DEFECT CHARACTERIZATION OF ANTIMONIDE-BASED TYPE-II
SUPERLATTICES FOR INFRARED DETECTION

BY

DANIEL YUAN ZUO

DISSERTATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:
Associate Professor Daniel M. Wasserman, Chair
Assistant Professor Can Bayram
Professor Jianming Jin
Eric A. Shaner, Sandia National Laboratories
ABSTRACT

Type II superlattices (T2SLs) have undergone a significant amount of development and progress since their initial proposed use for the detection of mid- to long-wavelength infrared (MWIR, LWIR) radiation. The type-II broken band gap alignment of these heterostructures allows electrons and holes to be confined in separate but adjacent, nanometer scale thickness, material layers. When arranged in a highly periodic structure comprised of thin layers of alternating materials, known as a superlattice (SL), the carriers form minibands that have an effective band gap lower than that of the constituent materials and suited for MWIR and LWIR applications. Traditionally, InAs/GaSb has been the preferred material system for T2SL IR detectors, but recent work has also explored the use of InAs/InAsSb as well. This thesis explores the characterization of carrier transport properties in T2SL devices through a variety of techniques, though primarily with the application of electron beam induced current (EBIC) measurements. For these EBIC measurements, using first-principle analysis of the low-injection diffusion of carriers generated by inelastic scattering of high-energy electrons, we develop theoretical models to characterize minority carrier diffusion length and surface recombination velocity in photodetectors along the growth direction. When applied to photodiodes with strain management layers incorporated in the InAs/GaSb SL, we used EBIC results to directly observe the improvement in carrier diffusion length and lifetime due to strain management. Furthermore, novel EBIC analysis of an nBn photodetector with InAs/InAsSb regions combined with lifetime measurements through time-resolved photoluminescence (TRPL) were demonstrated, forming a comprehensive characterization of the vertical transport
qualities, including mobility and diffusivity. Additionally, we propose a sample design incorporating a T2SL in a bulk p-n diode for the application of deep level transient spectroscopy (DLTS) to measure defect energy levels within the SL band gap. Contained in this work are our initial results in testing our first sample designs. Defects specific to the T2SL have yet to be observed, but we were able to measure shallow states in bulk GaSb.
ACKNOWLEDGMENTS

This dissertation could not have been written without the help of many others. First and foremost, I would like to thank my former advisor, Professor Shun Lien Chuang. Without his guidance, leadership, support, and encouragement I would not be the person I am today, let alone a doctoral candidate. I hope to one day set a similar example for others in the way that he did for me and many other students. I also wish to thank my current advisor, Professor Daniel Wasserman, whose patience and guidance allowed me to continue the work I started and complete the dissertation. Additionally, thanks to Professors Can Bayram, Jianming Jin, Xiuling Li, and Dr. Eric Shaner for serving on my committee.

Many fellow students and friends have lent their insight, expertise, and guidance over the years. Thanks to Benjamin Kesler, Runyu Liu, Chien-Yao Lu, Martin Mandl, Akira Matsudaira, Thomas O’Brien, Pengfei Qiao, Guan-Lin Su, and Lan Yu for their various support, collaboration, and intellectually stimulating discussions.

Finally, I owe much of my intellectual curiosity and motivation to my parents. Their constant support has encouraged me to challenge myself in my work and take satisfaction in my accomplishments.
TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION ................................................................................................. 1

History and Motivation .......................................................................................................... 1

Semiconductor Physics ........................................................................................................... 4

Equilibrium and Non-Equilibrium Carrier Dynamics ............................................................ 7

Recombination Mechanisms .................................................................................................. 11

Radiative Recombination ...................................................................................................... 12

Shockley-Read-Hall Recombination ....................................................................................... 13

Auger Recombination ............................................................................................................ 15

Light Detection via Stimulated Optical Absorption ................................................................. 16

Strain Management .............................................................................................................. 19

State-of-the-art ..................................................................................................................... 21

Outline of Thesis .................................................................................................................. 24

CHAPTER 2 ELECTRON BEAM INDUCED CURRENT ...................................................... 26

Introduction .......................................................................................................................... 26

EBIC Background ............................................................................................................... 28

EBIC Imaging of Defects ...................................................................................................... 29

Theory of EBIC Junction Study ............................................................................................. 33

Generation Volume .............................................................................................................. 34
Collection Probability ................................................................. 36
Experimental Procedure............................................................. 40
Extraction of Diffusion Length and Surface Recombination .............. 42

CHAPTER 3 INTERFACIAL TREATMENT IN TYPE-II SUPERLATTICE
PHOTODETECTORS ...................................................................... 47
Introduction.................................................................................. 47
Band Structure Simulation............................................................ 49
Fourier Transform Infrared Spectroscopy ........................................ 51
Experimental Results .................................................................... 55
Modeling of EBIC in Photodiodes with Long Absorber Regions .......... 57

CHAPTER 4 COMPREHENSIVE VERTICAL TRANSPORT
CHARACTERIZATION OF NBN TYPE-II SUPERLATTICE DETECTORS ...... 62
Introduction.................................................................................. 62
nBn Photodetectors ....................................................................... 63
Time-Resolved Photoluminescence ................................................. 65
Optical Characterization of InAs/InAsSb nBn Photodetectors .......... 66
Minority Carrier Lifetime Characterization....................................... 69
EBIC Characterization of nBn Detectors........................................... 71
Results and Discussion ................................................................ 75
CHAPTER 1

INTRODUCTION

History and Motivation

Photoconductive devices play an important role in a host of optical systems for optical communication, imaging, and sensing, with varying applications including telecommunications, astronomy, and military. Some light lies within the visible spectrum seen by the human eye, and can be detected using devices with active regions formed with bulk semiconductors. For such applications, the light has high enough energy that it can be absorbed by Si and group III-V compound semiconductors. Light with wavelengths longer than visible light, spanning from 700 nm to as much as 1 mm, is described as being in the infrared (IR) regime. There are several fundamental reasons for the significance of infrared light. For example, all objects emit radiation innately, the nature of which is described by Planck’s law of radiation and formulated as

\[
I(\lambda, T) = \frac{2hc^2}{\lambda^5} \exp\left(\frac{hc}{\lambda kT}\right) - 1,
\]

where \( h = 6.626 \times 10^{-34} \) J·s is Planck’s constant, \( c = 3 \times 10^{10} \) cm/s is the speed of light, \( k = 8.617 \times 10^{-5} \) eV/K or \( 1.381 \times 10^{-23} \) J/K is Boltzmann’s constant, and \( T \) is the temperature of the object in question, in Kelvin. \( I(\lambda, T) \) gives the energy radiated at a given wavelength for an object at a given temperature in units of W/sr/cm\(^3\), and \( I(\lambda, T)d\lambda d\Omega \) describes the power per unit surface area delivered to an infinitesimally small solid
angle $d\Omega$ by an infinitesimally small selection of wavelengths centered on $\lambda$. Figure 1 plots $I(\lambda, T)$ for a selection of various temperatures.

Based on this plot it can be seen that for objects at room temperature the peak of this radiation occurs close to 10 $\mu$m. An object at 200 K (dotted line) emits primarily in the LWIR, useful for relatively cool stellar bodies for example. While most stars have temperatures of several thousand Kelvin, and thus peak wavelength thermal emission at much shorter wavelengths that can be seen with the naked eye, infrared imaging equipment suspended above the Earth’s atmosphere (to avoid infrared absorption due to water and other elements present) can be used to observe a variety of additional interstellar phenomena. At the typical human body temperature (310 K, dashed line) the majority of the radiance is close to 10 $\mu$m. Detectors that are made to target this range can thus be used to image people in conditions that are unfavorable to visible light (extreme weather, cloud cover, and nighttime activities, for instance) and are thus highly valuable in military and
defense applications. Warmer bodies, such as those at 400 K, still emit primarily in the infrared and this trend can be used in infrared systems to determine the temperature of various objects via imaging, among other applications.

Many molecules absorb light within narrow bands of wavelength and convert the energy into vibrational movement based on the energy of the light and the structure of the molecule. These modes of vibration are determined by the particular interactions within a molecule, which depend on the masses of the elements and the strength of the bonds holding them together. The aggregate modes, called resonant frequencies, are thus unique to the molecules that produce them. When the molecules are present in a significant concentration, they will absorb light passing through with those frequencies. Thus an observation of these missing bands can identify the composition of the environment under study. Figure 2 shows a selection of gases and their spectra in the infrared, which can be

![Figure 2: Spectral signatures for various gases in the infrared spectrum.](image)
used as “fingerprints” to identify their presence in sensing measurements using IR radiation [1]. Based on these attributes, infrared detectors can be applied to a wide variety of scientific and industrial applications to analyze the chemical makeup and concentration of many environments.

Infrared detection is rooted in early efforts by Herschel in 1800 discovering the presence of invisible radiation beyond red light, finding that a thermometer held past the red end of light exiting a prism would read hotter than when held in the visible portion. In 1821 Seebeck produced the first thermocouple, which uses voltages generated by temperature differences in dissimilar metals. The bolometer, which measures incoming radiation by sensing the heating of a broadband light-absorbing element connected to a thermal reservoir by a weak thermal conductor, was later produced by Langley in 1878. However, it is the development of semiconductor devices which has spurred the greatest leaps in photodetector technology over the last decades.

**Semiconductor Physics**

An atom can be described in classical terms as a set of negatively charged electrons orbiting a positively charged nucleus of proton and neutrons. In quantum mechanics, it is understood [2, p. 30] that the electrons are bound to the nucleus and can only exist at specific energy levels, in contrast to a classical particle experiencing attractive force which can possess an arbitrary kinetic and potential energy. The possible energies in the semiconductor can be determined via Schrödinger’s equation:

$$\frac{i\hbar}{\hbar} \frac{\partial \psi(r, t)}{\partial t} = \left( -\frac{\hbar^2}{2m} \nabla^2 + V(r) \right) \psi(r, t),$$

(2)
where $\hbar$ is the Planck constant ($\hbar = 6.64 \times 10^{-34} J \cdot s$) divided by $2\pi$, $m$ is the mass of the particle, $\psi(r, t)$ is the position- and time-dependent wave function of the electron, and $V(r)$ is the potential energy of the system. The wave function describes the state of a particle and its evolution over time, and the other terms in the right side of Equation (2) comprise the “energy operator,” known as the Hamiltonian and typically abbreviated as $H$.

This nomenclature becomes intuitive with the time-independent Schrödinger equation:

$$H\psi(r) = E\psi(r), \quad (3)$$

where $E$ is the total energy of the particle described by $\psi$ at a fixed point in time.

