		T_{As}	380 °C	T_{Sb}	510 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	220 mil	$\varnothing_{Sb}$	100 %	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	R_{gr}	0.3 ML/s	R_{gr}	0.30 ML/s	R_{gr}	1.0 ML/s
		x_t	373 nm	x_t	2.1 ML	x_t	5 nm	x_t	204 nm
A1128	680	T_{sub}	660 °C	T_{sub}	530 °C	T_{sub}	465 °C	T_{sub}	660 °C
		T_{Ga}	955 °C	T_{Ga}	903 °C	T_{Ga}	903 °C	T_{Ga}	955 °C
		T_{As}	380 °C	T_{Sb}	510 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	220 mil	$\varnothing_{Sb}$	100 %	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	R_{gr}	0.3 ML/s	R_{gr}	0.30 ML/s	R_{gr}	1.0 ML/s
		x_t	373 nm	x_t	2.1 ML	x_t	5 nm	x_t	204 nm

Growth Step	Duration (s)
GaAs buffer deposition	See Table C.4
T_{sub} and T_{Ga} reduced under As flux	180
In$_{0.15}$Ga$_{0.85}$As QW deposition	See Table C.4
GaAs barrier deposition	See Table C.4
T_{sub} increased under As flux	180
GaAs barrier deposition	See Table C.4
T_{sub} reduced to 250 °C under As flux	840
As valve closed	20

Table C.3: Growth protocol for the In$_{0.15}$Ga$_{0.85}$As/GaAs QW-only samples: A1336, A1396, A1397. The Ga shutter closes at the end of deposition. The ternary growth rate is the sum of InAs and GaAs growth rates. The cyan steps correspond to the repeated sequence (x3).

Table C.4: Growth parameters used for the $In_{0.15}Ga_{0.85}As/GaAs$ QW-only samples.

Sample ID	Desorp. T (°C)	Buffer Layer		QW Layers		Cap Layers		Barrier Layers	
A1336	710	T_{sub}	680 °C	T_{sub}	640 °C	T_{sub}	640 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	947 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	200 mil	T_{cr}	850 °C	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	200 mil	R_{gr}	0.85 ML/s	R_{gr}	0.85 ML/s
		x_t	204 nm	R_{gr}	1.0 ML/s	x_t	102 nm	x_t	102 nm
				x_t	10 nm				
A1396	690	T_{sub}	680 °C	T_{sub}	640 °C	T_{sub}	640 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	947 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	200 mil	T_{cr}	850 °C	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	200 mil	R_{gr}	0.85 ML/s	R_{gr}	0.85 ML/s
		x_t	1221 nm	R_{gr}	1.0 ML/s	x_t	102 nm	x_t	102 nm
				x_t	15 nm				
A1397	700	T_{sub}	680 °C	T_{sub}	640 °C	T_{sub}	640 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	947 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{As}	380 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{cr}	850 °C
		$\varnothing_{As}$	200 mil	T_{cr}	850 °C	$\varnothing_{As}$	200 mil	$\varnothing_{As}$	200 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	200 mil	R_{gr}	0.85 ML/s	R_{gr}	0.85 ML/s
		x_t	204 nm	R_{gr}	1.0 ML/s	x_t	102 nm	x_t	102 nm
				x_t	20 nm				

Growth Step	Duration (s)
GaAs buffer deposition	See Table C.6
T_{sub} and T_{Ga} reduced under As flux	480
$In_{0.15}Ga_{0.85}As$ QW deposition	See Table C.6
As valve closed	15
T_{Ga} reduced	300
Sb valve open	30
GaSb deposition	See Table C.6
T_{sub} reduced and T_{Ga} increased under Sb flux	180
As valve open	20
Sb valve closed	2
$In_{0.15}Ga_{0.85}As$ cap deposition	See Table C.6
T_{sub} increased under As flux	480
GaAs barrier deposition	See Table C.6
T_{sub} reduced under As flux	480
T_{sub} increased under As flux	240
GaAs barrier deposition	See Table C.6
T_{sub} reduced under As flux	300
$In_{0.15}Ga_{0.85}As$ QW deposition	See Table C.6
As valve closed	15
T_{Ga} reduced	300
Sb valve open	30
GaSb deposition	See Table C.6
T_{Ga} increased under Sb flux	180
T_{sub} reduced to 450 °C under Sb flux	240
Sb valve closed and T_{sub} reduced to 250 °C	240

Table C.5: Growth protocol for the symmetric $In_{0.06}Ga_{0.94}As$/GaSb/GaAs RWELL samples: A1350, A1364, A1363 and A1366. The Ga shutter closes at the end of each deposition step. The cyan steps correspond to the repeated sequence (x3).

Table C.6: Growth parameters used for the symmetric $In_{0.06}Ga_{0.94}As/GaSb/GaAs$ RWELL samples.

Sample ID	Desorp. T (°C)	Buffer Layer		QW Layers		QR Layers		Cap Layers		Barrier Layers	
A1350	670	T_{sub}	680 °C	T_{sub}	520 °C	T_{sub}	520 °C	T_{sub}	475 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	903 °C	T_{In}	778 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{Sb}	520 °C	T_{Ga}	947 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C
		$\varnothing_{As}$	170 mil	T_{cr}	850 °C	$\varnothing_{Sb}$	100 %	T_{cr}	850 °C	$\varnothing_{As}$	170 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.3 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.85 ML/s
		x_t	644 nm	R_{gr}	1.0 ML/s	x_t	2.1 ML	R_{gr}	1.0 ML/s	x_t	102 nm
				x_t	5 nm			x_t	5 nm		
A1364	680	T_{sub}	680 °C	T_{sub}	560 °C	T_{sub}	560 °C	T_{sub}	475 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	903 °C	T_{In}	778 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{Sb}	530 °C	T_{Ga}	947 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C
		$\varnothing_{As}$	170 mil	T_{cr}	850 °C	$\varnothing_{Sb}$	100 %	T_{cr}	850 °C	$\varnothing_{As}$	170 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.3 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.85 ML/s
		x_t	204 nm	R_{gr}	1.0 ML/s	x_t	2.1 ML	R_{gr}	1.0 ML/s	x_t	102 nm
				x_t	5 nm			x_t	5 nm		
A1363	690	T_{sub}	680 °C	T_{sub}	580 °C	T_{sub}	580 °C	T_{sub}	475 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	903 °C	T_{In}	778 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{Sb}	530 °C	T_{Ga}	947 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C
		$\varnothing_{As}$	170 mil	T_{cr}	850 °C	$\varnothing_{Sb}$	100 %	T_{cr}	850 °C	$\varnothing_{As}$	170 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.3 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.85 ML/s
		x_t	204 nm	R_{gr}	1.0 ML/s	x_t	2.1 ML	R_{gr}	1.0 ML/s	x_t	102 nm
				x_t	5 nm			x_t	5 nm		
A1366	660	T_{sub}	680 °C	T_{sub}	620 °C	T_{sub}	620 °C	T_{sub}	475 °C	T_{sub}	680 °C
		T_{Ga}	955 °C	T_{In}	778 °C	T_{Ga}	903 °C	T_{In}	778 °C	T_{Ga}	947 °C
		T_{As}	380 °C	T_{Ga}	947 °C	T_{Sb}	530 °C	T_{Ga}	947 °C	T_{As}	380 °C
		T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C	T_{As}	380 °C	T_{cr}	850 °C
		$\varnothing_{As}$	170 mil	T_{cr}	850 °C	$\varnothing_{Sb}$	100 %	T_{cr}	850 °C	$\varnothing_{As}$	170 mil
		R_{gr}	1.0 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.3 ML/s	$\varnothing_{As}$	170 mil	R_{gr}	0.85 ML/s
		x_t	204 nm	R_{gr}	1.0 ML/s	x_t	2.1 ML	R_{gr}	1.0 ML/s	x_t	102 nm
				x_t	5 nm			x_t	5 nm		

Growth Step	Duration (s)
GaAs buffer deposition	See Table C.9
T_{sub} and T_{Ga} reduced under As flux	360
As valve closed	60
Growth interruption	300
Sb valve open	30
GaSb deposition	See Table C.9
T_{sub} reduced under Sb flux	120
As valve open	20
Sb valve closed	2
GaAs cap deposition	See Table C.9
T_{sub} and T_{Ga} increased under As flux	360
GaAs barrier deposition	See Table C.9
T_{sub} and T_{Ga} reduced under As flux	360
As valve closed	60
Growth interruption	300
Sb valve open	30
GaSb deposition	See Table C.9
Growth interruption under Sb flux	120
T_{sub} reduced to 250 °C under As flux	240
Sb valve closed and T_{sub} reduced to 250 °C	240

Table C.7: Growth protocol for the reference GaSb/GaAs QD/QRs sample: A1447. The Ga shutter closes at the end of deposition. The cyan steps correspond to the repeated sequence (x3).

Growth Step	Duration (s)
GaAs buffer deposition	See Table C.9
T_{sub} and T_{Ga} reduced under As flux	360
As valve closed	60
Growth interruption	300
Sb valve open	30
GaSb deposition	See Table C.9
T_{sub} reduced under Sb flux	120
As valve open	20
Sb valve closed	2
$In_{0.15}Ga_{0.85}As$ cap deposition	See Table C.9
GaAs protector deposition	See Table C.9
T_{sub} increased under As flux	360
T_{sub} reduced under As flux	300
$In_{0.15}Ga_{0.85}As$ QW deposition	See Table C.9
GaAs protector deposition	See Table C.9
T_{sub} and T_{Ga} increased under As flux	360
GaAs barrier deposition	885
T_{sub} and T_{Ga} reduced under As flux	240
As valve closed	60
Growth interruption	300
Sb valve open	30
GaSb deposition	See Table C.9
T_{sub} reduced under Sb flux	120
As valve open	20
Sb valve closed	2
$In_{0.15}Ga_{0.85}As$ cap deposition	See Table C.9
GaAs protector deposition	See Table C.9
T_{Ga} increased under As flux	300
As valve closed and T_{sub} reduced to 450 °C under Sb flux	240
Sb valve closed and T_{sub} reduced to 250 °C	240

Table C.8: Growth protocol for the asymmetric $In_{0.06}Ga_{0.94}As/GaSb/GaAs$ RWELL samples: A1458, A1459, A1464. The Ga shutter closes at the end of deposition. The cyan steps correspond to the repeated sequence (x3).

