High-Quality Extended-Wavelength Materials for Optoelectronic Applications

by

Ankur R. Sharma

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved February 2013 by the
Graduate Supervisory Committee:

Shane Johnson, Chair
Michael Goryll
Ronald Roedel

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

Photodetectors in the 1.7 to 4.0 μm range are being commercially developed on InP substrates to meet the needs of longer wavelength applications such as thermal and medical sensing. Currently, these devices utilize high indium content metamorphic $Ga_{1-x}In_xAs$ ($x > 0.53$) layers to extend the wavelength range beyond the 1.7 μm achievable using lattice matched GaInAs. The large lattice mismatch required to reach the extended wavelengths results in photodetector materials that contain a large number of misfit dislocations. The low quality of these materials results in a large nonradiative Shockley Read Hall generation/recombination rate that is manifested as an undesirable large thermal noise level in these photodetectors. This work focuses on utilizing the different band structure engineering methods to design more efficient devices on InP substrates.

One prospective way to improve photodetector performance at the extended wavelengths is to utilize lattice matched GaInAs/GaAsSb structures that have a type-II band alignment, where the ground state transition energy of the superlattice is smaller than the bandgap of either constituent material. Over the extended wavelength range of 2 to 3 μm this superlattice structure has an optimal period thickness of 3.4 to 5.2 nm and a wavefunction overlap of 0.8 to 0.4, respectively. In using a type-II superlattice to extend the cutoff wavelength there is a tradeoff between the wavelength reached and the electron-hole wavefunction overlap realized, and hence absorption coefficient achieved. This tradeoff and the subsequent reduction in performance can be overcome by two methods: adding

bismuth to this type-II material system; applying strain on both layers in the system to attain strain-balanced condition. These allow the valance band alignment and hence the wavefunction overlap to be tuned independently of the wavelength cutoff.

Adding 3% bismuth to the GaInAs constituent material, the resulting lattice matched $Ga_{0.516}In_{0.484}As_{0.970}Bi_{0.030}/GaAs_{0.511}Sb_{0.489}$ superlattice realizes a 50% larger absorption coefficient. While as, similar results can be achieved with strain-balanced condition with strain limited to 1.9% on either layer. The optimal design rules derived from the different possibilities make it feasible to extract superlattice period thickness with the best absorption coefficient for any cutoff wavelength in the range.

DEDICATION

This work is dedicated to my loving mother and father, for all their sacrifices and the blessings they have bestowed on me for years.

ACKNOWLEDGEMENTS

I am greatly indebted to my advisor Dr. Shane Johnson for his extended support and intellectual guidance during my graduate studies at Arizona State University. I have and will continue to benefit from his teachings, not limited to the realm of scientific work. I am also thankful to Dr. Ronald Roedel and Dr. Michael Goryll for their assistance and the interest they have committed towards the completion of this work. I take this opportunity and remember Dr. Dieter Schroder who instilled the interest of semiconductors in me with his encyclopedic knowledge and vast experience.

My graduate work would not have been a journey of joy without the support of many others. Chaturvedi Gogineni and Nathaniel Riordan have been a great source of information throughout my graduate school. I am grateful to have a research peer like Huan Liang who has motivated me throughout the duration of this work. My work has also been enriched with the inputs Preston Webster has provided with all simulations carried out in this work.

Last but not the least I would like to thank my family and friends who have always believed in me and nourished me with all the strength essential for completion of this work. I wish to extend a special thanks to my sister, Monisha who has always revitalized me from my failures.

TABLE OF CONTENTS

Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

LIST OF ACRONYMS

BAC – band anticrossing

RBS – Rutherford backscattering

XRD – X-Ray diffraction

PL - Photoluminescence

1 INTRODUCTION

Tremendous amount of research has been carried out in the field of semiconductor alloys to suffice the need of low cost and efficient optoelectronic devices. Optoelectronic devices are broadly classified as emitters and detectors. Emitters such as light emitting diodes emit photons at a wavelength proportional to the bandgap of the semiconductor material used to form the PN junction. Similarly, a solar cell absorbs photons from the sunlight with energy greater than or equal to the bandgap of the material utilized in the absorbing layer to generate an equivalent electrical signal. When the source of photon energy is not necessarily sunlight the device is generally termed as a photodetector. Thus the operation wavelength of optoelectronic devices is largely dependent on the bandgap of the semiconductor material used. It is important to have materials with bandgaps that can cover the entire light spectrum; from the ultraviolet into the infrared wavelengths to design efficient devices for a given wavelength.

This work concentrates on probing semiconductor materials for the near and mid infrared wavelength (1.7 to 4.0 μm) for detection purposes. Because of their detection wavelength these detectors are also called infrared photodetectors. Infrared photodetectors are used in numerous fields like remote gas sensing, biological imaging, molecular spectroscopy medical diagnosis, electrical characterization, temperature measurements, astronomy, fault isolation, insulation testing, night vision devices and military target acquisition and detection.

The cutoff wavelength beyond which a photodetector ceases to detect is defined by the material's bandgap. Thus it is possible to detect a broad spectrum

of wavelengths with a narrow bandgap material. However, the noise of a photodetector scales with the dark current which increases exponentially with reduction in bandgap. To reduce the dark current it is critical to design a photodetector with bandgap which corresponds to the desired wavelength of detection.

1.1 Band structure engineering for device applications

Band structures have been an integral part of semiconductor research owing to their capability to tune material parameters like bandgap, effective masses and conduction and valence band edge alignments of the material. Reconfiguration of band structures in a controlled manner with aid of various techniques is termed as band structure engineering. Alloying between compounds offers the ability to tailor the electronic structure, where the resultant change can be controlled by the mole fraction of the compounds in the final alloy. This method allows engineering a broad range of bandgaps between the traditionally known III-V compounds some of which are represented in Figure 1. The complete range of alloys between the two end point compounds shown by the solid lines represents direct bandgap alloys. For the direct bandgap alloys the momentum of electrons in conduction band equals that of the holes in valence band thus permitting optical transitions in these alloys. This is not true in the case of indirect alloys shown by the dashed red curves, where an electron has to transfer its momentum to recombine with a hole.

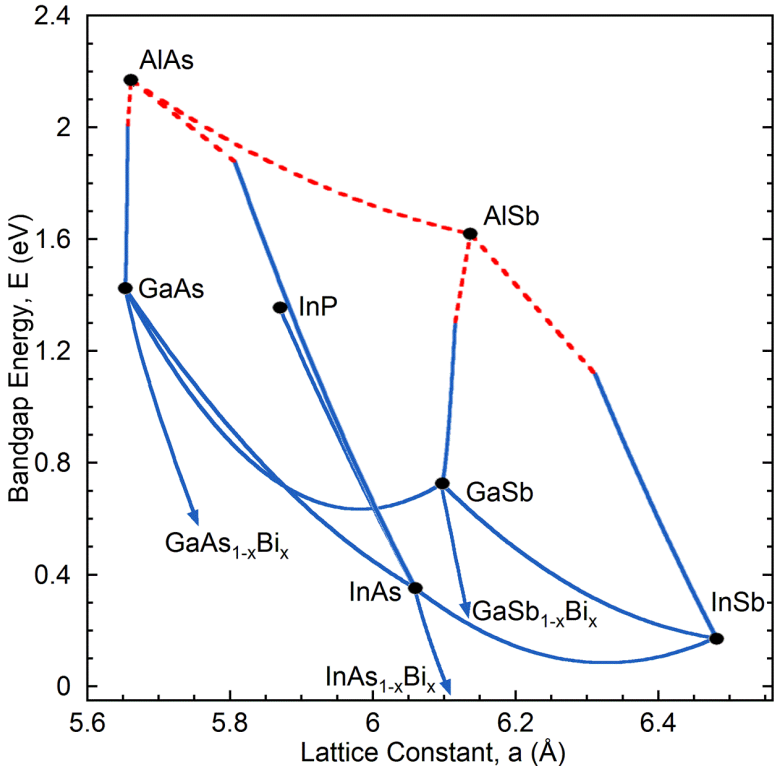


Figure 1: Bandgap versus lattice constant for few conventional compound semiconductors and bismuth based highly mismatched alloy.

Due to miscibility gap between the end compounds it was not possible to engineer all these alloys with simple chemical admixtures. It is now feasible to grow these alloys with the aid of epitaxial growth. With the advent of techniques such as molecular beam epitaxy and metal organic vapor phase epitaxy it is possible to control device growth at the atomic scale permitting the formation of complex structures and alloys.

These alloys are capable of covering large range of wavelengths in the infrared spectrum but are held back due to lack of availability of high quality, inexpensive substrates. The lattice mismatch between the alloy and substrates is proportional to the strain on the final structure. If these strain levels are not limited to the critical values they can deleteriously affect the growth quality

consequently producing crystallographic defects. Crystallographic defect like misfit dislocations manifests themselves as non-radiative transitions thus deteriorating the efficiency of the detector. Hence the use of alloying for bandgap engineering is limited to alloys with lattice constant in the vicinity of these readily available substrates.

The electronic structure of a bulk semiconductor is defined by the periodically varying potential over the crystal lattice, which in turn defines the permissible energy levels for charge carriers. It is possible to design two-dimensional structures where the potential profile changes in the direction of growth, thus limiting the movement of the charge carriers to two directions. An example of such a structure is a quantum well in which a thin well layer is sandwiched between two barrier layers with larger bandgap. The well layer is so thin that the carrier wavefunction cannot be restricted within the well region. This change modifies the confinement energy of the constituent consequently producing energy levels different than that of the bulk. These energy levels define the effective bandgap of the two-dimensional design hence adding another degree of freedom to bandgap engineering. Shift observed in the energy levels are largely dependent on the band edge alignment between the participating alloys. The figure below shows the various alignments available to form the two dimensional structures. These structures have been heavily studied and utilized in all optoelectronic device applications. Furthermore, these structures also aid in improving the optical properties by increasing the confinement of charge carriers in a layer.

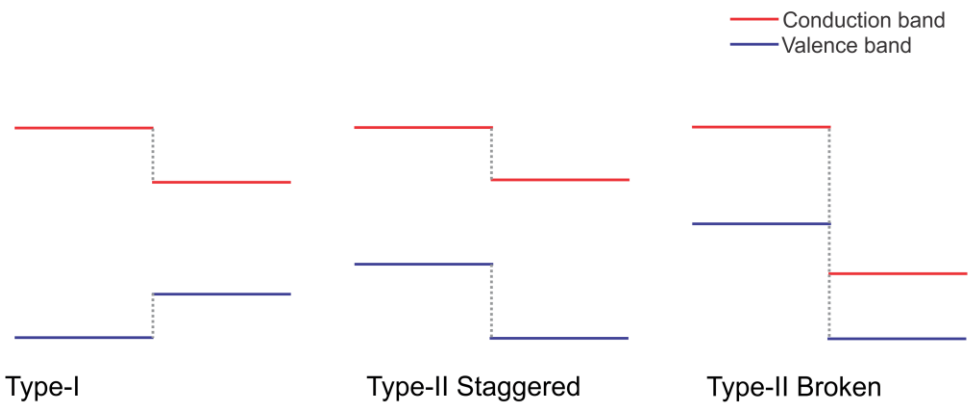


Figure 2: Different band edge alignments possible between alloys

Another, interesting technique to engineer bandgap is to use multiple layers of the constituents and switching between the layers much faster than a quantum well. Due to the proximity of well layers the wavefunction of two adjacent layers interact hence giving rise to minibands. Such a structure where the switch between lattices of different alloys occurs in very short intervals is called a superlattice.

Besides the above mentioned techniques a fairly new method to tailor band structure in line with alloying is the use of highly mismatched alloys. Some of these mismatched alloys with bismuth have been marked in Figure 1. These alloys constitute of elements which differ largely in their chemical and physical properties. These differences bring about drastic changes in the electrical and optical properties of the resultant alloy [1-3]. They not only show anomalously large reduction in bandgap but are also responsible for the direct bandgap nature of GaN_xP_{1-x} [4].

An additional degree freedom in band structure engineering can be realized by adding strain on the layers for the aforementioned structures. Strain

shifts the conduction and valence band edge as well as removes the degeneracy in the valence band thus reconfiguring the band structure of an alloy. It is vital to limit the strain-levels in a bulk alloy below the critical values. But in the two-dimensional structure if we grow strain-balanced structures it is possible to retain the effect of strain within the layers, hence suppressing the chances of misfit dislocation due to the lattice mismatch with the substrate. By strain balance we mean compensating the strain on one layer with an equal and opposite strain on the adjacent layer. In this work we utilize these different degrees of freedom of band structure engineering to design efficient photodetectors for an extended range of infrared wavelengths.

1.2 Infrared photodetectors for medical and industrial applications

In order to define the detection wavelength for a specific application it is vital to take into consideration the atmospheric transmission. This helps to minimize the effect of random elements like gases and water vapor present in the path of detection. The wavelengths where the effect of these obstructions is negligible or minimal are generally referred as atmospheric transmission windows. These windows are crucial for communication and all other detection purposes which have low signal to noise ratio. Gases like carbon dioxide (CO_2) and nitrous oxide (N_2O) can be detected with ease at their respective fundamental absorption lines in the atmospheric spectrum. Figure 3 shows the atmospheric transmission spectrum with some of its major components.

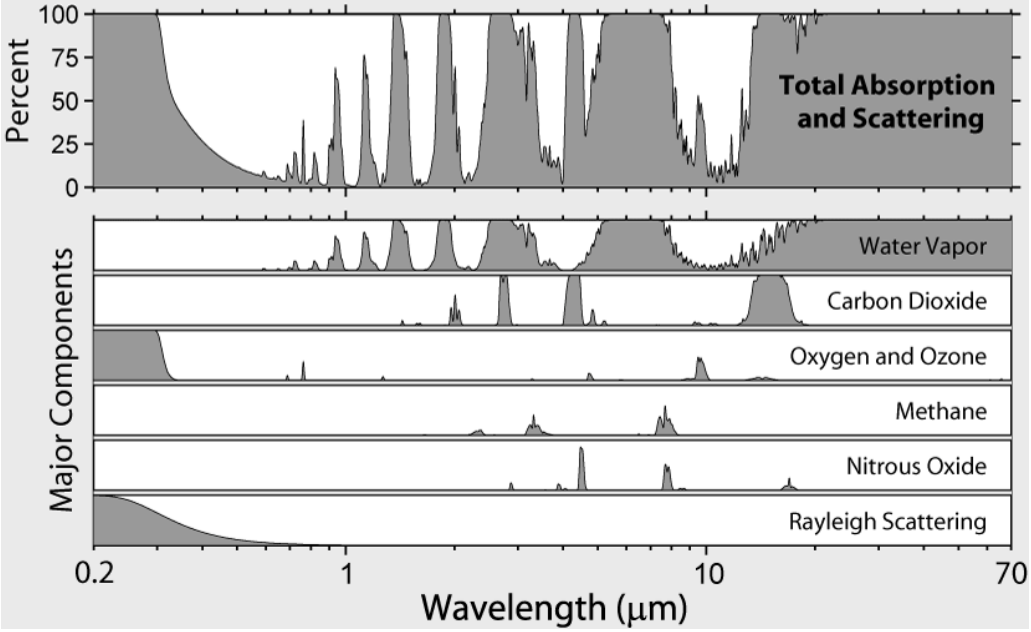


Figure 3: Atmospheric transmission spectrum of radiation. *Courtesy*: Robert A. Rohde / Global Warming Art [5].

From the above figure it is evident that there exist multiple absorption windows in the 1 to 4 μm region which can be utilized for free space communications. In this work we are interested in designing efficient photodetectors for the 1.7 to 4.0 μm wavelengths. The 2 to 3 μm range has been of large interest for multiple device applications. The 2.9 μm wavelength has the highest absorption coefficient in water which is useful for medical applications considering the high water content found in human tissues. Some of them are removal of damaged skin, cavity treatment by the heat radiated from these infrared devices and cancer diagnosis by studying change in cell morphology. Glucose has a strong absorption line at 2.5 μm making this range suitable for live-cell imaging in biological research. All commercial grade pyrometers use the 2.7 μm wavelength for temperature measurement in the industry. Also the presence of emission lines for gases like carbon dioxide, methane and carbon monoxide in

the range makes it very important for innovative gas sensing applications. The detectors in this range can also be utilized for target acquisition purposes in military which is largely done in the 3 to 5 μm spectrum.

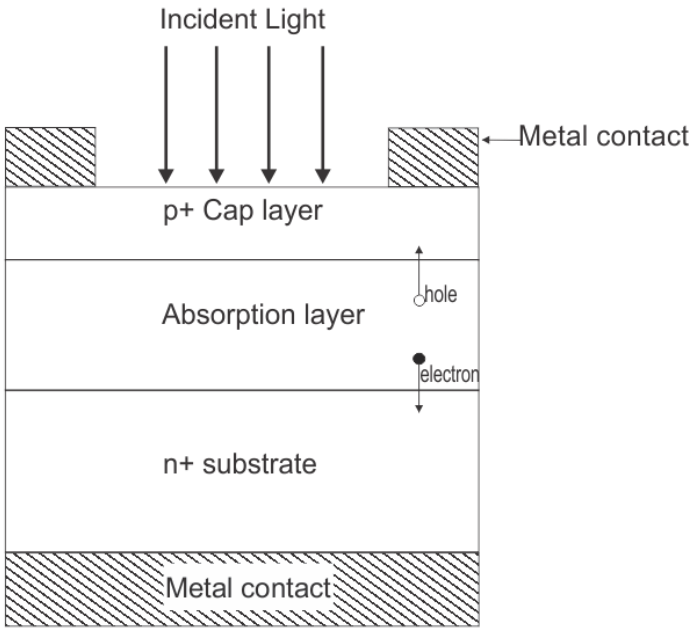


Figure 4: Schematic diagram of a photodetector.

Semiconductor photodetectors work on the principle of photogenerated electron hole pairs where the energy from the absorbed light excites an electron from the valence band to the conduction band. The energy of the light shone to the detector should be equal or greater than the semiconductor's bandgap for the light to be absorbed. The created electron and hole pair are then extracted to generate an equivalent electrical signal. Figure 4 depicts a simple diagram marking the important layers for a semiconductor photodetector. The active region is surrounded by the heavily doped regions which collect the photo excited carriers. The heavy doping on these layers assists in forming an ohmic contact which reduces the resistance at the junction. In general the substrates used are

doped to either n+ or p+ conditions forming heavily doped region below the active region. Bandgap of the material used in the absorption layer defines the working wavelength of the detector. The absorption layer, also the active region can comprise of multiple layers of bulk alloys. Besides it can also utilize the different degrees of freedom discussed earlier for engineering the desired bandgap. The thickness of the absorption layer is proportional to the amount of light absorbed, hence it is possible to improve the efficiency of the device with thicker layers. However these designs can turn out to be expensive and bulky due to the extra material utilized.

Detectors available today in the 1.7 to 3 μm range use bulk $Ga_{1-x}In_xAs$ ($x > 0.53$) for the active region. But as the indium content is increased there is mutual increment in the lattice constant of the alloy. This leads to lattice mismatch between the substrate and the layer large enough to produce misfit dislocations. The low quality of these materials dramatically decreases the nonradiative lifetime and hence increases the undesirable thermal noise in these photodetectors due to Shockley Read Hall generation recombination.

Over the years the type-II band alignment between GaInAs/GaAsSb has been exploited to improve the photodetector performance in this range [6, 7]. This work improves the efficiency over the entire extended wavelength range by utilizing type-II superlattice designs with the GaInAs/GaAsSb system.

Bismuth is added in the GaInAs layer to effectively reduce the bandgap. The $Ga_{1-y}In_yAs_{1-x}Bi_x$ alloy for the entire 1.7 to 4.0 μm spectrum can be grown at the InP lattice thus eradicating any chances of misfit dislocations. Similar results

are achieved with the use of strain-balanced designs for the GaInAs/GaAsSb alloy system. Both compressive and tensile strain are applied on each layer and compensated with an equivalent opposite strain in the other layer. Figure 5 shows both of these methods as they cover the entire spectrum of interest in the vicinity of the InP lattice constant.

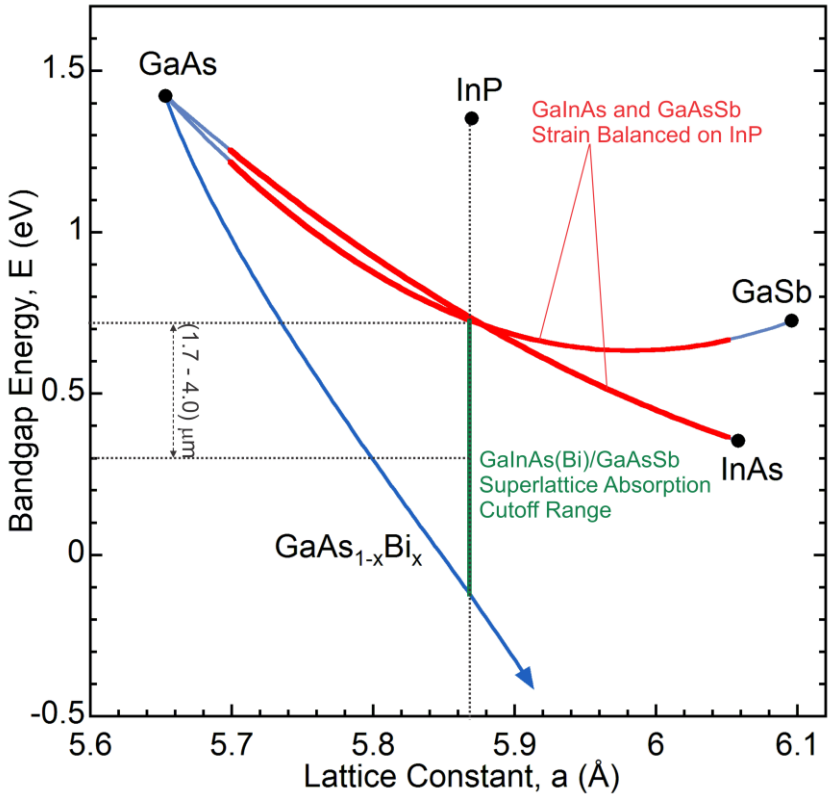


Figure 5: Bandgap versus lattice constant demonstrating strain-balance and bismuth based alloys for the GaInAs-GaAsSb superlattice system.

Red curves atop the alloy curves (in blue) represent the range of alloys on InP which are limited to strains below 3%. The alloys with lattice constants smaller than InP are tensilely strained while as the larger lattice constant will induce a compressive strain on the layer. The two adjacent layers in the GaInAs/GaAsSb system can be strain balanced to realize efficient device structures for the extended wavelengths covered by the system. Alloying GaAsBi

with GaInAs can also aid in covering a broad spectrum of infrared wavelengths at the InP lattice constant. The resultant quaternary GaInAsBi is shown in green which transforms from a semiconductor to a semi-metal as it approaches GaAsBi.

2 FUNDAMENTALS OF HIGHLY MISMATCHED ALLOYS

Alloying has proven to be the most widely used technique to realize semiconductors capable of covering a broad spectrum of wavelengths. Yet there lays a scope of improvement in the quality of these alloys for spectral ranges such as the 'green gap' [8] (520-570 nm). Conventional alloys are fabricated with constituents having similar chemical and physical properties limiting the variations in their final parameters to a very small range. The class of 'highly mismatched alloys' depict a distinct trend in their behavior because of the huge differences in the properties of their constituents.

2.1 Highly mismatched alloy background

In the periodic table N and Bi are at the two extremes of Group V elements and when alloyed with other widely known III-V compounds such as GaAs, GaSb they tend to form highly "mismatched" alloys. The substantial difference between their properties such as atomic size, electronegativity and ionization energy makes it extremely difficult to fabricate single-phase alloys. But with these dissimilarities between the constituent elements the band properties of the final alloy result in large deviations from the host compound. Another notable point is that it takes only dilute quantities of N and Bi to bring about these large variations. Nitrogen alloys have been a topic of interest for shorter wavelengths of the light spectrum and a lot of research was undertaken [9-11] in early 1970's to grow these with other available technologies. Nitrides picked up pace after they were first successfully grown using metal organic chemical vapor deposition [12] due to their popularity for the visible spectrum.

The highly mismatched alloy $GaAs_{1-x}N_x$ has recently been studied for infrared detection applications [13].

On other hand traces it was not until Ma *et al.* [14] who first reported photoluminescence spectra from Bi containing III-V alloys. Before these several attempts were made to grow InSbBi and InAsSbBi as they were expected to be a good substitute for other conventional alloys in the 8 to 12 μm wavelength [15-19]. The proposal of $InAs_{1-x}Bi_x$ [20] alloys for the 3 to 5 μm was followed by succession of growths [21, 22] realizing Bi based alloys for the first time. These growths were testified for good structural quality and composition using techniques like RBS, XRD and PL. The $InAs_{1-x}Bi_x$ and $InSb_{1-x}Bi_x$ alloys can be lattice matched with GaSb and InSb respectively with dilute quantities of bismuth ($x < 6\%$) thus favoring good crystal quality.

In the year 1998 K. Oe and H. Okamoto grew a novel semiconductor $GaAs_{1-x}Bi_x$ [23] using metal organic vapor phase epitaxy with 2% bismuth in the alloy. They depicted the temperature insensitivity of the alloy which is beneficial for fabricating lasers whose lasing wavelength should remain constant for all ambient temperature variations. Such lasers are specifically used in communications hence this property acted as a catalyst for $Ga_{1-y}In_yAs_{1-x}Bi_x$ lasers, which were expected to emit at the 1.55 μm wavelength used in fiber optic telecommunication. Tremendous amount of research has been carried out on $GaAs_{1-x}Bi_x$ properties because of its capability to cover a wide range of wavelengths in the near and short infrared region. Some of the important properties observed included the anomalously large reduction bandgap and large

increase in spin – orbit splitting energy [1, 2, 24]. The huge mismatch in these alloys has always acted as a hindrance to accomplish good quality growth with larger quantities of bismuth. Yet in the recent past there has been success in growing multi quantum wells with high bismuth content (x= 25%) in $GaAs_{1-x}Bi_x$ [3]. In the recent past another highly mismatched alloy of interest has been $GaSb_{1-x}Bi_x$ which is expected to operate in the mid-infrared range and has several potential benefits over other alloys presently deployed in the range [25].

To implement these mismatched alloys and realize photodetectors in the near and mid-infrared wavelength it is necessary to estimate the drastic deviations from their host semiconductor. Thus it is important to understand the underlying basic principles which bring about these changes. In the following sections these principles and some important trends observed in highly mismatched alloys are discussed.

2.2 Localized impurities

The constituents of an alloy can be classified into cation or anion species. Consider the III-V compounds, where the element from group III belongs to the cation species and the latter belongs to the anion species. Atoms from the primary elements comprising the alloy are called host atoms. Any other element besides the host species can be considered to be an impurity. Impurities with an equal number of valence electrons as the host are called isovalent impurities. These impurities can belong to either the cation or anion group. When these isovalent impurities are implanted into the host semiconductor they tend to replace the respective host atom species consequently forming an alloy. Although

it is important to remember that an element behaves as an impurity only when it is in the limit of dilute quantities when compared to the host atoms.

The amount of deviation observed in the properties of a host semiconductor when an impurity is added is defined by the differences in them. The differences are not limited to but include properties such as electronegativity, atomic radius and ionization energy. Therefore to estimate these changes it is important to understand the behavior of an impurity in the host semiconductor. If the properties of an impurity are similar to the host atom it substitutes, then the impurity tends to lie close to the host semiconductor bands and hence it is called a shallow level impurity. The energy level of an impurity in the semiconductor can be calculated using the effective mass approximation [26] method.

Effective mass approximation assumes the perturbing potential of the impurity to be weak and slowly varying in space. Hence the impurity can be treated as an integral part of the host lattice. The impurity electron can now be considered to behave like the host atom's electron in the crystal. The envelope function for the impurity can be calculated from Schrödinger's equation defined by the method where the envelope function encompasses the changes in the host wavefunction due to the impurity as a perturbation. The continuous eigenvalues derived from the Schrödinger's equation are used to describe the localized and delocalized states of the impurity. The states which are bound to the impurity and play no role in the impurity-host interaction are called localized states. On the contrary, ones available for interaction with the host atom are termed as delocalized states.

However some impurities fail to adhere to the aforementioned approximations thus invalidating application of the method. These impurities are called deep level impurities. These impurities tend to have a strong potential and in addition do not lie close to the host bands thus allowing them to bind a state locally. This localized nature of the impurity in the semiconductor makes it impossible to construct the wavefunction within the realm of the Bloch function defining the host's wavefunction. Moreover, the uncertainty of predicting the energy level of these impurities in a semiconductor increases. In order to extract the energy levels of these impurities Hjalmarson *et al.* [27] utilized a modified Koster - Slater model to understand the interaction between the localized and delocalized states of the impurity and semiconductors with a zinc blende crystal structure.