When two atoms are brought together, the bound states interact with each other and create two closely separated energy levels. This splitting of states compounds when more and more atoms are added to the system and interact with each other. In a semiconductor, the vast amount of atoms (~$10^{23}$ in a cubic centimeter) effectively turns the discrete states into continuous energy bands that allow electrons to travel through the lattice with varying momentum and energy. The arrangement of the atoms in a semiconductor forms a periodic lattice, such that for many given vectors $R$, $V(r + R) = V(r)$. The result of this, known as Bloch’s theorem, is that the wave function too must obey the same periodic relationship.

The general solution to Equation (3) with this condition is

$$\psi_{nk}(r) = e^{ikr}u_{nk}(r), \quad (4)$$

where

$$u_{nk}(r + R) = u_{nk}(r). \quad (5)$$
The resulting energies that correspond to each wave function are $E_n(k)$, and $k$ is known as the wave vector. The integer $n$ designates separate energy bands. Lower energy bands, having wave functions that are highly localized to the atom, can be thought of as containing the ‘bound’ electrons in the lattice. Typically, it is the higher energy bands (denoted as the valence and conduction bands) which interact with incident photons. As can be seen in Equation (4), the wave vector determines the direction in which the electron travels. Figure 3 shows a basic depiction of the energy versus the wave vector, also known as the dispersion relation, of several (higher energy) bands near $k = 0$. The upper band with a positive dispersion is known as the conduction band, and is separated from the lower bands which have negative dispersion. These lower, negative dispersion bands are the light
hole, heavy hole, and spin-orbital bands. Near the Brillouin zone center \( \mathbf{k} = 0 \), the curvature of \( E(\mathbf{k}) \) can be approximated as a parabolic curve of the form

\[
E(\mathbf{k}) \approx E_0 + \frac{\hbar^2 \mathbf{k}^2}{2m_{e,h}^*},
\]

where \( m_{e,h}^* \) is the particle’s effective mass within that region (\( e \) for electrons, \( h \) for holes), and \( E_0 \) is the minimum energy of the band. The effective mass is typically expressed as a factor of the rest mass of an electron, which is given as \( m_e = 9.11 \times 10^{-31} \) kg.

In an intrinsic semiconductor, the conduction band is mostly unoccupied and contains few electrons. Conversely, the valence bands are highly populated, being of lower energy, and contain a few “holes,” the absence of an electron that can be viewed as its own discrete particle with an opposite charge (and therefore opposite energy gradient). The minimum separation between the conduction and valence bands is known as the band gap, and denoted by \( E_g \). Electrons and holes are able to traverse the bands, and can even make transitions between them, but in every case energy and momentum (which is directly related to the wave vector in a Bloch wave function) must be conserved.

**Equilibrium and Non-Equilibrium Carrier Dynamics**

A semiconductor that is not subject to external forces and stimuli will maintain a steady concentration of free carriers. This equilibrium concentration is the result of a natural balance between generation and recombination of electron-hole pairs. The concentrations of electrons and holes are expressed via Fermi-Dirac statistics as

\[
n_0 = N_C e^{-\frac{E_C - E_F}{k_B T}} \quad \text{for electrons, and}
\]

\[
n_h = N_V e^{-\frac{E_V - E_F}{k_B T}} \quad \text{for holes,}
\]

where \( E_C, E_V \) are the conduction and valence band edges, \( E_F \) is the Fermi energy, \( k_B \) is Boltzmann’s constant, and \( T \) is the temperature.
\[ p_0 = N_\nu e^{-\frac{E_F-E_V}{k_BT}} \text{ for holes,} \]  
\[ N_{C,V} = 2 \left( \frac{m^*_e h k_B T}{2\pi \hbar^2} \right)^{\frac{3}{2}}, \]

where \( N_C \) and \( N_\nu \) are the conduction and valence band densities of states [3, p. 35], respectively, \( E_C \) and \( E_\nu \) represent the conduction and valence band edge minimum and maximum, \( k_B \) denotes Boltzmann’s constant, and \( T \) is the temperature of the semiconductor in Kelvin. \( E_F \) is referred to as the Fermi level, and is equivalent to the chemical potential in statistical mechanics at absolute zero: electrons cannot have an energy above \( E_F \) at 0 K, but with increasing temperature the threshold is not absolute and some portion of them have a probability of existing above it with a decreasing probability as a function of distance from \( E_F \).potential in statistical mechanics. As an auxiliary relation,

\[ n_0 p_0 = n_i^2 = N_C N_\nu e^{\frac{E_g}{k_BT}}. \]

The term \( n_i \) is called the intrinsic carrier concentration. Note that its only dependencies aside from temperature are properties intrinsic to the semiconductor.

When the semiconductor is subject to external excitation that results in excess electron and hole concentrations, recombination processes will work to bring the electron and hole levels back to equilibrium values. Moreover, if the non-equilibrium values are not uniformly distributed within the semiconductor, they will experience flux along the directions of variance, a phenomenon known as diffusion. Free carriers are also subject to Lorentz force \( \mathbf{F} = q\mathbf{E} \) from any electric fields that are present due to a net space charge.
distribution. The redistribution of charge due to these mechanisms is totaled to express net current densities for electrons and holes:

\[ J_n = q(n \mu_e E + D_e \nabla n) \]  \hspace{1cm} (11)

\[ J_p = q(p \mu_h E - D_h \nabla p), \]  \hspace{1cm} (12)

where \( q = 1.602 \times 10^{-19} \text{ C} \) is the elementary charge, \( n \) and \( p \) are the total carrier densities for electrons and holes, \( \mu_e \) and \( \mu_h \) are the carrier mobility coefficients, and \( D_e \) and \( D_h \) are the carrier diffusivity coefficients. The carrier mobilities relate to how fast they are transported by an electric field, with units of area per volt per second. The diffusivities describe the ease at which they spread from areas with a concentration gradient with units of area per second. The current densities are expressed with units of amperes per unit area.

The mobility coefficient is used in the Drude model, a classical view of the carriers as particles scattering as a result of interactions with vibrations in the lattice, impurities, and other electrons at regular intervals. In this system, the mobility is derived as

\[ \mu = \frac{q \tau_{coll}}{m}, \]  \hspace{1cm} (13)

where \( \tau_{coll} \) is the mean time between collision events and \( m \) is the effective mass of the particle. As long as the carrier concentrations do not exceed degenerate levels (when the Fermi level is within \( k_B T \) of, or above, the band edge), the mobility is linked to the diffusivity by Einstein’s relation:

\[ D = \frac{\mu k_B T}{q}. \]  \hspace{1cm} (14)

The overall change in excess carrier concentration numbers can be found by combining the net generation and recombination rates with the carrier flux via Equations (11) and (12). In
a closed region, the net change in carrier concentration is given by the rate at which carriers are generated in the sample minus the rate at which carriers recombine or flow out of the region. This continuity must apply at all times. For electrons, the differential equation is

$$\frac{d\delta n}{dt} = G_n - R_n + \frac{1}{q} \nabla \cdot J_n,$$

where $G_n$ represents the net generation rate per second, $R_n$ represents the net recombination rate per second, and $\delta n$ is the excess electron concentration per unit volume. Similarly for holes,

$$\frac{d\delta p}{dt} = G_p - R_p - \frac{1}{q} \nabla \cdot J_p.$$

It is useful to express the net recombination rate $R - G$ in terms of a time constant called the recombination lifetime $\tau$. For electrons

$$R_n - G_n = \frac{\delta n}{\tau_n},$$

and the relationship is analogous for holes. For regions with negligible electric field, Equation (15) becomes

$$\frac{d\delta n}{dt} = -\frac{\delta n}{\tau_n} + D_n \nabla^2 \delta n,$$

where $D_n$ denotes the diffusivity of electrons, and for a steady-state condition the formula becomes

$$\nabla^2 \delta n = \frac{\delta n}{D_n \tau_n},$$

$$L_n = \sqrt{D_n \tau_n}.$$
where $L_n$ is the electron diffusion length, with an analogous $L_p$ existing for holes in the same manner. The diffusion length is a measure of how far the carrier is expected to travel before recombining.

**Recombination Mechanisms**

The lifetime for a given carrier (electrons/holes) can be expressed as a sum of multiple distinct recombination mechanisms:

$$\frac{1}{\tau} = \frac{1}{\tau_{\text{rad}}} + \frac{1}{\tau_{\text{SRH}}} + \frac{1}{\tau_{\text{Auger}}},$$  \hspace{1cm} (21)

where $\tau_{\text{rad}}$ is the radiative recombination lifetime, $\tau_{\text{SRH}}$ is the Shockley-Read-Hall lifetime for trap-assisted recombination, and $\tau_{\text{Auger}}$ is the lifetime due to Auger recombination. Similarly, the overall recombination rate can be described as a sum of individual recombination rates:

$$R = R_{\text{rad}} + R_{\text{SRH}} + R_{\text{Auger}}.$$  \hspace{1cm} (22)
All three processes are diagrammed in Figure 4. An overview of each mechanism follows.

**Radiative Recombination**

In radiative recombination (Figure 4a), an electron in the conduction band recombines with a hole in the valence band. The energy lost by the electron is converted to a photon of equivalent wavelength \( \lambda = h c / E_g \) if the transition is between band minima) which is then emitted. This process can be observed with an external detection method and is commonly used to characterize the band transitions and/or the carrier lifetime, as is discussed in Chapter 4. The radiative recombination rate is given \([4]\) as

\[
R_{rad} = B (np - n_i^2),
\]

where \( B \) is the coefficient of radiative recombination. If the carrier concentrations are expressed in terms of the equilibrium values \( n = n_0 + \delta n \) and \( p = p_0 + \delta p \), then using Equation (17) to convert to lifetime gives
\[ \tau_{rad} = \frac{1}{B(n_0 + p_0 + \delta n)}. \]  \hspace{1cm} (24)

Note that \( \delta n = \delta p \) in the case of transitions between the valence and conduction bands, which always create a hole in the valence band and an electron in the conduction band.

**Shockley-Read-Hall Recombination**

Shockley-Read-Hall (SRH) recombination occurs in semiconductors which contain defects and impurities in the lattice structure. These can introduce localized energy states within the semiconductor band gap. In SRH recombination, the electron is captured by a local trap state. From this intermediate energy, it can then recombine in the valence band to complete the process, leaving the “trap” state open for another capture (see Figure 4b). The excess energy is emitted via photons or phonons (vibrational wave packets within the semiconductor lattice). Since the recombination process relies on the electron in the trap state recombining with a hole in the valence band before being remitted to the conduction band via thermal energy, trap states that are closer to the middle of the band gap are more effective recombination centers.