Table C.9: Growth parameters used for the asymmetric $In_{0.06}Ga_{0.94}As$/GaSb/GaAs RWELL samples.

A1447 — Desorp. T = 710 °C

	Buffer Layer	QR Layers	Cap Layers	QW Layers	Barrier Layers
T_{sub}	670 °C	540 °C	475 °C		670 °C
T_{Ga}	950 °C	899 °C	899 °C		950 °C
T_{As} / T_{Sb}	400 °C	445 °C	400 °C		400 °C
T_{cr}	850 °C	850 °C	850 °C	-	850 °C
$\varnothing_{As}$ / $\varnothing_{Sb}$	120 mil	100 %	120 mil		120 mil
R_{gr}	1.0 ML/s	0.3 ML/s	0.30 ML/s		1.0 ML/s
x_{t}	254 nm	2.1 ML	5 nm		250 nm

A1458 — Desorp. T = 700 °C

	Buffer Layer	QR Layers	Cap Layers	QW Layers	Barrier Layers
T_{sub}	670 °C	540 °C	475 °C		670 °C
T_{In}			706 °C		
T_{Ga}	950 °C	899 °C	899 °C		950 °C
T_{As} / T_{Sb}	400 °C	445 °C	400 °C	-	400 °C
T_{cr}	850 °C	850 °C	850 °C		850 °C
$\varnothing_{As}$ / $\varnothing_{Sb}$	120 mil	100 %	120 mil		120 mil
R_{gr}	1.0 ML/s	0.3 ML/s	0.32 ML/s		1.0 ML/s
x_{t}	339 nm	2.1 ML	5 nm		250 nm

A1459 — Desorp. T = 700 °C

	Buffer Layer	QR Layers	Cap Layers	QW Layers	Barrier Layers
T_{sub}	670 °C	540 °C	475 °C	540 °C	670 °C
T_{In}			706 °C	706 °C	
T_{Ga}	950 °C	899 °C	899 °C	899 °C	950 °C
T_{As} / T_{Sb}	400 °C	445 °C	400 °C	400 °C	400 °C
T_{cr}	850 °C	850 °C	850 °C	850 °C	850 °C
$\varnothing_{As}$ / $\varnothing_{Sb}$	120 mil	100 %	120 mil	120 mil	120 mil
R_{gr}	1.0 ML/s	0.3 ML/s	0.32 ML/s	0.32 ML/s	1.0 ML/s
x_{t}	407 nm	2.1 ML	5 nm	5 nm	250 nm

A1464 — Desorp. T = 700 ± 20 °C

	Buffer Layer	QR Layers	Cap Layers	QW Layers	Barrier Layers
T_{sub}	670 °C	540 °C	475 °C	540 °C	670 °C
T_{In}			706 °C	706 °C	
T_{Ga}	950 °C	899 °C	899 °C	899 °C	950 °C
T_{As} / T_{Sb}	400 °C	445 °C	400 °C	400 °C	400 °C
T_{cr}	850 °C	850 °C	850 °C	850 °C	850 °C
$\varnothing_{As}$ / $\varnothing_{Sb}$	120 mil	100 %	120 mil	120 mil	120 mil
R_{gr}	1.0 ML/s	0.3 ML/s	0.32 ML/s	0.32 ML/s	1.0 ML/s
x_{t}	373 nm	2.1 ML	5 nm	10 nm	250 nm

Bibliography

[1] Department for Levelling Up, Housing and Communities, HM Government, "Levelling Up the United Kingdom," *Chapter* 3 *The Policy Programme,* ISBN 978-1-5286-3017-7, CP 604, Feb. 2022.

[2] K. Noll, "10G-EPON vs. XGS-PON: Are They Really All That Different?," *CableLabs,* [Online] https://www.cablelabs.com/blog/generative-ai-for-network-operations-ai-cablelabs-expert [Last Accessed:] 19.10.2024.

[3] X. Chen, Y. Huo, S. Cho, B.-G. Park, and J. S. Harris, "Surface Treatment of Ge Grown Epitaxially on Si by Ex-Situ Annealing for Optical Computing by Ge Technology," *IEIE Trans. Smart Process. Comput.,* vol. 3, no. 4, pp. 331-337, Oct. 2014.

[4] P. Karioja, and D. Howe, "Diode-laser-to-waveguide butt coupling," *Appl. Opt.,* vol. 35, no. 3, pp. 404-416, Jan. 1996.

[5] A. Marent *et al.,* "10^6 years extrapolated hole storage time in GaSb/AlAs quantum dots," *Appl. Phys. Lett.,* vol. 91, no. 24, 242109, Dec. 2007.

[6] T. Egawa, T. Soga, T. Jimbo and M. Umeno, "Room-temperature continuous-wave operation of AlGaAs-GaAs single-quantum-well lasers on Si by metalorganic chemical-vapor deposition using AlGaAs-AlGaP intermediate layers," *IEEE J. Quantum Electron.,* vol. 27, no. 6, pp. 1798-1803, Jun. 1991.

[7] M. Akiyama, Y. Kawarada and K. Kaminishi, "Growth of single domain GaAs layer on (100)-oriented Si substrate by MOCVD," *Jpn. J. Appl. Phys.,* vol. 23, no. 11A, pp. L843-L845, Nov. 1984.

[8] Z. I. Kazi, P. Thilakan, T. Egawa, M. Umeno, and T. Jimbo, "Realization of GaAs/AlGaAs lasers on Si using epitaxial lateral overgrowth by metalorganic chemical vapor deposition," *Jpn. J. Appl. Phys.,* vol. 40, no. 8, pp. 4903-4906, Aug. 2001.

[9] C. S. C. Barrett *et al.,* "Effect of bulk growth temperature on antiphase domain boundary annihilation rate in MOCVD-grown GaAs on Si(001)," *J. Cryst. Growth,* vol 450, pp. 39-44, Sept. 2016.

[10] N. A. Shepelin, Z. P. Tehrani, N. Ohannessian, C. W. Schneider, D. Pergolesi, and T. Lippert, "A practical guide to pulsed laser deposition," *Chem. Soc. Rev.,* vol. 52, no. 7, pp. 2294-2321, Mar. 2023.

[11] M. Hayne and P. D. Hodgson, "Vertical-cavity surface-emitting laser," WO. Patent GB2016/050550, 9 Sept., 2016.

[12] D. Mahgerefteh *et al.,* "Techno-economic comparison of silicon photonics and multimode VCSELs," *IEEE J. Lightwave Technol.,* vol. 34, no. 2, pp. 233-242, Jan. 2016.

[13] H. Li, P. Wolf, P. Moser, G. Larisch, J. A. Lott, and D. Bimberg, "Vertical-cavity surface-emitting lasers for optical interconnects," *SPIE Newsroom,,* DOI: 10.1117/2.1201411.005689, Nov, 2014.

[14] A. Larsson, "Advances in VCSELs for communication and sensing," *IEEE J. Sel. Top. Quantum Electron.,* vol. 17, no. 6, pp. 1552-1567, Apr. 2011.

[15] D. G. Deppe, M. Li, X. Yang, and M. Bayat, "Advanced VCSEL Technology: Self-Heating and Intrinsic Modulation Response," *IEEE J. Quantum Electron.,* Vol. 54, no. 3, 2400209, Apr. 2018.

[16] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 μm," *Electron. Lett.,* vol. 15, no. 4, pp. 106-108, Feb. 1979.

[17] B. Chomycz, "Chromatic Dispersion," in *Planning Fibre Optic Networks.* New York, NY, USA: McGrawHill, ch. 4, pp. 106, 2009.

[18] S. Mokkapati, and C. Jagadish, "III-V compound SC for optoelectronic devices," *Mater. Today,* vol. 12, no. 4, pp. 22-32, Apr. 2009.

[19] L. C. Kimerling, "Recombination enhanced defect reactions," *Solid State Electron.,* vol. 21, no. 11-12, pp. 1391–1401, Nov. 1978.

[20] D. Jung *et al.,* "Impact of threading dislocation density on the lifetime of InAs quantum dot lasers on Si," *Appl. Phys. Lett.,* vol. 112, no. 15, 153507, Apr. 2018.

[21] A. Y. Liu, R. W. Herrick, O. Ueda, P. M. Petroff, A. C. Gossard and J. E. Bowers, "Reliability of InAs/GaAs quantum dot lasers epitaxially grown on

silicon," *IEEE J. Sel. Top. Quantum Electron.,* vol. 21, no. 6, 1900708, Apr. 2015.

[22] P. D. Hodgson, R. J. Young, M. A. Kamarudin, Q. D. Zhuang and M. Hayne, "Hole migration and optically induced charged depletion in GaSb/GaAs wetting layers and quantum rings," *Phys. Rev. B,* vol. 88, no. 15, 155322, Oct. 2013.

[23] F. Hatami *et al.,* "Carrier dynamics in type-II GaSb/GaAs quantum dots," *Phys. Rev. B,* vol. 57, no. 8, pp. 4635-4641, Feb. 1998.

[24] B. Bansal, M. Hayne, M. Geller, D. Bimberg and V. V. Moshchalkov, "Excitonic Mott transition in type-II quantum dots," *Phys. Rev. B,* vol. 77, no. 24, 241304(R), Jun. 2008.

[25] I. L. Kuskovsky *et al.,* "Optical Aharonov-Bohm effect in stacked type-II quantum dots," *Phys. Rev. B,* vol. 76, no. 3, 035342 Jul. 2007.

[26] P. D. Hodgson, M. Hayne, A. J. Robson, Q. D. Zhuang and L. Danos, "GaSb quantum rings in GaAs/Al$_x$Ga$_{1-x}$As quantum wells," *J. Appl. Phys.,* vol. 119, no. 4, 044305, Jan. 2016.