Delocalized or extended states of a host semiconductor construct the wavefunction of the valence electrons hence defining the dispersion relations between energy and wave-vector for the conduction and valence band. The study adjusts the Hamiltonian for a perturbed semiconductor by removing all its matrix elements related to neighbors beyond the first neighbors and yet reproduce the empirically known band structure for the semiconductor. The tight binding analysis within the Koster-Slater method of this Hamiltonian yields the energy level of the substitutional impurity. The theory also assumes that the localized states of the impurity participating in the interaction either have A_1 symmetry or T_2 symmetry where they interact with *s*-orbit or *p*-orbit of the host semiconductor respectively. This interaction with a specific orbit consequently defines the

location of the impurity in the band alignment. This assumption reveals the conformity of the model with the experimental findings, giving weight to the nature of the impurity. These findings approximate the energy level of the impurity to lie further from the bands and thus the name deep level or deep center impurity.

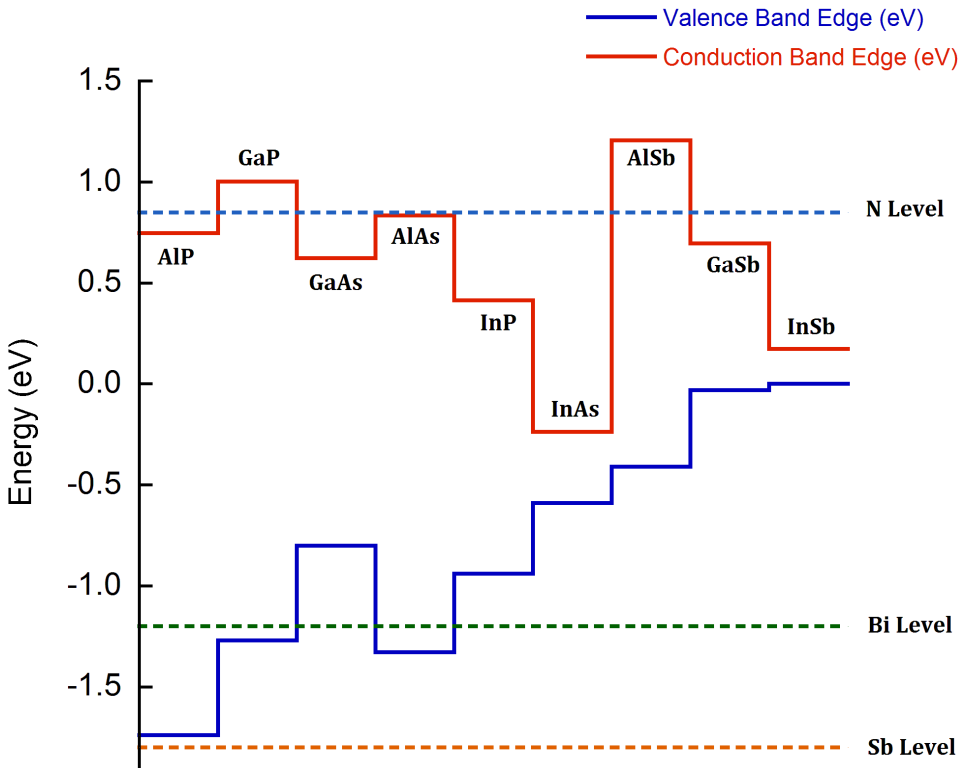


Figure 6: Impurity levels for group V elements aligned with the well known III-V semiconductors. (Courtesy: Chaturvedi Gogineni)

The energy levels for some of the group V elements have been marked and compared to a few III-V compound semiconductors in Figure 6. Some of these results for the nitrogen impurity level in the GaP lattice were verified experimentally [28]. We can therefore infer that impurity with large differences in properties with respect to the host species is likely to behave as a deep level

trap. Due to their nature to bind a state locally we refer to them as localized impurities.

The physical nature of the impurity when compared to the replacing atom defines the location in the host semiconductor. Isoelectronic impurities more electronegative in nature when compared to the anion species of the host atom act as weak acceptors and follow an A_1 symmetry. The A_1 symmetry impurities tend to lie closer to the conduction band of the semiconductor thus interacting with the electrons in the band. Nitrogen serves as a perfect example for III-V semiconductors as it's the most electronegative of the group V and consequently when it replaces any of its complementary anion species it interacts with the conduction band. This conforms with its proximity to the conduction band of most semiconductors shown in Figure 6. While impurities more metallic in nature act as weak donors and hence interact with the holes in valence band. These impurities are predicted to have T_2 symmetry. Heavier or larger atoms in the group like bismuth or antimony fall into this category.

2.3 Characteristic nature of highly mismatched alloys

An alloy inherits its chemical and physical properties from the constituents, governed by their mole fraction. Highly mismatched alloys deviate from this universal law due to the huge mismatch between its constituents and thus exhibit drastic changes in properties with small changes in composition. Large reduction in bandgap, increasing spin-orbit split-off energy, temperature independence, splitting of bands and variations in effective mass of carriers are some of the characteristic properties that highly mismatched alloys demonstrate.

An important point of emphasis is that these variations are a strong function of the impurity in the highly mismatched alloys. As an example, the A_1 symmetry like impurities would affect the electron effective mass as they interact with the conduction band in comparison to the T_2 symmetry impurities which would affect the hole effective mass. It is the localized nature of these impurities that is responsible for the abnormality in the properties of a mismatched alloy. The studies about some of these deviations are crude and require a more rigorous process to be in place to understand them. The properties for "well-matched" alloys can be predicted with ease using virtual crystal approximation [29] but this does not stand true for the "mismatched" alloys. The band anticrossing model has been able explain these abnormalities to an extent.

The bandgap of these alloys has been observed to reduce drastically with increasing mole fraction of impurity. This has been observed experimentally in all III-V semiconductors. To begin with addition of 1% of nitrogen in GaAs lattice reduces the bandgap by 180 meV [30] and 1% bismuth in GaAs reduces the bandgap by 60 meV [31]. The band anticrossing model (BAC) [32] explains this anomalous reduction due to the interaction of the host semiconductor band with the impurity. This anticrossing interaction causes the bands to split thus effectively producing two bands on either side of the original band. The new sub-bands can be marked as E_+ and E_- lying above and below the original band respectively. In the case of conduction band interaction the effective bandgap reduction is justified by the E_- band below the conduction band. Whereas, the E_+ band above the valence band in impurities like bismuth defines the new bandgap

of the alloy. This behavior due to anticrossing is the prime reason for large bandgap reductions. This has been depicted time and again using multiple experimental set-ups like photoluminescence and photoreflectance. We predict similar results in our simulations for the other III-V highly mismatched alloys which are presented in the following chapter.

Another notable property observed in the *p*-like or T_2 symmetry is the rise in spin-orbit splitting. The BAC model predicts a downward movement in the spin-orbit band with growing concentrations of impurity in the alloy. This augments to the upward shift in heavy and light hole bands thus enlarging the spin-orbit split-off energy (Δ_{so}). Increasing the Δ_{so} reduces the auger recombination process hence suppressing non-radiative transitions like conduction to heavy hole band/split-off to heavy hole band and inter valence band absorption which are governed by transition of carriers between the light/heavy hole and spin-orbit band.

These processes have been experimentally proven [33] to be the dominant reason for reduced efficiency in the InP based near wavelength lasers (1.3μm-1.6μm). Auger recombination is expected to increase dramatically with decrease in bandgap or increased temperature. This leads to an added factor of temperature dependence in these devices thus requiring constant cooling for stable performance. With the increased spin-orbit splitting these transitions are suppressed thus improving the efficiency of the device. This has been verified for antimony based alloys which exhibit reduced sensitivity to temperature in the mid and long-infrared range due to their large Δ_{so} [13]. Figure 7 depicts a trend of

steady increment in the spin-orbit splitting as we substitute higher atomic number elements for the anion site in the III-V binary alloys.

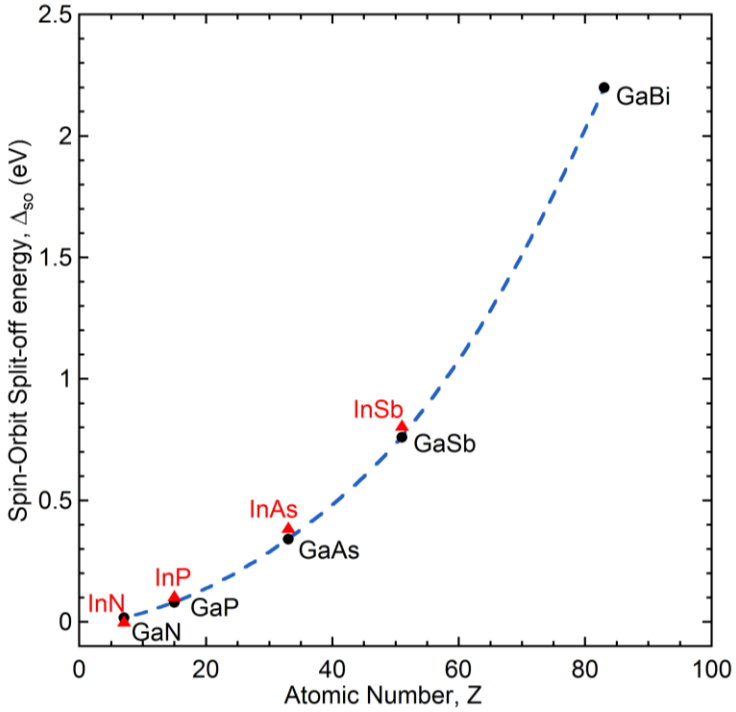


Figure 7: Spin-orbit splitting energy for various III-V compounds depicting dependence on atomic number of the anion species. Ref [31, 34]

These results clearly imply that we can expect a larger Δ_{so} for bismuth based alloys as predicted by BAC, thus opening new possibilities for temperature independent devices in the near and mid infrared wavelengths.

2.4 Bismuth as a constituent in III-V materials

Bismuth has been an element of interest for compound semiconductors and its presence can be found as early as early 1960s [35]. However, its evolution in this field has been hampered by the difficulty to incorporate substantial quantities into the lattice. This crisis arises due to the large atomic size of bismuth which is responsible for its surface segregation while growth. However, researchers have tried to utilize this property as a surfactant [36, 37] for growth of

other mismatched alloys like $GaAs_{1-x}N_x$. Nonetheless, intensified research on growth of III-V bismuth alloys has made generous improvement in growing thin films with substantial amounts of bismuth [3, 38]

Bismuth alloys have extensively been grown and characterized to operate in the near infrared wavelengths (1 to 2 μm). $GaAs_{1-x}Bi_x$ has been a forerunner in this sphere since the last decade, though the initial growths recorded for bismuth alloys were mainly on $InAs_xSb_yBi_{1-x-y}$ or $InAs_{1-x}Bi_x$. Indium based bismuth compounds can be a good substitute for the market dominant materials like mercury telluride (HgCdTe) detectors in the mid and long wavelength infrared region. But due to the iterating issues encountered in the growth of these alloys for over a decade (1984-1998) the researchers began to search for a better mismatched alloy. Though $GaAs_{1-x}Bi_x$ alloys have been successful in achieving wavelengths in the near infrared regime any attempts to go deeper into longer wavelengths have been allied to increased amount of bismuth in the alloy, which has been a detrimental technique to realize these wavelengths.

Due to the distinctive characteristics of highly mismatched alloys like GaAsBi we explore further to prove their huge, yet unrealized, potential. However to understand their properties it is important to be able to model their band structures.

3 BAND ANTICROSSING MODEL AND BAND STRUCTURE

The simplest formulation to predict these parameters for the alloy is to scale the values linearly between the constituent compounds also well known as the Vegard's law [39]. Though Vegard believed that the lattice constant can be best characterized as a linear relation between the constituents parameters governed by their molecular volume there was no mention about the band structure. The virtual crystal approximation [29] follows this law to calculate the band structure by replacing all crystal potentials observed by the free electron in the lattice with potentials from individual constituents subjective to their compositional fraction in the final compound. To account for the small degree of divergence from the virtual crystal approximation or the linear trend a bowing factor was added. In cases where the composition of the constituents is the main reason of potential fluctuations the bowing factor is small and negligible compared to the bandgap energies of the alloy compounds. The virtual crystal approximation along with the addition of a bowing factor is given in Equation 1 to describe the bandgap of the alloy formed between the two constituent materials

$$E_g^{AB}(x) = xE_g^A + (1-x)E_g^B - bx(1-x) \quad (1)$$

where, E_g^A and E_g^B are the bandgap energies of the constituent materials and x is the mole fraction of compound A in the alloy and b represents the bowing parameter that describes the deviation from linear. The bowing factor quantifies the effect of increasing disorder in the formation of 'random alloys' [40]. In this context, 'random alloys' refers to those alloys that lack periodicity in the arrangement of the constituent atoms in the lattice but still have a well-defined

crystalline structure. But this first-order approach fails when the highly mismatched alloys form clusters, where the atoms from the constituents are localized around a lattice site. These differences lead to large deviation from the linear interpolation and cannot be described by a single bowing factor.

The traditionally studied alloys are composed of cation or anion species with similar chemical and physical properties. This is not the case with highly mismatched alloys where the anion species such as arsenic have large differences in properties when compared to its counterpart bismuth. The distinct trend observed in these mismatched elements instigates the need of an improved model predicting the band parameters of these alloys with more accuracy.

3.1 Band Anticrossing Model (BAC)

In an attempt to explain this behavior many theories have been put forth. First of these attempts was when S.A. Barnett tried to predict the bandgap for InBi alloys based on quantum dielectric theory [20]. In his paper he reduced the Equation (1) by proving the bowing factor to be negligible thus making the bandgap of the alloy dependent on the alloy added. In another attempt Wei and Zunger utilized the first-principles local density approximation to explain the large bandgap reduction in $GaAs_{1-x}N_x$ [41]. They were the first to suggest the localized nature of the foreign anion species added in the host semiconductor which altered the band structure of the alloy. There assumption to use a composition dependent bowing to define the degree of localization for these foreign atoms was not sufficient to estimate the bandgap. To overcome these shortcomings the local density approximation method was replaced by an

empirical pseudopotential method with an increased focused on the interaction of the nitrogen-clusters in nitrogen containing alloys [13, 42].

As an alternative to this a BAC model for $Ga_{1-y}In_yAs_{1-x}N_x$ was proposed by Shan *et al.* [32] established on a two level band anticrossing. The BAC model treats the foreign atoms as isoelectronic impurities to the host anion species, as these have the same number of valence electrons. These impurities act as localized states in the host semiconductor. The interaction between these localized trap states and the host semiconductor band is responsible for the reconfiguration of the band structure.

The interaction of the impurity with the host's conduction or valence band is defined by the position of the impurity in the semiconductor. Either interaction yields disparate effects and hence has to be treated separately. Wu *et.al* [43] proposed the BAC model which restructures the conduction band utilizing a two level anticrossing interaction between the delocalized states of the host semiconductor's conduction band and the localized states of the A_1 symmetry impurity.

The valence band anticrossing model [31] is based on the same anticrossing. It utilizes the interaction between the valence band of the host and the localized states of the T_2 symmetry impurities to predict the band structure for the resultant alloy. The model simulates the energy-wave vector (E-k) dispersion curves and can be extended over the entire Brillouin zone to define other important attributes besides the bandgap of the alloy. In the following sections

we have dealt with both of these models and also predict the reconfigured structure for $InAs_{0.96}Bi_{0.04}$ and $GaSb_{0.85}Bi_{0.15}$.

3.2 Conduction Band Anticrossing Model

The BAC model predicts anticrossing between the host semiconductor and the impurity based on the impurity energy level. The trap level of any impurity can be defined by its chemical nature and the symmetry of the substitutional element. In the model used by Hjalmarson *et.al,* they classify the impurity into A_1 symmetry or *s*-like and T_2 symmetry or *p*-like.

The conduction BAC is based on a modified version of the single site model Anderson used to predict the interaction of the s-d orbitals for iron group elements dissolved in non-magnetic elements [44]. A many-impurity Anderson model was established to understand the properties of a semiconductor when it interacts with low concentration of impurities [45-47]. For the conduction BAC model the s-d interaction is replaced with the interaction between localized states of the impurity atom and the extended band states of the host semiconductor. The model predicts the effect of the band anticrossing on the final band structure on solving a three term Hamiltonian consisting of perturbed and unperturbed energies of the electron in the vicinity of the interaction.

$$H = \sum_k E_k^c c_k^+ c_k + \sum_j E_j^d d_j^+ d_j + \frac{1}{\sqrt{N}} \sum_{j,k} \left(e^{ik \cdot j} V_{kj} c_k^+ d_j + h.c \right) \qquad (2)$$

where, the first two terms represent the energy of electrons in the band or extended states at the conduction band E^c and the electrons in the localized states of the impurity level E^d. The third term represents the change in energy due to

the interaction between the two band levels with the interaction strength defined by the parameter V_{kj}. Solutions to this equation can be found using a Green's function approach [4]. Considering that impurities are present in dilute concentrations, the model assumes a random placement of impurity atoms in the semiconductor's lattice. To average the effect of these randomly placed impurity atoms on the electron wavefunction the coherent potential approximation is utilized. The coherent potential approximation describes the theory of how waves scatter in a material with spatial inhomogeneity. This averaging restores the space transitional invariance for the wave-vector, k. On solving the above Hamiltonian (2) using the now transitional invariant Green's function we can reduce the equation into a 2 × 2 matrix for a varying mole fraction of the impurity in the alloy represented by x.[4]

$$\begin{bmatrix} E_k^c - E(k) & V\sqrt{x} \\ V\sqrt{x} & E^d + i\Gamma_d - E(k) \end{bmatrix} = 0 \qquad (3)$$

The parameter Γ is derived from the imaginary part of the Green's solution which is expected to explain the effects on electron mobility and band broadening from the anticrossing. It is defined as $\Gamma_d = \pi\beta V^2 \rho_o(E^d)$ where the terms included are, density of states ρ_o of the conduction band, hybridization parameter V and an empirically determined constant β. Since the many-impurity Anderson model utilized in here is built for low impurity concentration ($0<x<<1$) thus it is vital to use the model only for dilute concentrations of the impurity. When the broadening parameter tends to zero or is negligible Equation (3) can be reduced to [4]

$$E_{\pm}(k) = \frac{1}{2}\left\{[E^C(k) + E^d] \pm \sqrt{[E^C(k) - E^d]^2 + 4V^2 \cdot x}\right\} \quad (4)$$

The mixing of states between the two energy levels predicts splitting of the band into two sub-bands. The lower sub-band (E_+) defines the new bandgap and plays a vital role in modifying other parameters for the band structure compared to the upper sub-band (E_-). The electronic structure over the entire Brillouin zone can be produced by using this conduction BAC model for a range of k values. In the case when broadening parameter is non zero or large we can approximate the solution to

$$E_{\pm}(k) = E_{\pm}(k) + i\Gamma_d \frac{[E_{\pm}(k) - E_k^c]}{[E_{\pm}(k) - E_k^c] + [E_{\pm}(k) - E^d]} \quad (5)$$

Alterations in the electronic structure have been experimentally depicted [43] and successfully explained with the conduction BAC model on $GaN_{0.05}As_{0.95}$. While the term Γ_d was non-zero, broadening was observed in both E_+ and E_-.with prominent effects in the vicinity of the N impurity level. It can also be inferred from the conduction BAC model that the hybridization will affect parameters like electron effective mass, density of states and the electron mobility due to changes in the curvature of the conduction band. Results from experiments [48, 49] conducted on GaN_xAs_{1-x} and GaN_xP_{1-x} have depicted the effect of interaction strength parameter or hybridization parameter V_{kj} is maximum at the conduction band minima hence removing chances of modifications in the any other high symmetry points unless the minima at these zones are close to the impurity level.

3.3 Valence Band Anticrossing Model

The conduction BAC model predicts the anticrossing between the highly electronegative impurities and the conduction band of the host semiconductor. Their electronegative nature defines their A_1 symmetry consequently positioning them in the vicinity of the conduction band. If the impurity species are more metallic in nature with respect to the host anion species the impurity tends to have a *p*-like or T_2 symmetry. The impurity level for most elements with T_2 symmetry has been found to lie near the valence band of the host semiconductors [27]. This proximity suggests the interaction of delocalized states from the valence band with the localized states from the band formed by these impurities.

The interaction of localized states from the T_2-symmetry impurity and the delocalized states of the host semiconductor cannot be explained by the conduction band anticrossing. To manifest the band anticrossing for the impurity level and the valence band interaction a new theory was introduced by K. Alberi *et al.*[31]. Valence band anticrossing successfully predicts the effect of band anticrossing between the host semiconductor's valence band and the energy level formed by impurity atoms. The valence BAC model is established on the hybridization between the *p*-states from the impurity band and the delocalized or extended states host atoms. Due to their metallic nature the impurity atoms have a tendency to behave as donors and thus weakly bind the holes.

Valence BAC model uses a modified **k.p** matrix to simulate the electronic structure for the anticrossing between the bands. The **k.p** model implements a perturbation theory on the Schrödinger's Equation (6) to predict the band

structure around a given value of k. It also incorporates the interaction of the band under consideration with other neighboring bands giving an extensive and accurate picture.

$$\mathrm{H}\Psi(\mathrm{r}) = \left[-\frac{\hbar^2}{2m_o}\nabla^2 + V(r)\right]\Psi(r) = E_{\mathrm{n}}(k)\Psi(r) \tag{6}$$

The Hamiltonian H is defined by the terms Planck's constant $\hbar$, free electron mass m_o, periodic potential experienced by an electron $V(r)$, the E-k dispersion of the n^{th} band $E_{\mathrm{n}}(\mathrm{k})$ and the periodic wavefunction of the electron $\Psi(\mathrm{r})$. The Equation (6) can be reduced using the Bloch's function with a periodic function $u_{\mathrm{nk}}(r)$ to Equation (7). After the perturbation is introduced into the reduced Equation (7), we get [50].

$$\left[\frac{p^2}{2m_o} + \frac{\hbar}{m_o}\mathbf{k}.\mathbf{p} + V(r)\right]u_{\mathrm{nk}}(r) = \left[E_{\mathrm{n}}(k) - \frac{\hbar^2k^2}{2m_o}\right]u_{\mathrm{nk}}(r) \tag{7}$$

$$E_n(k) = E_n(0) + \frac{\hbar^2k^2}{2m_o} + \frac{\hbar}{m_o}\mathbf{k}.\mathbf{p}_{nn'} + \frac{\hbar^2}{m_o^2}\sum_{n'\neq n}\frac{|\mathbf{k}.\mathbf{p}_{\mathrm{nn}'}|^2}{E_n(0) - E_{n'}(0)} \tag{8}$$

The new term introduced represents the interaction of a neighboring band with the band being studied. This equation can be utilized to form matrix equations which are further solved to calculate the bands. The complexity of the solutions is increased by the number of bands included which in turn results in more precise and accurate band structure.

$$H_v = \begin{vmatrix} H & \alpha & \beta & 0 & \frac{i\alpha}{\sqrt{2}} & -i\sqrt{2}\beta & V(x) & 0 & 0 & 0 & 0 & 0 \\ \alpha^* & L & 0 & \beta & \frac{iD}{\sqrt{2}} & i\sqrt{\frac{3}{2}}\alpha & 0 & V(x) & 0 & 0 & 0 & 0 \\ \beta^* & 0 & L & -\alpha & -i\sqrt{\frac{3}{2}}\alpha^* & \frac{iD}{\sqrt{2}} & 0 & 0 & V(x) & 0 & 0 & 0 \\ 0 & \beta^* & -\alpha^* & H & -i\sqrt{2}\beta^* & \frac{-i\alpha^*}{\sqrt{2}} & 0 & 0 & 0 & V(x) & 0 & 0 \\ \frac{-i\alpha^*}{\sqrt{2}} & \frac{-iD}{\sqrt{2}} & i\sqrt{\frac{3}{2}}\alpha & i\sqrt{2}\beta & S & 0 & 0 & 0 & 0 & 0 & V(x) & 0 \\ -i\sqrt{2}\beta^* & -i\sqrt{\frac{3}{2}}\alpha^* & \frac{-iD}{\sqrt{2}} & \frac{i\alpha}{\sqrt{2}} & 0 & S & 0 & 0 & 0 & 0 & 0 & V(x) \\ V(x) & 0 & 0 & 0 & 0 & 0 & E_{imp} & 0 & 0 & 0 & 0 & 0 \\ 0 & V(x) & 0 & 0 & 0 & 0 & 0 & E_{imp} & 0 & 0 & 0 & 0 \\ 0 & 0 & V(x) & 0 & 0 & 0 & 0 & 0 & E_{imp} & 0 & 0 & 0 \\ 0 & 0 & 0 & V(x) & 0 & 0 & 0 & 0 & 0 & E_{imp} & 0 & 0 \\ 0 & 0 & 0 & 0 & V(x) & 0 & 0 & 0 & 0 & 0 & E_{imp-so} & 0 \\ 0 & 0 & 0 & 0 & 0 & V(x) & 0 & 0 & 0 & 0 & 0 & E_{imp-so} \end{vmatrix} \quad (9)$$

Equation (8) is used in the well established 6 × 6 **k.p** matrix to describe the Γ_7 and Γ_8 valence bands of a semiconductor. For the valence BAC model another 6 × 6 matrix is augmented in the original matrix to incorporate the localized states from the impurity. The resulting 12 × 12 matrix adopted from K. Alberi *et al.* [31] is expressed in Equation (9).

Terms from the standard 6 × 6 matrix were appended to incorporate the linear change of the valence band maximum and the large split-off in the spin-off band.

$$H = \frac{\hbar^2}{2m_o}\left[\left(k_x^2 + k_y^2\right)\left(\gamma_1 + \gamma_2\right) + k_z^2\left(\gamma_1 - 2\gamma_2\right)\right] + \Delta E_{VBM}x \tag{10a}$$

$$L = \frac{\hbar^2}{2m_o}\left[\left(k_x^2 + k_y^2\right)\left(\gamma_1 - \gamma_2\right) + k_z^2\left(\gamma_1 + 2\gamma_2\right)\right] + \Delta E_{VBM}x \tag{10b}$$

$$\alpha = \sqrt{3}\frac{\hbar^2}{m_o}\left[k_z\left(k_x - ik_y\right)\gamma_3\right] \tag{10c}$$

$$\beta = \frac{\sqrt{3}}{2}\frac{\hbar^2}{m_o}\left[\left(k_x^2 - k_y^2\right)\gamma_2 - 2ik_xk_y\gamma_3\right] \tag{10d}$$

$$D = L - H \tag{10e}$$

$$S = \frac{1}{2}(L + H) - \Delta_o - \Delta E_{so}x \tag{10f}$$

$$V(x) = C_a\sqrt{x} \tag{10g}$$

The terms ΔE_{VBM} and ΔE_{so} represent the energy difference between the valence band maximum and the spin-orbit split-off energies of the two end point compounds respectively. $\gamma_1, \gamma_2, \gamma_3$ are the Luttinger parameters for the host compound and E_{imp-so} represents the spin-orbit split-off energy level for the impurity. The remaining sparse matrices appended account for the changes in the

electronic structure due to the band anticrossing between the localized and delocalized from the impurity and host respectively.

Since the valence BAC also assumes random placement of the impurity atoms in the host lattice the coherent potential approximation is utilized to restore the spatial transitional invariance of the crystal. The coherent potential approximation yields a modified hybridization parameter V defined by the impurity fraction x and coupling parameter or anticrossing strength C_a, also shown in Equation (10g). The coupling parameter is largely dependent on the mismatch between the impurity and the host species.

The valence BAC model predicts results analogous to the conduction BAC, where each resulting bands in the electronic structure splits into E_+ and E_- sub-bands. Valence BAC anticipates anticrossing between all the three bands in the valence band namely heavy hole, light hole and the spin-orbit split-off band with the band formed by the localized states of the impurity. This interaction results in six doubly degenerate bands after the interaction. It should be emphasized that three distinct anticrossing interactions is seen at values of k near the Brillouin zone center or the Γ point [51]. Another important detail to be considered here is regarding the reference energy level for all parameters in the valence BAC calculations is the valence band maximum of the host semiconductor and not InSb. A noticeable feature in the residual band structure is the large splitting energy between the heavy and light holes and the spin-orbit split-off.

The valence BAC has been extensively used to study the $GaAs_{1-x}Bi_x$ alloy system where the bandgap is expected to reduce from 1.42 eV [34] to -1.45 eV [52] as x varies from 0 to 1. Figure 8 shows a comparison between the bandgap values calculated using the virtual crystal approximation and valence BAC for the alloy. Using the valence BAC model we evaluate some other bismuth based highly mismatched alloys of interest for infrared detectors. Bismuth being the largest element is more metallic in nature than rest of the elements in the group.

