UNIVERSITY OF LEEDS

Quantum Well Based Group-IV SiGeSn Semiconductor Laser and Optoelectronic Devices

Zhichao Chen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Engineering

School of Electronic and Electrical Engineering

February 2022

Intellectual Property

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Signed: 

Acknowledgements

Firstly, I have to thank Dr. Zoran Ikonic to offer me this wonderful chance to work on this exciting and challenging project about the cutting edge Group IV semiconductor devices at the University of Leeds. Zoran was also my supervisor during my study for the masters degree, it was his patient guidance and teaching that gave me courage and confidence to continue the study for my Ph.D. Also, I would like to express my sincere gratitude to Dr. Dragan Indjin and Prof. Robert W. Kelsall for their excellent supervision together with Zoran. The professional advice and suggestions from them were of great importance for the success of this project. The intense and illuminating discussions during the supervision meetings gave me many enlightening ideas that have been applied to the project. Most importantly, the knowledge and personal qualities that I have learnt from them are invaluable for my future life and career.

Secondly, I am grateful to all the colleagues in Pollard Institute for providing a pleasant research atmosphere and the help they gave me. Especially I want to thank Yang Zhou, Rowan Parker Jervis, Hang Yu and Hanshuo Wu, not only as colleagues but also as friends.

Additionally, I need to thank to all my friends in Edinburgh, Manchester, Bristol and Plymouth for their accompany online during the tough pandemic period.

Finally, I want to give my best love and thanks to my fiancee and my parents. I could not go this far without their unconditional and consistent support.

Abstract

Group IV photonics is attracting more and more attention these days in order to realise large scale optoelectronic integration. Although many prototype devices have been demonstrated, few of them can be put into practical applications. One of the major challenges is the lack of the optimized design and theoretical models. In this project, a computational model based on 8-band **k.p** theory is built to simulate the Group IV quantum well based semiconductor, and in this thesis it is applied to two different types of devices, lasers and modulators. For lasers, we focus on the tensile strained structure as it can reduce the Sn content needed to reach direct bandgap structure, moreover, because of the splitting of light and heavy holes, the low density of states at the valence band top will be beneficial to lasing. Since Ge has indirect band gap, and the L valley may still be close to Γ valley even for direct gap GeSn, effective mass method is used for electrons in the L-valleys. The optimal range of well widths for transverse magnetic (TM) gain is found to be around 13-16nm. With constant well width (14 nm), the optimal choice of both inter-band gain and net gain with varying carrier density for different photon transition energy is found, by doing calculations throughout the parameter space of Sn and strain in the well. The inter-valence band absorption with split-off band was found to be an important loss mechanism that seems to be rarely discussed in the literature. A large inter- valence band absorption was found around 0.4-0.5 eV, when the bandgap is equal to the difference between the top valence band and split off band. And because of the influence of such loss mechanism the optimal choices of net gain have changed from that of the inter-band gain. Using the optimal Sn content and strain combined with waveguide design, the threshold current was estimated to be 1.19 kA/cm^{-2}, comparable to conventional III-V quantum well lasers. For modulators, a novel way using an intra-step quantum well was applied here to improve the performance of the GeSn quantum well electroabsorption modulator. Using SiGeSn as the barrier and GeSn as the material for the well layers, an intra-step well can be made by using different Sn contents in two

intra-layers. The band structure is also calculated by the **k.p** method, and the exciton effect was considered by variational method. Without increasing the total well width, and compared to the square quantum well, a much larger quantum confined Stark effect can be realised with intra-step quantum well. By considering the figures of merit related to practical performance, $\Delta\alpha/F$ and $\Delta\alpha/F^2$, the intra-step quantum well, compared to square quantum well, brings about 44% improvement on the bandwidth per unit applied voltage and 46% reduction on the power consumption per bit data transmitted. The model presented in this thesis can still be improved. It can include other realistic effects such as carrier transport, and on the other hand, the accuracy can also be improved by considering 'better' choice of parameters and perhaps higher order **k.p** theory. Other applications are also possible based on the existing code, e.g. light emitting diodes and photodetectors.

Contents

Publications

- Zhichao Chen, Zoran Ikonic, Dragan Indjin and Robert W. Kelsall, "Design optimization of tensile-strained SiGeSn/GeSn quantum wells at room temperature", *Journal of Applied Physics* 129, 123102 (2021).
- Zhichao Chen, Zoran Ikonic, Dragan Indjin and Robert W. Kelsall, "Design considerations of intra-step SiGeSn/GeSn quantum well electroabsorption modulators", *Journal of Applied Physics* 130, 153103 (2021).

List of Figures

List of Tables

List of Abbreviations

API	Application Programming Interface
CB	Conduction Band
CMOS	Complementary Metal Oxide Semiconductor
CVD	Chemical Vapor Deposition
DFT	Density Functional Theory
DKK	Dressehaus-Kip-Kittel
DOS	Density of States
EAM	Electro-Absorption Modulator
EPIC	Electronic-photonic Integrated Circuits
EPM	Empirical Pseudopotential Method
FCA	Free Carrier Absorption
FCC	Face-centred Cubic
FDM	Finite Difference Method
FWHM	Full Width Half Maximum
HH	Heavy Hole
HPC	High Performing Computing
IBA	Inter-band Absorption
IBG	Inter-band Gain
IVBA	Inter-valence-band Absorption
LAPACK	Linear Algebra PACKage
LED	Light Emitting Diode
LH	Light Hole
MBE	Molecular Beam Epitaxy
MFDM	Matrix Finite Difference Method

MIR	Mid-infrared
MME	Momentum Matrix Element
MQW	Multi Quantum Well
PC	Personal Computer
PD	Photodetector
PL	Photoluminescence
QCSE	Quantum Confined Stark Effect
QW	Quantum Well
RAM	Random Access Memory
SO	Split Off
TE	Transverse Electric
TM	Transverse Magnetic
TMM	Transfer Matrix Method
TRM	Tunnelling Resonance Method
VB	Valence Band
VBO	Valence Band Offset

Chapter 1

Introduction

In order to realise large-scale optoelectronic integration, Group IV materials have attracted significant attention in the recent years. As a widespread material of Group IV in the nature, silicon has already been commonly used as wafer and in lightwave circuits. With the existing complementary metal-oxide-semiconductor (CMOS) technology it has many advantages such as low cost, incredibly low defect density and a good thermal conductivity. But unlike the conventional III-V group materials, which have direct band structure, silicon has a much lower Δ valley than the Γ valley conduction band, which makes it an 'indirect' material and inhibit its capability to be an efficient light emitter. Nevertheless, a way was found for lasing in Silicon by the stimulated Raman scattering effect. However, such Silicon Raman laser needs additional light sources, which include a pump beam to stimulate the atom to a higher energy level and a signal beam to amplify the Stokes transition process [1]. As the Silicon Raman laser needs additional light source, it is not suitable to be a photonic component in an optoelectronic integrated circuit in the current CMOS framework. Alternatively, the group III-V is a good choice to be the material of the active component in the photonic integrate circuits. Since the stimulated emission was obtained in GaAs in 1960s [2], research of III-V materials has progressed very well. Combined with quantum well (QW) structure, many mature and even commercialized devices with low threshold, long life time, fast modulation speed and small footprint has already been demonstrated [3–5]. The well-established III-V devices can be directly transferred to a Silicon substrate by heterogeneous integration technique such as wafer bonding and flip-chip bonding, however these methods are still facing challenges regarding the alignment accuracy and fabrication cost [3]. Another way to integrate III-V devices on the Silicon substrate is monolithic

integration using molecular beam epitaxy (MBE) or chemical vapor deposition (CVD). However this approach suffers from the physical and chemical incompatibility between the Group III-V and Group IV materials, which will result in high density of dislocations, micro thermal crack on the interface and so on [1, 3]. The performance of the laser will be degraded due to these effects.

Although it seems that the integration of III-V devices on Silicon is more practical in the near future, but because Group IV semiconductors are covalently bonded and have similar chemical and physical properties, in the long term, the 'everything-IV' seems to be an optimal solution to realise completely monolithic group IV photonics integration [6]. Germanium as a Group IV element was proposed at early stage to be a possible material for active devices. Although it is also an indirect material, but its L-valley is only about 150 meV lower than the Γ valley, which increases the chance of the radiative recombination between the Γ valley and the valence band (VB). Lasing in Ge-on-Si at room temperature has been demonstrated in 2010 [7]. A Ge waveguide was epitaxially grown on Si substrate, by a thermally induced strain, the band structure becomes more 'direct' and assisted by n-doping of phosphorous, the indirect band structure was compensated and lasing was realised. In the meantime, other photonic devices based on Ge-on-Si such as photodetector [8] and modulator [9] have been demonstrated.

A further step based on Ge could be based on the binary and ternary alloy of the Group IV material (i.e. GeSn and SiGeSn). There are a couple of benefits from this emerging material system. Firstly, GeSn is a direct-band material when the Sn composition is large enough. This idea has been repetitively proved by many researchers since it was first proposed in 1982 [10]. During the past decades, the value where the indirect to direct transition takes place was also refreshed, from the initial prediction of 20% [11] to the recent value from 6% to 10% [12]. The solubility of Sn in Ge is low (only around 0.5% [13]), but using the cutting edge epitaxial process, MBE and CVD, the GeSn alloy can be fabricated in almost full composition range, according to a recent statistics [14]. However, challenge still exist to fabricate the material with high crystalline quality. The above trend indicates that GeSn is an even more promising direct band material to build monolithic integrated lasers. Secondly, apart from the indirect to direct transform, the bandgap energy of (Si)GeSn can potentially cover a large range of wavelengths, from short wave infrared to long wave infrared. This proved the possibility for the 'everything-IV' devices to be used in multiple areas, for example telecommunications, med-

ical, environmental monitoring, entertainment applications and military. Finally, regarding the semiconductor heterostructures, SiGeSn was found to be suitable as a barrier material for lasers and other photonic devices. Although the existence of Si will make the band structure more 'indirect', but even for a laser, this is acceptable because the recombination will happen mainly in the GeSn region. Compared to using just Ge as a barrier, the appropriate choice of Si and Sn content will give a much larger barrier height for both conduction band and valence band. This makes strain to be individually engineered without too much concerns about the band alignment problems [15]. Such material brings about a larger freedom of the parameter choice when designing semiconductors heterostructures like a quantum well or multi-quantum wells. From the experience with the III-V system, a quantum well structure has been proven to be a more efficient laser material. The step like density of state makes it easier to have gain and lower threshold can be expected. Other devices, such as modulators also benefit from quantum well structure, as the quantum Stark effect enables modulation at room temperature. Moreover, compared to single quantum well, the multi-quantum well structure can provide a larger optical confinement factor, and it usually helps to provide a better performance of optoelectronic devices.

The group IV photonics are coming to its age as many devices based on Si, Ge, Sn, and especially on their binary and ternary alloys (Si)GeSn, have been demonstrated during the past years. For light sources, the first working GeSn optically pumped laser was demonstrated in 2015 [16], with a working wavelength around 2.3 μm, such laser can work up to 90 K at 1000 kW/cm^2 peak excitation. Many researches have devoted effort to push the limits of GeSn laser, aiming to make it applicable in real life. For example, in [17] the operating temperature was increased to 230K by fabricating high quality GeSn with large Sn fraction for the active region. Apart from that, it was found that the threshold can be reduced to 25 and 65 kW/cm^2 at 10 and 77 K by using multi-quantum well (MQW) structure [18]. Introducing strain is another way out, so a low threshold lasing was realized [19] with relatively low Sn content (5.4%) in the active region. Quite recently, the first electrically pumped GeSn laser was also successfully demonstrated [20]. The first electric pumped laser has a working wavelength of $\sim$2.3 μm at 100 K, a threshold of 0.6 kA/cm^2 and an peak power of 2.7 mW/facet at 10K was reached. Although the performance of such laser need to be improved in the future, it is still a great step to the light source for optoelectronic integrated circuits. Other light sources based on GeSn heterostructures, such as LED, have also been demonstrated in the last decade, with a good progress on reducing the

current density. Around 30-50 A/cm^2 is needed to realize light emission in these LEDs [21–23]. For photodetectors (PD), because a larger range of direct band gap can be achieved in GeSn, the detection range has been extended to 1.8 μm). Photodectors with improved dark current density, larger detection range and operating speed have also been developed using GeSn as the basic material [24–26]. For electro-absorption modulators (EAM), many theoretical designs based on quantum confined Stark effect (QCSE) or Franz-Keldysh effect were reported [27–29]. A recent experimental demonstration has shown that modulation with good figure of merit within mid-infrared region (1.8 - 5.0 μm) is achievable by adjusting the Sn content[30]. This provides a promising future for EAM to be ready to integrate with other on-chip optoelectronic devices also using (Si)GeSn as their material.

In order to obtain a deeper understanding of the properties of the (Si)GeSn alloy, various methods have been applied to study the band structure of the Group IV materials. The methods based on density functional theory (DFT), as a first-principle way, are used to obtain the material properties based quantum mechanics. Such methods have been frequently used to study the Group IV semiconductors and alloys since 1990s [31–35]. Apart from DFT, the empirical pseudopotential method (EPM) is another useful tool to calculate the semiconductor band structure, which only needs a few empirical inputs like the form factors. P. Moontragoon et al have used this method to calculate the unstrained SiGeSn bandgap in full the composition range [36]. S. Gupta et al use similar non-local EPM to calculate the band structure and the bandgap dependence for GeSn on the Sn content and strain. [37]. More importantly for this work, the EPM have also been used to extract the Luttinger parameters by fitting the EPM band structure to the **k.p** results [38, 39], which helps to determine the dependence of the Luttinger parameters on the material composition. Then rather than using linear Vegard's law for the Luttinger parameters, the accuracy of the **k.p** method is greatly improved. Although the above methods need only a small number of inputs, they are usually time consuming. The **k.p** method requires more input parameters but requires less computation time. Also it is easy to apply it to the calculation of the semiconductor heterostructures by integrating it with finite difference method such as transfer matrix method, or matrix diagonalisation. The investigation of the material parameters in the last few years enables **k.p** calculations with good accuracy, and **k.p** method is a good choice for the simulation and design of group IV optoelectronic devices, since it offers both a good accuracy and calculation speed. Nevertheless, a level of uncertainty still exists related to the uncertainty of the parameters dependence on alloy composition and

temperature. Currently 8-band **k.p** is commonly used for the simulation of quantum well based (Si)GeSn devices [40–43]. As a method based on perturbation theory, the 8-band k.p is only accurate in a limited range around the center of the Brillouin zone, because only the conduction band, light hole(LH), heavy hole(HH) and split off(SO) band are explicitly included, and other remote bands are all treated as perturbation. But this is enough, as the band structure around the Brillouin zone center is most important, and if used together with simple approximations such as effective mass method for other remote valleys, the model can accurately estimate the performance of devices such as lasers and modulators. Additionally, higher order **k.p** method (e.g. 15-band **k.p** or 30-band **k.p**) can describe the band structure in the full Brillouin zone [44], and can be employed for the heterostructure calculations.

From the discussion above, it can be seen that Group IV photonic devices have a prospective future. But there is still a long way to go to fully realize the everything-IV optoelectronic integrated circuits. Apart from the realistic challenges, such as growing high-quality structure with accurate control of the desired alloy composition and strain and maintaining the device integrity during the manufacturing process, more important for this work is, what is still missing in the theoretical part. In this thesis, we will focus on two particular applications, lasers and modulators, and attempt to bring some improvements to the existing research results. For lasers, although some analysis or design considerations are given for GeSn/SiGeSn QW laser, most of them are compressively strained, and these lasers need a large Sn content to maintain the direct band gap, which is difficult to achieve together with a high crystal quality because of the low solubility of Sn in Ge. Liu et al. [42] have considered on the tensile strained GeSn QW laser, but only with few parameter choices. Experimentally, strain can be induced in different ways, like epitaxial grown on a substrate with larger lattice constant which can be a GeSn layer with larger Sn content [45], or by using Silicon nitride stressor[19]. A bridge-like suspension platform has also been proposed for optoelectronics with tensile strain[46]. Recently a tensile strain as large as 2% was realised by growing Ge on a GaAs substrate[47]. The current trend of technology brings about more degrees of freedom on the choice of Sn and strain, which makes the optimization of such lasers an open topic, that we will focus on in this thesis. For GeSn QW modulator operating in mid-infrared (MIR) region, design and optimization have been done regarding the QCSE, band alignment, and waveguide coupling [27, 48]. But they are all based on the simple square QW. Here we found that the performance can be further improved by using asymmetric intra-step structure. In Chapter 2 and 3, the theoretical framework used

for modelling is introduced. Chapter 2 will focus on the electronic properties. To obtain the bandstructue, the basic effective mass method, the 8-band **k.p** method and the finite difference method which makes them applicable to the Quantum well heterostructures are introduced. The numerical issues such as result accuracy, and spurious solutions are also discussed. Chapter 3 focus on the calculation of optical properties. The simplex method is used to make the code reasonably efficient on a single-processor personal computer. In Chapter 4 the tensile strained SiGeSn/GeSn QW laser is considered, the optimization of the well width, Sn content and strain in the well region is performed. Additionally, frequently missing in other research papers, the impact from inter-valence band absorption is also considered. The optimal parameters combination with or without this effect are given. In Chapter 5, the different QCSE response between the square QW and intra-step QW was discussed based on perturbation theory. Also, the performance of the intra-step QW modulator was analysed based on various figures of merit, they are compared with the square QW and a better performance was found for our optimized intra-step QW. In the last chapter, the results of this work are summarised, together with some suggestions on possible further improvements in the future.

Chapter 2

Semiconductor Band Structure

This chapter will introduce the methods used in this work to calculate the electronic properties, the wave functions and band structure, of the quantum well structure. The simplest way to calculate electronic properties is the effective mass method. By solving the one-band Schrödinger equation, the energy levels and wave functions can be obtained by treating different bands separately. Then the E-k relation is approximated by the parabolic function which involves the effective mass. This model is not precise enough to describe the scenario when inter-band coupling are very strong, but if one band is far away from others, such as the conduction band (CB) in large bandgap semiconductors, this model is good enough. In this work, to estimate the carrier distribution in L and Δ valley, the effective mass method is used. For the subband dispersion in the Γ valley, the multi-band $\mathbf{k} \cdot \mathbf{p}$ method is used. Different from the single band model, the 8-band Hamiltonian includes the coupling of conduction band, light hole, heavy hole and split off band. The non-parabolic properties are clearly observed from the band structure. By comparing with the experimental results from the photoluminescent measurements, the 8-band $\mathbf{k} \cdot \mathbf{p}$ method shows a good accuracy. Recent research has extended the $\mathbf{k} \cdot \mathbf{p}$ method to 30-band [44]. Compared with 8-band, the 30-band Hamiltonian will generally give a more precise material parameters, and more accurate description of the energy dispersion away from the Γ point and maybe good enough when the Sn-content is large. However, for both lasers and modulators, we are mostly interested in the transitions near the center of the Brillouin zone, and 8-band $\mathbf{k} \cdot \mathbf{p}$ method is sufficient for this purpose. Apart from the theoretical part, the numerical method used to solve the Schrödinger equation and 8-band Hamiltonian is also introduced in this chapter, together with the discussion of numerical issues and the corresponding solutions

we used, such as eliminating the spurious solution.

2.1 Single-band Effective Mass Model

2.1.1 Bound States of Finite Square Quantum Well

The time dependent Schrödinger equation of a single particle is:

$$\left[-\frac{\hbar^2}{2m_0}\boldsymbol{\nabla}^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) \tag{2.1}$$

where m_0 is the free electron mass. $V(\mathbf{r},t)$ is the potential. The potential is independent on time, and the wave function can be written in the form with separation of variables, time t and coordinate $\mathbf{r}$:

$$\Psi(\mathbf{r},t) = \Psi(\mathbf{r})e^{-iEt/\hbar} \tag{2.2}$$

By substituting Eq. 2.2 into Eq. 2.1, we obtain the time independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m_0}\boldsymbol{\nabla}^2 + V(\mathbf{r})\right]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{2.3}$$

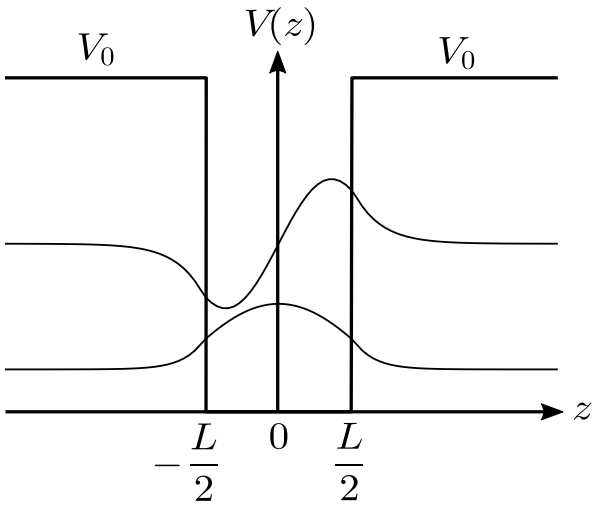

Figure 2.1: A square quantum well with a width of L. The barrier has a potential of V_0. The first and second bound states corresponds to the first even and odd solution.

To simplicity, we consider a quantum well shown in Fig. 2.1 and assume the structure is pseu-

domorphically grown along the z-direction [001]. The potential in the barrier is V_0 and the potential inside the well is zero. Further, because the effective masses m^* are usually different within each layer, Eq. 2.3 becomes:

$$\left[-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z}+V(z)\right]\psi_n(z)=E_n\psi_n(z) \tag{2.4}$$

Following the standard boundary condition that the wave function should decay to zero when it approaches $\pm\infty$, we use exponentials to describe the wave function outside the well and trigonometric function inside the well. Then for the even solution, the solution has the form:

$$\psi(z)=\begin{cases} B\exp(\kappa z) & z\leq -L/2 \\ A\cos(kz) & \mid z\mid < L/2 \\ B\exp(-\kappa z) & z\geq L/2 \end{cases} \tag{2.5}$$

And for odd solution:

$$\psi(z)=\begin{cases} -B\exp(\kappa z) & z\leq -L/2 \\ A\sin(kz) & \mid z\mid < L/2 \\ B\exp(-\kappa z) & z\geq L/2 \end{cases} \tag{2.6}$$

where

$$k=\sqrt{2m_w^*E}/\hbar \tag{2.7a}$$

$$\kappa=\sqrt{2m_b^*(V_0-E)}/\hbar \tag{2.7b}$$

Using the boundary condition at the interface, where the wave function ψ and its derivative divided by the effective mass $(1/m^*)\psi'$ are continuous, the quantization condition are obtained:

$$\begin{aligned} \kappa &= \frac{m_b^*k}{m_w^*}\tan\left(\frac{kL}{2}\right) && \text{(even)} \\ \kappa &= -\frac{m_b^*k}{m_w^*}\cot\left(\frac{kL}{2}\right) && \text{(odd)} \end{aligned} \tag{2.8}$$

The above equation can't be solved analytically. There are two ways to get the solution. Since k and κ both depend on energy, we can move the non-zero part to one side and rewrite Eq. 2.8

as a function of energy:

$$f_{\text{even}}(E) = \frac{m_b^*\sqrt{2m_w^*E}}{m_w^*\hbar}\tan\left(\frac{\sqrt{2m_w^*E}L}{2\hbar}\right) - \frac{\sqrt{2m_b^*(V_0-E)}}{\hbar} = 0 \tag{2.9a}$$

$$f_{\text{odd}}(E) = -\frac{m_b^*\sqrt{2m_w^*E}}{m_w^*\hbar}\cot\left(\frac{\sqrt{2m_w^*E}L}{2\hbar}\right) - \frac{\sqrt{2m_b^*(V_0-E)}}{\hbar} = 0 \tag{2.9b}$$

As an example, for a 15nm SiGe/Ge/SiGe QW, with Si content in the barrier of 12%, gives the conduction band potential of the barrier $V_0 = 0.247$ eV. The effective mass of the well and barrier layers are $m_w^* = 0.042m_0$ and $m_b^* = 0.066m_0$. The shape of Eq. 2.9a and Eq. 2.9b are shown in Fig. 2.2. The solution of such equations can be found by numerical methods such as bisection method and Newton-Raphson method. We should note that both Eq. 2.9a and Eq. 2.9b have infinite discontinuity because of the existence of tan and cot, which may lead to failure when using bisection method to incrementally search for the solution from 0 to V_0. So one always need to manually check if the solution is at the discontinuity points.

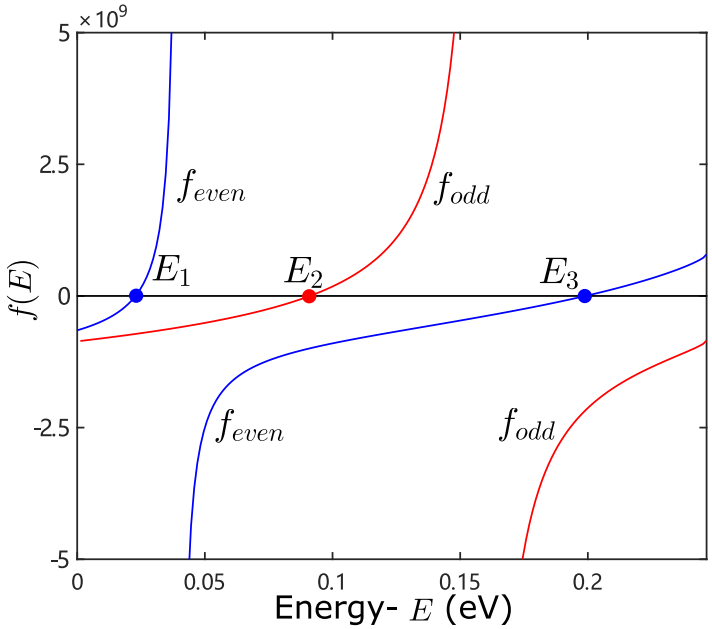

Figure 2.2: The plot of Eq. 2.9a (blue) and Eq. 2.9b (red) for a 15nm SiGe/Ge quantum well. The zero points correspond to the quantized energy levels.

Another similar method to find the solution of the Schrödinger equation is the graphical approach. According to Eq. 2.7 and Eq. 2.8, the quantization condition can be plotted on the $\kappa' L_w/2$ versus $kL_w/2$ plane:

$$\left(k\frac{L_w}{2}\right)^2 + \left(\kappa'\frac{L_w}{2}\right)^2 = \frac{2m_wV_0}{\hbar^2}\left(\frac{L_w}{2}\right)^2 \tag{2.10}$$

and

$$\kappa' \frac{L_w}{2} = \sqrt{\frac{m_b^*}{m_w^*}} k \frac{L_w}{2} \tan\left(k \frac{L_w}{2}\right) \quad \text{(even)} \tag{2.11a}$$

$$\kappa' \frac{L_w}{2} = -\sqrt{\frac{m_b^*}{m_w^*}} k \frac{L_w}{2} \cot\left(k \frac{L_w}{2}\right) \quad \text{(odd)} \tag{2.11b}$$

where $\kappa' = \sqrt{m_w/m_b}\kappa$. The solution can be obtained from the intersection point of the circle with a radius of $\sqrt{2m_w^* V_0}(L/2\hbar)$ and the function Eq. 2.11a in the first quadrant. Fig. 2.3 shows plots of Eq. 2.10 and Eq. 2.11a for the same SiGe/Ge/SiGe quantum well, discussed above. The eigenvalue of k and κ of the first to the third states can be read directly from the plot. It is more clear in this figure that no matter how small the barrier potential and the well width are, there is at least an even solution.

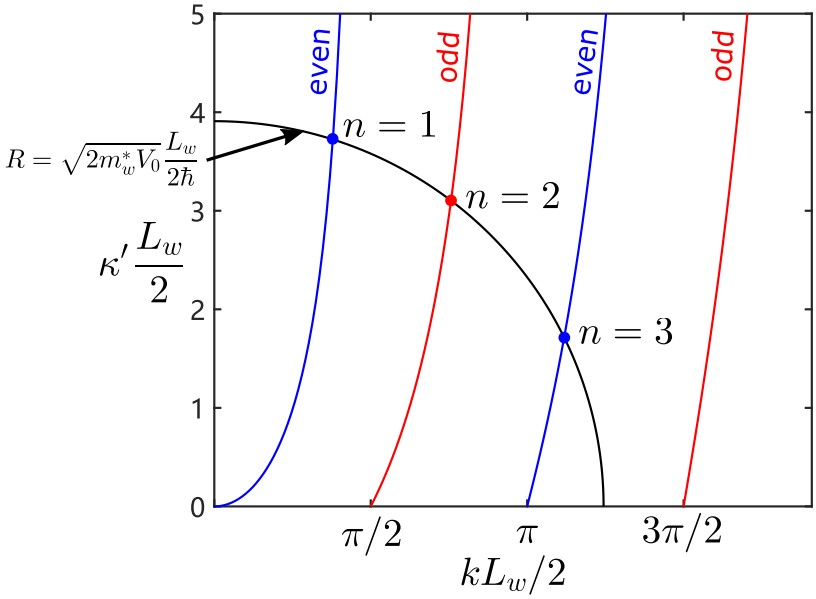

Figure 2.3: The graphical method to find the solution of k. In a plane of $kL_w/2$ and $\kappa' L_w/2$ the eigenvalue of k of each state is the intersection of the curve Eq.2.11a and a circle with radius $\sqrt{2m_w^* V_0}(L/2\hbar)$.

With either approach, we can obtain the eigenvalues of k or energy. Then the wave function should be normalized according to the normalization condition:

$$\int_{-\infty}^{\infty} |\psi_n(z)|^2 dz = 1 \tag{2.12}$$

by using the boundary condition again to eliminate the coefficient B, and substitute Eq. 2.8

into the integration:

$$\begin{aligned}
&2\int_{-\infty}^{-L_w/2} B^2\exp^2(\kappa z)dz + \int_{-L_w/2}^{L_w/2} A^2\cos^2(kz)dz \\
=&B^2\frac{1}{2\kappa}\exp(-\kappa L_w) + \frac{A^2}{2}\left(L_w + \frac{1}{k}\sin(kL_w)\right) \\
=&\frac{A^2}{2}\left[L_w + \frac{1}{\kappa}\left(\frac{\kappa}{k}\sin(kL_w) + \cos(kL_w) + 1\right)\right] \\
=&\frac{A^2}{2}\left[L_w + \frac{2}{\kappa}\left(\frac{m_b^*}{m_w^*}\sin^2\frac{kL_w}{2} + \cos^2\frac{kL_w}{2}\right)\right] = 1
\end{aligned} \tag{2.13}$$

Hence, for even solutions:

$$A = \sqrt{\frac{2}{L_w + \frac{2}{\kappa}\left(\frac{m_b^*}{m_w^*}\sin^2\frac{kL_w}{2} + \cos^2\frac{kL_w}{2}\right)}} \tag{2.14}$$

similarly, for odd solutions:

$$A = \sqrt{\frac{2}{L_w + \frac{2}{\kappa}\left(\frac{m_b^*}{m_w^*}\cos^2\frac{kL_w}{2} + \sin^2\frac{kL_w}{2}\right)}} \tag{2.15}$$

To this point, the solution of a simple single square QW is given. However, the above method is not suitable to solve more complicated cases, such as with the applied electric field, or self-consistent solution for a doped QW, and multiple QWs. Hence, in the following sections, two methods are discussed for solving the single-band model with arbitrary potential profile.

2.1.2 Transfer Matrix Method

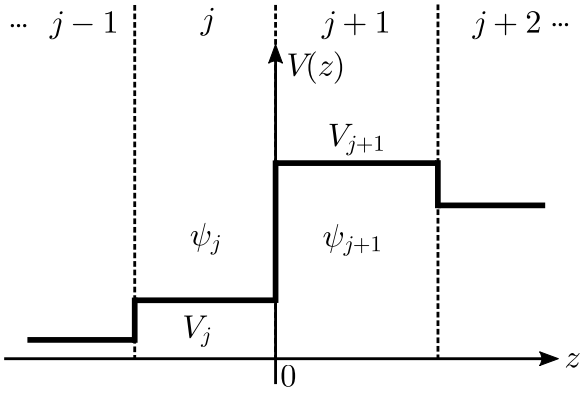

Figure 2.4: An example of potential profiles of different layers. The interface of the j and $j+1$ layers is located at $z = 0$.

The transfer matrix method (TMM) was used in an early model to simulate the arbitrary shape QW. However, it was not used in the final program here, due to its complexity when applied to 8-band k.p method and it has many numerical instabilities such as missing solution when using incremental search method. This leads to its incapability of dealing with the degenerate bands which may originate from spin or in MQW system. Here we still show how this method is implemented in one-band effective mass model, as this was used in early stage, and also used to compare with the later model. Also, we give some explanations of potential difficulties when using this method.

To derive the discontinuity matrix, we consider an interface located at $z = 0$ (see Fig. 2.4). Hence, considering the Schrödinger equation Eq. 2.4, a general description of wave function by Euler's formulas can be written in matrix form:

$$\psi_l(z) = \begin{bmatrix} \exp(ik_j z) & \exp(-ik_j z) \end{bmatrix} \begin{bmatrix} A_j \\ B_j \end{bmatrix} \tag{2.16}$$

in which:

$$k_j = \sqrt{2m_j^*(E - V_j)}/\hbar \tag{2.17}$$

where the subscript j is the index of a certain layer. Then by applying the boundary condition that ψ and $(1/m^*)\psi'$ are continuous at the interface, we find:

$$A_j + B_j = A_{j+1} + B_{j+1} \tag{2.18a}$$

$$\frac{ik_j}{m_j^*}(A_j - B_j) = \frac{ik_{j+1}}{m_{j+1}^*}(A_{j+1} - B_{j+1}) \tag{2.18b}$$

The coefficient on the two sides can be related by a discontinuity matrix:

$$\begin{bmatrix} A_{j+1} \\ B_{j+1} \end{bmatrix} = D_{j+1,j} \begin{bmatrix} A_j \\ B_j \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + \dfrac{m_{j+1}^* k_j}{m_j^* k_{j+1}} & 1 - \dfrac{m_{j+1}^* k_j}{m_j^* k_{j+1}} \\ 1 - \dfrac{m_{j+1}^* k_j}{m_j^* k_{j+1}} & 1 + \dfrac{m_{j+1}^* k_j}{m_j^* k_{j+1}} \end{bmatrix} \begin{bmatrix} A_j \\ B_j \end{bmatrix} \tag{2.19}$$

Note that, different from Eq. 2.7, where both k and κ are real numbers, the k_j in Eq. 2.17 could be complex numbers when $E < V_j$, which leads to complex coefficients in different layers. Apart from the discontinuity matrix, the propagation matrix should also be considered, because for

discritized arbitrary potential profile, the interfaces are located at different points. To use the discontinuity matrix the interface should be located at the original point. Then the original coordinate system should be transformed to a new system whose origin is located at the new interface. For example, in Fig. 2.5, the origin of the first and second coordinate system is at the left and right interface of the j-th layer. Assume the length of the j-th layer is h_j, then $z = z' + h_j$, the wave function of the second coordinate system is denoted by $\phi(z')$ with coefficient A' and B'. By equating the wave function in two different coordinate system, we have:

$$\begin{aligned} \phi(z') &= \psi(z) \\ A'e^{ikz'} + B'e^{-ikz'} &= Ae^{ikz} + Be^{-ikz} \\ A'e^{ikz'} + B'e^{-ikz'} &= Ae^{ik(z'+h_j)} + Be^{-ik(z'+h_j)} \end{aligned} \tag{2.20}$$

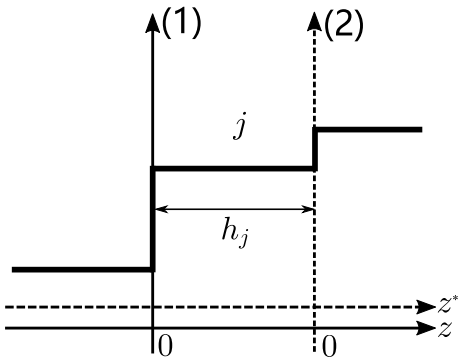

Figure 2.5: Two coordinate systems which have origin located at the left and right interface of the j-th layer. The length of the layer is h_j

It is easily found that $A' = Ae^{ikh_j}$ and $B' = Be^{-ikh_j}$. So the wave function coefficient in the first coordinate system is transformed into the coefficient in the second coordinate system by the propagation matrix of the j-th layer:

$$\begin{bmatrix} A' \\ B' \end{bmatrix} = \begin{bmatrix} e^{ikh_j} & 0 \\ 0 & e^{-ikh_j} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = P_j \begin{bmatrix} A \\ B \end{bmatrix} \tag{2.21}$$

To find the eigenvalues for an arbitrary potential, we can first discretise the potential profile into N step-like layers with constant potentials and effective masses. For convenience the width of each layer are set to have the same value h. The 0-th and $(N+1)$-th layers could be the

leftmost and rightmost barriers. Then the coefficients in an arbitrary layer j could be found from the coefficients in the zero-th layer:

$$\begin{bmatrix} A_j \\ B_j \end{bmatrix} = D_{j,j-1}P_{j-1}D_{j-1,j-2}\cdots D_{2,1}P_1D_{1,0}\begin{bmatrix} A_0 \\ B_0 \end{bmatrix} = T\begin{bmatrix} A_0 \\ B_0 \end{bmatrix} \tag{2.22}$$

Note that the coefficients A_j and B_j are not based on the original coordinate system which is located at the interface between zero-th and first layers (denoted as zero-th coordinate system). Instead, they are the coefficients in the $(j-1)$-th coordinate system, whose origin is at the interface of the $(j-1)$-th and j-th layers. Thus the wave function in the original coordinate system is:

$$\psi_j(z) = A_j e^{ik_j(z-z_{j-1})} + B_j e^{-ik_j(z-z_{j-1})} \tag{2.23}$$

where $z_{j-1} = h(j-1)$. Since the wave function in region 0 and $N+1$ must decay to zero when approaching to $-\infty$ and ∞, the coefficients A_1 and B_{N+1} are zero, so we obtain:

$$\begin{bmatrix} A_{N+1} \\ 0 \end{bmatrix} = \begin{bmatrix} T_{11}(E) & T_{12}(E) \\ T_{21}(E) & T_{22}(E) \end{bmatrix}\begin{bmatrix} 0 \\ B_0 \end{bmatrix} \tag{2.24}$$

Thus, to find the bound states eigenvalues, one needs to find the solutions of $T_{22} = 0$. Note that T_{22} is a complex number, but we can still use its necessary condition $\mathrm{Re}(\mathrm{T}_{22}) = 0$. Then check the imaginary part to avoid the spurious solution. For bound states, we use bisection method and incrementally search for solutions within the interval from the QW bottom to the lower barrier. Additionally, the zeros can be visualized by plotting the $\ln|T_{22}(E)|$. The solutions can be seen from the dips, i.e. minima of this function. Finally, again, the wave function should be normalized according to Eq. 2.12.

Fig. 2.6 shows an example of implementing TMM. The potential profile and effective mass has a parabolic dependence on the coordinate z. The potential at the bottom of the QW ($z = 5nm$) is zero and for the left and right barriers, the potentials V_b are both 1.05 eV. Similarly, the effective mass at $z = 5nm$ is $0.067m_0$ (the effective mass of GaAs). For the barrier, it is $0.15m_0$ (AlAs). The whole structure is divided into 22 layers. The potential and effective mass in each layer is using the average value at left and right interface. To make it clear, the wave functions shown in Fig. 2.6a are offset by their energies. Also the four spikes in Fig. 2.6b indicate the

energy of the four states. Note that the dips, i.e. the minimum spikes only appear within the interval $(0, V_b)$. Above V_b, $\ln|T_{22}(E)|$ only slightly oscillates around 0. Since Eq. 2.24 holds only for bound states, the wave function is a plane wave when the energy is higher than the barrier. Hence TMM is not able to give the unbound states solutions.

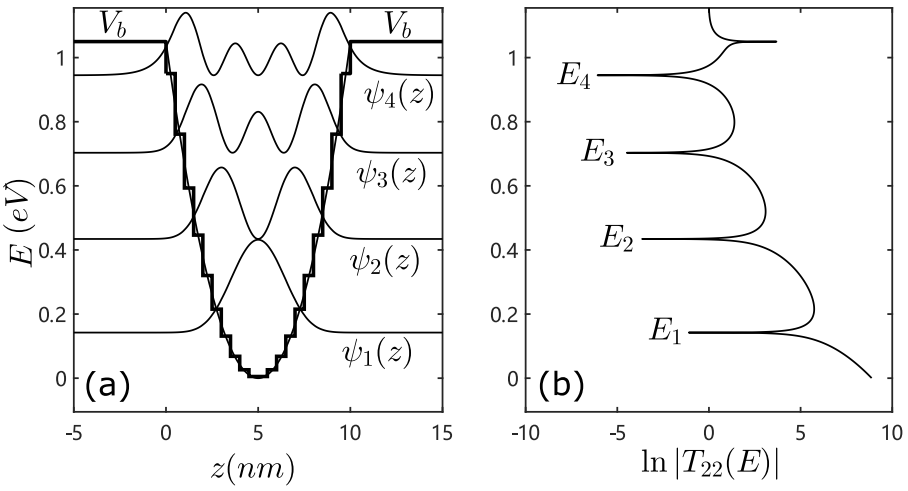

Figure 2.6: Solutions in a parabolic quantum well using transfer matrix method. The barrier height V_b is 1.05 eV and the effective mass also follows a parabolic relation, being $0.067m_0$ at the bottom and $0.15m_0$ at the barrier boundary. (a)The shape of the original and discretised potential (22 layers), the wave functions are offset by the energy levels; (b)The plot of $\ln|T_{22}(E)|$, solutions are indicated by the minimum spikes of the curve.

Although TMM can also be implemented in the **k.p** method by a relatively complex procedure, as shown in Chapter 14 of [49], it was not finally applied to this work for the following reasons. (1)In this work, it turns out that some good structure will not have a very deep QW. To accurately describe the carrier distribution, it maybe important to also include some free states, but the boundary condition applied to TMM does not find these solutions. (2)To find the eigenvalue, the problem transforms to a zero-finding problem. But from the aspect of programming, using incremental method to automatically find the solution will always have potential hazard for degenerate bands, or for two bands that have energies very close to each other, regardless of the spacing. (3)Another case that may cause missing solutions is when one of the energy levels E_n is very close to one of the discrete potential values V_j. The k_j will then approach to zero. Since k_j appears in the denominator, the matrix element could then be too large and overflow, which breaks the transfer matrix and results in the missing solutions. These scenarios happen randomly, and they are hard to avoid by numerical or programming technique. Hence TMM is not a good choice for the purpose of this project.

2.1.3 Matrix Finite Difference Method

A more comprehensive way to solve the one-band Schrödinger equation is the matrix finite difference method (MFDM). Instead of dividing the the potential profile into several layers, the MFDM treats the wave function as a discrete variable. This method may be slower than TMM in some simple cases, for example for a simple single square quantum well. The TMM only needs three layers to establish the transfer matrix, but MFDM needs more discrete points for the wave function and then finds the eigenvalues. However, the MFDM is a safer method to use because it does not have problems with missing solution. For degenerate states, it will give same number of solutions according to the degree of degeneracy. Moreover, the MFDM will include the free states which is higher than the boundary. Although such states are not truly continuous, they are still quantized states, but it can help to better describe the carrier distribution than TMM.

To implement the MFDM, we first divide the whole region (include the width of barrier and well) into N layers with width of Δz. For conduction band, consider the Schrödinger equation in Eq. 2.4, and expand the derivative at z_j:

$$-\frac{\hbar^2}{2}\frac{\frac{1}{m^*_{j+1/2}}\psi'_{j+1/2}-\frac{1}{m^*_{j-1/2}}\psi'_{j-1/2}}{\Delta z}+V_j\psi_j=E\psi_j \tag{2.25}$$

where the subscript $j+1/2$ is the middle point between z_j and z_{j+1}. Then we have:

$$\frac{1}{m^*_{j+1/2}}=\frac{1}{2}\left(\frac{1}{m^*_j}+\frac{1}{m^*_{j+1}}\right) \tag{2.26}$$

$$\psi'_{j+1/2}=\frac{\psi_{j+1}-\psi_j}{\Delta z} \tag{2.27}$$

and similarly for $j-1/2$, and by substituting Eq. 2.26 and Eq. 2.27 into Eq. 2.25 we obtain:

$$-\frac{\hbar^2}{4(\Delta z)^2}\left[\left(\frac{1}{m^*_{j+1}}+\frac{1}{m^*_j}\right)\psi_{j+1}-\left(\frac{1}{m^*_{j+1}}+\frac{2}{m^*_j}+\frac{1}{m^*_{j-1}}\right)\psi_j+\left(\frac{1}{m^*_j}+\frac{1}{m^*_{j-1}}\right)\psi_{j-1}\right]$$
$$+V_j\psi_j=E\psi_j \tag{2.28}$$

By applying the above equation at each discrete point:

$$b_1\psi_1+a_1\psi_2=E\psi_1$$

$$
\begin{aligned}
a_1\psi_1 + b_2\psi_2 + a_2\psi_3 &= E\psi_2 \\
&\vdots \\
a_{j-1}\psi_{j-1} + b_j\psi_j + b_j\psi_{j+1} &= E\psi_j \\
&\vdots \\
a_{N-1}\psi_{N-1} + b_N\psi_N &= E\psi_N
\end{aligned} \tag{2.29}
$$

The set of equations can be written in a matrix form:

$$
\begin{bmatrix}
b_1 & a_1 & 0 & \cdots & \cdots & 0 \\
a_1 & b_2 & a_2 & & & \vdots \\
0 & a_2 & \ddots & \ddots & & \vdots \\
\vdots & & \ddots & \ddots & a_{N-2} & 0 \\
\vdots & & & a_{N-2} & b_{N-1} & a_{N-1} \\
0 & \cdots & \cdots & 0 & a_{N-1} & b_N
\end{bmatrix}
\begin{bmatrix}
\psi_1 \\ \psi_2 \\ \vdots \\ \vdots \\ \psi_{N-1} \\ \psi_N
\end{bmatrix}
= E
\begin{bmatrix}
\psi_1 \\ \psi_2 \\ \vdots \\ \vdots \\ \psi_{N-1} \\ \psi_N
\end{bmatrix} \tag{2.30}
$$

in which:

$$
a_j = -\frac{\hbar^2}{4(\Delta z)^2}\left(\frac{1}{m^*_{j+1}} + \frac{1}{m^*_j}\right) \tag{2.31a}
$$

$$
b_j = \frac{\hbar^2}{4(\Delta z)^2}\left(\frac{1}{m^*_{j+1}} + \frac{2}{m^*_j} + \frac{1}{m^*_{j-1}}\right) + V_j \tag{2.31b}
$$

Note that, different from the TMM which takes the most outer barriers as one layer each, the matrix method divides the barriers into several layers. Hence the barrier width should be reasonably wide to let the wave function gradually decay to zero. Also, for $j = 1$ and $j = N$, it also involves $\psi(0)$ and $\psi(N+1)$, but they do not appear in Eq. 2.30. This is because by applying the matrix method, by default they are $\psi_0 = \psi_{N+1} = 0$, and the model is then equivalent to having the whole structure embedded in an infinite quantum well. If the barriers are wide enough (i.e the total structure length is also wide), the approximation of the standard boundary condition is accurate enough to calculate the bound states. For the free states, they are actually the 'bound states' within the infinite quantum well. These discrete free states, to some extent, can estimate the carrier distribution over the free states if the bound states are already well occupied by carriers.

The elements of Hamiltonian of Eq. 2.30 are all real numbers. So the eigenvalue problem can be

solved by LAPACK 'dstevd' subroutine [50]. The subroutine will gives N eigenvalues, starting from the first bound state. The eigenvalues will cover all the bound states if the number of grid points is larger than the number of bound states. Since there are usually only a few bound states confined in the QW, which is much smaller than the number of guid points, most of the eigenvalues are actually free states. Fig. 2.7 shows an example calculation result using MFDM. The QW is for the L-valley of the 14 nm $Si_{0.15}Ge_{0.8}Sn_{0.05}/Ge_{0.95}Sn_{0.05}$ QW. The number of grid points is 81 in this case. The first three states are bound states of the 14 nm QW, below the barrier, and the other states are considered as free states. We can see that these states are still quantized. As the barrier width chosen here is 20 nm, the 'free states' are the exact solutions for a 54 nm wide infinite QW with an 14 nm intra-QW located in the middle of it. Nevertheless, by including these 'quasi-free-states' we can still reasonably estimate the carrier distribution of the realistic structures.

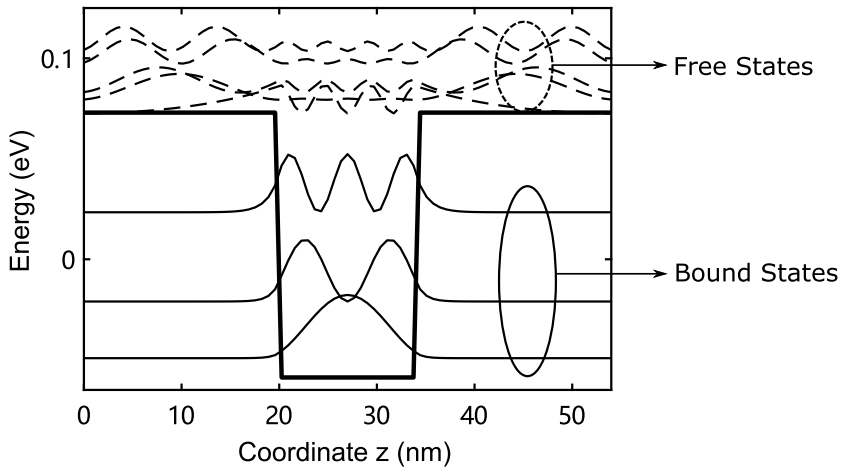

Figure 2.7: For a 14 nm $Si_{0.15}Ge_{0.8}Sn_{0.05}/Ge_{0.95}Sn_{0.05}$ QW (L-valley), bound states and free states calculated by matrix finite difference method. Barrier widths are set to be 20 nm on both sides.

To evaluate the accuracy, and the extent to which these calculated free states can represent the real situation, we consider a square infinite QW. The energy level are $E_n = (\hbar^2/2m^*)(n\pi/L_t)^2$, so the spacing of these 'quasi-free states' only depend on the length the whole structure (L_t) if the effective mass is constant. Then for the infinite QW, the density of states (DOS) is:

$$\rho_{2D}(E) = \frac{m^*}{\pi\hbar^2 L_t} \sum_i^N \Theta(E - E_i) \tag{2.32}$$

where $\Theta()$ is the step function and L_t is the total length of the QW and the two barriers. In

reality, the free states are continuous, and the DOS of continuous states can be expressed by the three-dimensional density of states:

$$\rho_{3D}(E) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2}\sqrt{E} \tag{2.33}$$

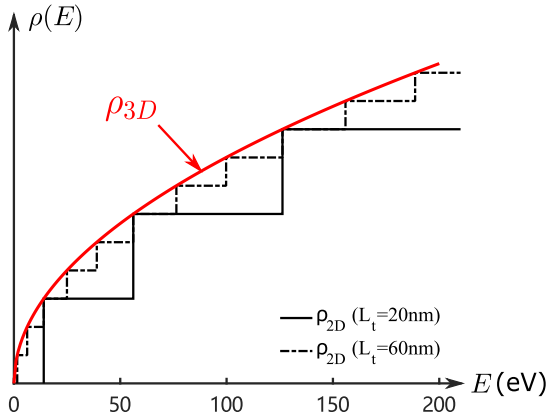

Figure 2.8: Conduction band density of states of bulk GaAs (solid red line), 20 nm infinite GaAs quantum well (solid black line) and 60 nm infinite quantum well (dashed black line)

Fig. 2.8 shows the comparison of 2D DOS of a 20 nm infinite GaAs QW and 60nm infinite GaAs QW. As can be seen in Fig. 2.8, the edges of the step function are located at the curve of the 3D DOS. And by comparing the 20nm and 60nm QW, the wider one has more steps that fill the blank space between 3D DOS and 2D DOS of the narrower QW. So we can conclude that the two-dimensional DOS in between two quantized states will be smaller than that of continuous states. However, the wider the infinite QW is, the smaller difference between 2D and 3D DOS will be. For our simulation, by using the MFDM, although the whole structure is not a perfect square infinite QW, but it also follows such pattern if we choose the barrier width to be reasonably wide (which leads to a large total structure width), the results from this approximation should be acceptable.

2.1.4 Effective Mass of Conduction Band

The $E-k$ energy dispersion is estimated by parabolic effective mass approximation:

$$E_{en}(k_t) = E_{en}(0) + \frac{\hbar^2 {k_x}^2}{2m_x} + \frac{\hbar^2 {k_y}^2}{2m_y} \tag{2.34a}$$

$$E_{en}(k_t) = E_{en}(0) + \frac{\hbar^2 {k_t}^2}{2m_t} \tag{2.34b}$$

where the in-plane wave vector and effective mass are $k_t = \sqrt{{k_x}^2 + {k_y}^2}$, $m_t = \sqrt{m_x m_y}$.