The SRH lifetime for holes in a majority n-type semiconductor where \( n_0 \gg p_0 \) [5] is completely dominated by the rate at which holes are captured, and is given as

\[ \tau_{p,SRH} = \frac{1}{v_p \sigma_p N_T}. \]  \hspace{1cm} (25)

where \( v_p \) is the thermal velocity of holes and \( \sigma_p \) is the capture cross section for holes and \( N_T \) is the trap concentration. Similarly, the SRH lifetime of electrons in a p-type semiconductor is
\[ \tau_{n,SRH} = \frac{1}{v_n \sigma_n N_T}, \]  
\[ (26) \]

where \( v_n \) is the thermal velocity of electrons, \( \sigma_n \) is the capture cross section for electrons of the trap states. The thermal velocities are a result of statistical mechanics, and are given as the mean velocity of the particles:

\[ v_n = \frac{\sqrt{8k_B T}}{\pi m_e} \quad \text{and} \quad v_p = \frac{\sqrt{8k_B T}}{\pi m_h}. \]  
\[ (27) \]

The net SRH recombination rate, per Shockley and Read’s analysis [5], is

\[ R_{SRH} = \frac{np - n_i^2}{\tau_{n,SRH}(n + n_1) + \tau_{p,SRH}(p + p_1)}, \]  
\[ (28) \]

where

\[ n_1 = N_C e^{-\frac{E_C - E_T}{k_BT}} \quad \text{and} \quad p_1 = N_V e^{-\frac{E_V - E_T}{k_BT}}. \]  
\[ (29) \]

Note that \( n_1 \) and \( p_1 \) are the equilibrium electron and hole concentrations if the Fermi level is aligned with the trap energy \( E_T \). From Equation (17) the net SRH lifetime is given as

\[ \tau_{SRH} = \delta n \left[ \frac{\tau_{n,SRH}(n_0 + \delta n + n_1) + \tau_{p,SRH}(p_0 + \delta n + p_1)}{np - n_i^2} \right], \]  
\[ (30) \]

which simplifies in the case where \( \delta n = \delta p \ll n_0, p_0 \) to

\[ \tau_{SRH} \approx \frac{\tau_{n,SRH}(n_0 + n_1) + \tau_{p,SRH}(p_0 + p_1)}{n_0 + p_0}. \]  
\[ (31) \]

In addition to trap states within the lattice, SRH recombination can occur at the surface of a semiconductor, where the lattice bonds abruptly terminate. These dangling bonds can have a number of active states within the band gap of the semiconductor. Additionally, the exposed surface is more readily introduced to impurities from the
environment. The recombination lifetime at the surface is nearly identical to the SRH lifetimes noted above, except that the trap density at the surface is two-dimensional.

Auger Recombination

Auger recombination refers to band-to-band impact ionization processes that involve the recombination of an electron-hole pair, with the energy of the annihilation being imparted to another carrier. In the example shown in Figure 4c, the electron-hole pair recombination occurs between the conduction and heavy hole bands, and an electron in the conduction band is kicked up to a higher energy. This is referred to as the electron Auger process, or Auger-1 using the classification of Beattie [6]. The analogous processes for holes in the valence bands, of which there are nine, are referred to as Auger-2 through Auger-10. The net Auger recombination rate [7] is

$$R_{\text{auger}} = \frac{C_n(n p - n_i^2)}{n_i^2} \left( \frac{n}{n_0} \right) + \frac{C_p(n p - n_i^2)}{n_i^2} \left( \frac{p}{p_0} \right),$$

where $C_n$ and $C_p$ are the coefficients for the electron and hole Auger processes, respectively. The Auger lifetime is expressed as

$$\tau_{\text{auger}} = \frac{n_i^4}{(n_0 + p_0 + \delta n)(C_n n p_0 + C_p p_0)}.$$
Modern light detectors take advantage of the optical absorption process in semiconductors, allowing for compact devices that, being similar in size, operation, and material structure with other semiconductor devices, offer integration with semiconductor opto-electronic and electronic components. A bulk semiconductor in equilibrium will have far more electrons in the valence band than the conduction band, and incoming photons can stimulate an electron to ‘jump’ the band gap and occupy the conduction band by transferring the photon energy to the electron. The energy of the photon is given by the product $h\nu$, where $h$ is Planck’s constant, and $\nu$ is the frequency of the photon (a photon can also be characterized by its wavelength $\lambda$, which is related to the frequency by $\lambda = \frac{c}{\nu}$).
The photon can only induce a jump across the band gap if this quantity is equal to or greater than the band gap. Consequently, devices that rely on this process have what is known as a cut-off frequency or wavelength, which corresponds to the lowest frequency (or longest wavelength) at which the device responds to optical excitation. This absorption process creates a free electron in the conduction band and a hole in the valence band. Other transitions are possible, such as intraband transitions where the excitation occurs between two separate bands in the same regime, e.g. from one conduction band to a higher energy conduction band. With the additional input of momentum from a phonon, a carrier can even be excited to a higher energy within a single band. Figure 5 diagrams the processes in a direct band gap semiconductor. The electron-hole pair (EHP) can then be detected as current if they are separated in an electric field and measured. Alternatively, the steady excess carrier concentrations created by the optical absorption process can be detected as an electric potential difference in a device, if the EHP is separated by a built-in field, as is the case in p-n junction based photodiodes (such as the photovoltaics utilized for solar energy harvesting).

Most semiconductor materials have bandgaps in the range of ~1 eV, allowing for detection of incident radiation in the visible and portions of the short-wavelength IR wavelength ranges. For longer wavelength infrared light detection, active regions with sufficiently small band gaps are required in order to generate free carriers. In 1959 Lawson et al. [8] initiated the development of HgCdTe (mercury cadmium telluride or MCT) alloys for IR detection. By adjusting the ratio of Hg to Cd, the band gap of the alloy can cover the wide range of 1-30 µm, allowing for short- to long-wavelength IR applications. They are
the most popular material for IR detectors, however the weak Hg–Te bond can lead to material quality and yield issues, especially for longer wavelengths [9]. Type-II superlattices (T2SLs) were initially proposed in 1977 by Sai-Halasz et al. [10]. Their proposal details InAs/GaSb superlattices with a short period the type-II band gap alignment, where the conduction band edge of the InAs lies below the valence band edge of the neighboring GaSb. This causes electrons and holes to be confined separately in the InAs and GaSb layers, respectively. Thus, the effective band gap of the superlattice is lower than the constituent materials and enables absorption of wavelengths in the mid- and long-wavelength infrared (MWIR and LWIR) regimes. Figure 6 shows a band diagram of an InAs/GaSb T2SL structure.

![Band diagram of an InAs/GaSb superlattice](image)

Figure 6. Band diagram of an InAs/GaSb superlattice with type-II band gap alignment, where the entirety of the InAs band gap (unshaded) lies below that of the GaSb band gap. Also shown is the separate confinement of electrons (black) and holes (white) leading to a smaller effective band gap.

The advantages of T2SLs over existing technologies including HgCdTe (MCT) alloys are a greater degree of control over the band gap via tuning of the material layer thicknesses and composition, good electron transport with effective masses still heavier than comparable MCT material leading to lower tunneling dark current, and leveraging of
the mature III-V growth technologies already in place for other optoelectronic applications. Through band gap engineering of the superlattice, the Auger recombination processes can be suppressed to achieve superior performance [11].

**Strain Management**

A major consideration in the growth of virtually all semiconductor devices at every scale is the lattice mismatch between the material being deposited and the substrate material. This mismatch is caused by forcing the new material (having a lattice constant $a_B$) to form atomic bonds with the existing material (having a lattice constant $a_A$). Though the materials might share the same crystallographic structure, the lattice constants of the two materials (i.e. their unit cell spacing, or periodicity) will not necessarily match ($a_A \neq a_B$) and so to match the substrate material, the deposited material will be forced to compress or stretch in order to match. This introduces strain, either compressive or tensile based on whether the substrate’s lattice constant is shorter or longer than the deposited material, respectively. The effects of strain on the band structure, particularly that of the heavy and light hole bands [3, pp. 132–144], must be taken into consideration when designing any semiconductor heterostructure device. T2SL structures are especially sensitive, as strain-induced shifts in the constituent materials’ band structure can have a dramatic effect on the low effective band gaps of the T2SL structures. In fact, for calculation of band structures and dispersion relations for infrared T2SL designs, it is necessary to consider coupling of spin-orbit split-off band with the conduction and valence (heavy hole and light hole) bands.
An additional concern arises from the mechanical behaviors that arise from the presence of strain. A strained material layer cannot be grown to arbitrary thickness; eventually the strain in the system is great enough such that the lattice undergoes localized deformations in order to eliminate or “relax” the stress. These so-called defects can manifest as imperfections as small as a single atom (the absence thereof, called a “vacancy,” or an exchange of two elements, called a “substitution” defect) but as large as entire lines of dislocations (characterized as “edge” and “screw” dislocations) [2, pp. 616–638]. Figure 7 depicts each of the basic types of defects. An important detail for defects in crystal structures is the requirement that dislocations form a contiguous loop, or terminate at the crystal surface [12, p. 153]. As a result, dislocations within semiconductors form interconnected networks that can span multiple dimensions. Dislocations that lie within a plane of growth (perpendicular to the direction of epitaxial growth) are called misfit dislocations, and the connecting dislocations that traverse the lattice vertically are called threading dislocations.
Figure 7. Illustrations of various types of defects. Vacancy (a) and interstitial (b) defects are referred to as point defects, and involve the absence of an atom from a lattice point or the insertion of an atom outside of one. In edge dislocations (c), entire planes become “slipped,” resulting in a plane terminating in the lattice. For screw dislocations (d), shifts occur in opposite directions at an interface (dashed lines and circles indicate the lattice underneath).

When these changes to the crystal structure occur, the solutions to Schrödinger’s equation are perturbed and new energy states become available. These defect states can reside within the band gap of the material, leading to Shockley-Read-Hall recombination as detailed previously. At long wavelengths, thick absorber regions are required to optimize the quantum efficiency of a detector device, which means that small amounts of strain can build up over ~microns of material growth. Thus, strain management in the layers of a superlattice are a very important aspect of design.