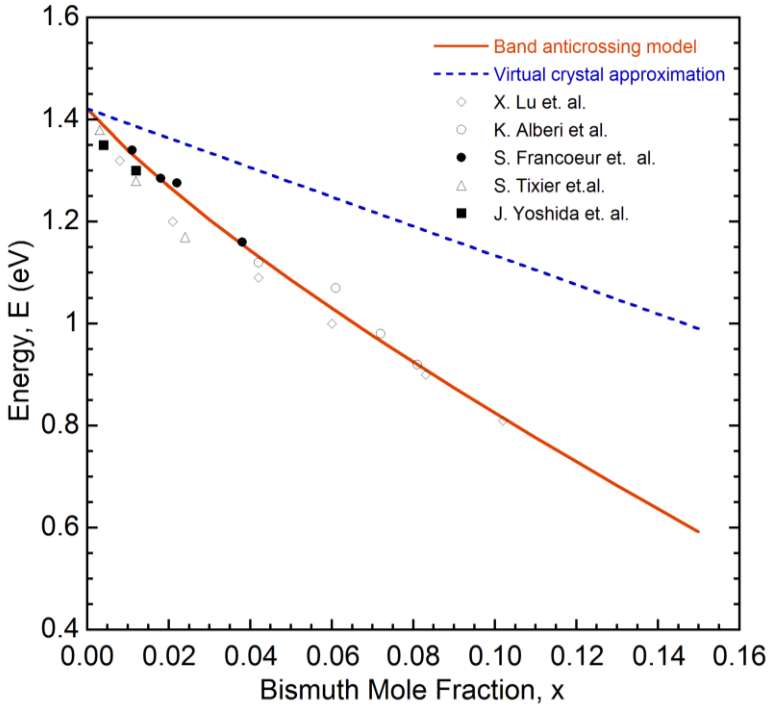


Figure 8: Experimentally determined values compared to the trend predicted by virtual crystal approximation for $GaAs_{1-x}Bi_x$ bandgap energies. Ref. [2, 31, 53-55]

This characteristic behavior of bismuth renders the T_2 or p-like symmetry to the impurity. Thus the impurity level of bismuth is placed close to the valence band for all well-studied semiconductors. The Bi impurity level generally lies below

the valence band for most semiconductors with an exception of GaN, where the impurity level lies within the bandgap.

3.4 Valence band structure for $InBi_xAs_{1-x}$

InAs inherently serves as a vital alloy for the mid wavelength infrared region due to its narrow bandgap. With an intension to cover all wavelengths beyond the InAs emission line, a highly mismatched alloy is simulated using the valence BAC model. It has been predicted that dilute amounts of bismuth bring about large reductions in the bandgap of the resultant alloy. Thus InAs is alloyed with InBi to form $InBi_xAs_{1-x}$ which transforms from a semiconductor into a semi-metal as the content of InBi is increased in the alloy.

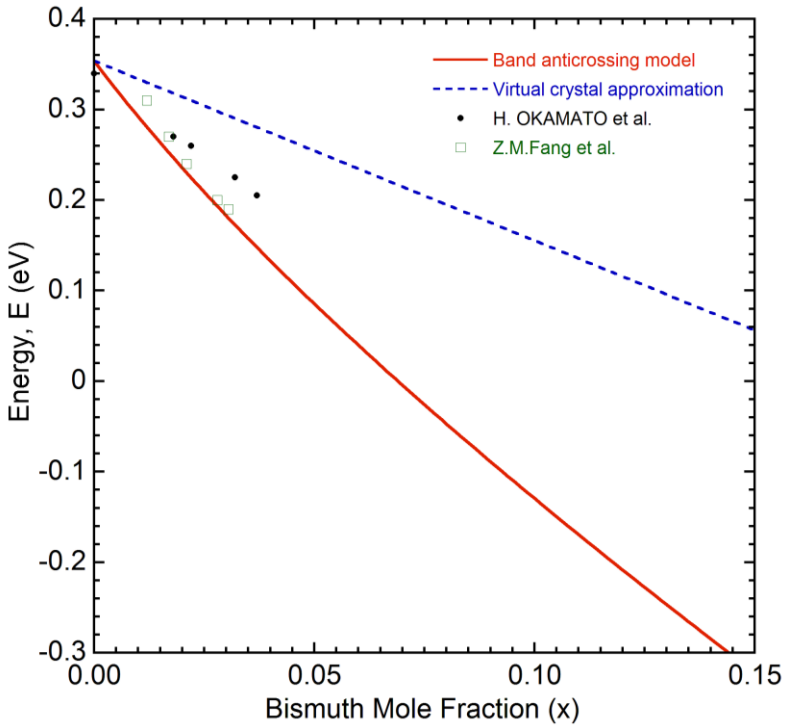


Figure 9: Comparison of $InBi_xAs_{1-x}$ bandgap from band anti crossing and virtual crystal approximation. Ref [24, 54]

Despite the huge miscibility gap in the alloy system successful growth with dilute quantities of bismuth [21, 56-58] has been reported over the last two decade using both metalorganic vapor phase epitaxy and molecular beam epitaxy. Some of the bandgap values extracted from these references are compared with the calculated bandgap from the valence BAC in Figure 9.

The calculations predicted a reduction of 40 meV/% of bismuth in bandgap which is in good agreement with the results acquired experimental [21, 58]. A zero bandgap alloy is expected with 7.2% bismuth in the lattice. The bismuth impurity level E_{Bi} and $E_{Bi\text{-}SO}$ was assumed 0.53 eV and 2.11 eV respectively, below the InAs valence band [31]. The band parameters for InAs and InBi were taken from Vurgaftman *et al.* [34] and Janotti *et al.* [52] respectively. The coupling parameter was approximated from the strength of anticrossing between Bi impurity level and GaAs. Owing to the proximity of GaAs and InAs valence bands the coupling parameter C_{Bi} was kept constant at 1.55 eV [51].

The choice of band parameters used in these calculations is critical for arriving at plausible results. Thus extreme care needs to be placed while selecting the parameters especially with the constituent alloy consisting bismuth in the anion species. To achieve consistent results we focus on using the same set of publications for all values to all possible limits. The conduction band is determined using the virtual crystal approximation considering bismuth impurity level to be distant enough for any hybridization. Since both constituent were direct bandgap alloys the conduction band minima remained at the Γ point. The

dispersion curves shown in Figure 10 are extracted using the valence BAC model depicting the splitting of the heavy hole, light hole and the spin-orbit spilt-off band into sub-bands where the zero energy level represents the valence band edge of InAs. Calculations conform to the increase in spin-orbit splitting energy expected for bismuth based alloys [31, 59]. With only 6% bismuth incorporated in the InAs lattice the hybridization raises the valence band edge beyond the InAs valence band maxima hence reducing the bandgap of InAs from 0.354 eV [21] to 0.051 eV.

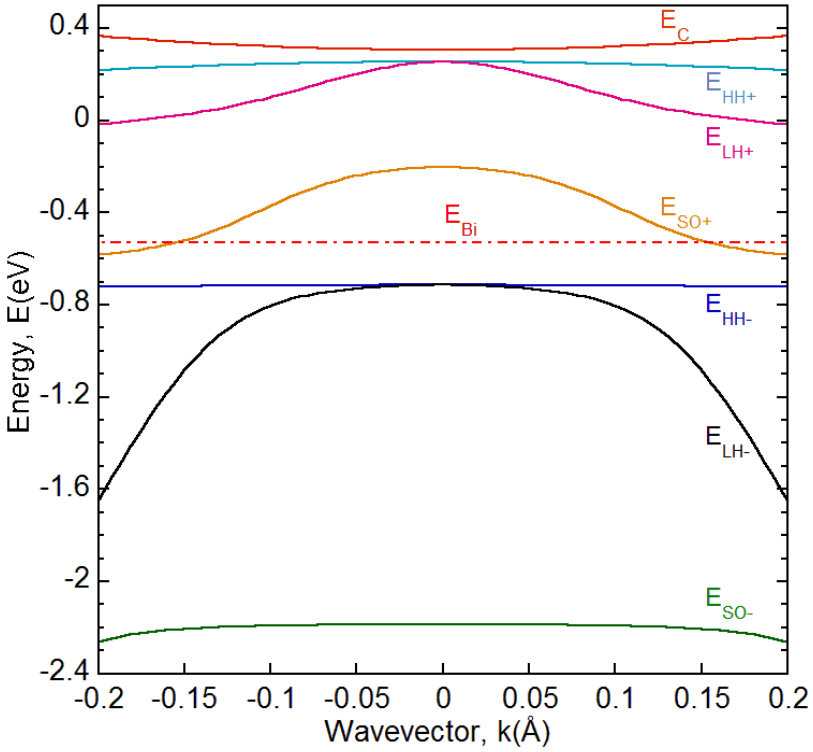


Figure 10: Energy-wave vector (E-k) dispersion curves for $InAs_{0.94}Bi_{0.06}$.

Although we don't show any calculations for the electrical properties of the carriers but we do expect modifications from the reconfiguration of the bands. For example, increment in the hole effective mass due to the increased state broadening with hybridization of the impurity and host bands, augmented with the

increased impurity scattering points a reduction in the hole mobility is evident [51].

3.5 Valence band structure for $GaBi_xSb_{1-x}$

The hybridization of the bands in this alloy system is not expected to be as strong as the InAsBi system. Firstly, the degree of mismatch between the electronegativity of the two group V element is rather modest. Furthermore the bismuth impurity level is 1.17 eV below the GaSb valence band minima compared to the 0.53eV in InAs. Though there has been no application of the alloy in real world devices yet the alloy can be utilized to realize long wavelength infrared detectors on the widely available GaSb substrates. Another important application can be to overcome the problem of small valence band offset in the strained InGaAsSb. The quaternary alloy is utilized for lasers with cut-off wavelength in the 3 to 3.3 μm range and the small valence band offset results in significant hole leakage [60]. This problem can be overcome with the large valence band bowing observed in bismuth based alloys.

Shown in Figure 11 is a plot of the theoretical values for the $GaSb_{1-x}Bi_x$ ($x<0.15$) alloy. For these calculations the valence band edge was determined with the valence BAC model where as the conduction band edge was calculated by linear interpolation between the two end point compounds, GaSb and GaBi. As expected the bandgap bowing is much smaller than InBiAs and hence its proximity to the virtual crystal approximation values. Our calculations predict a 30 meV/% of bismuth reduction in the bandgap compared to an 18 meV /% of bismuth for virtual crystal approximation. This result aligns well with the

findings observed for other alloys with lower degree of mismatch, like $GaAs_{1-x}Sb_x$ [31].

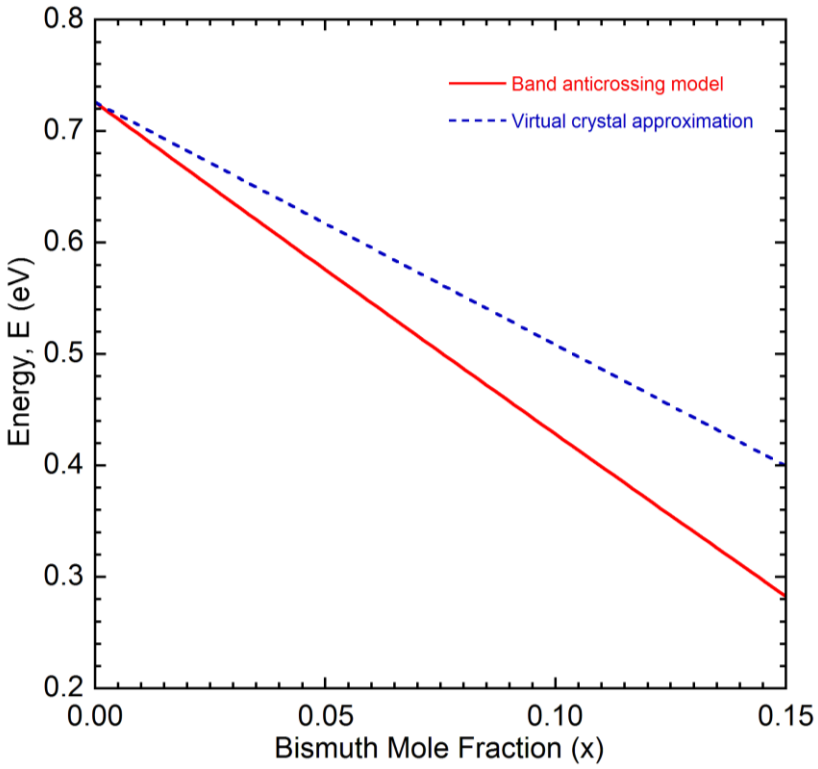


Figure 11: Comparison of bandgap values determined by band anti crossing and virtual crystal approximation for dilute quantities of Bi in $GaBi_xSb_{1-x}$.

The values for impurity level (E_{imp}) and the spin-orbit split-off energy (E_{imp-so}) for bismuth were assumed to be 1.17 eV and 2.67 eV respectively, below the GaSb valence band maximum. The conduction band, valence band and spin-orbit split-off bands for the semi-metal GaBi were predicted to be 0.03 eV, -2.15 eV and -1.11 eV. All of the above were based on the values from the work published by Alberi *et al.* [31]. The coupling parameter was set to 1 eV comparable to the 1.05 eV for GaAsSb [51] bearing in mind the similarity of the anticrossing strength. Though growing GaSbBi has not been an effortless task even amidst the presence of techniques like molecular beam epitaxy [25] but

some success has been achieved in incorporating Bi into the GaSb lattice using liquid phase epitaxy [61].The electronic structure for the alloy at the Γ-point have been depicted in Figure 12.

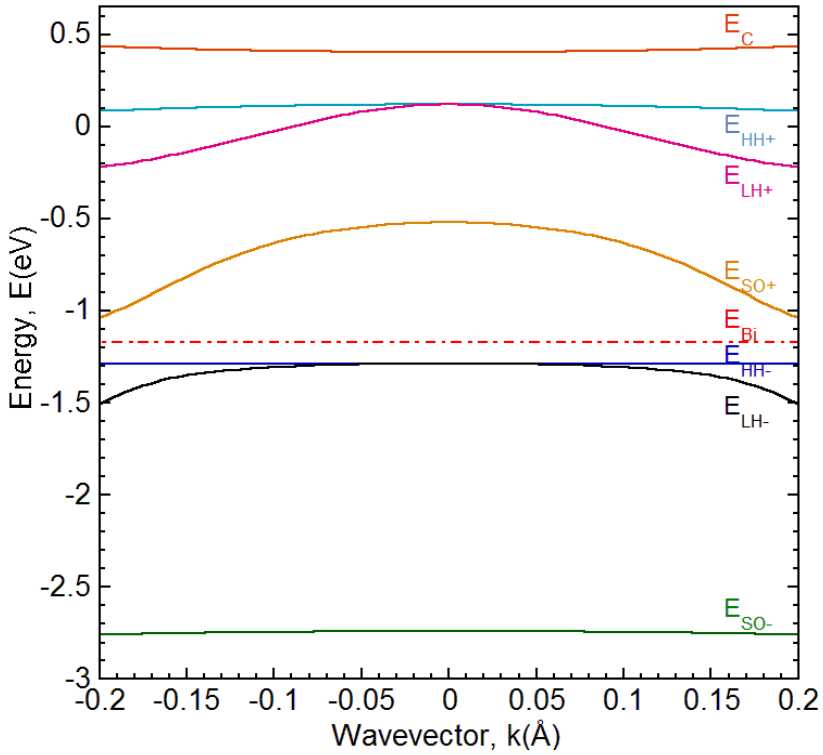


Figure 12: Energy-wave vector (E-k) dispersion curves with the expected splitting of sub-bands for $GaBi_{0.15}Sb_{0.85}$.

Another notable feature when compared to the InAsBi alloy is the spin-orbit splitting energy remains approximately constant in GaSbBi over the dilute bismuth range calculated. Increase in splitting energy aids in suppressing auger recombination processes like CHSH (conduction-heavy hole spin-orbit-heavy hole) (recombination of an excited spin-orbit electron with a vacant electronic state in the heavy hole band) and inter-valence band absorption thus improving the efficiency of the design. Due to the lack of good quality samples there has been no study on the hole effective mass and hole mobility for the alloy but due to

the large energy gap between the GaSb valence band edge and bismuth impurity level we expect the hybridization to have minimal effects.

In the next chapter we investigate applications of bismides to cover the short and mid infrared wavelengths on the extensively used substrates like InP.

4 NOVEL DEVICE STRUCTURES BASED ON HIGHLY MISMATCHED BISMUTH ALLOYS

In 1977, Tsu and Esaki proposed an engineered semiconductor with a one-dimensional periodic potential [62]. This proposal had advocated using advanced thin-film growth technique to engineer the optical and electrical properties of semiconductors. In 1987, Smith and Mailhiot [63] proposed Type-II superlattice materials for IR detection. The superlattice gives the additional degree freedom of controlling the ground state transition energy by varying the thickness of constituent layers.

Semiconductors like mercury telluride (HgCdTe) and InSb are traditionally used to cover a broad spectrum of infrared wavelengths in detection applications. But they are seldom used in the short and mid infrared wavelengths due to their inefficient performances. The inefficiency can be attributed to factors like inferior material quality due to lattice mismatch, thermal losses and temperature dependent output characteristics. Bandgap engineering has proven to be the most efficient method to overcome these drawbacks while achieving the desired bandgap. In this chapter we utilize the type-II alignment between GaInAs/GaAsSb at the InP lattice constant to design superlattice structures whose thickness can be adjusted to cover an extended range of wavelengths. Compared to bulk where the transition energy is represented by the bandgap of the material the transition energy in a type-II is governed by the relative conduction and valence band alignment in the adjacent layers.

The property of bismuth based alloys to reduce the bandgap more effectively than the high indium content in the alloy aids in covering wavelengths longer than 1.7 μm at the InP lattice constant. Different bismuth compositions are compared to achieve optimal design solution for a given cutoff wavelength.

4.1 Kronig-Penney Calculations

The Kronig-Penney model [64] is extensively used to approximate the band structures of semiconductors. It assumes a periodic potential arising from the close packed atoms in the crystalline structure of a solid. This periodic potential behaves as a potential well for the carriers. The model calculates the behavior of an electron in the average potential of the well. The average potential includes the interactions with all ions and other electrons. Any interactions between the electron and hole bands are neglected thus making the model less complex when compared to other band structure calculation models. The Schrödinger equation for these potentials is solved to define the wavefunction of an electron or a hole, describing its wave like property in the crystal.

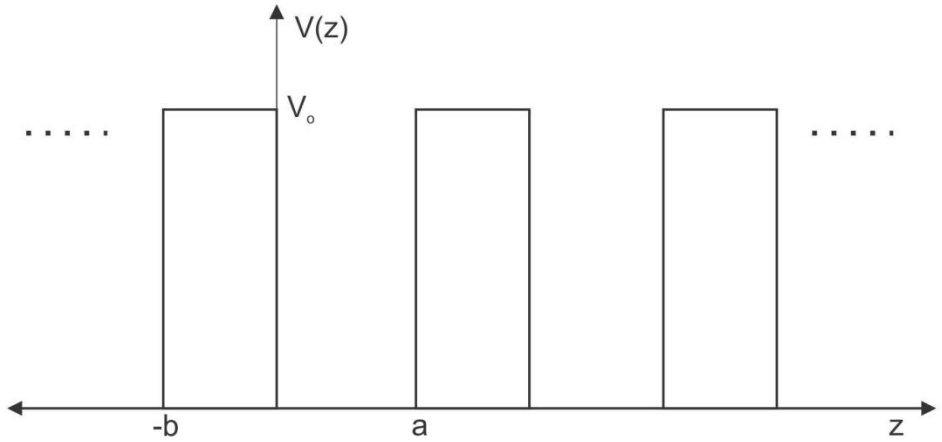


Figure 13: Schematic view of the periodic square potential where 'a' is the well thickness, 'b' is the barrier thickness and 'V_o' is the barrier potential.

The solution for the Schrödinger equation describing the wavefunction for the potential well shown above can be written as:

$$\Psi(z) = \begin{cases} Ae^{ikz} + Be^{-ikz}, & 0 \leq z \leq a \\ Ce^{ik_b(z-b)} + De^{-ik_b(z-b)}, & -b \leq z \leq 0 \end{cases} \quad (11)$$

where, the wave-vectors are defined by

$$k = \frac{2m}{\hbar^2} E \qquad for\ the\ well, \quad (12)$$

$$k_b = \frac{2m_b}{\hbar^2}(E - V_o) \qquad for\ the\ barrier \quad (13)$$

The model stands true for wave-vectors close to zero due to the parabolic approximation assumed in the energy bands. The conventional formalism was set-up for bulk materials where the effective mass was constant throughout the system, which is untrue in the case of superlattice structures. Superlattice structures constitute of alloys having different effective masses and thus a modification of the model is necessary. Though Mukherji and Nag [65] accounted for the change of effective mass between the layer but the model was still inconsistent as they assumed the wavefunction to be continuous at the interfaces of these layers. It was not until Cho and Prucnal [66] that a generalized formalism of the model was proposed for the superlattice structures. The overlap of carrier wavefunctions into the adjacent layers due to their proximity results in the development of a sub-band through the superlattice structure. These sub-bands are generally referred as minibands.

The envelope wavefunction approximation [67] is employed to replace the continuity of wavefunction at the interfaces. Figure 14 shows the transition of a

single wavefunction in a quantum well to the envelope wavefunction in a superlattice structure due to the coupling between them. The gray and yellow rectangles in the superlattice structure represent the bandgap of the two layers.

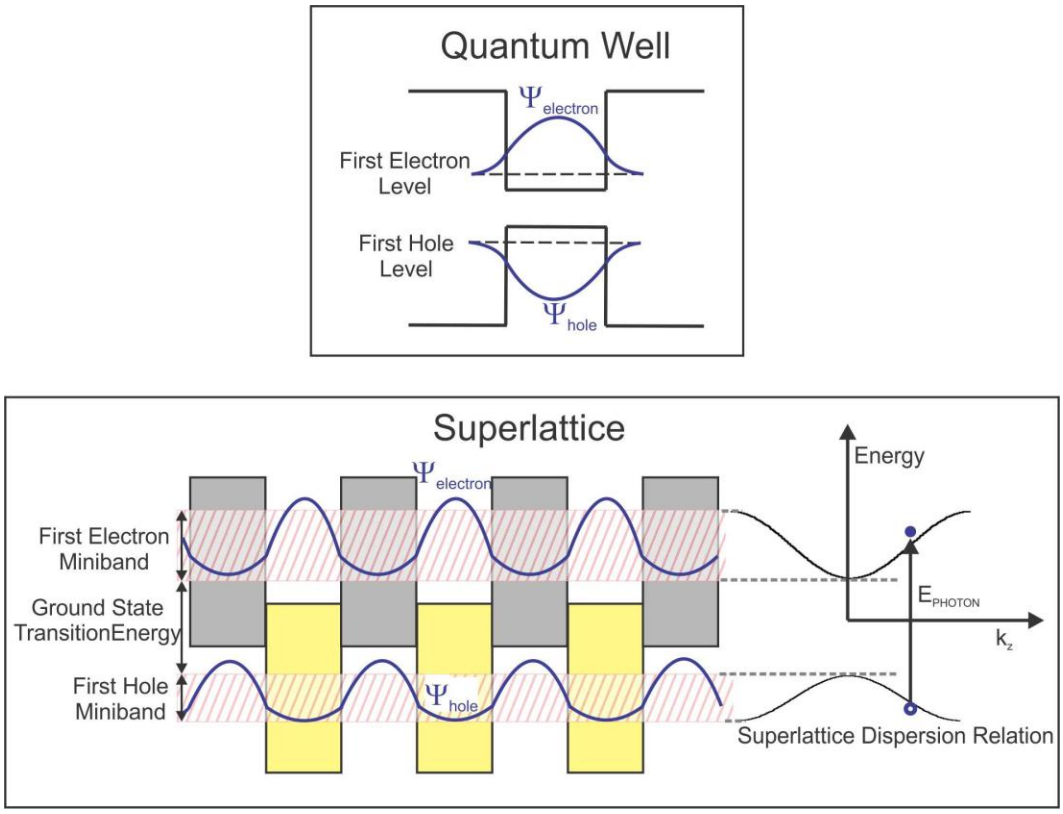


Figure 14: Schematic comparing the wave-functions of a quantum well and superlattice.

Using the modified boundary conditions the energy bands for the superlattice structures can be classified into odd-band and even-band solution each having their maximum and minimum energy level. The energy levels of these bands can then be defined as [66]:

$$\tan\left[\frac{a}{2\hbar}(2mE_{min})^{1/2}\right] - \left[\frac{m}{m_b}\left[\frac{V}{E_{min}} - 1\right]\right]^{1/2} \tan\left[\frac{b}{2\hbar}(2m_b(V - E_{min})^{1/2}\right] = 0 \quad (14a)$$

$$\tan\left[\frac{a}{2\hbar}(2mE_{max})^{1/2}\right]-\left[\frac{m}{m_b}\left[\frac{V}{E_{max}}-1\right]\right]^{1/2}\coth\left[\frac{b}{2\hbar}(2m_b(V-E_{max})^{1/2}\right]$$

$$=0 \tag{14b}$$

$$\cot\left[\frac{a}{2\hbar}(2mE_{min})^{1/2}\right]-\left[\frac{m}{m_b}\left[\frac{V}{E_{min}}-1\right]\right]^{1/2}\tanh\left[\frac{b}{2\hbar}(2m_b(V-E_{min})^{1/2}\right]$$

$$=0 \tag{14c}$$

$$\cot\left[\frac{a}{2\hbar}(2mE_{max})^{1/2}\right]-\left[\frac{m}{m_b}\left[\frac{V}{E_{max}}-1\right]\right]^{1/2}\coth\left[\frac{b}{2\hbar}(2m_b(V-E_{max})^{1/2}\right]$$

$$=0 \tag{14d}$$

The equations 14(a) and 14(b) define the minimum energy and maximum energy level of each band respectively. The smallest solution of 14(a) and 14(b) corresponds to the first miniband (ground miniband) in the superlattice and the second solution corresponds to the third miniband and hence forth all solutions represent the odd-index miniband like 5, 7, etc. Similarly, the Equation 14(c) and 14(d) represent the minimum and maximum energy levels for all even-indexed bands. The proportional relationship between the miniband levels and thickness of the well and barrier adds an extra degree of freedom in bandgap engineering in superlattices. The ground state transition energies can now be controlled with the thickness of the layers.

4.2 Superlattice absorption

The amount of light absorbed by the detector is characterized by its absorption co-efficient. The absorption co-efficient of a detector is defined as the

ratio of number of photons absorbed per unit volume R to the number of photons injected per unit area per second [50].

$$\alpha(\hbar\omega) = \frac{R}{\left(S/\hbar\omega\right)} \tag{15}$$

The number of photons injected per unit area per second can be derived from the photons per unit area per second S and the energy of a photon $\hbar\omega$. The absorption rate of the photons per unit of volume V, can be defined in terms of the initial (E_a) and final (E_b) energies of the state between which the electron travels and the probability of finding a electron in these states is denoted by f_a and f_b [50]

$$R_{a\to b} = \frac{2}{V}\sum_{k_a}\sum_{k_b}|H'_{ba}|^2\delta\left(E_b - E_a - \hbar\omega\right)f_a(1-f_b) \tag{16}$$

The term H'_{ba} is termed as the interband optical matrix element and can be represented using the product of momentum matrix and vector potential of the optical electric field $A(r) = \hat{e}A_o e^{ik.r}/2$.

$$H'_{ba} = -\frac{eA_o}{2m_o}\hat{e}.\,p_{ba} \tag{17}$$

Since the absorption refers to transition of electron from the valence to conduction we can refer to the momentum matrix as p_{cv}. Though the momentum matrix in bulk is defined by the complete envelop wavefunction of the electron and hole but in the case of interband transition it depends only on the periodic parts of the Bloch functions.

$$\Psi_{a,b}(r) = u_{v,c}(r)\frac{e^{ik_{v,c}.r}}{\sqrt{V}} \tag{18}$$

The term $u_{v,c}(r)$ represents the periodic part for the valence or conduction band for the periodic potential *V(r)* of the crystal. Then the interband momentum matrix and optical matrix element can be approximated by the Equation (19) and (20) respectively [50].

$$p_{cv} = \int_{\Omega} u_c^*(r) \frac{\hbar}{i} \nabla \mathrm{u_v(r)} \frac{\mathrm{d^3 r}}{\Omega} \tag{19}$$

$$H'_{ba} = -\frac{\mathrm{eA_o}}{2\mathrm{m_o}} \hat{\mathrm{e}}.\mathrm{p_{cv}} \delta_{\mathrm{k_c},\mathrm{k_v}} \tag{20}$$

The term $\delta_{\mathrm{k_c},\mathrm{k_v}}$ represents the conservation of momentum between the transitions.