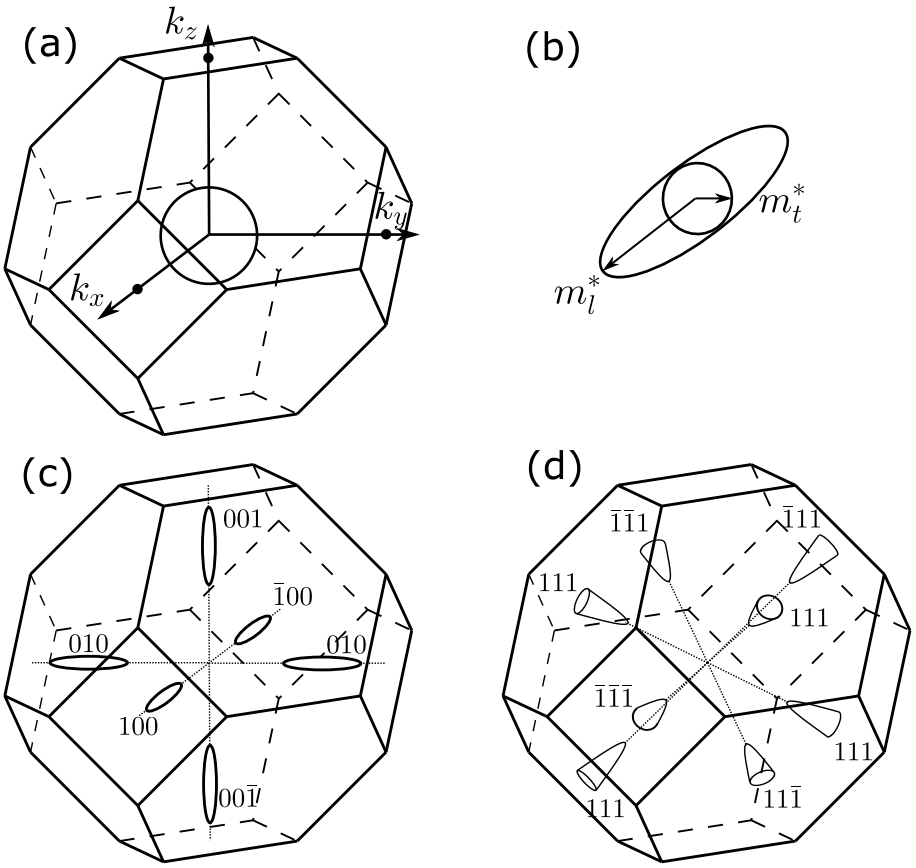

Figure 2.9: The fist Brillouin zone of Face-centred Cubic (FCC) lattice, and energy contour line of different valleys. (a)Γ valley; (b)The constant energy ellipsoid and axes of effective masses; (c)Δ valley; (d)L valley

Fig. 2.9 shows the first Brillouin zone of face centred cubic lattice and the equi-energy surface ('contour') of Γ, L and Δ valleys. For conduction band in Γ-valley, the energy contour is a sphere, as is shown in Fig. 2.9a. Hence the effective masses along x, y and z axis are the same, which means that the same effective mass is used to solve the quantization problem in Eq. 2.4 and the $E-k$ relation in Eq. 2.34b. However for other valleys, the contour is an ellipsoid (Fig. 2.9b). The energy dispersion of the ellipsoid is described by the longitudinal effective mass m_l (along principle axis) and the transverse effective mass m_t (perpendicular to the principle axis). The principle axis of the Δ valley coincides with the x, y and z axis, then the effective mass m_x, m_y and m_z are equivalent to the longitudinal or transverse effective masses. However, for the L valley, the principle axis does not coincide with any x, y, z axes. Then the effective masses for the L valley are obtained by unitary transformation [51]. The effective mass in the z direction used in Eq. 2.4 is also known as the quantization effective mass m_q and the effective

mass in the x-y plane is equivalent to the density of states effective mass, and is calculated from the expression $m_d = \sqrt{m_x m_y}$. The density of states effective mass is then used to calculate the two-dimensional carrier density and quasi-Fermi level. The effective masses components and degeneracy of different valleys shown in Fig. 2.9 for the structure grown along [001] direction are listed in Table 2.1. Note that the eight L-valleys plotted in Fig. 2.9 have only a half of each one inside the first Brillouin zone, which is equivalent to four full ellipsoids, hence the degeneracy of L-valley is four.

Table 2.1: Effective mass along x, y and z axis, quantization and longitudinal effective mass and subband degeneracy [51]

Valley	Orientation	m_x	m_y	m_z	m_q	m_d	Degeneracy
Δ_{I}	[001]	m_t	m_t	m_l	m_l	m_t	2
Δ_{II}	[100]	m_l	m_t	m_t	m_t	$\sqrt{m_t m_l}$	2
	[010]	m_t	m_l	m_t	m_t	$\sqrt{m_t m_l}$	2
L	[111]	$m_t \frac{2m_l+m_t}{m_l+2m_t}$	$\frac{m_l+2m_t}{3}$	$\frac{3m_l m_t}{2m_l+m_t}$	$\frac{3m_l m_t}{2m_l+m_t}$	$\sqrt{\frac{m_t(2m_l+m_t)}{3}}$	4

2.1.5 Effective Mass of Valence Band

For valence band, the quantized states at $\mathbf{k}_{||} = 0$ are found from the Schrödinger equation for holes:

$$\left[+\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m_h^z}\frac{\partial}{\partial z} + V_h(z)\right]\varphi(z) = E\varphi(z) \tag{2.35}$$

where the effective mass m_h^z is defined by the Luttinger parameters:

$$m_{hh}^z = \frac{m_0}{\gamma_1 - 2\gamma_2} \tag{2.36a}$$

$$m_{lh}^z = \frac{m_0}{\gamma_1 + 2\gamma_2} \tag{2.36b}$$

Note that for holes the sign of the second derivative term is positive, which means that by using matrix method the matrix element in Eq. 2.30 becomes:

$$a_j^h = \frac{\hbar^2}{4(\Delta z)^2}\left(\frac{1}{m_{j+1}^{h,z}} + \frac{1}{m_j^{h,z}}\right) \tag{2.37}$$

$$b_j^h = -\frac{\hbar^2}{4(\Delta z)^2}\left(\frac{1}{m_{j+1}^{h,z}} + \frac{2}{m_j^{h,z}} + \frac{1}{m_{j-1}^{h,z}}\right) + V_j \tag{2.38}$$

The potential in Eq. 2.35 is the band edge energy of heavy hole, light hole and split off bands. The heavy and light hole bands are degenerated at $\mathbf{k} = 0$ without strain. Since the single-band model treats each band separately, the parabolic estimation can be used if the coupling between these bands is neglected:

$$E_{hm}(k_t) = E_{hm}(0) - \frac{\hbar^2 {k_t}^2}{2m_h^t} \tag{2.39}$$

the in-plane effective mass can also be estimated from the Luttinger parameters [52]

$$m_{hh}^t = \frac{m_0}{\gamma_1 + \gamma_2} \tag{2.40a}$$

$$m_{lh}^t = \frac{m_0}{\gamma_1 - \gamma_2} \tag{2.40b}$$

The conduction band is far from other bands and is not degenerate with other bands. Hence it may be good enough to use the parabolic approximation in many cases. However, valence bands are strongly coupled with each other and non-parabolicity appears. Then it may not very accurate to use such approximation. There are only some extreme situations when using such approximation is acceptable. For example, when large strain exists the heavy holes and light holes are split significantly, which weakens the coupling. Then, in an acceptable range of the in-plane k vectors the effective mass approximation is accurate. But when the k vector becomes larger, the energy of one may get close to another band and then the parabolic energy dispersion will lose its accuracy. Fig. 2.10 shows the comparison of the subband energy dispersion calculated by the parabolic approximation and 8-band **k.p** method. The structure is a $Si_{0.15}Ge_{0.85}$/ Ge QW. In Fig. 2.10a, the strain in the well layer is zero. The light and heavy holes are strongly coupled, hence the accuracy of the effective mass approximation is very poor in this case. In Fig. 2.10b, the compressive strain is increased to -1%, so the heavy hole is split from the light hole band. Then we can see that the effective mass approximation fits better to the results of **k.p**. But the accuracy is reduced for larger in-plane k_x and k_y vectors. This trend also applies to higher subbands, as the higher subbands have lower (deeper) quantized energy, which is close to the LH bands. For the conduction band, effective mass method generally has a good agreement with the k.p method. But differences still exist, even for the quantized states when $k_t = 0$. This is because, even when $k_t = 0$, the conduction band is still coupled with valence band in the quantization direction via momentum matrix element (MME). Overall, the

parabolic effective mass approximation fits better with the **k.p** method when the coupling between bands is weak. To have a better accuracy for the valence band, coupling should always been considered, which makes **k.p** method a better choice.

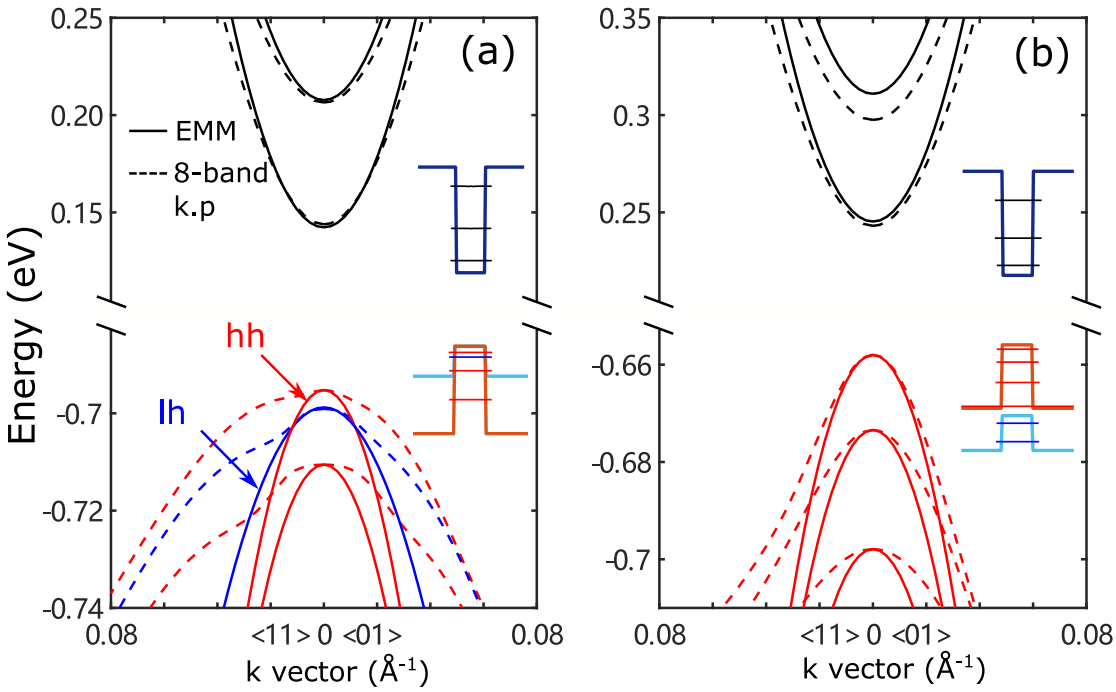

Figure 2.10: Band structure of $Si_{0.15}Ge_{0.85}$/ Ge QW calculated by effective mass method and eight-band **k.p** method. (a)Biaxial strain $\varepsilon = 0\%$ in the well; (b)$\varepsilon = -1\%$ in the well.

2.2 Eight-band k · p Method

The **k** · **p** method is a semi-empirical method based on the perturbation theory. Low order **k.p** methods like four-band, six-band and eight-band **k.p** are very useful for calculation of the electrical properties at the center Brillouin zone (Γ point). In this work, we follow Bahder's work [53], which extended the works of Kane [54] and Pikus and Bir [55]. This eight-band model includes the interaction between conduction band, heavy hole, light hole and split off bands. The matrix elements of the interaction matrix are treated by the Löwdin perturbation theory. Also the strain effects coming from the orbital part and spin-orbit part are both included. Although from the results we observed just a small difference from the ispin-orbit part when the strain is less than 2%, t still remains here for better accuracy of the results, and also to be included in case of larger strain, which may be explored in the future.

2.2.1 Hamiltonian in the Absence of Strain

Consider the Hamiltonian near the center of the Brillouin zone $\mathbf{k}=0$ without the presence of strain. The Schrödinger equation with the spin-orbit interaction for the Bloch function is:

$$\left[\frac{p^2}{2m_0}+V(\mathbf{r})+\frac{\hbar}{4{m_0}^2c^2}(\boldsymbol{\nabla}V\times\mathbf{p})\cdot\boldsymbol{\sigma}\right]\psi_{n\mathbf{k}}(\mathbf{r})=E_n(\mathbf{k})\psi_{n\mathbf{k}}(\mathbf{r}) \tag{2.41}$$

where the third term on the left represents the spin-orbit interaction and $\boldsymbol{\sigma}$ is the Pauli's spin matrix. Its components and the results of them operating on spins are:

$$\sigma_x=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\quad \sigma_y=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\quad \sigma_z=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{2.42a}$$

$$\begin{aligned}&\sigma_x\uparrow=\downarrow \qquad \sigma_y\uparrow=i\downarrow \qquad \sigma_z\uparrow=\uparrow\\ &\sigma_x\downarrow=\uparrow \qquad \sigma_y\downarrow=-i\uparrow \qquad \sigma_z\downarrow=-\downarrow\end{aligned} \tag{2.42b}$$

The wave function $\psi_{n\mathbf{k}}(\mathbf{r})$ can be written as:

$$\psi_{n\mathbf{k}}(\mathbf{r})=e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}) \tag{2.43}$$

By substituting Eq. 2.43 to Eq. 2.41, the Schrödinger equation of periodic function $u_{n\mathbf{k}}$ reads:

$$\begin{aligned}&\left[\frac{p^2}{2m_0}+V(\mathbf{r})+\frac{\hbar^2k^2}{2m_0}+\frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p}+\frac{\hbar}{4{m_0}^2c^2}(\boldsymbol{\nabla}V\times\mathbf{p})\cdot\boldsymbol{\sigma}+\frac{\hbar^2}{4{m_0}^2c^2}\boldsymbol{\nabla}V\times\mathbf{k}\cdot\boldsymbol{\sigma}\right]u_{n\mathbf{k}}(\mathbf{r})\\ &=E_n(\mathbf{k})u_{n\mathbf{k}}(r)\end{aligned} \tag{2.44}$$

The last term of the above equation can be reasonably ignored because the crystal momentum $\hbar\mathbf{k}$ is much smaller than the particle momentum $\mathbf{p}$. For the convenience, we let:

$$H_0=\frac{p^2}{2m_0}+V_0(\mathbf{r}) \tag{2.45a}$$

$$H_k=\frac{\hbar^2k^2}{2m_0} \tag{2.45b}$$

$$H_{\mathbf{k}\cdot\mathbf{p}}=\frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p} \tag{2.45c}$$

$$H_{s.o.}=\frac{\hbar}{4{m_0}^2c^2}(\boldsymbol{\nabla}V\times\mathbf{p})\cdot\boldsymbol{\sigma} \tag{2.45d}$$

$$H=H_0+H_k+H_{\mathbf{k}\cdot\mathbf{p}}+H_{s.o.} \tag{2.45e}$$

Then Eq. 2.44 becomes

$$H\left|u_{n\mathbf{k}}\right\rangle = E\left|u_{n\mathbf{k}}\right\rangle \tag{2.46}$$

For the eight-band k.p method, the Γ_6 conduction band, Γ_8 heavy hole and light hole band and Γ_7 split off band including spin are considered as the Class-A. Then $u_{n\mathbf{k}}(\mathbf{r})$ can be represented in the orthonormal basis of Class-A:

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{n} c_n(\mathbf{k}) u_{n0}(r) \tag{2.47}$$

By substituting Eq. 2.47 into Eq. 2.46, left multiplying $\langle u_{mo}|$, and using the orthogonal normalization relation $\langle u_{mo}|u_{n0}\rangle = \delta_{mn}$, the problem reduced to the following eigen-equation:

$$\sum_{n}^{A} \left(H_{mn} - E\delta_{mn}\right) c_n(\mathbf{k}) = 0 \tag{2.48}$$

where $H_{mn} = \langle u_{m0}|H|u_{n0}\rangle$ represents the interaction between $|u_{n0}\rangle$ and $|u_{m0}\rangle$. The above equation includes the coupling between the Class-A bands but ignores the effects of remote bands. These remote bands in **k.p** theory are denoted as Class-B bands. To include the effects from the remote bands, we use the Löwdin's method [56] which treats the influence from Class-B as perturbation terms. Then, instead of Eq. 2.48, the following equation should be solved:

$$\sum_{n}^{A} \left(U_{mn} - E\delta_{mn}\right) c_n(\mathbf{k}) = 0 \tag{2.49}$$

where

$$U_{mn} = H_{mn} + \sum_{\gamma\neq m,n}^{B} \frac{H_{m\gamma}H_{\gamma n}}{E - E_\gamma} \tag{2.50}$$

Without strain, the renormalized operator U is equal to the interaction Hamiltonian $H_{\mathbf{k}}$ without strain, and the matrix element of this Hamiltonian is calculated as:

$$\langle u_{m0}|H_{\mathbf{k}}|u_{n0}\rangle = \langle u_{m0}|H|u_{n0}\rangle + \sum_{\gamma\neq n,m}^{B} \frac{\langle u_{m0}|H_{\mathbf{k}\cdot\mathbf{p}}|u_{\gamma 0}\rangle \langle u_{\gamma 0}|H_{\mathbf{k}\cdot\mathbf{p}}|u_{n0}\rangle}{E - E_\gamma} \tag{2.51}$$

Note that the matrix elements here are obtained by ignoring the perturbation terms coming from the spin-orbit interaction $H_{s.o}$. Since $\gamma \notin A$, H_0 and H_k do not contribute to the perturbation,

so only the **k.p** operator $H_{\mathbf{k}\cdot\mathbf{p}}$ appears in the last term in Eq. 2.51.

For the zincblende and diamond crystal, the lowest conduction band has s-like orbital and the topmost valence bands have p-like orbitals, they contribute to the outermost valence electrons, for example $4s$, $4p_x$, $4p_y$ and $4p_z$ for Ga and As. These orbitals will have sp^3 hybridization when forming a crystal (like GaAs), and the following basis functions are then suitable to be used in the irreducible representation:

$$|S\uparrow\rangle \quad |X\uparrow\rangle \quad |Y\uparrow\rangle \quad |Z\uparrow\rangle \quad |S\downarrow\rangle \quad |X\downarrow\rangle \quad |Y\downarrow\rangle \quad |Z\downarrow\rangle \tag{2.52}$$

However, instead of using the above basis, $|j, m_j\rangle$ basis is used here, where j is the total angular momentum quantum number and m_j is its projection on z-axis. As for the basis in Eq. 2.52, H_{so} is not diagonal, but its diagonal in the $|j, m_j\rangle$ basis and the Hamiltonian is a diagonal matrix when $\mathbf{k} = 0$. This basis corresponds to Γ_6 conduction band, Γ_8 light hole, Γ_8 heavy hole and Γ_7 split-off band including spin respectively. With $|j, m_j\rangle$ basis these states are decoupled at $\mathbf{k} = 0$, which is convenient for this problem. After unitary transformation, the $|j, m_j\rangle$ basis is expressed by the basis in Eq. 2.52 as:

$$\begin{aligned}
|u^e_{10}\rangle &= \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_e = |S\downarrow\rangle \\
|u^e_{20}\rangle &= \left|\frac{1}{2}, \frac{1}{2}\right\rangle_e = |S\uparrow\rangle \\
\left|u^{lh}_{30}\right\rangle &= \left|\frac{3}{2}, \frac{1}{2}\right\rangle = -\frac{i}{\sqrt{6}}\left(|X\downarrow\rangle + i\,|Y\downarrow\rangle\right) + i\sqrt{\frac{2}{3}}\,|Z\uparrow\rangle \\
\left|u^{hh}_{40}\right\rangle &= \left|\frac{3}{2}, \frac{3}{2}\right\rangle = \frac{i}{\sqrt{2}}\left(|X\uparrow\rangle + i\,|Y\uparrow\rangle\right) \\
\left|u^{hh}_{50}\right\rangle &= \left|\frac{3}{2}, -\frac{3}{2}\right\rangle = -\frac{i}{\sqrt{2}}\left(|X\downarrow\rangle - i\,|Y\downarrow\rangle\right) \\
\left|u^{lh}_{60}\right\rangle &= \left|\frac{3}{2}, -\frac{1}{2}\right\rangle = \frac{i}{\sqrt{6}}\left(|X\uparrow\rangle - i\,|Y\uparrow\rangle\right) + i\sqrt{\frac{2}{3}}\,|Z\downarrow\rangle \\
|u^{so}_{70}\rangle &= \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = -\frac{i}{\sqrt{3}}\left(|X\uparrow\rangle - i\,|Y\uparrow\rangle\right) + \frac{i}{\sqrt{3}}\,|Z\downarrow\rangle \\
|u^{so}_{80}\rangle &= \left|\frac{1}{2}, \frac{1}{2}\right\rangle = -\frac{i}{\sqrt{3}}\left(|X\downarrow\rangle + i\,|Y\downarrow\rangle\right) - \frac{i}{\sqrt{3}}\,|Z\uparrow\rangle
\end{aligned} \tag{2.53}$$

Then, in this basis, from Eq. 2.51, the interaction matrix independent on the strain can be

extracted: [53]

$$H_{\mathbf{k}} = \begin{bmatrix} E_c+P_c & 0 & V^\dagger & 0 & \sqrt{3}V & -\sqrt{2}U & -U & \sqrt{2}V^\dagger \\ & E_c+P_c & -\sqrt{2}U & -\sqrt{3}V^\dagger & 0 & -V & \sqrt{2}V & U \\ & & -P+Q & -W^\dagger & R & 0 & \sqrt{\tfrac{3}{2}}W & -\sqrt{2}Q \\ & & & -P-Q & 0 & R & -\sqrt{2}R & \sqrt{\tfrac{1}{2}}W \\ & & & & -P-Q & W^\dagger & \sqrt{\tfrac{1}{2}}W^\dagger & \sqrt{2}R^\dagger \\ & & & & & -P+Q & \sqrt{2}Q & \sqrt{\tfrac{3}{2}}W^\dagger \\ & H.C. & & & & & Z & 0 \\ & & & & & & & Z \end{bmatrix} \tag{2.54}$$

where

$$P_c = E_c + \gamma_c \frac{\hbar^2}{2m_0}\left({k_x}^2 + {k_y}^2 + {k_z}^2\right) \tag{2.55a}$$

$$U = \frac{1}{\sqrt{3}} P_0 k_z \tag{2.55b}$$

$$V = \frac{1}{\sqrt{6}} P_0 (k_x - i k_y) \tag{2.55c}$$

$$P = -E_v + \gamma_1 \frac{\hbar^2}{2m_0}\left({k_x}^2 + {k_y}^2 + {k_z}^2\right) \tag{2.55d}$$

$$Q = \gamma_2 \frac{\hbar^2}{2m_0}\left({k_x}^2 + {k_y}^2 - 2{k_z}^2\right) \tag{2.55e}$$

$$R = -\sqrt{3}\frac{\hbar^2}{2m_0}\left[\gamma_2\left({k_x}^2 - {k_y}^2\right) - 2i\gamma_3 k_x k_y\right] \tag{2.55f}$$

$$W = 2\sqrt{3}\gamma_3 \frac{\hbar^2}{2m_0} k_z \left(k_x - i k_y\right) \tag{2.55g}$$

$$Z = E_v - \Delta_{so} - \gamma_1 \frac{\hbar^2}{2m_0}\left({k_x}^2 + {k_y}^2 + {k_z}^2\right) \tag{2.55h}$$

The lower triangle matrix in Eq. 2.54 is the conjugate of the upper matrix (i.e $\mathcal{H}_{nm} = {\mathcal{H}_{mn}}^\dagger$, $n \neq m$). E_v is the potential profile of the valence band. The conduction band potential is then defined as the sum of valence band potential and the bandgap at Γ point $E_c = E_v + E_g^\Gamma$. The reference of the potentials are arbitrarily chosen. For example, we can take E_v of Ge as zero, E_v for other materials are then obtained from the valence band offset (VBO) measured from this reference. $k_{x,y,z}$ are the components of the wave vector $\mathbf{k}$. Δ_{so} is the spin split-off energy describing the spin-orbit interaction, using the definition:

$$\Delta_{so} = -3i\left(\frac{\hbar^2}{4{m_0}^2 c^2}\right)\langle X|(\nabla V \times \mathbf{p})_y|Z\rangle \tag{2.56}$$

and P_0 is the inter-band momentum matrix element describing the coupling of the valence bands and conduction bands. P_0 and its energy form E_p are defined by:

$$P_0 = -\frac{i\hbar}{m_0}\langle S|p_x|X\rangle = \sqrt{\frac{\hbar^2}{2m_0}E_p} \tag{2.57a}$$

$$E_p = \frac{2m_0}{\hbar^2}{P_0}^2 \tag{2.57b}$$

The eight-band Luttinger parameters (γ_1, γ_2 and γ_3) in Eq. 2.54 are different from the six-band Luttinger parameters used by Luttinger and Kohn in the six-band model (γ_1^L, γ_2^L and γ_3^L) [57]. It is because the six-band model treats the lowest conduction band as Class-B (i.e. a remote band), but the eight-band model includes the conduction band into Class-A, so the contribution from the conduction band in the six-band model should be subtracted off in the eight-band model [58]:

$$\gamma_1 = \gamma_1^L - \frac{E_p}{3E_g^\Gamma + \Delta_{so}} \tag{2.58a}$$

$$\gamma_2 = \gamma_2^L - \frac{E_p}{2\left(3E_g^\Gamma + \Delta_{so}\right)} \tag{2.58b}$$

$$\gamma_3 = \gamma_3^L - \frac{E_p}{2\left(3E_g^\Gamma + \Delta_{so}\right)} \tag{2.58c}$$

The eight-band Luttinger parameters are related to the Kane parameters L', M, and N' as [54]:

$$\gamma_1 = -\frac{2}{3}\frac{m_0}{\hbar^2}(L' + 2M) - 1 \tag{2.59a}$$

$$\gamma_2 = -\frac{1}{3}\frac{m_0}{\hbar^2}(L' - M) \tag{2.59b}$$

$$\gamma_3 = -\frac{1}{3}\frac{m_0}{\hbar^2}N' \tag{2.59c}$$

in which:

$$L' = \frac{\hbar^2}{{m_0}^2}\sum_\gamma^B \frac{|\langle X|p_x|u_{\gamma 0}\rangle|^2}{E_v - E_\gamma} \tag{2.60a}$$

$$M = \frac{\hbar^2}{{m_0}^2}\sum_\gamma^B \frac{|\langle X|p_y|u_{\gamma 0}\rangle|^2}{E_v - E_\gamma} \tag{2.60b}$$

$$N' = \frac{\hbar^2}{{m_0}^2}\sum_\gamma^B \frac{\langle X|p_x|u_{\gamma 0}\rangle\langle u_{\gamma 0}|p_y|Y\rangle + \langle X|p_y|u_{\gamma 0}\rangle\langle u_{\gamma 0}|p_x|Y\rangle}{E_v - E_\gamma} \tag{2.60c}$$

Note that, similar to the Luttinger parameters, the Kane parameters L' and N' differ from the Dresselhaus-Kip-Kittel (DKK) parameters L and N [59], because Dresselhaus's six-band model uses $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis states, but Kane's model includes the conduction band basis function $|S\rangle$. The parameter M is the same for these two models because, according to the symmetry properties of the basis functions, $\langle X|p_y|S\rangle = 0$. In Eq. 2.60, the Class-B does not include the lowest conduction band. By subtracting the contribution from the conduction band, L' and N' are related to L and N by:

$$L = L' - \frac{\hbar^2}{{m_0}^2}\frac{|\langle S|p_x|X\rangle|^2}{E_g^\Gamma + \Delta_{so}/3} \tag{2.61a}$$

$$N = N' - \frac{\hbar^2}{{m_0}^2}\frac{|\langle S|p_x|X\rangle|^2}{E_g^\Gamma + \Delta_{so}/3} \tag{2.61b}$$

Additionally, we define $1/\gamma_c$ as the corrected effective mass. This parameter is defined as:

$$\gamma_c = \frac{2m_0}{\hbar^2}A' + 1 \tag{2.62a}$$

$$A' = \frac{\hbar^2}{{m_0}^2}\sum_{\gamma}^{B}\frac{|\langle S|p_x|u_{\gamma 0}\rangle|^2}{E_c - E_\gamma} \tag{2.62b}$$

Apart from its definition, we can also use the experimentally measured conduction band effective mass m_e to evaluate γ_c, using the relation [60]:

$$\gamma_c = \frac{m_0}{m_e} - E_p\frac{E_g^\Gamma + \frac{2}{3}\Delta_{so}}{E_g^\Gamma(E_g^\Gamma + \Delta_{so})} \tag{2.63}$$

From the above relation we can see that the 'correction' to the effective mass comes from the coupling between the conduction band and valence band. If the conduction band and valence band are decoupled, it would mean that $E_p = 0$, and the inverse of γ_c is just equal to the effective mass. Then the conduction band will have parabolic dispersion, which leads to the same results as from the effective mass method. Moreover, such relation of momentum matrix element and effective mass is also used to rescale the momentum matrix element entering the matrix, to avoid spurious solutions. This will be discussed in detail in the section 2.3.2.

2.2.2 Hamiltonian for Strained Semiconductor

The crystal under uniform deformation will remain periodic. But its electronic properties, such as bandgap and energy dispersion in the first Brillouin zone, will be altered. For diamond and

zincblende crystals the light hole and heavy hole bands are degenerate without strain. Such degeneracy is broken under strain, if the crystal symmetry is reduced under strain. Following Pikus and Bir's theory [55], this section will briefly describe how the strain Hamiltonian is introduced in the eight-band **k.p** model in [53].

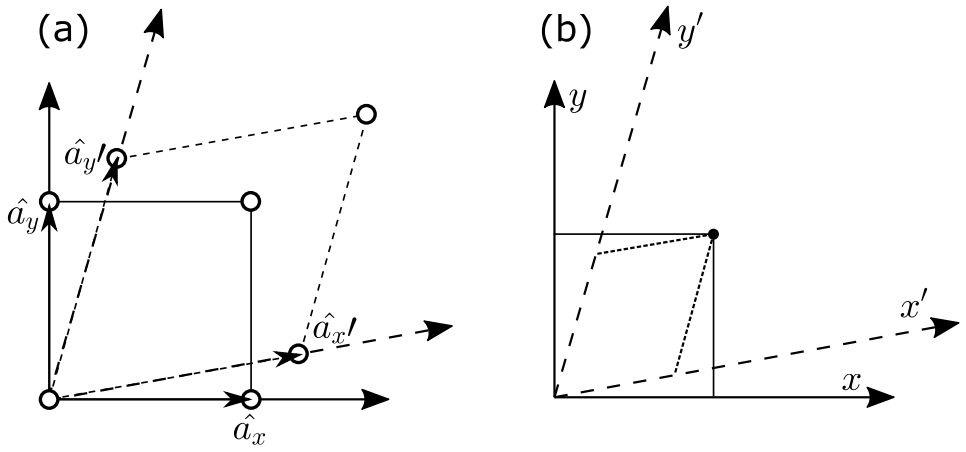

Figure 2.11: Coordinate transformation of strained semiconductor. (a) Schematic diagram of the unstrained unit cell (solid lines) and strained unit cell (dashed lines); (b) The coordinates of the same point under the unstrained(solid line) and strained(dashed lines) coordinate system

As is shown in Fig. 2.11a, the lattice vector of the strained crystal is related to that of unstrained crystal by the strain tensor:

$$\hat{\boldsymbol{a}}' = (1+\boldsymbol{\varepsilon})\hat{\boldsymbol{a}} \tag{2.64}$$

where

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix} \tag{2.65}$$

Therefore, under strain the periodicity of the potential $V(\mathbf{r})$ is changed, and then also the periodicity of wave functions. Therefore, directly applying the perturbation method on the potential is inappropriate because the perturbed wave function is the superpersition of the unperturbed wave functions which have the original periodicity. In order to retain the periodicity of potential and also the Bloch spinor in Eq. 2.41 and Eq. 2.44, we use the new lattice vector $\hat{\boldsymbol{a}}'$ as the basis in the strained coordinate system. As shown in Fig. 2.11b, for the same position

in space the coordinate of the strained system $\mathbf{r}' = (x', y', z')$ is related to that of the original system $\mathbf{r} = (x, y, z)$ by:

$$\mathbf{r} = (1 + \boldsymbol{\varepsilon})\mathbf{r}' \tag{2.66}$$

which leads to the operator $\mathbf{p} = -i\hbar\nabla$ in the unstrained coordinate system represented by that in the strained coordinate system $\mathbf{p}'$ as [55]:

$$\mathbf{p} = (1 - \boldsymbol{\varepsilon})\mathbf{p}' \tag{2.67}$$

$$p^2 \approx p\prime^2 - p'^2 - 2\mathbf{p}'\boldsymbol{\varepsilon}\mathbf{p}' \tag{2.68}$$

Similarly, the differential operator ∇ should be replaced by $(1 - \boldsymbol{\varepsilon})\nabla$. For the potential profile under strain, $V_\varepsilon(\mathbf{r})$ has a different period from that of $V_0(\mathbf{r})$. By substituting Eq. 2.66, $V_\varepsilon(\mathbf{r})$ is transformed into $V_\varepsilon((1 + \boldsymbol{\varepsilon})\mathbf{r}') = V'_\varepsilon(\mathbf{r}')$, which has the same period as $V_0(\mathbf{r}')$. The difference between the strained and unstrained potential is expanded in a series in terms of $\boldsymbol{\varepsilon}$:

$$V_{\boldsymbol{\varepsilon}}(\mathbf{r}) = V_{\boldsymbol{\varepsilon}}((1 + \boldsymbol{\varepsilon})\mathbf{r}') = V_o(\mathbf{r}') + \sum_{ij} V_{ij}(\mathbf{r}')\varepsilon_{ij} \tag{2.69}$$

where

$$V_{ij}(\mathbf{r}') = \frac{1}{2 - \delta_{ij}} \lim_{\boldsymbol{\varepsilon}\to 0} \frac{V_{\boldsymbol{\varepsilon}}((1 + \boldsymbol{\varepsilon})\mathbf{r}\prime) - V_0(\mathbf{r}')}{\varepsilon_{ij}} \tag{2.70}$$

The coefficient $1/(2 - \delta_{ij})$ exists because the identical terms $V_{ij}(\mathbf{r}')\varepsilon_{ij}$ and $V_{ji}(\mathbf{r}')\varepsilon_{ji}$ are added in Eq 2.69. Finally, the Bloch spinor is transformed into $u'_{nk}(\mathbf{r}') = u_{nk}((1 + \boldsymbol{\varepsilon})\mathbf{r}')$. Then, by substituting Eq. 2.66 - Eq. 2.70 into Eq. 2.44, the eigen-problem in the strained coordinates becomes:

$$(H' + D)u'_{nk} = E_n(\mathbf{k})u'_{nk}(\mathbf{r}') \tag{2.71}$$

where H' is the same as Eq. 2.45 with $\mathbf{r}$ and $\mathbf{p}$ directly replaced by $\mathbf{r}'$ and $\mathbf{p}'$. The additional term comes from the coordinate transformation D is then:

$$D = D_0 + D_{\mathbf{k}\cdot\mathbf{p}} + D_{s.o} \tag{2.72}$$

where

$$D_0 = \sum_{ij} D_{ij}\varepsilon_{ij} = \sum_{ij}\left(-\frac{1}{m_0}p_i' p_j' + V_{ij}(\mathbf{r}')\right)\varepsilon_{ij} \tag{2.73a}$$

$$D_{\mathbf{k}\cdot\mathbf{p}} = -\frac{\hbar}{m_0}\mathbf{k}\boldsymbol{\varepsilon}\mathbf{p}' \tag{2.73b}$$

$$\begin{aligned} D_{s.o} =& \frac{\hbar}{4{m_0}^2c^2}\Bigg(\nabla\sum_{ij}\varepsilon_{ij}V_{ij}(\mathbf{r}')\times\mathbf{p}' - \nabla V_0(\mathbf{r}')\times\boldsymbol{\varepsilon}\mathbf{p}' - \nabla\sum_{ij}V_{ij}(\mathbf{r}')\varepsilon_{ij}\times\boldsymbol{\varepsilon}\mathbf{p}' \\ & -\boldsymbol{\varepsilon}\nabla V_0(\mathbf{r}')\times\mathbf{p}' - \boldsymbol{\varepsilon}\nabla\sum_{ij}V_{ij}(\mathbf{r}')\varepsilon_{ij}\times\mathbf{p}' + \boldsymbol{\varepsilon}\nabla V_0(\mathbf{r}')\times\boldsymbol{\varepsilon}\mathbf{p}' + \boldsymbol{\varepsilon}\nabla\sum_{ij}V_{ij}(\mathbf{r}')\varepsilon_{ij}\times\boldsymbol{\varepsilon}\mathbf{p}'\Bigg)\boldsymbol{\sigma} \\ \approx& \frac{\hbar}{4{m_0}^2c^2}\Big(-\nabla V_0(\mathbf{r}')\times\boldsymbol{\varepsilon}\mathbf{p}' - \boldsymbol{\varepsilon}\nabla V_0(\mathbf{r}')\times\mathbf{p}'\Big)\boldsymbol{\sigma} \end{aligned} \tag{2.73c}$$

Note that in Eq. 2.73c only two terms remain in the final expression. This is because V_{ij} is much smaller than V_0, as indicated in [53]. Then, after coordinate transformation, the Löwdin renormalized operator in Eq. 2.50 becomes:

$$U'_{mn} = H'_{mn} + \sum_{\gamma\neq m,n}^{B}\frac{H'_{m\gamma}H'_{\gamma n}}{E-E_\gamma} + D_{mn} + \sum_{\gamma\neq m,n}^{B}\frac{H'_{m\gamma}D_{\gamma n} + D_{m\gamma}H'_{\gamma n}}{E-E_\gamma} \tag{2.74}$$

The first two terms are identical to the matrix in Eq. 2.54. The last term represents the interaction between class-A and class-B through strain. This term is neglected in [53, 61]. Thus, the strain part of the Hamiltonian is calculated as:

$$\langle u_{m0}|D|u_{n0}\rangle = \langle u_{m0}|D_0 + D_{\mathbf{k}\cdot\mathbf{p}}|u_{n0}\rangle + \langle u_{m0}|D_{s.o}|u_{n0}\rangle \tag{2.75}$$

With $|j, m_j\rangle$ basis, the explicit matrix form of the first term in Eq. 2.75 reads:

$$D_{\mathbf{k}} = D_0 + D_{\mathbf{k}\cdot\mathbf{p}} = \begin{bmatrix} a_c\varepsilon & 0 & -v^\dagger & 0 & -\sqrt{3}v & \sqrt{2}u & u & -\sqrt{2}v^\dagger \\ & a_c\varepsilon & \sqrt{2}u & \sqrt{3}v^\dagger & 0 & v & -\sqrt{2}v & -u \\ & & -p+q & -w^\dagger & r & 0 & \sqrt{\frac{3}{2}}w & -\sqrt{2}q \\ & & & -p-q & 0 & r & -\sqrt{2}r & \sqrt{\frac{1}{2}}w \\ & & & & -p-q & w^\dagger & \sqrt{\frac{1}{2}}w^\dagger & \sqrt{2}r^\dagger \\ & & & & & -p+q & \sqrt{2}q & \sqrt{\frac{3}{2}}w^\dagger \\ & H.C. & & & & & a_v\varepsilon & 0 \\ & & & & & & & a_v\varepsilon \end{bmatrix} \tag{2.76}$$

in which

$$\varepsilon = \mathrm{Tr}(\boldsymbol{\varepsilon}) \tag{2.77a}$$

$$u = \frac{1}{\sqrt{3}} P_0 \sum_j \varepsilon_{zj} k_j \tag{2.77b}$$

$$v = \frac{1}{\sqrt{6}} P_0 \sum_j \left(\varepsilon_{xj} - i\varepsilon_{yj}\right) k_j \tag{2.77c}$$

$$p = -a_v \varepsilon \tag{2.77d}$$

$$q = b\left[\varepsilon_{zz} - \frac{1}{2}\left(\varepsilon_{xx} + \varepsilon_{yy}\right)\right] \tag{2.77e}$$

$$r = \frac{\sqrt{3}}{2} b\left(\varepsilon_{xx} - \varepsilon_{yy}\right) - id\varepsilon_{xy} \tag{2.77f}$$

$$w = -d\left(\varepsilon_{xz} - i\varepsilon_{yz}\right) \tag{2.77g}$$

The **k.p** terms depend linearly on strain in $D_{\mathbf{k}\cdot\mathbf{p}}$, which leads to additional interaction between conduction band and valence band. The a_c and a_v are the deformation potentials of the conduction band and valence band, b and d are the shear deformation potentials. The deformation potentials are defined as:

$$a_c = \langle S|D_{xx}|S\rangle \tag{2.78a}$$

$$a_v = \frac{1}{3}\langle X|D_{xx}|X\rangle + \frac{2}{3}\langle X|D_{yy}|X\rangle \tag{2.78b}$$

$$b = \frac{1}{3}\langle X|D_{xx}|X\rangle - \frac{1}{3}\langle X|D_{yy}|X\rangle \tag{2.78c}$$

$$d = \frac{2}{\sqrt{3}}\langle X|D_{xz}|Z\rangle \tag{2.78d}$$

in which D_{xx} is D_{ij} when $i = j = x$ in Eq. 2.73a. The second term in Eq. 2.75 is the spin-orbit interaction from strain. The strain in this work is limited to below 2%. Although from our calculations this term makes small difference to the result when strain is small, it is still included, for better accuracy. This matrix reads:

$$D_{s.o} = \frac{\Delta_{so}}{3}\begin{bmatrix} 0_{2\times 2} & 0 & 0 \\ & N^{(1)}_{4\times 4} & N^{(2)}_{4\times 2} \\ & & N^{(3)}_{2\times 2} \end{bmatrix} \tag{2.79}$$

where

$$N^{(1)}_{4\times4} = \begin{bmatrix} -\frac{\varepsilon}{3}-\varepsilon_{zz} & -\frac{2}{\sqrt{3}}(\varepsilon_{xz}+i\varepsilon_{yz}) & \sqrt{\frac{1}{3}}(2i\varepsilon_{xy}+\varepsilon_{yy}-\varepsilon_{xx}) & 0 \\ & \varepsilon_{zz}-\varepsilon & 0 & \sqrt{\frac{1}{3}}(2i\varepsilon_{xy}+\varepsilon_{yy}-\varepsilon_{xx}) \\ & & \varepsilon_{zz}-\varepsilon & \frac{2}{\sqrt{3}}(\varepsilon_{xz}+i\varepsilon_{yz}) \\ H.C. & & & -\frac{\varepsilon}{3}-\varepsilon_{zz} \end{bmatrix}$$

$$N^{(2)}_{4\times2} = \begin{bmatrix} -\sqrt{\frac{1}{2}}(\varepsilon_{xz}-i\varepsilon_{yz}) & -\frac{1}{3\sqrt{2}}(3\varepsilon_{zz}-\varepsilon) \\ \frac{1}{\sqrt{6}}(2i\varepsilon_{xy}+\varepsilon_{yy}-\varepsilon_{xx}) & -\frac{1}{\sqrt{6}}(\varepsilon_{xz}-i\varepsilon_{yz}) \\ -\frac{1}{\sqrt{6}}(\varepsilon_{xz}+i\varepsilon_{yz}) & \frac{1}{\sqrt{6}}(2i\varepsilon_{xy}-\varepsilon_{yy}+\varepsilon_{xx}) \\ \frac{1}{3\sqrt{2}}(3\varepsilon_{zz}-\varepsilon) & -\sqrt{\frac{1}{2}}(\varepsilon_{xz}+i\varepsilon_{yz}) \end{bmatrix}$$

$$N^{(3)}_{2\times2} = \begin{bmatrix} \frac{4}{3}\varepsilon & 0 \\ 0 & \frac{4}{3}\varepsilon \end{bmatrix} \tag{2.80}$$

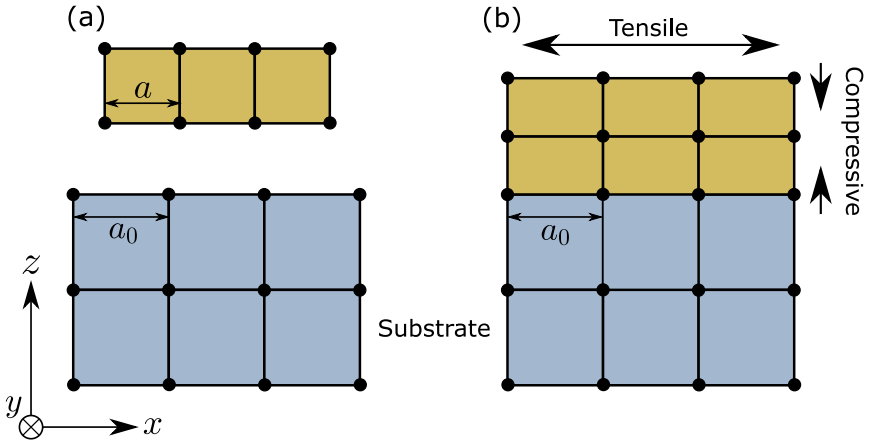

Figure 2.12: Schematic diagram of a material with smaller lattice constant pseudomorphically grown on a substrate with larger lattice constant. (a)Before growth and unstrained; (b)pseudomorphically strained.

The strain in this work is introduced in an ideal and simple way, we assume that the lattice is pseudomorphically grown along [001] orientation (z-direction). As can be seen in Fig. 2.12 the material itself has lattice constant a, and the substrate has the lattice constant of a_0. When the material is pseudomorphically grown on substrate, the material is strained and the in-plane lattice constant is the same as that of the substrate. By further restricting the strain to biaxial

strain, the strain tensor components are:

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a_0 - a}{a} \neq \varepsilon_{zz}$$
$$\varepsilon_{xy} = \varepsilon_{yz} = \varepsilon_{zx} = 0 \tag{2.81}$$

The stress is related to strain by stiffness tensor $\boldsymbol{\tau} = \mathbf{C}\boldsymbol{\varepsilon}$. The stress in the z-direction is $\tau_{zz} = C_{12}(\varepsilon_{xx} + \varepsilon_{yy}) + C_{11}\varepsilon_{zz}$. For a layer grown on a substrate there is no external stress in z-direction, i.e. $\tau_{zz} = 0$, and the strain of z-direction is obtained:

$$\varepsilon_{zz} = -\frac{2C_{12}}{C_{11}}\varepsilon_{xx} \tag{2.82}$$

To this point, we have obtained the explicit matrix forms of the three parts of the total Hamiltonian. The problem now becomes solving the eigenvalue of the sum of Eq. 2.54, Eq. 2.76 and Eq. 2.79:

$$\mathcal{H} = H_{\mathbf{k}} + D_{\mathbf{k}} + D_{s.o} \tag{2.83}$$

In summary, the total Hamiltonian $\mathcal{H}$ has two variables, $\mathbf{k}$ and $\boldsymbol{\varepsilon}$. The strain tensor in a system where a material is pseudomorphically grown on a substrate is determined by their lattice constants a_0 and a. Apart from these two variables, the Hamiltonian matrix is defined by the following constant parameters: (1)Band gap at Γ point E_g^{Γ}; (2)Spin-orbit split-off energy Δ_{so}; (3)Momentum matrix element E_p or rescaled momentum matrix element according to the electron effective mass m_e; (4)valence band offset E_v; (5)Luttinger parameters γ_1, γ_2 and γ_3; (6)Deformation potentials a_c, a_v, b, d. These constants are material parameters that are obtained from atomistic calculations or from the comparison of **k.p** results with experimental measurement.

2.2.3 Band Edge Energy

The band edge energies in the absence of strain are just the diagonal elements of the Hamiltonian in Eq. 2.54, and when $\mathbf{k} = 0$ the off-diagonal parts are all zero. Thus the band edge energies of CB, LH, HH and SO are E_c, E_v, E_v and $E_v - \Delta_{so}$ respectively. Under biaxial strain, and when $\mathbf{k} = 0$, off-diagonal terms still exist in $\langle u_{lh}|D_{\mathbf{k}}|u_{so}\rangle$ and $\langle u_{lh}|D_{s.o}|u_{so}\rangle$. The band edge energies of CB and HH are still the diagonal parts, because they are not coupling with other

bands when $\mathbf{k}=0$. However, the band edge energies of LH and SO are no longer the diagonal matrix elements. In order to obtain the values of the band edge energies, we can reduce the eight by eight total Hamiltonian $\mathcal{H}$ to a two by two matrix which only contains the LH and SO bands. The eigen problem becomes:

$$\begin{bmatrix} \mathcal{H}_{ll} & \mathcal{H}_{ls} \\ \mathcal{H}_{ls} & \mathcal{H}_{ss} \end{bmatrix} \begin{bmatrix} c_{lh}(\mathbf{k}=0) \\ c_{so}(\mathbf{k}=0) \end{bmatrix} = E(\mathbf{k}=0) \begin{bmatrix} c_{lh}(\mathbf{k}=0) \\ c_{so}(\mathbf{k}=0) \end{bmatrix} \tag{2.84}$$

where $\mathcal{H}_{ll} = \langle u_{lh}|\mathcal{H}(\mathbf{k}=0)|u_{lh}\rangle$, $\mathcal{H}_{ss} = \langle u_{so}|\mathcal{H}(\mathbf{k}=0)|u_{so}\rangle$ and $\mathcal{H}_{ls} = \langle u_{lh}|\mathcal{H}(\mathbf{k}=0)|u_{so}\rangle$, which gives:

$$\mathcal{H}_{ll} = E_v - p + q - \frac{\Delta_{so}}{3}\left(\frac{\varepsilon}{3} + \varepsilon_{zz}\right) \tag{2.85a}$$

$$\mathcal{H}_{ss} = E_v - \Delta_{so} - p + \frac{\Delta_{so}}{3}\frac{4}{3}\varepsilon \tag{2.85b}$$

$$\mathcal{H}_{ls} = \pm\left[\sqrt{2}q + \frac{\Delta_{so}}{3}\frac{1}{3\sqrt{2}}(3\varepsilon_{zz} - \varepsilon)\right] \tag{2.85c}$$

The $\pm$ sign in Eq. 2.85c is for different spins and it makes no difference to the eigenvalue. The p and q have the same definition as Eq. 2.77. The eigenvalues of the two by two matrix are trivial to find as we simply need to solve the equation:

$$\begin{vmatrix} \mathcal{H}_{ll} - E & \mathcal{H}_{ls} \\ \mathcal{H}_{ls} & \mathcal{H}_{ss} - E \end{vmatrix} = 0 \tag{2.86}$$

After solving the above equation, together with HH and CB, the band edge energies are:

$$E^{\varepsilon}_{cb}(\mathbf{k}=0) = E_c + a_c\varepsilon \tag{2.87a}$$

$$E^{\varepsilon}_{lh}(\mathbf{k}=0) = \frac{1}{2}\left(\mathcal{H}_{ll} + \mathcal{H}_{ss} + \sqrt{(\mathcal{H}_{ll} - \mathcal{H}_{ss})^2 + 4\mathcal{H}_{ls}^2}\right) \tag{2.87b}$$

$$E^{\varepsilon}_{hh}(\mathbf{k}=0) = E_v - p - q + \frac{\Delta_{so}}{3}(\varepsilon_{zz} - \varepsilon) \tag{2.87c}$$

$$E^{\varepsilon}_{so}(\mathbf{k}=0) = \frac{1}{2}\left(\mathcal{H}_{ll} + \mathcal{H}_{ss} - \sqrt{(\mathcal{H}_{ll} - \mathcal{H}_{ss})^2 + 4\mathcal{H}_{ls}^2}\right) \tag{2.87d}$$

In some books and papers the spin-orbit interaction from strain is ignored, hence $D_{s.o}$ is dropped. If we also drop the last term in Eq. 2.85a, Eq. 2.85b and Eq. 2.85c which come from the $D_{s.o}$, and use Eq. 2.86 again, the result for the band edge of LH and SO will be identical to that in,

for example, [62] and [52], which gives:

$$E^{\varepsilon}_{lh}(\mathbf{k}=0) \approx E_v - p + \frac{1}{2}\left(q - \Delta_{so} + \sqrt{{\Delta_{so}}^2 + 2\Delta_{so}q + 9q^2}\right) \tag{2.88a}$$

$$E^{\varepsilon}_{so}(\mathbf{k}=0) \approx E_v - p + \frac{1}{2}\left(q - \Delta_{so} - \sqrt{{\Delta_{so}}^2 + 2\Delta_{so}q + 9q^2}\right) \tag{2.88b}$$

2.2.4 Band Structure of Bulk Material

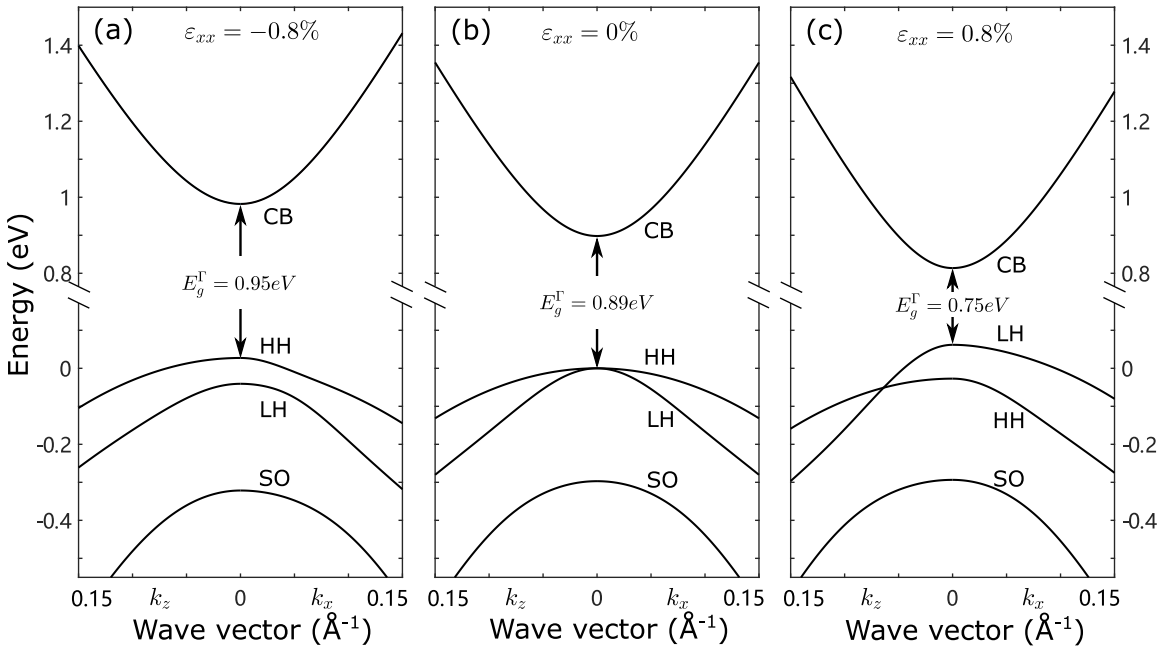

Figure 2.13: Bulk band structure of Ge under different biaxial strain. (a)$\varepsilon = -0.8\%$; (b)$\varepsilon = 0\%$; (c)$\varepsilon = 0.8\%$

To illustrate how the band structure changes under strain, we use Ge as an example, the band structures are plotted for wave vectors upto 0.15 Å^{-1}, which is acceptable for eight band k.p as this is only $\sim$ 13% of the whole span of the Brillouin zone. The strain is considered as biaxial strain, equal in x and y directions. As can be seen in Fig. 2.13, without the strain the band structure is identical along [001] (k_z-direction) and [100] (k_x-direction). Heavy holes have a larger effective mass than light holes in both directions. When $\varepsilon = -0.8\%$ (i.e compressive strain), the degeneracy of heavy holes and light holes is lifted, heavy hole band becomes higher than the light hole band. In k_z direction, the heavy hole band keeps its larger effective mass than light hole band. However the heavy hole under compressive strain in k_x direction have a lighter effective mass near the center, and becomes larger when k_x increases. When the tensile strain is applied, the light hole becomes higher than the heavy hole band. In k_z direction the

light holes have a smaller effective mass than heavy holes, so they will intersect at some point. In k_x direction, the light holes have larger effective mass than heavy holes, which leads to the in-plane effective mass approximation in Eq. 2.40a, that for the in-plane effective mass near Brillouin zone center, the heavy holes will actually have smaller effective mass and light holes will have larger effective mass under strain. The anisotropies and the strain dependence of the effective mass has been discussed in detailed in [63].