[27] G. Kiršanskė, P. Tighineanu, R. S. Daveau, J. Miguel-Sánchez, P. Lodahl and S. Stobbe, "Observation of the exciton Mott transition in the photoluminescence of coupled quantum wells," *Phys. Rev. B,* vol. 94, no. 15, 155438, Oct. 2016.

[28] F. Qiu *et al.,* "An investigation of exciton behavior in type-II self-assembled GaSb/GaAs quantum dots," *Nanotechnology,* vol. 27, no. 6, 065602, Dec. 2015.

[29] T. Nowozin *et al.,* "800 meV localization energy in GaSb/GaAs/Al$_{0.3}$Ga$_{0.7}$As quantum dots," *Appl. Phys. Lett.,* vol. 102, no. 5, 052115, Feb. 2013.

[30] M. J. Pullin *et al.,* "Room-temperature InAsSb strained-layer superlattice light-emitting diodes at $\lambda = 4.2$ μm with AlSb barriers for improved carrier confinement," *Appl. Phys. Lett.,* vol. 74, no. 16, pp. 2384-2386, Apr. 1999.

[31] K.-S. Lee, E.-H. Lee, C.-D. Lee, S. K. Noh, and Y. Kim, "Influence of thin AlAs interface layers on exciton transitions in InGaAs/GaAs quantum wells," *Microelectron. Eng.,* Vol. 43–44, pp. 237-242, Aug. 1998.

[32] W. Wang, X. Gan, Y. Xu, T. Wang, H. Wu and C. Liu, "High-quality n-type aluminum gallium nitride thin films grown by interrupted deposition and in-situ thermal annealing," *Mater. Sci. Semicond. Process.,* vol. 30, pp. 612-617, Feb. 2015.

[33] N. Fernández-Delgado *et al.,* "Effect of an in-situ thermal annealing on the structural properties of self-assembled GaSb/GaAs quantum dots," *Appl. Surf. Sci.,* vol. 395, pp. 136-139, Feb. 2017.

[34] J. Z. Li *et al.,* "Defect reduction of GaAs epitaxy on Si (001) using selective aspect ratio trapping," *Appl. Phys. Lett.,* vol. 91, no. 2, 021114, Jul. 2007.

[35] V. K. Yang, M. Groenert, C. W. Leitz, A. J. Pitera, M. T. Currie and E. A. Fitzgerald, "Crack formation in GaAs heteroepitaxial films on Si and SiGe virtual substrates," *J. Appl. Phys.,* vol. 93, no. 7, pp. 3859-3865, Apr. 2003.

[36] M. E. Groenert *et al.,* "Monolithic integration of room-temperature cw GaAs/AlGaAs lasers on Si substrates via relaxed graded GeSi buffer layers," *J. Appl. Phys.,* vol. 93, no. 1, pp. 362-367, Jan. 2003.

[37] T. Soga, T. Jimbo and M. Umeno, "Low etch pit density GaAs on Si grown by metalorganic chemical vapor deposition," *Appl. Phys. Lett.,* vol. 56, no. 15, pp. 1433-1435, Apr. 1990.

[38] L. M. Giovane, H. C. Luan, A. M. Agarwal and L. C. Kimerling, "Correlation between leakage current density and threading dislocation density in SiGe p-i-n diodes grown on relaxed graded buffer layers," *Appl. Phys. Lett.,* vol. 78, no. 4, pp. 541-543, Jan. 2001.

[39] W. Y. Uen *et al.,* "Heteroepitaxial growth of GaAs on Si by MOVPE using a-GaAs/a-Si double-buffer layers," *J. Cryst. Growth,* vol. 295, no. 2, pp. 103-107, Oct. 2006.

[40] K. Papatryfonos *et al.,* "Low-Defect Quantum Dot Lasers Directly Grown on Silicon Exhibiting Low Threshold Current and High Output Power at Elevated Temperatures," *Adv. Photonics Res.,* 2400082. Jul. 2024.

[41] C. Shang *et al.,*"High-temperature reliable quantum-dot lasers on Si with misfit and threading dislocation filters," *Optica,* vol. 8, no. 5, pp. 749-754, May 2021.

[42] N. El-Marsy, J. C. L. Tarn, T. P. Humphreys, N. Hamaguchi, N. H. Karam and S. M. Bedair, "Effectiveness of strained-layer superlattices in reducing defects in GaAs epilayers grown on silicon substrates," *Appl. Phys. Lett.,* vol. 51, no. 20, pp. 1608-1610, Nov. 1987.

[43] R. Fischer *et al.,* "Material properties of high-quality GaAs epitaxial layers grown on Si substrates," *J. Appl. Phys.,* vol. 60, no. 5, pp. 1640-1647, Sept. 1986.

[44] S. Sharan, J. Narayan and J. C. C. Fan, "Dislocation density reduction in GaAs epilayers on Si using strained layer superlattices," *J. Electron. Mater.,* vol. 20, pp. 779-784, Jul. 1991.

[45] H. Kim, D.-M. Geum, Y.-H. Ko and W.-S. Han, "Effects of High-Temperature Growth of Dislocation Filter Layers in GaAs-on-Si," *Nanoscale Res. Lett.,* vol. 17, no. 1, 126, Dec. 2022.

[46] A.-S. Mahajumi *et al.,* "Rapid thermal annealing and photoluminescence of type-II GaSb single monolayer quantum dot stacks," *J. Phys. D Appl. Phys.,* vol. 46, 305104, Jul. 2013.

[47] M. A. Kamarudin, M. Hayne, R. J. Young, Q. Zhuang, T. Ben and S. I. Molina, "Tuning the properties of exciton complexes in self-assembled GaSb/GaAs quantum rings," *Phys. Rev. B,* vol. 83, no. 11, 115311, Mar 2011.

[48] S. Kiravittaya, Y. Nakamura and O. G. Schmidt, "Photoluminescence linewidth narrowing of InAs/GaAs self-assembled quantum dots," *Phys. E Low Dimens. Syst. Nanostruct.,* vol. 13, no. 2-4, pp. 224-248, Mar. 2002.

[49] S. R. Bank *et al.,* "Recent Progress on 1.55-µm dilute-nitride lasers," *IEEE J. Quantum Electron.,* vol. 43, no. 9, Sept. 2007.

[50] W. Pan, L. Wang, Y. Zhang, W. Lei, and S. Wang, "MBE growth strategy and optimization of GaAsBi quantum well light emitting structure beyond 1.2 µm," *Appl. Phys. Lett.,* vol. 114, no. 15, 152102, Apr. 2019.

[51] S. J. Sweeney, "Bismide-alloys for higher efficiency infrared semiconductor lasers," in *22nd IEEE Int. Semicond. Laser Conf.,* Kyoto, Japan, 2010, pp. 111-112.

[52] I. P. Marko and S. J. Sweeney, "Progress Toward III–V Bismide Alloys for Near- and Midinfrared Laser Diodes," *IEEE J. Quantum Electron.,* vol. 23, no. 6, 1501512, Jun. 2017.

[53] M. A. Majid *et al.,* "Toward 1550-nm GaAs-based lasers using InAs/GaAs quantum dot bilayers," *IEEE J. Sel. Top. Quantum Electron.,* vol. 17, no. 5, pp. 1334-1342, Sept. 2011.

[54] W.-H. Lin, K.-W. Wang, S.-Y. Lin and M.-C. Wu, "Temperature-dependent photoluminescence and carrier dynamics of standard and coupled type-II GaSb/GaAs quantum rings," *J. Cryst. Growth,* vol. 378, pp. 426-429, Sept. 2013.

[55] F. Schäfer, B. Mayer, J. P. Reithmaier, and A. Forchel, "Improved carrier confinement in GaInAs/AlGaAs lasers by MBE grown short period superlattice quantum well barriers." *J. Cryst. Growth,* vol. 201-202, pp. 914-918, May 1999.

[56] P. J. Carrington, A.-S. Mahajumi, M. C. Wagener, J. R. Botha, Q. Zhuang, A. Krier, "Type II GaSb/GaAs quantum dot/ring stacks with extended photoresponse for efficient solar cells," *Phys. B Condens. Matter,* vol. 407, no. 10, pp. 1493-1496, May 2012.

[57] P. J. Carrington, R. J. Young, P. D. Hodgson, A. M. Sanchez, M. Hayne and A. Krier, "Long-Wavelength photoluminescence from stacked layers of high quality type-II GaSb/GaAs quantum rings," *Cryst. Growth Des.,* vol. 13, no. 3, pp. 1226-1230, Jan. 2013.

[58] V. Lopes-Oliveira *et al.,* "Structural and magnetic confinement of holes in the spin-polarized emission of coupled quantum ring–quantum dot chains," *Phys. Rev. B,* vol. 90, no. 12, 125315, Sept. 2014.

[59] W.-H. Lin, K.-W. Wang, S.-Y. Lin, and M.-C. Wu, "Long-wavelength electroluminescence of InGaAs-capped type-II GaSb/GaAs quantum-rings at room temperature," *J. Cryst. Growth,* vol. 378, pp. 571-575, Sept. 2013.

[60] D. A. B. Miller, "Device requirements for optical interconnects to silicon chips," *Proc. IEEE,* vol. 97, no. 7, pp. 1166-1185, Aug. 2009.

[61] P. M. Petroff and R. L. Hartman, "Rapid degradation phenomenon in heterojunction GaAlAs-GaAs lasers," *J. Appl. Phys.,* vol. 45, no. 9, pp. 3899-3903, Sept. 1974.

[62] T. Wang, H. Liu, A. Lee, F. Pozzi and A. Seeds, "1.3 μm InAs/GaAs quantum-dot lasers monolithically grown on Si substrates," *Opt. Exp.,* vol. 19, no. 12, pp. 11381-11386, May 2011.

[63] A. Lee, Q. Jiang, M. Tang, A. Seeds and H. Liu, "Continuous-wave InAs/GaAs quantum-dot laser diodes monolithically grown on Si substrate with low threshold current densities," *Opt. Exp.,* vol. 20, no. 20, pp. 22181-22187, Sept. 2012.

[64] S. Chen *et al.,* "Electrically pumped continuous-wave III-V quantum dot lasers on silicon," *Nat. Photon.,* vol. 10, pp. 307-311, Mar. 2016.