The interband momentum matrix for a superlattice includes the complete envelope wavefunction ($\Phi_{\mathrm{n,m}}(\mathrm{z})$) due to potential vector variations in the growth direction, having an initial wave vector $\mathrm{k_t}$. The envelope wavefunctions are considered for the conduction sub-band 'n' and valence sub-band 'm'.

$$\mathrm{p_{cv}} = \langle u_c | p | u_v \rangle \delta_{\mathrm{k_t},\mathrm{k_t'}} \Psi_{\mathrm{m}}^{\mathrm{n}} \tag{21}$$

$$\Psi_{\mathrm{m}}^{\mathrm{n}} = \int_{-\infty}^{\infty} \mathrm{dz}\ \Phi_{\mathrm{n}}^*(\mathrm{z})\ \Phi_{\mathrm{m}}(\mathrm{z}) \tag{22}$$

The Equation (22) represents the wavefunction overlap between the two bands [50]. Using this value of $\mathrm{p_{cv}}$ into Equation (20) and substituting the interband optical matrix Equation (16) R can be written as,

$$R_{a\to b} = \sum_{n,m} |\Psi_{\mathrm{m}}^{\mathrm{n}}|^2 \frac{2}{V} \sum_{k_t} |\hat{\mathrm{e}}.\mathrm{p_{cv}}|^2 \delta\,(E_b - E_a - \hbar\omega) f_a (1 - f_b) \tag{23}$$

The relation of photon absorption rate and the square of wavefunction overlap derived in Equation (23) can be utilized to compare various device designs. The energy conservation function between the initial (E_a) and final (E_b)

carrier energies in the relation restricts the comparison to qualitative analysis only. These energy levels are defined by the density of states. Density of states refers to the number of quantum states per unit of energy and is directly proportional to the mass of the carrier. This implies that for the same wavefunction overlap the strength of a transition involving a light hole will be much lower when compared to the heavy-hole transition. Thus we will classify the transitions involving light hole. This will aid in better qualitative analysis of the design.

In the following sections we utilize the aforementioned principles of quantum theory to find an optimal solution for detectors in the 1.7 to 4 μm range by exploiting the type-II alignment between GaInAs and GaAsSb.

4.3 Engineering the band structure of GaInAs-GaAsSb superlattices using Bi

In the year 1977 Halasz *et.al* [62] reported the type-II alignment for the $Ga_{1-y}In_yAs/GaAs_{1-x}Sb_x$ alloy system. Since then there have been multiple reports on the growth of this material system to realize devices for the short and mid wavelength infrared range [68-71]. Most of these attempts utilized a multiple quantum well structure for the active region of the device. Only few of them exploited the type-II alignment for superlattice design with no attempt to optimize the design. We utilize the type-II superlattice design in an attempt to improve the wavefunction overlap beyond the multiple quantum well structure and also provide an optimized solution for the short and mid infrared wavelengths.

The longest wavelength achievable is restricted to 1.69 µm for $Ga_{0.47}In_{0.53}As$ and 1.73 µm in the case of $GaAs_{0.51}Sb_{0.49}$ by their bulk bandgap. Whereas the type-II alignment between $Ga_{1-y}In_yAs/GaAs_{1-x}Sb_x$ overcomes this restriction to realize wavelengths beyond that of either constituent on InP substrates. The minibands simulated using the Kronig-Penney model for the $Ga_{0.47}In_{0.53}As/GaAs_{0.51}Sb_{0.49}$ type-II alignment, are shown in Figure 15. The layer with 2 nm thickness represents the $Ga_{0.47}In_{0.53}As$ and the adjacent with 8 nm thickness demarcates the $GaAs_{0.51}Sb_{0.49}$ layer. The bands shown in green represent the electron minibands where the second miniband is not completely confined and extends beyond the highest conduction band level of GaAsSb. The heavy holes are marked in blue while as the bands in pink represent the light hole minibands.

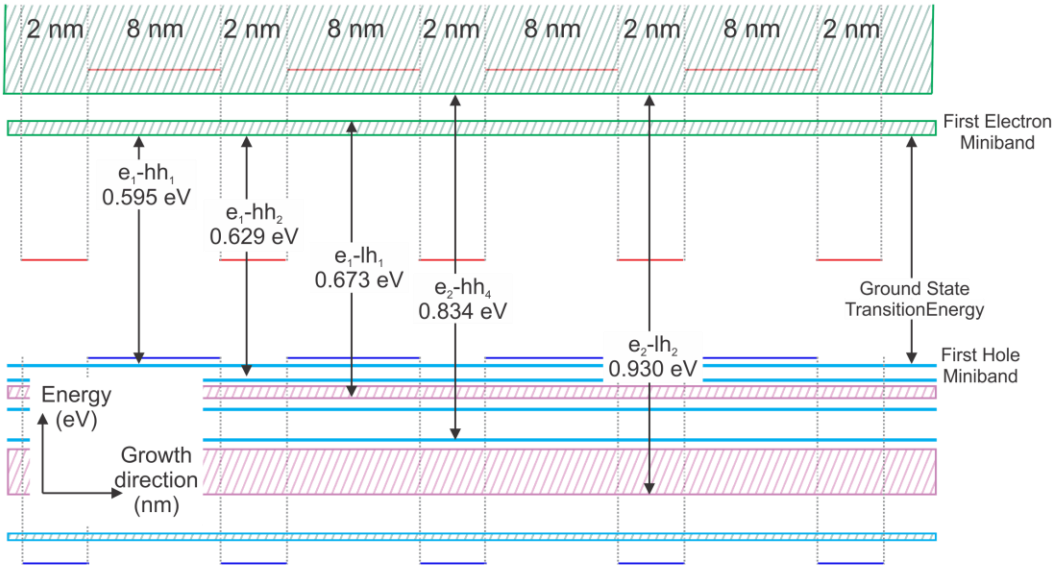


Figure 15: Schematic diagram showing the minibands formed in the $Ga_{0.47}In_{0.53}As$ - $GaAs_{0.51}Sb_{0.49}$ type-II superlattice. Electron minibands are shown in green shaded boxes, heavy holes by light blue and light holes by pink. The black arrows represent some of the possible transitions between the different minibands with each transition marked by the minibands involved.

Due to the difference in masses the light hole level tends to shift faster towards lower energy levels. The figure also depicts some of the several possible transition energies between these minibands. As the electron minibands shift to higher energy the confinement energy is reduced and thus the minibands are broader. Similarly, the hole minibands are thicker at lower energies. The difference in effective mass for light hole and heavy hole explains the larger shift observed in light hole minibands. The heavier effective mass of heavy holes also results in lower transport factor consequently reducing the penetration into adjacent layers. This explains the thinner heavy hole minibands compared to the light hole or electron minibands. This work concentrates on the effective bandgap of the superlattice or the ground state transition energy (0.595 eV) defined by the first electron and hole miniband and hence.

Figure 16 shows a contour plot for the extended range of wavelengths that can be covered by varying the thicknesses of the layers. The dashed black and blue line represents the bandgap for $GaAs_{0.51}Sb_{0.49}$ and $Ga_{0.47}In_{0.53}As$ respectively with the green dashed line marking the smallest possible transition in the superlattice. The family of curves shown in red is extracted by varying the thickness of the $GaAs_{0.51}Sb_{0.49}$ for a constant thickness of the $Ga_{0.47}In_{0.53}As$ layer. The larger bandgap is observed for the superlattice with smallest period due to the enhancement of the two-dimensional confinement energy in the superlattice system.

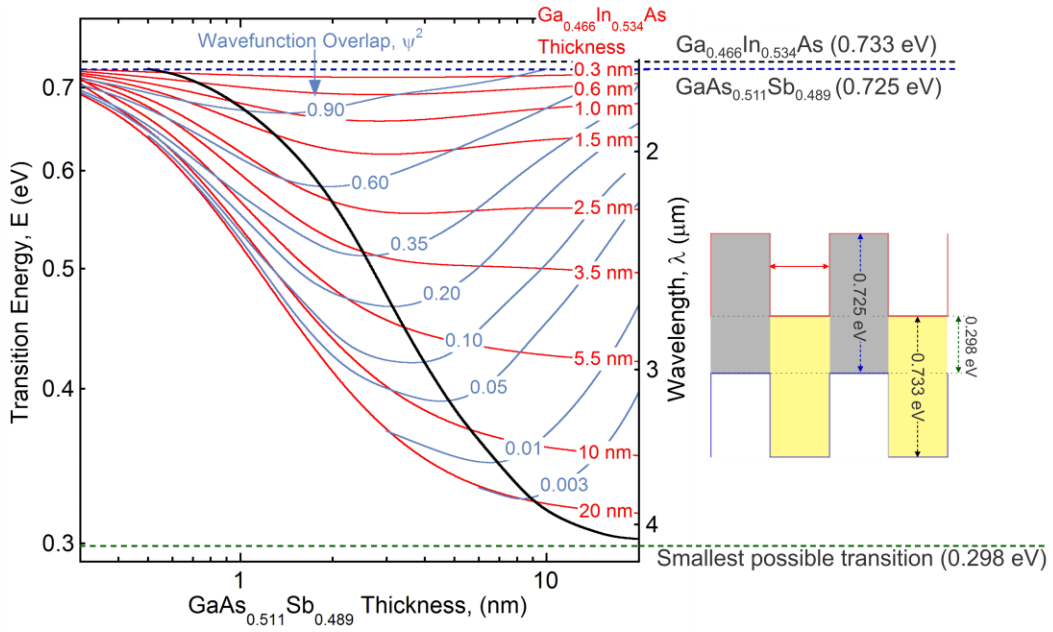


Figure 16: Family of curves for transition energy versus GaAsSb thickness at different $Ga_{0.47}In_{0.53}As$ thicknesses represented by red curves. The solid blue curves denote the constant squared wavefunction overlap where the largest overlap for different wavelengths are connected by the solid black line also called the optimal design. The inset shows the type-II band alignment between the GaInAs/GaAsSb material system with the constituent bandgaps and the smallest possible transition between them.

The minibands are continuous energy levels approximated by replacing the fast changing potentials with an effective potential derived from the two layers in the superlattice. This can be observed in the slope of family of curves for layer thicknesses. The slope approaches zero as one of the layers is thicker in comparison to the other and the slope is higher when they are similar, with exception of reaching the maximum or minimum possible transition energy. The layer thicknesses near the minimum possible transition energy are too thick thus diminishing the effect of superlattices. Thus at these thicknesses the ground state minibands saturate near the band edge of the layers behaving like bulk alloys. The energy level of the minibands also relies more on the thicker layer. This

explains the nearly flat curves at very thin GaInAs and thick GaAsSb layers or near the maximum possible transition.

With an interest to improve the wavefunction overlap for each wavelength it is must to have the wavefunction overlap plotted in the same graph. This is accomplished by plotting the contour maps for the squared wavefunction overlap atop the family of curves, which are represented by the light blue curves. Each curve marks the overlap achieved for a specific transition at different layer thicknesses. In general the wavefunction overlap in a superlattice system is increased by decreasing the superlattice period. This explains the high wavefunction overlap achieved at the very short periods of the superlattice. As the thickness of the layers is increased the layers begin to behave like bulk alloys confining the electrons to GaInAs and holes to GaAsSb layer. This increase in carrier confinement in the adjacent layers reduces the wavefunction overlap between the carriers hence deteriorating the optical performance of the superlattice. The lowest point on each of these curves corresponds to the highest overlap that can be achieved for a specific wavelength. The solid black line connects these minima over the entire spectrum thus demarcating the best achievable squared wavefunction overlap for all wavelengths. Since the line defines the optimum design parameters it is here on referred as the optimal design line.

Though the superlattice system is capable of covering the entire 1.7 to 4.0 μm wavelength but the results obtained with incorporation of bismuth in the $Ga_{1-y}In_yAs$-$GaAs_{1-x}Sb_x$ alloy system look promising. Improvement in wavefunction

overlap is observed with each percent increase of bismuth in the superlattice system. These results and some other major findings from similar simulations are discussed below.

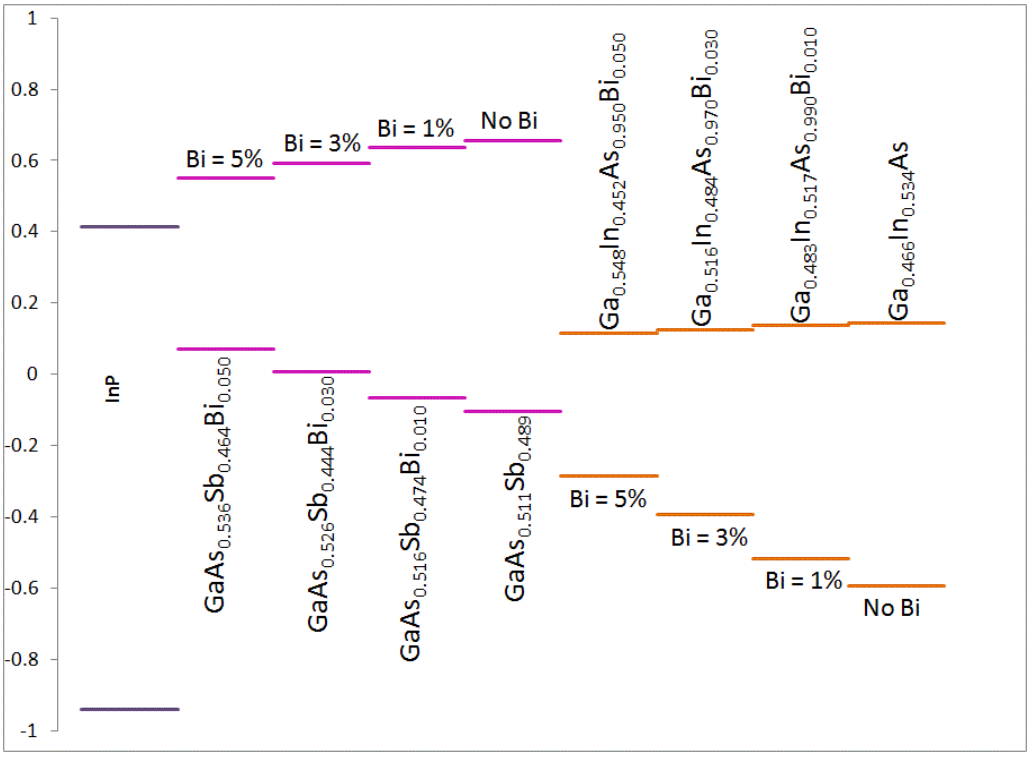


Figure 17: Band edge alignment for GaAsSb(Bi) and GaInAs(Bi) with different bismuth concentrations lattice matched to InP.

The impurity level of bismuth is positioned 0.61 eV below the valence band of $Ga_{0.47}In_{0.53}As$. The highly localized nature of the impurity level leads to subsequent interaction with the host semiconductors valence band. This interaction is simulated using the BAC model which consequently predicts the formation of sub-bands. The conduction band was linearly interpolated between $Ga_{0.47}In_{0.53}As$ and $GaAs_{0.68}Bi_{0.32}$. It is these sub-bands which result in an upward shift of the valence band minima with respect to its original position. The resultant band edge alignment for $Ga_{1-y}In_yAs_{1-x}Bi_x$ reduces the bandgap by 71

meV per % of bismuth. Similar calculations were also performed on $GaAs_{0.51}Sb_{0.49}$ where the impurity level is 1.64 eV below the host's valence band. The simulated energy dispersion curves show a downward shift in the conduction band edge and upward shift in the valence band edge. These shifts result in bandgap reduction by 59 meV per % of bismuth for the novel quaternary $GaAs_{1-x-y}Sb_xBi_y$. Figure 17 summarizes the band edge alignment obtained for the two quaternaries when lattice matched to InP.

Due to reasons such as larger reduction in bandgap and interaction of bismuth mainly with valence band in $Ga_{1-y}In_yAs_{1-x}Bi_x$ compared to $GaAs_{1-x-y}Sb_xBi_y$ we decide to carry further the remainder of the studies using the former alloy. Increase in the lattice constant due to larger bismuth atoms replacing arsenic in $Ga_{1-y}In_yAs$ is compensated by reducing the indium concentration and replacing it with the smaller gallium atoms, preserving the lattice match condition. Whereas when bismuth replaces antimony in $GaAs_{1-x}Sb_x$ it is achieved by increasing the mole fraction of arsenic. Furthermore, the drastic reduction of bandgap with dilute quantities of bismuth in the $Ga_{1-y}In_yAs$ layer also suggests the use of bulk $Ga_{1-y}In_yAs_{1-x}Bi_x$ for the entire 1.7 to 4.0 μm range.

Bulk alloy have an added advantage of ideal wavefunction overlap since all the carriers will be confined to a single layer. Figure 18 shows the bandgap for varying bismuth concentration in $Ga_{1-y}In_yAs_{1-x}Bi_x$. The 4.0 μm wavelength can be realized with 6.8% of Bi in the alloy. However it must be noted that it is extremely difficult to incorporate bismuth into the lattice due to large atomic size and hence a trade-off has to be meet while fabricating the alloy.

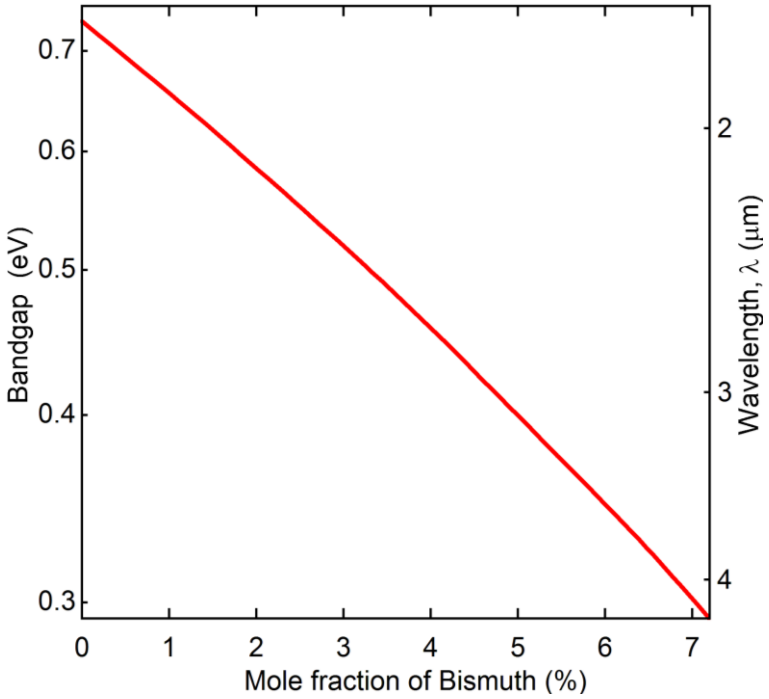


Figure 18: The bandgap of $Ga_{1-y}In_yAs_{1-x}Bi_x$ for increasing bismuth mole fraction in the alloy.

This problem can be resolved by utilizing smaller quantities of bismuth in the GaInAs/GaAsSb superlattice system. Figure 19 shows the alignment of the two alloys and few important transition energies with and without bismuth. The GaAsSb layer enacts as the hole well and GaInAsBi layer as the electron well. The smallest possible superlattice transition is defined by the lowest conduction band level and the highest valence band level in the superlattice system. Other transition energies possible from various energy levels formed by the minibands are marked in the picture as conduction band and valence band ground state range.

Since bismuth works mainly with the valence band of GaInAs thus in the figure we see the valence band shifted up with bismuth but the conduction movement is not discernible. With 5% bismuth in the GaInAs(Bi) layer the

transition energy reduces from 0.733 eV to 0.399 eV and the smallest transition possible between the two layers reduces from 0.298 eV to 0.273 eV.

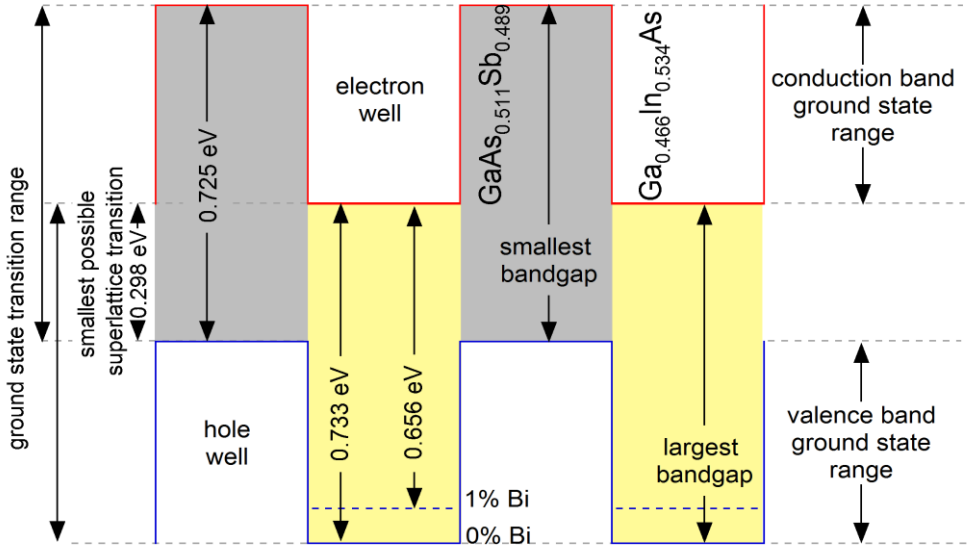


Figure 19: Schematic showing possible transition energies and bandgap in the GaInAs(Bi)-GaAsSb type-II superlattice. The dashed blue line represents the shift observed in the valence band edge for GaInAs with 1% bismuth.

To compare the impact of bismuth in the superlattice system we plot the squared wavefunction contour plots at different bismuth concentrations. Figure 20 shows the contour plot for 1% bismuth in the GaInAs lattice. The reduced bandgap of GaInAs with addition of Bi can be seen from the shift of the dashed black line to a lower energy level. There is also a minute reduction in the smallest possible transition energy from 0.298 eV to 0.297 eV. Square of the wavefunction overlap at 2.7 μm improved to 0.23 with 1% of bismuth compared to 0.19 without bismuth. Another noticeable feature with addition of bismuth into the superlattice was the formation of a new region bound by the bandgap of the two constituent materials where the wavefunction overlap attained values close to 1.

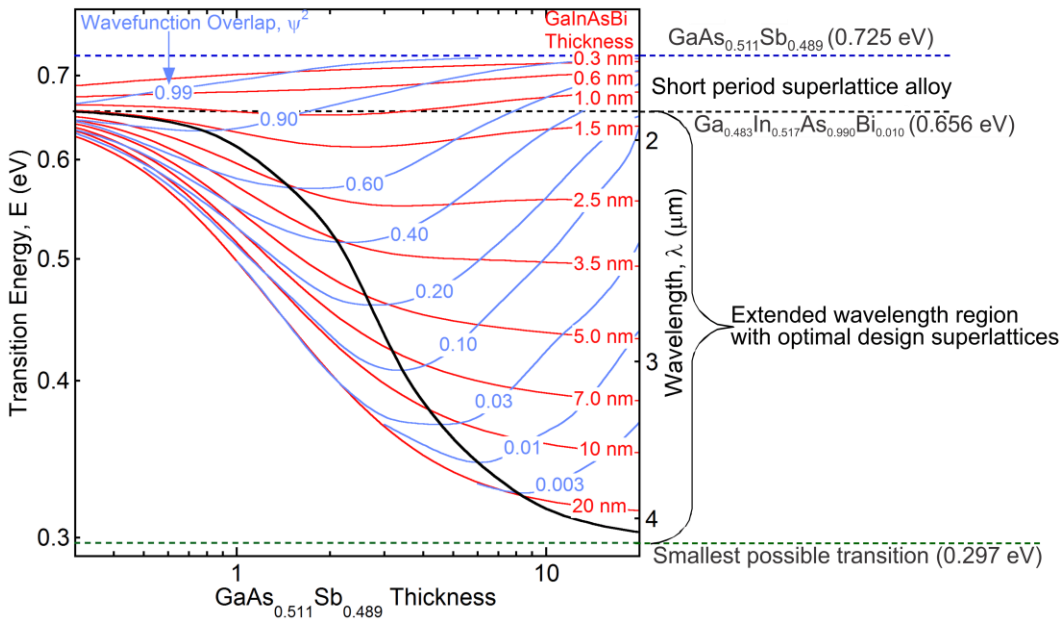


Figure 20: Family of curves for transition energy versus GaAsSb thickness at different $Ga_{0.48}In_{0.52}As_{0.99}Bi_{0.01}$ thicknesses represented by red curves. The solid blue curves denote the constant squared wavefunction overlap where the largest overlap for different wavelengths are connected by the solid black line also called the optimal design line.

The entire range of wavelengths covered by the superlattice system is divided into short period superlattice region and extended wavelength region. The short period superlattice alloy region is seen for the thinner period superlattices. This implies that the superlattice period at very thin layers act as "one" layer hence behaving like a digital alloy. The confinement of the carrier wavefunctions within "one" layer explains the high wavefunction overlap observed in the short period superlattice. For the short period superlattice region the wavefunction improves continuously as the superlattice period is reduced which is not true in the case of extended wavelength region. The maximum wavefunction overlap attained for a given wavelength is defined by a point on the optimal design line.

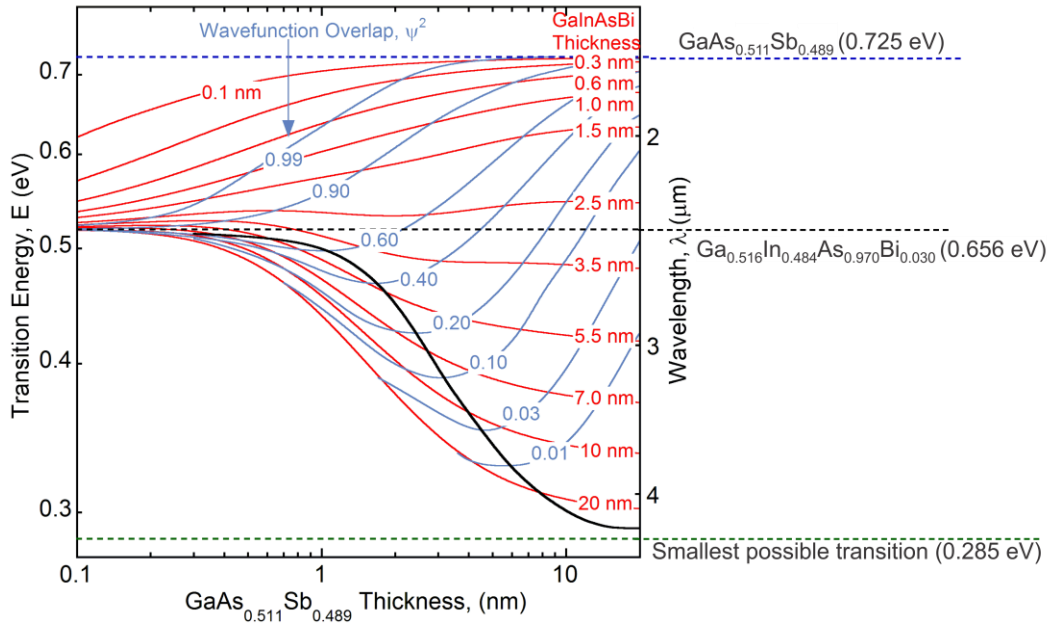


Figure 21: Family of curves for transition energy versus GaAsSb thickness at different $Ga_{0.52}In_{0.48}As_{0.97}Bi_{0.03}$ thicknesses represented by red curves. The solid blue curves denote the constant squared wavefunction overlap where the largest overlap for different wavelengths are connected by the solid black line also called the optimal design line.

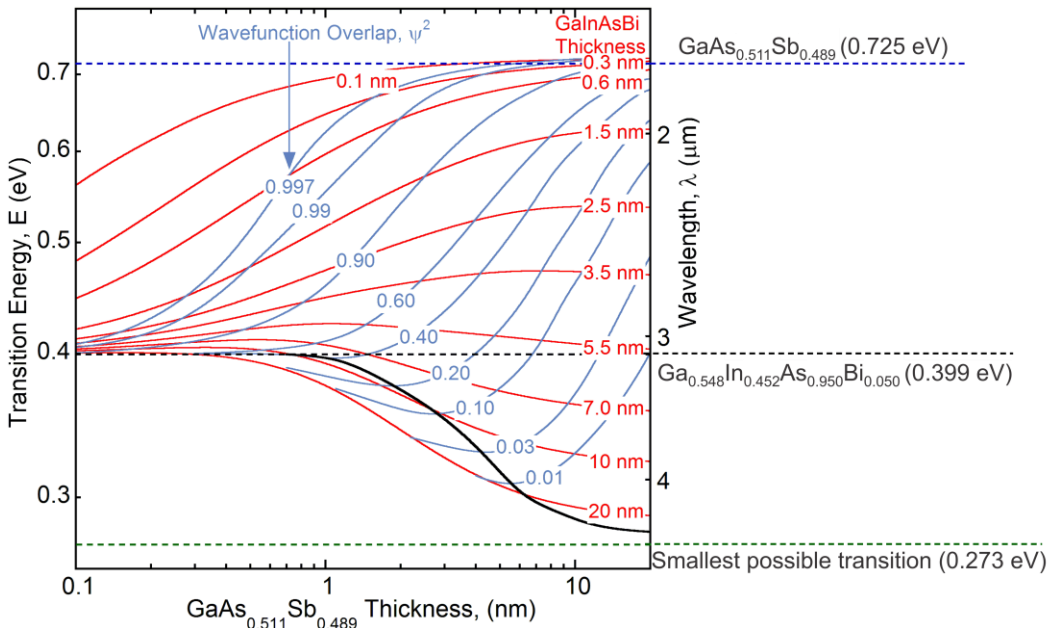


Figure 22: Family of curves for transition energy versus GaAsSb thickness at different $Ga_{0.55}In_{0.45}As_{0.95}Bi_{0.05}$ thicknesses represented by red curves. The solid blue curves denote the constant squared wavefunction overlap where the largest overlap for different wavelengths are connected by the solid black line also called the optimal design line.