Under biaxial strain, the band gap is:

$$\begin{aligned} E^{\Gamma}_{g,\varepsilon} &= E^{\varepsilon}_{cb}(\mathbf{k}=0) - E^{\varepsilon}_{hh}(\mathbf{k}=0), \quad \varepsilon_{xx} < 0 \\ E^{\Gamma}_{g,\varepsilon} &= E^{\varepsilon}_{cb}(\mathbf{k}=0) - E^{\varepsilon}_{lh}(\mathbf{k}=0), \quad \varepsilon_{xx} > 0 \end{aligned} \tag{2.89}$$

Generally, for both signs of strain the top of the valence band is slightly higher than that without strain, but since the deformation potential of the conduction band is much larger than that of valance band, it is the conduction band deformation potential that mainly determines the behaviour of the band gap under stain. Therefore with $\varepsilon > 0$ for tensile strain, and because $a_v < 0$, the band gap of Ge will be smaller for tensile strain. And similarly, the band gap will be larger for compressive strain. It is worth mentioning that, although Germanium is an indirect band gap material, the L-valley is only ∼150meV below the Γ-valley. The deformation potential of the L-valley is much smaller than that of Γ-valley, hence by introducing tensile strain it can transformed into a direct band gap material. The widely accepted value to achieve this transformation is about 1.7% [64, 65].

2.3 Implementation of $\mathbf{k}\cdot\mathbf{p}$ Method in Heterostructure Semiconductors

In a two dimensional heterostructure, the material parameters are different in different layers. More importantly, the potential profile is offset by heterostructure alignment. With a lower potential region sandwiched by two higher potential barriers, a quantum well is formed, which enables carriers to be confined inside the layer with lower potential and the energy of these carriers is usually quantized. Such carrier confinement and quantization of energy leads to a different density of states from that of the bulk material, thus better performance can be

expected for various optical devices such as lasers and modulators. In this section, the numerical implementation of **k.p** method for quantum well structures is introduced, together with explanations of how the spurious solutions are eliminated in this work.

2.3.1 Discretization of Hamiltonian

Different from the bulk material, the two dimensional material system has discrete solutions in the quantization direction. Assuming the growth direction is z-direction, the wave vector k_z should be replaced by the quantum mechanical operator $-i\partial/\partial z$, because the wave function is quantized in the z-direction. The k_x and k_y should remain the same as continuous positive values because the wave function in xy plane is a plane wave. To simplify the problem, we recast the total Hamiltonian matrix (Eq. 2.83) as a linear combination of terms with different orders of k_z:

$$\mathcal{H} = \mathcal{H}^{(2)} + \mathcal{H}^{(1)} + \mathcal{H}^{(0)} \tag{2.90}$$

The superscript numbers (2), (1) and (0) correspond to the matrix with terms with square of k_z, linear with k_z and without k_z. For example for $\mathcal{H}^{(2)}$ it is a diagonal matrix and its explicit form is:

$$\mathcal{H}^{(2)} = \frac{\hbar^2}{2m_0} \cdot \begin{bmatrix} \gamma_c & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_c & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -(\gamma_1+2\gamma_2) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -(\gamma_1-2\gamma_2) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -(\gamma_1-2\gamma_2) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -(\gamma_1+2\gamma_2) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_1 \end{bmatrix} k_z^2 \tag{2.91}$$

Similarly, we can easily write the other two terms $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(0)}$. The matrix elements in each term can then be expressed by the general form as $P{k_z}^2$, Pk_z and P. The P is a general notation for the sum of the parameters multiplying ${k_z}^2$, ${k_z}^1$ and ${k_z}^0$, and it is usually a function

of coordinate.

For the second order term ${k_z}^2$, if we keep the sequence of the operator as it was, then $(P{k_z}^2)^\dagger = {k_z}^2 P^\dagger \neq P{k_z}^2$, hence the hermiticity is broken, and because the second order terms are on the diagonal, its necessary to maintain the hermiticity of these terms. To retain the general hermiticity, this term is then be replaced by $k_z P k_z$. Same as was done in the effective mass method, the whole structure is discretised into N layers with the same width Δz, and denote the envelope wave function across the whole structure in z-direction as $F(z)$. Then, when $k_z P k_z$ operates on the wave function $F(z)$, the discrete expansion at z_j reads:

$$\begin{aligned}(k_z P k_z F)_{z_j} =& -\frac{\partial}{\partial z} P \frac{\partial}{\partial z} F\Big|_{z_j} \\ =& \frac{(-P_{j-1}-P_j)F_{j-1} + (P_{j-1}+2P_j+P_{j+1})F_j + (-P_j - P_{j+1})F_{j+1}}{2\Delta z^2}\end{aligned} \tag{2.92}$$

The standard boundary condition gives $F_0 = 0$ and $F_{N+1} = 0$. Then, the approach similar to that applied to Eq. 2.29 and Eq. 2.30 is again used. Thus, each matrix element $\mathcal{H}^{(2)}_{mn}$ is further expanded to a $N \times N$ matrix using the following expression:

$$\mathcal{H}^{(2)}_{mn} = k_z P k_z = \frac{1}{2\Delta z^2} \cdot \begin{bmatrix} P_0+2P_1+P_2 & -P_2-P_1 & 0 & \cdots & 0 \\ -P_1-P_2 & P_1+2P_2+P_3 & -P_2-P_3 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & -P_{N-2}-P_{N-1} & P_{N-2}+2P_{N-1}+P_N & -P_{N-1}-P_N \\ 0 & \cdots & 0 & -P_{N-1}-P_N & P_{N-1}+2P_N+P_{N+1} \end{bmatrix}$$
$$(m = n) \tag{2.93}$$

Note that P_j is the parameter at z_j point. $P_0 = P_1$ and $P_{N+1} = P_N$ in the above equation. And because the second order terms appear only on the diagonal, $\mathcal{H}^{(2)}_{mn}$ is just $N \times N$ matrices with zero element when $m \neq n$.

The first order terms with k_z only appear in off-diagonal elements of the Hamiltonian. For the operator Pk_z, the customary method is symmetrizing this operator as well, thus $Pk_z \to (Pk_z + k_z P)/2$, so that $\mathcal{H}^{(1)}_{mn} = (\mathcal{H}^{(1)}_{nm})^\dagger$. Such method was argued by Luttinger [66] and Foreman [67], because although the whole matrix is hermitian, the individual elements do not

need to be hermitian symmetric. Thus, simply symmetrizing the operator may lead to non-physical results. Specifically, there are two parameters involved in the first order of k_z terms, the momentum matrix element P_0 and Luttinger parameter γ_3. From the physical definition, $P_0 k_z$ should retain the order as it was. The Luttinger parameter γ_3 is more complicated, it is linearly related to the Kane's parameter N' by $\gamma_3 = (-m_0/3\hbar^2)N'$ and contains the contribution from different remote bands. From Foreman's theory [67], the order of the operator should be determined by the symmetry type of the remote band. Thus:

$$\langle X|H_{\mathbf{k}}|Z\rangle = k_x N'_+ k_z + k_z N_- k_x = -3\frac{\hbar^2}{m_0}(k_x\gamma_{3+}k_z + k_z\gamma_{3-}k_x) \tag{2.94}$$

in which N'_+ and N_- have the same definition of the first and second terms in Eq. 2.60c, representing the coupling with different symmetry type bands. In [68], Foreman has derived the explicit form of the six band Hamiltonian $H_{6\times6}$ and the operators order in $\left\langle u^{lh}\middle|H_{6\times6}\middle|u^{hh}\right\rangle$, and $\left\langle u^{lh}\middle|H_{6\times6}\middle|u^{so}\right\rangle$ is modified. In these two matrix elements, $\gamma_3^L k_z$ then respectively becomes:

$$\gamma_3^L k_z \to (\sigma-\delta)\left(-i\frac{\partial}{\partial z}\right) + \left(-i\frac{\partial}{\partial z}\right)\pi, \qquad lh \leftrightarrow hh \tag{2.95a}$$

$$\gamma_3^L k_z \to \left[\frac{1}{3}(\sigma-\delta)+\frac{2}{3}\pi\right]\left(-i\frac{\partial}{\partial z}\right) + \left(-i\frac{\partial}{\partial z}\right)\left[\frac{2}{3}(\sigma-\delta)+\frac{1}{3}\pi\right], \qquad lh \leftrightarrow so \tag{2.95b}$$

where σ, δ and π are the Foreman's parameters which can be calculated from the six-band Luttinger parameters. Moreover, non-zero terms are also added to the coupling in the light hole and split off bands in Foreman's Hamiltonian, which are usually zero in other forms of the Hamiltonian matrix [53, 57, 69]. It seems that, derived from the definition of the Foreman's parameter [68], such transformation can be directly implemented to the 8-band Hamiltonian by simply subtracting the Γ_6 conduction band contribution from σ, because σ represents the contribution from all s-like orbitals:

$$\sigma' = \sigma - \frac{E_p}{2\left(3E_g^\Gamma + \Delta_{so}\right)} \tag{2.96}$$

However, with the lack of literature for explicit form of eight-band Hamiltonian with $|j, m_j\rangle$ basis developed from Foreman's theory, it is uncertain if implementing such operator ordering is acceptable or not. Without considering the orders of the operators, the Hamiltonians developed for bulk material is still frequently used in the modelling of quantum wells [40, 41, 52, 70, 71].

For these reasons, we still retain the order of $\gamma_3 k_z$ as the traditional way. Indeed, simply applying the original order of the operator will cause the numerical instability (i.e. the spurious solutions). But in quantum well, only one direction is quantized, it will not generate additional spurious bands in the quantization direction k_z because the matrix elements containing γ_3 are equal to zero when $k_x = k_y = 0$. Although some deviations may be observed when in-plane wave vectors are not zero, they are usually small and negligible, as far as we can see from the cases in this work.

Next, since the Luttinger parameter γ_3 and the momentum matrix element E_p will have the same operator order as in Eq. 2.55, Pk_z operating on the envelope wave function at z_j can be easily written using the central difference scheme:

$$(Pk_zF)_{z_j} = P\left(-i\frac{\partial}{\partial z}\right)F\Big|_{z_j} = \frac{-i}{2\Delta z}(P_jF_{j+1} - P_jF_{j-1}) \tag{2.97a}$$

$$[(Pk_z)^\dagger F]_{z_j} = -i\frac{\partial}{\partial z}P^\dagger F\Big|_{z_j} = \frac{-i}{2\Delta z}(P^\dagger_{j+1}F_{j+1} - P^\dagger_{j-1}F_{j-1}) \tag{2.97b}$$

The first term $\mathcal{H}^{(1)}_{mn}$ is expanded by:

$$\mathcal{H}^{(1)}_{mn} = Pk_z = \frac{-i}{2\Delta z}\begin{bmatrix} 0 & P_1 & 0 & \cdots & 0 \\ -P_2 & 0 & P_2 & & \vdots \\ \vdots & \ddots & \ddots & \vdots & \\ \vdots & & P_{n-1} & 0 & P_{n-1} \\ 0 & \cdots & 0 & P_n & 0 \end{bmatrix} \tag{2.98a}$$

$$(\mathcal{H}^{(1)}_{mn})^\dagger = (Pk_z)^\dagger = \frac{-i}{2\Delta z}\begin{bmatrix} 0 & P_2^\dagger & 0 & \cdots & 0 \\ -P_1^\dagger & 0 & P_3^\dagger & & \vdots \\ \vdots & \ddots & \ddots & \vdots & \\ \vdots & & P^\dagger_{n-2} & 0 & P_n^\dagger \\ 0 & \cdots & 0 & P^\dagger_{n-1} & 0 \end{bmatrix} \tag{2.98b}$$

Finally, for the zero order term, the expanded matrices are directly given:

$$\mathcal{H}^{(1)}_{mn} = P = \begin{bmatrix} P_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & P_n \end{bmatrix}, \qquad (\mathcal{H}^{(1)}_{mn})^\dagger = P^\dagger = \begin{bmatrix} P_1^\dagger & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & P_n^\dagger \end{bmatrix} \tag{2.99}$$

Finally, we can see that each matrix element of 8×8 Hamiltonian is expand to $N\times N$ tridiagonal matrices. The whole Hamiltonian for the heterostructure semiconductor is now a $8N\times 8N$ matrix. The problem now becomes solving the eigenfunction:

$$\begin{Bmatrix} \mathcal{H}_{11} & \cdots & & \mathcal{H}_{1n} & & \cdots & \mathcal{H}_{18} \\ \vdots & & & \vdots & & & \vdots \\ & & \begin{bmatrix} \ddots & \ddots & & \\ \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \end{bmatrix}_{mn} & & & \\ \mathcal{H}_{m1} & \cdots & & & & \cdots & \mathcal{H}_{m8} \\ & & & & & & \\ \vdots & & & \vdots & & & \vdots \\ \mathcal{H}_{81} & \cdots & & \mathcal{H}_{8n} & & & \mathcal{H}_{88} \end{Bmatrix} \begin{Bmatrix} F_1 \\ \vdots \\ \begin{bmatrix} F_n^1 \\ \vdots \\ F_n^N \end{bmatrix} \\ \vdots \\ F_8 \end{Bmatrix} = E(k_x,k_y) \begin{Bmatrix} F_1 \\ \vdots \\ \begin{bmatrix} F_n^1 \\ \vdots \\ F_n^N \end{bmatrix} \\ \vdots \\ F_8 \end{Bmatrix} \tag{2.100}$$

To make it clear, F_n^j is the envelope wave function with the $|u_{n0}\rangle$ basis at z_j point. The total wave function of the l-th subband is given by:

$$\Psi_{l\mathbf{k}}(z) = \sum_{n=1}^{8} F_{ln\mathbf{k}}(z)\,|u_{n0}\rangle \tag{2.101}$$

With a complex matrix, the LAPACK subroutine 'zheevd' is used to solve the eigenproblem. The eigenvalues obtained are the subband energies at (k_x, k_y) point. The corresponding eigenvectors are the envelope wavefunctions of these subbands. Since such matrix has $8N$ eigenvalues, only those near the bandgap are useful, so subroutines are designed to filter the unnecessary results to maintain the computational efficiency.

2.3.2 Spurious Solutions Problem

The spurious solution is a common problem that frequently appears in high order eigenvalue problems. This problem in **k.p** model has been discussed in many papers [72–76]. Since in **k.p** method the bands are strongly coupled with each other, the spurious solutions will pollute the correct results and make them inappropriate to be used for further calculation of optical properties. In quantum well structures, the envelope wave functions of the spurious solutions have strange shapes, and they usually oscillate much stronger than normal. Simply removing them by examining the shape of the wave function at the programming level is not enough, because the coupling will also result in oscillations of the normal wave functions, which makes them no longer smooth. Thus, we need to prevent the generation of spurious solutions at the

level of their mathematical source.

In this work, we found that the major source of the spurious solutions is the non-monotonicity of the $E - k_z$ relation. To explain this, a similar analysis to that in [76] is made starting from a two band Hamiltonian (CB, VB), assuming $\mathbf{k} = k_z$:

$$H_{2\times 2} = \begin{bmatrix} A_c k_z^2 + E_g & iPk_z \\ -ik_z P & A_v k_z^2 \end{bmatrix} \tag{2.102}$$

The eigenvalue of such matrix is obtained by solving

$$\begin{vmatrix} A_c k_z^2 + E_g - E & iPk_z \\ -ik_z P & A_v k_z^2 - E \end{vmatrix} = 0$$
$$\longrightarrow \quad a_2 k_z^4 + a_1 k_z^2 + a_0 = 0 \tag{2.103}$$

where

$$a_2 = A_c A_v \tag{2.104a}$$
$$a_1 = E_g A_v - (A_c + A_v)E - P^2 \tag{2.104b}$$
$$a_0 = E(E - Eg) \tag{2.104c}$$

The a_2, a_1 and a_0 are the coefficients of k_z^4, k_z^2 and k_z^0 terms. First we consider the situation when $A_c A_v = a_2 > 0$. The schematic diagram of the $E - k$ relation is plotted in Fig. 2.14a. As can be seen, for conduction band, if $A_c + A_v < 0$ and $E_g < E < E_{cr}$, gives to the coefficient $a_1 < 0$, where $E_{cr} = (E_g A_v - P^2)/(A_c + A_v)$. The sign of the coefficients (a_2, a_1, a_0) are $(+, -, +)$. According to Descartes' rule of signs, when the sign of a polynomial changes twice, there are two or zero positive roots. For the conduction band at least one root exists, hence there will be two roots of k_z^2 within the interval (E_g, E_{cr}). The one with a higher k_z is thus the spurious solution. When $E > E_{cr}$ no positive solution will exist. We can easily see that the energy relation is no longer monotonic against k_z. Due to the non-monotonicity, spurious solution can also be found in the band gap and in valence band. Such conclusion can also be justified by the rule of signs. As the sign for the band gap and valence band are $(+, -, -)$ and $(+, -, +)$, there is one positive root for k_z^2 in the band gap and there are two different roots in the valence band. Those with high k_z are the spurious solutions. One the other hand, when

$A_c + A_v > 0$, non-monotonicity will appear in the valence band, as shown in Fig. 2.14b. Since E_{cr} now becomes negative and the condition leads to $a_1 < 0$, the spurious solution will appear when energy is larger than E_{cr}.

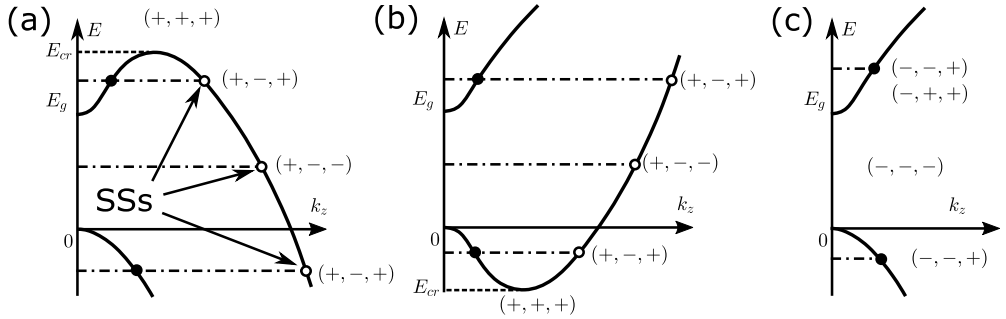

Figure 2.14: Schematic band structure of two-band model. (a)$a_3 > 0$, $A_c + A_v < 0$ the conduction band is non-monotonic; (b)$a_3 > 0$, $A_c + A_v > 0$ the valence band is non-monotonic; (c)$a_3 < 0$ and the $E - k_z$ relation is monotonic, the spurious solution is eliminated.

The $E - k$ relation for $a_2 < 0$ is plotted in Fig. 2.14c. In the conduction band there are two possibilities for the signs, these are $(-,-,+)$ and $(-,+,+)$. Both of them have only one change in sign, and yield only one positive solution. The $E(k)$ is monotonic under such circumstances. In the band gap the signs are $(-,-,-)$, indicating that no real solution can be found in this region. In valence band the signs are $(-,-,+)$, so only one positive roots exists. Clearly, in Fig. 2.14c, the $E-k$ relation is monotonic and no spurious solution will be found in this case. In summary, the criteria for a monotonic $E-k$ diagram for a two by two Hamiltonian is $A_c A_v < 0$. Similarly, the criteria for eight-band **k.p** is derived in [74]:

$$a_3 = \gamma_c(\gamma_1 + 4\gamma_2)(\gamma_1 - 2\gamma_2) > 0 \tag{2.105}$$

If this condition is fulfilled, the spurious solution originating from this source can be avoided. Next, the idea to be implemented here is to modify the input parameters without altering too much the physical properties of the system. Thus we follow the method used in [77]. Firstly, the parameter γ_c at is set to 1 as a constant. From the definition in Eq. 2.62a, this implies that the coupling between the remote band and the lowest conduction band $\langle S|p_x|u_{\gamma 0}\rangle$ is dropped. Then the rescaled momentum matrix element E_p is calculated from the relation in Eq. 2.63, which gives:

$$E_p = \left(\frac{m_0}{m_e} - 1\right)\frac{E_g^\Gamma(E_g^\Gamma + \Delta_{so})}{E_g^\Gamma + \frac{2}{3}\Delta_{so}} \tag{2.106}$$

For consistency, the eight band Luttinger parameters γ_1, γ_2 and γ_3 should also be recalculated from the six band Luttinger parameters γ_1^L, γ_2^L and γ_3^L and the rescaled momentum matrix elements. Although, from the mathematical point of view, such method does not ensure the criterium Eq. 2.105 is fulfilled, we found that with the material parameter sets that we used, it is satisfied with most of the alloy parameters calculated by Vegard's law, within the necessary parameter space of alloy compositions. The value of E_p, after rescaling, will show some difference from the experimental values, e.g. the value of E_p in [78] is 26.3 eV, and the rescaled value is 24.36 eV. This is acceptable because the change is just 7.4% for the major component (Ge) in the structure, and the band structure is not very sensitive to E_p, so it is worthwhile to sacrifice some accuracy in order to eliminate spurious solutions.

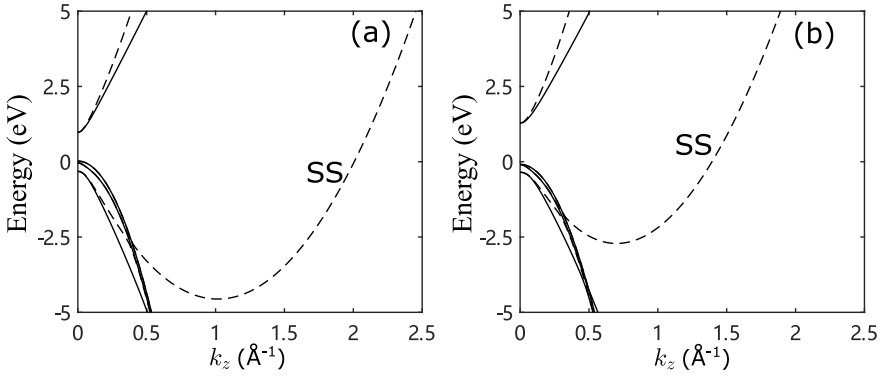

Figure 2.15: Band structure of (a)Ge and (b)$Si_{0.15}Ge_{0.85}$, the solid and dashed lines represent the band structure after and before E_p is rescaled.

Some examples of the spurious solutions are given next. Fig. 2.15 shows the calculated bulk band structures of $Si_{0.15}Ge_{0.85}$ and Ge. Although the eight-band **k.p** method is valid only in a very limited range, the band structures are still plotted in a larger range in order to show the origin of the spurious solution. In Fig. 2.15, the solid lines are calculated using the rescaled E_p while the dashed lines are the results with the original E_p from the experimental data used in [78]. Both examples set $\gamma_c = 1$. By using the original data from [78], a_3 is -12.8 and -7.2 for $Si_{0.15}Ge_{0.85}$ and Ge respectively, which do not fulfill the criterium in Eq. 2.105. The light hole bands of both materials are then non-monotonic, thus spurious solution is expected. After rescaling E_p, values of a_3 are 7.9 and 6.6 for $Si_{0.15}Ge_{0.85}$ and Ge. Then the $E - k$ relations for all bands are monotonic. The spurious solution for large k_z is eliminated.

Fig. 2.16c shows the band structure of a 12nm $Si_{0.15}Ge_{0.85}$/Ge quantum well on the $Si_{0.15}Ge_{0.85}$ substrate. Again the dashed lines represent the results with E_p without rescaling, and the solid

lines are for that after rescaling. Because of the nature of the matrix FDM, the whole structure can be viewed as being inside an infinite quantum well. The k_z is quantized, and the spurious solution will appear in the form of unphysical subbands. We can see that, apart from the spurious solution, the band structures of valence band are almost identical, and only a small difference observed in the conduction band, in which the rescaled results will have larger effective mass. Such difference is even smaller, and ignorable, for the first conduction subband, which is what we are most interested in. In bulk material, without rescaling, the non-monotonicity appears in both the barriers and well layer materials, and the spurious solutions discussed here can come from both of them. Moreover, since the non-monotonicity appears in valence band in bulk, and the boundary of k_z is proportional to the number of discrete point used in the finite difference method (FDM), such spurious solutions will only appear in valence band if a small number of points is chosen. However, they will appear in bandgap and conduction band if we increase the number of points. To have a good wave function, the number of points can not be too small, so these unphysical subbands in the band structure would be unavoidable, and it is necessary to remove them, while keeping the accuracy of the wavefunction.

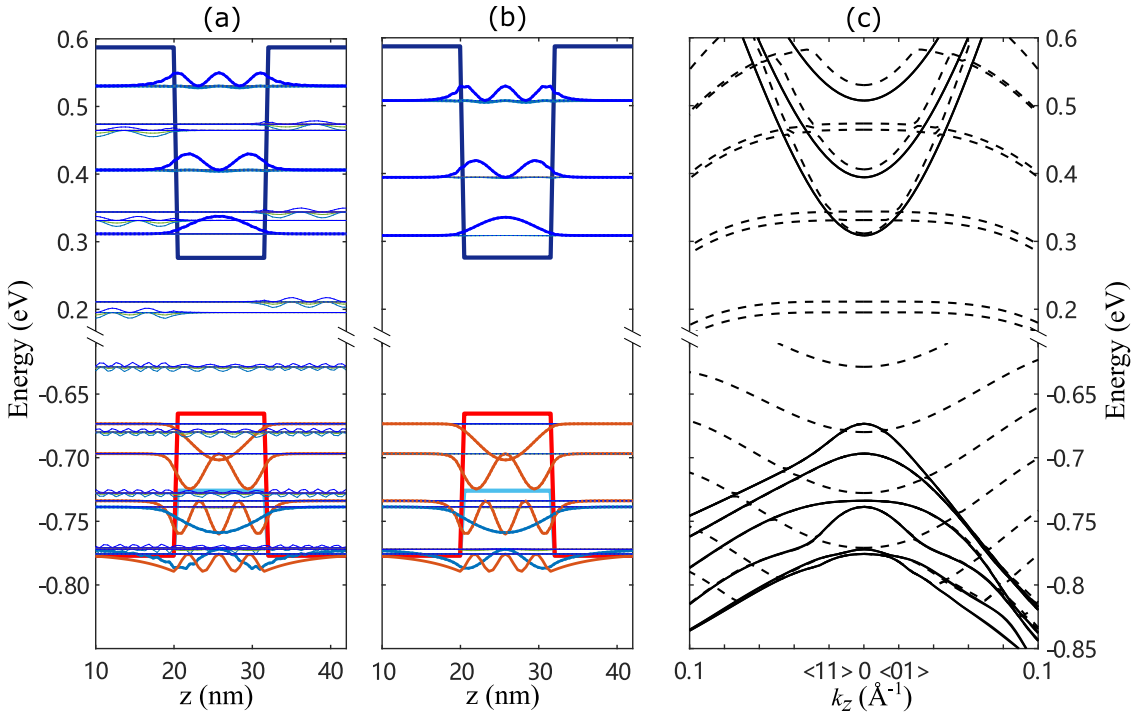

Figure 2.16: Wave functions and band structures of normal and spurious solutions in a 12nm $Si_{0.15}Ge_{0.85}$/Ge quantum well. (a)The wave function for $k_t = 0$ with spurious solution; (b)The wave function for $k_t = 0$ without spurious solution; (c)The band structure before(dashed line) and after(solid line) the method to avoid spurious solution is used.

Fig. 2.16a and Fig. 2.16b show the wave function for $k_t = 0$ before and after we used the method to remove the spurious solutions, with 100 discrete points. The spurious solution wave function can be easily identified. if it is near a normal band, it can also influence the shape of the normal wavefunction. For example, the second light hole band (cyan solid line) near the edge of the quantum well, although it has generally reasonable shape, it still has small oscillations in the envelope wave function. The heavy hole wave function is not affected here, because it does not couple to any other band when $k_t = 0$, but for $k_t \neq 0$ the wave function can still be affected. From Fig. 2.16b, we can see that the method can perfectly remove these unphysical subbands, and the wave functions are smoothly confined in the quantum well.

The non-monotonicity is not the only source of the spurious solution. A special case was found when we rescaled γ_c according to the experimental value of E_p, instead of rescaling E_p. For Ge, γ_c will have a negative value, but the criterion of Eq. 2.105 is fulfilled. The $E - k_z$ relation is monotonic, but spurious solution can still appear for quantum well structure if the number of points is small. When using a large number of points, these spurious solutions in conduction and valence band can be removed, but there remains a stationary unphysical band in the band gap. The wavefunction of this band is just a horizontal line with some oscillations at the boundary of the whole structure (i.e. the boundary of the infinite well). This led us to suspect that the source of this kind of spurious solution comes from the discrete differential scheme. More discussion about the differential scheme can be found in [73, 74]. In fact, it is challenging for a code to guarantee the problem-free solutions, because spurious solution can come from different sources. One has to always check the band structure, and even the wave function, to identify the spurious solution. However since the method from [77] is good enough to solve this problem in all cases considered in this work, we will not give any further discussion about other possible sources of spurious solutions.

2.3.3 Band Structure of QW and Model Validation

To add more validity to the model, and to exclude any possible errors at the programming level, our model is used to reproduce the experimentally measured band structure of a $Al_{0.315}Ga_{0.685}As$ /GaAs QW in Fig. 2.17 [79]. The band structure was measured in [79] by photoluminescence (PL) method. The wave vector k is determined by measuring the wavelength of photon generated by recombination of hot electron and neutral acceptors, thus the wavevector can be

calculated taking the parabolic energy dispersion of conduction band. The anisotropy of the band structure was obtained by using different orientations of the excitation light. In Fig. 2.17, despite a small split for different spin (probably because of the operator order of $\gamma_3 k_z$ and the differential scheme used in this work), the calculated results using our model show a good agreement with the experimental results. It should be noticed that, since the parameters used in AlGaAs/GaAs system are well established, such a comparison should only verify the reliability of this model. As a semi-empirical method, the accuracy of the model is further determined by the choice of material parameters when applying it to the SiGeSn/GeSn system.

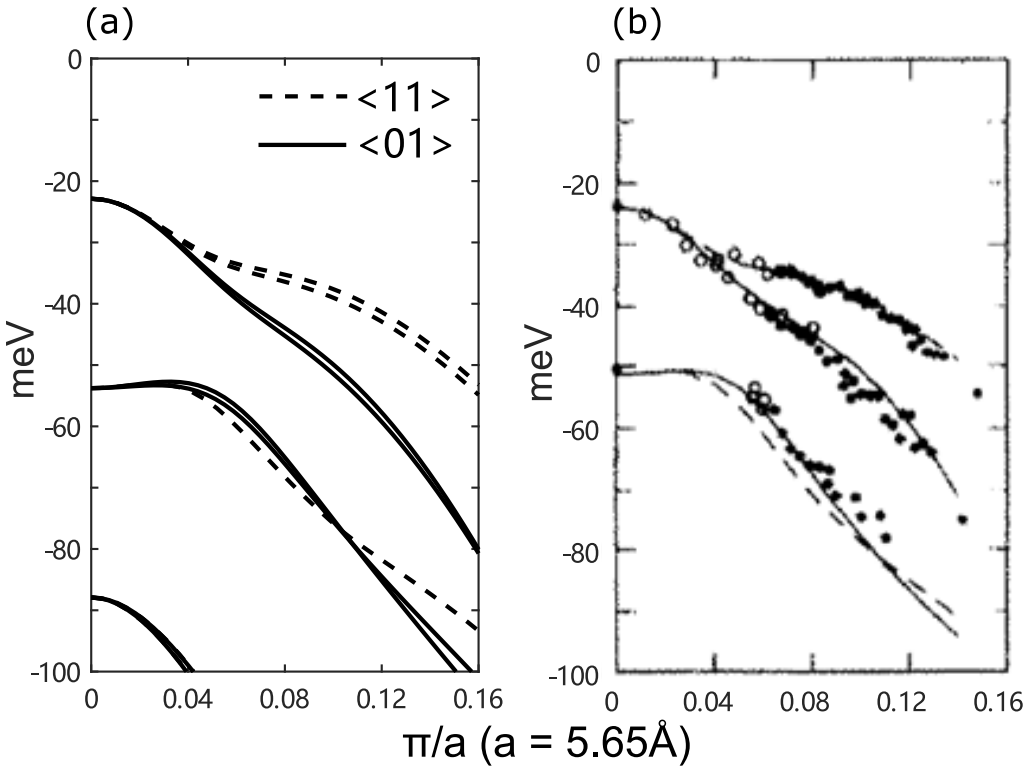

Figure 2.17: Band structure of $Al_{0.315}Ga_{0.685}As$/GaAs 5.1nm QW (a)calculated in this work, (b)experimentally measured in [79] (Note that figure is reproduced from [79]).

.

2.4 Summary

Effective mass method and **k.p** method are derived in detail, starting from basic quantum mechanics. The numerical methods to implement them are also described here. Since interband transitions mainly involve states around Γ valley, a good accuracy of the band structure around Γ valley is important, so the **k.p** method is used for this part. It includes the coupling between bands, the electronic properties calculated by **k.p** method are much more accurate

than if the effective mass method was used. as for the L and Δ valley, they do not take part in the inter-band transitions. They are included only to describe the carrier distribution over different valleys, so we used effective mass method for the band structure of L and Δ valleys, which is good enough for this purpose.

Chapter 3

Optical Properties

This chapter will discuss optical processes in semiconductors. Starting from interaction of semiconductor and electromagnetic radiation, with the optical transitions described by Fermi's golden rule, we describes the calculation of the absorption coefficient and gain. Furthermore, the numerical issues and methods used to improve the computation efficiency are also discussed.

3.1 Electron-Photon Interaction and Transition Rate

The Hamiltonian which describes the interaction between photons and electrons can be obtained by replacing the momentum operator in the time dependent Schrödinger equation (Eq.2.1) [80]:

$$\mathbf{p} \rightarrow \mathbf{p} - q_e \mathbf{A} \tag{3.1}$$

where $\mathbf{p} = (\hbar/i)\nabla$, q_e is the electron charge and $\mathbf{A}$ is the magnetic vector potential for the electromagnetic field. Then the Hamiltonian becomes:

$$\begin{aligned} H &= \frac{(\mathbf{p} - q_e\mathbf{A})^2}{2m_0} + V(\mathbf{r}) \\ &\simeq \frac{p^2}{2m_0} + V(\mathbf{r}) - \frac{q_e}{2m_0}(\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p}) \end{aligned} \tag{3.2}$$

Note that in the above equation, the quadratic term with A is dropped because $|q_e\mathbf{A}| \ll |\mathbf{p}|$ as the wavelength of light is usually much larger than the size of the lattice constant. Then we take the sum of the first and second term to be the unperturbed Hamiltonian H_0 and the last

term the perturbation from light, H'. The commutator of operator $\mathbf{p}$ and $\mathbf{A}$:

$$[\mathbf{p}, \mathbf{A}] = \mathbf{p} \cdot \mathbf{A} - \mathbf{A} \cdot \mathbf{p} = -i\hbar \nabla \cdot \mathbf{A} \tag{3.3}$$

so when $\nabla \mathbf{A} = 0$ (Coulomb gauge), we have $\mathbf{p} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{p}$. Then the time dependent perturbation term is:

$$H'(\mathbf{r}, t) = -\frac{q_e}{m_0} \mathbf{A}(\mathbf{r}, t) \cdot \mathbf{p} \tag{3.4}$$

Assuming $\mathbf{A}(\mathbf{r}, t)$ has the form [52]:

$$\mathbf{A}(\mathbf{r}, t) = \hat{\mathbf{e}} \frac{A_0}{2} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t} + \hat{\mathbf{e}} \frac{A_0}{2} e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega t} \tag{3.5}$$

we can rewrite the time dependent perturbation as the sum of two terms:

$$H'(\mathbf{r}, t) = H'(\mathbf{r}) e^{-i\omega t} + H'^{\dagger}(\mathbf{r}) e^{+i\omega t} \tag{3.6}$$

When the interaction time is long enough, according to Fermi's golden rule if an electron is initially in i-state, the probability it will transit to a final f-state per unit time is[52]:

$$W_{i \to f} = \frac{2\pi}{\hbar} |\langle f | H'(\mathbf{r}) | i \rangle|^2 \delta(E_f - E_i - \hbar\omega) + \frac{2\pi}{\hbar} |\langle f | H'^{\dagger}(\mathbf{r}) | i \rangle|^2 \delta(E_f - E_i + \hbar\omega) \tag{3.7}$$

When $E_f > E_i$ the electron transits upwards, only the first term exists, which corresponds to the photon absorption process. When $E_f < E_i$ only the second term exists, which corresponds to the photon emission and the electron transits downwards. Then we consider two states, a and b, and assume $E_b > E_a$. To estimate the total transition rate, the Fermi-Dirac distribution is introduced, denoted as f_a and f_b, which represent the probabilities that state a and b are occupied. Similarly, $(1 - f_a)$ and $(1 - f_b)$ are the probabilities that such states are empty. Next, the total upwards and downwards transition rates per unit volume are:

$$R_u = \frac{1}{V} \sum_{\mathbf{k}_a} \sum_{\mathbf{k}_b} \frac{2\pi}{\hbar} |\langle b | H'(\mathbf{r}) | a \rangle|^2 \delta(E_b - E_a - \hbar\omega) f_a (1 - f_b) \tag{3.8a}$$

$$R_d = \frac{1}{V} \sum_{\mathbf{k}_a} \sum_{\mathbf{k}_b} \frac{2\pi}{\hbar} |\langle a | H'^{\dagger}(\mathbf{r}) | b \rangle|^2 \delta(E_a - E_b + \hbar\omega) f_b (1 - f_a) \tag{3.8b}$$

Since $|\langle b|H'(\mathbf{r})|a\rangle| = |\langle a|H'^{\dagger}(\mathbf{r})|b\rangle|$, the net transition rate per unit volume from state-a to state-b is:

$$R_{abs} = R_u - R_d = \frac{1}{V}\sum_{\mathbf{k}_a}\sum_{\mathbf{k}_b}\frac{2\pi}{\hbar}|\langle b|H'(\mathbf{r})|a\rangle|^2\delta(E_b - E_a - \hbar\omega)(f_a - f_b) \tag{3.9}$$

3.2 Absorption and Gain Coefficient

3.2.1 General Expression

The absorption coefficient is equal the photon absorption rate per unit volume over the photon injection rate per unit area. The photon absorption rate is the net upward transition rate per unit volume introduced in Section 3.1. The photon injection rate is calculated by the energy flux density over the photon energy. The energy flux density is defined by the Poynting vector [52]:

$$\begin{aligned}\mathbf{S}(\mathbf{r},t) &= \mathbf{E}(\mathbf{r},t)\times\mathbf{H}(\mathbf{r},t)\\ &= -\frac{\partial}{\partial t}\mathbf{A}\times\left(\frac{1}{\mu}\nabla\times\mathbf{A}\right)\\ &= \mathbf{k}_{\mathrm{op}}\frac{\omega {A_0}^2}{\mu}\sin^2(\mathbf{k}_{\mathrm{op}}\cdot\mathbf{r}-\omega t)\end{aligned} \tag{3.10}$$

The energy flux density is the average value of Poynting vector:

$$P = |\langle\mathbf{S}(\mathbf{r},t)\rangle| = \frac{k_{\mathrm{op}}\omega {A_0}^2}{2\mu} = \frac{n_r c\epsilon_0\omega^2 {A_0}^2}{2} \tag{3.11}$$

where n_r is the refractive index, ϵ_0 is the vacuum permittivity, c is the speed of light. The matrix element of $H'(\mathbf{r},t)$ is calculated using dipole approximation which retains only the first term of the exponential $\exp(i\mathbf{k}\cdot\mathbf{r}) \simeq 1$. Moreover, the k-selection rule is used. Finally the momentum matrix element H'_{ba} reads:

$$H'_{ba} = \left\langle b\middle|-\frac{q_e A_0}{2m_0}\exp(i\mathbf{k}\cdot\mathbf{r})\hat{e}\cdot\mathbf{p}\middle|a\right\rangle = -\frac{q_e A_0}{2m_0}\hat{e}\cdot\mathbf{p}_{ba}\delta_{\mathbf{k}_a,\mathbf{k}_b} \tag{3.12}$$

Then, using Eq. 3.9, Eq. 3.11 and Eq. 3.12, the absorption coefficient can be written as:

$$\alpha(\hbar\omega) = \frac{R_{abs}}{P/\hbar\omega} = \frac{2\hbar\omega}{n_r c\epsilon_0\omega^2 {A_0}^2}R_{abs}$$

$$= \frac{\pi {q_e}^2}{n_r c\epsilon_0 {m_0}^2 \omega} \frac{1}{V} \sum_{\mathbf{k}} |\hat{e} \cdot \mathbf{p}_{ba}|^2 \delta(E_b - E_a - \hbar\omega)(f_a - f_b) \tag{3.13}$$

where $|\hat{e}\cdot\mathbf{p}_{ba}|$ is the momentum matrix element, which can be calculated from the wave functions obtained by **k.p** method. Also, we can see that A_0^2 in the photon absorption rate and in the power intensity are cancelled, which indicate that the absorption coefficient is independent on the optical intensity. For quantum well structure, the sum over the $\mathbf{k}$ is replaced by the summation of the s,t-subbands and the in-plane wave vectors k_x and k_y. Since the electron is free in xy direction, the wave vectors k_x and k_y are continuous, and the summation should be replaced by integration:

$$\frac{1}{V}\sum_{\mathbf{k}} \rightarrow \frac{1}{V}\sum_{s,t}\sum_{k_x}\sum_{k_y} = \frac{1}{V}\sum_{s,t}\int \frac{dk_x}{(2\pi/L_x)}\int \frac{dk_y}{(2\pi/L_y)} = \frac{1}{L_z}\sum_{s,t}\iint \frac{dk_x dk_y}{(2\pi)^2} \tag{3.14}$$

Then the explicit form of absorption coefficient of quantum well structure is:

$$\alpha(\hbar\omega) = \frac{\pi {q_e}^2}{n_r c\epsilon_0 {m_0}^2 \omega L_z} \sum_{s,t} \iint \frac{dk_x dk_y}{(2\pi)^2} |\hat{e} \cdot \mathbf{p}_{st}|^2 \delta(E_s - E_t - \hbar\omega)(f_t - f_s) \tag{3.15}$$

in the above equation, both E_s, E_t depend on the in-plane wave vector. The dependences $E_s(k_x, k_y)$ and $E_t(k_x, k_y)$ are calculated using **k.p** method, which was discussed in Chap. 2. The Fermi-Dirac distribution f_s and f_t depends not only on the wave vector but also on temperature. The Fermi-Dirac function reads:

$$f_s(k_x, k_y) = \left[1 + \exp\left(\frac{E_s(k_x, k_y) - E_{fs}}{k_B T}\right)\right]^{-1} \tag{3.16a}$$

$$f_t(k_x, k_y) = \left[1 + \exp\left(\frac{E_t(k_x, k_y) - E_{ft}}{k_B T}\right)\right]^{-1} \tag{3.16b}$$

For inter-band transitions, E_s is the energy of conduction subbands, E_t is the energy of valence subbands. $E_{fs} = E_{fc}$ is the quasi-Fermi level for conduction band and $E_{ft} = E_{fv}$ is the quasi-Fermi level for valence band. Things are a little different for calculation of inter-valence-band-absorption (IVBA), as E_s is the energy of LH and HH subbands, E_t is the energy of SO subbands. E_{fs} and E_{ft} now are both for the valence band quasi-Fermi level E_{fv}.

The quasi-Fermi levels for conduction band and valence band are generally different. Under optical or electrical pumping, both electrons and holes exist. The population of conduction band and valence band have to be described by two independent Fermi levels. The gain will

exist when the more carriers are in higher state, which is called 'Population inversion'. For the inter-band gain, in Eq. 3.15, the criteria of the population inversion is described by Fermi-Dirac inversion factor $f_v - f_c$. The absorption coefficient will become negative when $f_v - f_c < 0$, the negative part of the absorption spectrum is thus the gain. Hence the gain coefficient is just the negative of the absorption:

$$g(\hbar\omega) = -\alpha(\hbar\omega) \tag{3.17}$$

Therefore, without considering any other loss mechanisms, in a quantum well, the minimum requirement to have gain is $f_{v1} - f_{c1} < 0$. $c1$ and $v1$ correspond to the first conduction subband and the first valence subband. This leads to the condition:

$$\Delta E_f = E_{fc} - E_{fv} > E_{c1} - Ev1 = \Delta E_1 \tag{3.18}$$

As shown in Fig. 3.1, in this situation the gain comes from carriers in the first conduction subband and first valence subband. The largest gain is observed in the interval $(\Delta E_1, \Delta E_f)$. although a small gain also appears below ΔE_1 because of broadening, but this part is usually ignorable because it is even smaller than some of the loss mechanisms. The peak of the gain can be found near ΔE_1. With even larger difference of the quasi-Fermi levels, higher subbands can also contribute to the gain and the gain peak could appear at a higher energy.

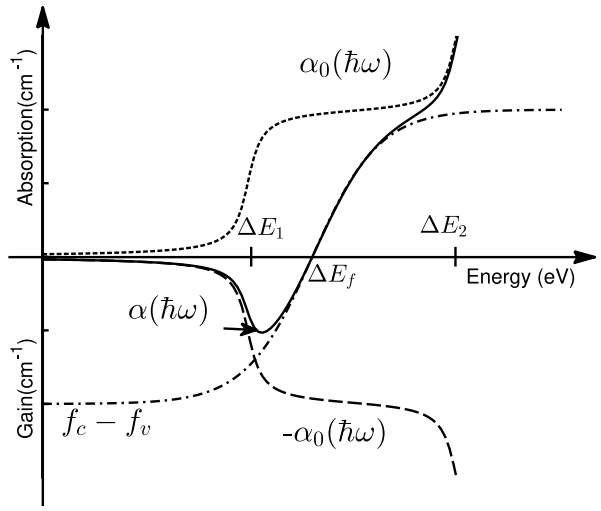

Figure 3.1: Schematic diagram for the absorption spectrum $\alpha(\hbar\omega)$. Gain can be observed between ΔE_1 and ΔE_f

.

The absorption spectrum is a step like function, corresponding to the integral of the delta function. However, in reality broadening exists due to many sources such as the carriers lifetime, scattering, and the fluctuations of the structure parameters. For these reasons, the delta function in Eq. 3.15 may be replaced by Lorentzian line shape function:

$$\begin{aligned} &\delta(E_s - E_t - \hbar\omega) \rightarrow L(E_s - E_t - \hbar\omega, \Gamma_w) \\ =&\frac{\frac{\Gamma_w}{2\pi}}{(E_s - E_t - \hbar\omega)^2 + \left(\frac{\Gamma_w}{2}\right)^2} \end{aligned} \tag{3.19}$$

where Γ_w is the full width at half maximum (FWHM) of the Lorentzian line shape. It will become delta function when Γ_w approaches zero.

3.2.2 Momentum Matrix Element in a Quantum Well

Instead of using the average value to estimate the momentum matrix element, it is obtained by exact calculation, using the wave functions and energy dispersion calculated from the **k.p** method. The momentum matrix element between s and t states is given as [70]:

$$M_{st}^{\hat{e}} = |\hat{e} \cdot \mathbf{p}_{st}| = \langle \Psi_{s\mathbf{k}} | \hat{e} \cdot \mathbf{p} | \Psi_{t\mathbf{k}} \rangle \tag{3.20}$$

Using the definition of total wave function in Eq. 2.101, the above equation is expanded to:

$$M_{st}^{\hat{e}} = \sum_{mn} \langle F_{sm\mathbf{k}}(z) | \hat{e} \cdot \mathbf{p} | F_{tn\mathbf{k}}(z) \rangle \, \delta_{mn} + \sum_{mn} \langle F_{sm\mathbf{k}}(z) | F_{tn\mathbf{k}}(z) \rangle \, \langle u_{m0} | \hat{e} \cdot \mathbf{p} | u_{n0} \rangle \tag{3.21}$$

in which n and m are the basis states. In detail, 1 and 2 correspond to the conduction band, 3 and 6 are for the light hole, 4 and 5 are the heavy hole and 7 and 8 are split-off bands. The momentum matrix element can further be approximated by ignoring the first term, because $\hbar k$ is much smaller than $(m_0/\hbar)P_0$. In the second term, the non-zero terms are those with appropriate symmetry of the basis. Only the following terms are none-zero:

$$\langle S | p_x | X \rangle = \langle S | p_y | Y \rangle = \langle S | p_z | Z \rangle = \frac{m_0 P_0}{-i\hbar} \tag{3.22}$$

Then the explicit expressions of the polarization dependent momentum matrix elements are:

For the transverse magnetic (TM) polarization $\hat{e} = z$:

$$\begin{aligned}\langle \Psi_{s\mathbf{k}}|p_z|\Psi_{t\mathbf{k}}\rangle &\simeq \sum_{mn} \langle F_{sm\mathbf{k}}(z)|F_{tn\mathbf{k}}(z)\rangle \langle u_{m0}|p_z|u_{n0}\rangle \\ &= -\frac{m_0}{\hbar}\Bigg[+\sqrt{\frac{2}{3}}\left(\langle F_{s1}|P_0(z)|F_{t6}\rangle + \langle F_{s6}|P_0(z)|F_{t1}\rangle\right) \\ &\quad + \sqrt{\frac{1}{3}}\left(\langle F_{s1}|P_0(z)|F_{t7}\rangle + \langle F_{s7}|P_0(z)|F_{t1}\rangle\right) \\ &\quad + \sqrt{\frac{2}{3}}\left(\langle F_{s2}|P_0(z)|F_{t3}\rangle + \langle F_{s3}|P_0(z)|F_{t2}\rangle\right) \\ &\quad - \sqrt{\frac{1}{3}}\left(\langle F_{s2}|P_0(z)|F_{t8}\rangle + \langle F_{s8}|P_0(z)|F_{t2}\rangle\right)\Bigg]_{\mathbf{k}}\end{aligned} \tag{3.23}$$

For the transverse electric (TE) polarization $\hat{e} = x, y$:

$$\begin{aligned}\langle \Psi_{s\mathbf{k}}|p_x|\Psi_{t\mathbf{k}}\rangle &\simeq \sum_{mn} \langle F_{sm\mathbf{k}}(z)|F_{tn\mathbf{k}}(z)\rangle \langle u_{m0}|p_x|u_{n0}\rangle \\ &= -\frac{m_0}{\hbar}\Bigg[-\sqrt{\frac{1}{6}}\left(\langle F_{s1}|P_0(z)|F_{t3}\rangle + \langle F_{s3}|P_0(z)|F_{t1}\rangle\right) \\ &\quad - \sqrt{\frac{1}{2}}\left(\langle F_{s1}|P_0(z)|F_{t5}\rangle + \langle F_{s5}|P_0(z)|F_{t1}\rangle\right) \\ &\quad - \sqrt{\frac{1}{3}}\left(\langle F_{s1}|P_0(z)|F_{t8}\rangle + \langle F_{s8}|P_0(z)|F_{t1}\rangle\right) \\ &\quad + \sqrt{\frac{1}{2}}\left(\langle F_{s2}|P_0(z)|F_{t4}\rangle + \langle F_{s4}|P_0(z)|F_{t2}\rangle\right) \\ &\quad + \sqrt{\frac{1}{6}}\left(\langle F_{s2}|P_0(z)|F_{t6}\rangle + \langle F_{s6}|P_0(z)|F_{t2}\rangle\right) \\ &\quad - \sqrt{\frac{1}{3}}\left(\langle F_{s2}|P_0(z)|F_{t7}\rangle + \langle F_{s7}|P_0(z)|F_{t2}\rangle\right)\Bigg]_{\mathbf{k}}\end{aligned} \tag{3.24}$$

An example of the calculated momentum matrix element is shown in Fig. 3.2. The structure is $Al_{0.3}Ga_{0.7}As/GaAs$ 8nm quantum well, Fig. 3.3 is the band structure of the quantum well. There are a few things worth to be discussed here. The MME of C1-HH1 (TM) is zero when $k_x = 0$. This is because the basis of heavy holes does not have Z-component so it does not contribute to the TM mode. But the momentum matrix elements starts to increase because of stronger coupling with light holes when k_x becomes larger, the light hole part of the wave function starts to appear in HH1 state. The ratio of matrix element C1-LH1 (TM) and C1-LH1 (TE) is four when $k_x = 0$ if we ignore the split off band in the wave function. It is not four in Fig. 3.2 because the k_z is quantized and even when the in-plane vector k_x, k_y is zero, the LH

and SO bands are still coupled. Then for LH1 state, the envelope wave function of SO basis is not exactly zero. Moreover, the MME of C1-LH1 (TM) decreases with k_x, indicating that the LH component decreases when k_x becomes larger.

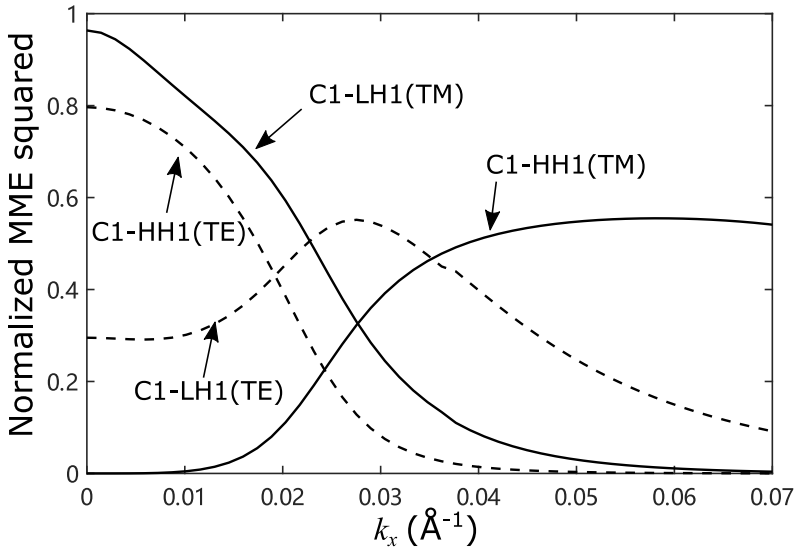

Figure 3.2: Momentum matrix elements in $Al_{0.3}Ga_{0.7}As/GaAs$ 8nm quantum well calculated from the wave function and band structure using **k.p** method.

.