State-of-the-art

Since the establishment of the theory describing the T2SL, there has been a great amount of research aimed at achieving practical demonstrations of T2SL infrared detectors
with superior performance over state-of-the-art MCT detectors. In theoretical predictions published by Grein et al. [13], the ideal performance of InAs/InGaSb superlattice detectors surpasses that of MCT thanks to a lower rate of Auger recombination. However, a non-ideal SL that includes a Shockley-Read-Hall (SRH) recombination lifetime of 35 ns performs worse than MCT at lower temperatures, where the contribution of Auger recombination is proportionally smaller. Substantial effort has been focused on implementing barrier configurations to mitigate sources of dark current [14]–[16]. Notable is the so-called “nBn” device configuration, which consists of an n-type photo-active region, a high band gap thin barrier layer that acts to block majority carriers while still allowing holes to flow due to a low valence band offset. A detailed discussion of the nBn detector archetype follows in Chapter 4. Figure 8 shows a collection of published detectors produced by the Navy Research Laboratory [17], compared to the “Rule 07” heuristic for MCT detector performance [18], [19]. Developed and proposed in 2008, “Rule 07” is an empirical fit based on state-of-the-art MCT devices for temperatures above 77 K and with a temperature-device cutoff wavelength product ranging from 1700 to 400 µm-K. Handy as a rule of thumb for estimating the optimal performance of a MCT detector for a given cutoff, “Rule 07” is a convenient reference for evaluating the performance of T2SL devices. More recently, research has been focused on T2SLs based on InAs/InAsSb, removing the Ga to eliminate the possibility of cation vacancy and swapping-related defects and improve strain management [20]. Such devices have demonstrated promising carrier transport qualities [21] and have been successfully integrated into detector configurations [22] as well as nBn detector devices [23]. To continue to develop T2SL
detector technology and achieve superior performance, it is necessary to study fundamental materials issues that impact T2SL-based applications (including IR light emitting diodes and lasers as well as photodiodes) in order to identify future venues for improvement.

Figure 8: Plot of the resistance-area product versus cutoff wavelength for several state-of-the-art T2SL photodetectors (symbols), compared to the state-of-the-art performance of MCT detectors as delimited by the heuristic “rule 07” formula (line).

A popular method for characterizing the quality of a T2SL absorber is to measure its minority carrier lifetime via optical stimulation. A widely used technique for optical devices called photoluminescence (PL) uses focused laser pulses to excite excess carriers in a sample while a spectrometer measures the resulting radiative recombination via collection optics. This allows for the determination of the primary transition energy of the excited carriers, which is usually a stand-in for the band gap of the material. If instead one measures the total output power from the sample over time, in what is called time-resolved PL (TRPL) [24], one can obtain an estimate for the minority carrier recombination lifetime by fitting a model to the data while the number of excess carriers remains low compared to the background concentration. While useful for gauging the performance quality of a
T2SL, it does not reveal the carrier diffusion characteristics of the device such as minority carrier diffusion length, mobility, and diffusivity, which are also essential components in evaluating the impact of defects in the superlattice.

Electron beam induced current (EBIC) is a measurement for determining the diffusion length in a sample by measuring the electrical current of a sample in a scanning electron microscope (SEM), which uses high-energy electrons in a tightly focused beam to bombard the sample and produce a large amount of scattering events thereby creating excess carriers in the sample via kinetic energy transfer. Early demonstrations of this process in the 1960s [25] were followed by comprehensive treatments of the diffusion theory in the 1980s [26]–[28] as SEM technology and its use became more widespread. It has the capability to allow for the modeling of diffusion lengths via fitting of the data using first principle analysis, and when combined with other measurement techniques, such as TRPL, allows for thorough study of a sample’s transport properties. Additionally, the measurement can be carried out on virtually any functioning device structure, allowing for added testing capability to be inserted with minimal disruption to an existing workflow.

Outline of Thesis

The research presented aims to demonstrate an enhancement to the methodology of characterizing defects in type-II superlattice structures used for mid- and longwave infrared detection, but is intended to be generally applicable to other heterostructures and applications. Its main focus is the development of EBIC theory for the determination of minority carrier diffusion length in photodetector devices, which enables the determination
of other important physical parameters when used in conjunction with other measurement techniques.

Chapter 2 presents an overview of the EBIC theory and its formulation, followed by the experimental procedures for acquiring EBIC images as well as the theoretical and computational techniques we applied in order to obtain results. We also briefly cover qualitative EBIC top-down imaging of defects in an undoped T2SL structure using Schottky contacts.

Chapter 3 details the extension of the existing diffusion probability function to cover the EBIC measurement of InAs/GaSb T2SL devices with an extended lightly doped absorber layer, and the results obtained from two samples which show the improvement in diffusion that was obtained by forcing thin InSb interfacial layers in the superlattice to balance strain and allow for further band gap tuning.

Chapter 4 covers the application of EBIC theory to the nBn structure, using the measurement of InAs/InAsSb T2SLs in the nBn device configuration to demonstrate both the usage of EBIC with the barrier configuration and the combination of EBIC analysis with lifetime data obtained from complimentary TRPL measurements to determine the vertical minority carrier mobility and diffusivity.

Chapter 5 describes initial efforts to apply capacitance-voltage measurement techniques to characterize defect energy levels within a T2SL band gap. A novel sample design is described for the experiment, and initial results are presented and discussed.

Chapter 6 summarizes the work and discuss potential future work to further improve device characterization and the understanding of defect analysis.
CHAPTER 2
ELECTRON BEAM INDUCED CURRENT

Introduction

The key operating principle of a photodetector is the collection of excess carriers generated by a targeted range of incident photon wavelengths. As a simple example, consider an abrupt junction formed by a semiconductor with two regions of contrasting dopant concentrations. Dopants are foreign atoms that are artificially introduced to the semiconductor lattice and have different numbers of electrons in their outer valence shells, and thus alter the equilibrium carrier concentrations of the material. For instance, replacing a single Sb atom in GaSb with Te, which is the next element in the periodic table, allows the Te atom to “donate” its extra electron, raising the electron concentration. This is reflected as a shift in the Fermi level (refer to Equations (7) and (8)). When two adjacent regions in a semiconductor have different equilibrium Fermi levels, the abrupt change in carrier concentration leads to diffusion and creates a region around the junction that is devoid of carriers (known as the depletion region), leaving charged dopant ions and a region with a standing electric field. This allows the Fermi potential to remain constant in equilibrium, as shown in Figure 9.

This chapter contains experimental results and media that have been previously published in a scholarly journal: [35]. Permission has been obtained from the copyright holder, Cambridge University Press, for reprinting of material for this thesis. The author acknowledges contributions by the co-authors, namely the measurement and presentation of data featured in Figure 11.
Figure 9. Band diagram of a p-n homojunction, showing the conduction and valence band edges ($E_C$ and $E_V$) and the Fermi level ($E_F$). In the depletion region at the junction, generated electron hole pairs (a) are separated by the electric field to produce current. Additionally, minority carriers outside of the depletion region (b) can diffuse to it and be swept to the opposite side.

This configuration has many fundamentally useful properties. When a bias voltage is applied to the junction, the result depends on the polarity of the voltage. When a positive potential is applied to the p-side, the two ends of the junction become less separated, the depletion region shrinks, and minority carriers on both sides diffuse across it, leading to an exponential increase in current. This condition is known as forward bias. When positive potential is applied to the n-type side instead, the depletion region widens as minority carriers are removed from each side. The only current that flows in this scenario is due to free carriers diffusing into the electric field (depicted in Figure 9b), which is not affected by the magnitude of the applied bias voltage. This condition is referred to as reverse bias. The contrast in current passing through in forward and reverse bias allows a p-n junction to act as a rectifying element, allowing current to pass in one direction but virtually blocking it in the other. However, the p-n junction can also act as a photodetector, since
electron hole pairs generated by incoming photons can be picked up by the electric field and separated to provide current, which occurs most efficiently when EHPs are generated within the depletion region of the junction (Figure 9a).

Traditional device designs, such as the p-i-n junction, incorporate long active regions (in the case of the p-i-n junction, an intrinsic semiconductor inserted in the middle lengthens the depletion region) where photocarriers are generated and collected by an electric field. Outside of the electric field, excess minority carriers can contribute to current if they are able to diffuse into the field after being generated. Thus the quantum efficiency can be enhanced by improving the minority carrier diffusion length in the p and n regions, to allow for more current to be generated from absorbed photons. This behavior allows for the diffusion length to be characterized by carefully controlling the location of generated carriers and observing the change in current. Examples of such experiments go back to 1951, when F. S. Goucher [29] of Bell Laboratories demonstrated the measurement of diffusion and lifetime by scanning a point of light across a germanium p-n junction.

**EBIC Background**

Electron beam induced current (EBIC) works on the same principle as Goucher’s experiment but operates with orders of magnitude more precision and in modern studies is typically accompanied by rigorous first-principle models to provide accurate results. Rather than a focused light beam, EBIC employs beams of high-energy electrons which can be controlled with nanometer precision in the vacuum chamber of a scanning electron microscope (SEM). As the electrons penetrate the surface of the sample and collide within the crystal lattice, inelastic collisions transfer energy and create electron-hole pairs [30],
[31] within a volume that is determined by the atomic properties and the initial electron energy, typically on the order of $10^3$ eV. Knowledge of this generation volume as well as the device layer structure allows for the determination of minority carrier diffusion length in a similar fashion to the work by Goucher. The work presented follows largely from the initial theory and demonstration by Donolato [26], [27] and Bonard and Ganière [32] respectively. In addition, previous work by Li et al. [33], [34] was instrumental in providing the basis for the numerical implementation of the theory describing the EBIC current for a p-n junction.

**EBIC Imaging of Defects**

In addition to using cross-sectional EBIC images to extract diffusion parameters, the technique can also be generally applied to qualitatively assess the presence of electrically active defects in a material. Generally speaking, if defect structures are present with capture cross sections and trap densities great enough to impact a significant portion of carriers generated by the electron beam, then the defects’ proximity to the point of impact of the electron beam will affect the total current picked up by the system. Thus, within a few diffusion lengths of a collecting junction, the defects can be imaged by scanning the electron beam’s position and observing the change in total induced current. In the following section, we discuss two examples of this methodology in practice. In the first example, EBIC imaging is used to analyze the depth variance of strain-induced defect structures within a Si/SiGe/Si heterostructure. In the second, the technique is applied to a T2SL structure comprised of InAs/InAsSb.
The Si/SiGe/Si heterostructure [35] was grown perpendicular to (100) with 1.9 µm of step graded strained SiGe deposited on a Si substrate. It was followed by 2.65 µm of relaxed and unstrained SiGe, and a 17 nm layer of Si as the top layer. The collecting junction for EBIC was created by depositing a 70 nm Al top contact, creating a Schottky diode. EBIC images were taken using the system previously detailed in this chapter at ~7 K with helium cooling. To vary the depth within the sample at which carriers were generated, separate beam energies were used: 3 keV and 20 keV. To predict the depth at which electron hole pairs would be generated for each beam energy, Monte Carlo simulations of a variety of beam energies were carried out with a matching simulated layer structure to determine the distribution of energy within the sample. The results are plotted in Figure 10.