[65] D. G. Deppe, K. Shavritranuruk, G. Ozgur, H. Chen and S. Freisem, "Quantum dot laser diode with low threshold and low internal loss," *Electronics Letters,* vol. 45, no. 1, pp. 54-56, Feb. 2009.

[66] A. Y. Liu *et al.,* "High performance continuous wave 1.3 µm quantum dot lasers on silicon," *Appl. Phys. Lett.,* vol. 104, no. 4, 041104, Jan. 2014.

[67] A. Du, C. Cao, S. Han, H. Wang, and Q. Gong, "Continuous-wave operation of Si-based 1.31 µm InAs/GaAs quantum-dot laser grown by molecular beam epitaxy," *Phys. Scr.,* vol. 98, no. 8, 085523, Jul. 2023.

[68] Z. Lv *et al.,* "Ultra-high thermal stability InAs/GaAs quantum dot lasers grown on on-axis Si (001) with a record-high continuous-wave operating temperature of 150 °C," *Opt. Express,* vol. 31, no. 15, pp. 24173-24182, Jul. 2023.

[69] C. Shang *et al.* "Perspectives on Advances in Quantum Dot Lasers and Integration with Si Photonic Integrated Circuits," *ACS Photonics,* vol. 8, no. 9, pp. 2555-2566, Aug. 2021.

[70] H. Liu *et al.,* "Long-wavelength InAs/GaAs quantum-dot laser diode monolithically grown on Ge substrate," *Nat. Photonics,* vol. 5, pp. 416-419, Jun. 2011.

[71] A. D. Lee, Q. Jiang, M. Tang, Y. Zhang, A. J. Seeds and H. Liu, "InAs/GaAs quantum-dot lasers monolithically grown on Si, Ge, and Ge-on-Si substrates," *IEEE J. Sel. Top. Quantum Electron.,* vol. 19, no. 4, 1901107, Jul. 2013.

[72] S. Zhu, B. Shi, Q. Li and K. M. Lau, "1.5 µm quantum-dot diode lasers directly grown on CMOS-standard (001) silicon," *Appl. Phys. Lett.,* vol. 113, no. 22, 221103, Nov. 2018.

[73] R. Beanland *et al.,* "Structural analysis of life tested 1.3 µm quantum dot lasers," *J. Appl. Phys.,* vol. 103, no. 1, 014913, Jan. 2008.

[74] T. Kageyama *et al.,* "Long-wavelength quantum dot FP and DFB lasers for high temperature applications," in *Proc. SPIE 8277, Novel In-Plane Semiconductor Lasers XI,* San Francisco, California, US, 82770C, Mar. 2012.

[75] Z. Yao *et al.,* "Recent Developments of Quantum Dot Materials for High Speed and Ultrafast Lasers," *Nanomaterials,* vol. 12, no. 7, 1058, Mar. 2022.

[76] J. Tatebayashi *et al.,* "Lasing characteristics of self-assembled quantum dots embedded in an InGaAs quantum well," *Appl. Phys. lett.,* vol. 90, no. 26, 261115, Jun. 2007.

[77] G. Roelkens *et al.,* "III-V-on-Silicon Photonic Devices for Optical Communication and Sensing," *Photonics,* vol 2, no. 3, pp. 969-1004, Sept. 2015.

[78] M. Liao *et al.,* "Monolithically integrated electrically pumped continuous-wave III-V quantum dot light sources on silicon," *IEEE J. Sel. Top. Quantum Electron.,* vol. 23, no. 6, 1900910, Dec. 2017.

[79] P. J. Carrington *et al.,* "InSb based quantum dot nanostructures for mid-infrared photonic devices," in *Proc. SPIE,* Sept. 2016, vol. 9919, 99190C.

[80] T. Nowozin *et al.,* "Linking structural and electronic properties of high-purity self-assembled GaSb/GaAs quantum dots," *Phys. Rev. B,* vol. 86, no. 3, 035305, Jul. 2012.

[81] W.-H. Lin, M.-Y. Lin, S.-Y. Wu and S.-Y. Lin, "Room-Temperature electroluminescence of type-II GaSb/GaAs quantum rings," *IEEE Photonics Technol. Lett.,* vol. 24, no. 14, pp. 1203-1205, Jul. 2012.

[82] E. F. Schubert, *Light-Emitting Diodes,* 3rd ed., New York, NY, USA: Cambridge University Press, 2018, ch. , pp.

[83] R. B. Laghumavarapu, A. Moscho, A. Khoshakhlagh, M. El-Emawy, L. F. Lester and D. L. Huffaker, "GaSb/GaAs type II quantum dot solar cells for enhanced infrared spectral response," *Appl. Phys. Lett.,* vol. 90, no. 17, 173125 Apr. 2007.

[84] T. Gu, M. A. El-Emawy, K. Yang, A. Stintz and L. F. Lester, "Resistance to edge recombination in GaAs-based dots-in-a-well solar cells," *Appl. Phys. Lett.,* vol. 95, no. 26, 261106, Dec. 2009.

[85] M. Geller, A. Marent, T. Nowozin, D. Bimberg, N. Akçay and N. Öncan, "A write time of 6ns for quantum dot–based memory structures," *Appl. Phys. Lett.,* vol. 92, no. 9, 092108, 2008.

[86] M. Hayne *et al.,* "The structural, electronic and optical properties of GaSb/GaAs nanostructures for charge-based memory," *J. Phys. D Appl. Phys.,* vol. 46, no. 26, 264001, Jun. 2013.

[87] C. Couteau *et al.,* "Applications of single photons to quantum communication and computing," *Nat Rev Phys,* vol. 5, no. 6, pp. 326-338, May 2023.

[88] I. Choi, R. Young, and P. Townsend, "Quantum information to the home," *New Journal of Physics,* vol. 13, no. 6, 063039, Jun. 2011.

[89] Z.-S. Chen, B. Ma, X. K. Shang, H.-Q. Ni, J.-L. Wang, Z.-C. Niu. "Bright Single-Photon Source at 1.3 μm Based on InAs Bilayer Quantum Dot in Micropillar," *Nanoscale Res. Lett.,* vol. 12, no. 1, 378, May 2017.

[90] G-C. Shan, Z-Q. Yin, C. H. Shek, W. Huang, Single photon sources with single semiconductor quantum dots. Frontiers of Physics 9(2), (2014).

[91] M. P. Young *et al.,* "Photoluminescence studies of individual and few GaSb/GaAs quantum rings," *AIP Adv.,* vol. 4, no. 11, 117127, Nov. 2014.

[92] M. A. Kamarudin, "Growth and optical properties of self-assembled GaSb/GaAs quantum dots," Ph.D. thesis, Dept. Phys., Lancaster Univ., Lancaster, 2010.

[93] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties* in Graduate Texts in Physics, 4th ed. Heidelberg, DE: Springer Berlin, 2010, p. 260.

[94] T. Edvinsson, "Optical quantum confinement and photocatalytic properties in two-, one- and zero-dimensional nanostructures," *R. Soc. Open Sci.,* vol. 5, no. 9, 180387, Sept. 2018.

[95] W. Pauli, *Wave Mechanics: Pauli Lectures on Physics* in Dover Books on Physics, vol. 5. C. P. Enz, Ed., Mineola, NY, USA: Dover Publications Inc., 2003, pp.

[96] J. Singh and R. T. Williams, Eds. *Excitonic and Photonic Processes in Materials* in Springer Series in Materials Science, vol. 203. Singapore, SG: Springer, 2014, pp.

[97] C. Kittel, P. McEuen and R. Lewis, *Introduction to Solid State Physics,* 8th ed. S. Ingrao and S. Johnson, Eds., New York, NY, USA: John Wiley & Sons, 2005, pp.

[98] N. W. Ashcroft and N. D. Mermin, "The Tight-Binding Method," in *Solid State Physics.* D. Garbose Crane, Ed., Philadelphia, PA, USA: Saunders College Publishing, 1976, ch. 10, p. 183.

[99] N. W. Ashcroft and N. D. Mermin, "Determination of Crystal Structures by X-Ray Diffraction," in *Solid State Physics.* D. Garbose Crane, Ed., Philadelphia, PA, USA: Saunders College Publishing, 1976, ch. 6, p. 99.

[100] J.-W. Luo, G. Bester, and A. Zunger, "Full-Zone Spin Splitting for Electrons and Holes in Bulk GaAs and GaSb," *Phys. Rev. Lett.,* vol. 102, no. 5-6, 056405, Feb. 2009.

[101] Y. P. Varshni, "Temperature dependence of the energy gap in semiconductors," *Phys.,* vol. 34, no. 1, pp. 149-154, 1967.

[102] M. A. Omar, *Elementary Solid State Physics: Principles and Applications* in Addison-Wesley series in solid state sciences. D. Lazarus and J. Rey, Eds., USA: Pearson Education, 1993.

[103] G. Patriarche, L. Largeau, J.-C. Harmand and D. Gollub, "Morphology and composition of highly strained InGaAs and InGaAsN layers grown on GaAs substrate," *Appl. Phys. Lett.,* vol. 84, no. 2, pp. 203-205, Jan. 2004.

[104] J.-S. Park, M. Tang, S. Chen and H. Liu, "Heteroepitaxial Growth of III-V Semiconductors on Silicon," *Crystals,* vol. 10, no. 12, 1163, Dec. 2020.

[105] T. Hom, W. Kiszenik and B. Post, "Accurate lattice constants from multiple reflection measurements. II. Lattice constants of germanium silicon, and diamond," *J. Appl. Cryst.,* vol. 8, no. 4, pp. 457-458, Aug. 1975.

[106] Y. Okada and Y. Tokumaru, "Precise determination of lattice parameter and thermal expansion coefficient of silicon between 300 and 1500 K," *J. Appl. Phys* vol. 56, no. 2, pp. 314-320, Jul. 1984.

[107] I. Vurgaftman, J. R. Meyer and L. R. Ram-Mohan, "Band parameters for III–V compound semiconductors and their alloys," *J. Appl. Phys.,* vol. 89, no. 11, pp. 5815-5875, Jun. 2001.