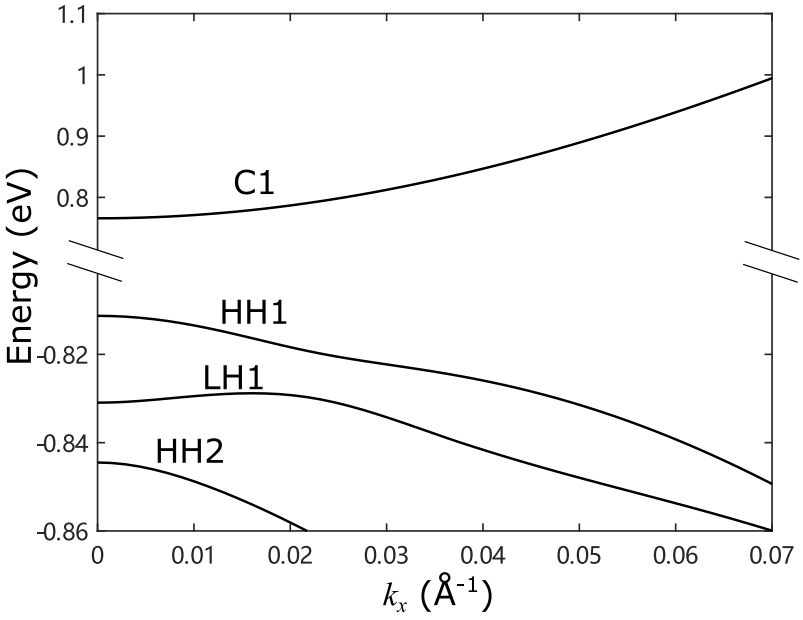

Figure 3.3: The band structure ([01] direction) of $Al_{0.3}Ga_{0.7}As/GaAs$ 8nm quantum well.

.

3.3 Simplex Method of Brillouin Zone Integration

In Eq. 3.15, if the delta function is replaced by Lorentzian line shape function, the integration of Lorentzian is conducted throughout the full k-space. This process is time consuming and with low efficiency. The tail of the Lorentzian is nearly zero, which does not contribute to the accuracy of the over all result. To reduce the calculation time we use the simplex method to estimate the integral with delta function in the Brillouin zone [81], with the general form of the integration in two-dimensional Brillouin zone:

$$I(E) = \iint f(k_x, k_y)\delta(E - E'(k))dk_x dk_y \tag{3.25}$$

Consider the equation of the absorption coefficient Eq. 3.15, the $f(k_x, k_y)$ in the above equation is just the momentum matrix element times the Fermi-Dirac inversion factor. $E'(k)$ is the energy difference between the conduction subband and valence subband in 2D Brillouin zone. Assume an energy surface in the Brillouin zone as shown in Fig. 3.4.

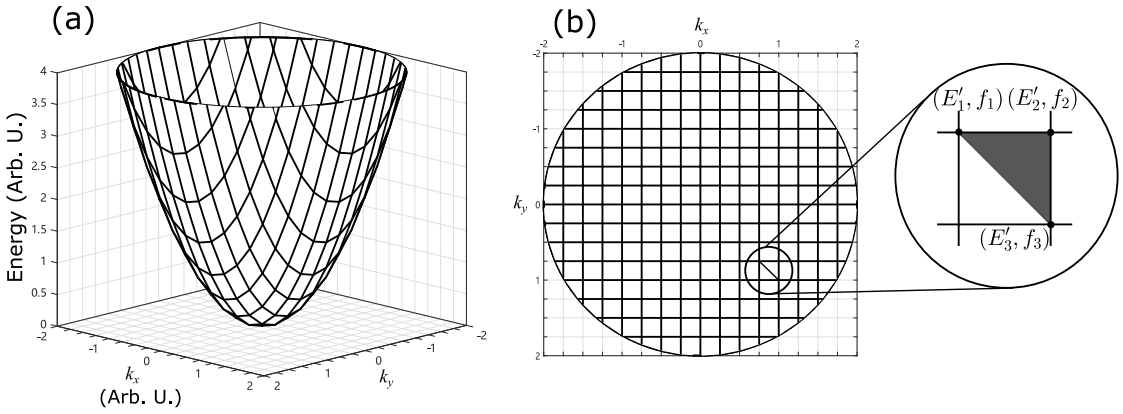

Figure 3.4: Example of an energy surface in a 2D Brillouin zone (a)3D view; (b)Top view, one of the elementary squares is zoomed in, Energy E_i' and function f_i are the value at the vertexes of the triangle.

.

Computationally, the continuous surface is discretized to the grid points. Each grid point in the Brillouin zone represents an energy $E(k)$ and a function $f(k)$. The nearest four grid points form a square in the k_x and k_y plane. Such a square can be further divided into two triangles. Thus the integration in the full Brillouin zone can be performed as the integration over all the sub-triangles. In Fig. 3.4b, we assume that the energy at the three points follows the sequence $E_1' < E_2' < E_3'$. The integration over a single triangle surface in Fig. 3.4b can be approximated

by the following expression:

$$I_\Delta(E) = \begin{cases} 0, & E < E_1', E > E_3' \\ S_\Delta \left[\dfrac{2(E-E_1')f_1}{(E_2'-E_1')(E_3'-E_2')} + \dfrac{(E-E_1')^2}{(E_2'-E_1')(E_3'-E_2')} \left(\dfrac{f_2-f_1}{E_2'-E_1'} + \dfrac{f_3-f_1}{E_3'-E_1'} \right) \right], & E_1' < E < E_2' \\ S_\Delta \left[\dfrac{2(E-E_3')f_3}{(E_1'-E_3')(E_2'-E_3')} + \dfrac{(E-E_3')^2}{(E_1'-E_3')(E_2'-E_3')} \left(\dfrac{f_1-f_3}{E_1'-E_3'} + \dfrac{f_2-f_3}{E_2'-E_3'} \right) \right], & E_2' < E < E_3' \end{cases} \tag{3.26}$$

in which $I_\Delta(E)$ is the integration in a single triangle, and $S_\Delta = \Delta k^2/2$ is the area of the triangle. The absorption coefficient calculated by Eq. 3.26 corresponds to using the delta function. The shape of the absorption spectrum is a step like curve. To take broadening into account, the convolution over Lorentzian function is performed in the energy domain. The convolution equation reads:

$$h(x) = \int f(x')L(x-x')dx' \tag{3.27}$$

in which, $f(x)$ is the original function, $h(x)$ is the convoluted function.

An example is shown in Fig. 3.5. The absorption of a $Al_{0.3}Ga_{0.7}As/GaAs$ 5nm QW is calculated by assuming the Fermi-Dirac inversion factor $f_v - f_c = 1$. The number of points in the z-direction is 100. The grid in k-space is 80×80. The results are plotted at 80 points in the photon energy axis. In Fig. 3.5, the dashed line is calculated by the simplex method. The solid line is the curve obtained after making convolution with Lorentzian. The dotted line is calculated by doing the integration in the full k-space. Doing the integration in the whole k-space is time consuming. For such calculation the elapsed time on a single core Intel i7 processor with 3.7GHz is 431 seconds. However by using the triangle simplex method, the elapsed time is only 54 seconds including the convolution. The results of both are very similar, the absorption curve calculated by triangle simplex method after Lorentzian convolution is slightly smaller than the full k-space integration. This is because the convolution is performed in a limited range of energy, which is smaller than the energy span to the borders of the chosen k-space, and Lorentzian line-shape function has a fairly long tail. The difference can be smaller for other line-shapes like Gaussian, because it drops to zero faster than the Lorentzian and the influence from its tails is smaller.

Overall, in this example we can see that the difference between results of these two approaches are quite acceptable. The simplex method dramatically improves the calculation efficiency, about 8 times, so this method is used in our program to calculate the absorption and gain.

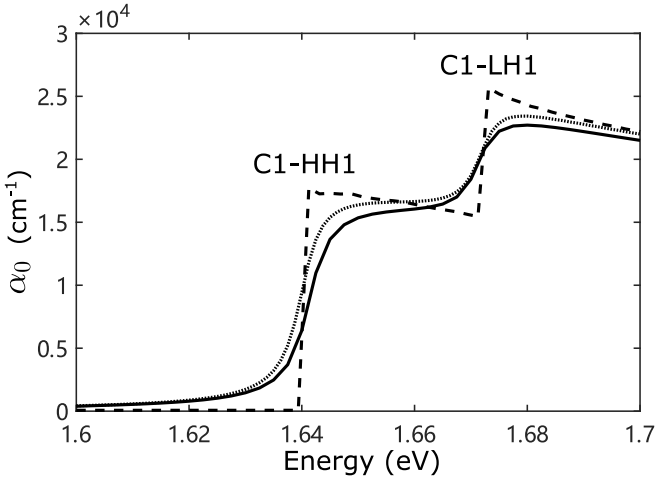

Figure 3.5: Absorption coefficient calculated by various methods. The dotted line is calculated using full-k-space integration over Lorentzian function; The dashed line is calculated using triangle simplex method; The solid line is obtained by Lorentzian convolution with the result from triangle simplex method (Dashed line)

.

3.4 Summary

In summary, this chapter has discussed the optical processes in semiconductors. The way to calculate the momentum matrix element and absorption coefficient in a quantum well was discussed in detail. The simplex method is used to reduce the computation time. In reality the overall optical properties depend on various parameters. In a quantum well laser, the gain spectrum depends on the quasi-Fermi levels of the conduction band and valence band. The quasi-Fermi levels depend on temperature, injection carrier density and doping density. The free carrier absorption should also be considered for a laser. For a modulator, the exciton effect is included, the indirect absorption is described using a bulk model, and the broadening function is also different from the Lorentzian. These additional calculations for the absorption(gain) coefficients are discussed in detail in Chapters 4 and 5.

Chapter 4

Design Optimization of Tensile Strained SiGeSn-GeSn Quantum well Laser

4.1 Introduction

As the last missing bit to realize large-scale optoelectonics integration, large effort is still continuously put into this area to find an efficient on-chip light source. Since the intermediate approach, to integrate III-V laser on Si is facing problems related to the fabrication process and high cost [16], group IV materials are now the most promising candidates to enable achieving this vision [82]. The major problem to be solved in this approach is that group IV semiconductors usually have indirect bandgap. Fortunately, with the unique band structure of Ge, in which L-valley is not far from the Γ valley, the indirect bandgap structure can be transformed into a direct bandgap structure.

Two approaches for the direct gap transformation have been proposed to make Ge an efficient light emitter. The first approach is by using GeSn as a lasing material. The first working GeSn laser was demonstrated in 2015 [16], although lasing was only observed upto 90K, but it was a big step indicating GeSn is a promising candidate for light emitting material. Following this, another team has proposed that with higher Sn content the working temperature could be further increased. The best data so far is 230K by using 16% Sn in the active layer [17].

However, all these lasers were optically pumped. The first electrically pumped laser with 11% Sn in the active region, operating up to 100K, was realised in 2020 [20], which further boosted the progress of on-chip Group IV systems. Another approach to achieve lasing in Ge is to introduce tensile strain. In [7, 83], lasing in Ge is realized in tensile strained Ge assisted by n-type doping. The tensile strain is introduced by thermal expansion mismatch which has a value of just 0.25%. Thus it adds some directness of the bandgap (i.e. reduces the difference between the bandgap at L and Γ valley), but does not make the material with a true direct bandgap. So, to compensate for the remaining difference of direct and indirect bandgap, n-type doping is used to achieve population inversion in Γ valley. To have a direct bandgap the tensile strain should be at least 1.7% [64, 65]. Challenges still exist in both approaches. As for the former, it is very hard to manufacture high Sn content material, a high defects density will appear due to the large lattice mismatch between Ge and Sn. The low solubility of Sn is another reason which inhibits the high Sn content. These defects act as electron and hole recombination and trapping centers and prevent higher temperature operation of the laser. As for the second approach, the strain is usually introduced by pseudomorphic growth. As the strain energy is stored in mismatched layers and is proportional to the layer thickness, the largest strain for a structure with some thickness is limited. Additionally, introducing relatively large strain is still an open topic to be resolved. A way out could be to combine both methods, i.e. to apply tensile strain on GeSn. In this way the Sn content and the tensile strain needed for indirect-direct transition are both reduced, and material with a large directness can be achieved with moderate Sn content and tensile strain. This trend has been proved in [19], in which tensile strain is introduced to GeSn by SiN stressor, and a ultra-low threshold was realized with only 5.4% Sn and around 1.4% strain in the active layer.

Apart from the material itself, experience from the traditional Group III-V system has indicated that MQW structures can bring improvements to the performance of the laser. In Group IV, it was found that SiGeSn is a good choice for the barrier material of QW because, compared with Ge or GeSn with lower Sn content as a barrier, SiGeSn gives a larger band offsets and provides more room for strain engineering [15]. SiGeSn/ GeSn QWs were characterized in previous research, indicating it was a promising structure for light emitters, both for light emitting diode (LED) and laser. E.g. it enabled a lower threshold for MQW SiGeSn/GeSn laser compared to bulk GeSn[84].

Here we discuss the use of SiGeSn/GeSn multiple QWs for the active region of the laser. Various measurements and calculations have been made on the band alignment, well width and optimum choice of the Sn content and strain in the well for a SiGeSn/GeSn QW with some well width to have the largest gain for a particular photon transition energy. In this work, the bandgap and wave functions are calculated using eight-band **k.p** method. Indirect valleys (L and Δ) are included in calculation using the effective mass method in order to describe the carrier distribution in these valleys. Moreover, inter-valence-band-absorption is also included. For larger Sn content and tensile strain the direct bandgap of GeSn becomes much smaller than that of pure Ge ($\sim$ 0.89eV), and when the bandgap is about $\sim$0.4eV the transition energy of QW becomes very close to the split off energy, which makes the IVBA a major loss mechanism. This, however, is rarely discussed in other works on Group IV QW modelling.

Next, a section is added to discuss the band alignment in the QW structures. For tensile strained SiGeSn/GeSn QW, a QW can usually be expected in the conduction band for Γ valley, but sometimes the type-I potential well in LH band cannot be found, and a type-II band alignment is found for SiGeSn/GeSn heterostructure. This section will give some reference information about the choice of the barrier material in order to obtain a good carrier confinement.

To give a brief outline of this chapter, combined with the description already given in theoretical Chapter 2 and 3, some additional details of the model will be introduced. The parameters used are given, including those for pure elements, the bowing parameters and quadratic interpolation parameters, followed by evaluation of band alignment and well width. These will give guidelines about choosing the structure parameters i.e. the barrier material and well width, Sn content and strain in the well. The practical parameters are then estimated for a wave guide structure.

4.2 Computational model

4.2.1 Band Structure Calculation

The band structure near the Γ point is calculated by solving the eigen problem of the Hamiltonian matrix, Eq. 2.83. For L and Δ valley, the effective mass method is used, Eq. 2.4. Both calculations regarding the solution of Schrödinger equation in 2D heterostructure are made by the matrix finite difference method. The details of these methods have already been discussed in Chapter 2. To reduce the computation time, the MQW structure is calculated by assuming

that the QWs are identical and separated by identical barriers. Then it can be equivalent to considering a single QW. Although **k.p** method can also be used to calculate the whole MQW structure, this would require much more discrete points to retain the accuracy of the results, which would need more computational power for such calculation. Moreover, the (almost) degenerate bands in these identical QWs will participate in the calculation of absorption spectrum, this will further increase the computation time of the whole process. Hence, for realistic computational efficiency, it is reasonable to use SQW model for the whole structure. When the barrier width of the MQW is sufficiently large the wave function will not tunnel through the barrier, and each individual QW is almost independent. The calculation results also show negligible difference in band structures and absorption spectra for the MQW structure and if considering an individual QW. The sparse matrix of the eight-band **k.p** model is solved by LAPACK 'zheevd.f' subroutine, since the Hermitian matrix elements are complex numbers [50]. For effective mass method the subroutine 'dstevd.f' is used for real symmetric matrix. Then, with 100 discrete points, the **k.p** matrix will generate 800 eigenvalues, and the effective mass matrix will give 100 eigenvalues. So, if the in-plane k-vector space has just 20 points in both k_x and k_y directions, and if we use symmetry conditions in the k-space to reduce the unnecessary calculations, the number of points in k-space is ∼200. So the total number of data cells in one calculation is about 1.6×10^5. And these data cells contain the information on the wavefunction with eight components. Thus, from the programming point of view, it is unwise to store all these data when doing calculations. Bands necessary for calculation need to be selected, and those with estimated small contribution to the final result should be ommited. Further details will be discussed in the next section.

4.2.2 Total absorption spectrum and band selection rules

The total absorption consists of three parts, the inter-band absorption (IBA) or inter-band gain (IBG), free carrier absorption (FCA), and inter-valence-band absorption (IVBA). To calculate the inter-band absorption or gain spectrum under a given carrier density, we need to determine the quasi-Fermi level. The probability that a state with energy E is occupied by electron is:

$$f^n = \frac{1}{\exp[(E - E_f)/k_B T] + 1} \tag{4.1}$$

On the other hand, the probability that a state is occupied by a hole is:

$$f^p = 1 - f^n = \frac{1}{\exp[(E_f - E)/k_B T] + 1} \tag{4.2}$$

Thus the number of carriers per unit volume is calculated by summing the number of states times the occupation probability. The 3D carrier density of elections and holes is:

$$N_{3\mathrm{D}} = \frac{1}{V}\sum_{\mathbf{k}} f_c^n(E(\mathbf{k})) = \frac{1}{L_z}\sum_{i}^{c}\iint \frac{dk_x dk_y}{(2\pi)^2} f_c^n(E_i(\mathbf{k}_{||})) \tag{4.3a}$$

$$P_{3\mathrm{D}} = \frac{1}{L_z}\sum_{j}^{v}\iint \frac{dk_x dk_y}{(2\pi)^2} f_v^p(E_j(\mathbf{k}_{||})) \tag{4.3b}$$

If L-valley is also accounted, the relationship between 2D carrier density and quasi-Fermi level for conduction band and valence band reads:

$$\begin{aligned} N_{2D} =& \sum_{i}^{n_\Gamma}\iint \frac{dk_x dk_y}{(2\pi)^2}\left[1+\exp\left(\frac{E_i^\Gamma(\mathbf{k}_\parallel) - E_{fc}}{k_B T}\right)\right]^{-1} \\ &+ 4\sum_{i}^{n_L} 2\frac{m_d^L k_B T}{2\pi\hbar^2}\ln\left[1+\exp\left(\frac{E_{fc} - E_i^L}{k_B T}\right)\right] \end{aligned} \tag{4.4a}$$

$$P_{2D} = \sum_{i}^{n_v}\iint \frac{dk_x dk_y}{(2\pi)^2}\left[1+\exp\left(\frac{E_{fv} - E_i^v(\mathbf{k}_\parallel)}{k_B T}\right)\right]^{-1} \tag{4.4b}$$

In the above equations, the first terms represent the carrier density in the Γ valley. The second term in Eq. 4.4a refers to the carrier concentration in the L-valley. The factor of 4 is the degeneracy of the L-valley. m_d^L is the density of state effective mass of L valley, it is calculated according to Table 2.1 and data in Table 4.1. Note that, since L valley states are calculated by single band effective mass method, the second term in Eq. 4.4a is obtained by assuming the parabolic energy dispersion, and the sum over spin is replaced by the factor of 2, thus it is not included in the summation. However, in the first term, as the band structure is calculated using eight-band **k.p** method, the degenerate bands with different spins are considered as two independent bands, and they are included in the summation instead of using the factor of 2.

The carrier density is determined by injection current and doping. Here we assume that the injection current gives the same carrier densities of electrons and holes, and the doping is accounted as additional carrier density of electrons (n-type) or holes (p-type). The carrier density is considered as an input parameter in the calculation. Then, with a given carrier

density, the quasi-Fermi level can be calculated by root finding methods.

After the quasi-Fermi level is determined. The inter-band absorption(gain) spectra is then calculated using Eq. 3.15. For the free carrier absorption, we assume that Ge-rich material will have a similar FCA as pure Ge, so the expression for bulk Ge at room temperature is used:

$$\alpha_{FCA}(\lambda) = C_n N_{3D} \lambda^a + C_p P_{3D} \lambda^b \tag{4.5}$$

in which C_n, C_p, a and b are constants, and λ is in the units of nm. N_{3D} and P_{3D} are 3D carrier density in the units of cm^{-3}, this is taken as the 2D carrier density divided by the structure thickness L_z. As most of the carriers are confined in the QW, L_z can be the thickness of the QW. The values of C_n, C_p, a and b are given in [83]. The constants a and b are $\sim$2, $a = 2.25$ and $b = 2.43$. For C_n and C_p, when carrier density is 10^{19} to 10^{20} cm^{-3}, $C_n = -3.4 \times 10^{-25}$ and $C_p = -3.2 \times 10^{-25}$. The 2D carrier density used in this work is of the order of 10^{12} cm^{-2}, and the well width is of the order of 10 nm (i.e. 10^6 cm). The 3D carrier density is then of the order of 10^{18} cm^{-3}. This value is one order of magnitude smaller, thus C_n was suggested to be halved for lower carrier densities[83]. Then in Eq. 4.5, $C_n = -1.7 \times 10^{-25}$.

The inter-valence-band-absorption can come from transitions between HH and LH, HH and SO, LH and SO bands. The energy between HH and LH is too small to be considered, so only the transitions between HH and SO, or LH and SO bands are calculated. For pure Ge, the split-off energy is just $\sim$0.29 eV, which is much smaller than the bandgap, and the IVBA then barely has any influence on the inter-band-gain. Sn has larger split off energy than Ge. By mixing with Sn, the split off energy increases with Sn concentration. Tensile strain will also make the difference between the top valence band, here LH, and the split off band become larger. These two effects, on the other hand, will decrease the bandgap simultaneously. And when the difference of LH and SO is close to the optical transition energy, the IVBA will become the major loss mechanism. For inter-band-gain in QW, since the density of states is a step like function, the gain will increase steeply to its peak and then decrease. The peak of the gain will usually be found near the transition energy between the first quantized states of conduction band and valence band, and depends slightly on the linewidth. As for inter-valence-band-absorption, it will not increase steeply at the transition energy between LH band and SO band. This is because the IVBA occurs between the s-like and p-like components of these bands. The valence band has no s-component at the center of the Brillouin zone. The s-like component only appears when

$k \neq 0$ and coupling becomes stronger. Also, under tensile strain the SO band will generally have a smaller in-plane effective mass than LH band. The absorption peak appears at $k' \neq 0$, and has a higher energy than that at the Γ point, as a combined result of the coupling and Fermi-Dirac distribution. The most difficult part in calculating the IVBA for QW structures using the current model comes from the numerical nature of the combination of **k.p** and MFDM, because for a finite total structure width, when the quantized states are calculated for each different type of bands, the program will also generate quasi-free bands. These bands are called 'quasi' because they are still subband of the infinitely deep QW with finite total structure width. In the valence band there are three different types of bands, LH, HH and SO. These quasi-free bands of LH and HH with lower energy will overlap with the quantized SO bands. This can also happen to a free LH state and a HH bound state when the tensile strain is large, because the LH will have much higher energy than HH, and even the barrier for LH will be higher than the well layer potential of HH.

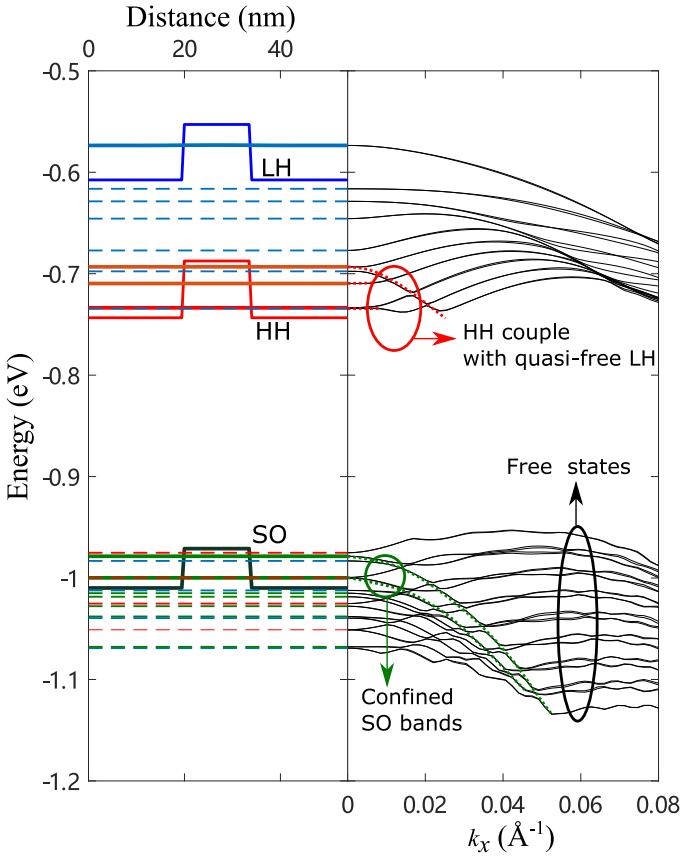

Figure 4.1: Band structure of 14nm $Si_{0.15}GeSn_{0.105}/GeSn_{0.065}$ QW with 1.14% tensile strain in the well, the bound states are marked as solid lines and the quasi free states are dashed lines. The LH, HH and SO bands are represented by blue, red and green lines.

.

Fig 4.1 shows an example of such a scenario. Because of the coupling between these bands, the band structure is also 'entangled' with each other and not continuous anymore. Because the eigenvalues given by the LAPACK subroutine are in the ascending order, when an intersection happens between two bands, the subbands indices will exchange. Each subband will only contain a small part of the confined real band. Therefore, many of these subbands are needed to form a complete SO band in the whole selected k-space. Storing all of them in the random access memory (RAM) is impractical for personal computer (PC). It could be realized in a high performing computing (HPC) cluster, but still with low efficiency. Thus, to improve the efficiency, these bands are filtered, and only those necessary for the calculation are kept. The process of selecting the SO bands is explained next.

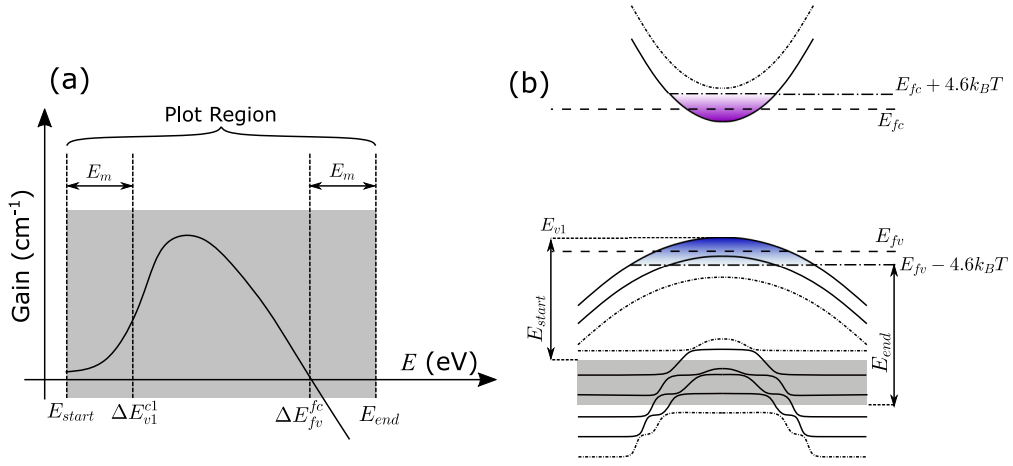

Figure 4.2: The schematic diagram of (a)The region of interest to plot the gain spectrum (b) The band structure calculated by FDM and graphical explanation of how the necessary SO bands are selected to plot Fig. 4.2a.

.

First consider the gain spectrum from Fig. 4.2a, as analysed in Chapter 3. The gain can be observed exactly between the transition energy ΔE_{v1}^{c1} and the difference between the quasi-Fermi levels ΔE_{fv}^{fc} without finite linewidth effects. With some broadening, the gain can also be found at energy lower than ΔE_{v1}^{c1} as the tail of the line shape function. Thus, with some margin E_m, the energy interval related to the gain is (E_{start}, E_{end}), in which $E_{start} = \Delta E_{v1}^{c1} - E_m, E_{end} = \Delta E_{fv}^{fc} + E_m$. For this reason, the upper boundary energy interval of interest in SO band will be $E_{v1} - E_{start}$. Next we consider the quasi-Fermi level. According to the definition of Fermi-Dirac distribution, the probability that a state with energy lower than $E_{fv} - 4.6k_BT$ is occupied by

hole is less than 1%, So it is very unlikely that transitions between this state with SO band can happen. Thus, we can use this as the lower boundary, which is $E_{fv} - 4.6k_BT - E_{end}$. Finally, as can be seen in Fig. 4.2b, only the bands in the gray and coloured areas will be included in the calculation. The quasi-Fermi level is also used to select the bands for inter-band gain calculation. For the valence band those subbands whose energy at Γ point is higher than $E_{fv} - 4.6k_BT$ are stored. Similarly, for the conduction band, the energies at Γ point lower than $E_{fc} + 4.6k_BT$ is stored and used to calculate the absorption. All the necessary bands selected in this way are marked as solid lines in Fig. 4.2b. These selected bands are sufficient to calculate the gain spectrum in the interval as shown in Fig. 4.2a. Those not selected will thus have either a larger or smaller transition energy, outside the plot region or have negligible contribution to the gain spectrum.

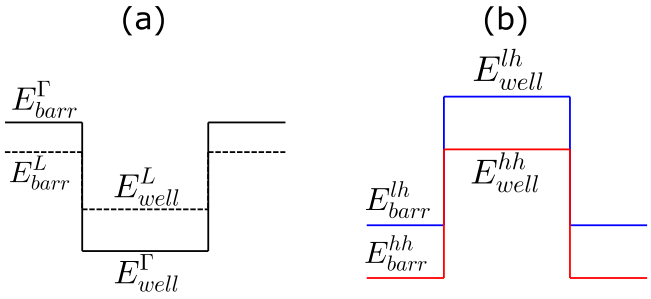

Figure 4.3: Schematic band diagram of (a)Conduction band (b)Top valence band.

.

To obtain the quasi-Fermi level, we assume the carriers are electrically pumped and the injected carrier density for electrons and holes are equal, while N-doping or P-doping provides additional carrier density of electrons and holes. The working temperature in this model are all set to be 300K. The quasi-Fermi level is calculated twice. The first calculation is in order to ensure that the necessary bands are stored, to accurately calculate the Fermi level and the optical properties. For the first calculation, a rough estimation is used, with a small number of points in k-space. Initially, we use the bound states for the first quasi-Fermi level calculation. More specifically, as shown in Fig. 4.3, for the conduction band at Γ valley, the subbands with Γ point energy between E^{Γ}_{well} and E^{Γ}_{barr} are used. For L valley, the subbands with quantized energy between E^{L}_{well} and E^{L}_{barr} will appear in the second term of Eq. 4.4a. For the valence band, bands with energy between $min(E^{lh}_{barr}, E^{hh}_{barr})$ and $max(E^{lh}_{well}, E^{hh}_{well})$ at Γ point are chosen. Note that when the strain is large, some quasi-free states between LH and HH are also included. This pre calculation of quasi-Fermi level is in order to verify if the selected bands used for the second calculation of quasi-Fermi level are sufficient, or whether any free states with higher energy also

need to be included. Here, we denote the first rough estimation of quasi-Fermi level as E'_{fc} and E'_{fv}. If $E'_{fc}+4.6k_BT$ or $E'_{fv}-4.6k_BT$ exceed the barrier, it would mean that these bound states are not enough for the calculation of the quasi-Fermi level, the quasi-free states with energy higher than $E'_{fc}+4.6k_BT$ in Gamma and L valley and those lower than $E'_{fv}-4.6k_BT$ also need to be included in calculation. For the accuracy of the quasi-Fermi level, this algorithm is only in precaution of some extreme case when the carrier density is very large. Usually, the bound states are sufficient for the calculation of quasi-Fermi level and the difference between two calculations is small. Another reason for this pre-calculation of Fermi level is to give a guide for the choice of SO bands. Since all the states at some k-point are calculated in parallel, it will exceed the RAM limit if all of them are stored. However, with a criteria to select bands, only those necessary for further calculation are stored, which dramatically improves the efficiency of the program. Moreover, we can further reduce the computation time by skipping the integration from the top valence band to the free states of LH and HH near the SO band, because SO have much lower energy than LH and HH, the free states of LH and HH around SO energy actually have a large k_z vector. These states will usually have zero inner product, and calculating them is time consuming. So in order to filter these free-states, the program will check if the states are bound by comparing the proportion of the wavefunction inside and outside of the well, and whether the envelope wavefunctions of $|u^{so}_{70}\rangle$ and $|u^{so}_{80}\rangle$ have a significant percentage of the total wave function. As an example, we use the same structure in Fig. 4.1, the filtered 2D band structure in k_x and k_y plane is shown in Fig. 4.4a

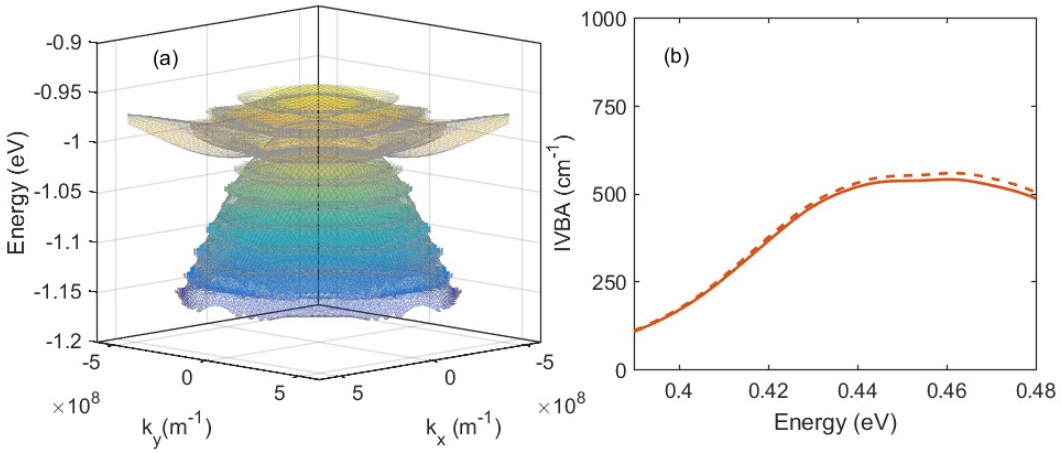

Figure 4.4: (a)Filtered 2D band structure of SO band. (b)IVBA calculated before(dashed line) and after (solid line) the filtering is applied

.

In Fig. 4.4a, apart from the paraboloid-like shape, some branch of free-states still exist at the

top of the SO subbands. This is because the coupling between HH, LH and SO makes these free-states to get some SO component of wavefunction, and the program uses a rather loose criteria in the filter subroutine to prevent the miscalculation in this part, although they may not be strictly necessary. Overall, this removes most of the free-states in the band structure, and the calculation time can be reduced by more than 2/3 by applying such method. Fig. 4.4b shows the calculated results before and after the free states are removed. Negligible difference is observed between these two, proving the reliability of such method.

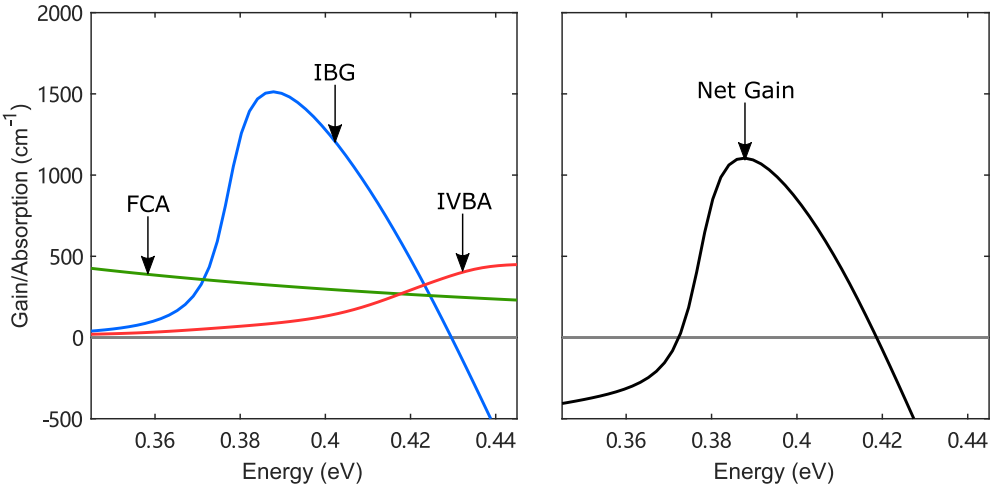

Figure 4.5: Example of the detailed information about the inter band gain and loss mechanism, together with the net gain spectra of a $Si_{0.15}GeSn_{0.129}/GeSn_{0.089}$ 14 nm QW with 1.07% tensile strain in the well.

.

In summary, the total absorption spectra is the sum of the inter-band absorption, free carrier absorption and inter-valence band absorption. Similarly the net gain can be calculated as:

$$g_{net}(\hbar\omega) = g_{ibg}(\hbar\omega) - \alpha_{fca}(\hbar\omega) - \alpha_{ivba}(\hbar\omega) \tag{4.6}$$

Fig. 4.5 shows the detailed results for a $Si_{0.15}GeSn_{0.129}/GeSn_{0.089}$ 14 nm QW with 1.07% tensile strain in the well. Fig. 4.5a explicitly shows the contribution from inter-band gain (blue), free carrier absorption (green) and inter-valence absorption (red), and Fig. 4.5b is the combined result. Due to the loss mechanisms the gain is reduced. In some other cases, it is possible that the loss is even larger than the inter-band gain. In such cases, although the population inversion condition is fulfilled, the net gain will still be below zero, and a larger carrier density will be

needed to have net gain. We can see in this example that the peak of IVBA appears at a higher energy than the IBG, thus it is not the major loss mechanism here. However, if it is close to the peak of IBG, it is possible that some structure has a larger IBG, but after considering the IVBA, this structure will no longer be the optimal choice. This shows the necessity of including the IVBA in the process of optimization. In addition, the FWHM is a somewhat arbitrary value, as discussed previously. The typical values of broadening used in laser modelling papers are around 5 meV - 20 meV (e.g. 7meV [41]; 20meV [85]). The FWHM here is set to 10 meV for interband gain and absorption, and 25meV for IVBA. The FWHM for IVBA is set to a much larger value in order to flatten the intense numerical fluctuations in the SO band. It is acceptable to do so since the real IVBA spectrum is much smoother than the IBA spectrum because of the slow increase of the s-like component in the valence band, therefore using larger FWHM will not affect the shape of the IVBA too much.

Last but not least, the whole flow chart of the program is shown in the appendix, which provides more details about how the whole program works.

4.3 Material Parameters

The **k.p** method is a semi-empirical method. A disadvantage of using such method is that it needs a lot of input parameters. Since Group IV with Sn is still in development, the parameters of this material system are not fully established. Various disagreements still exist in different papers. Although the parameters for pure elements are rather reliable. The major challenge exists when it comes to the binary or ternary alloy of Si, Ge and Sn. More studies address the non-linear parameter dependence on Sn composition for GeSn (e.g. [35, 38, 39])than for the other two alloys SiGe and SiSn. The properties of the ternary alloy SiGeSn are estimated using three bowing parameters of SiGe, GeSn and SiSn. Because of the lack of data, some of them can only be estimated by linear interpolation. This may result in some inaccuracy if the non-linear dependence on the composition is very large. The parameters used in this chapter are obtained by comparing those from the literature and other simulations, if small difference is found, the more recent value is used, and in the case of large variations we have used the more frequently used values. In this section, the parameters of pure elements, and also the bowing and interpolation parameter for alloys are given. The equations and methods of how

these parameter enter the model is introduced. Then, some parameters such as the band gap of L and Γ valley can be verified by comparing our simulation outcomes with some experimental and widely accepted results. They show reasonable accuracy with the data extracted from measurement results.

Table 4.1: Parameters of Si, Ge and Sn.

Parameter	Symbol	G_{Si}	G_{Ge}	G_{Sn}
Lattice constant	a_{lat} (Å)	5.4307[a]	5.6579[a]	6.489[a]
Deformation potential(CB)	a_c(eV)	-10.06[a]	-8.24[a]	-6[a]
Deformation potential(VB)	a_v(eV)	2.46[a]	1.24[a]	1.58[a]
Deformation potential(shear)	b(eV)	-2.1[a]	-2.9[a]	-2.7[a]
Deformation potential(shear)	d(eV)	-5.3[b]	-4.8[b]	-4.1[b]
Deformation potential(L)	a_L(eV)	-0.66[a]	-1.54[a]	-2.14[a]
Stiffness constant	C_{11}(GPa)	165.77[b]	128.53[b]	69[b]
Stiffness constant	C_{12}(GPa)	63.93[b]	48.26[b]	29.3[b]
Bandgap (Γ)	E_g^{Γ}(eV)	4.185[a]	0.8981[a,c]	-0.408[a]
Bandgap (L)	E_g^{L}(eV)	2.716[a]	0.744[a]	0.1202[a]
Split-off Energy	Δ_{so}(eV)	0.044[a]	0.297[a,d]	0.8[a,d]
Luttinger(6-bands)	γ_1^L	4.285[a]	11.11[e]	-12[a]
Luttinger(6-bands)	γ_2^L	0.339[a]	3.252[e]	-8.45[a]
Luttinger(6-bands)	γ_3^L	1.446[a]	4.689[e]	-6.84[a]
MME	E_p(eV)	21.6[a]	25.61[e]	23.8[a]
VBO	E_v(eV)	-1.27[f]	-0.69[a,d]	0
Effective mass(Γ)	m_e^{Γ}	0.241[c]	0.042[e]	-0.058[a]
Longi. Effective mass(L)	m_l^L	1.418[b]	1.5929[e]	1.478[b]
Trans. Effective mass(L)	m_t^L	0.13[b]	0.0922[e]	0.075[b]
Refractive index	n_r	3.4[g]	4[g]	6.18[h]
Varshni Parameter	α_{vshni}^{Γ}(eV/K)	$3.91 \cdot 10^{-4}$[i]	$5.82 \cdot 10^{-4}$[i]	0[a]
Varshni Parameter	β_{vshni}^{Γ}(K)	125[i]	296[i]	0[a]
Varshni Parameter	α_{vshni}^{L}(eV/K)	$4.774 \cdot 10^{-4}$[i]	$4.774 \cdot 10^{-4}$[i]	0[a]
Varshni Parameter	β_{vshni}^{L}(K)	235[i]	235[i]	0[a]

Note The parameters here have some difference with those used in [86]
[a] Reference [78]
[b] Reference [87]
[c] Reference [88]
[d] Reference [89]
[e] Reference [39]
[f] Reference [90]
[g] Reference [91]
[h] Reference [92]
[i] Reference [93]

4.3.1 Pure Elements and Linear Vegard's Law

The material parameters of pure elements are used to estimate the material parameters for $Si_xGe_{1-x-y}Sn_y$ alloy in the parameter space of composition, x and y. Without bowing, the parameter of an alloy can be calculated using the simple linear interpolation from the Vegard's

law. For convenience, let $x = \chi_{\mathrm{Si}}$, $1 - x - y = \chi_{\mathrm{Ge}}$, $y = \chi_{\mathrm{Sn}}$. Then the linear interpolation reads:

$$G_{\mathrm{SiGeSn}} = \chi_{\mathrm{Si}} G_{\mathrm{Si}} + \chi_{\mathrm{Ge}} G_{\mathrm{Ge}} + \chi_{\mathrm{Sn}} G_{\mathrm{Sn}} \tag{4.7}$$

in which G denotes a particular parameter, such as the deformation potential or stiffness constant. Hence G_{SiGeSn} is the estimated value of the parameter in SiGeSn alloy with x Si, $1-x-y$ Ge and y Sn. G_{Si}, G_{Ge} and G_{Sn} are the values of this parameter for pure elements. For binary alloy, for example SiGe, we just need to set $z = \chi_{\mathrm{Sn}} = 0$, then Eq. 4.7 is equivalent to:

$$G_{\mathrm{SiGe}} = \chi_{\mathrm{Si}} G_{\mathrm{Si}} + \chi_{\mathrm{Ge}} G_{\mathrm{Ge}} \tag{4.8}$$

Table 4.1 shows the parameters of pure Si, Ge and Sn. The linear Vegard's law described by Eq. 4.7 is used when the bowing parameter is missing in literature or when the bowing is small enough to be ignored.

The following parameters are using linear Vegard's Law in Table 4.1: deformation potentials a_c, a_v, b, d, a_L, stiffness constants C_{11} and C_{12}, valence band offset E_v, split-off energy Δ_{so}, refection index n_r and Varshni parameters α^{Γ}_{vshni}, β^{Γ}_{vshni}, α^{L}_{vshni} and β^{L}_{vshni}. Some uncertainty still exists about using this linear interpolation. For example, although the non-linear dependence of the deformation potentials in Group IV were rarely discussed, in III-V material system the bowing has been found for deformation potential for InGaAs alloy using first-principles calculation [94]. It is likely that the bowing of deformation potential can exist in SiGeSn material system. However, since linear Vegard's law is frequently used for this parameter (e.g. [40, 41, 78]), and also in view of the fact that Si and Sn are contents that are small in SiGeSn or GeSn alloys, it is acceptable to use linear approximation here, and essentially the error would be small. For stiffness constant, [95] and [96] show a small and negligible bowing in GeSn and SiGe, and because of the lack of data for SiSn, this parameter is also calculated by linear interpolation. The refractive index is one of the coefficients in equation for the absorption, the refection index is a constant in Eq. 3.15. So it is calculated as the average value across the active region. The temperature dependence of bandgaps is calculated from the Varshni formula:

$$E_g^i(T) = E_g^i(0) - \frac{\alpha^i T^2}{T + \beta^i} \tag{4.9}$$

in which $i = L$ or Γ and T is the temperature in the units of K.

As shown in [97], although the measured result for pure Ge match this dependence quite well, disagreement appears when the Sn composition dependent Varshini parameters are needed. Here we use the same way as in [78] and simply set Varshni parameters of Sn to zero. It seems that a recent research [98] also supports this to some extent, because for Sn β^{Γ}_{vshni} is only 11K and α^{Γ}_{vshni} is three orders of magnitude smaller than that for Ge, which indicates that such a linear interpolation could be acceptable. For the valence band offset and split off energy the linear interpolation is used because these are widely accepted values and frequently used in simulations [40, 78, 89, 99].

Again, the linear interpolation is mostly used because insufficient data has been found, and with further the research this may be replaced by some other choice which delivers a better match in the parameter space of interest.

4.3.2 Non-linear Parameter Dependence for Mixed Materials

The parameters for a mixed material do not necessarily have linear dependence on the composition. The non-linear dependence on the composition is usually described by bowing parameter. The material parameter calculated using bowing parameter and Vegard's law may be used to estimate this parameter in the full composition range. It still uses the values for the pure elements as a basis, and the values at boundaries of the composition (i.e. 0% and 100% of the composition) are just the parameters for pure elements. However, for some parameters, such as momentum matrix element and Luttinger parameters, the results are not fitted in the whole parameter space of the compositions (i.e. such estimates are only valid for a limited range of composition, for example less than 15% Sn). Finally, as a special case, the SiGe bowing of Luttinger parameters is not described by a simple value or an equation. For this reason, the parameters for ternary alloy SiGeSn are rescaled according to the ratio of values for SiGe with the same Si content, obtained by interpolation without and with bowing, respectively.

To include Bowing in full, the Vegard's law for ternary alloy needs three correction terms with appropriate bowing parameters. The Eq. 4.7 then becomes:

$$G_{\mathrm{SiGeSn}} = \chi_{\mathrm{Si}} G_{\mathrm{Si}} + \chi_{\mathrm{Ge}} G_{\mathrm{Ge}} + \chi_{\mathrm{Sn}} G_{\mathrm{Sn}} - \chi_{\mathrm{Si}}\chi_{\mathrm{Ge}} B_{\mathrm{SiGe}} - \chi_{\mathrm{Ge}}\chi_{\mathrm{Sn}} B_{\mathrm{GeSn}} - \chi_{\mathrm{Si}}\chi_{\mathrm{Sn}} B_{\mathrm{SiSn}} \quad (4.10)$$

Note that the sign of the bowing terms is negative. Hence for a positive bowing parameter the curve or surface will be bowed down, and vice versa. The parameters that use Eq. 4.10 and the bowing parameters are listed in Table 4.2

Table 4.2: Bowing Parameters of Si, Ge and Sn.

Parameter	Symbol	B_{SiGe}	B_{GeSn}	B_{SiSn}
Lattice constant	a_{lat} (Å)	0.026[a]	-0.041[a]	0[a]
Bandgap (Γ)	E_g^{Γ}(eV), c^{Γ}	0.21[a]	2.55[b]; $-4 \cdot 10^{-4} T$[a,c]	3.915[a]
Bandgap (L)	E_g^{L}(eV), c^{L}	0.335[a]	0.89[b]; $-7 \cdot 10^{-4} T$[a,b]	2.124[a]

[a] Reference [78]
[b] Reference [100]
[c] Reference [101]

Note that the temperature dependence is also included for the bowing parameter of GeSn in Table 4.2. The bowing parameter of GeSn at temperature T was calculated as:

$$B^{i}_{\mathrm{GeSn,g}}(T) = B^{i}_{\mathrm{GeSn,g}}(0) + c^{i}T \tag{4.11}$$

The Sn content necessary for indirect-direct transition is an important parameter for the simulation of Lasers based on GeSn material. As can be seen in Fig. 4.6, a large difference for this value appears depending on including the bowing parameter or not. If linear Vegard's law is used the predicted value for this transition is 22.5%, this matches the early research estimates ∼20%. However, recent research has found this value to be much smaller, and the prediction using the data from Table. 4.1 and bowing from Table. 4.2 gives 7% for the Sn content, which matches well with other recent data [16, 37, 100]. Also we should note that for the lattice constant, the zero bowing for SiSn was not taken because of the lack of data, but because it was found that this value really is very near zero, as indicated in [102].

It is worth mentioning that bowing may also exist for those parameters for which we are here using linear interpolation. For example, the calculations in [35] indicate the bowing for valence band offset and split off energy of GeSn. However the difference from linear interpolation is small when Sn composition is not large.

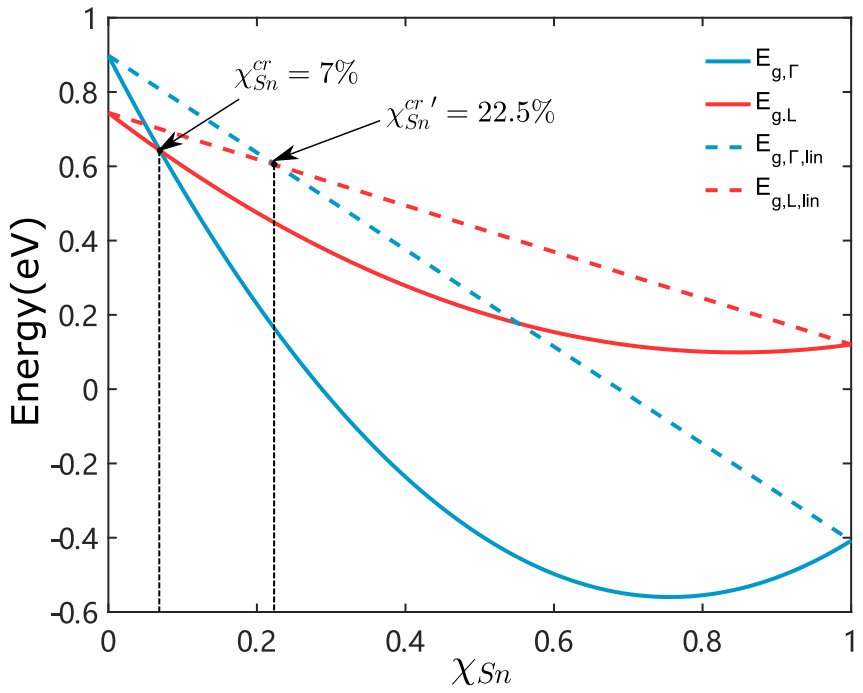

Figure 4.6: Energy gap at Γ and L valley calculated by linear interpolation (dashed lines) and with bowing included (solid lines) at 0K.

.

For parameters present in the band structure, like Luttinger parameter and effective mass, some early papers and even some recent ones still use the simple weighted average between Si, Ge and Sn [40, 41, 85, 89, 103]. It was recently found that linear interpolation could be inappropriate to estimate the composition dependence of these parameters. As indicated in [38] and [39], the Luttinger parameters fitted from empirical pseudopotentials calculation actually increase with the Sn content when it is in a limited range ($\leq$ 15%-20%). But the pure Sn has negative Luttinger parameters. Simple linear interpolation would cause the Luttinger parameters decrease with Sn content, which is opposite of the recently found trend. In this work these band structure parameter for GeSn are taken from a more recent work [39]. The parameters of $Ge_{1-y}Sn_y$ are calculated by the quadratic interpolation equation which reads:

$$G_{GeSn} = a_0 + a_1 y + a_2 y^2 \tag{4.12}$$

and the interpolation parameters are given in Table 4.3

Table 4.3: Interpolation Parameters of GeSn [39]

Parameter	Symbol	a_0	a_1	a_2
Luttinger	γ_1^L	11.11	28.53	105.2
Luttinger	γ_2^L	3.252	13.95	52.75
Luttinger	γ_3^L	4.689	14.37	52.57
Effective mass(Γ)	m_e^Γ	0.042	-0.1436	0.1026
Longi. Effective mass(L)	m_l^L	1.5929	-0.0642	-0.1087
Trans. Effective mass(L)	m_t^L	0.0922	-0.0561	-0.0146

As a coefficient of zero order term, a_0 is the parameter for pure Ge. Note that the quadratic interpolation method and the data in Table 4.3 are only valid when Sn content is less than 15%. Note that by using quadratic interpolation the parameters of pure Sn in Table 4.1 are not used in our calculations. Also, to calculate the material parameters of ternary alloy $Si_xGe_{1-x-y}Sn_y$, the GeSn is considered as one 'unit', and linear interpolation is then used for its 'mixture' with Si:

$$G_{\mathrm{SiGeSn}} = \left[a_0 + a_1\frac{\chi_{\mathrm{Sn}}}{1-\chi_{\mathrm{Si}}} + a_2\left(\frac{\chi_{\mathrm{Sn}}}{1-\chi_{\mathrm{Si}}}\right)^2\right](1-\chi_{\mathrm{Si}}) + G_{\mathrm{Si}}\chi_{\mathrm{Si}} \tag{4.13}$$

By using Eq. 4.13, we assume that the bowing parameters of SiGe and SiSn are zero. For the effective mass, m_e^Γ, m_l^L and m_t^L, Eq. 4.13 is used because the lack of data for bowing factors for SiGe and SiSn. However, as indicated in [88] and [104], a strong non-linear dependence of Luttinger parameter with Si composition exists. Hence, to estimate such bowing two additional material parameters are introduced here: E_p' and E_g', which are defined as the momentum matrix element and energy gap between the topmost valence band and Γ_{15} conduction band. The values of E_p' and E_0' are given in Table 4.4.