![Figure 10](image_url)

**Figure 10.** Plot of electron energy density in arbitrary units versus distance from top surface, as simulated via Monte Carlo simulation. Five different electron beam energies were used: 3, 5, 10, 20, and 30 keV. Also shown are the material layers, separated by dashed vertical lines (the top Si layer, which is 17 nm thick, is not shown). The Si substrate lies beyond the strained SiGe layer.
Based on the distributions shown, it can be seen that the lower energy electron beams (3, 5, and 10 keV) do not impart significant energy deep within the sample, having energy densities which fall by four orders of magnitude (from their peak, near surface, value) before making it even 1μm through the relaxed SiGe layer. Only the two highest simulated energies (20 and 30 keV) are capable of imparting significant energies to the strained SiGe layer, and therefore only EBIC images with those energies would be expected to have evidence of influence from defects residing within the strained region itself. Assuming that the relaxed SiGe would have many fewer naturally occurring defects originating in it, features in an EBIC image at 3 keV but not 20 keV would be expected to be located within or near the top strained Si layer, and not the strained SiGe. The reverse would hold for features visible in the 20 keV image, but not the 3 keV image. The results of the EBIC imaging are shown in Figure 11.
Figure 11. Comparison of EBIC images at (a) 3 keV and (b) 20 keV in a Si/SiGe/Si heterostructure. (c, d) A schematic plot to show the correlation between features in the images and hypothesized strain-induced dislocations. Dark points in the EBIC images are produced by vertical threading dislocations (TD), and dark lines are produced by misfit dislocations (MD).

The top-left image, Figure 11a, shows the EBIC image that results from a beam energy of 20 keV. Prominent in the image are cross-hatched lines, which correspond to a high density of dislocations. The same pattern of distinct bands in the (110) directions has been demonstrated in combined EBIC and TEM studies to originate from clusters of misfit dislocations (MDs) with threading dislocation (TD) “arms” [36]. When a TD penetrates through a large length of the sample, its corresponding “shadow” in the EBIC image is darker as a result of a greater density of defects at that beam position, leading to darker
pinhole-shaped points in the 20 keV image. Figure 11c shows an illustration of MDs and TDs originating from the strained SiGe layer that corresponds to the features seen in the EBIC image. In contrast, the image in Figure 11b shows the same location scanned with a 3 keV beam. Traces of the MD clusters from the 20 keV image are shown on both with red dashed lines, demonstrating that the linear features in Figure 11b similarly arise from MDs. Dots in the image correspond to TDs: at 3 keV, visible features must be near the top Si layer. Two scenarios for TDs at this depth exist: in the first, TDs connect MDs in the Si layer to the top surface. In the second, TDs connect from defects in the Si to the strained SiGe layer deeper within the sample. When this occurs, we see a dot at the end of the MD line in the low-energy EBIC image. In the first scenario, the resulting TD would be too short to be seen in the EBIC image, so these can be ruled out. The image of Figure 11d is consistent with this scenario. Some black dots do not connect to a line, suggesting that TDs have penetrated to the sample surface from the SiGe without forming MD lines at the top Si layer. Other dots are located at the end of MD lines and some MDs end with a black dot, while other ends do not. These observations provide unique and detailed information about defect character and distributions at the top of the strained Si surface.

**Theory of EBIC Junction Study**

The amount of current collected in the device due to the electron beam is treated as the product of two behaviors: the first is the spatial distribution of carriers generated by the electron beam relative to its point of impact on the sample, called the generation volume. This distribution is represented as \( h(x, y, z) \), with the origin representing the point where the center of the beam impacts the sample. The second behavior is the probability that a
given carrier contributes to the EBIC current by reaching a collecting boundary via
diffusion. This probability is represented as $\phi(x, y, z)$ using a reference frame that is fixed
based on the sample geometry. The two quantities are multiplied and convoluted to yield
an expression for the EBIC signal based on the electron beam position:

$$\eta(x) = \iiint h(x' - x, y', z')\phi(x', y', z')dx'dy'dz'. \quad (34)$$

**Generation Volume**

The generation volume is dependent on the composition of the sample under study,
as well as the energy of the electrons that impact it. Given that finding the distribution
based on analytical treatment is impractical due to the vast amount of scattering bodies in
the crystal lattice, representing the distribution is usually done with empirical formulae or
aggregate statistical methods. The latter involves simulating the exact trajectories of a large
number of electrons based on random number generation and compiling the results. This
technique is generally referred to as the “Monte Carlo” method, and for the generation
volume of an electron beam the model developed by Hovington et al. [37] and Drouin et
al. [38] is used to dictate the scattering length and angle of the electrons. A Gaussian shape
is assumed for the electron beam with a specified beam width. For our work, this parameter
is set to 1 nm in a rough approximation of the resolution capability of the JEOL 7000F
SEM. The initial position of the electron is randomly determined as

$$x_0 = \frac{d\sqrt{\log(R_1)}}{2 \times 1.65} \times \cos 2\pi R_2,$$

$$y_0 = \frac{d\sqrt{\log(R_1)}}{2 \times 1.65} \times \cos 2\pi R_3,$$  \quad (35)
where $d$ is the beam width, and $R_{1,2,3}$ are random numbers between zero and one. Once within the sample, the collisions of the electrons within the material are based on the scattering cross section $\sigma_i$, where $i$ has integer values to represent each element out of $n$ in the material. Values for $\sigma_i$ are determined from pre-calculated values stored in the computing program [37]. The distance between elastic collision events is calculated as

$$L_{\text{coll}} = \lambda_{\text{el}} \log(R_4), \quad (36)$$

$$\frac{1}{\lambda_{\text{el}}} = \rho N_0 \sum_{i=1}^{n} \frac{C_i \sigma_i}{A_i}, \quad (37)$$

where $\rho$ is the overall density of the material, $N_0$ is Avogadro’s constant, and $C_i, A_i$ are the weight fraction and atomic weight of element $i$, respectively. The scattering angle $\theta$ (for elastic collisions) is determined by numerically integrating the differential cross section $\delta \sigma$:

$$R_\theta = \frac{\int_{0}^{\theta} \delta \sigma \sin(\theta) \, d\theta}{\int_{0}^{\pi} \delta \sigma \sin(\theta) \, d\theta} \quad (38)$$

$R_\theta$ gives the probability of a scattering angle and can be used to determine the angle randomly. Between collisions, inelastic collisions are assumed to provide a continuous loss of energy without affecting the trajectory. The energy change per unit length between elastic collisions is given as

$$\frac{dE}{dS} = -7.85 \times 10^{-3} \rho \frac{E}{E} \sum_{j=1}^{n} \frac{C_j Z_j}{F_j} \ln \left( 1.116 \left( \frac{E}{J_j + K_j} \right) \right), \quad (39)$$

where $E$ is the energy at the previous elastic collision, $Z_j$ is the atomic number of element $j$, $J_j$ is the mean ionization potential, and $K_j$ is a variable derived from $Z_j$ [38]. A distance
and angle are calculated, with an accompanying energy loss, until the primary electron reaches a threshold energy value (typically 50 eV) and is discarded. The process then repeats with another simulated electron until the total number of electrons requested have been simulated. A variety of spatial information, such as the energy of electrons and the total energy lost as a function of position relative to the initial beam entry position, can then be recorded for a large amount of simulated electrons.

The results can be used in their raw form for a numerical evaluation of Equation (39), or fit to an empirical function. For the latter approach, we follow the form described by Bonard and Ganière [32]:

\[
h(x, y, z) = a \times \exp\left(-\frac{x^2}{\sigma_1^2}\right) z^2 \exp\left(-\frac{z}{\sigma_2}\right),
\]

which provides an analytically approachable form to the shape of the generation volume and is generally accurate in describing the overall shape of the distribution. Importantly, it allows for a nearly completely analytical solution to the total EBIC current distribution.

**Collection Probability**

The excess minority carriers, once generated, must then be collected as current. For low-level injection the assumption is that the carriers will undergo motion purely via diffusion. The problem is to then determine the likelihood of a carrier reaching a given collection boundary based on its relative position. The derivation starts with the diffusion equation (Equation (18)), which determines the distribution of carriers in steady state or in the presence of a source. Following the work of Donolato [28], we evaluate the Green’s function solution when a Dirac delta function is used as the generating source:
\[ D \nabla^2 (r, r') - \frac{1}{\tau} G(r, r') = -\delta(r - r'), \quad (41) \]

where \( G \) is Green’s function and \( \delta(r, r') \) is the Dirac delta shifted to a location \( r' \). In general there are two significant boundaries in the problem: the collection boundary representing a junction edge (e.g. an electric field, or some feature where current collection occurs), and the boundary between the material and vacuum, at which the electron beam impacts and surface recombination occurs. The two-dimensional problem is diagrammed in Figure 12. The coordinate reference frame is situated such that the x axis is lateral to the electron beam, which is parallel to the z axis and travels towards positive z. The material under study occupies the space \( z > 0 \), and the collection boundary is situated at \( x = 0 \).

![Figure 12. Depiction of the two-dimensional boundary equation for carrier collection probability. The sample (gray shading) occupies the positive half of the z-axis and is normal to the incident electron beam (red). The junction edge (\( x = 0 \)) and top surface of the sample (\( z = 0 \)) are assigned recombination velocities \( s \) and \( s_0 \), respectively.](image)

The boundary conditions for the Green’s function are
\[
D \frac{\partial G}{\partial x} \bigg|_{x=0} = sG, \text{ and }
\]
\[
D \frac{\partial G}{\partial z} \bigg|_{z=0} = s_0G.
\]

Given that the source for the Green’s function is a unit point source, the probability of collection at the junction edge is given by integrating the total current flowing through it, i.e.

\[
\phi(\mathbf{r}') = D \int_0^\infty \frac{\partial G}{\partial x} \, dz = Ds \int_0^\infty G \, dz.
\]  

(43)

If we designate a separate solution to Equation (41) without the source term as \( F(\mathbf{r}) \), then Green’s representation theorem yields

\[
F(\mathbf{r}') = D \int_0^\infty \left( F \frac{\partial G}{\partial x} - G \frac{\partial F}{\partial x} \right) \, dz + D \int_0^\infty \left( F \frac{\partial G}{\partial z} - G \frac{\partial F}{\partial z} \right) \, dx.
\]  

(44)

If \( F \) has an identical boundary condition at \( z = 0 \) to \( G \), then the second term will vanish. Additionally, if the boundary condition for \( x = 0 \) is modified to be

\[
\frac{\partial F}{\partial x} \bigg|_{x=0} = s(1 - F),
\]  

(45)

then Equation (44) becomes identical to Equation (43).