[108] S. Adachi, *Thermal Properties,* in "Physical Properties of III-V Semiconductor compounds,". New York, NY, USA: John Wiley & Sons Inc., 1992, ch. 4, pp. 53, ISBN: 0-471-57329-9.

[109] T. A. Nilsen, M. Breivik, G. Myrvågnes; B.-O. Fimland, "Thermal expansion of GaSb measured by temperature dependent x-ray diffraction," *J. Vac. Sci. Technol. B,* vol. 28, no. 3, pp. C3I17–C3I20, May 2010.

[110] G. S. Parkinson *et al.,* "Antiphase domain boundaries at the $Fe_3O_4(001)$ surface," *Phys. Rev. B,* vol. 85, no. 19, 195450, May 2012.

[111] L. Vegard, "Die konstitution der mischkristalle und die raumfüllung der atome". *Z. Physik,* Vol. 5, no. 1, pp. 17-26, Jan. 1921.

[112] K. T. Jacob, S. Raj, and L. Rannesh, "Vegard's law: a fundamental relation or an approximation?," *Int. J. Mat. Res.,* vol. 98, no. 9, pp. 776-779, Sept. 2007.

[113] A. Mascarenhas, S. Francoeur and S. Yoon, "Physics of Isoelectronic Dopants in GaAs," in *Dilute Nitride Semiconductors,* M. Henini Ed., Elsevier, ch. 6, pp. 179-221, 2005.

[114] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers. I. Misfit dislocations," *J. Cryst. Growth,* vol. 27, pp. 118-125, Dec. 1974.

[115] E. A. Fitzgerald, "Dislocations in strained-layer epitaxy: theory, experiment, and applications," *Mater. Sci. Rep.,* vol. 7, no. 3, pp. 87-142, Nov. 1991.

[116] T. J. Gosling and J. R. Willis, "The energy of arrays of dislocations in an anisotropic half-space," *Phil. Mag. A*, vol. 69, no. 1, pp. 65-90, 1994.

[117] D. J. Dunstan, "Strain and strain relaxation in semiconductors," *J. Mater. Sci. Mater. Electron.,* vol. 8, pp. 337-375, Dec. 1997.

[118] A. Braun, K. M. Briggs, P. B oni, "Analytical solution to Matthews' and Blakeslee's critical dislocation formation thickness of epitaxially grown thin films," *J. Crys. Growth,* vol. 241, no. 1-2, pp.231-234, Jan. 2002.

[119] C. G. Van de Walle, "Band lineups and deformation potentials in the model-solid theory," *Phys. Rev. B,* vol. 39, no. 3, pp. 1871-1883, Jan. 1989.

[120] E. Bauer, "Phänomenologische theorie der kristallabscheidung an oberflächen. I," *Z. Kristallogr.,* vol. 110, no. 1-6, pp. 372-394, Jan. 1958.

[121] R. Kaischew, Equilibrium shape and work of formation of crystalline nuclei on a foreign substrate. Commun. Bulg. Acad. Sci, (1950).

[122] I. Markov, *Crystal Growth for Beginners: Fundamentals of Nucleation, Crystal Growth and Epitaxy,* Singapore World Scientific, 1995. ch. , pp. .

[123] R. Peierls, "Clustering in adsorbed films," *Phys. Rev. B,* vol. 18, no. 4, pp. 2013-2015, Aug. 1978.

[124] T. Young, "An Essay on the Cohesion of Fluids," *Phil. Trans. R. Soc. Lond.,* vol. 95, pp. 65-87, Jan. 1805.

[125] E. G. Shafrin and W. A. Zisman, "Constitutive relations in the wetting of low energy surfaces and the theory of the retraction method of preparing monolayers," *J. Phys. Chem.,* vol. 64, no. 5, pp. 519-524, May. 1960.

[126] S. Azizian, M. Khosravi, Advanced Low-Cost Separation Techniques in Interface Science, Interface Science and Technology 2019, ch. , pp. .

[127] R. J. Good, "Contact angle, wetting, and adhesion: a critical review," *J. Adhes. Sci. Technol.,* vol. 6, no. 12, pp. 1269-1302, 1992.

[128] J. E. Prieto and I. Markov, "Stranski–Krastanov mechanism of growth and the effect of misfit sign on quantum dots nucleation," *Surf. Sci.,* vol. 664, pp. 172-184, Oct. 2017.

[129] J. Tersoff and F. K. LeGoues, "Competing relaxation mechanisms in strained layers," *Phys. Rev. Lett.,* vol. 72, no. 22, pp. 3570-3573, May 1994.

[130] A. Stemmann, Ch. Heyn, T. Köppen, T. Kipp and W. Hansen, "Local droplet etching of nanoholes and rings on GaAs and AlGaAs surfaces," *Appl. Phys. Lett.,* vol. 93, no. 12, 123108, Sept. 2008.

[131] H. Yonezu, "Dislocation-free III-V-N Alloy Layers on Si Substrates and Their Device Applications," in *Dilute Nitride Semiconductors,* M. Henini Ed., Elsevier, 2005, ch. 14, pp. .

[132] E. P. Smakman, J. K. Garleff, R. J. Young, M. Hayne, P. Rambabu and P. M. Koenraad, "GaSb/GaAs quantum dot formation and demolition studied with cross-sectional scanning tunneling microscopy," *Appl. Phys. Lett.,* vol. 100, no. 14, 142116, Apr. 2012.

[133] T. Walther, A. G. Cullis, D. J. Norris and M. Hopkinson, "Nature of the Stranski-Krastanow Transition during Epitaxy of InGaAs on GaAs," *Phys. Rev. Lett.,* vol. 86, no. 11, pp. 2381-2384, Mar. 2001.

[134] S.-H. Bae *et al.,* "Graphene-assisted spontaneous relaxation towards dislocation-free heteroepitaxy," *Nat. Nanotechnol.,* vol. 15, no. 4, pp. 272-276, Feb. 2020.

[135] W. Hagen and H. J. Queisser, "In situ x-ray topography of epitaxial Ge layers during growth," *Appl. Phys. Lett.,* vol. 32, no. 5, pp. 269-270, Mar. 1978.

[136] J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara and I. K. Robinson, "Ge_xSi_{1-x}/Si strained-layer superlattice grown by molecular beam epitaxy," *J. Vac. Sci. Technol. A,* vol. 2, no. 2, pp. 436-440, Apr. 1984.

[137] C. L. M. Daunt *et al.,* "Compact electroabsorption modulators for photonic integrated circuits, using an isolated pedestal contact scheme," *IEEE Photonics Technol. Lett.,* vol. 24, no. 5, pp. 356-358, Mar. 2012.

[138] J. Villain, "Various types of growth instabilities in coherent epitaxy," *J. Cryst. Growth,* vol. 275, no. 1-2, pp. e2307-e2311, Feb. 2005.

[139] S. I. Molina *et al.,* "Direct imaging of quantum wires nucleated at diatomic steps," *Appl. Phys. Lett.,* vol. 91, no. 14, 143112, Oct. 2007.

[140] F. Bastiman, "Epitaxy Explained, Little known MBE facts: High Purity Material," *Dr Faebian Bastiman, Epitaxy Explained (Wordpress).* Apr. 2015. [Online]. Available: https://faebianbastiman.wordpress.com/2015/04/12/little-known-mbe-facts-high-purity-material/ [Accessed: 6 Jun. 2023].

[141] A. H. Cottrell, *The Mechanical Properties of Matter,* in "Wiley Series on the Science and Technology of Materials," vol. 24. New York, NY, USA: John Wiley & Sons, 1964, pp. 67-74.

[142] R. L. Headrick, I. K. Robinson, E. Vlieg and L. C. Feldman, "Structure determination of the Si(111):B($\sqrt{3} \times \sqrt{3}$)R30° surface: Subsurface substitutional doping," *Phys. Rev. Lett.,* vol. 63, no. 12, pp. 1253-1256, Sept. 1989.

[143] A. F. J. Levi, S. L. McCall and P. M. Platzman, "Nonrandom doping and elastic scattering of carriers in semiconductors," *Appl. Phys. Lett.,* vol. 54, no. 10, pp. 940-942, Mar. 1989.

[144] E. F. Schubert, G. H. Gilmer, R. F. Kopf and H. S. Luftman, "Maximum concentration of impurities in semiconductors," *Phys. Rev. B,* vol. 46, no. 23, 15078, Dec. 1992.

[145] H. Raebiger, H. Nakayama and T. Fujita, "Control of defect binding and magnetic interaction energies in dilute magnetic semiconductors by charge state manipulation," *J. Appl. Phys.,* vol. 115, no. 1, 012008, Jan. 2014.

[146] T. F. K. Weatherley *et al.,* "Imaging Nonradiative Point Defects Buried in Quantum Wells Using Cathodoluminescence," *Nano Lett.,* vol. 21, no. 12, pp. 5217-5224, Jun. 2021.

[147] S. Birindelli *et al.,* "Peculiarities of the hydrogenated In(AsN) alloy," *Semicond. Sci. Technol.,* vol. 30, no. 10, 105030, Oct. 2015.

[148] C. W. Passchier and R. A. J. Trouw, *Microtectonics,* 2nd ed., Heidelberg, DE: Springer Berlin, 2005. ch. , pp. ,

[149] A. Mandal, U. Verma and S. Chakrabarti, "Effects of ex situ annealing on quaternary alloy (InAlGaAs) capped InAs/GaAs quantum dot heterostructures on optimization of optoelectronic and structural properties with variation in growth rate, barrier thickness, and seed quantum dot monolayer coverage," *Superlattices Microstruct.,* vol. 58, no. , pp. 101-119, Jun. 2013

[150] N. Fernández-Delgado *et al.,* "Atomic-column scanning transmission electron microscopy analysis of misfit dislocations in GaSb/GaAs quantum dots." *J. Mater. Sci.,* vol. 51, pp. 7691-7698, May 2016.

[151] A. Ourmazd, "The electrical properties of dislocations in semiconductors," *Contemp. Phys.,* vol. 25, no. 3, pp. 251-268, 1984.