On further increment of the bismuth mole fraction in the alloy a similar trend of changes are observed. The bandgap of the GaInAsBi further reduced and the area the short superlattice period increased as seen from Figure 221, 22. The maximum wavefunction overlap for the 2.7 μm wavelength increased to 0.38 and 1 with 3% and 5% bismuth respectively.

The optimal designs from the different bismuth mole fractions are plotted together in order to define the best design. Figure 23 plots the optimal design line for transition energy as a function of varying GaAsSb layer thickness. This is repeated for all concentrations of Bi in the GaInAs layer. Since for the transition energies larger than the minimum bandgap we observe the short period superlattice region which do not have an optimal design the curves are restricted by the bandgap of the GaInAs(Bi) layer. Similarly, the GaInAs(Bi) layer thickness for different wavelengths on the optimal design line are plotted against the corresponding GaAsSb layer thickness in Figure 24.

On comparing the optimal design for different bismuth mole fractions in the superlattice system a set of parametric equations are deduced with an aim to define the best superlattice design at each wavelength. Since the curve fit on each optimal design line is utilized to determine the system of equations, they are termed as optimal design rules. Eventually, the resultant parameters obtained from these optimal design rules are bound to give the best wavefunction overlap for the selected cut-off wavelength. To derive the best superlattice period for a cut-off wavelength a two step method has to be followed. First with defined transition energy (E) or wavelength of interest Equation (24) can be utilized to.

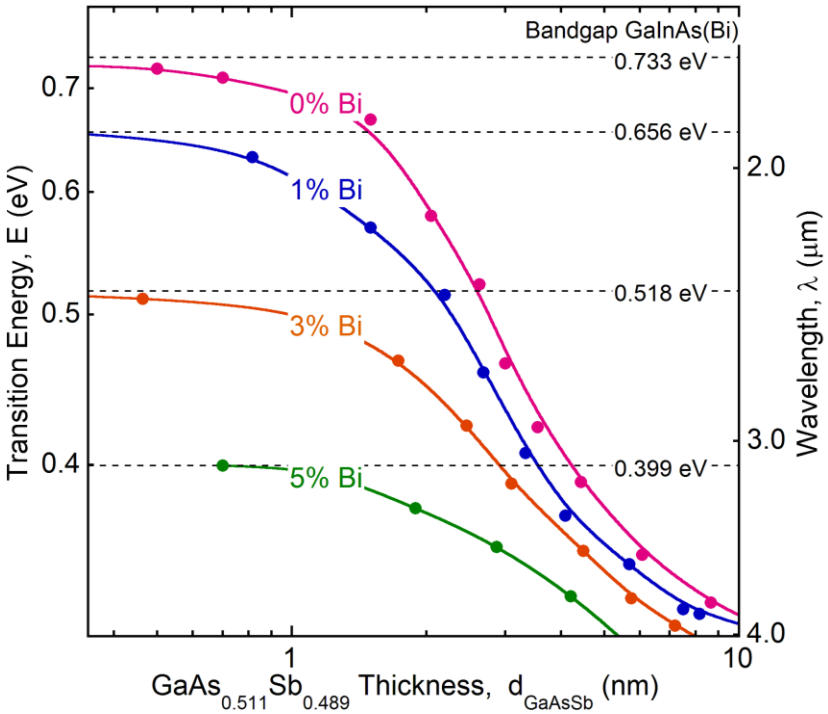


Figure 23: Comparison of optimal designs for the GaInAs(Bi)/GaAsSb superlattice system at different GaAsSb layer thicknesses over the 1.7 to 4.0 um range for different bismuth concentraions in the GaInAs layer.

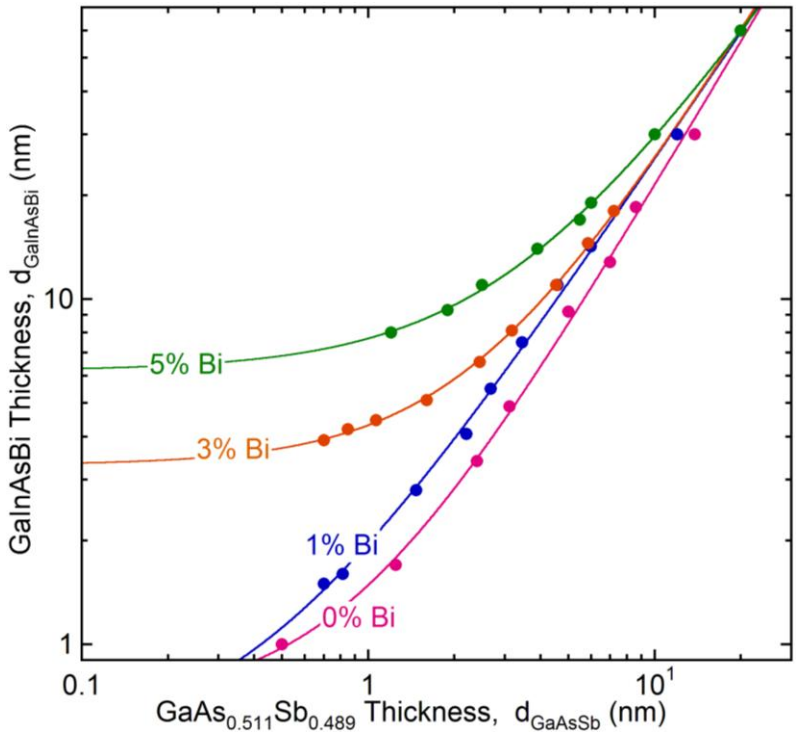


Figure 24: GaInAs(Bi) versus GaAsSb optimal thickness curves for different bismuth concentrations in the GaInAs(Bi)/GaAsSb superlattice system.

deduce the thickness of the GaAsSb layer. In the second step this thickness is substituted in Equation (25) to assess the best thickness for the GaInAsBi layer

$$E = E_{\mathrm{Bi1}} + E_{\mathrm{Bi2}}\left[exp\left(\frac{-d_A}{d_{\mathrm{Bi1}}}\right)\right] \tag{24}$$

$$d_B = d_{\mathrm{Bi2}}\left(1 + \left(\frac{d_A}{d_{\mathrm{Bi3}}}\right)^{\mathrm{p_{Bi}}}\right) \tag{25}$$

where, d_A and d_{B} represent the thickness of GaAsSb and GaInAsBi respectively.

Table 1: Constants defined in the optimal design rules for different dilute bismuth concentrations in GaInAs.

Transition Energy versus thickness				**GaInAs(Bi) versus GaAsSb thickness**		
Bi (%)	E_{Bi1} (eV)	E_{Bi2} (eV)	d_{Bi1} (nm)	d_{Bi2} (nm)	d_{Bi3} (nm)	p_{Bi}
0	0.3013	0.1783	2.902	0.8718	0.6707	1.4078
1	0.3007	0.2928	2.525	0.4092	0.4903	1.2310
3	0.2886	0.4582	2.999	2.4121	3.3063	1.3483
5	0.2803	0.4929	2.937	3.2855	6.1989	1.1962

Table 1 presents the constants for the Equation (24) and (25) for each of the above presented bismuth percentages. It can be inferred from the table that E_0 is closely related to the minimum bandgap achievable.

Although higher percentages of bismuth in the superlattice system is proportional to higher wavefunction overlap but it is difficult to predict which solution of bismuth will be the most efficient considering the complexity involved in adding bismuth in the alloy. The graph in Figure 25 shows a comparison for the best squared wavefunction overlap achieved for the wavelengths covered by the

superlattice with the different bismuth concentration in the GaInAs layer. The dashed black line represents the bandgap of the GaInAs(Bi) layer which in turn defines a cut-off wavelength for the design. The cut-off wavelength shifts to longer wavelengths as the fraction of bismuth is increased. The wavelengths longer than the cut-off wavelength fall in the extended wavelength region and are defined by optimal design line. The squared wavefunction overlap approaches unity as we move closer to the cut-off wavelength for the respective design. All wavelengths shorter than the cut-off wavelength fall under the short period superlattice region where the ideal wavefunction overlap can be realized with extremely thin layers.

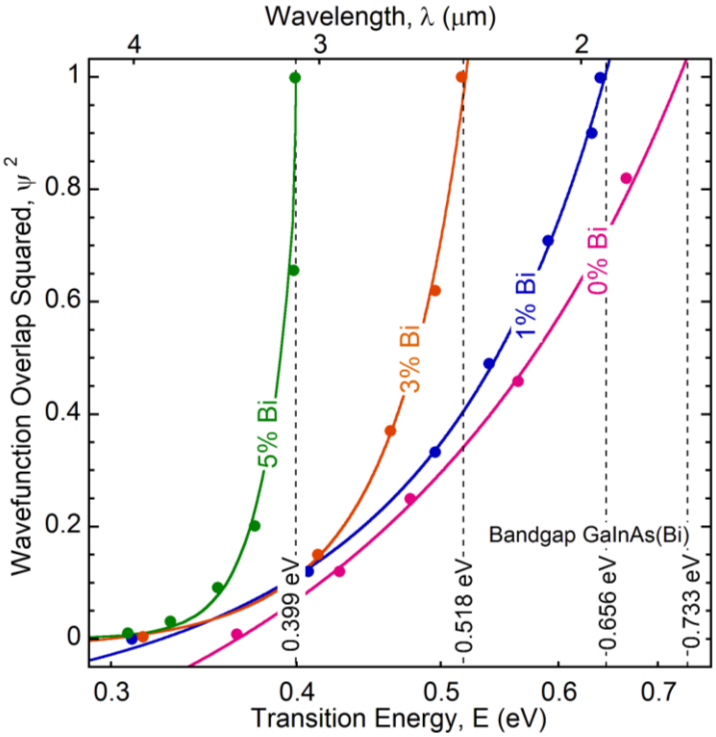


Figure 25: Best squared wavefunction overlap achievable for the MWIR spectrum at different bismuth concentrations.

Equation (26) best describes the curves obtained in Figure 25 at different bismuth concentrations, where the constants defined in the equation are listed in Table 2.

$$\Psi^2 = A_{\mathrm{Bi2}}\left(\exp\left(\frac{\mathrm{E}}{E_{\mathrm{Bi3}}}\right)\right) - A_{\mathrm{Bi1}} \qquad (26)$$

Table 2: Constants defined in the design equation for wavefunction overlap at different dilute bismuth concentrations in GaInAs.

Bi (%)	**A_{Bi1}**	**A_{Bi2}**	**E_{B3} (eV)**
0	0.9957	0.4611	0.5008
1	0.2346	0.0436	0.1928
3	0.0310	$1.916*10^{-4}$	0.0605
5	$1.243*10^{-5}$	$5.128*10^{-9}$	0.0214

The wavefunction overlap doubles as the bismuth increases from 0% to 3% in the superlattice. Thus from our calculations it is evident that dilute quantities of bismuth can form lattice matched alloys with an improved wavefunction overlap. Although the important point to remember is the amount of bismuth that can be incorporated into these structures. For bismuth concentration above 6.8% the complete 1.7 to 4 μm can be achieved with the bulk GaInAsBi. For any value below this concentration the optimal design rules derived need to be utilized to attain the best wavefunction overlap with the superlattice system. In the following chapter we apply strain on the

GaInAs/GaAsSb system to improve the efficiency of the system without having to add Bi in the layers.

5 STRAIN-BALANCED SUPERLATTICE DESIGN

Large strains due to lattice mismatch between the superlattice layers can adversely affect the quality of growth. The enormous force applied by these strains on the layers can result in misfit dislocations when the layer tends to relax to its original lattice. If the strain in a layer is counterbalanced by an opposite strain in the adjacent layer these forces acting at the interfaces can be nullified. Such a condition is defined as a strain-balanced condition. This is feasible in the GaInAs/GaAsSb superlattice system if the strain on the GaInAs layer can be compensated by an equal and opposite strain on the GaAsSb layer. Alternating this pattern in the design can assist in the growth of thicker superlattice stacks with negligible chance of relaxation on either ends of the stack. The difference in strain levels of the adjacent layer can also be compensated by varying the thickness of the layer.

Presence of strain in the layers also modifies the band edge alignment of the system. Tensile strain shifts the conduction band down towards lower energies and the degenerate valence band splits, with the light hole shifting up and the heavy hole shifting down. On the other hand compressive strain reverses the direction of these shifts. The conduction band and heavy hole shift upwards to higher energies while as the light hole shifts down. The resultant band edge alignment from the strain-balance condition reconstructs the superlattice minibands hence modulating the ground state transition energy. This chapter details the effect of various strain levels on parameters like band offsets and

wavefunction overlap in the GaInAs-GaAsSb superlattice on InP substrates for the 1.7 to 4.0 μm spectrum.

5.1 Coherently strained superlattices

The thickness of a layer below which the mismatch between the substrate and layer lattice constant can be accommodated by strain is called critical thickness. When the layer thickness exceeds this critical dimension the force exerted by the mismatch is large enough to relax the layer lattice constant to its original value consequently producing misfit dislocations in the strained layer. These misfit dislocations act as non radiative recombination centers which adversely affect the optical efficiency of a device.

Matthews and Blakeslee [72] developed a model to estimate the critical thickness of a strained layer. They utilized the force exerted by the misfit strain and tension in the dislocation line to explain the change of symmetry at interfaces between the layers. The thickness at which the interface loses its coherency due to strain can be expressed by [72].

$$h_c = \frac{a_o\left(1 - \frac{\nu}{4}\right)\left[\ln\left(\frac{\sqrt{2}h_c}{a_0}\right) + 1\right]}{\sqrt{2}|f|(1+\nu)} \tag{27a}$$

$$f = \frac{(a_o - a_l)}{a_l} \tag{27b}$$

$$\nu = \frac{C_{12}}{C_{11} + C_{12}} \tag{27c}$$

where a_o and a_l denote the substrate and layer lattice constant respectively, f is the layer strain, C_{11}, C_{12} are the elastic constants of the layer and h_c is critical

thickness of the layer. With the layer limited to the critical thickness it is feasible to grow a strained layer and modulate the band edge alignment of the alloy without misfit dislocations.

Strain reconfigures the band edge alignment thus modulating the bandgap of an alloy. With the intent to improve efficiency using strain over the extended range of wavelengths covered by the GaInAs/GaAsSb type-II superlattice system it is necessary to estimate the shift in the conduction band and valence band edge due to the strain. The shift in conduction band δE_c, light hole δE_{v_LH} and heavy hole δE_{v_HH} can be described by the following set of equations.

$$\varepsilon_{\parallel} = \frac{a_o}{a_l} - 1 \tag{28a}$$

$$a_{\perp} = a_l\left(1 - 2\frac{C_{12}}{C_{11}}\,\varepsilon_{\parallel}\right) \tag{28b}$$

$$\varepsilon_{\perp} = \frac{a_{\perp}}{a_l} - 1 \tag{28c}$$

$$\frac{\Delta\Omega}{\Omega} = 2\varepsilon_{\parallel} + \varepsilon_{\perp} \tag{28d}$$

$$\Delta E = 2b_l(\varepsilon_{\perp} - \varepsilon_{\parallel}) \tag{28e}$$

$$\delta E_c = a_c\frac{\Delta\Omega}{\Omega} \tag{28f}$$

$$\delta E_{v\_HH} = a_v\frac{\Delta\Omega}{\Omega} - \frac{\Delta E}{E} \tag{28g}$$

$$\delta E_{v\_LH} = a_v\frac{\Delta\Omega}{\Omega} + \frac{\Delta E}{E} \tag{28h}$$

$\varepsilon_{\parallel}$ and $\varepsilon_{\perp}$ are the parallel and perpendicular component of strain on the layer. a_c, a_v represent the hydrostatic deformation potentials and b_l is the shear

deformation potential.

It is important to have a consistent method to calculate the strain-balance condition for the superlattice system. Daukes *et.al* [73] published a comprehensive paper comparing the available methods to calculate strain-balance conditions. The paper categorizes them as average lattice method, thickness weighted method and zero-stress method. The average lattice method calculates the final lattice parameter a_0 from the thickness average of the lattice constants for the compressive and tensile layers. Equation (29) represents this method numerically where the thicknesses of the two layers are represented by d_A, d_B and their respective lattice constants by a_A, a_B. Second technique equates the strain-thickness products for the tensile and compressive layers. The basic form for this method is shown in Equation (30a) and (30b) which neglects the differences in the elastic constants (A_A, A_B) defined by deformation potentials of the layer. The Equation (30c) and (30d) takes into account the differences between the elastic constants.

$$a_o = \frac{d_A a_A + d_B a_B}{d_A + d_B} \tag{29}$$

$$d_A \varepsilon_A + d_B \varepsilon_B = 0 \tag{30a}$$

$$a_o = \frac{(d_A + d_B) a_A a_B}{d_A a_B + d_B a_A} \tag{30b}$$

$$A_A d_A \varepsilon_A + A_B d_B \varepsilon_B = 0 \tag{30c}$$

$$a_o = \frac{(A_A d_A + A_B d_B) a_A a_B}{A_A d_A a_B + A_B d_B a_A} \tag{30d}$$

$$A_A d_A \varepsilon_A a_B + A_B d_B \varepsilon_B a_A = 0 \tag{31a}$$

$$a_o = \frac{(A_A d_A a_A a_B^2 + A_B d_B a_B a_A^2) a_A a_B}{A_A d_A a_B^2 + A_B d_B a_A^2} \quad (31b)$$

The paper presents zero stress method as the most rigorous derivation for finding the strain-balance condition. The biaxial strain on the layers induces an effective in-plane stress which causes the lattice distortion. Hence the method considers the effect of stress when strain-balancing two layer with opposite strain. The zero average in-plane stress can be derived from the average strain energy density with respect to strain in one layer [73]. This is then used to deduce the final lattice constant as shown in (31b).

To compare the different techniques a strain-balanced condition for 1% tensilely strained $Ga_{0.61}In_{0.39}As$ and 1% compressively strained $GaAs_{0.38}Sb_{0.62}$ layers was calculated using all the four methods. For a 1.5 nm thick $Ga_{0.61}In_{0.39}As$ layer the equivalent thicknesses yielded for the $GaAs_{0.38}Sb_{0.62}$ by the four methods are listed in the table below.

Table 3: Comparison between different strain-balancing methods.

	Average lattice	**Basic thickness weighted**	**Thickness weighted with elastic constants**	**Zero-Stress**
d_{GaAsSb} (nm)	1.47	1.50	1.58	1.62

Since the differences in the resultant values for the four methods are practically insignificant the basic form of thickness weighted method is used to calculate the strain-balance.

5.2 Band alignment for strained GaInAs-GaAsSb

The band edge alignment of a compound is modulated with the introduction of strain into its lattice. Thus to realize its potential it is important to analyze these alignments. The band edge alignment for the compressively strained GaAsSb layer and the tensilely strained GaInAs layer for three different strains is shown in Figure 26.

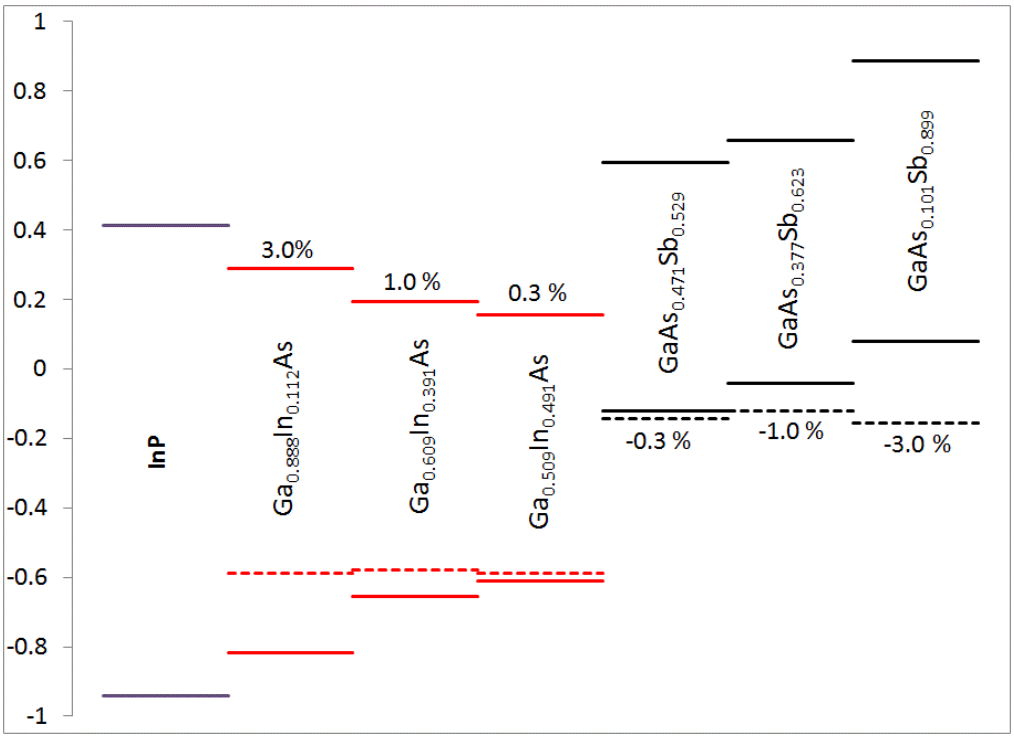


Figure 26: Band edge alignment for tensilely strained GaInAs on InP is shown by the red lines. Similarly the black lines denote the compressively strained GaAsSb layer on InP. The dashed lines represent the light hole energy level for each alignment.

In order to compressively strain $GaAs_{1-x}Sb_x$ on InP substrates the lattice constant for the tertiary alloy has to be greater than the lattice matched conditions ($x > 0.49$). This is achieved by decreasing the mole fraction of smaller arsenic atom and replacing it with the larger antimony atom. To introduce tensile strain in the $Ga_{1-y}In_yAs$ layer the lattice constant for the layer is reduced below the InP

lattice constant by increasing the gallium mole fraction in the alloy ($y < 0.53$). Since these alignments have a positive bandgap the uppermost solid line in Figure 26 stands for the conduction band and lower solid lines mark the heavy holes. The dashed lines represent the light hole level for each alignment. The alignment becomes more type-II with increase in strains on the layers thus implying validity of the strain method to improve the efficiency of the system.

Bandgap of the GaInAs layer increases from 0.73 eV to 0.74 eV with 0.3% strain. Although a reduction in the bandgap is expected but the tensile strain also results in increased Ga fraction. As the mole fraction of Ga is increased we move closer to the GaAs binary which has a much larger bandgap than the lattice matched $Ga_{0.47}In_{0.53}As$ and hence the increase in bandgap.

For the Figure 27 the GaAsSb is tensilely strained and the GaInAs is compressively strained. The lattice constant for $GaAs_{1-x}Sb_x$ is decreased below the lattice matched condition ($x < 0.49$) to apply tensile strain on the layer. To maintain the strain-balanced condition the strain on GaInAs has to be compressive. The lattice constant for $Ga_{1-y}In_yAs$ is increased beyond the lattice matched condition ($y > 0.53$) to compressively strain the layer.

Similar to the case of tensile strain in GaInAs, the tensile strain on GaAsSb also observes an increase in the bandgap due to higher content of GaAs in the alloy. It can be seen in Figure 27 that with increase in, tensile strain on GaAsSb layer and compressive strain on GaInAs the alignment becomes less type-II. This will defeat the very purpose of using type-II superlattice to reach bandgaps smaller than either constituent.

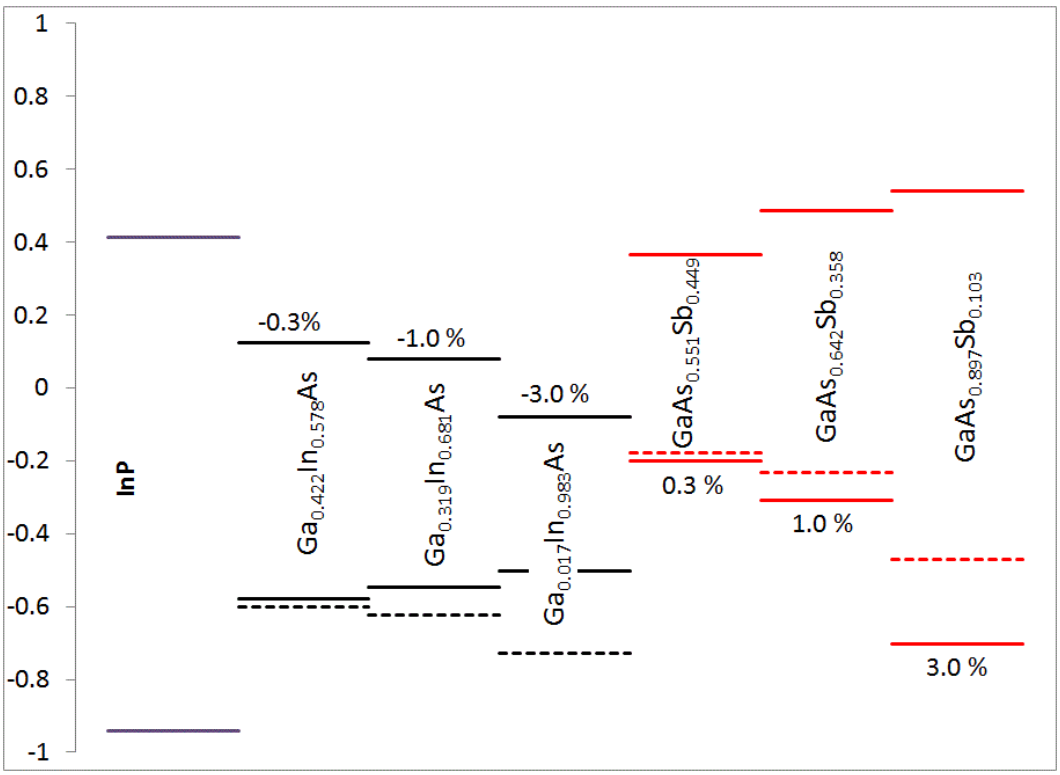


Figure 27: Band edge alignment for compressively strained GaInAs on InP is shown by the black lines. Similarly the red lines denote the tensilely strained GaAsSb layer on InP. The dashed lines represent the light hole energy level for each alignment.

In addition, even with state of the art growth techniques it is impossible to grow perfect crystalline layer with strains larger than 3%. Thus for all further calculations we restrict the maximum value of strain on either layer to 3%. Nevertheless the strain-balance method can still be utilized to improve the performance of the superlattice system for the extended wavelengths.

5.3 Extended wavelengths for strain-balanced GaInAs-GaAsSb

The transition energy for an unstrained superlattice system is dependent on the minibands which in turn is governed by the thicknesses of the two constituent layers. With induction of strain in the system the list is appended to include the strain on either layer thus increasing the complexity of the problem. Each set of

values for these four variables has to satisfy the thickness weighted relation in Equation (30a) to attain a strain-balanced condition. To tackle this problem we classify the problem into a two-fold process. First, find the optimal thickness design for all possible strain-balanced conditions. In the second fold each of these optimal designs for a specific strain balance condition are compared to pick the best design.

Since for the strain-balanced condition the transition energy is defined by the thickness and strain from each layer the possibility of realizing specific transition energy turns into a multi-dimensional problem. This problem is solved by slicing it into two parts, where for the first part the strain on a layer is kept constant and the remaining three variables are changed over a defined range in a plot. The constant strain on one layer is compensated by a varying strain on the other layer to achieve the strain balanced condition over a range of thicknesses. Since the strain and thickness for a layer are related by the strain-balanced condition they can be plotted on a single axis. Thus the transition energy can be again plotted as a function of the two layer thicknesses.

There are four ways to plot the varying layer thicknesses and strains for the system. In the first case, a family of curves for different thicknesses with a varying compressive strain on one layer is plotted while the layer with constant tensile strain is held at the x-axis. Second method swaps the strain between the layers so that a family of curves is plotted with a varying tensile strain and the constant compressive strain layer is held at the x- axis. The varying strain for the above mentioned methods can also be held at the x-axis. This variation adds two

more methods of plotting where the strain on the layer at the horizontal axis can be varied or kept constant.