Thus we conclude that the collection probability can be expressed as a differential equation with the same form as the diffusion equation:

\[
\nabla^2 \phi - \frac{\phi}{L^2} = 0.
\]  

(46)
In a two-dimensional example where one lateral dimension is assumed to be infinite (as is the case in the experiment, where the lateral dimensions dwarf the epitaxial structure), boundary conditions are set as

\[
\left. \frac{\delta \phi}{\delta x} \right|_{x=0} = s\phi(0,z) - s, \quad (47)
\]

\[
\phi|_{x=\infty} = 0, \quad (48)
\]

\[
\left. \frac{\delta \phi}{\delta z} \right|_{z=0} = s_0\phi(x,0), \text{ and} \quad (49)
\]

\[
\left. \frac{\delta \phi}{\delta z} \right|_{z=\infty} = 0. \quad (50)
\]

In an ideal sample \( s \) approaches infinity, corresponding to a perfect collection interface. We usually assume this to be the case given the prevalence of an electric field. The solution for an infinite \( s \) is given as

\[
\phi(x,z) = \frac{1}{\pi} \left( \frac{s_0}{D} \right) \int_0^\infty \frac{1}{k^2 + (s_0/D)^2} \exp[-\mu(k)x] \cos(kz) \] 

\[
+ \left( \frac{s_0}{Dk} \right) \sin(kz) dk, \quad (51)
\]

with \( \mu(k) = \sqrt{k^2 + L^{-2}} \), \( L \) being the carrier diffusion length. \( k \) is a wave vector parameter that arises from the Eigen function solution to the differential equation in the \( z \) direction. Note that \( k \) in this context differs from the wave vector parameter in the Bloch theorem. The integration along \( k \) must be carried out numerically, though typically it extends to \(~100 \mu m^{-1} \) before further evaluation is smaller than double-precision floating point storage can represent and can be safely discarded. For devices with a collection boundary in the middle of two layers, a separate instance of Equation (51) can be applied in the opposite \( x \)
direction, with a distinct set of $L$ and $S/D$ parameters. It is also common to assume that $\phi = 1$ for a small distance at the boundary, corresponding to the built-in electric field of the junction.

**Experimental Procedure**

The experimental setup is shown in Figure 13. Samples are typically fabricated post-epitaxy by etching rectangular mesas from the device layers. To accomplish this, a photoresist is spun onto pieces of the wafer and patterned using ultra-violet radiation. By masking areas of the sample from the radiation, specific shapes within the photoresist can be kept unexposed. The portions of photoresist that receive a dosage of radiation become soluble in a developing solution and are washed away. Afterwards, the samples are placed in a solution of acids to remove material from the areas not covered in photoresist. Metal can then be evaporated onto the etched and unetched portions of the sample to create electrical contacts on both sides of the devices. Once fabricated, the sample is cleaved to expose a smooth surface that intersects with the etched portion of one or more devices. To facilitate in achieving a clean break, the substrate of the sample can be mechanically lapped to reduce the total thickness. The left image in Figure 13 shows the end result of the fabrication process; the yellow regions represent evaporated metals.
Figure 13. Experimental setup for EBIC measurement. Samples are cleaved to expose the junction side wall (left) and mounted on a block (right) to allow for the electron beam to strike the surface normally. Contact pads allow for wire bonds to measure the device current.

Fabricated samples are placed in the specimen chamber of a JEOL 7000F SEM, with an exposed junction side wall facing upwards normal to the incident electron beam. To accomplish this, the substrate is bonded to a rectangular brass block using indium paste and placed at a 90 degree angle on a cryogenic stage. The device is contacted with gold wires ball-bonded to the metallized surface, leading to contact pads where wires are soldered for external measurement. The setup requires that the SEM include a set of feedthrough connectors to allow for an electrical current to be passed from the sample inside of the chamber to an outside receiver. A current amplifier (Stanford Research Systems SR570) is then used to convert the EBIC from the sample into a voltage signal that can be fed into scanning hardware used by the SEM. Using the same hardware configuration that forms secondary electron images, an EBIC image is formed with the pixels representing a fixed position with a brightness value corresponding to the current read by the amplifier. Figure 14 shows an example of such an image. The variance in the
EBIC signal based on beam position along the growth direction can be found using image analysis software to compile the positions and pixel values to form a plot. By averaging multiple pixel lines together, the effects of variance due to small-scale imperfections and contamination on the sample surface, as well as noise in the system, can be mitigated.

![Image](image.png)

Figure 14. Example of images taken during an EBIC measurement. The secondary electron image (top half) is aligned with the EBIC image (bottom half) to show the relative location of the signal. Annotations on the secondary electron image show the locations of the top contact metal and the p-n junction.

**Extraction of Diffusion Length and Surface Recombination**

Extracting the diffusion length from the data can be carried out by calculating the proper solution to Equation (34) via numerical modeling and matching the results to the data with other physical parameters fixed. Figure 15 shows an example simulation with the effects of varying the diffusion length and surface recombination parameters, using the same beam energy value (as reflected in the values of $\sigma_1$ and $\sigma_2$). Based on the similarity
of progression between extremes, it can be reasoned that a given trend can be approximated closely by more than one pair of values for $L$ and $s_0/D$. Since the results depend on both the diffusion length and the magnitude of the ratio of surface recombination velocity to diffusivity, it is impractical to attempt a fit of a single data curve to extract both quantities. When possible, taking multiple EBIC images at varying beam energies can help to fit both quantities. Doing so allows for multiple models using generation volumes that change based on beam energy, with higher energies creating curves that depend more on the diffusion length than the surface recombination due to a majority of the carriers being generated further from the surface. To illustrate this, Figure 16 shows another set of example curves, this time showing the changes due to beam energy in two hypothetical sample—one with a low surface recombination ratio and the other with a high ratio. In this way, the separation of curves due to beam energy can be used to isolate and determine the surface recombination parameter.
Figure 15. Example EBIC simulation curves plotted for a hypothetical p-n junction device with a 1 µm depletion width (-1 < x < 0). Shown are the variations in a single n-type region (x < -1) for various minority carrier diffusion lengths (top) and for various surface recombination velocity/diffusivity ratio values (bottom).
Figure 16. Example EBIC plots of curves with variance in the beam energy, reflected as different values of beam generation parameters ($\sigma_1, \sigma_2$). All other parameters are fixed between the two graphs, except for the surface recombination to diffusivity ratio. In the upper graph, the ratio is 1 and in the lower graph the ratio is 100. The difference in behavior due to this change shows the effectiveness of using multiple beam energies to obtain a fit.

To fully determine the carrier diffusion lengths of a p-n junction, measurements of the EBIC signal should be made at multiple beam energies to cover the variation due to
surface effects. Depending on the sample dimensions, the characteristics of the p-type region can affect the EBIC behavior in the n-type region, and vice versa. Thus the distance covered by the measurement should be long enough such that there are portions of the data where a single region is dominant, generally on the order of $\sigma_1$. If multiple beam data are not available, an estimation of the order of magnitude of $S/D$ can be used to fit a single curve for a fixed beam energy. In an ideal scenario, the diffusivity of the material or the lifetime of the minority carriers can be used to supplement the EBIC results to provide the surface recombination velocity itself and the carrier mobilities along the direction of the EBIC scan.
CHAPTER 3
INTERFACIAL TREATMENT IN TYPE-II SUPERLATTICE PHOTODETECTORS

Introduction

An existing approach [39], [40] to managing strain in superlattices is to introduce monolayers of material between the existing superlattice layers in order to offset the lattice mismatch. This approach results in the balancing of compressive and tensile strains, allowing for increased absorber region thicknesses. In particular for the family of InAs/InGaSb superlattices, effects at the layer interface can have a drastic impact on device quality. Figure 17 shows the lattice constants plotted with respect to band gap energy for various semiconductor materials and compounds.

![Figure 17. Band gap versus lattice constant diagram for various semiconductor materials. Figure reprinted with permission from [41] under “open access” policy.](image)

This chapter contains experimental results and media that have been previously published in a scholarly journal; see [43]. Permission has been obtained from the copyright holder, American Institute of Physics, for reprinting of material for this thesis. Portions of this work relating to k-p band simulation were carried out by Pengfei Qiao.
Note the locations of InAs and GaSb on the figure: they are very close to each other (6.05 and 6.09 Å, respectively), meaning that a pure InAs/GaSb superlattice can be grown on a GaSb substrate with minimal tensile strain. However, the actual process of growing the layers on top of one another complicates matters. When switching between the two compounds, the identity of both the anion and cation are changed. This means that the interface transition must be a different compound, either InSb (Ga-Sb-In-As) or GaAs (In-As-Ga-Sb), both of which have a much larger difference in lattice constant. In addition to the single bond between the two compounds, anions from one layer can vaporize and exchange with those of the layer being grown, allowing the interfacial layers to have widths that exceed a monolayer. It has been shown [42] that this process occurs more readily when InAs is grown on GaSb, owing to a lower surface free energy for Sb. The strain caused by the unwanted interfacial behavior can lead to an increase in formation of dislocations in the sample, and ultimately lead to sub-optimal performance.

We set out in our work [43] to demonstrate the performance benefits of introducing intentional interfacial layers between InAs and GaSb, as well as the applications in altering the overall band gap of the T2SL. To do so, we designed a control InAs/GaSb T2SL for longwave IR detection, and created another design with the same InAs/GaSb thicknesses, but with an additional layer of InSb. T2SL samples with InAs/GaSb binary layers were grown via molecular beam epitaxy (MBE) to compare differences caused by interfacial layers during growth. For the control sample, a 500 nm GaSb p-type (1 × 10^{18} cm^{-3}) contact layer was grown on a GaSb substrate, followed by the main superlattice structure which was comprised of 24 Å GaSb / 45 Å InAs layers. 80 periods of graded p-type doped
superlattice were grown on top of the GaSb contact layer, followed by 300 periods of lightly p-doped absorber region superlattice, and another 80 periods of graded n-type doped superlattice. The top n-type (1 × 10^{18} \text{cm}^{-3}) contact layer was 20 nm of InAs. A second set of samples (referred to as the ‘interfacial sample’) were also grown with the same layer configuration, save for a forced 2.4 Å InSb interfacial layer following the GaSb layer in the superlattice. With the substrate as GaSb, the shorter lattice constant of InAs is offset by the much greater lattice constant of InSb, whereas a binary InAs/GaSb system has no mechanisms for deliberate strain management. Indeed, this is a strong reason for the consideration of a ternary system such as InAs/InGaSb.