Table 4.4: Energy gap and momentum matrix element of Si and Ge between the top valence band and Γ_{15} conduction band [88]

Parameter	Symbol	Si	Ge
MME($\Gamma_{25} \rightarrow \Gamma_{15}$)	E_p'	14.4	17.5
Energy gap($\Gamma_{25} \rightarrow \Gamma_{15}$)	E_g'	3.4	3.124

The Luttinger parameters are calculated from the following equations:

$$\gamma_1 = \frac{E_p}{3E_g} + \frac{2E_p'}{3E_g'} + \bar{\gamma_1} \tag{4.14a}$$

$$\gamma_2 = \frac{E_p}{6E_g} - \frac{E_p'}{6E_g'} + \bar{\gamma_2} \tag{4.14b}$$

$$\gamma_3 = \frac{E_p}{6E_g} + \frac{E'_p}{6E'_g} + \bar{\gamma_3} \tag{4.14c}$$

in which E_p, E'_p, E_g and E'_g of Si_xGe_{1-x} are estimated by linear interpolation. $\bar{\gamma_1}$, $\bar{\gamma_2}$ and $\bar{\gamma_3}$ should not be confused with Luttinger parameter, they are just linear interpolation values between Si and Ge. Although in [88] the bowing of E'_g is also mentioned, this bowing is small and it can be considered as second order bowing, compared to the bowing of Luttinger parameters. The difference between taking linear interpolation of Luttinger parameters and using Eq.4.14 is shown in Fig. 4.7. A strong bowing can be easily seen in this figure. To introduce such bowing to ternary $Si_xGe_{1-x-y}Sn_y$ alloy, we denote the value from linear interpolation and the one from Eq.4.14 as γ_{lin} and γ_{bow}. Then the Luttinger parameters calculated by Eq. 4.13 are rescaled by the ratio of γ_{bow} and γ_{lin} for the same Si concentration, which reads:

$$\gamma_{\mathrm{SiGeSn}} = \left\{ \left[a_0 + a_1 \frac{\chi_{\mathrm{Sn}}}{1-\chi_{\mathrm{Si}}} + a_2 \left(\frac{\chi_{\mathrm{Sn}}}{1-\chi_{\mathrm{Si}}} \right)^2 \right] (1-\chi_{\mathrm{Si}}) + \gamma_{\mathrm{Si}} \chi_{\mathrm{Si}} \right\} \frac{\gamma_{bow}}{\gamma_{lin}} \tag{4.15}$$

We should mention again that the above equation will work only for Ge rich alloy. It becomes inaccurate when Si or Sn composition becomes large.

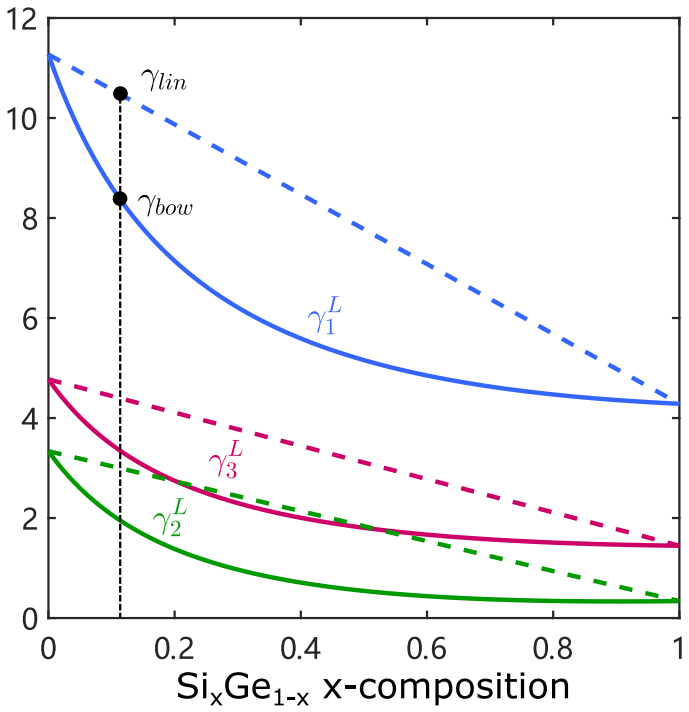

Figure 4.7: Luttinger parameters calculated using Eq.4.14. A strong bowing can be seen for all three parameters.

It is also worth mentioning that the parameters choice may also involve numerical issues. Fig. 4.8 shows the calculated band structure of $Si_{0.1}Ge_{0.9}$/Ge 14nm QW after (a) and before (b) the Luttinger parameter bowing of SiGe is included. The band structure in Fig. 4.8a looks fine. However, when the bowing of SiGe is not taken and only the linear interpolation between Si and Ge is used, the unusual splitting of the different spins appears. A possible explanation could come from the differential scheme, because the local inversion symmetry is broken at the interface, and ignorance of the bowing will overestimate the contribution from the remote bands. However, a more precise reason of the source of such splitting and the reason why it is removed after the bowing is accounted remains a problem.

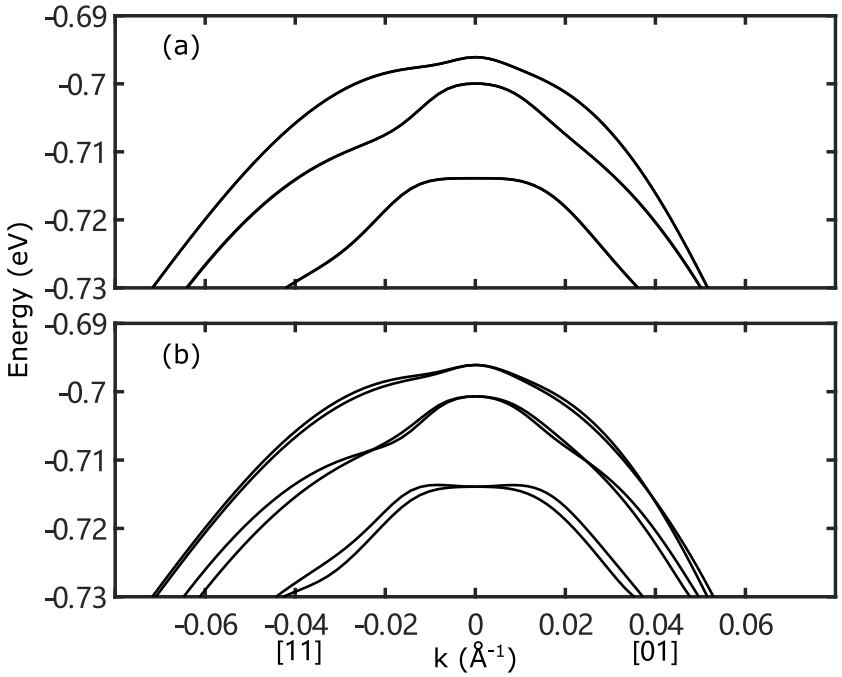

Figure 4.8: Band structure of $Si_{0.1}Ge_{0.9}$/Ge 14nm QW. (a)Before bowing of SiGe is applied. (b) After bowing is applied.

4.3.3 Discussion on The Parameters Choice

The choice of parameters is always a tricky process in the field of modelling, especially when using empirical or semi-empirical methods, because they have many input parameters and it is challenging to trace down the source for each input. The parameters come from different sources, some of them are experimentally measured and others may be calculated in a more fundamental way like first-principle density functional theory. These data may fit the conditions they are measured or derived for, which at the same time represents a possible limit of their

validity. With further development of this field, the parameters will certainly be 'refreshed', which brings a better accuracy of the predicted results from computational models.

Next, some experimental results from different papers are presented and compared with the calculated results using our model. As can be seen in Table 4.5, a rather good agreement is achieved regarding the transition energy with different strains and temperatures. This can add some validity to the parameters choice, but it is hard to be sure that all the inputs are fully reliable, because the results may be less sensitive to some of the parameters. More research needs to be done on the parameters for the ternary alloy SiGeSn, although it is here a barrier material, it usually has a smaller influence on results, but sometimes the barrier properties are very important, for example the band alignment and carrier confinement of LH for tensile strained laser. The accurate valence band offset will directly influence the modelling results of carrier confinement for the valence band. The carrier confinement is equally important for LH as for CB. Some other properties of the semiconductor structure will be influenced by the band alignment, such as carrier transport. But as this topic has not been studied in this work, we will only discuss about the band discontinuity using the existing parameters from other works. Some detailed calculations and discussion will be presented in section 4.4.

Table 4.5: Comparison of experimental measured and calculated photon transition energy (eV) of high Sn content QWs

Structure	Well width (nm)	Strain in well ε_{xx}	Temp.	Transition band	Exp.	Calculated
$Si_{0.5}GeSn_{0.13}$/ $GeSn_{0.13}$	22	-0.68%[f]	4K	C1-HH1	0.51[a,d]	0.498
$Si_{0.5}GeSn_{0.63}$/ $GeSn_{0.138}$	22	-1.05%[f]	10K	C1-HH1	0.5[a,d]	0.497
$Si_{0.15}Ge$/ Ge	10	-0.29%[g]	300K	C1-HH1	0.88[c,e]	0.879
				C1-LH1	0.91[c,e]	0.9126

[a] Reference [84]
[b] Reference [18]
[c] Reference [105]
[d] Transition energy is extracted from the peak of photoluminescence spectrum
[e] Transition energy is extracted from the exciton peak of effective absorption spectrum, more detail about the comparison of absorption spectrum can be seen in Chap. 5.
[f] Strain is obtained from the experiments measurements
[g] Strain is obtained by assuming the $Si_{0.1}Ge$ substrate has tensile strain of 0.15%.

4.4 Band Alignment

The band alignment problem arises because for tensile strained lasers, recombination happens mainly between CB and LH. But for SiGeSn, GeSn heterostructure light holes usually have a poorer carrier confinement than HH because of a different dependence on strain. However, this problem does not exist in some earlier demonstrated compressively strained SiGeSn/GeSn QW (e.g. [89]), because these lasers use HH as a lasing band, and when Si content in the barrier is ∼1%, the heavy hole discontinuity at the SiGeSn/GeSn interface will usually be larger than ∼100 meV, which is sufficient for a good carrier confinement. The situation is different for LH, because Si has a smaller lattice constant, and when the Sn contents in the barrier and well are similar or the same, alloying with Si will make the barrier layer more tensile strained, which 'pushes up' the band edge energy of LH in the barrier. This will give a shallow LH well and a poor carrier confinement can be expected if the barrier are carelessly chosen. A problem which is related to this poor carrier confinement, such as carrier leakage, is possible to happen especially at high operating temperatures.

To make it clearer, Fig. 4.9 shows three possible band alignment situations in the parameter space. For type I (Fig. 4.9a), the energy of Γ CB in the GeSn layer is lower than in the SiGeSn barrier, and the LH in GeSn layer has a higher energy than in the barrier. Thus the electrons and holes can be confined in the same layer, which is necessary for an efficient laser material. Here we denote the band discontinuity in type I as positive as are the values of V_Γ and V_{lh}. For type II alignment (Fig. 4.9b,c), it is possible that either V_Γ or V_{lh} is negative, in such a case electrons and holes are not confined in the same layer and this reduces the efficiency of recombination. So, what we are interested here is type I with sufficiently large discontinuity to maintain a good carrier confinement for both electrons and holes.

The widely used valence band offset value from [40, 78, 89, 99] gives the VBO between Ge and Sn of 0.69eV, and 0.48eV between Si and Ge. These values were first demonstrated in 1988, but are still be frequently used nowadays. However, it is worth mentioning that some disagreements exist in some more recent research. As indicated in [106], the 'natural' unstrained valence band offset between Si and Ge could be as large as 0.8eV. and the VBO between Ge and Sn is around 0.9eV [35]. In Table 4.1, an intermediate setting is used. Offset between Ge and Sn is the frequently used value (0.69eV), and for the VBO between Si and Ge the value 0.58eV in [90] is used.

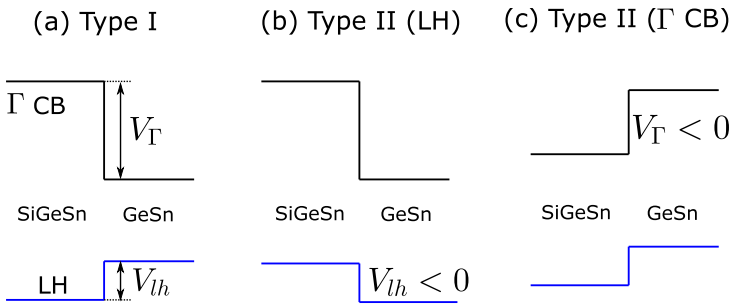

Figure 4.9: Three different band alignment types which can appear in the parameter space.

.

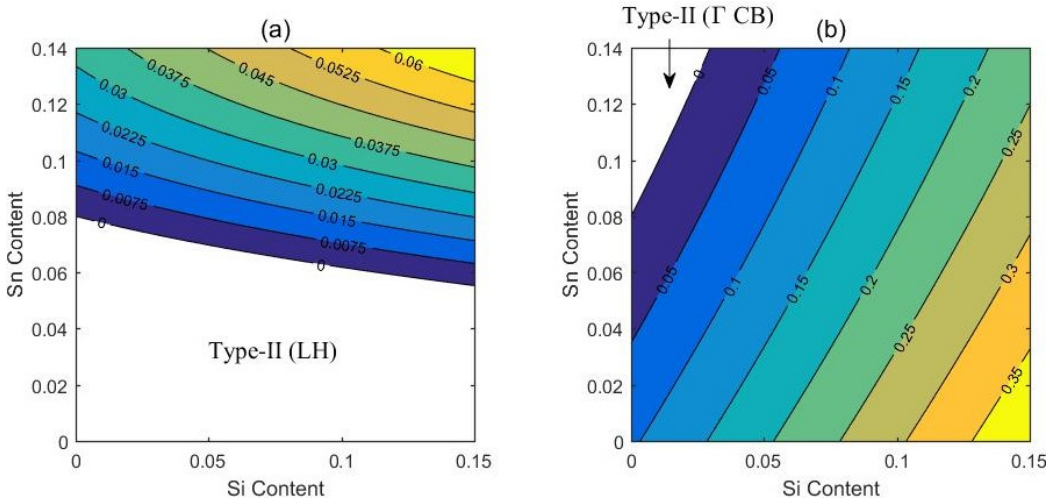

Figure 4.10: Band discontinuity (eV) at $Si_xGeSn_y/Ge_{0.92}Sn_{0.08}$ interface for (a) LH band (b) Γ conduction band. The substrate is $Ge_{0.85}Sn_{0.15}$, and the results are calculated using data from Table 4.1.

.

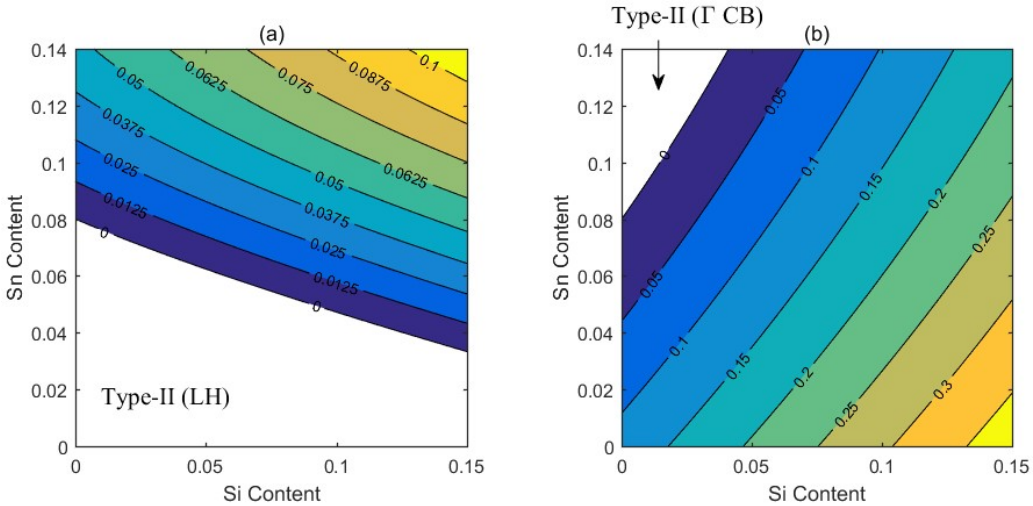

Figure 4.11: Band discontinuity (eV) at $Si_xGeSn_y/Ge_{0.92}Sn_{0.08}$ interface for (a) LH band (b) Γ conduction band. The substrate is $Ge_{0.85}Sn_{0.15}$, and the valence band offset is from [35, 106].

.

To compare how the different VBO input will influence the band alignment in the parameter space, the results calculated using VBO set from both Table 4.1 ($\Delta E_v^{Si,Ge} = 0.58$ eV, $\Delta E_v^{Ge,Sn} = 0.69$ eV) and [35, 106] ($\Delta E_v^{Si,Ge} = 0.84$ eV, $\Delta E_v^{Ge,Sn} = 0.91$ eV $B_{vbo}^{GeSn} = 0.59$ eV) with other parameters remaining the same, are shown in Fig. 4.10 and Fig. 4.11. The QW layer is here set as $Ge_{0.92}Sn_{0.08}$, the whole structure is assumed to be pseudomorphically grown on a $Ge_{0.85}Sn_{0.15}$ substrate. This gives a 1.05% tensile strain in the well layer. Some similar trends can be seen in both figures. Firstly, increasing the Si content in the barrier layer can increase the V_Γ and V_{lh}. Secondly, different from increasing the Si content, a larger Sn content in the barrier will reduce the discontinuity of Γ CB but will increase that of LH. This is because, although Sn has larger E_v than Si and Ge, and mixing with Sn reduces the V_{lh}, but because of a much larger lattice constant of Sn, higher Sn content in SiGeSn will make the alloy have a larger natural lattice constant, and the strain is then reduced. This reduction of strain then becomes the dominant effect that pulls down the bandedge energy of LH in the barrier, further increasing V_{lh}. Then we can see that Fig. 4.10 and Fig. 4.11 give similar trend of V_Γ in the parameter space, the major difference is found for V_{lh}. With smaller valence offset between Si and Ge, a much narrower region of type-I band alignment exists. From both figures, we can see that for the CB carrier confinement is less problematic, because for a 10% Si in the barrier, the discontinuity at SiGeSn/GeSn interface is at least 100 meV for Sn content in the barrier less than 14%. The major concern comes for the LH. For the large VBO between Si and Ge, as in Fig. 4.11, a 10% Si and Sn in the barrier will give about ∼50meV carrier confinement, however the same composition in Fig. 4.10 only gives ∼30meV discontinuity. Then to achieve a ∼50meV carrier confinement, the barriers are chosen with a Si content of 15% and Sn content is 4% larger than in the well layer. A similar setup for the barrier can be also found in [42]. With this setting, similar discontinuity (> 50 meV) can be found for other parameters choices for Sn and strain in the well layer, and a reasonable carrier confinement is achieved.

The experimental characterization of the band alignment in tensile strained SiGeSn/GeSn QW is still missing for now. So all the results shown here are just theoretical prediction based on the parameters choice. These predictions may have some error compared with reality, hence more experimental work is needed to justify these predictions.

4.5 Optimization Consideration

For GeSn, increasing Sn content and introducing larger tensile strain will both result in a more direct bandstructure, and large separation of LH and HH will give a relatively independent LH band, thus a lower overall density of state near the valence band top can be found. Both of these effects are beneficial for efficient light emitting material. Thus we may speculate that the larger the Sn content and tensile strain are, the easier it will be to achieve lasing in such material. However, the value of Sn content and strain in reality is limited by fabrication challenges. Also, some applications have requirements on the lasing wavelength. Hence, a more practical optimization direction would be to find the most suitable combinations of Sn content and strain for a certain photon transition energy in a reasonable parameter space.

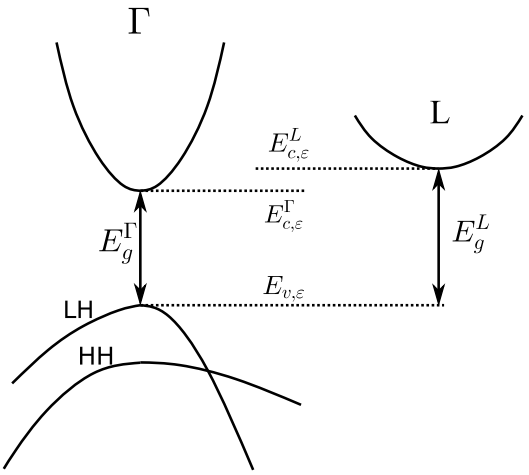

Figure 4.12: Schematic diagram of the band structure of bulk material

.

To start with, we look into the bandgap dependence of the bulk GeSn in the two dimensional parameter space of strain and Sn content. The definition of the bandgap to different valleys under strain can be seen in the schematic diagram in Fig. 4.12. Note that the valence band energy $E_{v,\varepsilon}$ is the energy of the valence band top at Γ point, and for tensile strain, as shown in Fig. 4.12, $E_{v,\varepsilon} = E_{lh,\varepsilon}$. For compressive strain, $E_{v,\varepsilon} = E_{hh,\varepsilon}$. Then the contour line of minimum bandgap at room temperature calculated in the ε_{xx}-Sn plane is shown in Fig. 4.13. The minimum bandgap is defined as:

$$E_g^{min} = \min(E_g^\Gamma, E_g^L) \tag{4.16}$$

The combinations of Sn content and ε_{xx} required to achieve a indirect-direct transition are also

shown in this figure as bold black solid line, it is clear that with the help of tensile strain, the Sn content to achieve a direct bandgap is lower. The critical Sn content to achieve this transition without strain is about 6.5%, and for pure Ge the strain needed for direct band is around 1.7%. These match the current values given in [37, 64, 65, 100]. Additionally, the Sn content dependence on in-plane strain, where the Γ bandgap is equal to the transition energy between the top valence band and SO band is shown as red solid line. A large IVBA loss could be expected around this line. Although bulk GeSn is different from the SiGeSn/GeSn QW structure, this can still give us a very good guidance to the following optimization process, and help to choose the part of parameter space of interest. The band edge energies of Γ-CB, L-CB, LH, HH, and SO are calculated by Eq. 2.87a-d.

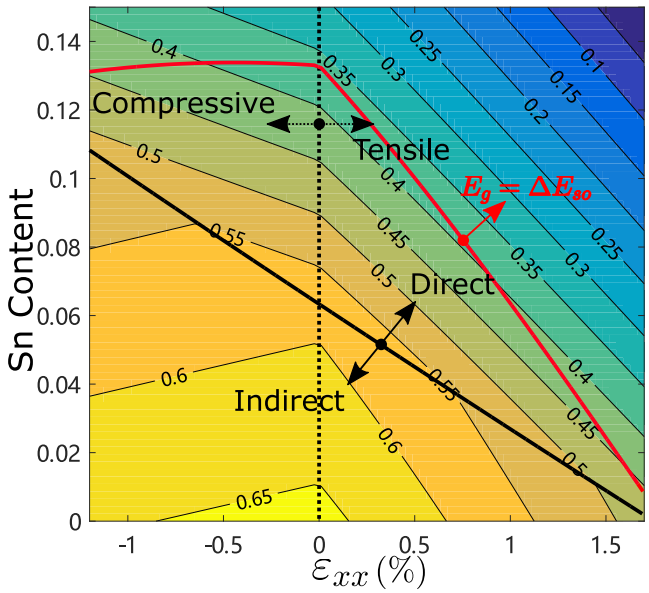

Figure 4.13: Minimum bandgap contour line of GeSn at 300K in the two dimensional space of Sn content and in-plane strain.

.

Next, we consider the quantum well structures. Similar to bulk material, the same photon transition energy can be achieved by adjusting the strain and Sn content in the well layer. Different from the bulk material, we also need to take quantization effects into consideration. For LH and Γ conduction band, the transition energy of SiGeSn/GeSn QW is usually ∼0.06 eV larger than the bulk GeSn, when tensile strain is in the well layer. For compressive strain, this difference will be smaller because heavy holes have a larger quantization effective mass. We focus on the tensile strain and three examples are given to illustrate some key design considerations.

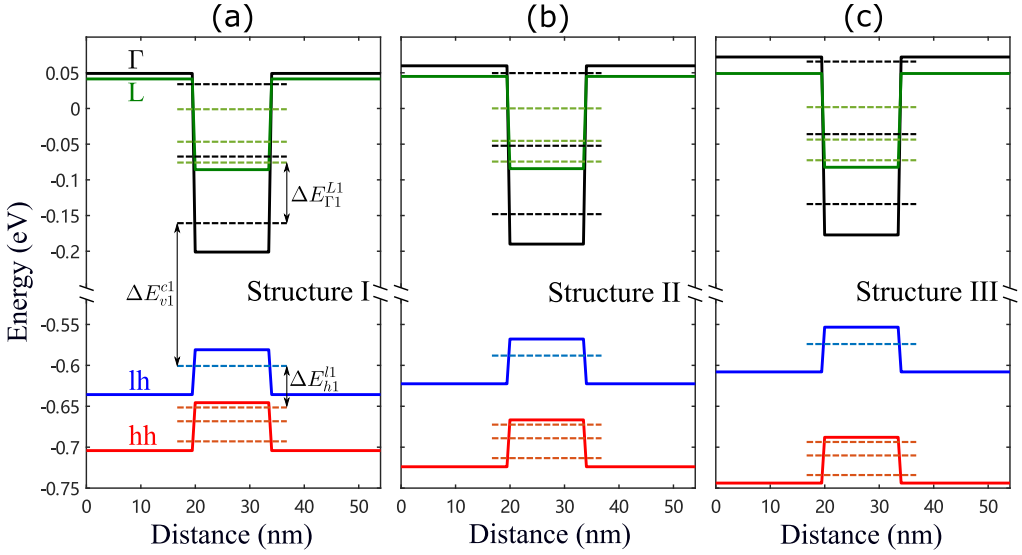

Figure 4.14: Band structure of $Si_{0.15}GeSn_{y+0.04}/GeSn_y$ and in-plane strain ε_{xx} in well, with 0.44eV photon transition energy (a)$\varepsilon_{xx} = 0.58\%$, $y = 9.55\%$; (b)$\varepsilon_{xx} = 0.86\%$, $y = 8\%$; (c)$\varepsilon_{xx} = 1.14\%$, $y = 6.44\%$.

.

The three structures shown in Fig. 4.14a, b and c are denoted as structure I, II, and III here. The well widths are all set to be 14nm. Again, the barrier are set to have 15% Si and the Sn content is 4% larger than in the well layer, as was explained in Section 4.4. Instead of setting a particular buffer layer, the in-plane strain ε_{xx} in the well is here assumed to be an independent variable, and then the strain in the barrier is calculated accordingly. All three structures have transition energy of 0.44eV. With ε_{xx} of 0.58%, 0.86% and 1.14% respectively for structure I, II and III, the Sn fractions in the well layer are 9.55%, 8%, 6.44% to achieve the same transition energy. The quality of such structures for laser applications can be characterized by two figures of merit, the spacing between the first quantized state of L-valley and Γ-valley $\Delta E_{\Gamma 1}^{L1}$, and the spacing between the first quantized state of LH and HH, ΔE_{h1}^{l1}. Since the Fermi-Dirac distribution is much less steep at room temperature than at low temperature, it is easier for the carriers to be distributed into these higher energy states, which points to some of difficulties in designing a laser for room temperature, and makes these figures of merit important aspects to be considered.

For $\Delta E_{\Gamma 1}^{L1}$, from structure I to III, the values are 85 meV, 73 meV and 61 meV. And the values of ΔE_{h1}^{l1} for structures I to III are 51 meV, 84 meV and 120 meV. So we can see a general trend that, for a particular transition energy, by increasing the tensile strain and reducing the Sn content, $\Delta E_{\Gamma 1}^{L1}$ will decrease with the strain, i.e. the band structure becomes less direct, which

reduces the quality of the structure. However, ΔE_{h1}^{l1} will increase with strain, which makes the near-top valence band to have a smaller density of states, which is favourable for a laser material. So there is a trade-off between these two effects to achieve the largest inter-band gain. If one of these figures of merit is already large, its further increase will not significantly improve the performance. However, if this causes the other figure of merit to decrease, and its value was already not large, the overall performance will degrade. Therefore, we do not expect that either too large or too small strain will result in the best performance, and a balanced, or optimal point exists somewhere in between.

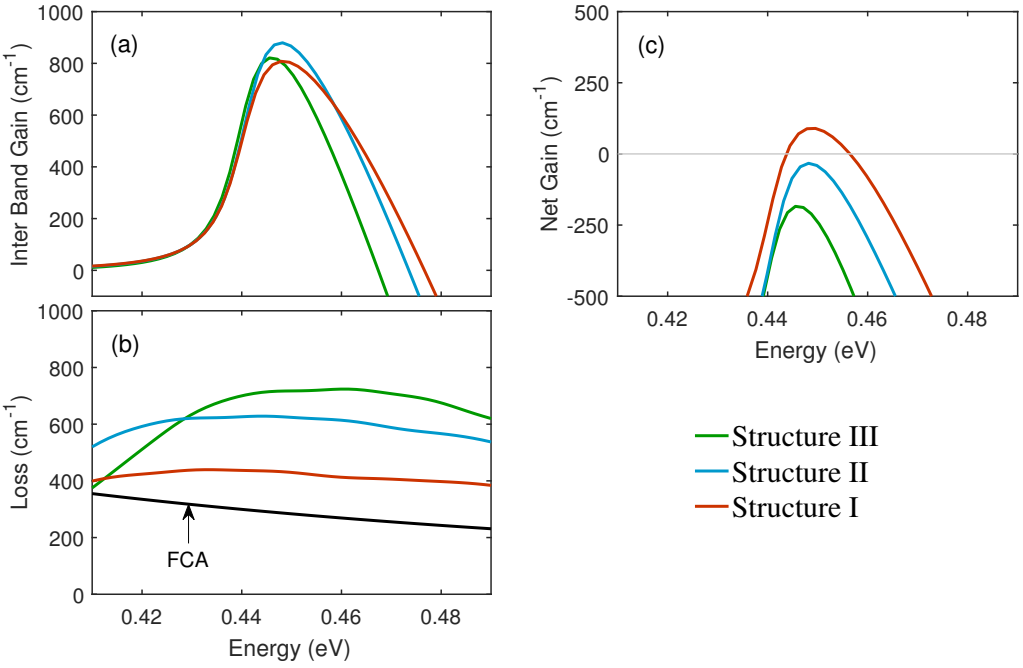

Figure 4.15: Gain and absorption spectrum of structures I-III in Fig. 4.14 with 5×10^{12} cm^{-2} carrier density. (a)Inter band gain; (b)Inter valence band absorption and FCA; (c)Net gain.

.

In Fig. 4.15, the gain spectra of structures I-III with 5×10^{12} cm^{-2} carrier density are plotted. For inter-band gain, as shown in Fig. 4.15a, the peak gain values are actually very near to each other, but structure II has the largest peak gain. This matches the expectation mentioned above. More importantly, these examples are chosen so that the transition energy is close to the difference of top valence band and split off band. Then, as expected, all of these three cases have relatively large IVBA, as shown in Fig. 4.15b. Among them, structure III has the largest loss at the gain peak. So the results after considering the loss mechanisms are very different. Structure I, which has the lowest inter band gain now becomes the only structure which has

positive net gain, and for the other two structures the inter-band gain is totally cancelled by the loss mechanism from IVBA and FCA.

In summary, to obtain an optimized structure of SiGeSn/GeSn QW for laser, we not only need to consider the inter band gain, which could be characterized by figures of merit of $\Delta E_{\Gamma 1}^{L1}$ and ΔE_{h1}^{l1}, but also the loss mechanisms from IVBA and FCA. The high IVBA can usually be found near the Sn content and strain combination where the transition energy is close to the difference between top valence band and split off band. To verify this assumption, the detailed calculations throughout the 2D parameter space of Sn content and tensile strain are made. This, combined with inter band gain in the two dimensional parameter space, will give us some idea about how this effect will influence the overall net gain and how the parameters choice should be modified accordingly. We expect this will make useful guidance for further research.

4.6 Well width

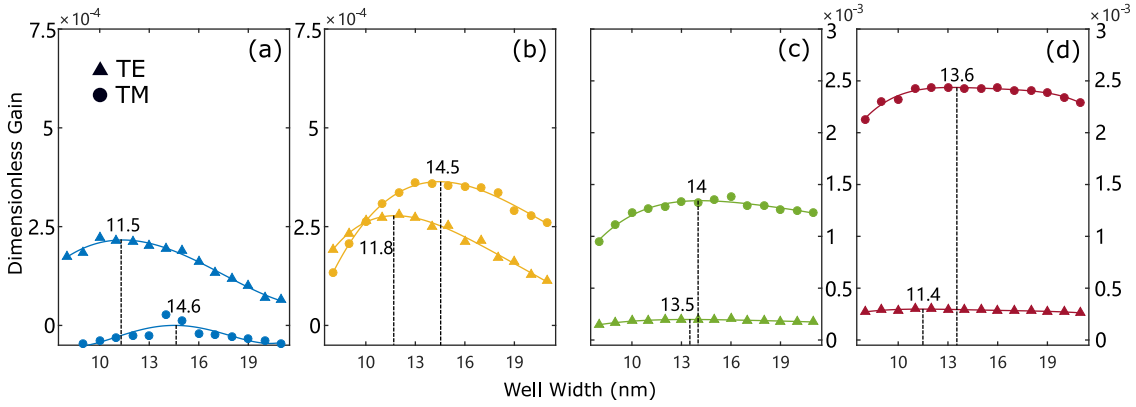

Figure 4.16: The maximum dimensionless gain vs. well width dependence (2D carrier density is 4×10^{12} cm^{-1}), curve are 4th order fitted with polynomial, with the maximum points denoted. (a)$\varepsilon_{xx} = 0\%$, Sn = 12%; (b)$\varepsilon_{xx} = 0.15\%$, Sn = 12%; (c)$\varepsilon_{xx} = 0.7\%$, Sn = 10%; (d)$\varepsilon_{xx} = 1\%$, Sn = 10%. (Calculated by parameter (c))

.

To reduce the number of variable, well width is determined in advance. Because it would be hard to give optimal well width for every combination of Sn content and strain, making well width as a separate input variable is more reasonable. Besides, different well widths will also change the photon transition energy. So it is convenient to know the rough interval of suitable well widths. When calculating the gain for a QW, the structural input L_z in Eq. 3.15 is equal to

the well width L_w. As a denominator in the equation, the different well widths will have a large effect on the absorption coefficient in units of cm^{-1}. Here we use and compare dimensionless gain, calculated as $\alpha \times L_w$. Also, in this section, only inter-band gain is studied, because the situation for the loss is different for different combinations of Sn content and strain, but here we only aim at giving a general range of the well widths. The well width may be reconsidered for a particular combination of Sn content and strain, and more specific design obtained by analysis of all effects together.

Four examples are shown here, the first two with low tensile strain and the last two with high tensile strain. The strains in the well from Fig. 4.16a to Fig. 4.16d are 0%, 0.15%, 0.7% and 1%. The Sn contents in the well for Fig. 4.16a and Fig. 4.16b are 12%, and For Fig. 4.16c and Fig. 4.16d they are 10%. The structure with low strain gives a similar TE and TM gain, because the LH and HH band are very close to each other. Highly strained structures have the TM gain much higher than TE, because LH is much higher than HH and becomes the dominant lasing band. Note that the barriers are using the setting discussed in section 4.4 to have a reasonable carrier confinement for the LH in the model. Since the calculations are done with step of 1nm, to avoid the numerical instability, the 4th order polynomial fitting is used to smooth the peak gain curve vs. well width dependence. The carrier density used here is 4×10^{12} cm^{-2}. As can be seen, all these curves, apart from some fluctuations caused by numerical effects, have a trend of gain increasing with well width for smaller well widths, and then decreasing after the peak point is reached. This is because when the well width is small, e.g. 8nm, the band structure is less direct than for a larger well width. Since the difference of quantization effects for Γ and L valley, and L valley have a much larger effective mass than the Γ valley, when the well becomes narrower, the quantized energy in Γ valley will increase much faster than for the L valley, hence reducing the figure of merit $\Delta E_{\Gamma 1}^{L1}$. This will result in more carriers in the L valley and less remaining in the Γ valley. According to the calculations, the percentages of carriers in the Γ valley for the four structures shown in Fig. 4.16 for 8nm well width are 19%, 22.4%, 21.4% and 28.4%, respectively. But when the well width is 21 nm the fraction of carriers in Γ valley increases to 30.3%, 35.4% 34.5% and 44.6%. On the other hand, a wider well width is not always good for gain, because larger well width will result in the step like density of states to become more like 3D density of states. This is equivalent to increasing the DOS near the top of the valence band. With these states also occupied by carriers, especially at room temperature, and with the peak gain occurring between the lowest CB and highest VB states, the large well

width will also make QW to lose its superiority over the bulk material.

Apart from the above examples, three additional structures with the same band gap energy are tested here to provide more support to the conclusions. With in-plane strain of 0.4%, 0.9% and 1.4%, Sn content of 10.52%, 7.79%, 4.96%, the band gap of GeSn is 0.38eV. The estimated optimal well width is 14.6nm, 15.8nm and 16.8nm. It seems that, for the same bandgap and similar transition energy, larger strain and smaller Sn content structures tend to have larger optimal well width. An explanation of this could be that larger strain and smaller Sn content to achieve same transition energy will make the difference $\Delta E_{\Gamma 1}^{L1}$ smaller, which prevents taking a smaller well width, because smaller well width will reduce the directness of the band structure.

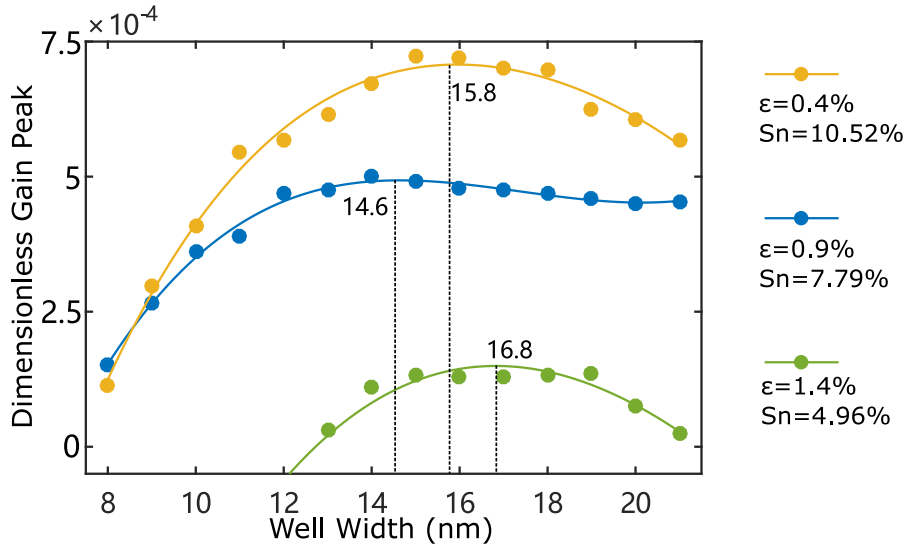

Figure 4.17: Dimensionless gain of SiGeSn/GeSn quantum well with different well widths. The band gap of GeSn is 0.38eV with different combinations of in-plane strain and Sn content in the GeSn layer.

.

Finally, if we consider the MQW structure, the barrier width is also another important factor to be considered. Thin barriers will increase the overall optical confinement in the QW active region and increase the model gain. But thin barriers will result in a poor carrier confinement, the wavefunction will penetrate through the barrier to the adjacent well. The influence of these effects (e.g. carrier leakage) on the overall performance of an electric pumped MQW laser needs additional calculations of carrier transport, which was not done in this work. Hence, here we just give a rough guide for the barrier width, based on the qualitative observation of the results. This is that a ∼10 nm barrier is good enough to separate CB and LH wavefunctions in each

well, and for the well width of a tensile strained QW, 13-16 nm is a good choice to deliver a good TM mode gain.

4.7 Optimization of Sn Content and Tensile Strain

For this optimization we are only interested in the TM polarization, because for tensile strained SiGeSn/GeSn QW large part of holes are in LH state, and this creates much larger absorption/-gain for TM than for TE.

4.7.1 Loss Mechanisms

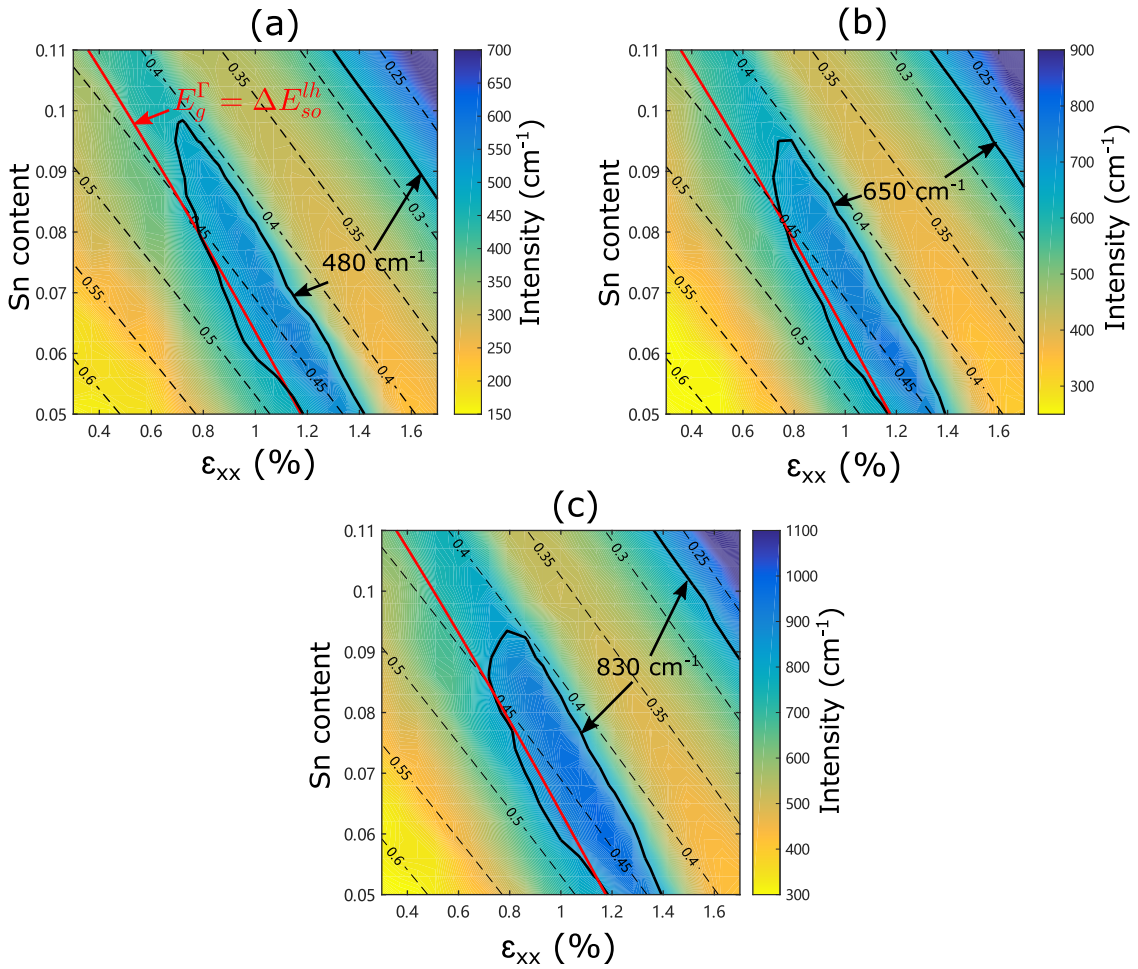

Figure 4.18: Loss mechanism from IVBA and FCA with carrier densities of (a)3×10^{12} cm^{-2}, (b)4×10^{12} cm^{-2}, (c)5×10^{12} cm^{-2}.

.

Here, a 14nm well width is used as a fixed variable, and the two dimensional space of in-plane strain ε_{xx} and Sn content in the well is explored. The range of ε_{xx} is from 0.3% to 1.7%, and the Sn content is from 5% to 11%, and this parameter space is discretised into 21×21 grid. Again the barrier for the $Ge_{1-y}Sn_y$ well is set to $Si_{0.15}Ge_{0.81-y}Sn_{y+0.04}$. The results for losses are shown in Fig. 4.18, using interpolation and colour representation.

In Fig. 4.18, the values of loss from IVBA and FCA are taken from the value at the peak of the inter-band gain, and those without gain are taken from the transition energy plus the half width half maximum of the Line-shape function (5 meV). The photon transition energy is also marked using dashed lines in the parameter space. These results are for three different 2D carrier densities 3×10^{12} cm^{-2}, 4×10^{12} cm^{-2}, 5×10^{12} cm^{-2}. The black solid lines are the contour line of 480 cm^{-1}, 650 cm^{-1}, 830 cm^{-1} loss for these carrier densities respectively. Also, the strain and Sn content for which the band gap is equal to the difference between the top valence band and SO band for bulk GeSn is marked as red solid line in these figures. The ridge-like high loss region in the middle is mainly caused by the IVBA, the higher loss at the upper right corner of each plot is caused by the FCA, because FCA tends to be lager for lower transition energy. As can be seen, the high loss region caused by IVBA appears in parallel with the red line, with lower energy, the difference between high loss ridge and the dependence when $E_g^\Gamma = \Delta E_{so}^{lh}$ is from the quantization effect of QW. Although, as explained before, the peak of the IVBA has higher energy than the transition energy between LH and SO for bulk material, but here the transition happens between quantized subbands. The quantization effect makes the inter-band transition energy much larger than the bandgap, implying that more Sn and strain is needed to make the inter-band transition energy and the inter valence band transition energy equal. We can also see that the high IVBA in this range of parameters is mainly located between 0.4-0.5eV. Within this energy interval, the part with larger strain but lower Sn content has larger IVBA. This is because large strain decreases DOS at the top of the valence band. This makes the quasi-Fermi level to move deeper into valence band. And since more holes can be found at the top of the valence band, this increases the probability for the absorption to take place between top valence band and the deeper SO band. Each point in the parameter space is calculated according to the band structure from the **k.p** method. Such calculation is time consuming, but from the pattern shown here, we can use the dependence found for $E_g^\Gamma = \Delta E_{so}^{lh}$ in bulk material to predict the rough combination of parameters where the high loss from IVBA could happen, also modified by some correction due to quantization effects. These properties are much easier

to obtain, one can almost immediately calculate this dependence from a few parameters of GeSn without solving the Schrödinger equation. The high loss ridge is rather stable in the parameter space for different carrier densities, with only some changes in the absorption values.

Finally, it it worth mentioning that, as indicated in [107], it seems the **k.p** calculation underestimate the IVBA from the experiment for GaAs. The reason for this was not discussed. But if we assume that this is also the case for GeSn, it is possible that even larger IVBA exists in reality, which makes it a more important aspect to be considered, because it could move some combinations of the strain and Sn content in the high loss region, although they may deliver high IBG.

4.7.2 Optimal Gain

For the same parameter space as in Sec. 4.7.1, in this section the peak of both inter-band gain and net gain at room temperature are shown. They are represented by the color in a different color space. Since the mesh grid in this parameter space is rather sparse, to obtain the optimal points for a particular energy, the interpolation method is used to find the maximum gain. Then, by comparing all the unit cells in the parameter space, the rough location of the optimal combinations regarding the gain at a particular transition energy can be found.

In Fig. 4.19, the left column (i.e. Fig. 4.19a, Fig. 4.19c and Fig. 4.19e) shows the inter-band gain peak. Similarly, the right column shows the net gain after including the IVBA and FCA. The three rows correspond to the 2D carrier densities 3×10^{12} cm^{-2}, 4×10^{12} cm^{-2} and 5×10^{12} cm^{-2}.

If only the inter-band gain is considered, the optimal points are arranged rather linearly in the parameter space. Some general pattens can be summarized: (1)The photon transition energy that can be realized in this parameter space is limited by the carrier density. In other words, the range of the emission wavelengths is generally restricted by the carrier density that can be achieved by injection or doping, the larger carrier density allows a gain at the larger photon energy (shorter wavelength) to be obtained. (2)For lower transition energy, the optimal points tend to have larger strain but smaller Sn content. When the photon transition energy is smaller than ∼0.33eV-0.35eV (depending on the carrier density), because of the limits imposed to of the parameter space, the optimal points will pile up at the edge of the graph (i.e. $\varepsilon = 1.7\%$). (3)The optimal points are also sensitive to the carrier density in the structure. Generally, for larger carrier density, the optimal points tend to shift to higher strain but lower Sn content. For

example, as can be seen in Fig. 4.19a, Fig. 4.19c and Fig. 4.19e, for 0.4eV, the optimal point for 3×10^{12} cm^{-2} is around $\varepsilon = 0.9\%$ and Sn = 9.2%, but for 5×10^{12} cm^{-2} the optimal point moves to $\varepsilon = 1.3\%$ and Sn = 7%. This makes the carrier density (or the current density) another uncertainty parameter which needs to be considered carefully when designing an optimized structure.

After including the absorption losses, the net gain peak is shown in Fig. 4.19b, Fig. 4.19d and Fig. 4.19f, for 2D carrier densities from 3×10^{12} cm^{-2} to 5×10^{12} cm^{-2}. It is obvious that the region where gain can be observed is suppressed by the losses. For purpose of comparison, the contour line of losses shown in Fig. 4.18 is also denoted here.

The loss mechanisms have an impact on the optimal parameters choice, especially between 0.4eV to 0.5 eV. For 3×10^{12} cm^{-2}, points with energy higher than 0.4eV vanish because no net gain is found due to the loss. For higher carrier density, 4×10^{12} cm^{-2} to 5×10^{12} cm^{-2}, since the high loss region 'penetrated' through the linear arrangement of the optimal points for the IBG, the optimal points of net gain are split into two parts. This effect is more obvious in the plot for 5×10^{12} cm^{-2}. For transition energy lower than $\sim$0.42eV, a relatively high strain $>\sim$1.3% and low Sn $\sim$ 0.65% is preferable. For photon energy higher than $\sim$0.42eV, lower strain $<\sim$0.7% and high Sn content $\sim$1% is preferable.

Although the FCA is relatively large for low photon transition energies, it will make no difference to the optimal choice, because in this model, FCA depends only on the wavelength, and not on strain and Sn content.

From the above results, it can be concluded that the loss mechanisms, especially from IVBA, are very important in the optimization of tensile strained SiGeSn/GeSn QWs. According to this calculation, this effect is extremely large and even larger than the IBG around 0.4-0.5 eV (2500nm - 3100nm wavelength). However, this has been largely ignored in the previous research on the GeSn QW lasers(e.g. [40, 42]). We should note that the calculations presented here are just theoretical predictions, and have not yet been compared to experiments. The results are also very sensitive to the empirical input parameters. So, more work needs to be done on the IVBA, both experimental and theoretical in the future, to validate the predictions of the theoretical calculations, and improve the reliability of the optimization results.

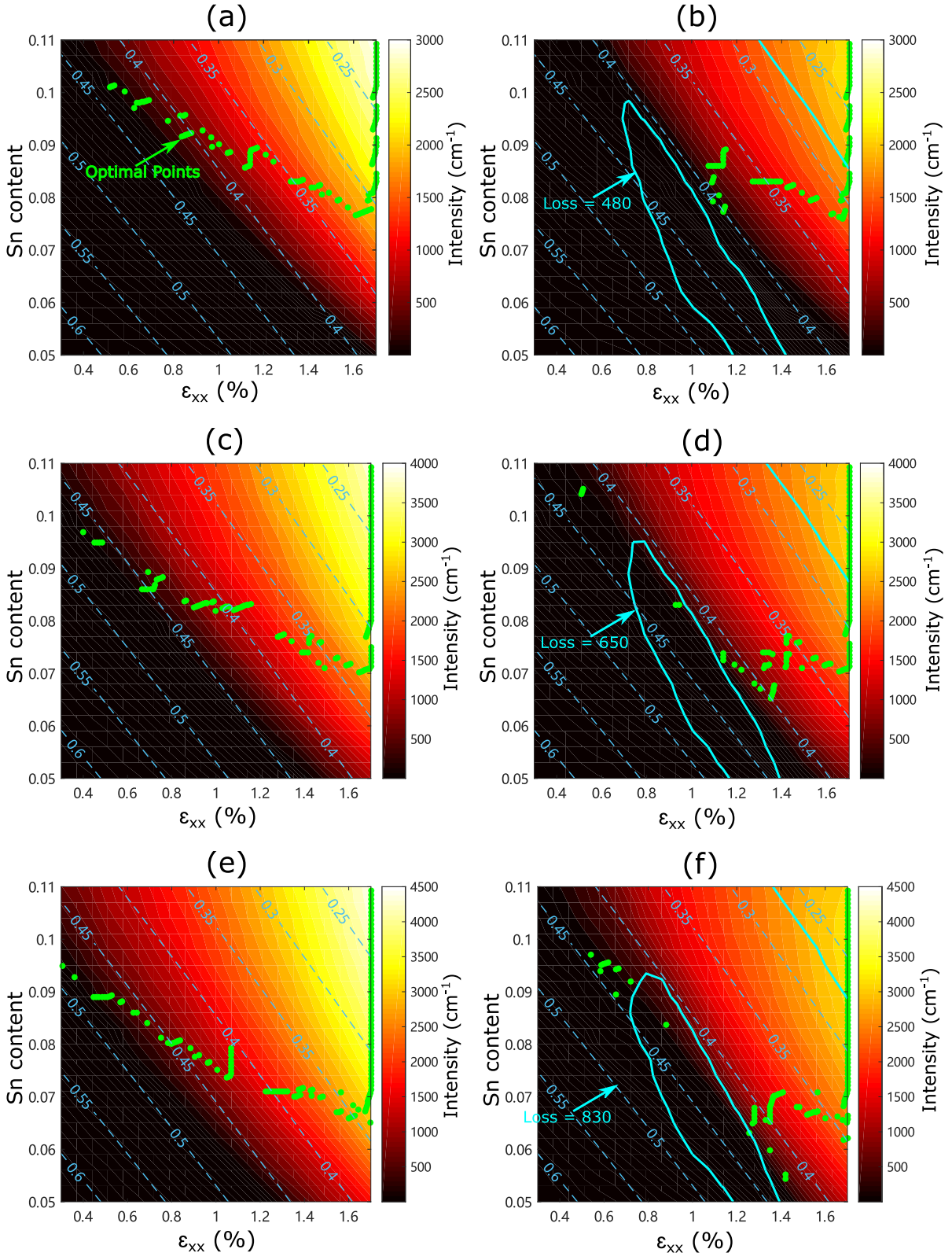

Figure 4.19: Peak inter-band gain and net gain for a 14nm $Si_{0.15}Ge_{0.81-y}Sn_{y+0.04}/Ge_{1-y}Sn_y$ in the parameter space of ε_{xx} and Sn content y at different carrier densities. The optimal points are marked with green dots. Energy contour line are in blue dashed line (a)IBG at 3×10^{12} cm^{-2}; (b)Net gain, at 3×10^{12} cm^{-2}; (c)IBG, at 4×10^{12} cm^{-2}; (d)Net gain, at 4×10^{12} cm^{-2} (e)IBG, at 5×10^{12} cm^{-2}; (f)Net gain, at 5×10^{12} cm^{-2}

.