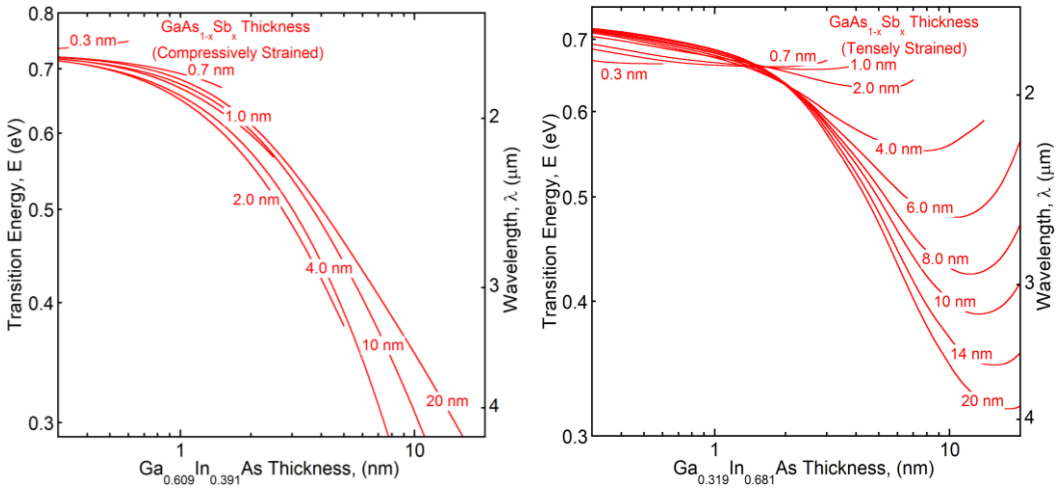
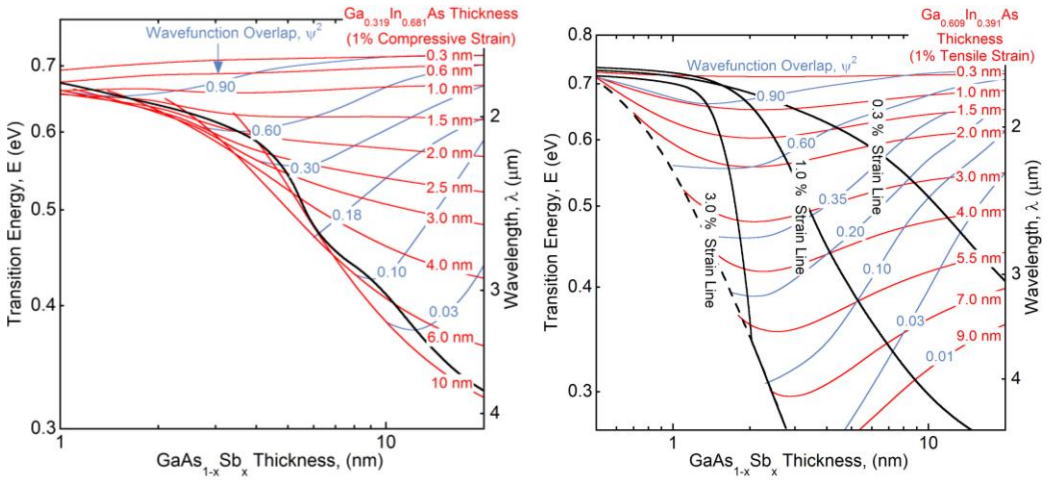


Figure 28: Transition energy versus GaInAs thickness curves for varying GaAsSb thickness. Each red curve indicates a variable a) compressive and b) tensile strain on GaAsSb layer for a constant 1% a) tensile and b) compressive strain on GaInAs layer.

Figure 28(a) depicts the first case where the $Ga_{0.319}In_{0.681}As$ layer having 1% constant tensile strain is held at the horizontal axis. The family of curves for different thickness is plotted with compressive strain below 3% on the GaAsSb layer. The image on the right swaps the strain to apply 1% compressive strain on $Ga_{0.319}In_{0.681}As$ layer whereas a varying tensile strain below 3% is applied on GaAsSb.

In Figure 29 the varying strain on GaAsSb is held at the x-axis. The $GaAs_{1-x}Sb_x$ layer is tensilely strained in Figure 29(a) and family of curves is then traced for 1% compressive strain on GaInAs. For Figure 29(b) the strain on the layers are swapped to apply compressive strain on the GaAsSb layer and tensile on GaInAs layer.



(a) (b)

Figure 29: Transition energy versus GaAsSb thickness curves for varying GaInAs thickness. Each red curve indicates a constant 1% a) compressive and b) tensile strain on GaInAs layer for a variable a) tensile and b) compressive strain on GaAsSb layer.

The plots in Figure 28 show that the family of curves with constant strain on the horizontal axis are too crammed and complex. This makes it extremely difficult to plot and compare the contours for the squared wavefunction overlap on these plots. Due to these complications we can rule out the methods which plot the constant strain on the horizontal axis. In Figure 29 we assess the method of plotting the layer with varying strain on the horizontal axis and have family of curves for the constant strain. These graphs look promising as the family of curves is well spaced and it is lot simpler to plot the wavefunction overlap contours. Figure 29(b) shows optimal design line to be limited by the maximum strain allowed on the varying strain layer therefore simplifying the comparison of various designs. Thus it can be deduced that it is best to plot the family of curves for the layer with constant tensile strain while the layer with varying compressive strain is held at the x-axis.

The above graphs were limited to 1% constant strain with either compressive or tensile strain on the layers. To realize the optimal solution for the strain balanced conditions it is required to explore these conditions over a range of strain levels. To compare the acquired results the range here is limited to 0.3%, 1% and 3% for either direction of strain.

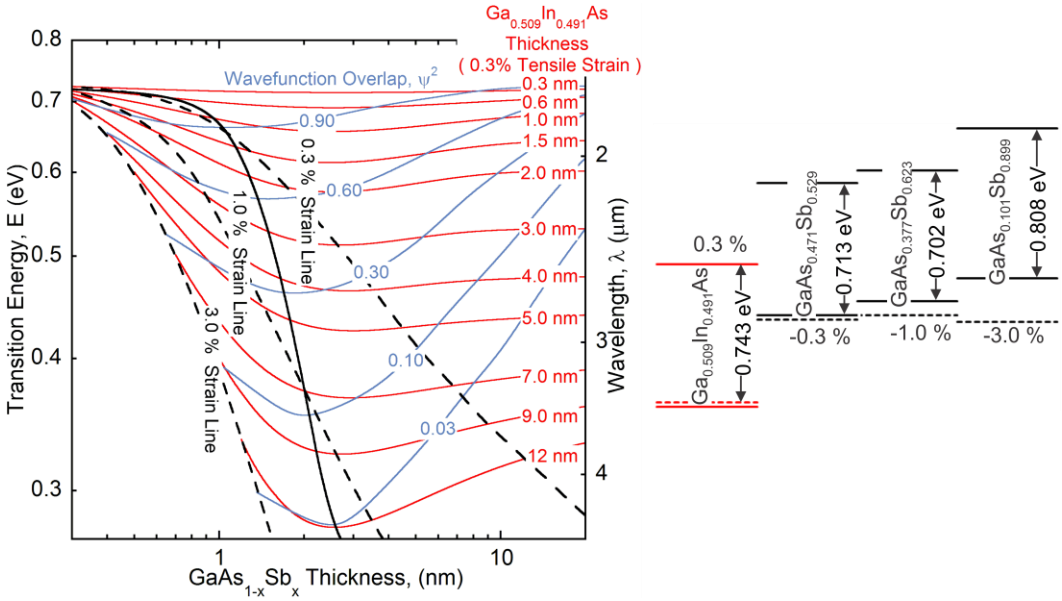


Figure 30: Transition energy versus GaAsSb thickness curves with constant squared wavefunction overlap curves in blue. Each red curve denotes a varying compressive strain on GaAsSb layer with a constant 0.3% tensile strain on GaInAs layer. The dashed lines represent curves of constant compressive strain on GaAsSb layer. The inset depicts the band edge alignment for various strain conditions marked on the plot.

To begin with we plot the squared wavefunction overlap contour plot for 0.3% tensilely strained $Ga_{0.509}In_{0.491}As$. The compressive strain on GaAsSb varies to attain a strain balanced condition for each pair of layer thicknesses on the red curves. The dashed black lines on the contour plot depict the constant strain lines for GaAsSb over the entire region. The band edge alignment for these strain lines, along with the tensilely strained GaInAs is shown to the right of the plot.

The solid black line represents the optimal design in the 1.7 to 4 μm range. Square of the wavefunction overlap at 2.7 μm improved from 0.19 to 0.27 as compared to the lattice matched $Ga_{0.47}In_{0.53}As$-$GaAs_{0.51}Sb_{0.49}$ system.

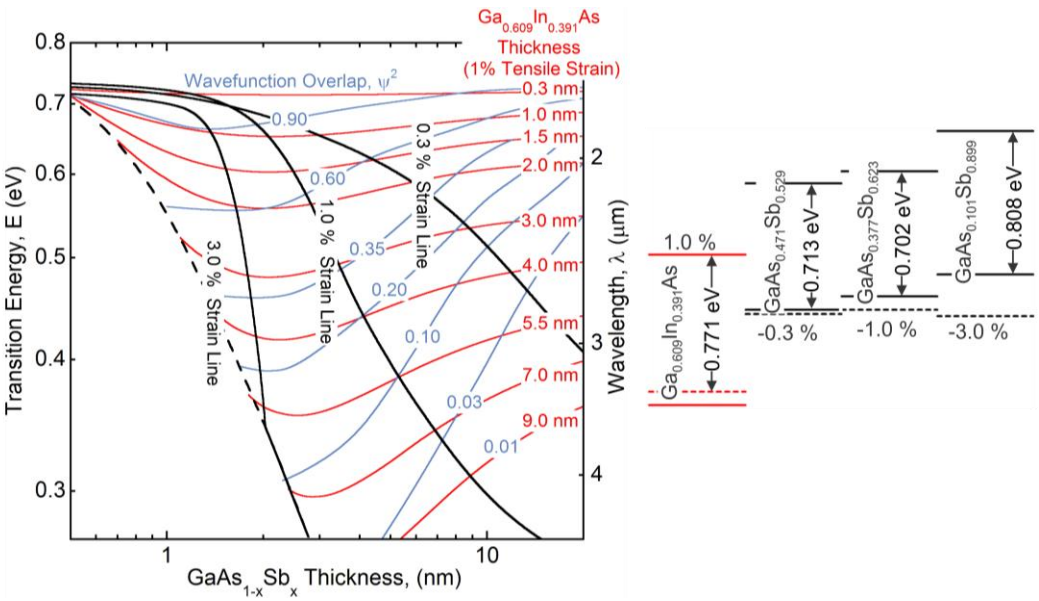


Figure 31: Transition energy versus GaAsSb thickness curves with constant squared wavefunction overlap curves in blue. Each red curve denotes a varying compressive strain on GaAsSb layer with a constant 1.0% tensile strain on GaInAs layer.

In Figure 31 the tensile strain on GaInAs is increased to 1% and a notable difference from the 0.3% tensile strain is the optimal design line. With 1% tensile strain on the GaInAs layer the optimal design line is defined by the strain on GaAsSb and hence each constant strain line is shown as a solid black line. Only for the 3% strain line the best wavefunction for the wavelengths are defined by a minimum point on the overlap curves. Thus the strain line is shown by the dashed black line until the optimal design line merges into it. Also square of the wavefunction overlap is observed to improve for wavelengths beyond 3.6 μm with increase in compressive strain on GaAsSb. This stands true for the shorter wavelengths if the strain is limited to 1% on the GaAsSb layer.

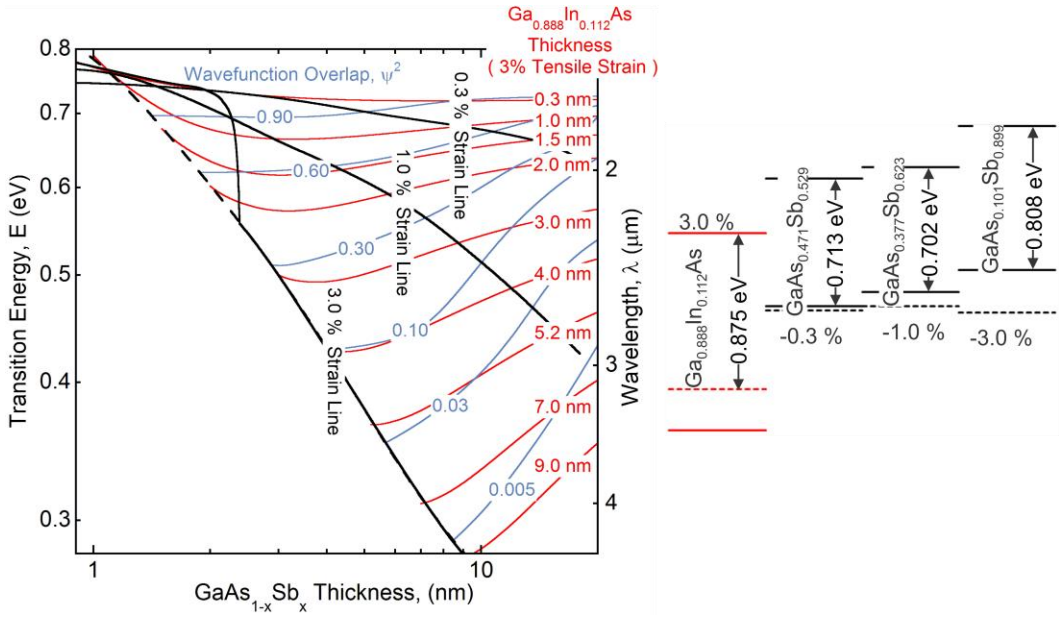


Figure 32: Transition energy versus GaAsSb thickness curves with constant squared wavefunction overlap curves in blue. Each red curve denotes a varying compressive strain on GaAsSb layer with a constant 3.0% tensile strain on GaInAs layer.

Table 4: Critical thicknesses for compressively strained GaAsSb layer and tensilely strained GaInAs layer.

Strain (%)	GaAsSb critical thickness (nm)	GaInAs critical thickness (nm)
0.3	340	330
1	54.4	53.51
3	13.95	13.83

As the light hole shifts to higher energy for 3% tensile strain on GaInAs layer it is possible for the ground transition to be defined by a electron and light hole miniband. However for light holes to define the highest miniband the GaInAs layer needs to much thicker compared to the GaAsSb. Since the GaAsSb thicknesses for which this condition is satisfied is extremely small and thus we

can neglect them. On comparing the best squared wavefunction overlap at 2.7 μm for 1% and 3% tensile strain on GaInAs we observe a reduction from 0.35 to 0.20. The critical thicknesses calculated using Equation (27a) for the two layers at the different strain levels are listed in Table 4.

The wavefunction overlap improved for the 0.3 % and 1 % tensile strain on the GaInAs but reduced for 3% strain. This result suggests an optimized strain level for the tensilely strained GaInAs between the 1% and 3% strain levels. At the same time it is important to reverse the strain between the layers and carry out similar simulations to characterize the best strain design for the system.

For the next group of simulations the GaAsSb layer is tensilely strained and a complementary compressive strain is applied on the GaInAs layer. Similar to the earlier plots the family of curves for different thickness is traced for the layer with constant tensile layer. The layer with varying compressive strain is held at the horizontal axis. Figure 33 shows the effect of 0.3% tensilely strained GaAsSb layer on the family of curves for varying compressive strain on the GaInAs layer at the horizontal axis. The dashed black lines represent the constant strain for the GaInAs layer. The valence band regains its degeneracy for all GaInAs thicknesses beyond 2nm when the GaAsSb layer is 12nm thick. Beyond this thickness of GaAsSb the light holes begin to define the ground state transitions.

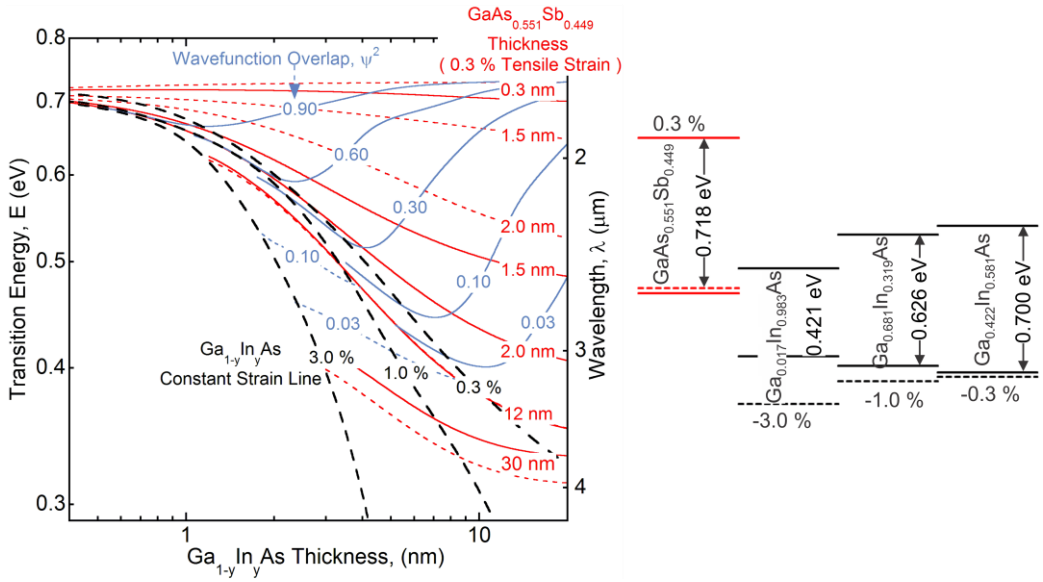


Figure 33: Transition energy versus GaInAs thickness curves with constant squared wavefunction overlap curves in blue. Each red curve denotes a varying compressive strain on GaInAs layer with a constant 0.3% tensile strain on GaAsSb layer. The dashed black lines represent curves of constant compressive strain on GaInAs layer. The dashed red and blue curves represent the lowest electron to light hole transition and its wavefunction overlap respectively.

In general, the light holes are less confined than heavy holes due to their higher transport factor. Hence the wavefunction overlap for the light holes can be larger than the heavy holes. However the density of states for light holes is lower than the heavy holes due to the difference in their masses which may result in reduced absorption. The superlattice density of states is directly proportional to the effective mass of the charge carrier [74]. Since the heavy hole for both layers is 10 times greater than the light hole we can approximate the density of states to be higher by a similar factor. Therefore the light hole governed transitions cannot be directly compared with the heavy hole transitions. To demarcate the difference between ground state transitions defined by the light hole each and heavy hole each thickness curve has been marked by dashed and solid lines respectively.

Furthermore due to the difference between the effective masses for the light hole and heavy hole there is a difference between the shifts observed in the respective minibands. This difference explain the large energy gap between the light and heavy hole ground state transitions as the strain on GaInAs layer reduces below 0.3%. In addition the energy levels of the minibands are also largely reliant on the thicker GaInAs layer which augments the downward shift of the light hole miniband in turn increasing the electron to light hole transition energy. Once the GaAsSb layer thickness is large enough this effect reduces and the light hole ground state transitions begin to define the superlattice bandgap.

The wavefunction overlap for the graph is also shown separately for the light hole and heavy hole ground transition. The squared wavefunction overlap for the heavy hole is shown by solid blue lines whereas the dashed blue lines represent the light hole transitions. The optimal design line is defined by the best wavefunction overlap achieved at each wavelength. Due to its proximity to the 0.3% tensile strain line for GaInAs it has been shown separately by the solid black line in the contour plot for Figure 34.

With the increase of tensile strain on the GaInAs to 1% the range of wavelengths covered by the heavy hole transition reduces. This can be seen in Figure 35 where the light holes begin to dominate the ground state transitions for GaAsSb layer thicknesses beyond 5 nm compared to the 12 nm with 0.3% tensile strain on GaAsSb. The larger downward shift of the heavy hole level with increasing tensile strain on GaAsSb layer results in shift of the heavy hole miniband to lower energy.

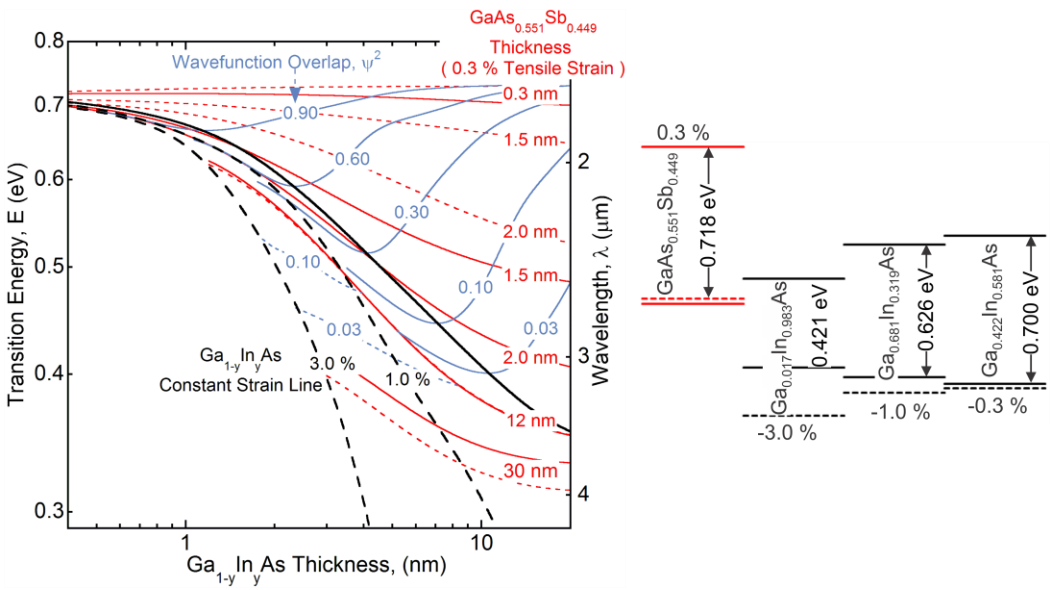


Figure 34: Transition energy versus GaAsSb thickness curves with optimal design line for varying compressive strain on GaAsSb layer with a constant 0.3% tensile strain on GaInAs layer.

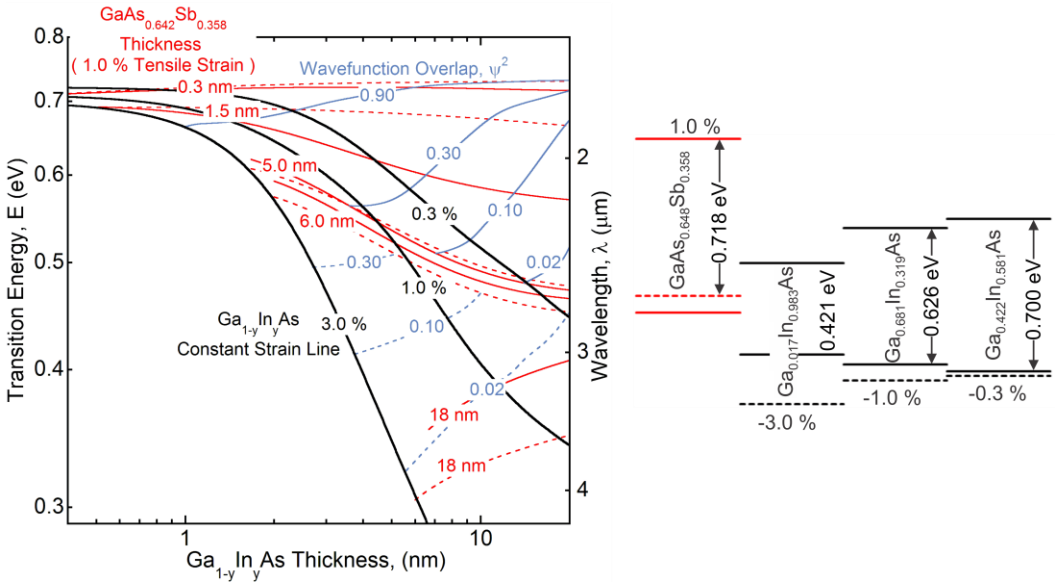


Figure 35: Transition energy versus GaInAs thickness curves with constant squared wavefunction overlap curves in blue. Each red curve denotes a varying compressive strain on GaInAs layer with a constant 1.0% tensile strain on GaAsSb layer.

Another noticeable change is the optimal design line which is now defined by the amount of compressive strain applied on the GaInAs layer. Square of the wavefunction overlap increased for 2.7 μm from 0.07 with 0.3% tensile strain on

GaAsSb to 0.19 at 1% tensile strain. Since the increment in wavefunction overlap is observed for the light hole transitions it is not appropriate to compare the improvement with the heavy hole transitions in the earlier designs.

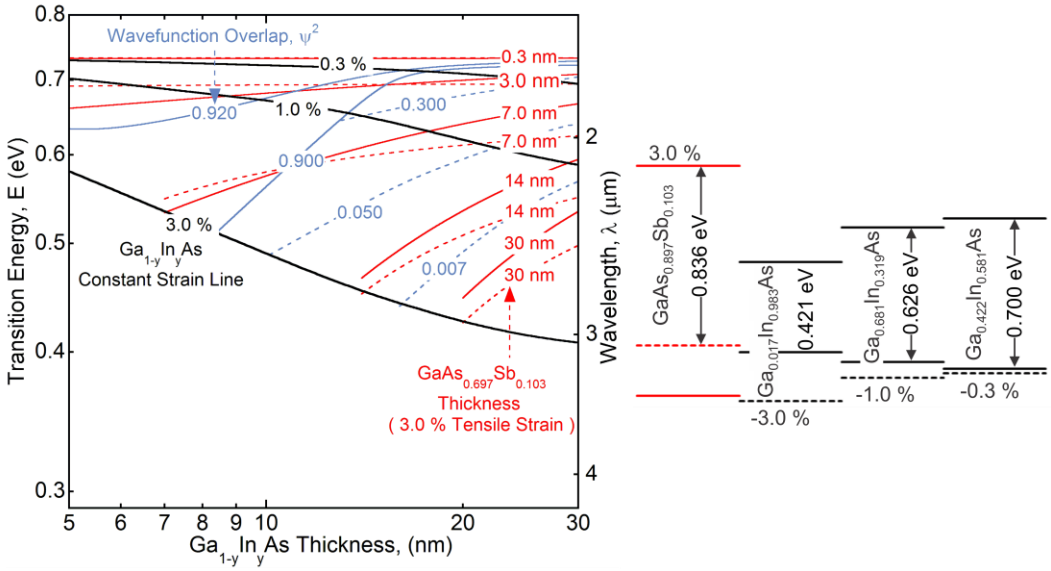


Figure 36: Transition energy versus GaInAs thickness curves with constant squared wavefunction overlap curves in blue. Each red curve denotes a varying compressive strain on GaInAs layer with a constant 3.0% tensile strain on GaAsSb layer.

As the tensile strain is increased to 3% the ground state transition is mostly dominated by light hole transitions. The reduction in the longest achievable wavelength can be associated with the reduced type-II alignment between the two layers. The increased GaAsSb bandgap due to larger GaAs fraction in the alloy shifts the valence band downwards thus reducing the type-II alignment with GaInAs. The minimum possible transition is then dominated by the compressively strained GaInAs layer. The heavy hole always forms a type-I alignment but the light hole remains type-II for all configurations. This explains the high wavefunction overlap achieved for the heavy hole transitions, marked by the solid light blue lines. On the other hand the wavefunction overlap drops

drastically as the transitions are defined by the light holes. The squared wavefunction overlap for the 2.7 μm wavelength defined by the light holes is reduced to 0.01 from 0.19 at 1% tensile strain.

Furthermore, the thicknesses utilized to cover the extended wavelength regime exceed the critical thickness values. Thus the longest wavelength that can be realized without creating misfit dislocations is limited to 2.72 μm. The critical thicknesses for the layers at different strain levels have been listed in Table5.

Table 5: Critical thicknesses for tensilely strained GaAsSb layer and compressively strained GaInAs layer.

Strain (%)	GaAsSb critical thickness (nm)	GaInAs critical thickness (nm)
0.3	338	328
1	54.33	52.84
3	13.89	13.27

Though the strain-balance designs shown here are limited to 0.3%, 1% and 3% strain on the layer but they were explored for a larger range of strains. To evaluate the optimum design from these designs their performances can be compared for a specific wavelength. The Figure 37 utilizes the 2.7 um wavelength for this comparison. The horizontal axis of the graph varies the antimony mole fraction in GaAsSb which in turn changes the strain while as each curve on the plot represents a constant strain level on the GaInAs layer.