[152] R. Hull, "Misfit Strain and Accommodation in SiGe Heterostructures," in *Semiconductors and Semimetals,* vol. 56. R. Hull and J. C. Bean, Eds., Elsevier, 1998, ch. 3, pp. 101-167.

[153] E. A. Stach and R. Hull, "Dislocations in Semiconductors," in *Encyclopedia of Materials: Science and Technology,* 2nd ed. K. H. J. Buschow *et al.,* Eds., Elsevier, 2001, pp. 2301-2312.

[154] A. R. Zanatta, "Revisiting the optical bandgap of semiconductors and the proposal of a unified methodology to its determination," *Sci. Rep.,* vol. 9, 11225, Aug. 2019.

[155] A. Haug, "Free Carrier Absorption in Semiconductor Lasers," *Semicond. Sci. Technol.,* vol. 7, no. 3, pp. 373-378, 1992.

[156] W. Grieshaber, E. F. Schubert, I. D. Goepfert, R. F. Karlicek Jr, M. J. Schurman and C. Tran, "Competition between band gap and yellow luminescence in GaN and its relevance for optoelectronic devices," *J. Appl. Phys.,* vol. 80, no. 8, pp. 4615-4620, Oct. 1996.

[157] P. J. Carrington, "Quantum nanostructures grown by molecular beam epitaxy for mid-infrared applications," Ph.D. thesis, Dept. Phys., Lancaster Univ., Lancaster, 2009.

[158] M. Hayne *et al.,* "Optically induced charging effects in self-assembled GaSb/GaAs quantum dots," *Phys. Rev. B,* vol. 70, no. 8, 081302(R), Aug. 2004.

[159] L. Müller-Kirsch *et al.,* "Many-particle effects in type II quantum dots," *Appl. Phys. Lett.,* vol. 78, no. 10, pp. 1418-1420, Mar. 2001.

[160] M. Shahmohammadi *et al.,* "Biexcitonic molecules survive excitons at the Mott transition," *Nat. Commun.,* vol. 5, 5251, Oct. 2014.

[161] B. Bansal, S. Godefroo, M. Hayne, G. Medeiros-Ribeiro and V. V. Moshchalkov, "Extended excitons and compact heliumlike biexcitons in type-II quantum dots," *Phys. Rev. B,* vol. 80, no. 20, 205317, Nov. 2009.

[162] M. Hayne *et al.,* "Electron localization by self-assembled GaSb/GaAs quantum dots," *Appl. Phys. Lett.,* vol. 82, no. 24, pp. 4355-4357, Jun. 2003.

[163] I. Pelant and J. Valenta, *Luminescence Spectroscopy of Semiconductors,* New York, NY, USA: Oxford University Press, 2012, pp. .

[164] M. Fox, *Optical Properties of Solids,* in Oxford Master Series in Condensed Matter Physics, 2nd ed., vol. 3. New York, NY, USA: Oxford University Press, 2010, pp. .

[165] M. Stern, V. Garminder, V. Umansky and I. Bar-Joseph, "Mott transition of excitons in coupled quantum wells," *Phys. Rev. Lett.,* vol. 100, no. 25, 256402, Jun. 2008.

[166] G. Duggan, and H. I. Ralph, "Exciton binding energy in type-II GaAs-(Al, Ga) As quantum-well heterostructures," *Phys. Rev. B,* vol. 35, no. , pp. 4152(R), Mar. 1987.

[167] X. L. Yang, S. H. Guo, F. T. Chan, K. W. Wong, and W. Y. Ching, "Analytic solution of a two-dimensional hydrogen atom. I. Nonrelativistic theory." *Phys. Rev. A,* vol. 43, no. 3, pp. 1186-1196, Feb. 1991.

[168] S. Dey *et al.,* "The confinement energy of quantum dots," 2012, arXiv:1212.2318.

[169] R. Heitz, F. Guffarth, I. Mukhametzhanov, M. Grundmann, A. Madhukar and D. Bimberg, "Many-body effects on the optical spectra of InAs/GaAs quantum dots," *Phys. Rev. B,* vol. 62, no. 24, 16881, Dec. 2000.

[170] P. D. Hodgson, *et al.,* "Blueshifts of the emission energy in type-II quantum dot and quantum ring nanostructures," *J. Appl. Phys.,* vol. 114, no. 7, 073519, Aug. 2013.

[171] T. Kawazu T and H. Sakaki "Effects of Sb/As intermixing on optical properties of GaSb type-II quantum dots in GaAs grown by droplet epitaxy," *Appl. Phys. Lett.,* vol. 97, no. 26, 261906, Dec. 2010.

[172] J. A. Keen, "Characterisation and simulation of InAs/InAsSb structures for mid-infrared LEDs," Ph.D. thesis, Dept. Phys., Lancaster Univ., Lancaster, 2018.

[173] H. Lipsanen, M. Sopanen, J. Ahopelto, J. Sandmann and J. Feldmann, "Effect Of Inp Passivation On Carrier Recombination in $In_xGa_{1-x}As$/GaAs Surface Quantum Wells," *Jpn. J. Appl. Phys.,* vol. 38, no. 25, pp. 1133-1134, Feb. 1999.

[174] P. J. Carrington, Q. Zhuang, M. Yin and A. Krier, "Temperature dependence of mid-infrared electroluminescence in type II InAsSb/InAs multi-quantum well light-emitting diodes," *Semicond. Sci. Technol.,* vol. 24, no. 7, 075001, May 2009.

[175] W. Liu *et al.,* "Impact of defects on Auger recombination in c-plane In-GaN/GaN single quantum well in the efficiency droop regime," *Appl. Phys. Lett.,* vol. 116, no. 22, 222106, Jun. 2020.

[176] J. R. Lindle, J. R. Meyer, C. A. Hoffman, F. J. Bartoli, G. W. Turner and H. K. Choi, "Auger lifetime in InAs, InAsSb, and InAsSb-InAlAsSb quantum wells," *Appl. Phys. Lett.,* vol. 67, no. 21, pp. 3153-3155, Nov. 1995.

[177] W. Shockley and W. T. Read Jr, "Statistics of the recombinations of holes and electrons," *Phys. Rev.,* vol. 87, no. 5, pp. 835-842, Sept. 1952.

[178] E. Finkeißen, M. Potemski, P. Wyder, L. Viña and G. Weimann, "Cooling of a semiconductor by luminescence up-conversion," *Appl. Phys. Lett.,* vol. 75, no. 9, pp. 1258-1260, Aug. 1999.

[179] J. R. Lakowicz, Principles of Fluorescence Spectroscopy (Kluwer Academic / Plenum Publishers p.10. ISBN 978-0-387-31278-1 (1999).

[180] S. Birner *et al.,* "nextnano: General Purpose 3-D Simulations," *IEEE Trans. Electron Dev.,* vol. 54, no. 9, pp. 2137-2142, Sept. 2007.

[181] T. Andlauer and P. Vogl, "Full-band envelope-function approach for type-II broken-gap superlattices," *Phys. Rev. B,* vol. 80, no. 3, 035304, Jul. 2009.

[182] T. Andlauer, R. Morschl and P. Vogl, "Gauge-invariant discretization in multi-band envelope function theory and g factors in nanowire dots," *Phys. Rev. B,* vol. 78, no. 7, 075317, Aug. 2008.

[183] J. M. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," *Phys. Rev.* vol. 97, no. 4, pp. 869-883, Feb. 1955.

[184] G. Bastard, "Superlattice band structure in the envelope-function approximation," *Phys. Rev. B,* vol. 24, no. 10, pp. 5693-5697, Nov. 1981.

[185] E. O. Kane, J. Phys. Chem. Solids 1, 249 (1957). See also, E. O. Kane, in Semiconductors and Semimetals, edited by R. K. Willardson and A. C. Beer Academic, New York, 1966), Vol. 1; E. O. Kane, in Handbook on Semiconductors, edited by T. S. Moss (North-Holland, Amsterdam, 1982).

[186] G. E. Pikus and G. L. Bir, Fiz. Tverd. Tela (Leningrad) 1, 1642 (1959) [Sov. Phys.—Solid State 1, 1502 (1960)

[187] G. L. Bir and G. E. Pikus, Symmetry and strain-induced effects in semiconductors, John Wiley & Sons, New York, 1974.

[188] M. G. Burt, "Fundamentals of envelope function theory for electronic states and photonic modes in nanostructures," *J. Phys. Condens. Matter.,* vol. 11, no. 9, pp. R53-R83, Mar. 1999.

[189] B. A. Foreman, "Elimination of spurious solutions from eight-band $\mathbf{k} \cdot \mathbf{p}$ theory," *Phys. Rev. B,* 56, R12748(R), Nov. 1997.

[190] F. Bastiman, "How to growth your first sample: What is a reconstruction?," *Dr Faebian Bastiman, Epitaxy Explained (Wordpress).* Apr. 2014. [Online]. Available at: https://faebianbastiman.wordpress.com/2014/04/11/how-to-growth-your-first-sample-what-is-a-reconstruction/ [Accessed: 29 Jun. 2023].

[191] F. Bastiman, "Little known MBE facts: Group V overpressure," *Dr Faebian Bastiman, Epitaxy Explained (Wordpress).* Nov. 2012. [Online]. Available at: https://faebianbastiman.wordpress.com/2012/11/21/little-known-mbe-facts-group-v-overpressure/ [Accessed: 29 Jun. 2023].

[192] A. Lorke, R. J. Luyken, J. M. Garcia and P. M. Petroff, "Growth and electronic properties of self-organized quantum rings," *Jpn. J. Appl. Phys.,* vol. 40, no. 3B, pp. 1857-1859, Mar. 2001.

[193] S. Kobayashi, C. Jiang, T. Kawazu and H. Sakaki, "Self-Assembled growth of GaSb type II quantum ring structures," *Jpn. J. Appl. Phys.,* vol. 43, no. 5B, pp. L662-L664, Apr. 2004.

[194] R. Timm *et al.,* "Self-Organized formation of GaSb/GaAs quantum rings," *Phys. Rev. Lett.,* vol. 101, no. 25, 256101, Dec. 2008.