To present some more detailed information, we will now focus on structures with transition energy of 0.41eV (3000 nm wavelength) to see how the performance of a structures changes with different choices of strain and Sn content in the well. We assume the strain is introduced by a buffer layer, then in order to introduce a tensile strain the buffer layer needs to have a larger lattice constant than the well layer. This can be achieved by a GeSn substrate with larger Sn content than the in well layer. In Fig. 4.20a, for a 14nm SiGeSn/GeSn QW, the dependence of Sn content in the well and buffer layer on the in-plane strain ε_{xx} in the well, needed to achieve 0.41eV transition energy is given. The Sn content in the well decreases with increasing strain, while the change of Sn content in the buffer layer is relatively small, from 0.6% to 1.6%, about ∼15% Sn is needed in the buffer layer. In Fig. 4.20b two figures of merit, the ΔE_{hh1}^{lh1} and $\Delta E_{\Gamma 1}^{L1}$ are given. They roughly describe the utility of such structure to have inter-band gain. A linear-like dependence is found for both of them, the difference between LH1 and HH1 increases from 50meV to 180meV for strain in the well between 0.6% and 1.6%. While the effective directness of the structure decreases from 100 meV to 59 meV in this interval, this opposite trend of the two figures of merit shows the trade off problem, and leads to an optimal point where the largest inter-band gain can be found.

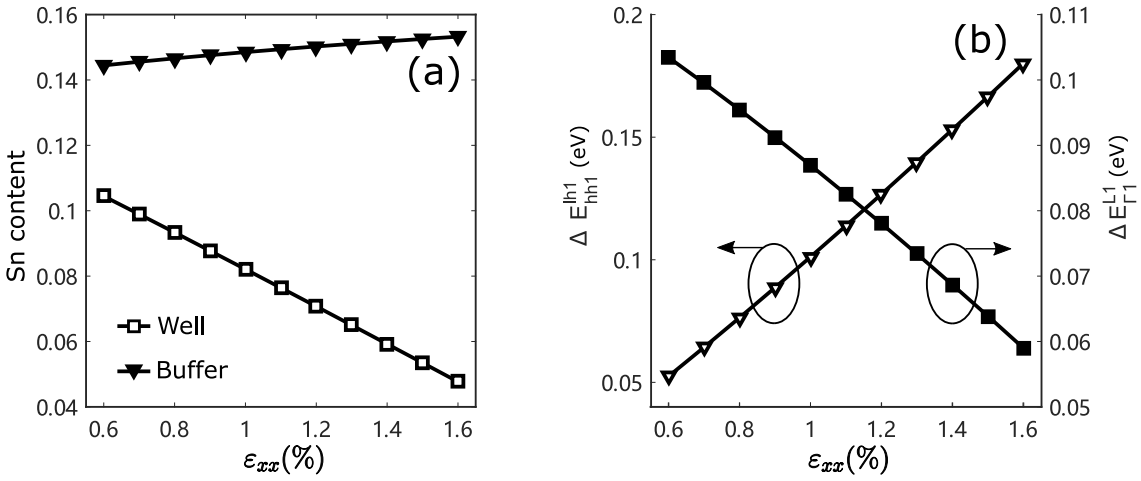

Figure 4.20: Material parameters and figures of merit for a 14nm $Si_{0.15}GeSn_{y+0.04}/GeSn_y$ quantum well with photon transition energy of 0.41eV. (a)The dependence of Sn content in the well and barrier on the in-plane strain ε_{xx} in the well to achieve the same transition energy (0.41eV); (b)ΔE_{hh1}^{lh1} and $\Delta E_{\Gamma 1}^{L1}$ against strain in the well. The corresponding Sn contents in the well are the same as in Fig.(a).

.

Fig. 4.21 shows the peak gain dependence on the strain for carrier densities of $3{\times}10^{12}$ cm^{-2} to 5×10^{12} cm^{-2}. The Sn contents correspond to different in-plane strains, needed to achieve

0.41eV transition energy, same as shown in Fig. 4.20a. As shown in Fig. 4.21, the maxima of inter-band gain are located at 0.84%, 1.01% and 1.11% in-plane strain for carrier density from 3×10^{12} cm^{-2} to 5×10^{12} cm^{-2}. This indicates that the influence of the two figures of merit weigh differently with varying carrier density. For lower carrier densities, the optimal structure tends to have a relatively larger $\Delta E_{\Gamma 1}^{L1}$ but smaller ΔE_{hh1}^{lh1}, thus a lower optimal strain is found. But for higher carrier densities a larger ΔE_{hh1}^{lh1} and lower $\Delta E_{\Gamma 1}^{L1}$ is preferred, this leads to a larger optimal strain. Moreover, by including the loss mechanism, it can be seen that the position of the optimal net gain is significantly changed. No net gain can be observed for carrier density 3×10^{12} cm^{-2}, while for 4×10^{12} cm^{-2} and 5×10^{12} cm^{-2}, since the curve of the loss vs. in-plane strain has a peak at a lower value of ε_{xx} than the peak gain, the maximum of the net gain is then shifted to higher ε_{xx}, and the maximum net gains are then found for strain values 1.2% and 1.41% respectively.

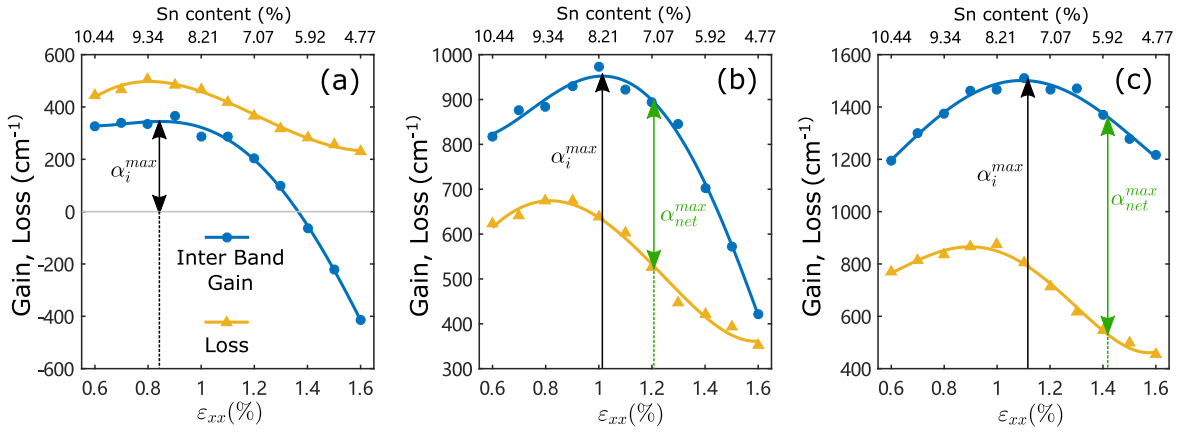

Figure 4.21: The inter-band peak gain and loss at this peak coming from IVBA and FCA for 14nm $Si_{0.15}GeSn_{y+0.04}/GeSn_y$ QW with 0.41eV transition energy obtained by different combinations of strain and Sn content in the well. (a)3×10^{12} cm^{-2}; (b)4×10^{12} cm^{-2}; (c)5×10^{12} cm^{-2}

.

Next, some mathematical explanations are given based on the above results. As can be seen in Fig. 4.21, the inter-band gain has a trend of increasing at lower strain and deceasing after it reaches a peak. But if we check the difference between the conduction band and valence band quasi-Fermi levels ($E_{fc} - E_{fv}$), a monotonic dependence is found as can be seen in Fig. 4.22a. These values actually decrease with increasing strain. The reason for this paradox comes from the asymmetric band structure of conduction band and valence band (i.e. different effective mass). Here we use a simplified model to explain this paradox. Assuming parabolic conduction

and valence bands, and setting the linewidth to zero, the peak gain can be calculated as:

$$g_p = g_0[f_c(E_{e1}^{k=0}) - f_v(E_{h1}^{k=0})] \tag{4.17}$$

where:

$$g_0 = C_0 \left|\hat{e} \cdot \mathbf{M}\right|^2 \rho_r^{2D} \tag{4.18a}$$

$$C_0 = \frac{\pi {q_e}^2}{n_r c \epsilon_0 {m_0}^2 \omega}, \quad \rho_r^{2D} = \frac{m_r^*}{\pi \hbar^2 L_z} \tag{4.18b}$$

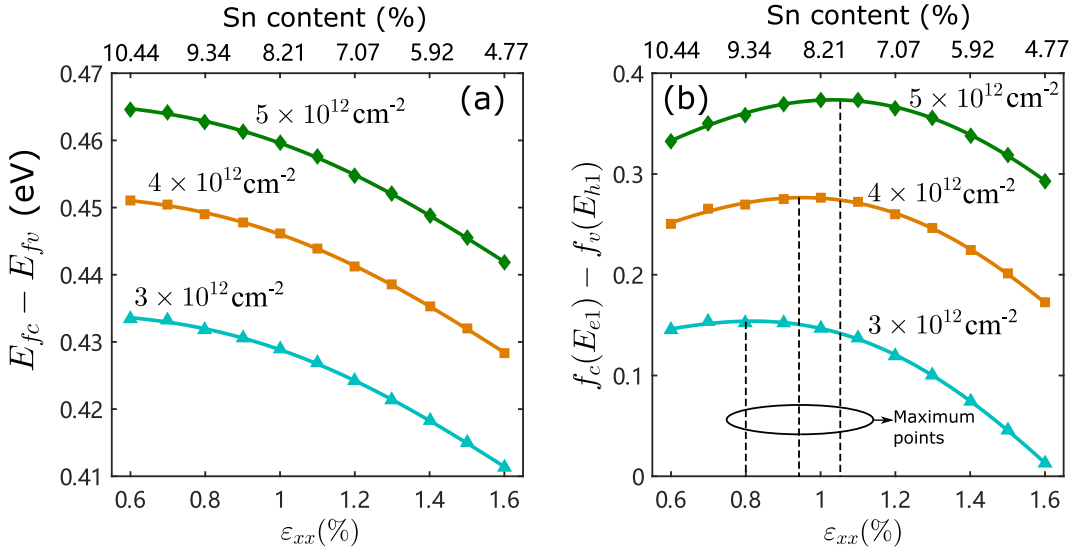

Figure 4.22: The properties of Fermi-Dirac distribution for carrier densities from 3×10^{12} cm^{-2} to 5×10^{12} cm^{-2} for 14nm $Si_{0.15}GeSn_{y+0.04}/GeSn_y$ QW with 0.41eV transition energy, obtained by different combinations of strain and Sn content in the well. (a)The difference between quasi-Fermi levels; (b)The Fermi-Dirac population inversion factor (with opposite sign).

.

To further simplify the model, we assume that the reduced effective mass m_r^* will not change a lot by varying the strain and Sn content in the well in a limited range to achieve a constant transition energy, and we take the effective mass to have a constant value. Hence g_0 becomes constant. Then the only parameter that determines the peak gain in the gain spectrum with zero linewidth is the Fermi-Dirac population inversion factor (with opposite sign i.e. $f_c - f_v$). The normalized peak gain is simply equal to $f_c - f_v$, as denoted in Fig. 4.23a. In some other scenario, like increasing the injection carrier density, larger difference of the quasi-Fermi levels $E_{fc} - E_{fv}$ will usually mean larger $f_c - f_v$. But this is not the case here. When increasing the

strain and reducing the Sn content in the well layer, it is possible that a structure with a higher ε_{xx} and lower Sn content has larger $f_c - f_v$, but smaller $E_{fc} - E_{fv}$, examples are plotted in Fig. 4.23a with red and green lines. This is because the conduction band in-plane effective mass is several times smaller than that of the LH band. For a structure with low strain and high Sn content in the well, the structure is already very 'direct', the conduction band quasi-Fermi level will be very sensitive to the carrier density, but the valence band Fermi level will be less sensitive due to the large effective mass. When the Bernard-Duraffourg population inversion condition is fulfilled, the quasi-Fermi level for holes will usually be close to but above the valence band top. But the electron quasi-Fermi level in this scenario will be much higher than the conduction band bottom. The schematic diagram of such a situation is plotted in Fig. 4.23b. The value of $f_c(E_{e1}^{k=0})$ is here almost 1, while the value of $f_v(E_{h1}^{k=0})$ is in the mid-range, around 0.5. Consider the figures of merit plotted in Fig. 4.20. When the strain in the well is increased, the larger ΔE_{H1}^{L1} and smaller ΔE_{hh1}^{lh1} leads to both electron and hole quasi-Fermi levels moving downwards. It is clear, because of their original location, that the slope or the derivative of Fermi-function against energy is different for electrons and holes, df_c/dE is much smaller than df_v/dE. As we can see in Fig. 4.22a, $E_{fc} - E_{fv}$ is decreasing with increasing the strain, indicating that E_{fc} deceases faster than E_{fv} per unit change of the strain. Then, with a small increment of strain the change in quasi-Fermi levels is also small, the paradox we mentioned above will appear when $\Delta E_{fc}/\Delta E_{fv} < df_v/df_c$ is fulfilled, $\Delta f_v = f_v - f_v' > f_c - f_c' = \Delta f_c$. Thus $f_c' - f_v' > f_c - f_v$ which leads to an increase of gain but a decrease of the value $E_{fc} - E_{fv}$ (from the green curve to the red curve in Fig. 4.23a). However, when the strain is already very large, then we can expect a low DOS at the valence band top, but a less direct conduction band. This will lead to the quasi-Fermi level closer to the conduction band bottom, and $\Delta E_{fc}/\Delta E_{fv} < df_v/df_c$ will no longer hold. Then further increasing of strain from this point will bring about a decrease of gain (from the red curve to the blue curve in Fig. 4.23a).

The calculated $f_c - f_v$ in this case is shown in Fig. 4.22b. The maxima of these plots for different carrier densities coincide with the maximum inter-band peak gains in Fig. 4.21. The small difference may come from the change of the effective mass for different combinations of strain and Sn content, and from the introduction of the linewidth from Lorentzian convolution.

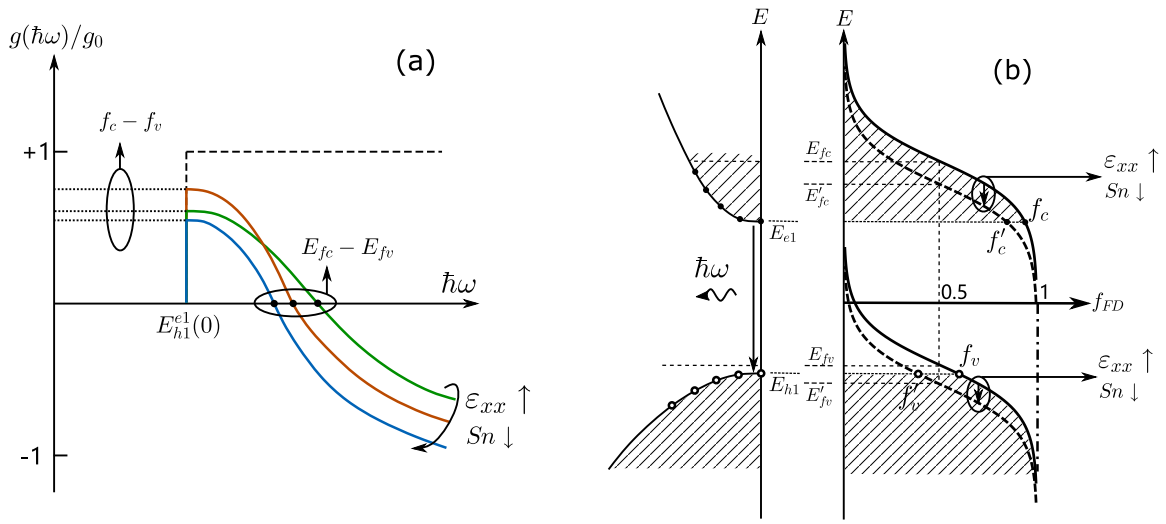

Figure 4.23: (a)Normalized gain spectrum with zero linewidth for different combinations of Sn content and strain in the well; (b)Band structure and Fermi-Dirac distribution, for the tensile strained structure. Increasing strain and reducing the Sn content to maintain the same transition energy is equivalent to decreasing the quasi-Fermi level for both electrons and holes.

.

4.8 Performance Estimation

Lastly, in order to give a rough idea about the practical performance of such a laser, we calculate the threshold current density for the optimal combination of parameter obtained above, for a 0.41 eV photon emission energy with $5 \times 10^{12}\text{cm}^{-1}$ carrier density, for a particular waveguide design. The injection current density can be derived from the recombination equation which describes the carrier dynamics:

$$\frac{dN}{dt} = G_c - A_{nr}N - R_\Gamma P(N\Theta_\Gamma) - R_L P(N\Theta_L) - C_n N^2 P - C_p N P^2 \tag{4.19}$$

where G_c is the carrier generation rate due to injection; N and P are the volume carrier densities of electrons and holes; A_{nr} is the non-radiative recombination rate; R_Γ and R_L are the direct and indirect recombination coefficients, C_n and C_p are the Auger recombination rates of electrons and holes.

In steady state in Eq. 4.19 $dN/dt = 0$, we assume $G_c = \eta_i J/(q_e d)$, then the relation between the injection current density and carrier density in Eq. 4.19 can be written as:

$$\eta_i J = q_e d \left[A_{nr}N + R_\Gamma P(N\Theta_\Gamma) + R_L P(N\Theta_L) + C_n N^2 P + C_p N P^2 \right] \tag{4.20}$$

in which η_i is the injection efficiency; J is the current density; q_e is the electron charge ; d is the length of the active region; Θ_Γ and Θ_L is the percentage of electrons in Γ and L valleys. The value of the recombination rates are taken as the values in pure Ge, these parameters are given in Table. 4.6.

Table 4.6: Recombination coefficient of Ge

Parameter (Unit)	A_{nr} (s^{-1})	R_Γ ($cm^{-3}s^{-1}$)	R_Γ ($cm^{-3}s^{-1}$)	C_n ($cm^{-6}s^{-1}$)	C_p ($cm^{-6}s^{-1}$)
Value	10^{-8}[a]	1.3×10^{-10}[b]	5.1×10^{-15}[b]	2.2×10^{-30}[c]	5.4×10^{-31}[c]

[a] Reference [85]
[b] Reference [46]
[c] Reference [108]

The threshold carrier density is calculated by making the modal gain equal to the modal loss plus the mirror loss, i.e. $G_m = L_m + L_{mi}$. The modal gain is calculated as the net gain times the optical confinement factor, $G_m = g \times \Gamma$. Here the waveguide described in [85] is used, with the active region replaced by our optimal results. The schematic diagram of the waveguide is shown Fig. 4.24. With 14nm well and 9.6 nm barrier the 4×MQW active region has the same thickness of 104 nm as that in [85]. Here, as a rough estimate we use the data in [85] and the optical confinement factor of the QWs is 6.98%, the sum of the modal loss and mirror loss is 82.16 cm^{-1}, and the injection efficiency is set as 75%. Again, since only the first subband is here used for lasing, and the TE gain from LH is very small compared with the TM mode gain, thus the TE mode is not considered here.

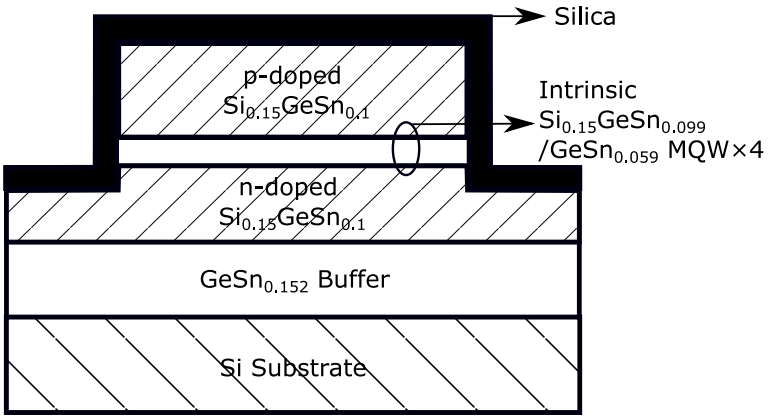

Figure 4.24: Schematic diagram of the waveguide design, the ridge is wrapped by silica in order to provide better optical confinement.

.

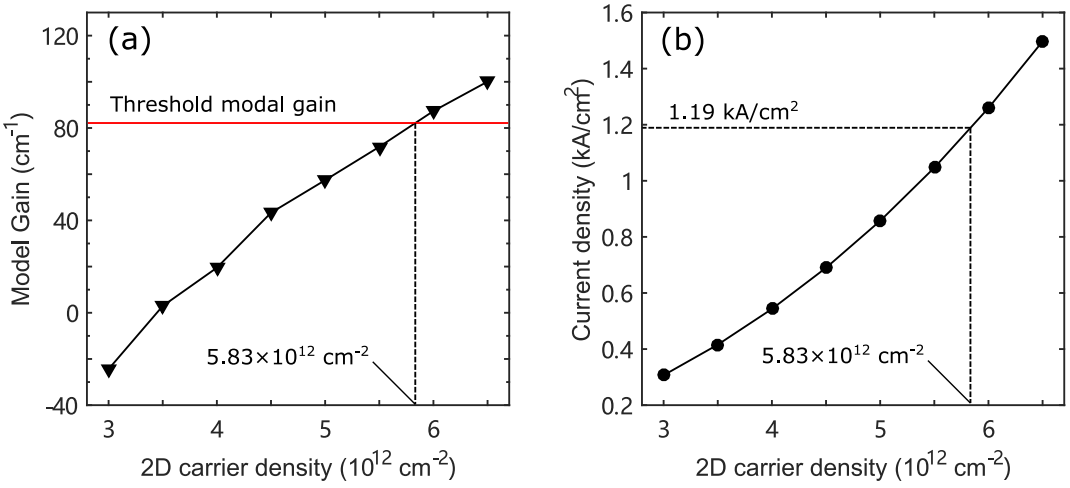

Figure 4.25: Lasing threshold calculation for $Si_{0.15}GeSn_{0.99}/GeSn_{0.059}$ MQW with 1.41% tensile strain in the well (a)peak model gain versus 2D carrier density. (b)current density versus 2D carrier density

.

Fig. 4.25 shows the calculation results for the threshold current density of $Si_{0.15}GeSn_{0.99}/GeSn_{0.059}$ MQW with 1.41% tensile strain in the well. As shown in Fig. 4.25a, the threshold 2D carrier density is $5.83 \times 10^{12} cm^{-2}$. The dependence of current density on the 2D carrier density of this structure is shown in Fig. 4.25b. The $5.83 \times 10^{12} cm^{-2}$ carrier density requires the threshold current density of 1.19kA/cm^2. This result is comparable to some conventional Group III-V lasers, for example 1.05kA/cm^2 for a $Al_{0.09}GaAs/GaAs$ 10×MQW laser [109], or 4.2 kA/cm^2 for $Al_{0.03}GaAs/GaAs$ 10×MQW laser in continuous wave operation [110].

4.9 Conclusion

Based on **k.p** method and effective mass method, the band structures of Γ and L valley states in QWs are calculated. The optimization of the tensile strained SiGeSn/GeSn QW laser is considered in several aspects. Firstly, in order to give a reasonable carrier confinement, the band alignment in the parameter space is done for the LH. It was found that 15% Si in the barrier and 4% Sn more than in the well layer provides a reasonable type I band alignment with $\sim$ 55meV discontinuity. Then by doing gain calculation with varying well width in different structures, 13-16 nm was found to be a good choice to get high TM mode gain. After that, the inter-valence band absorption was found to be an important loss mechanism, although it was rarely discussed in early research on modelling the SiGeSn/GeSn laser. High IVBA can be observed when the spacing between the top valence band and split-off band is equal to

the bandgap. For a 14nm $Si_{0.15}Ge_{0.81-y}Sn_{y+0.04}/Ge_{1-y}Sn_y$ QW, we calculated the loss at the peak gain in the parameter space of Sn content and strain in the well. A high loss ridge was found due to the IVBA, this ridge is in parallel and close to the link of Sn and strain when $E_g^\Gamma = \Delta E_{so}^{lh}$. For different parameters combinations, there are trade-offs regarding the difference between Γ, L valley and LH, HH. In order to find the optimal combination, the peak gain in the same parameter space is also calculated, and optimal points are rather continuous when only inter-band gain is considered. The optimal points were also found to be sensitive to the carrier density. However after including the losses, a great change is observed for these optimal points. The combination of Sn content and strain where the gain is observed is suppressed by the loss mechanisms, and the lines of optimal points are also separated by the high loss region. A more detailed calculation was done for the structures with 3μm emission wavelength. With the link of strain and Sn content giving 0.41eV(3μm) transition energy, a clear peak is observed for the gain against the strain(Sn content). This looks somewhat contradictory with the monotonic dependence of difference between electron and hole quasi-Fermi levels, but because of different properties of conduction and valence band, it is possible for some structure to deliver a larger gain with a smaller difference of electron and hole quasi-Fermi levels. It was found that the dependence of Fermi-Dirac inversion factor on strain agrees very well with the inter-band gain, which further supports this view. The optimal values of strain, after including the loss for 0.41eV photon transition energy are $\sim 1.2\%$ for 4×10^{12} cm^{-2} and $\sim 1.4\%$ for 5×10^{12} cm^{-2} carrier density. No net gain was found for 3×10^{12} cm^{-2}. Finally, with the waveguide design combined with the optimal design for 5×10^{12} cm^{-2}, the threshold current density for such MQW structure is 1.2 kA/cm^{-2}. This is comparable with some conventional III-V laser designs.

Chapter 5

Intra-step SiGeSn/GeSn Quantum Well Electroabsorption Modulators

5.1 Introduction

With the fast development of Silicon-based electronic-photonic integrated circuits (EPICs), light sources have been demonstrated in the past few years, especially optically pumped and electric pumped laser e.g. [16, 20]. Here we consider the corresponding on-chip modulators with similar material composition and wavelengths. These light sources are using GeSn as a lasing material, and they usually have a larger wavelength (>2000 nm) than the standard telecommunication wavelength where the optical fibers have low transmission loss and signal distortion (1310-1550nm). Previous research on the electroabsorption modulators both theoretical (e.g. [70, 85, 111, 112]) and experimental (e.g. [9, 105, 113]), focused mainly on the traditional wavelengths, and mostly used Ge as the active material. Recently, as GeSn is attracting increasing attention as a material for modulators in the MIR region [27, 29, 30, 48], this research is promising for on-chip transceiver design, in line with the popular GeSn laser research nowadays.

QWs based structures have already shown their superiority over the bulk material for both Group IV and III-V material [30, 114, 115]. Based on the quantum confined Stark effect (QCSE), a much larger absorption change is observed when varying the applied field, which makes it possible for an electroabsorption modulator to work even at room temperature. In Group IV, many attempts have already been done to optimize the performance of the EAM. For

example, in [70] it was proposed to use a rather wide QW to increase the QCSE, thus a lower driving voltage is needed for modulation. The problem here is that, although the QCSE is strong for wide well, a lower absorption change and extinction ratio is also expected simultaneously, because the exciton effect is weaker in a wider well. Another example is [116] where, with strain engineering, a strong carrier confinement, and a larger absorption is achieved. But in this paper, it was also argued that a stronger carrier confinement may also inhibit the high-speed modulation. As can be seen from these examples, in optimizing a square QW structure, there are always some trade-offs between the extinction ratio, power consumption and modulation speed.

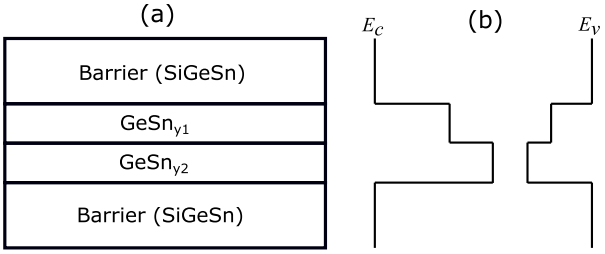

Figure 5.1: (a) Schematic diagram of the intra-step QW (Note that to form an intra-step, the Sn content $y_1 < y_2$), (b) The corresponding energy profile for conduction band and valence band.

Here, instead of focusing on the traditional square QW, we propose to improve the performance by a different geometry - the intra-step QW. Fig. 5.1 shows the schematic diagram of a typical intra-step QW. Such structure is not new for III-V material system, as it was found in the III-V system that intra-step QW has a much larger QCSE than a square QW [117, 118]. Another useful property of the intra-step QW is to delay the on-set of the red shift of QCSE to higher electric field [119], thus a high-optical power can be modulated. These two functionalities can be achieved by applying opposite directions of the electric field in asymmetric intra-step QW. The former is realised when the field direction is from the intra-barrier to the intra-well, and the latter is realised in the opposite case. In this work we focus on the former, and here a low driving voltage is required for on-chip devices. As is indicated in [120], for the intra-step QW the advantage of having a large QCSE may also be accompanied by a reduced absorption change or extinction ratio, indicating a careful optimization is needed to obtain a practical improvement of the overall performance. Indeed, by optimisation of SiGeSn/GeSn/GeSn/SiGeSn intra-step QW we have found the structure which has both larger QCSE and larger absorption change than a square QW. The intra-step QW is realized by using different Sn contents in

the GeSn/GeSn heterostructure. Some important figures of merit, such as quantum confined Stark shift ΔE_{QCSE}, the absorption coefficient change $\Delta\alpha$, the absorption change per unit field $\Delta\alpha/F$, and the absorption change per unit field squared $\Delta\alpha/F^2$ are investigated and compared to optimize the design.

In order to have the bandgap in the MIR region, the Sn contents in the two intra-layers should be in the range from 6% to 12%. Furthermore, 10% Si is used in the barrier to provide a reasonable carrier confinement. The substrate is assumed to be fully relaxed and is chosen so that the GeSn layers will have compressive strain and the SiGeSn barrier layer will have tensile strain.

Since the application of the structure is field-direction sensitive, we denote the two intra layers as layer-1 and layer-2, and assume the electric field is pointing from layer 1 to layer 2. Then the intra-step QW is formed by assuming the Sn content in the first layer (y_1) is smaller than that of the second layer (y_2). Then the depth of the intra-well can be controlled by varying the difference between y_1 and y_2 (Δy). To reduce the number of variables in calculations, we can either set one of the Sn contents (e.g. y_1) as constant and vary the other $y_2 = y_1 + \Delta y$, or set the mid-point y_m fixed, and then the Sn contents in the two layers are $y_1 = y_m - \Delta y/2$ and $y_2 = y_m + \Delta y/2$. Apart from the Sn content, the width ratio $R_w = w_1/w_2$ is another variable. Since we want to increase the QCSE by geometry instead of changing the well width, we assume the total length $w_1 + w_2$ to be same as for the square QW. Then, in terms of the figure of merit $\Delta\alpha/F$, the optimal structure in the parameter space is the one which is found to have both larger QCSE and absorption change than the square QW with the same total width. Another structure that have an even larger QCSE but a smaller absorption change is also used for comparison. By doing so we found the optimal choices could be different for different figures of merit. According to the results, a lower power consumption and a higher bandwidth can be achieved by adjusting (keeping) the extinction ratio at the minimal required value.

5.2 Modelling of Electroabsorption Modulator

In this section the computational model of the electroabsorption modulator is introduced and, different from the laser, the electric field needs to be considered as it is the key variable for quantum confined Stark effect (QCSE). Moreover, the exciton effect plays a vital role in QW

EA modulator, a peak can be found at the exciton transition energy, leading to large absorption change by the applied field. The variational method is used to calculate exciton binding energy and the effective Bohr radius. For the absorption, the 1s and the continuum contributions from excitons are treated separately. The parameters used in this model are given in [121]. The calculation of these parameters is the same as discussed in Section 4.3.

5.2.1 Band Structure

The 8-band **k.p** method is also used in this Chapter. With electric field perpendicular to the heterostructure plane(z-direction), the potential profile is replaced by $V(z) + q_e F z$. Then each $N \times N$ sub-matrix has $q_e F \times diag(z_1, z_2, \cdots, z_N)$ added to the diagonal. The band structure is obtained by finding the eigenvalues of the Hamiltonian:

$$\mathbf{H}_{8\times 8} = \mathbf{H}(\mathbf{k}) + \mathbf{D}(\mathbf{k}) + \mathbf{D}_{so} + q_e F z \tag{5.1}$$

The method of solving the Hamiltonian is already described in Chapter 3. It is worth to mention that because of the nature of the FDM, when electric field exists, a numerical problem appears, coming from the artificial QW in this system. In this model, one has to set a barrier width, which is finite, and to assume that the derivatives of wave functions are zero at the outer edges of the barriers. This is fine for zero electric field (i.e. square potential profile). However, the electric field will make an additional 'well' at one side of the barrier, as illustrated in Fig. 5.2.

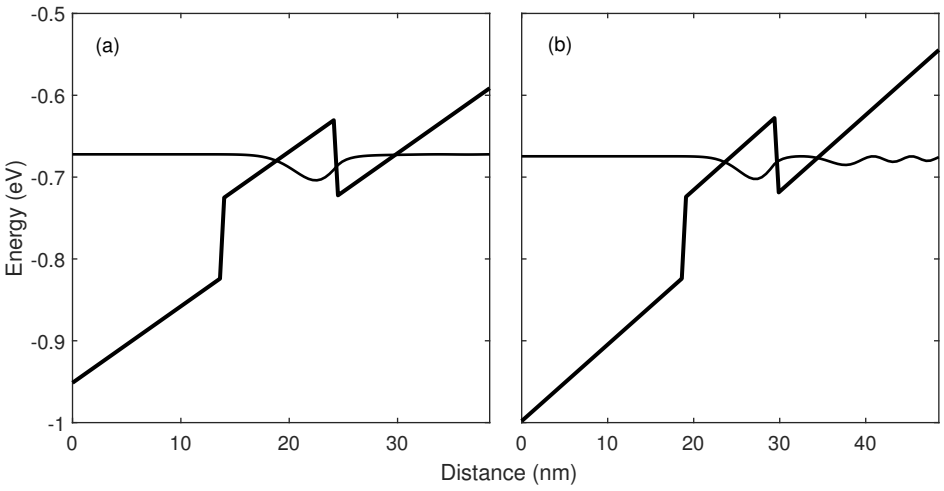

Figure 5.2: Example of numerical 'instability' appearing in the modelling

The two plots are both for the HH band structure of a 10 nm QW. The artificial well will not always cause trouble, as is shown in Fig. 5.2a. The wave function of the first heavy hole state looks ok when the barrier width is taken to be 14 nm. But when the barrier width is set to 19 nm, the bound state in the artificial well on the right side will be resonate with the wave function in the real well. This will not have a great impact on the state energy, however it will reduce the overlap with the conduction band wave function and reduce the contribution of this state to the absorption. Moreover, because of the incorrect wave function, the exciton binding energy and effective Bohr radius will be incorrectly calculated. Since this numerical instability only occasionally happens, it will not bring about a big trouble in the simulations. It can be easily identified by some sudden change in the exciton binding energy, or by checking the wavefunctions for each electric field. Then we just need to slightly change the outer barrier width (it should be distinguished from the actual barrier width in the MQW). The 'bound' states in this artificial well are then changed and no longer resonant with the wave function in the actual well.

5.2.2 Exciton effects

In order to be integrated into the existing 8-band **k.p** method, the variation method is applied [52, 122]. It is worth to mention that some other methods may also be used for this problem, such as many-body theory [116] and Green's function formalism [123]. All these methods showed reasonable agreement with the experiments [88, 112, 116, 123].

To start with, the wave function of 1s state has the general form:

$$\phi(\boldsymbol{\rho}) = \sqrt{\frac{2}{\pi}}\frac{1}{\lambda}\exp\left(\frac{-\rho}{\lambda}\right) \tag{5.2}$$

where $\boldsymbol{\rho}$ is vector representing the separation between electron and hole in the 2D plane of the heterostructure, λ is the variational parameter, also called effective Bohr radius. The 1s-like states are generally observed in experiment, the higher, for example 2s, states of exciton are hardly spotted [122]. Therefore in this model only 1s state is considered. The variational parameter is determined by minimizing the exciton binding energy:

$$E_{ex}(\lambda) = \left\langle \phi \middle| -\frac{\hbar^2}{2m_r}\nabla_\rho^2 - V(\rho) \middle| \phi \right\rangle$$

$$= \frac{\hbar^2}{2m_r\lambda^2} - \frac{q_e^2}{4\pi\epsilon_s}\int dz_e \int dz_h |f_n(z_e)|^2 |g_m(z_h)|^2 \left(\frac{2}{\lambda}\right) G\left(\frac{2|z_e - z_h|}{\lambda}\right) \tag{5.3}$$

in which $V(\rho)$ is the interaction potential, q_e is the electron charge, ϵ_s the permittivity of semiconductor, and m_r is the reduced effective mass which is calculated as:

$$m_r = \frac{m_e^* m_{h,||}^*}{m_e^* + m_{h,||}^*} \tag{5.4}$$

The in-plane heavy hole and light hole effective masses are calculated from the Luttinger parameters using Eq. 2.40a, z_e and z_h are the coordinates in the growth direction, f_n and g_m are the nth and mth state wave function of conduction band and valence band. The square of the wave function is obtained by summing up all eight components of the normalized wave function when $k = 0$. Finally, $G(x)$ is defined as:

$$G(x) = \int_0^\infty dt \frac{\exp(-xt)}{(1+t^2)^{3/2}} \tag{5.5}$$

With this integration it would be time consuming to consider the exciton problem, so here we used the polynomial approximation proposed in [122], and $G(x)$ is then written as:

$$G(x) = \begin{cases} 1 + Ax + Bx^2 \ln\left(\frac{x}{2}\right) + Cx^2 + Dx^3 + Ex^4 \ln\left(\frac{x}{2}\right), & (0 < x \leqslant 6.8) \\ \frac{1}{x} - \frac{3}{x^3} + \frac{45}{x^5} - \frac{1575}{x^7}, & (x > 6.8) \end{cases} \tag{5.6}$$

The values of the coefficients are listed in Table.5.1.

Table 5.1: Values of coefficients in Eq. 5.6

Coefficient	Value
A	-8.9707×10^{-1}
B	-2.5262×10^{-1}
C	2.2576×10^{-1}
D	3.2372×10^{-2}
E	-4.1369×10^{-4}

The G-function is plotted in Fig. 5.3. The discontinuity when $x = 6.8$ is ignorable. According to [122] the error of this approximation is only 1.5%, but it greatly improves the calculation speed. Then, using this approximation, the dependence of the variational parameter can be calculated. The $Si_{0.15}Ge_{0.85}$/Ge QW is used as an example here. As can be seen in Fig. 5.4a,

the e1-hh1 exciton binding energy has a minimum of ∼-4 meV when λ is around 22 nm. This minimum can be found automatically in the program by using, for example, bisection method to find the zero of the first derivative. The result in our work is very close to that in [124], as λ was found to be ∼23 nm for a Ge QW when no electric field is applied. Fig. 5.4b shows the dependence of binding energy and effective Bohr radius on the applied field in the structure. The exciton binding energy decreases and the effective Bohr radius increases with increasing field. This is because, for non-zero electric field, the wave functions of electrons and holes will move in opposite directions. This separation weakens the exciton binding, so the energy is reduced and the in-plane extension is increased.

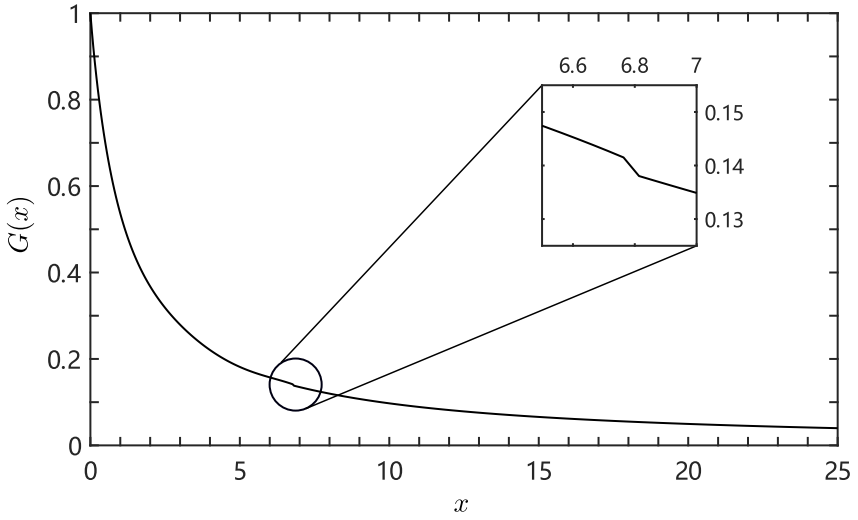

Figure 5.3: Plot of $G(x)$ using the polynomial approximation

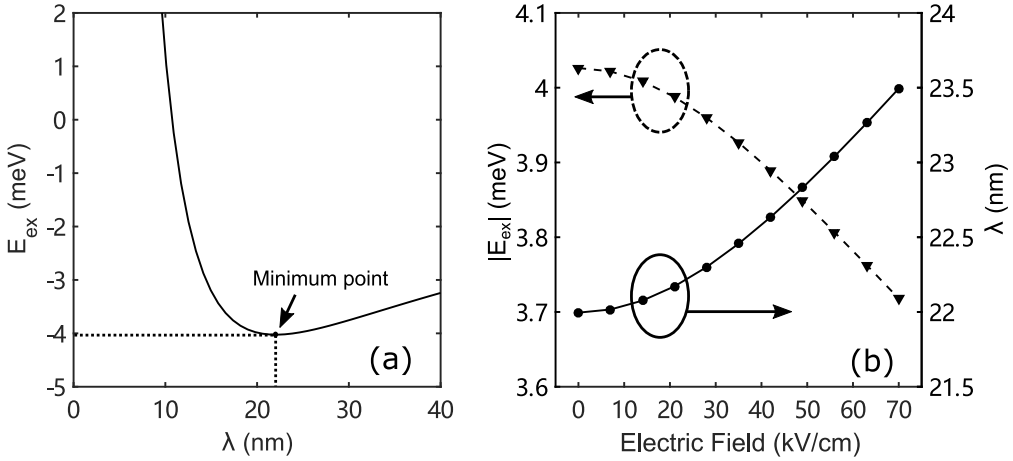

Figure 5.4: Exciton solution for a 10nm $Si_{0.15}Ge_{0.85}$/Ge QW. (a)The dependence of exciton binding energy E_{ex} (e1-hh1) versus the variational parameter λ; (b)Absolute value of the exciton binding energy of HH1 and effective Bohr radius against the electric field (0-70 kV/cm) applied to the structure.

5.2.3 Absorption Spectrum

The general formula for absorption spectrum is given in Chapter.3 (Eq.3.15). However, for the modulator application, we have to take the exciton contribution into account. Hence the absorption equation is modified to include the contribution from the excitons. Following the derivation in [52], the 1s exciton contribution to the total absorption spectrum is obtained by summing contributions of each pair of electron and hole subbands, and the expression for this part reads:

$$\alpha_{1s}(\hbar\omega) = \frac{\pi q_e^2}{n_r c \epsilon_0 m_0^2 \omega L_z} \sum_{n,m} \left(\frac{2}{\pi \lambda_{nm}^2}\right) \left|\left\langle \Psi_n^{k=0} \middle| \hat{e} \cdot \mathbf{p} \middle| \Psi_m^{k=0} \right\rangle\right|^2 V(\hbar\omega - E_x^{nm}, \Gamma_L, \Gamma_G) \tag{5.7}$$

the constants q_e, n_r, c, ϵ_0, m_0 have the same definition as those in Eq. 3.15. The L_z here is defined as the period of the MQW, i.e. the sum of barrier and well layer widths. n and m are the conduction and valence subbands indices respectively. Again, here the degenerate bands for spin up and down are treated separately. The momentum matrix elements are here calculated at the center of the Brillouin zone. E_x is the exciton transition energy, defined as:

$$E_x^{nm} = E_{ex} + E_{en} - E_{hm} \tag{5.8}$$

Since the exciton binding energy is negative, the transition energy is slightly smaller than the effective band gap of the QW. Here, instead of Lorentzian or Gaussian profile, Voigt profile is used. Assuming the FWHM of the Lorentzian is Γ_L, and the FWHM for Gaussian is Γ_G, the Voigt profile is defined as the convolution of Lorentzian line-shape with a Gaussian function:

$$V(x, \Gamma_L, \Gamma_G) = \int_{-\infty}^{\infty} G(x', \Gamma_G) L(x - x', \Gamma_L) dx' \tag{5.9}$$

The Lorentzian line shape function $L(x, \Gamma_L)$ is given by Eq. 3.19, and the Gaussian line-shape reads:

$$G(x, \Gamma_G) = \frac{1}{\sqrt{\ln(2)^{-1}}\pi(\Gamma_G/2)} \exp\left(\frac{-x^2}{\ln(2)^{-1}(\Gamma_G/2)^2}\right) \tag{5.10}$$

This time consuming calculation can be simplified by the method introduced in [125]. Then instead of doing the integration, the following expression is used as an approximation:

$$V(x,\Gamma_L,\Gamma_G) = \frac{\Gamma_L}{\Gamma_G} a_L \sqrt{\pi \ln(2)} \sum_{i=1}^{4} \frac{C_i(Y-A_i) + D_i(X-B_i)}{(Y-A_i)^2 + (X-B_i)^2} \tag{5.11}$$

in which a_L is the amplitude of the Lorentzian line-shape. For a normalized Lorentzian line-shape $a_L = 2/(\pi\Gamma_L)$. The input parameters X and Y are defined as:

$$\begin{aligned} X &= \frac{2\sqrt{\ln(2)}}{\Gamma_G} x \\ Y &= \frac{\Gamma_L}{\Gamma_G} \sqrt{\ln(2)} \end{aligned} \tag{5.12}$$

and the input parameter of A_i, B_i, C_i and D_i are given in Table.5.2.

Table 5.2: Parameters for the Voigt function approximation [125]

i	A_i	B_i	C_i	D_i
1	-1.2150	1.2359	-0.3085	0.0210
2	-1.3509	0.3786	0.5906	-1.1858
3	-1.2150	-1.2359	-0.3085	-0.0210
4	-1.3509	-0.3786	0.5906	1.1858

Using the Voigt profile, both homogeneous broadening described by Lorentzian and inhomogeneous described by Gaussian can be included simultaneously. For the homogeneous broadening, it depends on the electron scattering time in the Γ valley. Because of the lack of data for GeSn and SiGeSn, the scattering time in Ge is used here. As is given in [126], the Γ valley scattering time for Ge is 110fs. Then, according to the Heisenberg time-energy uncertainty relation $\Delta E \Delta t \geqslant \hbar/2$, the corresponding half width half maximum is 3 meV, so $\Gamma_L = 6$ meV. It is also indicated in [127] that its dependence on temperature is rather weak. For the inhomogeneous broadening, it is influenced by aspects of epitaxial growth such as structural fluctuations, defects, and it also usually depends on the electric field. For the simulation, it is hard to determine the exact fabricating conditions, so Γ_G is a rather arbitrary value here. Thus, the linear variation of the Γ_G between 6.64 and 10 meV from 0 to 160 kV/cm is used [122].

The indirect absorption is calculated using a Ge bulk model [128, 129], also used in [70]. The

expression of indirect absorption is:

$$\alpha_{ind}(\hbar\omega) = A_{obs}\left[\frac{1}{1-\exp\left(-\frac{E_{ph}}{k_bT}\right)}\left(\frac{\hbar\omega - E_g^L - E_{ph}}{\hbar\omega}\right)^2 + \frac{1}{1-\exp\left(\frac{E_{ph}}{k_bT}\right)}\left(\frac{\hbar\omega - E_g^L + E_{ph}}{\hbar\omega}\right)^2\right] \tag{5.13}$$

where the constant $A_{obs} = 1150\ \mathrm{cm}^{-1}$ is determined in [129] by fitting the results from Eq. 5.13 to the experiment, k_b is the Boltzmann constant, T is the temperature in K. The Ge optical phonon energy E_{ph} is 27.8 meV [128]. Although this approximation of using the expression for bulk Ge will deliver a smaller absorption than realistic in QW, the indirect absorption contribution is small compared to the total absorption, which makes it less important for the overall result, and if Sn is added into the well region, the increased directness of such structure will further reduce the indirect absorption. The purpose of this component is to add some insertion loss when photon energy is smaller than the bandgap, together with the tail of the line shape.

As for the contribution of continuum states of exciton, it can be related to the inter-band absorption spectrum by the Coulomb enhancement factor S_{mn}:

$$\alpha(\hbar\omega) = \frac{\pi {q_e}^2}{n_r c\epsilon_0 {m_0}^2 \omega L_z}\sum_{n,m}\iint \frac{dk_x dk_y}{(2\pi)^2} S_{nm}\left|\langle\Psi_n|\hat{e}\cdot\mathbf{p}|\Psi_m\rangle\right|^2 V(\hbar\omega - E_{hm}^{en}(\mathbf{k}), \Gamma_L, \Gamma_G) \tag{5.14}$$

in which, $E_{hm}^{en}(\mathbf{k})$ is the transition energy between nth conduction band and mth valence band when the wavevector is equal to $\mathbf{k}$. Note that, as explained in Section.3.3, the absorption is first calculated using triangle simplex method, then convolution is done with the Voigt profile. The total absorption spectrum is then the sum of Eq. 5.7, Eq. 5.13 and Eq. 5.14. The Coulomb enhancement factor for a pure 2D exciton is:

$$S_{nm} = \frac{2}{1+\exp\left(\frac{-2\pi}{\sqrt{E_t(\mathbf{k})/R_y}}\right)} \tag{5.15}$$

in which the Rydberg energy is defined as $R_y = m_r q_e^4/(2\hbar^2(4\pi\epsilon_s)^2)$, and $E_t(\mathbf{k}) = (\hbar k)^2/(2m_r)$. The Coulomb enhancement factor has the value of 2 that gradually approaches to 1 when energy tends to infinity. However, as the quantum well is not a pure 2D structure but has an effective

dimensionality between 2D and 3D, using Eq. 5.15 will overestimate the continuum contribution. In fact, as indicated in [123], using the pure 2D enhancement factor will make continuous part of the absorption spectrum of a 10nm GaAs QW overestimated by about 50% and result in a dip at the onset of the continuum which is not observed in the experiment. For this reason, assuming that such overestimation also exists in Group IV QW, the Coulomb enhancement factor in our case takes a constant value of 1.33 because we are only interested in the absorption near the first exciton peak. The results calculated by both Eq. 5.15 and with $S_{nm} = 1.33$ are shown in Fig. 5.5 for a 10nm Ge/$Si_{0.15}$Ge QW. Comparison with the experimentally measured absorption from [105], shows that using 1.33 as the Coulomb enhancement factor provides better fit with the experiment. Two exciton peaks for e1-hh1 and e1-lh1 are clearly resolved. However, by using the pure 2D Coulomb enhancement factor, because of the overestimation of the continuous part the first peak of e1-hh1 is less obvious, and instead of it, a step like curve was found, which makes 1.33 a better choice for Coulomb enhancement factor.

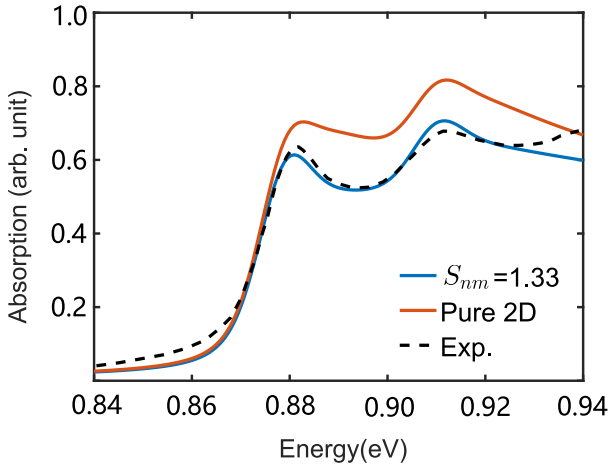

Figure 5.5: Comparison of the calculated absorption spectra of a 10nm Ge/$Si_{0.15}$Ge QW when using the pure 2D enhancement factor and 1.33. The experimental results are plotted from [105]

5.2.4 Comparison with Experiment

The simulation results from this model under different bias values are compared with those in [105] for a 10nm Ge QW. The input parameter of this model is the applied field in kV/cm, while in the experiment the voltage is used [105], And the corresponding applied field is not explicitly given. The values for 0 V, 1 V, 2 V, 3 V and 4 V were obtained from Fig.4 in [105]. The applied fields corresponding to these voltages are 15.03, 37.33, 58.95, 81.81 and 102.45 kV/cm.

The $Si_{0.15}Ge_{0.85}$ MQW structure was grown on a $Si_{0.1}Ge_{0.9}$ substrate. Directly using this structure in the model will results in a larger separation of the peak between hh1 and lh1 than in the experiment, this is because in reality the buffer layer is not fully relaxed after annealing process, it will usually have a tensile strain from 0.15% to 0.2% [112]. But this has not been mentioned in [105]. To resolve this effect and to better match the experiment, the way used in other simulation papers is also applied here [88, 112, 116], and we found that the substrate with 0.15% tensile strain will have the best match with the experiment. The FWHM of the first heavy hole exciton peak was $\sim$ 16meV in the experiment. Hence, just for the purpose of validation, here we used the same homogeneous broadening $\Gamma_L = 6meV$ in Section5.2.3, and the inhomogeneous broadening Γ_G is set to 10 meV as a constant value. The FWHM of the Voigt profile is then close to the experiment broadening ($\sim$8 meV).