**Band Structure Simulation**

To inform the design of the structures, we implemented computational routines to solve Schrödinger’s equation (3) numerically in order to determine the hole and electron wave functions as well as their corresponding energies. We employed the k·p method [3, pp. 126–130], [44] with an 8 × 8 Hamiltonian matrix, based on Kane’s formulation [45] with the Luttinger-Kohn model for degenerate bands [46], and Pikus-Bir’s model for strain effects [47]. With the Block theorem, the full equation is

\[
Hu_k(r) = E(k)u_k(r),
\]

\[
H = \frac{p^2}{2m} + V(r) + \frac{\hbar^2 k^2}{2m} + \frac{\hbar}{m_0} k \cdot \left( p + \frac{\hbar}{4mc^2} \sigma \times \nabla V \right) + \frac{\hbar}{4m^2c^2} \nabla V \times p \cdot \sigma,
\]

where \(p\) is the momentum vector, and \(\sigma\) is the Pauli spin matrix. By considering coupling between the conduction, light hole, heavy hole, and spin orbit bands, the Hamiltonian is block-diagonalized into an 8 × 8 matrix that includes the effects of strain as a perturbation.
For further details on the construction of the Hamiltonian, we refer to the works of Mou [48] and Qiao et al. [44].

Eigenvalues and corresponding eigenvectors of the matrix were calculated using MATLAB software to implement a finite difference method solution across a one-dimensional grid of material layers corresponding to the T2SL structures. Figure 18 shows the simulated band structure of the control as well as the calculated solutions for the first electron and heavy hole minibands. The cutoff wavelength for each calculated T2SL structure was estimated as the energy between the first heavy hole and conduction minibands.

Figure 18. Band diagrams of an InAs/GaSb superlattice (a) without interfacial treatment and (b) with a forced 2.4 Å InSb interface. Strained conduction and valence band edges are shown. The first conduction (C1) and heavy-hole (HH1) minibands are calculated by the 8-band k·p method and displayed as horizontal lines.
As is evident in the figure, the presence of the thin InSb layer shifts the T2SL valence band upwards, shrinking the effective band gap. This is due to the compressively strained InSb having a valence band higher than that of the GaSb. The effect on the conduction band states is less pronounced, as the InSb conduction band edge lies between that of the GaSb and InAs. For the control sample, the cutoff wavelength of the absorber was calculated as 10 µm, and the interfacial sample had a cutoff wavelength of 11.8 µm. These theoretical calculations were compared to experimental data to evaluate their accuracy.

**Fourier Transform Infrared Spectroscopy**

The response of a photodiode to external optical stimulus must be measured over a range of wavelengths in order to determine the cutoff wavelength. This can be done by selecting a wavelength at a time and measuring the electrical response calibrated to the incoming power, but such a method is time-consuming and would require high-precision instrumentation in order to provide the necessary resolution and breadth of wavelengths.
Fourier transform infrared spectroscopy (FTIR) is a technique that allows for the evaluation of many wavelengths simultaneously in a single measurement. What follows is a basic description of the measurement technique and its underlying mechanics. The heart of the measurement is a Michelson interferometer, shown in Figure 19. Incoming light of a given wavelength from a source (a globar in the figure, which is a SiC rod heated to ~1000 °C- it behaves similar to a black body as described in Chapter 1) is split by a beam splitter and bounced back by two mirrors. The two split beams rejoin and are collected by a detector, which provides the total intensity of the combined beam. When the two beams of light rejoin, the resulting beam will have a total intensity that depends on the difference
in total lengths traveled. If the difference is zero or an integer multiple of the wavelength, then the interference is constructive and the original light is restored. If instead the path difference is an odd multiple of half the wavelength, the interference is destructive and the two beams cancel each other out. The intensity seen at the output by the detector can be quantified as

\[ I(\lambda, \delta) = \frac{1}{2} I_0(\nu) D(\nu) F(\nu) (1 + \cos(2\pi\nu\delta)), \]  

(54)

where \( \nu \) is the wavenumber \( (1/\lambda) \), \( I_0(\nu) \) is the intensity spectrum of the light source, \( D(\nu) \) is the response spectrum of the detector, \( F(\nu) \) is the loss spectrum of the instrumentation (reflection losses at the mirrors and beamsplitter, etc.), and \( \delta \) is the path difference between the two mirror arms. In an FTIR, one of the mirrors is mounted on a one-dimensional motion track so that \( \delta \) can be modulated. The total signal observed by the detector as a function of the path difference is

\[ I(\delta) = \int_0^\infty I(\nu, \delta) d\nu \]

\[ = \frac{1}{2} \int_0^\infty I_0(\nu) D(\nu) F(\nu) d\nu \]

\[ + \frac{1}{2} \int_0^\infty I_0(\nu) D(\nu) F(\nu) \cos(2\pi\nu\delta) d\nu. \]  

(55)

\( I(\delta) \) is commonly referred to as the interferogram. Note that the first term of the result is a constant with respect to the path difference. If the constant term is removed, the combined spectral response can be recovered via a Fourier transform:

\[ I_0(\lambda) D(\lambda) F(\lambda) = \frac{4}{\pi} \int_0^\infty I_{AC}(\delta) \cos(2\pi\nu\delta) d\nu, \]  

(56)
where \( I_{AC}(\delta) \) denotes the signal measured with the constant term subtracted. If \( I_0(\lambda) \) and \( F(\lambda) \) are known, or determined to be constant over the range of interest, then the Fourier transform recovers the response function of the detector. Thus, the spectral response of the detector can be recovered in a single measurement by sweeping \( \delta \). This also has the benefit of utilizing the detector’s response to a wide range of wavelengths, meaning that the overall impact of noise is lower compared to a monochromatic measurement.

In our measurements, because the exact response of the source and FTIR system was not known the absolute responsivity of the devices were estimated with an external calibration method. A Mikron M 305 black body instrument, operating at 1000 °C, was used in conjunction with a 4.8 µm optical spike filter to excite the detector with a known intensity at the specified wavelength. To isolate the excitation of the black body from other background radiation, an optical chopper set to 1 kHz frequency was used to modulate the black body radiation and a Stanford Research Systems SR530 lock-in amplifier measured the current of the detectors. The optical power incident on the detector is given as

\[
\Phi_d = \pi \left( \frac{a}{2d} \right)^2 A_{opt} c m t_w \int \frac{2hc^2}{\lambda^5} \frac{t_f(\lambda)}{\exp\left(\frac{hc}{\lambda kT}\right) - 1} d\lambda,
\]

where \( \Phi_d \) is the optical power in watts, \( a \) is the black body aperture diameter, \( d \) is the distance from the detector to the black body aperture, \( A_{opt} \) is the optical area of the detector, \( c_m \) is the optical chopper modulation factor (~0.35), \( t_w \) is the transmission coefficient of the ZnSe window (0.7), and \( t_f(\lambda) \) is the measured transmission spectrum of the spike filter. The limits of the integration are largely determined by \( t_f(\lambda) \), which is negligible outside of a 1 µm window around 4.8 µm. The detector responsivity \( R \) can then
be estimated at 4.8 µm as $i_p/\Phi_d$ (A/W), where $i_p$ is the total current measured via lock-in amplifier. This value can then be used to normalize the spectra obtained via FTIR, but care must be taken to note that wavelengths where the system and source responses are not flat will not reflect the true responsivity of the device.

**Experimental Results**

The response spectra of the control and interfacial samples were measured with a Bomem DA-8 FTIR, using fabricated photodetector samples connected to the system via a current amplifier (Stanford Research Systems S570). The samples were cooled to 77 K with liquid nitrogen in a sealed cryogenic container (“cryostat”) with ZnSe optical windows to allow transmission of light. A globar was used as the infrared source (~4 to 15 µm), combined with a beam splitter made with KBr (~0.2 to 28 µm). Calibrated responsivity spectra for 400 × 400 µm$^2$ detector samples are shown in Figure 20. In the data, the control sample is referred to as “IFA” and the interfacial sample is referred to as “IFRA.” Note that the data shows a fall-off of responsivity for smaller wavelengths; this behavior can be ignored since it is due to the decrease in emission for small wavelengths by the globar source.
The data shows a clear shift in cut-off wavelengths due to the forced InSb layer. From the estimated wavelength at which the responsivity falls to zero based on fits to the data, the cutoff for the control sample is measured as 10 µm and the cutoff for the interfacial sample is measured as 11.7 µm, in close agreement with the modeled band structure of the T2SLs. In addition, the peak responsivities of the devices are clearly different. Despite having a longer cut-off wavelength the interfacial sample has a peak responsivity of ~2 A/W, compared to ~1.5 A/W for the control sample. To investigate the transport properties of the devices and demonstrate a direct observation of the improvement due to interfacial control, we sought to characterize the minority carrier diffusion lengths in the doped T2SL layers via an EBIC measurement.
Modeling of EBIC in Photodiodes with Long Absorber Regions

To model the EBIC behavior in both the control and interfacial samples, we developed an extension to the theory published by Donolato, Bonard and Ganière. In a heterostructure device with more than two distinct regions, especially when the middle region is too long for the depletion approximation to be applied for the full length, the additional boundary requires consideration. This scenario was modeled by considering a device setup equivalent to Figure 12, but with a boundary at \( x = W_\pi \) corresponding to the lightly p-type (referred to in shorthand as “\( \pi \))” absorber region in the device structure. At this boundary, the doping concentration shifts from lightly p-type to strong p-type. There, the diffusion length of minority electrons is assumed to change from \( L_\pi \) in the lightly doped region to \( L_n \) in the heavily doped region. The collection probability then behaves as described in (51), except that for \( x > W_\pi \), the diffusion length is changed and the value of \( \phi \) at the boundary must be continuous. Thus, the probability of collection across the entire structure is given as

\[
\phi(x, z) = \frac{2}{\pi} \left( \frac{S_h}{D_h} \right) \int_0^\infty \frac{1}{k^2 + (S_h/D_h)^2} \exp[+\mu_h(k)x] \left[ \cos(kz) + \left( \frac{S_h}{D_h k} \right) \sin(kz) \right] dk, \quad (x \leq 0),
\]

\[
\phi(x, z) = \frac{2}{\pi} \left( \frac{S_e}{D_e} \right) \int_0^\infty \frac{1}{k^2 + (S_e/D_e)^2} \exp[-\mu_e(k)x] \left[ \cos(kz) + \left( \frac{S_e}{D_e k} \right) \sin(kz) \right] dk, \quad (0 < x < W_\pi),
\]
\[ \phi(x, z) = \frac{2}{\pi} \left( \frac{S_e}{D_e} \right) \int_0^\infty \frac{1}{k^2 + \left( \frac{S_e}{D_e} \right)^2} \exp[-\mu_\pi(k)W_{\pi} - \mu_\pi(k)L_{\pi}] \]

\[ \times \left[ \cos(kz) + \left( \frac{S_e}{D_e} \right) \sin(kz) \right] dk , \quad (x \geq W_{\pi}). \]

An example of this probability distribution with the integral along \( k \) carried out numerically is shown in Figure 21. In the example, \( W_{\pi} \) is set as 0.5 \( \mu \text{m} \) to aid in the visualization. Arbitrary values were used for the transport parameters, but the diffusion length in the lightly doped absorber region was kept larger than that of the heavily doped p-type region. From the plot it can be seen that the probability of collection remains continuous at all points, but the change in collection over a given distance along \( x \) changes along the \( \pi \)-p boundary.