[195] M. A. Kamarudin *et al.,* "GaSb quantum dot morphology for different growth temperatures and the dissolution effect of the GaAs capping layer," *J. Phys. D Appl. Phys.,* vol. 43, no. 6, 065402, Jan. 2010.

[196] W. S. Lee, G. W. Yoffe, D. G. Schlom and J. S. Harris Jr, "Accurate measurement of MBE substrate temperature," *J. Cryst. Growth,* vol. 111, no. 1-4, pp. 131-135, May 1991.

[197] A. Jasik *et al.,* "The influence of the growth temperature and interruption time on the crystal quality of InGaAs/GaAs QW structures grown by MBE and MOCVD methods," *J. Cryst. Growth,* vol. 310, no. 11, pp. 2785-2792, May 2008.

[198] O. Tejayadi, Y. L. Sun, J. Klem, R. Fischer, M. V. Klein and H. Morkoç, "Effects of MBE growth conditions on carbon contamination in GaAs," *Solid State Commun.,* vol. 46, no. 3, pp. 251-254, Apr. 1983.

[199] P. Hodgson, "The optical properties of self-assembled quantum dots, rings and posts," Ph.D. thesis, Dept. Phys., Lancaster Univ., Lancaster, 2014.

[200] Teledyne Princeton Instruments, "InGaAs Cameras: The Basics Educational Notes," *princetoninstruments.com*. [Online] https://www.princetoninstruments.com/learn/camera-%20fundamentals/ingaas-sensors-the-basics#:%20:text=The%20InGaAs%20two%2Ddimensional% [Accessed: 30 Jun. 2023].

[201] Fosco., "Optical Fiber Loss and Attenuation," *fiberoptics4sale.com*. [Online] https://www.fiberoptics4sale.com/blogs/archive-posts/95048006-optical-fiber-loss-and-attenuation [Last Accessed:] 09.02.2022.

[202] Oxford Instruments. LN2 or TE cooling for NIR and SWIR-optimised detectors?. https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fandor.oxinst.com%2Flearning%2Fview%2Farticle%2Fln2-versus-tecooling&data=04%7C01%7Credmanc%40live.lancs.ac.uk%7C6ca682cb99bd4c63356608d958f8beae%7C9c9bcd11977a4e9ca9a0bc734090164a%7C0%7C0%7C637638648251458686%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=1%2B7t8os1wmRxCVo4l0zMHH9hsKI3HXnvGbJ8CDZ3lzI%3D&reserved=0 [Last Accessed:] 06/08/2021. Website has since changed.

[203] Andor Technology. IDus InGaAs Detector manual, DU492A 1.7, pp.2 (2007).

[204] Oxford Instruments, Andor. Spectroscopy Resolution and Wavelength Calculator. https://andor.oxinst.com/resolution-calculator. Miramar Communications Ltd (2023). [Last Accessed: 28.06.23].

[205] Richardson Grating Laboratory. Thermo RGL. Efficiency Characteristics of Diffraction Gratings. http://astro.if.ufrgs.br/levato/instrum9/redes/chap9.htm#:~:text=P-plane%20or%20TE%20polarized,see%20Figure%209-2).&text=Both%20the%20S%20and%20P%20components%20are%20perpendicular%20to%20the%20propagation%20directions. [Last Accessed:] 28.06.2023 (2000).

[206] *0.22 NA Hard Polymer Buffer, Silica/Silica Multimode Fiber,* Thorlabs, 12239-S01, Rev E, Mar. 2011.

[207] J. P. Cleveland, S. Manne, D. Bocek and P. K. Hansma, "A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy," *Rev. Sci. Instrum.,* vol. 64, no. 2, pp. 403–405, Feb. 1993.

[208] M.-K. Yeh, N.-H. Tai and B.-Y. Chen, "Effects of anisotropic material property on the spring constant and the resonant frequency of atomic force microscope cantilever," *Rev. Sci. Instrum.,* vol. 80, no. 4, 043705, Apr. 2009.

[209] N. L. Tolman, R. Bai and H. Liu, "Hydrocarbons in the Meniscus: Effects on Conductive Atomic Force Microscopy," *Langmuir,* vol. 39, no. 12, pp. 4274-4281, Mar. 2023.

[210] G. F. Harrington and J. Santiso, "Back-to-Basics tutorial: X-ray diffraction of thin films," *J. Electroceram,* vol 47, pp. 141-163, Oct. 2021.

[211] T. B. Bahder, "Eight-band $\boldsymbol{k}{\cdot}p$ model of strained zinc-blende crystals," *Phys. Rev. B,* vol. 41, no. 17, pp. 11992-12001, Jun. 1990. Erratum *Phys. Rev. B,* vol. 46, no. 15, pp. 9913-9913, Oct. 1992.

[212] T. B. Bahder, R. L. Tober, and J. D. Bruno, "Pyroelectric effect in semiconductor heterostructures," *Superlattices Microstruct.,* Vol. 14, no. 2-3, pp. 149-152, Sept. 1993.

[213] D. L. Smith and C. Mailhiot, "Piezoelectric effects in strained-layer superlattices," *J. Appl. Phys.,* Vol. 63, no. 8, pp. 2717-2719, Apr. 1988.

[214] 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," *Appl. Phys. Lett.,* vol. 107, no. 6, 062104, Aug. 2015.

[215] S. de Gironcoli, S. Baroni, and R. Resta,"Piezoelectric Properties of III-V Semiconductors from First-Principles Linear-Response Theory," *Phys. Rev. Lett.,* vol. 62, no. 24, pp. 2853-2856, Jun. 1989.

[216] R. Songmuang, S. Kiravittaya, M. Sawadsaringkarn, S. Panyakeow and O. G. Schmidt, "Photoluminescence investigation of low-temperature capped self-assembled InAs/GaAs quantum dots," *J. Cryst. Growth,* vol. 251, no. 1-4, pp. 166-171, Apr. 2003.

[217] O. G. Schmidt *et al.,* "Self-assembled semiconductor nanostructures: climbing up the ladder of order," *Surf. Sci.,* vol. 514, no. 1-3, pp. 10-18, Aug. 2002.

[218] Y. Chen and J. Washburn, "Structural Transition in Large-Lattice-Mismatch Heteroepitaxy," *Phys. Rev. Lett.,* vol. 77, no. 19, pp. 4046-4049, Nov. 1996.

[219] S. I. Molina *et al.,* "High resolution electron microscopy of GaAs capped GaSb nanostructures," *Appl. Phys. Lett.,* vol. 94, no. 4, 043114, Jan. 2009.

[220] P. Offermans *et al.,* "Atomic-scale structure of self-assembled In(Ga)As quantum rings in GaAs," *Appl. Phys. Lett.,* vol. 87, no. 13, 131902, Sept. 2005.

[221] R. Timm *et al.,* "Quantum ring formation and antimony segregation in GaSb/GaAs nanostructures," *J. Vac. Sci. Technol. B,* vol. 26, no. 4, pp. 1492-1503, Aug. 2008.

[222] B. Dodson and J.Y. Tsao, "Relaxation of strained-layer semiconductor structures via plastic flow," *Appl. Phys. Lett.,* vol. 51, no. 17, pp. 1325-1327, Oct. 1987.

[223] T. Sasaki *et al.,* "Growth temperature dependence of strain relaxation during InGaAs/GaAs(001) heteroepitaxy," *J. Cryst. Growth,* vol. 323, no. 1, pp. 13-16, May 2011.

[224] Q. D. Zhuang *et al.,* "Effect of rapid thermal annealing on InGaAs/GaAs quantum wells," *J. Cryst. Growth,* vol. 212, no. 1-2, pp. 352-355, 2000.

[225] E. A. Zibik *et al.,* "Effects of alloy intermixing on the lateral confinement potential in InAs/GaAs self-assembled quantum dots probed by intersublevel absorption spectroscopy," *Appl. Phys. Lett.,* vol. 90, no. 16, 163107, Apr. 2007.

[226] S. Fujimoto, M. Aoki and Y. Horikoshi, "X-Ray analysis of In distribution in molecular beam epitaxy grown InGaAs/GaAs quantum well structures," *Jpn. J. Appl. Phys.,* vol. 38, no. 4A, pp. 1872-1874, Apr. 1999.

[227] B. Deveaud, J. Shah, T. C. Damen and W. T. Tsang, "Capture of electrons and holes in quantum wells," *Appl. Phys. Lett.,* vol. 52, no. 22, pp. 1886-1888, May 1988.

[228] C. H. Lin, W. W. Pai, F. Y. Chang and H. H. Lin, "Comparative study of InAs quantum dots with different InGaAs capping methods," *Appl. Phys. Lett.,* vol. 90, no. 6, 063102, Feb. 2007.

[229] P. D. Hodgson *et al.,* "Optical and structural properties of InGaSb/GaAs quantum dots grown by molecular beam epitaxy," *Semicond. Sci. Technol.,* vol. 33, no. 12, 125021, Nov. 2018.

[230] M. Mashita, Y. Hiyama, K. Arai, B.-H. Koo and T. Yao, "Indium Reevaporation during molecular beam epitaxial growth of InGaAs layers on GaAs substrates," *Jpn. J. Appl. Phys.,* vol. 39, no. 7B, pp. 4435-4437, Jul. 2000.

[231] R. R. Pelá, L. K. Teles, M. Marques and S. Martini, "Theoretical study of the indium incorporation into III-V compounds revisited: The role of indium segregation and desorption," *J. Appl. Phys.,* vol. 113, no. 3, 033515, Jan. 2013.

[232] F. Turco and J. Massies, "Strain-induced In incorporation coefficient variation in the growth of $Al_{1-x}In_xAs$ alloys by molecular beam epitaxy," *Appl. Phys. Lett.,* vol. 51, no. 24, pp. 1989-1991, Dec. 1987.

[233] K. Muraki, S. Fukatsu, Y. Shiraki and R. Ito, "Surface segregation of In atoms during molecular beam epitaxy and its influence on the energy levels in In-GaAs/GaAs quantum wells," *Appl. Phys. Lett.,* vol. 61, no. 5, pp. 557-559, Aug. 1992.