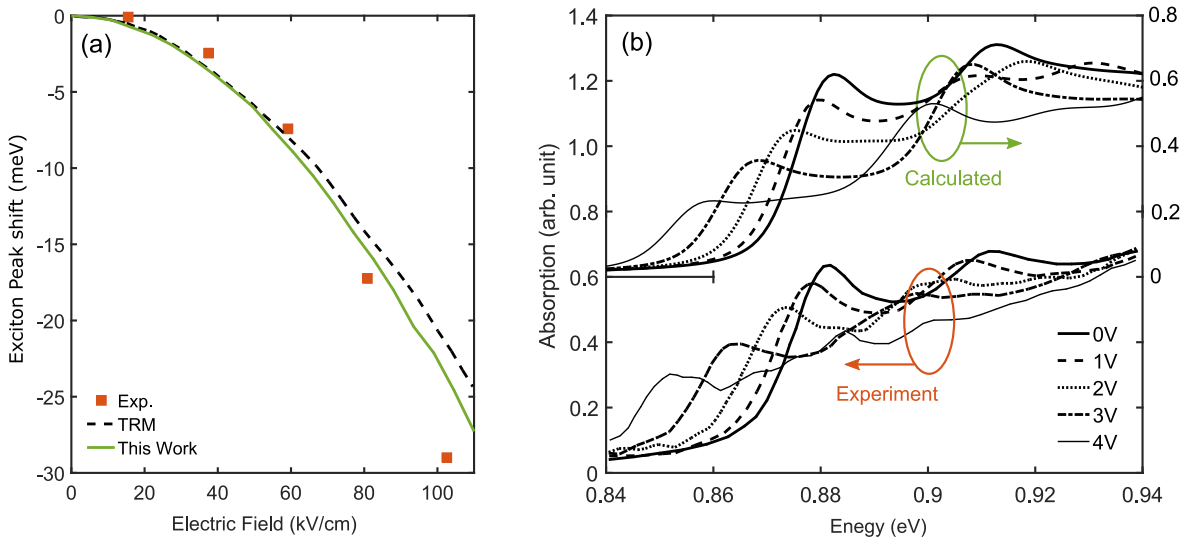

Figure 5.6: Comparison between experimental and simulation results for a $Si_{0.15}Ge_{0.85}$/ Ge MQW grown on $Si_{0.1}Ge_{0.9}$ substrate. (a)c1-hh1 exciton peak against the electric field from this work(green), tunnelling resonance method(dash line) and experiment; (b)Absorption spectrum with bias voltage from 0-4 V from this work and experiment, the calculated spectrum is offset for clarity

.

Fig.5.6a shows the exciton peak shift caused by QCSE. The five red square dots are from the experimental measurements. Compared with the results from our model and the results calculated by the tunnelling resonance method (TRM) [105], our calculated results have about 5 meV difference on the peak shift at 5 V ($\sim$100 kV/cm), this difference is even larger for TRM. This difference was explained in [116] as due to the unintentional background doping, which causes additional non-uniform electric field, which is not considered in the modelling. But

overall, the **k.p** model seems to have a better agreement with the experiment than the TRM. For the absorption spectrum, the model has a reasonable agreement with the experimental results, especially at 0V. The exciton peaks are clearly resolved for e1-hh1 and e1-lh1, positioned at 0.88 and 0.911 eV, which only have 1meV difference from the experiment. For rather high electric fields the exciton peaks are less distinct, but they can still be identified from the experiment. In fact, such difference was also found in [88], and it was explained as the result of increased indirect absorption when the bias is high, also related to larger heating at high electric fields, which will boost the indirect absorption process.

Generally, the model can have an acceptable match to the experiment after some adjustments to the structural input, which provides sufficient validity for it to be used in the following design consideration. Since the differences between experiment and theory are rather consistent for all cases that are simulated, this will not inhibit the main idea of such design considerations.

5.3 Quantum Confined Stark Effect in Intra-step Quantum Well

Compared with the conventional square QW, the QCSE is different in intra-step QW, and with some freedom regarding its geometry, the QCSE can be engineered in a limited range. To understand the difference, the perturbation theory based on infinite QW is here used to qualitatively analyse the QCSE.

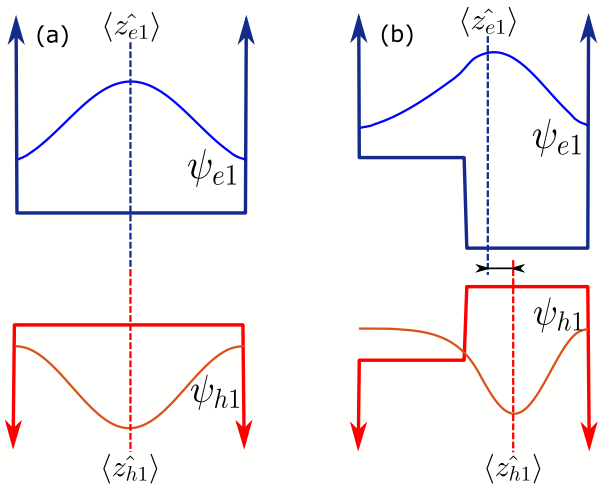

Figure 5.7: The wave function and potential profile in (a)Infinite square QW, (b)Infinite intra-step QW

.

Fig. 5.7a shows a typical square QW with infinite barrier heigh. When an electric field is applied, the term $H' = q_e F z$ is treated as a perturbation. Following the derivation in [52], the subband energy within the second order perturbation is:

$$\begin{aligned} E &= E_n^{(0)} + H'_{nn} + \sum_{m \neq n} \frac{H'_{nm}}{E_n^{(0)} - E_m^{(0)}} \\ &= E_n^{(0)} + \left\langle \psi_n^{(0)} \middle| q_e F z \middle| \psi_n^{(0)} \right\rangle + C_n \frac{(\pi q_e F L_{eff})^2}{2E_1^{(o)}} \\ &= E_n^{(0)} + \langle \hat{z} \rangle q_e F + C_n \frac{(\pi q_e F L_{eff})^2}{2E_1^{(o)}} \end{aligned} \tag{5.16}$$

The wavefunctions and energy with the superscript '(0)' are the unperturbed solutions (i.e., with no electric field), n is the index of the subband, $\langle \hat{z} \rangle$ is the average value of the particle coordinate, L_{eff} is the effective well width, C_n is the coefficient of the nth band. The explicit expression of C_n is derived in [52], and for the first band $C_1 \simeq -2.19 \times 10^{-3}$. Then, by applying Eq. 5.16 to both electrons and holes, the bandgap variation with the applied field can be written as:

$$E_{h1}^{e1} \simeq E_{h1}^{e1}(F=0) - \left[\left(\langle \hat{z_{h1}} \rangle - \langle \hat{z_{e1}} \rangle \right) q_e F + 2.19 \times 10^{-3} \frac{(m_e^* + m_h^*) q_e^2 L_{eff}^4}{\hbar^2} F^2 \right] \tag{5.17}$$

The second term represents the change in the energy gap which is due to QCSE. For the convenience of explanation, the QCSE can be further written as a quadratic function of the electric charge times the electric field:

$$\Delta E_{QCSE} = S_1 (q_e F) + S_2 (q_e F)^2 \tag{5.18}$$

where the coefficients are:

$$S_1 = \langle \hat{z_{h1}} \rangle - \langle \hat{z_{e1}} \rangle, \quad S_2 = 2.19 \times 10^{-3} \frac{(m_e^* + m_h^*) L_{eff}^4}{\hbar^2} \tag{5.19}$$

S_1 is the difference between the average coordinate of the hole and electron wavefunctions, and S_2 is determined by electron and hole effective mass and the effective width. For the square QW, as is shown in Fig. 5.7a, the unperturbed average position for both electron and hole are both located at the center, due to the symmetry. Hence, S_1 is zero and only the second term exists. Thus for a square QW, the QCSE is only determined by the sum of effective masses $m_e^* + m_h^*$

and the effective well width L_{eff}. Since the material composition in the well is not varying a lot, the effective mass will also not vary a lot because of the material change. The coefficient S_2 is proportional to $m_e^* + m_h^*$ to the power one, and to L_{eff} to the power of four, so the QCSE is more sensitive to the well width. The test results also support this point of view. As can be seen in Fig. 5.8, for the $Si_{0.1}Ge_{0.835}Sn_{0.065}$/GeSn QW grown on $Si_{0.1}Ge_{0.82}Sn_{0.08}$ substrate, when the Sn content in the well varies from 8.25% to 10.15%, a slightly larger QCSE is found because of a larger sum of effective masses. However, this difference is very small compared to the difference brought by the change of the well width. The Stark shift is increased significantly by increasing the well width from 10 nm to 14 nm, as expected from the above analysis. So, for the square QW the most efficient way to increase the QCSE is to use QW with larger well width. However, such a structure will also have a smaller absorption change, because the carriers are easier to separate to the opposite sides of the QW under electric field, which reduces the overlap of the wavefunctions, and the optical matrix element becomes small.

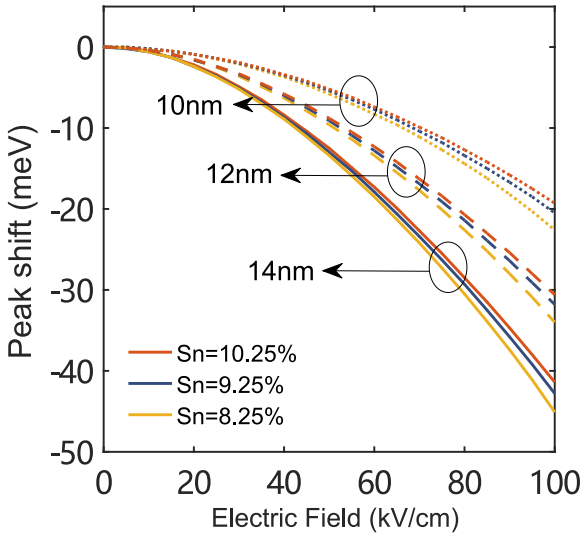

Figure 5.8: Stark shift vs electric field in square QW with different well widths and Sn contents in the well.

.

For the intra-step QW, as is shown in Fig. 5.7b, things are different. A similar analysis can be conducted through Eq. 5.18. For a typical example given in Fig. 5.7b, with the presence of the intra well, the wavefunctions for both electrons and holes are more or less confined in the intra well. Since holes have a much larger effective mass than electrons, holes are more confined to the intra well, while the electron energy is usually above the intra well and less

confined wavefunction is observed. This difference will cause different average position of the electrons and holes, which makes the coefficient S_1 no longer to be zero. For this reason, instead of having a parabolic-like QCSE dependence on the applied field, for intra-step QW it will be more linear. Although this change of geometry will also cause some change in the coefficient S_2, a larger Stark shift can always be found for intra-step QW when the electric field is small. Although there is no analytical expression for S_1, it can always be engineered by adjusting the ratio of width and the depth of the intra well by controlling the Sn content in the intra layers. Finally, a larger QCSE will be obtained within the range of applied field used for modulation.

As the first step of engineering the intra-step QW, in this section, only the QCSE is considered. The overall performance, that is also affected by the absorption change, will be discussed in the next section. To give an idea about the difference between intra-step QW and square QW, and how a large QCSE can be achieved by engineering the geometry of the structure, six typical structures are used here for comparison. As explained previously, we only want the QCSE to be enhanced by geometry engineering, not by changing the well width. So a constant geometry length of $w_1 + w_2 = 14$nm is applied here. The width and the Sn contents of the intra layers of the six structures are listed in Table.5.3

Table 5.3: Width of the intra layers and Sn contents ($Ge_{1-y}Sn_y$) of structures A-F

	w_1 (nm)	w_2 (nm)	y_1^{Sn} (%)	y_2^{Sn} (%)
Structure A	14	–	9	–
Structure B	7	7	8	10.2
Structure C	13	1	8	10.2
Structure D	11	3	8	10.2
Structure E	9.5	4.5	8	9.2
Structure F	12	2	8	12

With $Si_{0.1}Ge_{0.835}Sn_{0.065}$ as the barrier and $Si_{0.1}Ge_{0.82}Sn_{0.08}$ as the substrate, the well region will have compressive strain, and the heavy hole potential is higher than light hole in the well. Then for clarity, only the band diagrams of conduction band and heavy hole band are shown. Together with the energy levels and wavefunctions of e1 and hh1, they are plotted in Fig. 5.9. Note that structure A is just a square QW for comparison. For other intra-step QWs, the Sn content in the intra-barrier is set to 8%, the depth of the intra-well is controlled by how much the Sn content in the intra-well layer is larger than that in the intra-barrier.

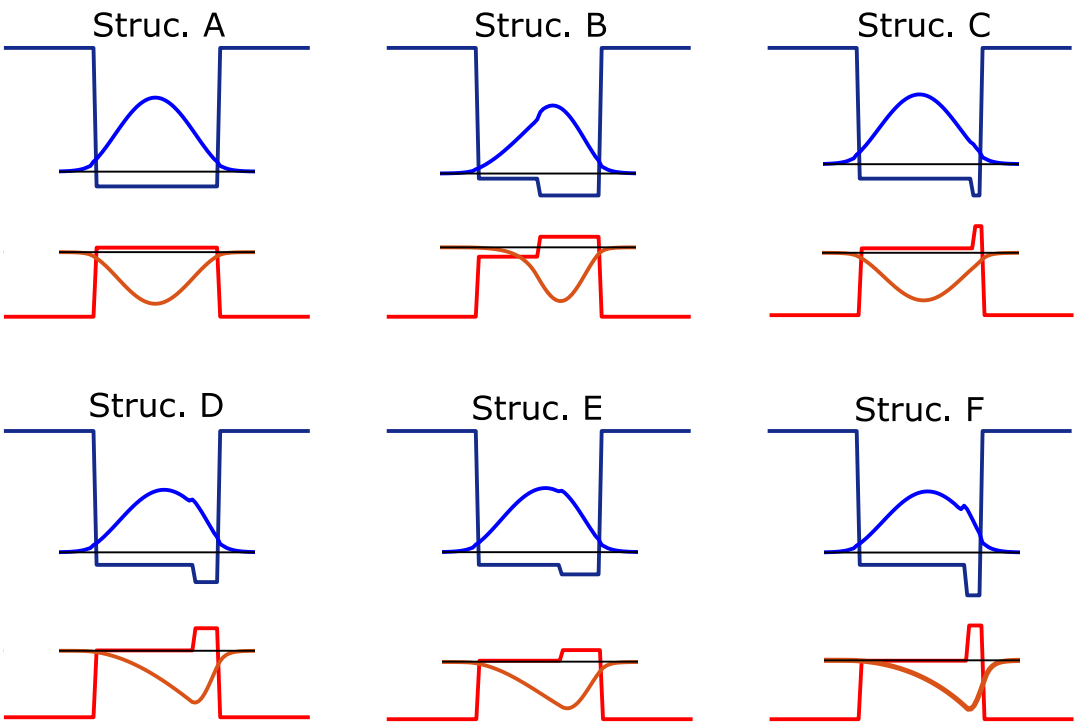

Figure 5.9: Conduction band and heavy hole band potential profiles, subbands and wavefunctions of structures A-F.

.

The Stark shift for the six structure versus the applied field is plotted in Fig. 5.10. Note that the direction of the positive electric field is from the intra-step barrier to the intra-step well. The intra-step QW has a non-zero coefficient S_1 (the first term in Eq. 5.18) according to perturbation theory, but the coefficient S_2 will also changed by modifying the geometry. The expression for S_2 in Eq. 5.19 is derived for a square QW. Although we did not give analytical expression for S_2 of intra-step QW, the calculations show that the intra-step QW will usually have a smaller S_2 than the square QW, this can be due to a smaller effective well width L_{eff}. This will bring about a scenario that if S_1 is not large enough, the intra-step QW will not always have a large QCSE in the range of fields which is needed for modulation. An example can be seen by comparing the QCSE of structure A and B. Because of a rather wide intra well for structure B, the difference between the average position of the electron and heavy hole is small, and then S_1 will be small in such a case. From Fig. 5.10a, we can see that, compared to the square QW (structure A), a larger QCSE can only be observed from 0-$\sim$60 kV/cm. Beyond 60 kV/cm, square QW will have a larger shift.

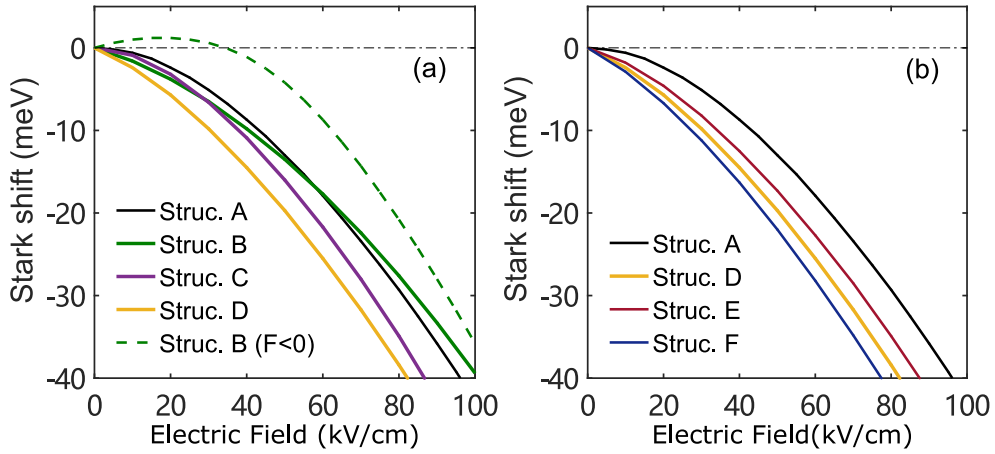

Figure 5.10: Comparison of the QCSE shift vs the applied field from 0-100 kV/cm between (a)structure A-D; (b)structure A, D, E and F

.

To achieve a large Stark shift from 0-100 kV/cm, we first use the same Sn content in the two intra-layers as structure B and change the width ratio (w_1/w_2). Structure C has a very narrow intra-well, just 1nm wide. For this reason, structure C has wavefunctions very similar to those in a square QW. From the calculation the difference between the average position of electron and hole (i.e. S_1) is only 0.7 nm, compared with 2nm for structure B. Because of this small value of S_1, it has a smaller QCSE than structure B at the start of the horizontal axis. But also, as it is more like a square QW, S_2 for structure C has not changed a lot compared to the square QW, because the energy levels of electron and hole are all 'above' the intra-well, so they have a similar effective well width. Thus it has a larger Stark shift after about 30 kV/cm, compared to B, and it has a generally larger QCSE from 0-100 kV/cm compared to the square QW. Within 0-100 kV/cm, we found that structure D, with a width ratio of 3.67, has the largest QCSE than any other width ratio. As can be seen in Fig. 5.10a, S_1 calculated here is 2.4nm, which is also the largest value for such a combination of Sn content in the intra-layers. From the results for structures B, C and D, it can be seen that for a given Sn combination of contents (i.e. a certain intra-well depth), the largest QCSE can be achieved by neither a wide intra-step well nor by a very narrow well, and an optimal width ratio is expected to exist.

Further testing is conducted by changing the well depth, by adjusting the Sn content in the intra-well. Then the optimal width ratio was found for each combination, it is denoted as R'_w here. The structure E has a shallower intra-well, the R'_w is found to be 2.11. And for structure F, R'_w is found to be 6. A general trend can be seen that for a deeper intra-well a larger width

ratio R'_w is needed for optimal structure. From Fig. 5.10b, structure F has the largest QCSE, indicating that a larger QCSE can be achieved by a deeper well, assisted by a narrower intra-well width. This can be also revealed by the coefficient of the first perturbation term. The values of S_1 for D, E and F are 2.4, 1.8 and 2.9 nm respectively. Note that an interesting pattern for the optimal width ratio for structure D-F is that all of their first hole states are located just at the edge of the intra-well. Although the theoretical explanation to this coincidence is still missing, this could be a possible criterion for a large Stark shift to be observed for a particular Sn combination in the intra layers.

We can keep increasing the QCSE by using deeper and narrower well, but it can be easily seen that such improvement from intra-step QW is limited by the total structure width. Since the first term of the perturbation contains $\langle \hat{z_{h1}} \rangle - \langle \hat{z_{e1}} \rangle$, and since in the intra-layer structure both the electron and hole average positions are located on the right side of the wider well, it is impossible for S_1 to be larger than $(w_1 + w_2)/2$. As fabrication difficulties also exist to grow a very thin layer with larger Sn difference than other layers, we do not expect these extreme cases to be practically useful.

Finally, it is worth mentioning that the Stark shift behaviour of square QW is independent of the electric field direction. However, in an asymmetric structure the different directions of the applied field bring about two different potential profiles. For an intra-step QW the sign of the first perturbation coefficient is positive, which means that when a forward applied field is applied ($F > 0$), the first and second term will have the same sign. If $F < 0$, the sign will be opposite for the two terms. In this scenario, by increasing the applied field, the transition energy will initially experience a blue shift, and then a red shift for a higher electric field. This is confirmed in our calculations. As can be seen in Fig. 5.10a, for structure B, when a reverse bias is applied the peak shift is initially increasing, and a blue shift is observed between 0 to 38 kV/cm. Thus the on-set field can be set to have a lager value, and the voltage interval used for modulation can be large as well. This particular feature is applied to modulate light with high optical power because high applied field will help to sweep the photogenerated carriers out of the well. More details of this application can be seen in [119, 130, 131]. But as it is not the topic of this work, no further discussion will be conducted here.

5.4 Electroabsorption Modulator Figures of Merit and Device Performance

In this section, we look into the general performance of a modulator in some comprehensive aspects. The following figures of merit are considered. (1)The exciton peak shift ΔE_x against the applied field, representing the QCSE of the structure. (2)The absorption intensity change $\Delta\alpha$. (3)the absorption change per unit applied field $\Delta\alpha/F$. (4)the absorption change per unit applied field squared $\Delta\alpha/F^2$. The third and forth figures of merit are proposed in [132] and [133]. They correspond to different practical performance index of maximum bandwidth to the swing voltage and power consumption per bit data. Some control variables are set for this test. The total structure width $w_1 + w_2$ is set to 12 nm, and the sum of the Sn contents in the two layers are set to 18% (i.e. $y_1^{Sn} + y_2^{Sn} = 0.18$). The buffer layer is using fully relaxed $Si_{0.1}Ge_{0.82}Sn_{0.08}$, and the barrier is $Si_{0.1}Ge_{0.835}Sn_{0.065}$ pseudomorphically grown on this buffer. The barrier width in the MQW structure is set to 10 nm. Then a 2.28% tensile strain is found in the barrier and compressive strain (around 6%) is found in the GeSn QW layers.

Apart from the control variables, we have two varying parameters, the width ratio R_w to determine the width of the intra-well and the difference of the Sn contents in the two layers $\Delta y^{Sn} = y_2^{Sn} - y_1^{Sn}$ which determines the intra-well potential depth. For the convenience, instead of directly varying R_w and Δy^{Sn}, we change the width and Sn content in the right intra-layer (i.e. w_2 and y_2^{Sn}), the width and Sn fraction in the left layer is determined by $y_1^{Sn} = 0.18 - y_2^{Sn}$ and $w_1 = 12 - w_2$. A searching procedure is conducted in this parameter space. The Sn content in the intra-layers is limited between 6% and 12% in order to control the operation wavelength and maintain a good carrier confinement for the given barrier material. Then a structure was found which has better performance in terms of all the four figures of merit mentioned above compared to a square QW with 9% Sn content in the well.

For clarity, here are some explanation of how these figures of merit are extracted from the absorption spectra of different structures with varying applied field. Some important components of these figures of merit are plotted in Fig. 5.11. For different intra-step structures, the transition energy at on-set applied field is different. In order to compare these structures, the absorption change at the exciton transition energy is plotted against the peak shift $\Delta E_x = E_x(0) - E_x(F_0)$ achieved by appropriate bias F_0. It should be noted that the definition of the exciton peak shift

ΔE_x is here slightly different from the Stark shift ΔE_{QCSE} mentioned before, but since the change of the exciton binding energy is negligible (only 1meV from 0-100 kV/cm) compared to the QCSE energy shift, we can take that $\Delta E_{QCSE} \simeq \Delta E_x$. As can be seen in Fig. 5.11, since the total absorption consists of the contributions from 1s exciton state and continuum state, the total absorption peak is not located exactly at the exciton transition energy E_x, but it is close to E_x, and the difference between the total absorption peak and exciton peak is rather constant. Therefore, this transition energy is a good choice for the operating photon energy, at which the modulation is performed, and a large absorption change is obtained. When a large Stark shift is achieved, the detuning energy ΔE_d can be chosen to be the same as the Stark shift energy (i.e. the operating energy $\hbar\omega_0$ is equal to the exciton transition energy E_x when a field F_0 is applied). Then the absorption is measured at $E_x(F_0)$ as $\Delta\alpha = \alpha(F_0) - \alpha(0)$.

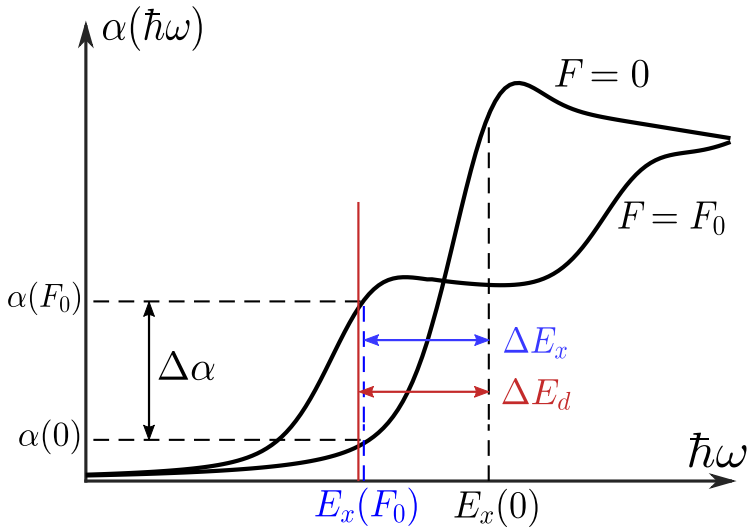

Figure 5.11: The absorption spectrum of a GeSn quantum well at two different biases 0 and F_0.

.

After using our **k.p** model to calculate the absorption spectra of structures with different applied fields, the program will follow the above description to obtain these figures of merit in this simulation. For the four figures of merit, ΔE_x is plotted against the applied field. $\Delta\alpha$, $\Delta\alpha/F$ and $\Delta\alpha/F^2$ are plotted versus the ΔE_x achieved by some applied field. The response of square QW with 9% Sn content in the well to different applied fields is chosen for comparison. For the intra-step QW, among these three figures of merit, we mainly focus on $\Delta\alpha/F$, and at the same time want to get larger values of the other three figures of merit. The optimal structure is found by checking the plots of these figures of merit. The Sn contents of the intra-layers

of such structure (denoted as structure II) are $y_1^{Sn} = 7.25\%$ and $y_2^{Sn} = 10.75\%$, and the widths are $w_1 = 8.5$ nm, $w_2 = 3.5$ nm. It should be noted that, from the simulation level, such structure does not necessarily have the largest $\Delta\alpha/F$. Similar to the way of obtaining a larger QCSE introduced in Section 5.3, this figure of merit can be further improved by using some extreme cases with narrower intra-well. These structures are not chosen for the following reasons. Firstly, insignificant improvement has been found by using a well width less than 3.5 nm. Although narrower structures can have a larger QCSE, such improvement becomes smaller by pushing the parameter choice to more extreme cases. At the same time, the overall absorption change will become smaller for more extreme case with large QCSE, they will be even smaller compared to the square QW, which violates one of the rules to find the optimal structure. Secondly, considering the realistic limits, an extremely narrow well with large variation of the Sn composition will bring about more difficulties into the manufacture process, which makes this kind of structure to lose its practical value. Hence, we think that structure II is a more practical example to be considered here.

Apart from the optimal structure, the figures of merit of two other structures are given. Since the transition energy of structure II has some difference from the control group, to better compare the practical performance later with a square QW, the Sn content of the square QW is changed from 9% to 8.5%. Then the square QW with zero applied field will have the same transition energy as structure II. The square QW here is denoted as structure I. On the other hand, structure II does not have the largest QCSE with its Sn content combination, so another structure was found to have the largest QCSE by using the same Sn content combination as structure II, but different width ratio. This structure has a narrower intra-layer, of 2 nm. Here it is denoted as structure III. This example is chosen to illustrate that a larger QCSE does not necessary deliver a better overall performance considering other figures of merit. The choice of the optimal structure depends on which figure of merit is most valuable to the design purpose. Note that, because of a narrower intra-well, structure III has a different transition energy from other two structures at the onset voltage. Thus we will not compare its practical performance when integrated with a real waveguide design. Nevertheless, we can still compare their figures of merit by assuming their on-set exciton peak at the original point and plot the figures of merit against the peak shift energy on one graph. In reality, to compensate for the difference between the required transition energy and design a structure with similar performance, one can probably refer to the alloy composition difference and width ratio of the original design,

then shift the Sn content to achieve the required transition energy, because the geometry inside the well is the major aspect that affects the optical behaviour of the structure, and slightly shift the Sn content which will only shift the potential profile up and down, and a small change of the final performance is expected from such modification.

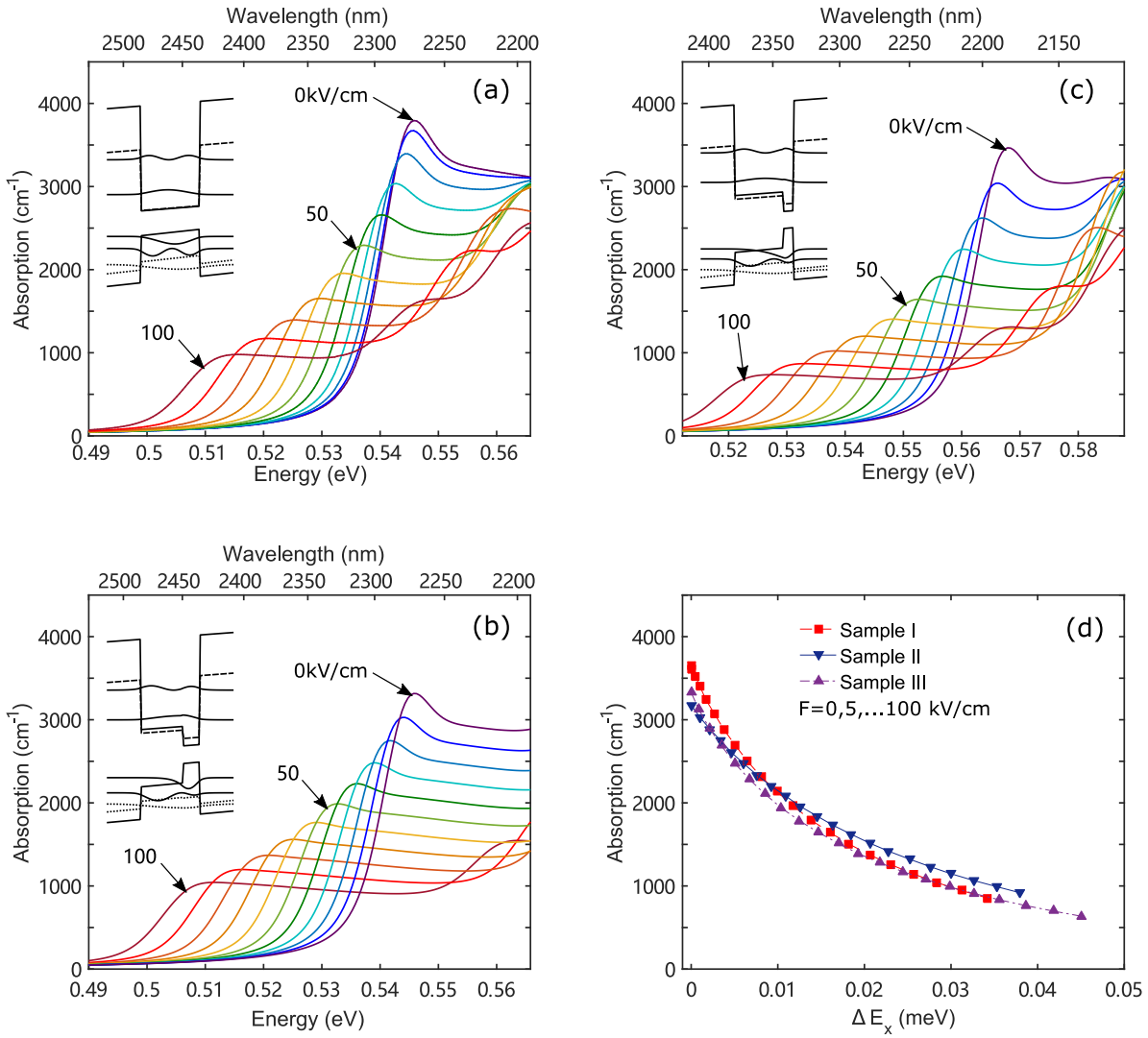

Figure 5.12: TE absorption spectrum at room temperature for (a)structure I, (b)structure II, (c)structure III. Schematic QW band structure and wavefunctions are shown in the insets, in which the dashed and dotted lines represent L-valley and LH respectively; (d)The absorption intensity at exciton peak versus the peak shift energy

.

In Fig. 5.12, the absorption spectrum of structures I-III with applied field from 0-100 kV/cm in steps of 5 kV/cm, and the comparison of the absorption at exciton peak are shown. Only the TE mode is considered here, as the heavy hole state is much higher than light hole state for these structures. From Fig. 5.12d, the absorption values at the exciton transition energy are

similar to each other, but they are still different in some details. When the applied field is low, for example 0 kV/cm, the absorption peaks of square QW are larger than those of intra-step QW. This is because the intra-step structure reduces the overlap between the wavefunction of conduction band and valence band. Nevertheless, the absorption intensity decreases slower for the intra-step QWs than for the square QW. For structure II, the absorption has a larger value than for structure I after about 10 meV shift. Another difference is that for square QW (structure I). The stark shift is not very obvious for low applied field. The shift per unit applied field becomes larger when the applied field is already fairly large (e.g. 40 kV/cm). But for the step QW, as explained in Section 5.3, because of the first perturbation term, the QCSE behaves more linearly than for square QW. Then we can see that even at low applied field, the edges are changing more rapidly and more uniformly with the increasing electric field. This is a good aspect for the on-chip modulators because even low applied field can be used for modulation without need to set the onset voltage to some larger value.

Fig. 5.13 compares the four figures of merit of structures I-III. The exciton peak shift is the energy difference between the exciton transition energy for a finite field F and the on-set field, here we assume the on-set applied field is 0 kV/cm. Also the absorption change is measured at the exciton transition energy at any given F from the residual absorption at zero field. This residual absorption for the on-set applied field is related to the insertion loss of the modulator. As can be seen in Fig. 5.13a, structure II has larger QCSE than structure I from 0-100 kV/cm, however structure III, compared with the other two, has the largest QCSE. This larger QCSE will also cause the reduction on the absorption change. As can be seen in Fig. 5.13b, structure III has the smallest absorption change. However, structure II is capable of maintaining a larger absorption change and QCSE at the same time compared with the square QW. In terms of the absorption change per unit of applied field $\Delta\alpha/F$ shown in Fig. 5.13c, structure II is much better than structure I and slightly better than structure III within the full range of the peak shift energy. Comparing structure II and III, we can see that, although a structure has a larger QCSE, it does not necessarily have an overall better performance if considering another important aspect- the absorption change, which, to some extent, corroborates the idea in [120]. But as we can see from comparison between structures I and II, it is still possible to have larger QCSE and larger absorption change simultaneously by engineering the intra-geometry of the QW, which makes the simple square QW definitely not the optimal solution in this scenario.

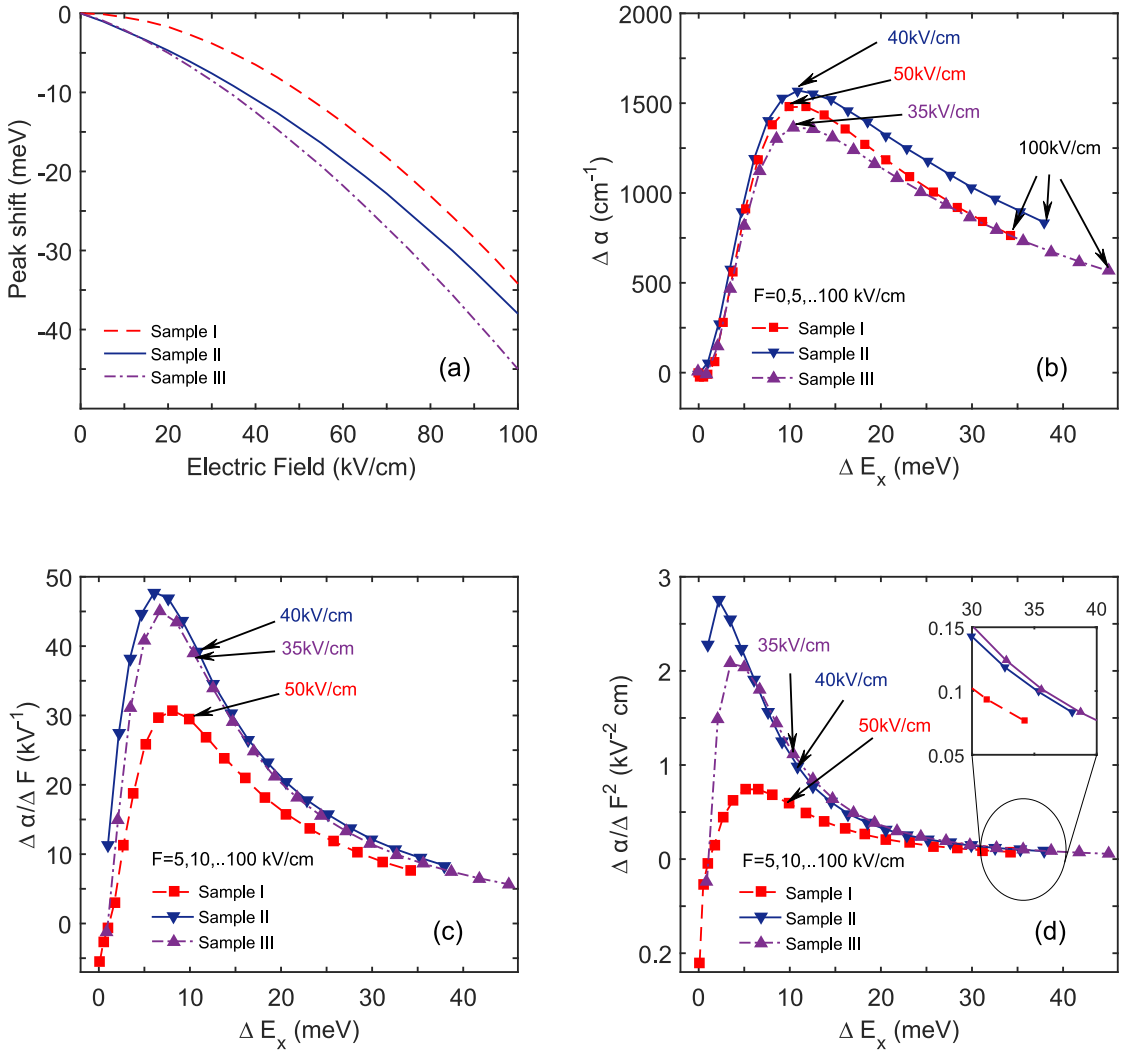

Figure 5.13: Comparison of figures of merit. (a)Peak shift against the electric field; (b)$\Delta\alpha$ (c)$\Delta\alpha/F$ and (d)$\Delta\alpha/F^2$ against the peak shift for electric field ranging from 0 to 100 kV/cm.

.

Additionally, another figure of merit that may be used to measure the usefulness of the modulator design is the absorption change per unit applied field squared $\Delta\alpha/F^2$. This figure of merit was proposed in [133], and maximizing it is equivalent to minimizing the power consumption per bit data transmitted when a certain contrast ratio is reached. As can be seen in Fig. 5.13d, for this figure of merit, the intra-step QW still has a large superiority compared to the square QW. Even at large Stark shift, shown in the magnified tail of the curve, the improvements of both square QWs are as large as 50%. Comparing structures II and III, different from $\Delta\alpha/F$, structure III has larger $\Delta\alpha/F^2$ than structure II. This is because structure III has a much larger QCSE, although a smaller absorption change is found in structure III, but since the denominator of this figure of merit is F^2, this makes it more sensitive to the applied field needed to achieve the same Stark shift. So it can be seen that the optimal design should be considered

according to a specific standard, which is similar to the objective function in the mathematical optimization process. Different figures of merit will bring about different optimal choices.

For $\Delta\alpha$, $\Delta\alpha/F$ and $\Delta\alpha/F^2$, they reach their peak at certain exciton peak shift energy. For $\Delta\alpha$ the maximum for the three structures can be found around 11 meV shift. $\Delta\alpha/F$ reaches its peak when ΔE_x is around 8 meV, and the peak of $\Delta\alpha/F^2$ appears around even smaller shift energy ($<\sim$5 meV). Although the goal of designing a modulator is to maximize these figures of merit, one should also consider whether the detuning energy of working device is further restricted by the extinction ratio (ER) and the insertion loss (IL). The operating wavelength is very close to the exciton transition energy at on-set (i.e. a small detuning), then a large IL and small ER can be expected, because the absorption change is small but the residual absorption is large. Here we use ΔE_x as the horizontal axis, and the operating photon energy is a relative parameter. But for the particular three designs here, if a particular photon energy is needed, one has to shift the parameter choice (Sn content, substrate, etc.) to ensure a proper detuning is reached, and the on-set transition energy is ideally away from the operating wavelength or photon energy.

The above results are all based on the assumption that the on-set applied field is set to be 0 kV/cm. Because the characteristics of the QCSE for the square QW, the Stark shift is very small when the applied field is initially low. A typical method used in modulation is to increase the on-set voltage to some non-zero value [70, 112]. Thus the swing voltage can be reduced. When the on-set is not zero, here we denote it as F', the peak shift ΔE_x is between the exciton peak at F' and at the applied field $F > F'$. The absorption change is measured between these two absorption spectra at the exciton transition energy at F. Based on the test on structure I, when the on-set field is 40 kV/cm, the largest figure of merit $\Delta\alpha/F$ can be found in the plot against the ΔE_x. When the on-set is larger than 40 kV/cm, it will cause a reduction of the figure of merit $\Delta\alpha/F$ because a smaller change of absorption is found. For $\Delta\alpha/F^2$ the improvement from a larger on-set is more obvious, and the reason is again because such figure of merit is more sensitive to the shift rate, and a larger interval of the modulation field will have a larger slope of the energy shift against the applied field. To make it clear, the figures of merit when the peak shift is 10 meV are plotted in Fig. 5.14, together with the figures of merit of structure II when 0 kV/cm on-set is applied. As can be seen in Fig. 5.14a, even without increase of on-set to some higher value, for $\Delta\alpha/F$ the intra-step QW keeps its superiority in most of the cases. The

maximum figure of merit that a square QW can achieve when using 40 kV/cm on-set is only 5% larger than that of intra-step QW using zero on-set. For another figure of merit, $\Delta\alpha/F^2$, it increases with increasing on-set applied field, and the intra-step QW is still a better choice if the on-set is lower than 20 kV/cm.

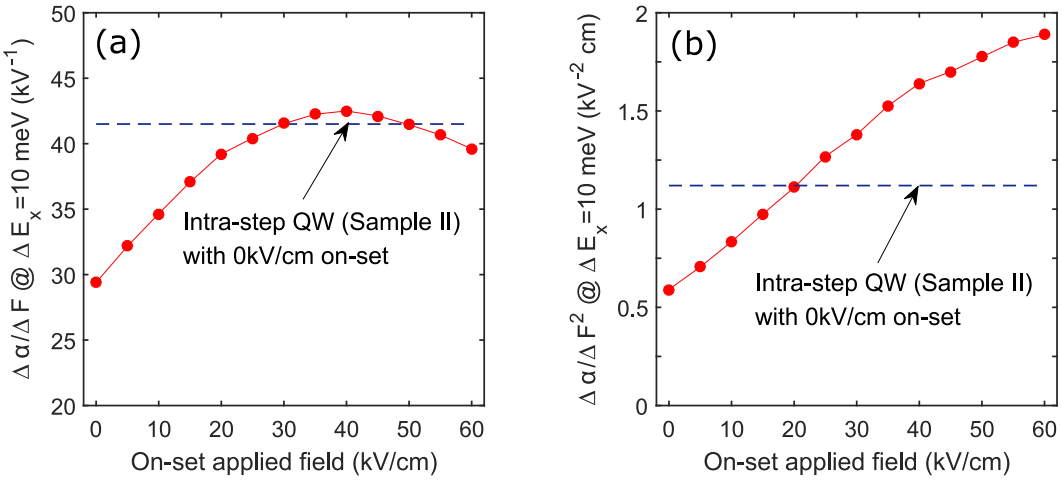

Figure 5.14: Figures of merit of the square QW at 10 meV peak shift when using different on-set applied fields. (a)$\Delta\alpha/F$ and (b)$\Delta\alpha/F^2$. For comparison, the figures of merit of the intra-step QW (structure II) with 0 kV on-set are also shown (dashed blue lines)

.

In fact, due to the second perturbation term in QCSE, the swing voltage can always be reduced by using a larger on-set voltage. This also applies for the intra-step QW cases, although the second term may be smaller than for a square QW with the same width. However using a larger on-set is not what we want to achieve in this design considerations, because for on-chip device a low operation voltage is always preferable. Presently GeSn is a still an under-development material, the tolerance of the maximum voltage it can handle is still uncertain. High voltage operation may cause multiple problems, like current leakage and even device breakdown. The linear and larger Stark shift gives a larger freedom for the voltage region to be used for modulation. The on-set can be set to zero without wasting any swing voltage for modulation. This makes the intra-step QW more suitable for low voltage operation and more favourable as an on-chip EAM design.

Note that in this model the built-in voltage is not included. The optimal structure I is proposed based on this assumption, and zero electric field can be found in the intrinsic region when zero voltage bias is used. Usually the built-in voltage will not be very large, and the corresponding

applied field will not be as large as that from the swing voltage, but the optimal structure can still be a bit different from the design resented here.

We can reasonably assume that regardless of what the initial conditions or the basic set up of the device, the performance can always be improved by varying the potential profile of the QW. Considering the current fabrication capabilities, this does not have to be limited to an intra-step QW, and more complicated structures can probably provide even better performance. Similar attempts have already been made for the III-V material system. For example, in [134], genetic algorithms are used to search for the broken-symmetry structures in order to achieve a particular functionality.

Finally, the device performance is estimated by combining it with a wave guide design. The operating wavelength is set to 2340 nm, corresponding to photon energy of 0.53 eV, with about 15 meV detuning energy for a zero onset for structures I and II. Fig. 5.15 shows the schematic diagram of the device. The whole structure was grown on a silicon substrate, followed by a fully relaxed p-doped $Si_{0.1}Ge_{0.82}Sn_{0.08}$ buffer layer, so the structure is pseudomorphically grown on such buffer layer. The active region consists of 10 quantum-wells of 12 nm width and 11 $Si_{0.1}Ge_{0.835}Sn_{0.065}$ barrier layers of 10 nm. For comparison, the quantum well is either the square quantum well (structure I) or intra-step quantum well (structure II). The total thickness of the structure is then 230 nm, and this active region is sandwiched between two 50 nm intrinsic spacer layers. Hence the total thickness of the intrinsic layer is 330 nm, denoted as d_i in the figure. n-doped region is above the intrinsic region and on the top is the cap cladding layer. The thickness of the n-doped and p-doped layer will influence the optical confinement factor of the active region, however this is not calculated in detail here, and instead of that, the optical confinement factor was chosen from the data in [135]. With a 0.1 μm n-doped layer and a 0.5 μm p-doped layer, and a similar active region thickness (210 nm), the optical confinement factor Γ_c of the whole active region can be as large as 0.36. The width of the waveguide ridge is 2 μm, and the length of the waveguide L_g is chosen according to the requirement of the extinction ratio (ER, R_{ER}). The cavity length L_g is then related to the extinction ratio by the following equation:

$$L_g = \frac{\ln(R_{ER})}{\Gamma_c \Delta\alpha} \tag{5.20}$$

The requirement for the extinction ration is set to 10 dB, and for different structure or different on-set for the same structure, this cavity length could be different, it is calculated accordingly. Finally, the load resistance is assumed to be 100 Ω.

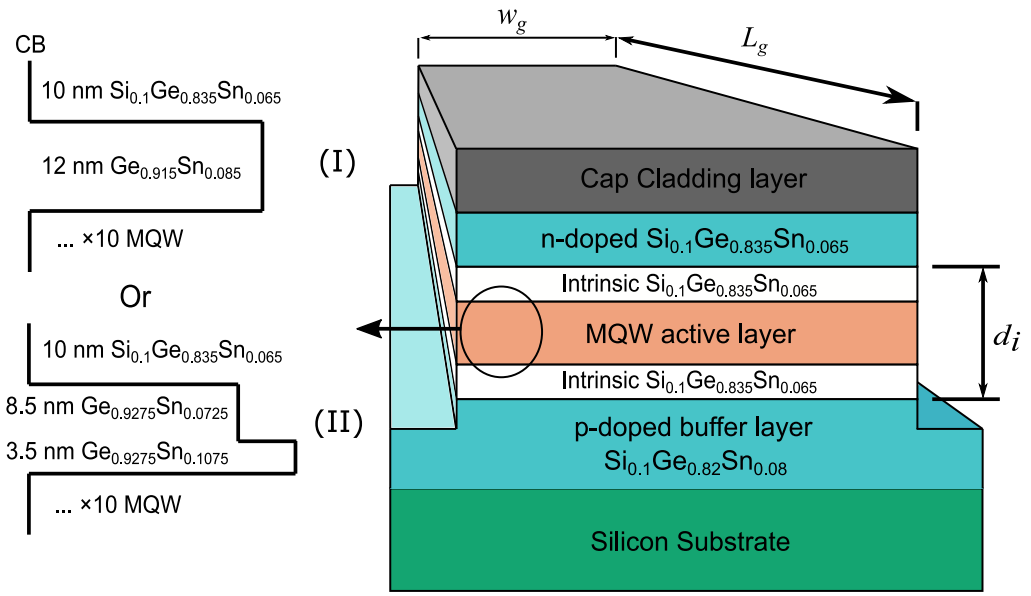

Figure 5.15: Schematic diagram of the waveguide design. The layer with ten MQWs is using either square QWs (structure I) or intra-step QWs (structure II), and they have the same total thickness.

.

For such structure, the two figures of merit, $\Delta\alpha/F$ and $\Delta\alpha/F^2$ can thus be related to the practical performance parameters. Firstly, $\Delta\alpha/F$ can can be related to the ratio of maximum bandwidth to the swing voltage [132]. By assuming that the maximum bandwidth is only related to the time constant R_LC, in which C is the capacitance of the intrinsic region, we have:

$$\Delta f = \frac{1}{2\pi R_L C} = \frac{d_i}{\epsilon_s \pi R_L w_g L_g} \tag{5.21a}$$

$$V_i = F d_i \tag{5.21b}$$

Then by substituting Eq. 5.20 into Eq. 5.21, the equation is independent on the cavity length. The relation between the absorption change per unit applied field and the maximum bandwidth per unit applied voltage is obtained:

$$\frac{\Delta f}{\Delta V} = \frac{\Gamma_c}{\epsilon_s \pi R_l w_g \ln(R_{ER})} \left(\frac{\Delta\alpha}{\Delta F}\right) \tag{5.22}$$

The absorption per unit applied field squared $\Delta\alpha/F^2$ can be related to the switching power per bit by a coefficient independent on L_g [133]:

$$\frac{P_{ac}}{N_B} = \frac{1}{2}C\Delta V^2 = \frac{1}{2}\epsilon_s w_g L_g d_i \Delta F^2 = \frac{\ln(R_{ER})\epsilon_s w_g d_i}{2\Gamma_c}\left(\frac{\Delta\alpha}{\Delta F^2}\right)^{-1} \tag{5.23}$$

From Eq. 5.22 and Eq. 5.23, when w_g and d_i are constant, and the extinction ratio has its minimum required value, $\Delta f/\Delta V$ is proportional to the figure of merit $\Delta\alpha/F$, and $\Delta P_{ac}/N_B$ is inversely proportional to $\Delta\alpha/F^2$. A better performance can thus be expected from higher value of these two figures of merit.

A more explicit plot of the figures of merit versus the photon energy is given in Fig. 5.16 to compare structure I and II which have the same transition energy at zero applied field.

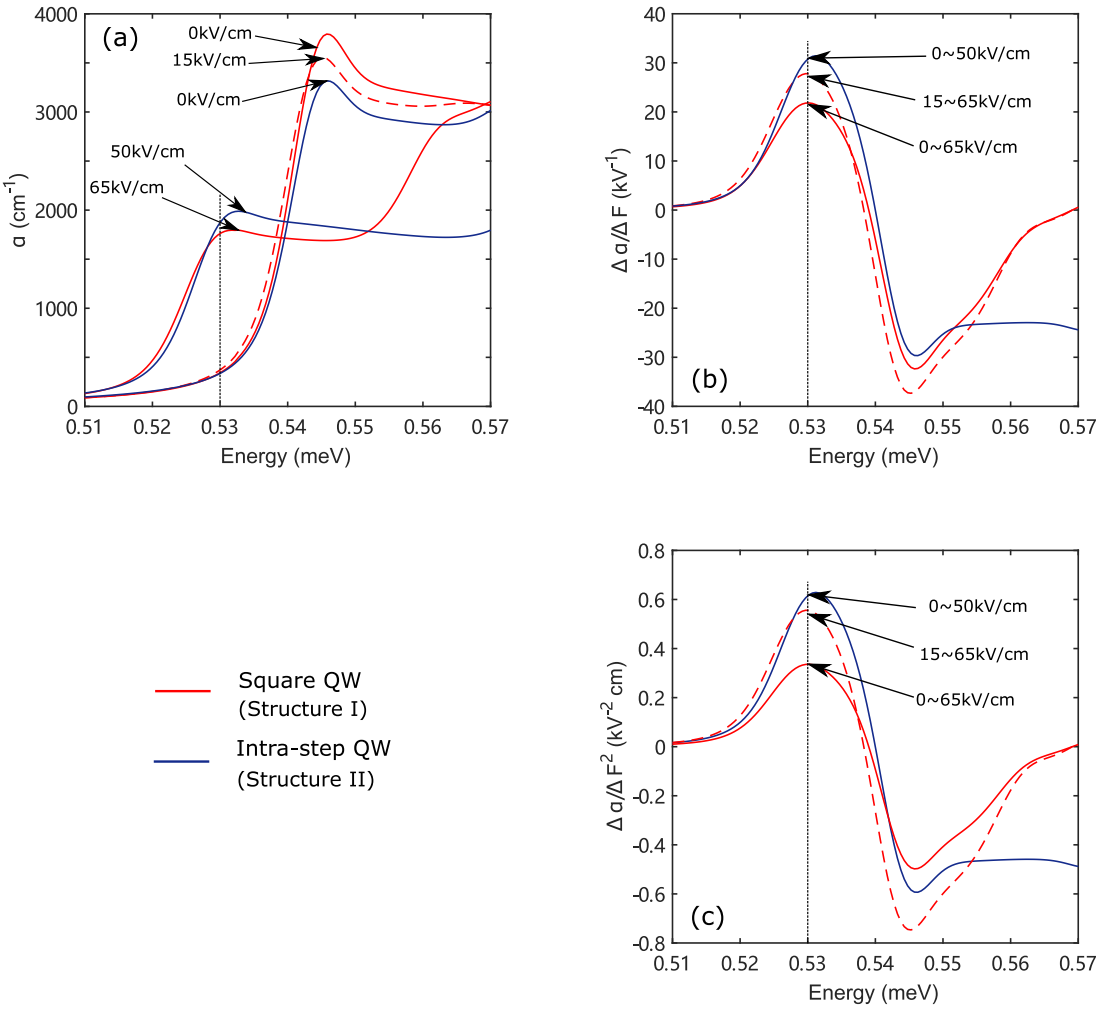

Figure 5.16: Figures of merit versus photon energy. (a)absorption spectra of structure I at 0, 15 and 65 kV/cm and structure II at 0 and 50 kV/cm. (b)The corresponding figure of merit $\Delta\alpha/F$. (c)$\Delta\alpha/F^2$ against photon energy

.