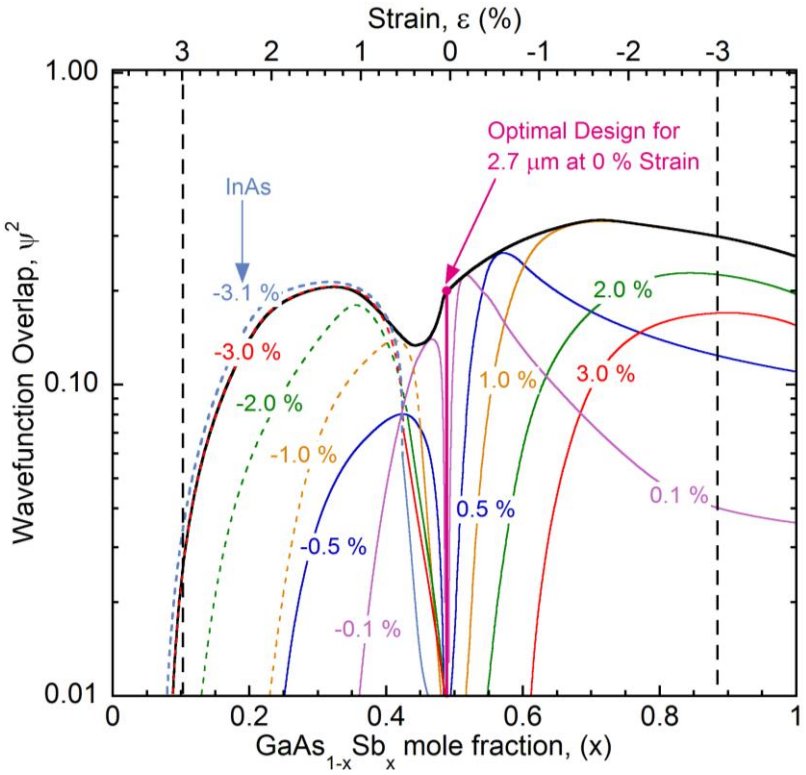


Figure 37: Comparison of optimal designs at different GaAsSb mole fraction for the 2.7 μm wavelength. Each curve represents a constant compressive and tensile strain on GaInAs layer respectively to the left and right of the zero strain line in pink.

The wavefunction overlap for the perfectly lattice matched condition on InP is shown in pink with the highest attainable squared wavefunction overlap for 2.7 μm marked as the optimal design for zero strain. As the value of x in $GaAs_{1-x}Sb_x$ reduces to the left of the lattice matched condition the tensile strain on the layer increases. The curves on this side correspond to the compressively strained GaInAs layer acquired from the strain-balanced contour plots. Thus each point on the curve is defined by a unique pair of layer thicknesses listed in Table 6. The amount of compressive strain on the layer is numerically marked in percentage. Similarly, to the right of the lattice matched condition GaAsSb layer is compressively strained and each curve represents the tensilely strained GaInAs

layer. The dashed black line represents the 3% strain in either direction on GaAsSb layer.

Table 6: Optimal layer thicknesses for different GaAsSb strain levels at the 2.7 μm wavelength.

		Tensile Strain (%)					Compressive Strain (%)				
Layer Thickness	**0**	**0.1**	**0.5**	**1.0**	**2.0**	**3.0**	**-0.1**	**-0.5**	**-1.0**	**-2.0**	**-3.0**
$Ga_{1-y}In_yAs$ (nm)	3.2	5.4	5.0	3.2	5.2	13.1	4.5	3.9	3.6	3.6	4.2
$GaAs_{1-x}Sb_x$ (nm)	5.3	3.2	9.0	9.3	7.8	13.1	3.8	3.9	3.6	1.8	1.4

The curves shown by dashes represent the overlap attained by a light hole transition. The valence band becomes degenerate at the point where these optimal design lines transform from heavy hole to a light hole transition. Thus we expect a better performance at this point due the higher density of states with degeneracy in the valence band. It is evident that the tensilely strained GaInAs and compressively strained GaAsSb is the best choice for strain balanced designs. Firstly, the squared wavefunction overlap is larger for all compressive strains on GaAsSb when compared to their equivalent complementary case. Secondly, the ground state transitions are never light hole based for a compressively strained GaAsSb layer. The overlap is best for tensile strains below 1% on GaInAs when compressive strain is limited to -0.9% on GaAsSb. However this trend changes for all compressive strains beyond -0.9% on GaAsSb where the best overlap is always achieved with 1% tensile strain on GaInAs layer. The maximum wavefunction overlap that can be realized from the various optimal strain balance conditions is traced by the solid black line. Thus the curve can be termed as the optimized design line for 2.7 μm over the entire range of possible strains on the

two layers. The highest squared wavefunction overlap of 0.35 is attained with 1.0% tensile strain on GaInAs and -1.88% compressive strain (x = 0.75) on GaAsSb layer. This point defines the best strain-balanced design for the GaInAs/GaAsSb system at 2.7 µm. The corresponding layer thickness for GaInAs and GaAsSb are 3.1 nm and 1.8 nm respectively.

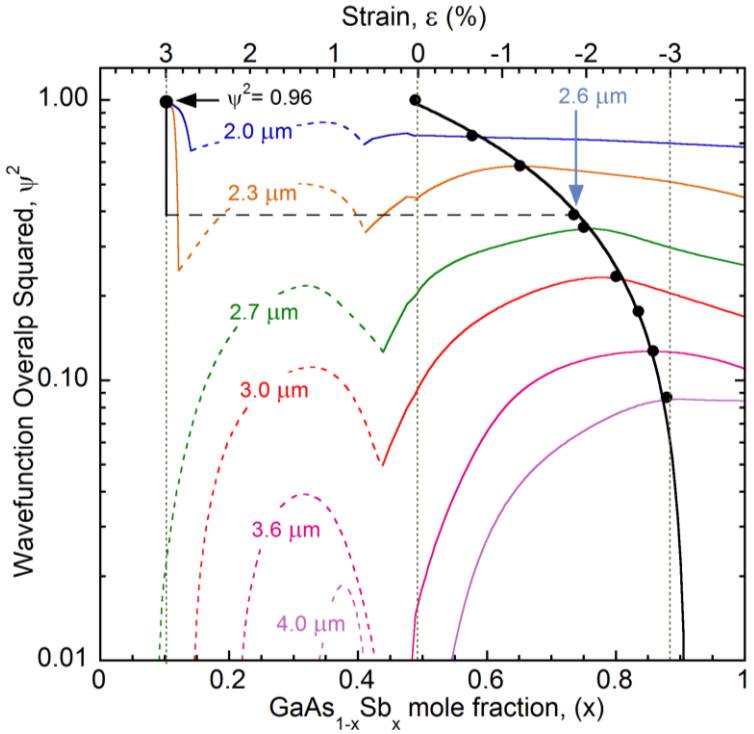


Figure 38: Optimal design curves for varying GaAsSb mole fraction for selective wavelengths in the 1.7 to 4.0 µm spectrum. Dashed curves and solid curves denote light hole and heavy hole based optimal designs.

Similarly, the optimal design across all possible strain balance design for different wavelengths in the 1.7 to 4.0 µm can be traced to find a generalized solution for the spectrum. The optimized design curves for selective wavelengths are shown in the **Error! Reference source not found.**. Similar to the 2.7 µm wavelength the optimal design was dominated by the light holes for tensilely strained $GaAs_{1-x}Sb_x$ ($x < 0.49$) and the heavy holes for compressively strained

designs The light holes and heavy holes designs have been demarcated by the dashed and solid lines respectively.

The discontinuity observed in the 2 μm and 2.3 μm curves at the zero strain level is due to the change of shift in band alignment with opposite strain on either side. The optimal design for each of the wavelengths is defined by the highest squared wavefunction overlap attained across the strain balanced conditions. The wavelengths beyond 2.6 μm behaved similar to the 2.7 μm curves for which, the best design is observed with a compressive strain on GaAsSb layer and a tensile strain on GaInAs layer.

The highest overlap for wavelengths below 2.6 μm is found with 3% tensile strain line on GaAsSb layer and very large opposite strain on GaInAs. The resultant band edge alignment between the superlattice layers switches from type-II to type-I alignment. This explains the proximity to ideal wavefunction overlap for wavelengths below 2.6 μm. Thus the optimal design line connecting the wavefunction overlap peaks for different wavelengths shown by the solid black curve is discontinuous at 2.6 μm. It resumes at the 3% tensile strain line, which can be shifted to higher tensile strains to realize an ideal wavefunction overlap for these wavelengths. Since it is difficult to have good growth quality with such high strain on the layer the optimal design line for these wavelengths has also been traced for the compressively strained GaAsSb. The optimal design line is curve fit to deduce a parametric relation between the squared wavefunction overlap Ψ^2 and mole fraction x of GaAsSb.

$$\Psi^2 = 2.10 - 2.33x, \quad \text{for } x \geq 0.489 \tag{32}$$

$$d_A = d_{s1}\left(exp\left(\frac{E_{s1}-E}{E_{s2}}\right) - 1\right) \quad (33a)$$

$$d_B = d_{s2}\left(1 - \frac{E}{E_{s3}}\right) \quad (33b)$$

$$x_B = A_{s1} - \frac{E}{E_{s4}} \quad (33c)$$

$$\Psi^2 = \left(\frac{E}{E_{s5}}\right)^{p_s} - A_{s2} \quad (33d)$$

Table 7: Constants defined in the optimized strain balance design rules.

GaInAs thickness versus Transition energy	**GaAsSb thickness versus Transition energy**	**Mole fraction versus Transition energy**	**Squared overlap versus transition energy**
d_{s1} = 1.348 nm	d_{s2} = 3.339 nm	A_{s1} = 1.195	E_{s5} = 0.657 eV
E_{s2} = 0.839 eV	E_{s3} = 0.973 eV	E_{s4} = 1.032 eV	A_{s2} = 0.18
E_{s3} = 0.296 eV			p_s = 1.78

Each point on the optimal design line in **Error! Reference source not found.** also comprises the thicknesses of the two layers which in turn define the optimal superlattice period for a given wavelength. These thicknesses can be utilized to form a parametric equation between the transition energy or wavelength and superlattice period thickness. Eventually these equations can be used to realize an optimized superlattice design for each wavelength in the spectrum. Equation (33a) fulfills this need by relating the transition energy with the thickness d_A and d_B, where A and B represent GaInAs and GaAsSb respectively. The relation is deduced from the curve traced between the various transition energies and thicknesses as shown Figure 39.

Furthermore for an optimal design at any wavelength it is necessary to define a relation between the transition energy and mole fraction which is accomplished in Equation (33c) deduced from the GaAsSb mole fraction curve in Figure 40. No relation is defined for the GaInAs mole fraction since the optimal design for all wavelengths above 2.1 µm is realized with 1% strain on GaInAs. For wavelengths below 2.1 µm the zero strain condition is chosen as the optimal design although the increment in peak overlap achieved for these wavelengths is practically insignificant for tensile strains below 1% on GaInAs. Equation (33d) is deduced in Figure 41 which defines the maximum squared wavefunction overlap that can be attained with these optimal design rules for a particular cut-off wavelength.

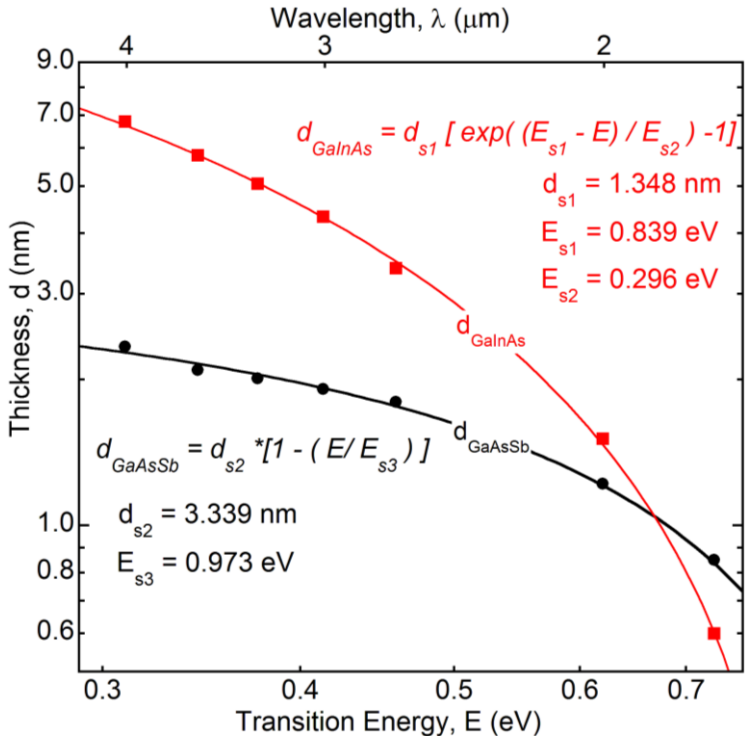


Figure 39: Optimal design curves for GaInAs and GaAsSb layer thickness at different wavelengths.

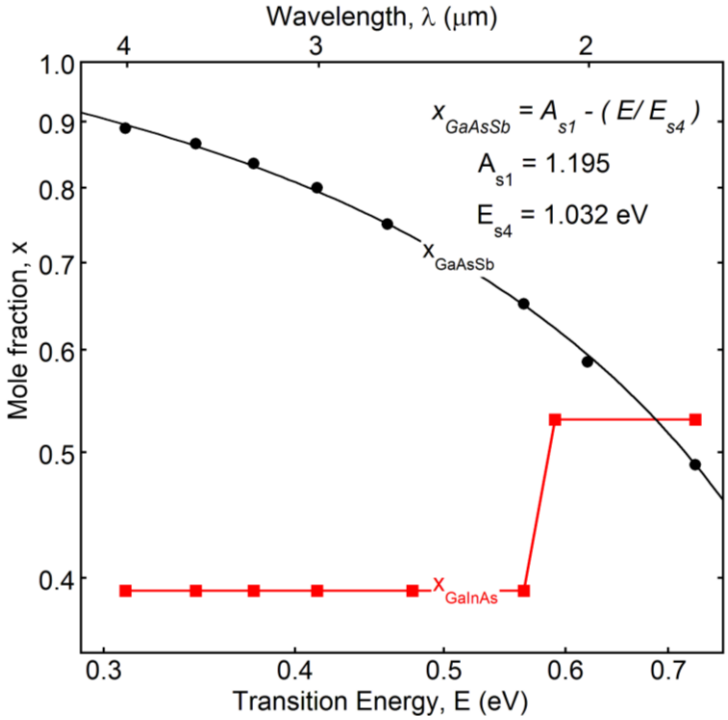


Figure 40: Optimal design curves for GaInAs and GaAsSb mole fraction at different wavelengths.

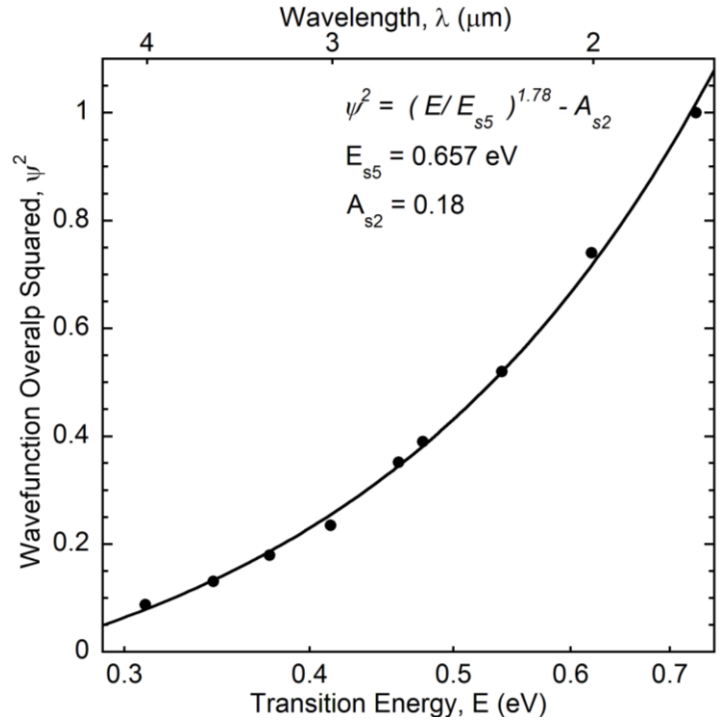


Figure 41: Largest squared wavefunction overlap realized with strain balanced GaInAs/GaAsSb superlattice system over the 1.7 to 4.0 μm wavelength range.

6 CONCLUSIONS

Type-II superlattice structures for GaInAs-GaAsSb have been thoroughly studied to design efficient photodetectors for the 1.7 to 4.0 μm spectrum. Highly mismatched alloys and strain-balanced structures are utilized to engineer the type-II band alignment between GaInAs and GaAsSb. The resultant alignments aid in improving the optical properties for the extended wavelength designs. Precise simulations were carried out using the valence band anticrossing model for the band structure of highly mismatched alloys, Kronig-Penney formulation for superlattice minibands and thickness weighted method for strain-balanced superlattice structures. The results from these calculations are then condensed into squared wavefunction contour plots to extract an optimal superlattice period for a given cutoff wavelength.

Perfectly lattice matched $Ga_{1-y}In_yAs_{1-x}Bi_x$ alloys on InP substrates can be realized for the entire wavelength range with less than 6.8% bismuth in the alloy due to the anomalous bandgap reduction observed in mismatched alloys. The absorption performance of GaInAs/GaAsSb type II superlattices on InP is enhanced by adding dilute amounts of bismuth to the GaInAs layer. The type-II band alignment is modified with the incorporation of bismuth which consequently improves the wavefunction overlap. The transitions in the type-II GaInAsBi/GaAsSb superlattice structure can be classified into two regimes. Firstly, the region with ideal wavefunction overlap that lies between the bandgaps of the two constituent materials also called the short period superlattice regime. The second region, with less than ideal wavefunction overlap is bound between

the energies smaller than either constituent's bandgap and the smallest possible transition that can be realized by the type-II alignment between GaInAs/GaAsSb. The wavelengths realized in this range fall beyond the limit of either constituent and hence the name extended wavelength regime.

The squared wavefunction overlap doubles when the bismuth mole fraction is increased from 0% to 3%. But with increase in bismuth mole fraction the short period superlattice regime widens as the bandgap of the GaInAsBi layer reduces consequently narrowing the extended wavelength regime. Since the entire spectrum can be covered with bulk GaInAsBi a trade-off has to be met between the wavefunction overlap and bismuth concentration. The optimal design rules formulated can be utilized to deduce the superlattice period, to achieve a maximum overlap at a cut-off wavelength with different bismuth concentrations.

Furthermore, the strain-balanced GaInAs/GaAsSb superlattice designs were investigated for optimal strain conditions. The presence of two tertiary alloys in the system makes it feasible to have compressive or tensile strain in either layer with various strain levels and compensate it with complementary strain in the adjacent layer largely increasing the number of possible designs.

The optimal designs for various strain balance conditions are compared for 2.7 μm. From the comparison, strain balance with compressively strained GaAsSb and tensilely strained GaInAs was found to give the best overlap. The highest squared wavefunction overlap improved to 0.35 in comparison to the 0.19 for no strain in either layer. This overlap was attained with 1.0% tensile strain on

GaInAs and -1.9% compressive strain on GaAsSb with a layer thickness of 3.4 nm and 1.8 nm respectively. Strain balance conditions for the tensilely strained GaAsSb merely reached a square wavefunction overlap equivalent to the lattice matched condition with -3% compressive strain on GaInAs and 1.29% on GaAsSb. Similarly, highest overlap conditions for other wavelengths in the 1.7 to 4.0 μm spectrum are extracted for all mole-fractions of GaAsSb. The maximum wavefunction overlap for the 2.6 to 4.0 μm range was obtained with 1% tensile strain on GaInAs while as the compressive strain on GaAsSb increases with the shift towards longer wavelengths. All wavelengths below 2.6 μm approach an ideal wavefunction overlap for 3% tensile strain on GaAsSb layer due to the type-I alignment between the layers. Parametric equations describing the relation between the layer thicknesses and transition energy are derived. These equations can be further used to design superlattices with maximum wavefunction overlap for a cut-off wavelength.

Another distinct feature observed with the different strain balance conditions was the type of ground state transition. A large part of the ground state transitions are governed by light holes for the tensilely strained GaAsSb layer where as these transitions are always defined by heavy holes for compressively strained GaAsSb layer. The heavy hole transitions are preferred for photodetector applications due to their higher density of states. Nevertheless, light hole based transitions could be utilized for applications that require higher hole transport.

Finally, the exploration of optimal designs with the aforementioned techniques yields promising results for the 1.7 to 4.0 μm range. It is highly

probable to fabricate devices with performance capabilities exceeding the presently available detector designs in the range.

REFERENCES

[1] B. Fluegel, S. Francoeur, A. Mascarenhas, S. Tixier, E. C. Young and T. Tiedje, "Giant spin-orbit bowing in $GaAs_{1-x}Bi_x$", *Phys. Rev. Lett.*, vol. 97, pp. 067205, 2006.

[2] S. Francoeur, M. -J. Seong, A. Mascarenhas, S. Tixier, M. Adamcyk and T. Tiedje, "Band gap of $GaAs_{1-x}Bi_x$, 0<x<3.6%", *Appl. Phys. Lett.*, vol. 82, pp. 3874, 2003.

[3] Y. Tominaga, K. Oe and M. Yoshimoto, "Low temperature dependence of oscillation wavelength in $GaAs_{1-x}Bi_x$Laser by photo-pumping", *Applied Physics Express*, vol. 3, pp. 062201, 2010.

[4] W. Walukiewicz, K. Alberi, J. Wu, W. Shan, K. M. Yu, and J. W. Ager, III *Electronic Band Structure of Highly Mismatched Semiconductor Alloys* Springer Berlin Heidelberg, 2008.

[5] R. A. Rohde, "Radiation Transmitted by the Atmosphere.Png." Global Warming Art http://www.globalwarmingart.com/wiki/File:Atmospheric_Transmission.png (accessed 01/10, 2013).

[6] M. Peter, R. Kiefer, F. Fuchs, N. Herres, K. Winkler, K. -H. Bachem and J. Wagner, "Light-emitting diodes and laser diodes based on a $Ga_{1-x}In_xAs/GaAs_{1-y}Sb_y$ type II superlattice on InP substrate", *Appl. Phys. Lett.*, vol. 74, pp. 1951, 1999.

[7] J. Y. T. Huang, "Design and characterization of strained InGaAs/GaAsSb type-II 'W' quantum wells on InP substrates for mid-IR emission", *Journal of Physics. D, Applied Physics*, vol. 42, pp. 025108; 025108, 2009.

[8] M. H. Crawford, "LEDs for solid-state lighting: Performance challenges and recent advances", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 15, pp. 1028, 2009.

[9] H. P. Maruska, D. A. Stevenson and J. I. Pankove, "Violet luminescence of mg-doped GaN", *Appl. Phys. Lett.*, vol. 22, pp. 303, 1973.

[10] J. I. Pankove, E. A. Miller and J. E. Berkeyheiser, "GaN blue light-emitting diodes", *J Lumin*, vol. 5, pp. 84, 1972.

[11] J. Pankove, H. P. Maruska and J. E. Berkeyheiser, "Optical absorption of GaN", *Appl. Phys. Lett.*, vol. 17, pp. 197, 1970.

[12] S. Nakamura, M. Senoh and T. Mukai, "P-GaN/N-InGaN/N-GaN double-heterostructure blue-light-emitting diodes", *Japanese Journal of Applied Physics. Part 2, Letters*, vol. 32, pp. L8, 1993.

[13] Y. Zhong and A. Mascarenhas, "Scaling of band-gap reduction in heavily nitrogen doped GaAs", *Phys. Rev. B.*, vol. 63, pp. 161303, 2001.

[14] K. Ma, Z. M. Fang, D. H. Jaw, R. M. Cohen, G. B. Stringfellow, W. P. Kosar and D. W. Brown, "Organometallic vapor phase epitaxial growth and characterization of InAsBi and InAsSbBi", *Appl. Phys. Lett.*, vol. 55, pp. 2420, 1989.

[15] B. Joukoff and A. M. Jean-Louis, "Growth of $InSb_{1-x}Bi_x$single crystals by czochralski method", *J. Cryst. Growth*, vol. 12, pp. 169, 1972.

[16] K. Oe, S. Ando and K. Sugiyama, "$lnSb_{1-x}Bi_x$Films grown by molecular beam epitaxy", *Japanese Journal of Applied Physics*, vol. 20, pp. L303, 1981.

[17] J. Zilko and J. E. Greene, "Growth and phase stability of epitaxial metastable $InSb_{1-x}Bi_x$films on GaAs. I. crystal growth", *J. Appl. Phys.*, vol. 51, pp. 1549, 1980.

[18] A. J. Noreika, J. Greggi, W. J. Takei and M. H. Francombe, "Properties of MBE grown InSb and $InSb_{1-x}Bi_x$", *Journal of Vacuum Science & Technology.A*, vol. 1, pp. 558, 1983.

[19] T. Humphreys, P. K. Chiang, S. M. Bedair and N. R. Parikh, "Metalorganic chemical vapor deposition and characterization of the in-as-sb-bi material system for infrared detection", *Appl. Phys. Lett.*, vol. 53, pp. 142, 1988.

[20] S. A. Barnett, "Direct E0 energy gaps of bismuth-containing III–V alloys predicted using quantum dielectric theory", *Journal of Vacuum Science & Technology.A, Vacuum, Surfaces, and Films*, vol. 5, pp. 2845, 1987.

[21] Z. M. Fang, K. Y. Ma, R. M. Cohen and G. B. Stringfellow, "Photoluminescence of InAsBi and InAsSbBi grown by organometallic vapor phase epitaxy", *J.Appl.Phys.*, vol. 68, pp. 1187, 1990.

[22] H. Okamoto and K. Oe, "Growth of metastable alloy InAsBi by low-pressure MOVPE", *Japanese Journal of Applied Physics.Part 2, Letters*, vol. 37, pp. 1608, 1998.

[23] K. Oe and H. Okamoto, "New semiconductor alloy $GaAs_{1-x}Bi_x$Grown by metal organic vapor phase epitaxy", *Japanese Journal of Applied Physics. Part 2, Letters*, vol. 37, pp. L1283, 1998.

[24] Y. Zhang, A. Mascarenhas and L. -W. Wang, "Similar and dissimilar aspects of III-V semiconductors containing bi versus N", *Phys. Rev. B*, vol. 17, pp. 155201, 2005.

[25] Y. Song, S. Wang, I. S. Roy, P. Shi and A. Hallen, "Growth of $GaSb_{1-x}Bi_x$ by molecular beam epitaxy", *Journal of Vacuum Science & Technology B*, vol. 30, pp. 1071, 2012.

[26] P. Y. Yu, M. Cardona, *Fundamentals of semiconductors*, Berlin; New York: Springer, 2001.

[27] H. Hjalmarson, P. Vogl, D. J. Wolford and J. D. Dow, "Theory of substitutional deep traps in covalent semiconductors", *Phys. Rev. Lett.*, vol. 44, pp. 810, 1980.

[28] D. J. Thomas, J. J. Hopfield and C. J. Frosch, "Isoelectronic traps due to nitrogen in gallium phosphide", *Phys. Rev. Lett.*, vol. 15, pp. 857, 1965.

[29] L. Nordheim, "The electron theory of metals", *Ann. Phys. Lpz.*, vol. 9, pp. 607, 1931.

[30] K. Uesugi, N. Morooka and I. Suemune, "Reexamination of N composition dependence of coherently grown GaNAs band gap energy with high-resolution x-ray diffraction mapping measurements", *Appl. Phys. Lett.*, vol. 74, pp. 1254, 1999.

[31] K. Alberi, J. Wu, W. Walukiewicz, K. M. Yu, O. D. Dubon, S. P. Watkins, C. X. Wang, X. Liu, Y. -J. Cho and J. K. Furdyna, "Valence band anticrossing in mismatched III-V semiconductor alloys", *Physica Status Solidi (C)*, vol. 4, pp. 1711, 2007.

[32] W. Shan, W. Walukiewicz, J. W. Ager, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson and S. R. Kurtz, "Band anticrossing in GaInNAs alloys", *Phys. Rev. Lett.*, vol. 82, pp. 1221, 1999.

[33] A. F. Phillips, S. J. Sweeney, A. R. Adams and P. J. A. Thijs, "The hydrostatic pressure dependence of the threshold current in 1.3 μm InGaAsP quantum well semiconductor diode lasers", *Physica Status Solidi.B, Basic Solid State Physics*, vol. 211, pp. 513, 1999.

[34] I. Vurgaftman, J. R. Meyer and L. R. Ram-Mohan, "Band parameters for III–V compound semiconductors and their alloys", *J.Appl.Phys.*, vol. 89, pp. 5815, 2001.

[35] M. Gershenzon, D. G. Thomas and R. E. Dietz, *Proceedings of the International Conference on the Physics of Semiconductors, Exeter,* pp. 752, 1962.