[234] R. L. S. Devine, "Photoluminescence characterisation of InGaAs/GaAs quantum well structures," *Semicond. Sci. Technol.,* vol. 3, no. 12, pp. 1171-1176, Dec. 1988.

[235] S. Weeke, M. Leyer, M. Pristovsek, F. Brunner, M. Weyers and W. Richter, "Segregation and desorption of antimony in InP (001) in MOVPE," *J. Cryst. Growth,* vol. 298, pp. 159-162, Jan. 2007.

[236] M. Sato and Y. Horikoshi, "Effect of indium replacement by gallium on the energy gaps of InAs/GaAs thin-layer structures," *J. Appl. Phys.,* vol. 69, no. 11, pp. 7697-7702, Jun. 1991.

[237] J. He, F. Bao and J. Zhang, "Capping effect of GaAsSb and InGaAsSb on the structural and optical properties of type II GaSb/GaAs quantum dots," *Appl. Phys. Lett.,* vol. 100, no. 17, 171914, Apr. 2012.

[238] B. L. Liang, Z. M. Wang, Y. I. Mazur and G. J. Salamo, "Photoluminescence of surface InAs quantum dot stacking on multilayer buried quantum dots," *Appl. Phys. Lett.,* vol. 89, no. 24, 243124, Dec. 2006.

[239] R. J. Young, E. P. Smakman, A. M. Sanchez, P. Hodgson, P. M. Koenraad and M. Hayne, "Optical observation of single-carrier charging in type-II quantum ring ensembles," *Appl. Phys. Lett.,* vol. 100, no. 8, 082104, Feb. 2012.

[240] K. Suzuki, R. A. Hogg and Y. Arakawa, "Structural and optical properties of type II GaSb/GaAs self-assembled quantum dots grown by molecular beam epitaxy," *J. Appl. Phys.,* vol. 85, no. 12, pp. 8349-8352, Jun. 1999.

[241] S. P. Bremner, K. Ghosh, L. Nataraj, S. G. Cloutier, C. B. Honsberg, "Influence of Sb/As soak times on the structural and optical properties of GaAsSb/GaAs interfaces," *Thin Solid Films,* vol. 519. no. 1, pp. 64-68, Oct. 2010.

[242] R. Timm *et al.,* "Structure and intermixing of GaSb/GaAs quantum dots," *Appl. Phys. Lett.,* vol. 85, no. 24, pp. 5890-5892, Dec. 2004.

[243] A. J. Martin *et al.,* "The disintegration of GaSb/GaAs nanostructures upon capping," *Appl. Phys. Lett.,* vol. 102, no. 11, 113103, Mar. 2013.

[244] C. E. Pryor and M. E. Pistol, "Band-edge diagrams for strained III–V semiconductor quantum wells, wires, and dots," *Phys. Rev. B,* vol. 72, no. 20, 205311, Nov. 2005.

[245] O. Madelung, Semiconductors: Data Handbook, 3rd Edition. Springer, New York (2004).

[246] Zon *et al.,* "Photoluminescence properties as a function of growth mechanism for GaSb/GaAs quantum dots grown on Ge substrates," *J. Appl. Phys.,* vol. 126, no. 8, 084301, Aug. 2019.

[247] F. Urbach, "The Long-Wavelength Edge of Photographic Sensitivity and of the Electronic Absorption of Solids," *Phys. Rev.,* vol. 92, no. 5, p. 1324, Dec. 1953.

[248] A. Esser, E. Runge, R. Zimmermann and W. Langbein, "Photoluminescence and radiative lifetime of trions in GaAs quantum wells," *Phys. Rev. B,* vol. 62, no. 12, Jun. 2000.

[249] M. Baranowski, M. Latkowska, R. Kudrawiec and J. Misiewicz, "Hopping excitons in GaInNAs - simulation of micro- and macrophotoluminescence spectra," *Acta Phys. Pol., A,* vol. 120 No. 5. Nov. 2011.

[250] B. R. Bennett, B. V. Shanabrook and R. Magno, "Phonons in self-assembled (In,Ga,Al)Sb quantum dots," *Appl. Phys. Lett.,* vol. 68, no. 7, pp. 958-960, Feb. 1996.

[251] K. Gradkowski *et al.,* "Coulomb effects in type-II Ga(As)Sb quantum dots," *Phys. Status Solidi B,* vol. 246, no. 4, pp. 752-755, Mar. 2009.

[252] K. Gradkowski *et al.,* "Coulomb-induced emission dynamics and self-consistent calculations of type-II Sb-containing quantum dot systems," *Phys. Rev. B,* vol. 85, no. 3, 035432, Jan. 2012.

[253] J. H. Davies, The Physics of Low-Dimensional Semiconductors (Cambridge University Press, 2006).

[254] N. N. Ledentsov *et al.,* "Radiative states in type-II GaSb/GaAs quantum wells," *Phys. Rev. B,* vol. 52, no. 19, 14058, Nov. 1995.

[255] C. Pryor, "Eight-band calculations of strained InAs/GaAs quantum dots compared with one, four, and six-band approximations," 1997, arXiv:cond-mat/9710304.

[256] S. Haldar *et al.,* "Effect of carrier confinement on effective mass of excitons and estimation of ultralow disorder in $Al_xGa_{1-x}As$/GaAs quantum wells by magneto-photoluminescence. *Sci. Rep.,* vol 7, no. 1, 4905, Jul. 2017.

[257] Y.-C. Chen, P. Wang, J. J. Coleman, D. P. Bour, K. K. Lee, and R. G. Waters, "Carrier recombination rates in strained-layer InGaAs-GaAs quantum wells," *IEEE J. Quantum Electron.,* vol. 27, no. 6, pp. 1451-1455, Jun. 1991.

[258] A. D. Andreev and E. P. O'Reilly, "Radiative recombination in InAs-based quantum dot lasers: dependence on carrier density and dot parameters," *Phys. Status Solidi C,* vol. 2, no. 4, pp. 1374-1378, 2005.

[259] P. Díaz-Arencibia and I. Hernández-Calderón, "Determination of the minimum island size for full exciton localization due to thickness fluctuations in $Zn_{1-x}Cd_xSe$ quantum wells." *Microelectron. J.,* vol. 33, no. 4, pp. 337-339, Apr. 2002.

[260] H. Wang *et al.,* "Influence of excitation power and temperature on photoluminescence in InGaN/GaN multiple quantum wells," *Opt. Express,* vol 20, no. 4, pp. 3932-3940, Feb. 2012.

[261] J. Orenstein and M. Kastner, "Photocurrent transient spectroscopy: measurement of the density of localized states in α-As_2Se_3," *Phys. Rev. Lett.,* vol. 46, no. 21, pp. 1421-1424, May 1981.

[262] P. G. Eliseev, "The red σ^2/kT spectral shift in partially disordered semiconductors," *J. Appl. Phys.,* vol. 93, no. 9, pp. 5404-5415, May 2003.

[263] D. M. Thomas, R. J. Hunt, N. D. Drummond, and M. Hayne, "Binding energies of excitonic complexes in type-II quantum rings from diffusion quantum Monte Carlo calculations," *Phys. Rev. B,* vol 99, no. 11, 115306, Mar. 2019.

[264] Z. Wang *et al.,* "Room-temperature InP distributed feedback laser array directly grown on silicon," *Nat. Photon.,* vol. 9, no. 12, pp. 837-842, Oct. 2015.

[265] B. Tian *et al.,* "Room temperature O-band DFB laser array directly grown on (001) silicon," *Nano Lett.,* vol. 17, no. 1, pp. 559-564, Jan. 2017.

[266] Q. Li, K. W. Ng, and K. M. Lau, "Growing antiphase-domain-free GaAs thin films out of highly ordered planar nanowire arrays on exact (001) silicon," *Appl. Phys. Lett.,* vol. 106, no. 7, 072105, Feb. 2015.

[267] R. Cipro *et al.,* "Low defect InGaAs quantum well selectively grown by metal organic chemical vapor deposition on Si(100) 300 mm wafers for next genera-

tion non planar devices," *Appl. Phys. Lett.,* vol 104, no. 26, pp. 262103, Jul. 2014.

[268] B.P. Zakharchenya, S.A. Permogorov, Excitons in Crystals, Encyclopedia of Condensed Matter Physics (2005).

Acronyms

NIR near-infra-red

MBE molecular beam epitaxy

QR quantum ring

RT room temperature

RWELL quantum ring-in-well

QW quantum well

AFM atomic force microscopy

PL photoluminescence

LT low temperature

HT high temperature

LD laser diode

IC integrated circuit

MOCVD metal-organic chemical vapour deposition

VCSEL vertical-cavity surface-emitting laser

CW continuous-wave

FWHM full width at half maximum

SLS strained-layer superlattice

TD threading dislocation

CB conduction band

LED light-emitting diode

MQW multiple quantum well

BZ Brillouin zone

DoS density of states

SO spin-orbit

VB valence band

NFE nearly-free electron

TB tight-binding

HH heavy hole

LH light hole

ZB zinc-blende

OICD optically-induced charge depletion

WL wetting layer

CTE coefficient of linear thermal expansion

APB antiphase boundary

MD misfit dislocation

ML monolayer

FM Frank-van-der-Merwe

VW Volmer-Weber

SK Stranski-Krastanov

XSTM cross-sectional scanning tunnelling microscopy

SRH Shockley-Read-Hall

BtB band-to-band

CHCC conduction-heavy conduction-conduction

CHHH conduction-heavy heavy-heavy

CHLH conduction-heavy light-heavy

CHSH conduction-heavy spin-orbit-heavy

EQE external quantum efficiency

CMOS complementary metal-oxide semiconductor

LN$_2$ liquid nitrogen

BEP beam equivalent pressure

RHEED reflection high-energy electron diffraction

CCD charge-coupled device

ADC analogue-to-digital converter

LHe-4 liquid helium-4

RTA rapid thermal annealing

XTEM cross-sectional transmission electron microscopy

BB band bending

CC capacitive charging

COM centre of mass

SEM scanning electron microscopy