At the operating wavelength 2350 nm (i.e. 0.53 eV), the maximum absorption change can be reached when the off state applied field is ∼65 kV/cm for the square QW and ∼50 kV/cm for the intra-step QW. When the on-set applied field is zero, the absorption coefficient for square QW at 'on transmitting state' is 342 cm^{-1} and in the 'off absorbing state' (65 kV/cm) it is 1760 cm^{-1}. For intra-step QW, these two values are 336 cm^{-1} for the on-state and 1900 cm^{-1} for the off-state. After accounting for the optical confinement factor 0.36, the absorption change $\Gamma_c \Delta\alpha$ of the square QW is 510 cm^{-1}, and 563 cm^{-1} for the intra-step QW. As a non-zero onset is usually applied in reality to reduce the swing voltage, Fig. 5.16 also presents the result when a 15 kV/cm on-set is used, for such structure the on-set voltage will be 0.5 V. By doing so, the swing voltage can be the same as the intra-step QW 1.65V. This on-set was chosen because the residual absorption will only slightly increase. Since the QCSE for square QW is stronger for larger applied field, further increasing of the on-set will significantly reduce the detuning and a larger insertion loss is expected. As can be seen in Fig. 5.16b and Fig. 5.16c, both figures of merit of the square QW are improved compared to the zero on-set case, but are still smaller than for the intra-step QW without altering the on-set voltage. The practical parameters calculated by Eq. 5.20, Eq. 5.21, Eq. 5.22 and Eq. 5.23 are listed in Table 5.4

Table 5.4: Practical parameters of the modulator design using structure I with 0 and 15 kV/cm on-set and structure II with only 0 kV/cm on-set

MQW type	L_g [a] (μm)	V_i^{on} (V)	ΔV (V)	IL (dB)	Δf (GHz)	$\Delta f/\Delta V$ [b] (GHz/V)	P_{ac}/N_B [b] (fJ/bit)
Square QW (I)	45.1	0	2.15	2.4	41.1	19.1	89
Square QW (I)	46.3	0.5	1.65	2.7	39.9	24.2	54.2
Intra-step QW (II)	40.8	0	1.65	2.1	45.3	27.5	47.8

[a] Calculated by Eq. 5.20 to achieve a 10 dB ER.
[b] Calculated according to the cavity length and assuming $R_L = 100\Omega$.

Comparing structures I and II, both with zero on-set voltage, the bandwidth per unit applied voltage is 44% higher, and the power consumption is reduced by 46% for the intra-step QW. By changing the on-set voltage to 0.5 V for the square QW, both of its performance indexes are improved, but an advantage of the intra-step QW still exists, as the bandwidth per unit applied field is still 21% higher and the switching power per bit is still 12% lower than that for the square QW. Moreover, the insertion loss is also smaller for the intra-step QW. It is clear that the intra-step QW can still outperform the square QW, even if its on-set is restricted to

zero. With the possibility of using it for modulators even with zero applied voltage, it can be seen that the intra-step QW is more suitable for low voltage operation and it is more favourable to be used in on-chip design.

5.5 Conclusion

In this chapter the electronic properties of the electroabsorption modulator are calculated by **k.p** model, and the variational method is used to obtain the exciton solutions, and include the contribution from the 1s state and continuum state of the exciton. The indirect absorption is also accounted for using a bulk Ge model. The model shows a fairly good reliability by comparing with the experimental results for SiGe/Ge QW. For our structure, SiGeSn is used as the barrier material, and the intra-layers are made for GeSn. The intra-step is formed by using GeSn with different Sn content. By analysing the QCSE using the perturbation theory, the QCSE behaviour of the square QW and intra-step QW is different due to the first perturbation term. A large QCSE can be expected when the first term is large enough. It was found that, for appropriate Sn content combinations in the intra-layers, a large QCSE from 0-100 kV/cm can be obtained when the width ratio of the intra-layers is chosen to make the first heavy-hole state at the top of the deeper intra-well (i.e. at the band edge of the shallower intra-layer). For a deeper intra-well, a narrower intra-well width is expected to have a large Stark shift, and the QCSE can be larger for more extreme cases. But as the deep intra-well needs to have a large difference in the material composition, and the width of the layer in reality is limited by fabrication problems, these extreme cases can hardly be recommended for real applications. Moreover, together with the QCSE, the overall performance of the modulator is considered by taking the absorption change into account. Also, the two important figures of merit $\Delta\alpha/F$ and $\Delta\alpha/F^2$ are calculated, as they are related to the achievable bandwidth per unit voltage and power consumption per bit when the same extinction ratio is achieved by a waveguide design. The optimized intra-step QW shows a much better performance compared with the square QW when zero on-set voltage is used. Although by using a larger on-set voltage for the square QW, 15 kV/cm (0.5V), the figures of merit can be improved, but the intra-step QW still retains its advantage by using only zero on-set, because for a square QW, by increasing the applied voltage the QCSE is relatively low at the beginning. Using large on-set is necessary to reduce the swing voltage. But as a trend for EPIC, low-voltage operation is then preferred because high voltage

may cause problems such as heating or device breakdown. The intra-step QW is thus obviously a more suitable choice for for the on-chip EA modulators. As the technology is developing, we may fully explore the modern fabrication and epitaxial growth capability. The intra-step QW is just a simple geometry that was considered to be compared with the square QW, and it shows some potential and superiority over the square QW. More complicated asymmetric structures are still waiting to be found to further push the limit of the performance of the EA modulator or some other devices.

Chapter 6

Conclusion and Suggestions for Future Work

In this project, a computational model was build to simulate the QW based semiconductor devices. The major theory were the effective mass method from the basic quantum mechanics and 8-band **k.p** from Bahder's paper [53] which not only includes the interband mixing but also accounts for the strain effects. The error in the initial paper was corrected in this model. Although such model was built to estimate the band structure for a bulk material, it can also be applied to heterostructure by using the finite difference method. Because of the complexity of transfer matrix method to be incorporated in the **k.p** model and incapability of using it for the calculation of free states, here we use matrix method based on finite difference scheme. For quantum well grown in the [001] direction we just need to replace the k_z wavevector with the operator $-i\partial/\partial z$. Similar method is also applied in the effective mass method. By using reasonably large number of discrete points and appropriate barrier width as an approximation for the boundary conditions, a good accuracy was found by comparing with experimental results. The major source of the spurious solution problem was also identified and solved by rescaling the momentum matrix element. It was found to be an efficient way to eliminate the spurious solutions without too much modifications of the existing model, and with only small changes in the finial results. For the optical properties, based on the equation for absorption coefficient, the explicit expression for the momentum matrix element was derived based on the basis used before. Since the computation time was found to be unacceptably long because of lots of calculation needed for the integration of the line-shape functions, the simplex method was used to reduce

the calculation time. For the same numerical settings the simplex method only needs 1/10 of calculation time and gives almost the same result with that obtained before. The calculation part of the model was written in Fortran, assisted by the LAPACK linear solver package. It also provides a graphical display using Fortran-Matlab API, which is also very convenient for examining the real-time results and debugging. The code uses modular design, which enables some future extensions and modifications like different input parameter, arbitrary structure, alternative Hamiltonian matrix and additional effects involved.

In chapter 4, the model is used to simulate and optimize the GeSn QW lasers. The conventional GeSn lasers usually have large Sn content and challenge exists in growing high quality layer with small defect density. As tensile strained laser with lower Sn content has already been proved to be efficient for bulk material, this idea is here extended to the QW structure. The overall goal of this research was to find how the gain coefficient depends on the choice of parameters such as well width, Sn content and strain, and also the behaviour of losses, affecting the net gain. The Γ valley is considered by the **k.p** method, the band structure near Γ point is used to obtain the quasi Fermi level and the optical properties. The L and Δ valleys were modelled by the effective mass method, in order to account for the electrons distributed in these valleys, and correct the electron quasi Fermi level. By considering the band alignment for the LH band, because a reasonable height of the barrier is needed to make the laser efficient at room temperature (e.g. >50 meV), it was found that for different Sn contents in the well using 15% Si and 4% larger Sn content in the barrier than in the well can give a good depth for the well. The range of optimal well widths was obtained by using different Sn compositions and strain in the well. By comparing the dimensionless gain for a couple of combinations across the parameter space, 13-16 nm QW width seems to be a good choice to deliver a large TM mode gain. Then, after setting the well width to a constant value (14 nm), we calculate the maximum gain that can be obtained in the two dimensional parameter space, Sn content and strain. Loss mechanisms are included, because at a particular carrier density the FCA depends only on the photon transition energy, and the major influence on the material net gain may come from IVBA. In the limited parameter space of 0.3% - 1.7% tensile strain and 5% - 11% Sn, a large IVBA can be found around 0.4-0.5 eV transition energy. The optimal choices before and after considering the IVBA show a rather large difference, especially when the carrier density is as large as 5×10^{12} cm^{-2}. A more detailed calculation is done when the wavelength is set to be 3μm, the peak gain dependence on different parameter choices is explicitly given. The optimal

choice for 5×10^{12} cm^{-2} carrier density is 1.41% strain in the well and 5.9% Sn in the well. By combining this choice with a waveguide design, the estimated threshold carrier density is 5.83×10^{12} cm^{-2} and the current density is 1.19 kA/cm^2, which is even comparable to that in some III-V lasers.

As an extension to the existing research on the GeSn QW EAM, in chapter 5 the improved performance of such devices by using an asymmetric geometry is considered. The asymmetric geometry was here limited to the simple intra-step structure, because it only has two important variables, the width ratio and the intra-layer discontinuity (i.e the depth of the intra-well). The variational method is used to deal with the exciton problem, with some calibration, and the calculated results show a good agreement with the experiment. By analysing the QCSE using the perturbation theory, the difference of the QCSE between the square QW and intra-step QW was discussed. By testing the QCSE from 0-100 kV/cm of intra-step QW with different intra-well depth and width ratio, it was found that the largest QCSE for a certain well-depth can be obtained approximately when the first HH state is located at the top edge of the intra-well. The overall performance is considered via the figures of merit $\Delta\alpha/F$ and $\Delta\alpha/F^2$. The optimal structure which gives the largest $\Delta\alpha/F$ is not necessarily the same as the one giving the largest QCSE. After comparing the optimal structure and square QW integrated with a waveguide design, for zero volts on-set, the maximum bandwidth per volt applied is improved by 44%, and the power consumption per bit is reduced by 46%. This improvement still exists when a non-zero on-set (0.5V) is used for the square QW, they are 21% and 12% respectively for these two performance indices. The intra-step QW is obviously more efficient for low voltage operation according to our tests, but considering the potential development of the epitaxial growth technology, the simple intra-step QW may not be the ultimate optimal solution. It is possible that a more complicated intra-layer geometry will bring about even higher performance. However, since it involves more varying parameters, finding it would need more complicated optimization algorithm such as, gradient descent and genetic algorithm. Also, these new geometries may also face some realistic issues, such as higher cost and fabrication limitations.

In summary, the objectives of this project are mostly fulfilled. By using our **k.p** model, we optimize the performance of the tensile strained SiGeSn/GeSn laser and give a reference to the optimal choice of the QW design. EAM using intra-step QW was proposed, which is found to be useful in low voltage on-chip operation. The contribution of this work to the nearly future can be

summarised as follows. Firstly, tensile strain has not been applied to the state-of-the-art GeSn lasers. For example, the first electric pumped laser still uses a high-Sn material in the active region. But by applying tensile strain, the Sn content can be reduced and improve the quality of the layer. With lower defects, the working temperature can be improved. Our calculation can provide some ideas of the choice of the material composition and strain in the structure when such laser is experimentally fabricated. Secondly, by using the intra-step QW, better figures of merit for low voltage operation compared with square QW were found. This indicates that the asymmetric structure could be a new direction to improve the performance of EAM. Future investigation followed by this idea can be conducted by designing more complicated asymmetric structures such as double-intra-step-QW. And finally, the model is a very good starting point of a comprehensive software to design the optoelectronic devices, which can potentially provide an easier approach for researchers to investigate and conduct optimization of different structures and types of devices.

More experimental work is needed in order to continue this research and make the theoretical design applicable to realistic. Starting from the fabrication and characterization of some simple structures such as single QW or bulk material with desired strain. The experimentally measured results using methods like photoluminescence should be compared with the calculation result in order to make sure the model can accurately predict the performance of different structures using similar material recipes. Then after considering carrier transport and self-consistent, more complicated structures such as MQW can be considered theoretically. And finally, we consider the realistic structures, such as a waveguide, and combining the QW design with applicable fabrication process, the prototype of the device (laser or modulator) can be built. Then based on this prototype, we can further test the performance of the laser under different temperatures. Last but not least, more work regarding structural optimization is needed to make the laser meet the demand of daily usage.

The finalization of everything-IV EPIC still has a long way to go, which needs time, funding and research efforts. As one of these steps, we presented a computational model based on 8-band **k.p** and effective mass method to compute and simulate the QW based semiconductor devices. With the simplex techniques for integration, such code can be easily run on a PC without the necessity for HPC clusters. This enables us to study and engineer the QW based devices. Also, with modular design, it can also be improved and potentially extended to some other

research directions without too many modifications of the source code. Within this framework, some improvements and extensions can be considered and conducted in the future, they will be summarised by the following points:

• **Theoretical**

The **k.p** method was originally developed for bulk material. It can also be applied to heterostructures by applying the finite difference scheme. However, problem still exists in the order of the wavevector operator k. As the method of symmetrizing of the operator was argued by Foreman and Luttinger, this work keeps the order of the operators as it was in the original **k.p** matrix, as it is also the case for many other theoretical papers using 8-band **k.p**. Although it seems to be fine to preserve the order with the momentum matrix element Pk_z, by Foreman's theory the order of k_z with the third Luttinger parameter γ_3 should be altered according to the symmetry of the remote bands [68]. He also derived the explicit form for 6-band **k.p** Hamiltonian for QW, but whether this theory can be transferred to our 8-band model and remains a question. Although it was not observed in our calculations, it is reported to be a potential source of the spurious solutions and may 'pollute' the physical results. As it was rarely emphasised in the modelling of Group IV semiconductor devices, more research may be needed on this topic. And surely the model should be modified to improve the accuracy of the results. Another important issue needing to be resolved is to eliminate the existing disagreement of the properties of the alloys in the Group IV. This could be parameters dependence on the alloy composition and temperature, and corrections after considering the realistic effects such as defects density. This will need both theoretical research and experimental work to calculate and characterize the material properties under different circumstances. With more accurate input parameters and maybe fitting functions regarding the alloy composition and temperature, the model will be able to better describe the reality. Lastly, additional extension could also be applied to the existing model. Modelling of realistic effects could also be included in the code to make it a more comprehensive design toolbox, such as to self-consistently calculate the potential profile, to model the carrier transport through the structure, to do mode profile calculation, etc. The extension of the order of the **k.p** matrix is also possible, replacing the effective mass method for other valleys, because it has already been demonstrated that 30-**k.p** is capable of doing the calculation in the whole Brillouin zone. Though it will need longer calculation, it is theoretically possible to apply high-order **k.p** to the quantum well structures.

• **Programming**

There are also some potential improvements in the code itself that can be done in the future. Firstly, the RAM usage. For the convenience of debugging the code, the way it used to store the data in the RAM is not of the highest efficiency. An example is the Hamiltonian after applying the finite difference scheme. In this code it was stored in a sparse matrix, in this way the final wavefunction represented by the eigenvector is continuous. i.e. the first N elements of the eigenvector represent the envelope wavefunction of the $|u_{10}\rangle$ state, the elements N+1 to 2N for the $|u_{20}\rangle$ state, and so on. The matrix can also be reshaped into a band matrix, as has been done in [136]. In a band matrix, all the non-zero values are near the diagonal. With the LAPACK 'zgbcon' subroutine, the input can be a matrix with only these none-zero values. The benefit of this method is to reduce the maximum RAM needed in the process, but the final eigenvector should be rearranged to represent each envelope wave function and it also makes the program hard to debug during the development process. Secondly, the calculation efficiency can be further improved by considering the multi-threading computing. It is tricky to make the code to be multi-threading because the logic is quite different from the traditional programming. This may involve some new techniques, and it may take time to apply them, for example allocating each loop for different thread and memory management. Nevertheless, it is quite important to finally apply these techniques in order to make full use of the high performance cluster in the future. Thirdly, the code now uses MATLAB to generate the graphical output. To make the code more user-friendly for changing the input, a graphical user interface may be considered. Fourthly, because more effects and even higher order **k.p** method may be applied in the future, it is important to make the code ready to be used in the HPC clusters, which may involve some issues such as data storage, interface and so on. Lastly, as a self-developed program, source code may need some standardization to make it easier to read, which is important for the follow-up development and maintenance.

• **Other Applications**

The application of this model does not have to be limited to lasers and modulators. It is also suitable for design of other QW based devices such as LED and photodetectors. With the band structure calculated by **k.p**, necessary optical properties such as spontaneous emission and intersubband absorption can easily be calculated by it by adding some subroutines to include these parts in calculation. Moreover, the **k.p** model can also be used for other types of heterostructures, like quantum wires and quantum dots. Matrix approaches for finite-difference

expansions of simple Schrödinger equation are introduced in some text books (e.g.[49]). Similarly they can be combined with the **k.p** theory. Although modelling of such structures would need more modifications of the code, this might be another interesting direction.

Appendix A

Material Parameters of (Al)GsAs

The material parameters of GaAs, AlAs and AlGaAs used in early validation process of the computational model of this project are given:

Table A.1: Parameters of GaAs, AlAs and AlGaAs [137]

Parameter	Symbol	GaAs	AlAs	B_{AlGaAs}
Lattice constant	a_{lat} (Å)	5.65325	5.66110	
Deformation potential(CB)	a_c(eV)	-7.17	-5.64	
Deformation potential(VB)	a_v(eV)	-1.16	-2.47	
Deformation potential(shear)	b(eV)	-2.0	-2.3	
Deformation potential(shear)	d(eV)	-4.8	-3.4	
Deformation potential(L)	a_L(eV)	-0.66	-1.54	
Stiffness constant	C_{11}(GPa)	1221	1250	
Stiffness constant	C_{12}(GPa)	566	534	
Bandgap (Γ)	E_g^{Γ}(eV)	1.519	3.099	$-0.127 + 1.310x$[a]
Bandgap (L)	E_g^{L}(eV)	1.815	2.46	0.055
Split-off Energy	Δ_{so}(eV)	0.341	0.28	
Luttinger(6-bands)	γ_1^L	6.98	3.76	
Luttinger(6-bands)	γ_2^L	2.06	0.82	
Luttinger(6-bands)	γ_3^L	2.93	1.42	
MME	E_p(eV)	28.8	21.1	
VBO	E_v(eV)	-0.80	-1.33	
Effective mass(Γ)	m_e^{Γ}	0.067	0.15	
Longi. Effective mass(L)	m_l^L	1.9	1.32	
Trans. Effective mass(L)	m_t^L	0.0754	0.15	
Refractive index	n_r	3.5	2.5	

[a] The bowing depends on the alloy composition. x is the content of Aluminium in $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ alloy

Appendix B

Program Layout

Fig. B.1 shows the general flow chart of the program. The part in the blue square is executed twice, the first is to roughly calculate the quasi-Fermi level and give a guidance to the band filter subroutine and to judge if more bands are needed to accurately calculate the quasi-Fermi level. The yellow square contains the parts used for the simulation for laser, and the green square is for the modulator modelling, to solve the exciton problem.

For most of the calculations, the code is written in FORTRAN, and the results are plotted using MATLAB. Data are transmitted from FORTRAN to MATLAB through Fortran Matrix API [138].

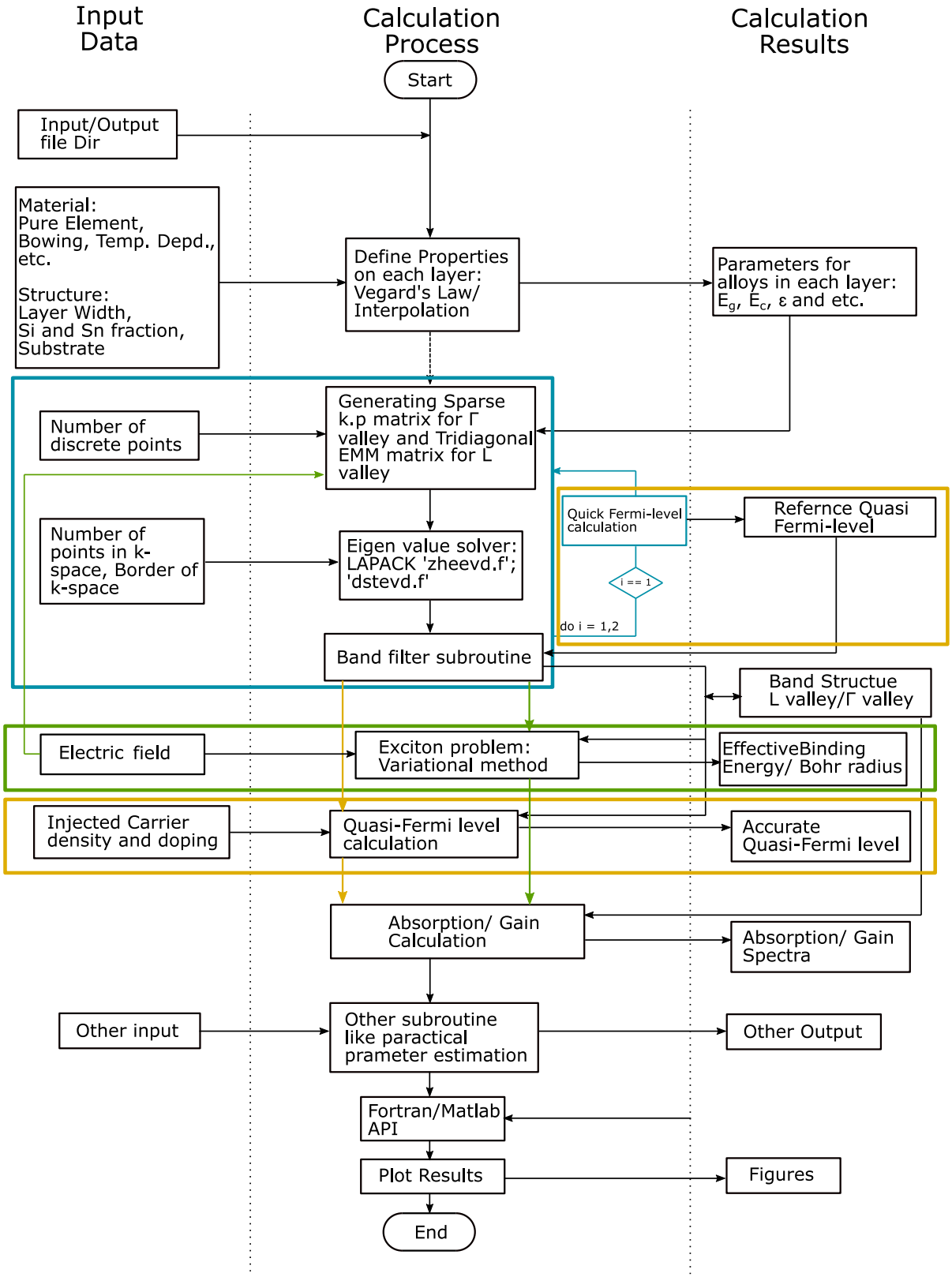

Figure B.1: Program layout of the computation model

.

References

[1] D. Liang and J. E. Bowers, "Recent progress in lasers on silicon," *Nature photonics*, vol. 4, no. 8, pp. 511–517, 2010.

[2] D. Nasledov, A. Rogachev, S. Ryvkin, *et al.*, *Recombination radiation of gallium arsenide*, 1962.

[3] M. Tang, J.-S. Park, Z. Wang, *et al.*, "Integration of III-V lasers on Si for Si photonics," *Progress in Quantum Electronics*, vol. 66, pp. 1–18, 2019.

[4] M. Aoki, M. Suzuki, H. Sano, *et al.*, "InGaAs/InGaAsP MQW electroabsorption modulator integrated with a DFB laser fabricated by band-gap energy control selective area MOCVD," *IEEE Journal of Quantum Electronics*, vol. 29, no. 6, pp. 2088–2096, 1993.

[5] G. Liu, A. Stintz, H. Li, *et al.*, "Extremely low room-temperature threshold current density diode lasers using InAs dots in $In_{0.15}Ga_{0.85}As$ quantum well," *Electronics Letters*, vol. 35, no. 14, pp. 1163–1165, 1999.

[6] R. Soref, D. Buca, and S.-Q. Yu, "Group IV photonics: driving integrated optoelectronics," *Optics and Photonics News*, vol. 27, no. 1, pp. 32–39, 2016.

[7] J. Liu, X. Sun, R. Camacho-Aguilera, *et al.*, "Ge-on-Si laser operating at room temperature," *Optics Letters*, vol. 35, no. 5, pp. 679–681, 2010.

[8] L. Vivien, J. Osmond, J.-M. Fédéli, *et al.*, "42 GHz pin Germanium photodetector integrated in a silicon-on-insulator waveguide," *Optics Express*, vol. 17, no. 8, pp. 6252–6257, 2009.

[9] J. E. Roth, O. Fidaner, R. K. Schaevitz, *et al.*, "Optical modulator on silicon employing germanium quantum wells," *Optics Express*, vol. 15, no. 9, pp. 5851–5859, 2007.

[10] C. Goodman, "Direct-gap group IV semiconductors based on tin," *IEE Proceedings I (Solid-State and Electron Devices)*, vol. 129, no. 5, pp. 189–192, 1982.

[11] D. W. Jenkins and J. D. Dow, "Electronic properties of metastable Ge_xSn_{1-x} alloys," *Physical Review B*, vol. 36, no. 15, p. 7994, 1987.

[12] J. Zheng, Z. Liu, C. Xue, *et al.*, "Recent progress in GeSn growth and GeSn-based photonic devices," *Journal of Semiconductors*, vol. 39, no. 6, p. 061 006, 2018.

[13] G. He and H. A. Atwater, "Interband transitions in Sn_xGe_{1-x} alloys," *Physical Review Letters*, vol. 79, no. 10, p. 1937, 1997.

[14] O. Moutanabbir, S. Assali, X. Gong, *et al.*, "Monolithic infrared silicon photonics: the rise of (Si) GeSn semiconductors," *Applied Physics Letters*, vol. 118, no. 11, p. 110 502, 2021.

[15] G. Abernathy, Y. Zhou, S. Ojo, *et al.*, "Study of SiGeSn/GeSn single quantum well toward high-performance all-group-IV optoelectronics," *Journal of Applied Physics*, vol. 129, no. 9, p. 093 105, 2021.

[16] S. Wirths, R. Geiger, N. Von Den Driesch, *et al.*, "Lasing in direct-bandgap GeSn alloy grown on Si," *Nature Photonics*, vol. 9, no. 2, pp. 88–92, 2015.

[17] Q. M. Thai, N. Pauc, J. Aubin, *et al.*, "GeSn heterostructure micro-disk laser operating at 230 K," *Optics Express*, vol. 26, no. 25, pp. 32 500–32 508, 2018.

[18] J. Margetis, Y. Zhou, W. Dou, *et al.*, "All group-IV SiGeSn/GeSn/SiGeSn QW laser on Si operating up to 90 K," *Applied Physics Letters*, vol. 113, no. 22, p. 221 104, 2018.

[19] A. Elbaz, D. Buca, N. von den Driesch, *et al.*, "Ultra-low-threshold continuous-wave and pulsed lasing in tensile-strained GeSn alloys," *Nature Photonics*, vol. 14, no. 6, pp. 375–382, 2020.

[20] Y. Zhou, Y. Miao, S. Ojo, *et al.*, "Electrically injected GeSn lasers on Si operating up to 100 K," *Optica*, vol. 7, no. 8, pp. 924–928, 2020.

[21] M. Oehme, K. Kostecki, T. Arguirov, *et al.*, "GeSn heterojunction LEDs on Si substrates," *IEEE Photonics Technology Letters*, vol. 26, no. 2, pp. 187–189, 2013.

[22] H. Tseng, K. Wu, H. Li, *et al.*, "Mid-infrared electroluminescence from a Ge/$Ge_{0.922}Sn_{0.078}$/Ge double heterostructure pin diode on a Si substrate," *Applied Physics Letters*, vol. 102, no. 18, p. 182 106, 2013.

[23] D. Stange, N. Von Den Driesch, D. Rainko, *et al.*, "Study of GeSn based heterostructures: towards optimized group IV MQW LEDs," *Optics Express*, vol. 24, no. 2, pp. 1358–1367, 2016.

[24] Y. Dong, W. Wang, D. Lei, *et al.*, "Suppression of dark current in germanium-tin on silicon pin photodiode by a silicon surface passivation technique," *Optics Express*, vol. 23, no. 14, pp. 18 611–18 619, 2015.

[25] M. Oehme, D. Widmann, K. Kostecki, *et al.*, "GeSn/Ge multiquantum well photodetectors on Si substrates," *Optics letters*, vol. 39, no. 16, pp. 4711–4714, 2014.

[26] J. Zheng, S. Wang, Z. Liu, *et al.*, "GeSn pin photodetectors with GeSn layer grown by magnetron sputtering epitaxy," *Applied Physics Letters*, vol. 108, no. 3, p. 033 503, 2016.

[27] M. Akie, T. Fujisawa, T. Sato, *et al.*, "GeSn/SiGeSn multiple-quantum-well electroabsorption modulator with taper coupler for mid-infrared Ge-on-Si platform," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 6, pp. 1–8, 2018.

[28] R. Soref, G. Sun, and H. Cheng, "Franz-Keldysh electro-absorption modulation in germanium-tin alloys," *Journal of Applied Physics*, vol. 111, no. 12, p. 123 113, 2012.

[29] R. Ponce, S. S. Azadeh, D. Stange, *et al.*, "Design of a high-speed germanium-tin absorption modulator at mid-infrared wavelengths," in *2017 IEEE 14th International Conference on Group IV Photonics (GFP)*, IEEE, 2017, pp. 19–20.

[30] Y.-D. Hsieh, J.-H. Lin, R. Soref, *et al.*, "Electro-absorption modulation in GeSn alloys for wide-spectrum mid-infrared applications," *Communications Materials*, vol. 2, no. 1, pp. 1–8, 2021.

[31] T. Brudevoll, D. Citrin, N. Christensen, *et al.*, "Calculated band structure of zinc-blende-type SnGe," *Physical Review B*, vol. 48, no. 23, p. 17 128, 1993.

[32] J. Shen, J. Zi, X. Xie, *et al.*, "Ab initio calculation of the structure of the random alloys Sn_xGe_{1-x}," *Physical Review B*, vol. 56, no. 19, p. 12 084, 1997.

[33] M. Sahnoun, R. Khenata, H. Baltache, *et al.*, "First-principles calculations of optical properties of GeC, SnC and GeSn under hydrostatic pressure," *Physica B: Condensed Matter*, vol. 355, no. 1-4, pp. 392–400, 2005.

[34] X. Zhang, C. Ying, Z. Li, *et al.*, "First-principles calculations of structural stability, elastic, dynamical and thermodynamic properties of SiGe, SiSn, GeSn," *Superlattices and Microstructures*, vol. 52, no. 3, pp. 459–469, 2012.

[35] M. Polak, P. Scharoch, and R. Kudrawiec, "The electronic band structure of $Ge_{1-x}Sn_x$ in the full composition range: indirect, direct, and inverted gaps regimes, band offsets, and the Burstein–Moss effect," *Journal of Physics D: Applied Physics*, vol. 50, no. 19, p. 195 103, 2017.

[36] P. Moontragoon, R. Soref, and Z. Ikonic, "The direct and indirect bandgaps of unstrained $Si_xGe_{1-x-y}Sn_y$ and their photonic device applications," *Journal of Applied Physics*, vol. 112, no. 7, p. 073 106, 2012.

[37] S. Gupta, B. Magyari-Köpe, Y. Nishi, *et al.*, "Achieving direct band gap in germanium through integration of Sn alloying and external strain," *Journal of Applied Physics*, vol. 113, no. 7, p. 073 707, 2013.

[38] K. Lu Low, Y. Yang, G. Han, *et al.*, "Electronic band structure and effective mass parameters of $Ge_{1-x}Sn_x$ alloys," *Journal of Applied Physics*, vol. 112, no. 10, p. 103 715, 2012.

[39] S.-Q. Liu and S.-T. Yen, "Extraction of eight-band k.p parameters from empirical pseudopotentials for GeSn," *Journal of Applied Physics*, vol. 125, no. 24, p. 245 701, 2019.

[40] Y. Zhu, Q. Xu, W. Fan, *et al.*, "Theoretical gain of strained $GeSn_{0.02}/Ge_{1-x-y}Si_xSn_y$ quantum well laser," *Journal of Applied Physics*, vol. 107, no. 7, p. 073 108, 2010.

[41] H. Maczko, R. Kudrawiec, and M. Gladysiewicz, "Material gain engineering in GeSn/Ge quantum wells integrated with an Si platform," *Scientific Reports*, vol. 6, no. 1, pp. 1–11, 2016.

[42] Y. Liu, C. Fang, X. Gao, *et al.*, "Theoretical Investigation of Tensile-Strained GeSn/SiGeSn Multiple Quantum Well Laser Wrapped in Si_3N_4 Linear Stressor," *IEEE Photonics Journal*, vol. 10, no. 1, pp. 1–9, 2017.

[43] B. Marzban, D. Stange, D. Rainko, *et al.*, "Modeling of a SiGeSn quantum well laser," *Photonics Research*, vol. 9, no. 7, pp. 1234–1254, 2021.

[44] Z. Song, W. Fan, C. S. Tan, *et al.*, "Band structure of $Ge_{1-x}Sn_x$ alloy: a full-zone 30-band **k.p** model," *New Journal of Physics*, vol. 21, no. 7, p. 073 037, 2019.

[45] S. Wirths, Z. Ikonic, A. Tiedemann, *et al.*, "Tensely strained GeSn alloys as optical gain media," *Applied Physics Letters*, vol. 103, no. 19, p. 192 110, 2013.

[46] J. R. Jain, A. Hryciw, T. M. Baer, *et al.*, "A micromachining-based technology for enhancing germanium light emission via tensile strain," *Nature Photonics*, vol. 6, no. 6, pp. 398–405, 2012.

[47] G. Dushaq, B. Paredes, and M. Rasras, "Strong enhancement of direct transition photoluminescence at room temperature for highly tensile-strained Ge decorated using 5 nm gold nanoparticles," *Nanotechnology*, vol. 31, no. 31, p. 315 201, 2020.

[48] N. Yahyaoui, N. Sfina, J.-L. Lazzari, *et al.*, "Stark shift of the absorption spectra in $Ge/Ge_{1-x}Sn_x/Ge$ type-I single QW cell for mid-wavelength infra-red modulators," *Superlattices and Microstructures*, vol. 85, pp. 629–637, 2015.

[49] P. Harrison and A. Valavanis, *Quantum wells, wires and dots: theoretical and computational physics of semiconductor nanostructures*. John Wiley & Sons, 2016.

[50] E. Anderson, *LAPACK Users' Guide, 3rd ed.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.

[51] A. Rahman, M. S. Lundstrom, and A. W. Ghosh, "Generalized effective-mass approach for n-type metal-oxide-semiconductor field-effect transistors on arbitrarily oriented wafers," *Journal of Applied Physics*, vol. 97, no. 5, p. 053 702, 2005.

[52] S. L. Chuang, *Physics of photonic devices*. John Wiley & Sons, 2012, vol. 80.

[53] T. B. Bahder, "Eight-band k.p model of strained zinc-blende crystals," *Physical Review B*, vol. 41, no. 17, p. 11 992, 1990.

[54] E. O. Kane, "Band structure of indium antimonide," *Journal of Physics and Chemistry of Solids*, vol. 1, no. 4, p. 249, 1957.

[55] G. E. Pikus and G. L. Bir, "Effect of deformation on the hole energy spectrum of germanium and silicon," *Soviet Physics-Solid State*, vol. 1, p. 1502, 1960.

[56] P.-O. Löwdin, "A note on the quantum-mechanical perturbation theory," *The Journal of Chemical Physics*, vol. 19, no. 11, pp. 1396–1401, 1951.

[57] J. M. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," *Physical Review*, vol. 97, no. 4, p. 869, 1955.

[58] C. R. Pidgeon and R. Brown, "Interband magneto-absorption and faraday rotation in insb," *Physical Review*, vol. 146, no. 2, p. 575, 1966.

[59] G. Dresselhaus, A. Kip, and C. Kittel, "Cyclotron resonance of electrons and holes in silicon and germanium crystals," *Physical Review*, vol. 98, no. 2, p. 368, 1955.

[60] J. Los, A. Fasolino, and A. Catellani, "Generalization of the k.p approach for strained layered semiconductor structures grown on high-index-planes," *Physical Review B*, vol. 53, no. 8, p. 4630, 1996.

[61] M. Willatzen and L. C. L. Y. Voon, *The k.p method*. Springer Berlin Heidelberg, 2009.

[62] C. Boztug, J. R. Sánchez-Pérez, F. Cavallo, *et al.*, "Strained-germanium nanostructures for infrared photonics," *ACS nano*, vol. 8, no. 4, pp. 3136–3151, 2014.

[63] J. C. Hensel and G. Feher, "Cyclotron resonance experiments in uniaxially stressed silicon: Valence band inverse mass parameters and deformation potentials," *Physical Review*, vol. 129, no. 3, p. 1041, 1963.

[64] Y. Huo, H. Lin, R. Chen, *et al.*, "Strong enhancement of direct transition photoluminescence with highly tensile-strained ge grown by molecular beam epitaxy," *Applied Physics Letters*, vol. 98, no. 1, p. 011 111, 2011.

[65] H. Tahini, A. Chroneos, R. W. Grimes, *et al.*, "Strain-induced changes to the electronic structure of germanium," *Journal of Physics: Condensed Matter*, vol. 24, no. 19, p. 195 802, 2012.

[66] J. M. Luttinger, "Quantum theory of cyclotron resonance in semiconductors: General theory," *Physical Review*, vol. 102, no. 4, p. 1030, 1956.

[67] B. A. Foreman, "Elimination of spurious solutions from eight-band k.p theory," *Physical Review B*, vol. 56, no. 20, R12748, 1997.

[68] B. A. Foreman, "Effective-mass hamiltonian and boundary conditions for the valence bands of semiconductor microstructures," *Physical Review B*, vol. 48, no. 7, p. 4964, 1993.

[69] S. L. Chuang, "Efficient band-structure calculations of strained quantum wells," *Physical Review B*, vol. 43, no. 12, p. 9649, 1991.

[70] J. Gao, H. Zhou, J. Jiang, *et al.*, "Design of low bias voltage Ge/SiGe multiple quantum wells electro-absorption modulator at 1550 nm," *AIP Advances*, vol. 7, no. 3, p. 035 317, 2017.

[71] C. Y.-P. Chao and S. L. Chuang, "Spin-orbit-coupling effects on the valence-band structure of strained semiconductor quantum wells," *Physical Review B*, vol. 46, no. 7, p. 4110, 1992.

[72] R. G. Veprek, S. Steiger, and B. Witzigmann, "Ellipticity and the spurious solution problem of k.p envelope equations," *Physical Review B*, vol. 76, no. 16, p. 165 320, 2007.

[73] Y. Jiang, X. Ma, Y. Xu, *et al.*, "Finite difference method for analyzing band structure in semiconductor heterostructures without spurious solutions," *Journal of Applied Physics*, vol. 116, no. 17, p. 173 702, 2014.

[74] X. Ma, K. Li, Z. Zhang, *et al.*, "Stable finite element method of eight-band k.p model without spurious solutions and numerical study of interfaces in heterostructures," *Journal of Applied Physics*, vol. 116, no. 23, p. 235 702, 2014.

[75] A. Meney, B. Gonul, and E. O'Reilly, "Evaluation of various approximations used in the envelope-function method," *Physical Review B*, vol. 50, no. 15, p. 10 893, 1994.

[76] K. Kolokolov, J. Li, and C. Ning, "**k.p** hamiltonian without spurious-state solutions," *Physical Review B*, vol. 68, no. 16, p. 161 308, 2003.

[77] S. Birner, "Modeling of semiconductor nanostructures and semiconductor-electrolyte interfaces," 2011.

[78] D. Rainko, Z. Ikonic, N. Vukmirović, *et al.*, "Investigation of carrier confinement in direct bandgap GeSn/SiGeSn 2D and 0D heterostructures," *Scientific Reports*, vol. 8, no. 1, pp. 1–13, 2018.

[79] J. Kash, M. Zachau, M. Tischler, *et al.*, "Optical measurements of warped valence bands in quantum wells," *Surface science*, vol. 305, no. 1-3, pp. 251–255, 1994.

[80] J. D. Jackson, *Classical electrodynamics*, 1999.

[81] R. Winkler, "A note on analytic quadratic Brillouin zone integration," *Journal of Physics: Condensed Matter*, vol. 5, no. 15, p. 2321, 1993.

[82] R. Soref, "Mid-infrared photonics in silicon and germanium," *Nature Photonics*, vol. 4, no. 8, pp. 495–497, 2010.

[83] J. Liu, X. Sun, D. Pan, *et al.*, "Tensile-strained, n-type Ge as a gain medium for monolithic laser integration on Si," *Optics Express*, vol. 15, no. 18, pp. 11 272–11 277, 2007.

[84] D. Stange, N. von den Driesch, T. Zabel, *et al.*, "GeSn/SiGeSn heterostructure and multi quantum well lasers," *ACS Photonics*, vol. 5, no. 11, pp. 4628–4636, 2018.

[85] G.-E. Chang and C.-O. Chang, "Tensile-strained Ge/SiGeSn quantum wells for polarization insensitive electroabsorption waveguide modulators," *IEEE Journal of Quantum Electronics*, vol. 48, no. 4, pp. 533–541, 2012.

[86] Z. Chen, Z. Ikonic, D. Indjin, *et al.*, "Design optimization of tensile-strained SiGeSn/GeSn quantum wells at room temperature," *Journal of Applied Physics*, vol. 129, p. 123 102, 2021.

[87] A. Sadao, "Properties of group-IV, III-V and II-VI semiconductors," *Hoboken (USA) Wiley & Sons*, 2005.

[88] D. Paul, "8-band **k.p** modelling of mid-infrared intersubband absorption in Ge quantum wells," *Journal of Applied Physics*, vol. 120, no. 4, p. 043 103, 2016.

[89] G. Chang, S. Chang, and S. L. Chuang, "Strain balanced $Ge_zSn_{1-z}/Si_xGe_ySn_{1-x-y}$ Multiple-Quantum-Well Lasers," *IEEE journal of Quantum Electronics*, vol. 46, no. 12, pp. 1813–1820, 2010.

[90] C. Van de Walle, "4.5 SiGe heterojunctions and band offsets," 1999.

[91] H. Li, "Refractive index of silicon and germanium and its wavelength and temperature derivatives," *Journal of Physical and Chemical Reference Data*, vol. 9, no. 3, pp. 561–658, 1980.

[92] A. Golovashkin and G. Motulevich, "Optical and electrical properties of tin," *Sov. Phys. JETP*, vol. 19, no. 2, pp. 310–317, 1964.

[93] M. S. Shur, *Handbook series on semiconductor parameters*. World Scientific, 1996, vol. 1.

[94] P. A. Khomyakov, M. Luisier, and A. Schenk, "Compositional bowing of band energies and their deformation potentials in strained ingaas ternary alloys: A first-principles study," *Applied Physics Letters*, vol. 107, no. 6, p. 062 104, 2015.

[95] H. Langueur, K. Kassali, and N. Lebgaa, "Density functional study of structural, mechanic, thermodynamic and dynamic properties of SiGe alloys," *Journal of Computational and Theoretical Nanoscience*, vol. 10, no. 1, pp. 86–94, 2013.

[96] N. Bouarissa and F. Annane, "Electronic properties and elastic constants of the ordered $Ge_{1-x}Sn_x$ alloys," *Materials Science and Engineering: B*, vol. 95, no. 2, pp. 100–106, 2002.

[97] K. Zelazna, M. Wełna, J. Misiewicz, *et al.*, "Temperature dependence of energy gap of $Ge_{1-x}Sn_x$ alloys with x $<$ 0.11 studied by photoreflectance," *Journal of Physics D: Applied Physics*, vol. 49, no. 23, p. 235 301, 2016.

[98] M. Bertrand, Q.-M. Thai, J. Chrétien, *et al.*, "Experimental Calibration of Sn-Related Varshni Parameters for High Sn Content GeSn Layers," *Annalen der Physik*, vol. 531, no. 6, p. 1 800 396, 2019.

[99] M. Jaros, "Simple analytic model for heterojunction band offsets," *Physical Review B*, vol. 37, no. 12, p. 7112, 1988.

[100] M.-Y. Ryu, T. R. Harris, Y. Yeo, *et al.*, "Temperature-dependent photoluminescence of Ge/Si and $Ge_{1-y}Sn_y$/Si, indicating possible indirect-to-direct bandgap transition at lower Sn content," *Applied Physics Letters*, vol. 102, no. 17, p. 171 908, 2013.

[101] N. von den Driesch, C. Stampfer, and S. Mantl, "Epitaxy of group IV Si-Ge-Sn alloys for advanced heterostructure light emitters," Fachgruppe Physik, Tech. Rep., 2018.

[102] J. Tolle, A. Chizmeshya, Y.-Y. Fang, *et al.*, "Low temperature chemical vapor deposition of Si-based compounds via SiH_3 SiH_2 SiH_3: Metastable SiSn/ GeSn/ Si (100) heteroepitaxial structures," *Applied Physics Letters*, vol. 89, no. 23, p. 231 924, 2006.

[103] S.-W. Chang and S. L. Chuang, "Theory of Optical Gain of Ge-$Si_xGe_ySn_{1-x-y}$ Quantum-Well Lasers," *IEEE Journal of Quantum Electronics*, vol. 43, no. 3, pp. 249–256, 2007.

[104] R. Winkler, M. Merkler, T. Darnhofer, *et al.*, "Theory for the cyclotron resonance of holes in strained asymmetric Ge-SiGe quantum wells," *Physical Review B*, vol. 53, no. 16, p. 10 858, 1996.

[105] Y.-H. Kuo, Y. K. Lee, Y. Ge, *et al.*, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," *Nature*, vol. 437, no. 7063, pp. 1334–1336, 2005.

[106] Y.-H. Li, A. Walsh, S. Chen, *et al.*, "Revised ab initio natural band offsets of all group IV, II-VI, and III-V semiconductors," *Applied Physics Letters*, vol. 94, no. 21, p. 212 109, 2009.

[107] M. Asada, A. Adams, K. Stubkjaer, *et al.*, "The temperature dependence of the threshold current of GaInAsP/InP DH lasers," *IEEE Journal of Quantum Electronics*, vol. 17, no. 5, pp. 611–619, 1981.

[108] J. van der Heide, N. Posthuma, G. Flamand, *et al.*, "Cost-efficient thermophotovoltaic cells based on germanium substrates," *Solar Energy Materials and Solar Cells*, vol. 93, no. 10, pp. 1810–1816, 2009.

[109] G. Weimann and W. Schlapp, "GaAs AlGaAs MQW and GRINSCH lasers grown by molecular beam epitaxy," *Physica B+ C*, vol. 129, no. 1-3, pp. 459–464, 1985.

[110] T. Egawa, Y. Murata, T. Jimbo, *et al.*, "Low-temperature continuous-wave operation of AlGaAs-GaAs vertical-cavity surface-emitting lasers on Si substrates," *IEEE Photonics Technology Letters*, vol. 9, no. 7, pp. 872–874, 1997.

[111] G.-E. Chang and C.-O. Chang, "Design of strain-free GeSn/SiGeSn quantum-well electroabsorption modulators at 1550 nm wavelength," in *7th IEEE International Conference on Group IV Photonics*, IEEE, 2010, pp. 87–89.

[112] L. Lever, Z. Ikonic, A. Valavanis, *et al.*, "Design of Ge–SiGe quantum-confined Stark effect electroabsorption heterostructures for CMOS compatible photonics," *Journal of Lightwave Technology*, vol. 28, no. 22, pp. 3273–3281, 2010.

[113] R. K. Schaevitz, J. E. Roth, S. Ren, *et al.*, "Material properties of Si-Ge/Ge quantum wells," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 14, no. 4, pp. 1082–1089, 2008.

[114] T. Wood, C. Burrus, D. Miller, *et al.*, "High-speed optical modulation with GaAs/-GaAlAs quantum wells in a p-i-n diode structure," *Applied Physics Letters*, vol. 44, no. 1, pp. 16–18, 1984.

[115] D. A. Miller, D. Chemla, T. Damen, *et al.*, "Band-edge electroabsorption in quantum well structures: The quantum-confined Stark effect," *Physical Review Letters*, vol. 53, no. 22, p. 2173, 1984.

[116] T. Fujisawa and K. Saitoh, "Quantum-confined Stark effect analysis of GeSn/SiGeSn quantum wells for mid-infrared Si-based electroabsorption devices based on many-body theory," *IEEE Journal of Quantum Electronics*, vol. 51, no. 11, pp. 1–7, 2015.

[117] M. Morita, K. Goto, and T. Suzuki, "Quantum-confined Stark effect in stepped-potential quantum wells," *Japanese Journal of Applied Physics*, vol. 29, no. 9A, p. L1663, 1990.

[118] T. Tütken, B. Hawdon, M. Zimmermann, *et al.*, "Large observed exciton shifts with electric field in InGaAs/InGaAsP stepped quantum wells," *Applied Physics Letters*, vol. 63, no. 8, pp. 1086–1088, 1993.

[119] D.-S. Shin, P. K. Yu, and S. Pappert, "High-power electroabsorption modulator using intra-step-barrier quantum wells," *Journal of Applied Physics*, vol. 89, no. 2, pp. 1515–1517, 2001.

[120] T. Woodward, J. Cunningham, and W. Jan, "Comparison of stepped-well and square-well multiple-quantum-well optical modulators," *Journal of Applied Physics*, vol. 78, no. 3, pp. 1411–1414, 1995.

[121] Z. Chen, Z. Ikonic, D. Indjin, *et al.*, "Design considerations of intra-step SiGeSn/GeSn quantum well electroabsorption modulators," *Journal of Applied Physics*, vol. 130, no. 15, p. 153 103, 2021.

[122] P. Mares and S. Chuang, "Modeling of self-electro-optic-effect devices," *Journal of Applied Physics*, vol. 74, no. 2, pp. 1388–1397, 1993.

[123] S.-L. Chuang, S. Schmitt-Rink, D. A. Miller, *et al.*, "Exciton Green's-function approach to optical absorption in a quantum well with an applied electric field," *Physical Review B*, vol. 43, no. 2, p. 1500, 1991.

[124] Y.-H. Kuo and Y.-S. Li, "Variational calculation for the direct-gap exciton in the Ge quantum well systems," *Physical Review B*, vol. 79, no. 24, p. 245 328, 2009.

[125] A. McLean, C. Mitchell, and D. Swanston, "Implementation of an efficient analytical approximation to the Voigt function for photoemission lineshape analysis," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 69, no. 2, pp. 125–132, 1994.

[126] C. Lange, N. Köster, S. Chatterjee, *et al.*, "Ultrafast nonlinear optical response of photoexcited Ge/SiGe quantum wells: Evidence for a femtosecond transient population inversion," *Physical Review B*, vol. 79, no. 20, p. 201 306, 2009.

[127] S. A. Claussen, E. Tasyurek, J. E. Roth, *et al.*, "Measurement and modeling of ultrafast carrier dynamics and transport in germanium/silicon-germanium quantum wells," *Optics Express*, vol. 18, no. 25, pp. 25 596–25 607, 2010.

[128] R. Schaevitz, D. Ly-Gagnon, J. Roth, *et al.*, "Indirect absorption in germanium quantum wells," *AIP Advances*, vol. 1, no. 3, p. 032 164, 2011.

[129] G. Macfarlane and V. Roberts, "Infrared absorption of germanium near the lattice edge," *Physical Review*, vol. 97, no. 6, p. 1714, 1955.

[130] D. Shin, W. Chen, S. Pappert, *et al.*, "Analysis of intra-step-barrier quantum wells for high-power electroabsorption modulators," in *2001 International Topical Meeting on Microwave Photonics. Technical Digest. MWP'01 (Cat. No. 01EX476)*, IEEE, 2002, pp. 17–20.

[131] C. Yuan-Bing, P. Jiao-Qing, Z. Fan, *et al.*, "High-power electroabsorption modulator using intrastep quantum well," *Chinese Physics Letters*, vol. 24, no. 7, p. 2128, 2007.

[132] M. Chin, "On the figures of merit for electroabsorption waveguide modulators," *IEEE Photonics Technology Letters*, vol. 4, no. 7, pp. 726–728, 1992.

[133] M. Chin and W. S. Chang, "Theoretical design optimization of multiple-quantum-well electroabsorption waveguide modulators," *IEEE Journal of Quantum Electronics*, vol. 29, no. 9, pp. 2476–2488, 1993.

[134] J. Thalken, W. Li, S. Haas, *et al.*, "Adaptive design of excitonic absorption in broken-symmetry quantum wells," *Applied Physics Letters*, vol. 85, no. 1, pp. 121–123, 2004.

[135] H. Takeuchi, K. Tsuzuki, K. Sato, *et al.*, "Very high-speed light-source module up to 40 Gb/s containing an MQW electroabsorption modulator integrated with a DFB laser," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 3, no. 2, pp. 336–343, 1997.

[136] M. Marchewka, "Finite-difference method applied for eight-band **k.p** model for $hg_{1-x}cd_x$ te/hgte quantum well," *International Journal of Modern Physics B*, vol. 31, no. 20, p. 1 750 137, 2017.

[137] I. Vurgaftman, J. á. Meyer, and L. á. Ram-Mohan, "Band parameters for III–V compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

[138] *MATLAB Fortran API Libraries*. [Online]. Available: `https://www.mathworks.com/help/matlab/matlab_external/matlab-fortran-api-libraries.html`.