[36] T. Liu, S. Chandril, A. J. Ptak, D. Korakakis and T. H. Myers, "Bismuth surfactant effects for GaAsN and beryllium doping of GaAsN and GaInAsN grown by molecular beam epitaxy", *J.Cryst.Growth*, vol. 304, pp. 402, 2007.

[37] S. Tixier, M. Adamcyk, E. C. Young, J. H. Schmid and T. Tiedje, "Surfactant enhanced growth of GaNAs and InGaNAs using bismuth", *J.Cryst.Growth*, vol. 251, pp. 449, 2003.

[38] X. Lu, D. A. Beaton, R. B. Lewis, T. Tiedje and M. B. Whitwick, "Effect of molecular beam epitaxy growth conditions on the bi content of $GaAs_{1-x}Bi_x$", *Appl. Phys. Lett.*, vol. 92, pp. 192110, 2008.

[39] L. Vegard, "The constitution of the mixed crystals and the space filling of the atoms", *Zeitschrift Für Physik A Hadrons and Nuclei*, vol. 5, pp. 17, 1921.

[40] J. Singh, *Electronic and optoelectronic properties of semiconductor structures*. Cambridge, U.K.; New York: Cambridge University Press, 2003.

[41] S. Wei and A. Zugner, "Giant and composition-dependent optical bowing coefficient in GaAsN alloys", *Physical Review Letters*, vol. 76, pp. 664, 1996.

[42] T. Mattila, S. Wei and A. Zunger, "Localization and anticrossing of electron levels in $GaAs_{1-x}N_x$alloys", *Physical Review.B*, vol. 60, pp. R11245, 1999.

[43] J. Wu, W. Shan and W. Walukiewicz, "Band anticrossing in highly mismatched III–V semiconductor alloys", *Semiconductor Science and Technology*, vol. 17, pp. 860, 2002.

[44] P. Anderson, "Localized magnetic states in metals", *Physical Review*, vol. 124, pp. 41, 1961.

[45] A. N. Kocharian, "Change of valence in rare earth semiconductors in the multi-impurity anderson model", *Soviet Physics, Solid State*, vol. 28, pp. 6, 1986.

[46] M. A. lvanov and Y. G. Pogorelov, "Electron spectrum of crystal in the presence of large radius impurity states", *Journal of Experimental and Theoretical Physics*, vol. 49, pp. 510, 1979.

[47] M. A. lvanov and Y. G. Pogorelov, "Electron properties of a system of two-parameter long-range impurity states", *Journal of Experimental and Theoretical Physics*, vol. 61, pp. 1033, 1985.

[48] J. Perkins, A. Mascarenhas, J. F. Geisz and D. J. Friedman, "Conduction-band-resonant nitrogen-induced levels in $GaAs_{1-x}N_x$with x<0.03", *Physical Review.B, Condensed Matter*, vol. 64, , 2001.

[49] J. Wu, W. Walukiewicz, K. M. Yu, J. W. Ager III, E. E. Haller, Y. G. Hong, H. P. Xin and C. W. Tu, "Band anticrossing in $GaP_{1-x}N_{x$}alloys", *Physical Review.B, Condensed Matter and Materials Physics*, vol. 65, , 2002.

[50] Chuang, S. L. *Physics of optoelectronic devices* New York: Wiley, 1995.

[51] K. M. Alberi. "Valence Band Anticrossing in Highly Mismatched Alloys." Ph.D., University of California, Berkeley, 2008.

[52] A. Janotti and Wei,S.,Zhang,S.B, "Theoretical study of the effects of isovalent coalloying of bi and N in GaAs", *Phys. Rev. B*, vol. 65, pp. 115203, 2002.

[53] X. Lu, D. A. Beaton, R. B. Lewis, T. Tiedje and Y. Zhang, "Composition dependence of photoluminescence of $GaAs_{1-x}Bi_x$alloys", *Appl. Phys. Lett.*, vol. 95, pp. 041903, 2009.

[54] S. Tixier, M. Adamcyk, T. Tiedje, S. Francoeur, A. Mascarenhas, P. Wei and F. Schiettekatte, "Molecular beam epitaxy growth of $GaAs_{1-x}Bi_x$", *Appl. Phys. Lett.*, vol. 82, pp. 2245, 2003.

[55] J. Yoshida, T. Kita, O. Wada and K. Oe, "Temperature dependence of $GaAs_{1-x}Bi_x$ Band gap studied by photoreflectance spectroscopy", *Japenese Journal of Applied Physics*, vol. 42, pp. 371, 2003.

[56] H. Okamoto and K. Oe, "Structural and energy-gap characterization of metalorganic-vapor-phase-epitaxy-grown InAsBi", *Japenese Journal of Applied Physics*, vol. 38, pp. 1022, 1998.

[57] P. Verma, K. Oe, M. Yamada, H. Harima, M. Herms and G. Irmer, "Raman studies on $GaAs_{1-x}Bi_x$and $InAs_{1-x}Bi_x$", *J.Appl.Phys.*, vol. 89, pp. 1657, 2001.

[58] S. P. Svensson, H. Hier, W. L. Sarney, D. Donetsky, D. Wang and G. Belenky, "Molecular beam epitaxy control and photoluminescence properties of InAsBi", *Journal of Vacuum Science & Technology B*, vol. 30, , 2011.

[59] C. A. Broderick, M. Usman, S. J. Sweeney and E. P. O'Reilly, "Band engineering in dilute nitride and bismide semiconductor lasers", *Semiconductor Science and Technology*, vol. 27, pp. 094011, 2012.

[60] T. Hosoda, G. Kipshidze, G. Tsvid, L. Shterengas and G. Belenky, "Type-I GaSb-based laser diodes operating in 3.1- to 3.3 μm wavelength range", *IEEE Photonics Technology Letters*, vol. 22, pp. 718, 2010.

[61] S. K. Dasa, T. D. Dasa, S. Dhara, M. d. l. Mareb and A. Krierb, "Infrared photoluminescence of diluteGaSb:Bi alloys grown by liquid phase epitaxy", *Journal of Crystal Growth*, 2012.

[62] G. Sai-Halasz, R. Tsu and L. Esaki, "A new semiconductor superlattice", *Appl. Phys. Lett.*, vol. 30, pp. 651, 1977.

[63] D. L. Smith and C. Mailhiot, "Proposal for strained type II superlattice infrared detectors", *J.Appl.Phys.*, vol. 62, pp. 2545, 1987.

[64] R. d. L. Kronig and W. L. Penney, "Quantum mechanics of electrons in crystal lattices", *Proceeding of the Royal Society of London*, vol. 130, pp. 499, 1930.

[65] D. Mukherji and B. R. Nag, "Band structure of semiconductor superlattices", *Phys. Rev. B.*, vol. 12, pp. 4338, 1975.

[66] H. -S. Cho and P. R. Prucnal, "New formalism of the kronig-penney model with applications to superlattices", *Phys. Rev. B.*, vol. 36, pp. 3237, 1987.

[67] G. Bastard, "Superlattice band structure in the envelope-function approximation", *Phys. Rev. B.*, vol. 24, pp. 5693, 1981.

[68] J. F. Klem, S. R. Kurtz and A. Datye, "Growth and properties of GaAsSb/InGaAs superlattices on InP", *J.Cryst.Growth*, vol. 111, pp. 628, 1991.

[69] M. Peter, R. Kiefer, F. Fuchs, N. Herres, K. Winkler, K. -H. Bachem and J. Wagner, "Light-emitting diodes and laser diodes based on a $Ga_{1-x}In_xAs/GaAs_{1-y}Sb_y$ type II superlattice on InP substrate", *Appl. Phys. Lett.*, vol. 74, pp. 1951, 1999.

[70] J. Y. T. Huang, L. J. Mawst, T. F. Kuech, X. Song, S. E. Babcock, C. S. Kim, I. Vurgaftman, J. R. Meyer and A. L. Holmes Jr, "Design and characterization of strained InGaAs/GaAsSb type-II 'W' quantum wells on InP substrates for mid-IR emission", *J.Phys.D*, vol. 42, pp. 025108, 2009.

[71] B. Chen, W. Jiang, J. Yuan, A. L. Holmes and B. M. Onat, "SWIR/MWIR InP-based p-i-n photodiodes with InGaAs/GaAsSb type-II quantum wells", *Quantum Electronics, IEEE Journal Of*, vol. 47, pp. 1244, 2011.

[72] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers: I. misfit dislocations", *J. Cryst. Growth*, vol. 27, pp. 118, 1974.

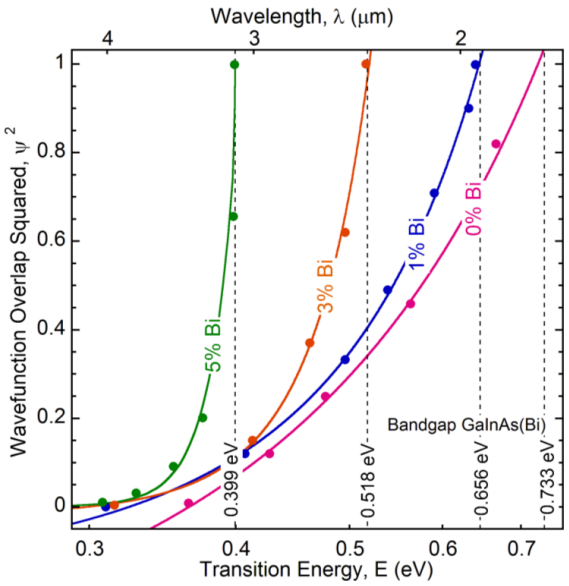
[73] N. Ekins-Daukes, K. Kawaguchi and J. Zhang, "Strain-balanced criteria for multiple quantum well structures and its signature in X-ray rocking curves", *Crystal Growth & Design*, vol. 2, pp. 287, 2002.

[74] M. W. Prairie and R. M. Kolbas, "A general derivation of the density of states function for quantum wells and superlattices", *Superlattices and Microstructures*, vol. 7, pp. 269, 1990.

APPENDIX A

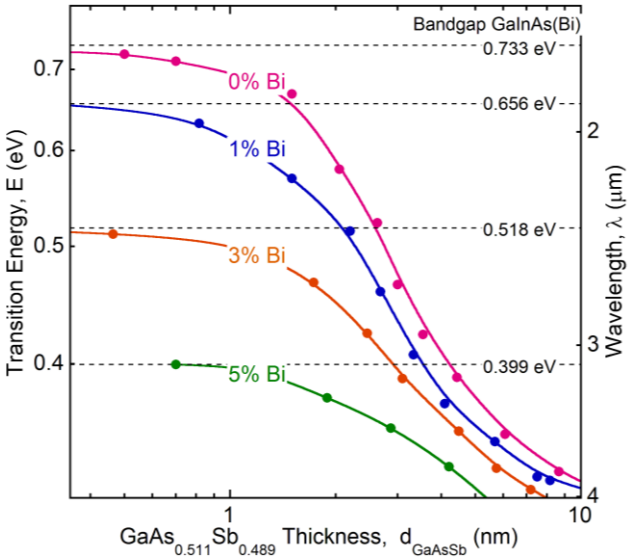
SUMMARY OF OPTIMAL DESIGN RULES FOR

$Ga_{1-y}In_yAs_{1-x}Bi_x$ /$GaAs_{1-x}Sb_x$ SUPERLATTICE STRUCTURES



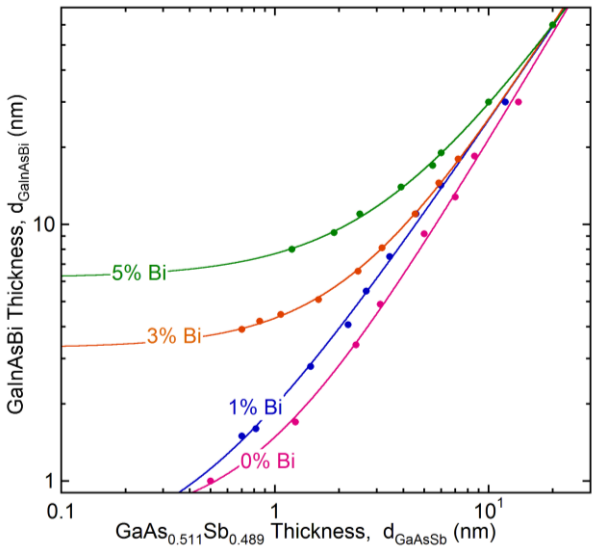
$$\Psi^2 = A_{\mathrm{Bi2}}\left(\exp\left(\frac{\mathrm{E}}{E_{\mathrm{Bi3}}}\right)\right) - A_{\mathrm{Bi1}}$$

Bi (%)	A_{Bi1}	A_{Bi2}	E_{Bi3} (eV)
0	0.9957	0.4611	0.5008
1	0.2346	0.0436	0.1928
3	0.0310	$1.916 \cdot 10^{-4}$	0.0605
5	$1.243 \cdot 10^{-5}$	$5.128 \cdot 10^{-9}$	0.0214



$$E = E_{\mathrm{Bi1}} + E_{\mathrm{Bi2}}\left[exp\left(\frac{-d_A}{d_{\mathrm{Bi1}}}\right)\right]$$

Bi (%)	E_{Bi1} (eV)	E_{Bi2} (eV)	d_{Bi1} (nm)
0	0.3013	0.1783	2.902
1	0.3007	0.2928	2.525
3	0.2886	0.4582	2.999
5	0.2803	0.4929	2.937

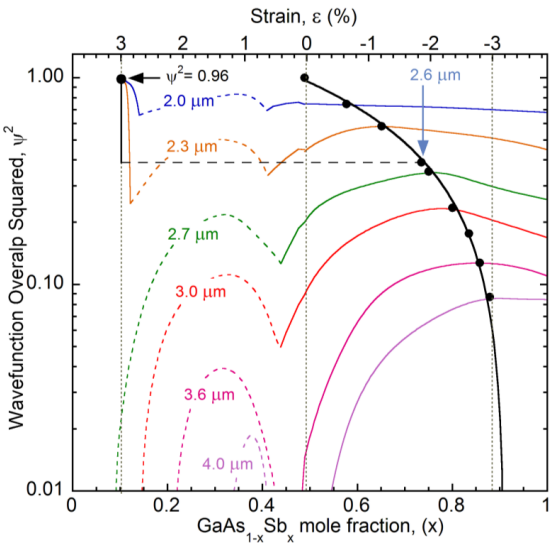


$$d_B = d_{\mathrm{Bi2}}\left(1 + \left(\frac{d_A}{d_{\mathrm{Bi3}}}\right)^{p_{\mathrm{Bi}}}\right)$$

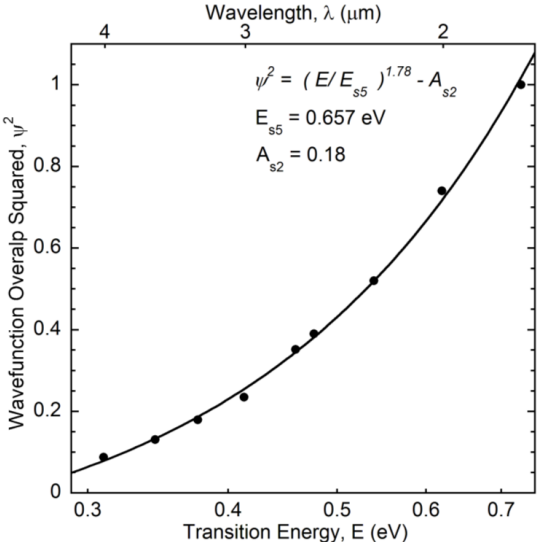
Bi (%)	**d_{Bi2} (nm)**	**d_{Bi3} (nm)**	**p_{Bi}**
0	0.8718	0.6707	1.4078
1	0.4092	0.4903	1.2310
3	2.4121	3.3063	1.3483
5	3.2855	6.1989	1.1962

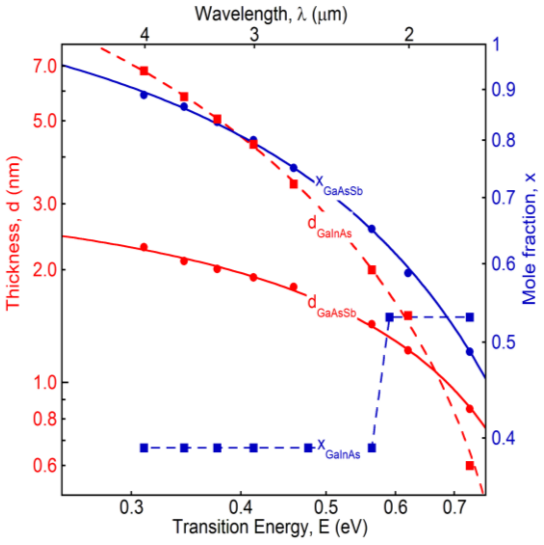
APPENDIX B

SUMMARY OF OPTIMAL DESIGN RULES FOR STRAIN BALANCED $Ga_{1-y}In_yAs$ /$GaAs_{1-x}Sb_x$ SUPERLATTICE STRUCTURES



$$\Psi^2 = 2.10 - 2.33x, \qquad \text{for } x \geq 0.489$$





GaInAs thickness versus Transition energy	GaAsSb thickness versus Transition energy	Mole fraction versus Transition energy	Squared overlap versus transition energy
$d_A = d_{s1}\left(exp\left(\frac{E_{s1}-E}{E_{s2}}\right)-1\right)$	$d_B = d_{s2}\left(1-\frac{E}{E_{s3}}\right)$	$x_B = A_{s1}-\frac{E}{E_{s4}}$	$\boldsymbol{\Psi^2 = \left(\frac{E}{E_{s5}}\right)^{p_s} - A_{s2}}$
d_{s1} = 1.348 nm E_{s1} = 0.839 eV E_{s2} = 0.296 eV	d_{s2} = 3.339 nm E_{s3} = 0.973 eV	A_{s1} = 1.195 E_{s4} = 1.032 eV	E_{s5} = 0.657 eV A_{s2} = 0.18 p_s = 1.78

APPENDIX C

MATERIAL PARAMETERS USED IN VALENCE BAND ANTICROSSING OF $GaAs_{1-x}Bi_x$, $InAs_{1-x}Bi_x$ and $GaSb_{1-x}Bi_x$

Parameter	Value	Unit	Reference
Luttinger Parameters			
InAs			
γ_1	20.0		
γ_2	8.5		[34]
γ_3	9.2		
GaSb			
γ_1	13.4		
γ_2	4.7		[34]
γ_3	6.0		
GaAs			
γ_1	6.98		
γ_2	2.06		[34]
γ_3	2.93		
Lattice Constants			
InAs	$6.0583 + 2.74\cdot10^{-5}\cdot(T-300)$	Å	[34]
GaSb	$6.0959 + 4.72\cdot10^{-5}\cdot(T-300)$	Å	[34]
GaAs	$5.6533 + 388\cdot10^{-5}\cdot(T-300)$	Å	[34]
InBi	6.686	Å	[52]
GaBi	6.324	Å	[52]
Bandgaps			
InAs	$0.417 - 0.276\cdot10^{-3}\cdot T^2/(T+93)$	eV	[34]
GaSb	$0.813 - 3.78\cdot10^{-3}\cdot T^2/(T+94)$	eV	[34]
GaAs	$1.519 + 5.405\cdot10^{-4}\cdot T^2/T+204)$	eV	[34]
InBi	-1.63	eV	[52]
GaBi	-1.45	eV	[52]
Spin-orbit Energies			
InAs	0.39	eV	[34]
GaSb	0.76	eV	[34]
GaAs	0.34	eV	[34]
InBi	2.27	eV	
GaBi	2.20	eV	[31]
Valence Band Offsets			
InAs	-0.59	eV	[34]
GaSb	-0.03	eV	[34]
GaAs	-0.80	eV	[34]
InBi	0.60	eV	

GaBi	0.00	eV	[31]
Coupling parameter			
$InAs_{1-x}Bix$	1.55	eV	
$GaSb_{1-x}Bi_x$	1.05	eV	
$GaAs_{1-x}Bi_x$	1.55	eV	[31]
Impurity Level of Bi			
InAs	-0.53	eV	
GaSb	-1.17	eV	
GaAs	-0.40	eV	[31]
Spin orbit energy for Bi			
InAs	-2.11	eV	
GaSb	-2.67	eV	
GaAs	-1.90	eV	[31]

APPNEDIX D

MATERIAL PARAMETERS USED TO CALCULATE THE $Ga_{1-y}In_yAs_{1-x}Bi_x$ /$GaAs_{1-x}Sb_x$ AND $Ga_{1-y}In_yAs/GaAs_{1-x}Sb_x$ SUPERLATTICE BANDGAPS

Parameter	Value	Unit	Reference
InAs	$6.0583 + 2.74 \cdot 10^{-5} \cdot (T-300)$	Å	[34]
GaSb	$6.0959 + 4.72 \cdot 10^{-5} \cdot (T-300)$	Å	[34]
GaAs	$5.6533 + 388 \cdot 10^{-5} \cdot (T-300)$	Å	[34]
$Ga_{1-y}In_yAs$	$a_{GaInAs} = (1-y) \cdot a_{GaAs} + y \cdot a_{InAs}$	Å	[34]
$Ga_{1-y}In_yAs_{1-x}Bi_x$	$a_{GaInAsBi} = (1-x) \cdot a_{GaInAs} + x \cdot a_{GaAsBi}$	Å	
$GaAs_{1-x}Sb_x$	$a_{GaAsSb} = (1-x) \cdot a_{GaAs} + x \cdot a_{GaSb}$	Å	[34]
Bandgaps			
InAs	$0.417 - 0.276 \cdot 10^{-3} \cdot T^2/(T+93)$	eV	[34]
GaSb	$0.813 - 3.78 \cdot 10^{-3} \cdot T^2/(T+94)$	eV	[34]
GaAs	$1.519 + 5.405 \cdot 10^{-4} \cdot T^2/T+204)$	eV	[34]
GaInAs Bowing Factor (C_{Eg_GaInAs})	0.477	eV	[34]
$Ga_{1-y}In_yAs$	$(1-y) \cdot E_{g\_GaAs} + y \cdot E_{g\_InAs} - y \cdot (1-y) \cdot C_{Eg\_GaInAs}$	eV	[34]
$Ga_{1-y}In_yAs_{1-x}Bi_x$	$(1-x) \cdot E_{g\_GaInAs} + x \cdot E_{g\_GaAsBi}$	eV	
GaAsSb Bowing Factor (C_{Eg_GaAsSb})	1.43	eV	[34]
$GaAs_{1-x}Sb_x$	$(1-x) \cdot E_{g\_GaAs} + x \cdot E_{g\_GaSb} - x \cdot (1-x) \cdot C_{Eg\_GaAsSb}$	eV	[34]
Spin-orbit Energies			
InAs	0.39	eV	[34]
GaSb	0.76	eV	[34]
GaAs	0.34	eV	[34]
GaInAs Bowing Factor (C_{SO_GaInAs})	0.15	eV	[34]
$Ga_{1-y}In_yAs$	$(1-y) \cdot \Delta_{SO\_GaAs} + y \cdot \Delta_{SO\_InAs} - y \cdot (1-y) \cdot C_{SO\_GaInAs}$	eV	[34]
$Ga_{1-y}In_yAs_{1-x}Bi_x$	$(1-x) \cdot \Delta_{SO\_GaInAs} + x \cdot \Delta_{SO\_GaAsBi}$	eV	
GaAsSb Bowing Factor (C_{SO_GaAsSb})	0.60	eV	[34]
$GaAs_{1-x}Sb_x$	$(1-x) \cdot \Delta_{SO\_GaAs} + x \cdot \Delta_{SO\_GaSb} - x \cdot (1-x) \cdot C_{SO\_GaAsSb}$	eV	[34]
Valence Band			

Offsets			
InAs	-0.59	eV	[34]
GaSb	-0.03	eV	[34]
GaAs	-0.80	eV	[34]
GaInAs Bowing Factor (C_{Ev_GaInAs})	-0.38	eV	[34]
$Ga_{1-y}In_yAs$	$(1-y)\cdot VBO_{GaAs} + y\cdot VBO_{InAs} - y\cdot(1-y)\cdot C_{Ev\_GaInAs}$	eV	[34]
$Ga_{1-y}In_yAs_{1-x}Bi_x$	$(1-x)\cdot VBO_{GaInAs} + x\cdot VBO_{GaAsBi}$	eV	
GaAsSb Bowing Factor (C_{Ev_GaAsSb})	-1.06	eV	[34]
$GaAs_{1-x}Sb_x$	$(1-x)\cdot VBO_{GaAs} + x\cdot VBO_{GaSb} - x\cdot(1-x)\cdot C_{Ev\_GaAsSb}$	eV	[34]
Effective Masses			
InAs			
Electron	0.023		
Light Hole	$1/(\gamma_{1\_InAs} + 2\cdot\gamma_{2\_InAs})$	m_o	[34]
Heavy Hole	$1/(\gamma_{1\_InAs} - 2\cdot\gamma_{2\_InAs})$		
GaSb			
Electron	0.04		
Light Hole	$1/(\gamma_{1\_GaSb} + 2\cdot\gamma_{2\_GaSb})$	m_o	[34]
Heavy Hole	$1/(\gamma_{1\_GaSb} - 2\cdot\gamma_{2\_GaSb})$		
GaAs			
Electron	0.067		
Light Hole	$1/(\gamma_{1\_GaAs} + 2\cdot\gamma_{2\_GaAs})$	m_o	[34]
Heavy Hole	$1/(\gamma_{1\_GaAs} - 2\cdot\gamma_{2\_GaAs})$		
$Ga_{1-y}In_yAs$			
GaInAs Bowing Factor (C_{me_GaInAs})	0.003		[34]
Electron	$(1-y)\cdot m_{e\_GaAs} + y\cdot m_{e\_InAs} - y\cdot(1-y)\cdot C_{me\_GaInAs}$		
Light Hole	$(1-y)\cdot m_{lh\_GaAs} + y\cdot m_{lh\_InAs}$	m_o	[34]
Heavy Hole	$(1-y)\cdot m_{hh\_GaAs} + y\cdot m_{hh\_InAs}$		
$Ga_{1-y}In_yAs_{1-x}Bi_x$			
Electron	$(1-x)\cdot m_{e\_GaInAs} + x\cdot 0.099$		

Light Hole	$(1-x)\cdot m_{lh\_GaInAs} + x\cdot 0.082$	m_o	
Heavy Hole	$(1-x)\cdot m_{hh\_GaInAs} + x\cdot 0.51$		
$GaAs_{1-x}Sb_x$			
GaAsSb Bowing Factor (C_{me_GaAsSb})	0.0258		[34]
Electron	$(1-x)\cdot m_{e\_GaAs} + x\cdot m_{e\_GaSb} - x\cdot(1-x)\cdot C_{me\_GaAsSb}$		
Light Hole	$(1-x)\cdot m_{lh\_GaAs} + x\cdot m_{lh\_GaSb}$	m_o	[34]
Heavy Hole	$(1-x)\cdot m_{hh\_GaAs} + x\cdot m_{hh\_GaSb}$		
Elastic Constants			
InAs			
C_{11}	832.9	GPa	
C_{12}	452.6	GPa	
b	-1.8	eV	[34]
a_c	-5.08	eV	
a_v	-1	eV	
GaSb			
C_{11}	884.2	GPa	
C_{12}	402.6	GPa	
b	-2	eV	[34]
a_c	-7.5	eV	
a_v	-0.8	eV	
GaAs			
C_{11}	1221	GPa	
C_{12}	566	GPa	
b	-2	eV	[34]
a_c	-7.17	eV	
a_v	-1.16	eV	
$Ga_{1-y}In_yAs$			
C_{11}	$(1-y)\cdot C_{11\_GaAs} + y\cdot C_{11\_InAs}$	GPa	
C_{12}	$(1-y)\cdot C_{12\_GaAs} + y\cdot C_{12\_InAs}$	GPa	
b	$(1-y)\cdot b_{GaAs} + y\cdot b_{InAs}$	eV	
GaInAs Bowing Factor (C_{ac_GaInAs})	2.61	eV	[34]

a_c	$(1-y)\cdot a_{c\_GaAs} + y\cdot a_{c\_InAs} - y\cdot(1-y)\cdot C_{ac\_GaInAs}$	eV	
a_v	$(1-y)\cdot a_{v\_GaAs} + y\cdot a_{v\_InAs}$	eV	
$GaAs_{1-x}Sb_x$			
C_{11}	$(1-x)\cdot C_{11\_GaAs} + x\cdot C_{11\_GaSb}$	GPa	
C_{12}	$(1-x)\cdot C_{12\_GaAs} + x\cdot C_{12\_GaSb}$	GPa	
b	$(1-x)\cdot b_{GaAs} + x\cdot b_{GaSb}$	eV	[34]
a_c	$(1-x)\cdot a_{c\_GaAs} + x\cdot a_{c\_GaSb}$	eV	
a_v	$(1-x)\cdot a_{v\_GaAs} + x\cdot a_{v\_GaSb}$	eV	