Figure 21. Plot of the modeled carrier collection probability based on carrier position for a representative transverse EBIC experiment. The \( x \) coordinate represents distance parallel to the device growth direction, and the \( z \) coordinate represents distance from the surface of the device. The area between \( x = 0 \) and \( x = 0.5 \) \( \mu \text{m} \) represents a lightly doped absorber region. At \( x = 0 \), the majority carriers switch from holes to electrons, and thus carriers along that boundary are considered to have a collection probability of 1.
EBIC measurements were performed in a JEOL 7000F scanning electron microscope (SEM) with a helium cooled stage. The samples were wire bonded to allow for electrical contact and measured with a Stanford Research System SR570 current preamplifier to generate EBIC profiles. Data was obtained for both types of samples (with and without InSb interfaces) while helium cooled to 6 K with a beam energy of 15 keV and a probe current of 1 nA. The data, as well as the fitted curves for each sample, are shown in Figure 22. In the plot, the EBIC values obtained in the measurement are normalized to a maximum value of 1 and translated to align the peak values for comparison. To align with the orientation of the sample in the SEM, the $x$ axis is reversed such that the n-type region corresponds to $x > 0$.

Figure 22. Measured EBIC signal vs. beam position for two InAs/Gasb T2SL p-$\pi$-n devices, grown using no interfacial treatment (square), and with a forced 2.4 Å InSb interfacial layer (cross). Also shown are the computed fits for each data set (dashed lines).
A qualitative analysis of the two data sets indicates that the sample with forced interfaces exhibits a decreased amount of roll-off, especially on the n-type side. Carrier mobilities were adapted from Bandara et al. [49] for similar samples, and estimated surface recombination to diffusion ratios, based on previous modeling by Li et al [33], were used to fit the EBIC behavior of both samples. Though it is a major assumption to use the same values for both samples, we wished to simplify the fitting of multiple parameters to provide meaningful comparison between the two samples. Additionally, the $S/D$ parameters ultimately have a small impact on the results obtained for diffusion length for the given beam energy. For minority electrons, the assumed parameters were $1100 \text{ cm}^2/\text{Vs}$ mobility and $400 \text{ cm}^{-1} S_e/D_e$. For minority holes, we assumed values of $100 \text{ cm}^2/\text{Vs}$ mobility and $1000 \text{ cm}^{-1} S_h/D_h$. Though recent published results on vertical mobility for similar lightly p-doped T2SL structures by Umana-Membreno et al. [50] have shown larger values for electron and hole mobilities, we believe that our values are reasonable given the higher doping concentrations in the contact superlattice regions. Given those values, we observed a significant improvement in the diffusion length (from 500 nm to 900 nm), and thus the estimated minority carrier lifetime (from 48.4 ns to 157 ns) of the holes in the n-region due to the interfacial treatment. A slight improvement for the minority electrons in the p-doped region is also noticed, but the margin of improvement is much smaller – 1.7 to 2 µm and 50.8 to 70.3 ns for diffusion length and lifetime, respectively.

In conclusion, we have derived an altered model to fit vertical EBIC measurements taken from p-π-n T2SL photodetector structures. Using this technique we have fit experimental data to extract minority carrier lifetimes. We predicted a 2 µm shift in cutoff
wavelength for InAs/GaSb T2SLs with an intentional InSb layer grown between the GaSb and InAs layers via $k \cdot p$ band structure analysis, and confirmed the prediction with responsivity spectra measured with FTIR. Additionally, a three-fold increase in minority carrier lifetime within the interfacial samples demonstrates an improvement in the transport properties of the T2SL on top of the benefits of additional band gap engineering. We expect that the cause of this improvement derives mainly from structural improvement in our GaSb layers resulting from the improved strain management, and the results are reflected in an increase in peak responsivity for the interfacial sample. Further study of strain effect with EBIC could focus on InAs/InAsSb structures, as well as the temperature dependence of vertical transport in strained T2SLs to characterize the impact of strain on high-temperature operation.
CHAPTER 4

COMPREHENSIVE VERTICAL TRANSPORT CHARACTERIZATION OF NBN TYPE-II SUPERLATTICE DETECTORS

Introduction

To further develop the characterization potential of the EBIC methodology, both as a general technique and as a method of studying the transport in T2SL-based photodetectors, we sought to apply the technique to state-of-the-art device designs. Moreover, we wished to support the diffusion length fitting with data from other measurements to reduce and/or eliminate the use of assumed material properties in the process. In order to accomplish this, a full EBIC characterization to determine the diffusion length and surface recombination to diffusivity ratio (as detailed in Chapter 2) would be required, with an additional source of the diffusivity value, or carrier lifetime. With this information, by virtue of the relationships given by equations (14) and (20), the vertical mobility of minority carriers can be determined.

For the following study we used T2SL detectors grown in an nBn device architecture. To obtain information about the carrier lifetime within the T2SL, a technique

This chapter contains experimental results and media that have been previously published in a scholarly journal: see [65]. Permission has been obtained from the copyright holder, American Institute of Physics, for reprinting of material for this thesis. Samples used in this study were provided by Arizona State University. Portions of this work relating to photoluminescence spectra were carried out by Runyu Liu. Experimental apparatuses used for time-resolved lifetime measurements were provided by Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under Contract No. DE-AC04-94AL85000.
known as time-resolved photoluminescence (TRPL) was employed. This chapter provides an overview of both the nBn device as well as the concepts behind the TRPL measurement.

**nBn Photodetectors**

Recent innovations have explored methods of engineering the device band structure of a T2SL photodetector beyond the basic p-n junction architecture in order to mitigate harmful leakage currents that arise in junctions formed with narrow bandgap T2SLs. One such structure is the nBn detector [51], a unipolar device comprised of an n-type absorber layer, a thin barrier layer, and a top n-type contact layer, as depicted schematically in Figure 23. The thin majority carrier barrier layer suppresses dark current by utilizing a high band gap blocking layer to block the flow of majority carriers while allowing optically generated minority carriers through due to a minimal valence band offset, which can be eliminated by applying a negative potential to the top contact layer. This suppresses SRH recombination by keeping the depletion of the absorber region to a minimum, since regions where the Fermi level lies close to the middle of the bandgap have a higher SRH recombination rate (28).
Figure 23. Band gap diagram of an nBn detector at zero bias (top) and with a negative bias applied to the rightmost layer (bottom). When electron hole pairs are generated in the n-type regions, holes are allowed to pass from the absorber through the barrier layer to the top contact while electrons are blocked from flowing into the absorber. A bias is applied to eliminate any band offset barrier in the valence band.

Other contributions to dark current (i.e. noise) are thermal and Auger generation processes in the absorber region. Holes generated through thermal processes are still able to diffuse to the barrier and cross to the top contact, meaning that this current will be more significant than SRH in the nBn detector. Following the reasoning of Maimon and Wicks [51], the temperature dependence of the device can be evaluated by examining the current in photodiodes compared to nBn detectors. In photodiodes, the current density due to SRH depends primarily on the depletion region, where the Fermi level activates mid-gap traps and the material is mostly intrinsic. The SRH-dominated current density is thus

\[ J_{SRH} \approx \frac{q n_i}{\tau_{SRH}} W_{dep} \]  

(61)

where \( W_{dep} \) is the width of the depletion region. For nBn detectors, holes in the n-type absorber region must diffuse to the barrier, and so the current density is
where $\tau_{diff}$ is the diffusion-based lifetime and $N_D$ is the n-type doping concentration. The temperature dependence of $n_i^2$ is $\propto \exp\left(\frac{E_g}{k_B T}\right)$ (see Equation (10)), so with all other considerations being roughly equal, e.g. $W_{dep}$ and $L_h$, $\tau_{SRH}$ and $\tau_{diff}$, then an nBn detector will operate with comparable dark current at a temperature twice that of a photodiode ($n_i$ vs. $n_i^2$).

Our work in this chapter deals with an nBn detector that incorporates a Ga-free T2SL absorber region. Recently, there has been growing interest in Ga-free InAs/InAsSb T2SL structures. Such devices have demonstrated reduced dark currents [23] and longer minority carrier lifetimes, hypothesized to result from the absence of native defects associated with GaSb [52]. Notably, a minority carrier lifetime of greater than 400 ns was reported by Steenbergen et al. in a bulk InAs/InAsSb T2SL designed for LWIR [20]. Further investigation by Olson et al. [21], [53] and Aytac et al. [54] explored the recombination mechanisms for the relatively unstudied system, showing that SRH recombination dominated the lifetimes in both unintentionally and intentionally doped bulk InAs/InAsSb T2SLs at low temperatures.

**Time-Resolved Photoluminescence**

The term “photoluminescence” (PL) refers to a process by which a material emits light after being subjected to illumination by an outside light source. Provided that the light used has enough energy per photon to excite carriers across the band gap, it will create excess carriers in the material, which will eventually recombine according to the processes...
described in Chapter 1 in the absence of an electric field. A portion of that recombination will be radiative, emitting photons back out of the material. Since the excess energy \( h\nu - E_g \) is lost extremely quickly to phonon scattering in the lattice [55], the photons emitted will primarily be those of the band gap, meaning that a spectral analysis of this luminescence can be used to experimentally estimate the band gap energy.

The time dependence of the excess carrier population following a temporally localized excitation is also significant. Based on the recombination mechanisms of Chapter 1, the overall lifetime for excess minority carriers in an n-type semiconductor in a low-injection limit [56] can be written as

\[
\frac{1}{\tau_h} = \frac{1}{\tau_{SRH}} + Bn_0 + Cn_0^2,
\]

where \( B \) is the radiative recombination coefficient, and \( C \) is the Auger recombination coefficient. Following an initial injection of excess carriers at a time \( t = 0 \), the concentration decays through recombination according to \( \delta p(t) = \delta p(0)\exp(-t/\tau_h) \). If the photoluminescence of the sample is measured over time, then the measured intensity will follow \( \delta p \) and thus by fitting intensity to a single exponential function when the remaining minority carrier concentration is low enough to satisfy the above conditions, the minority carrier lifetime can be recovered. This technique is referred to as time-resolved photoluminescence (TRPL).

**Optical Characterization of InAs/InAsSb nBn Photodetectors**

The sample used for measurements was provided by Arizona State University; a description of its growth follows. The sample was grown via molecular beam epitaxy on
an undoped GaSb substrate. The epitaxial layers consist of a 500 nm GaSb buffer, a 950 nm n-type T2SL bottom contact layer, a 2.4 µm (256 periods) InAs/InAsSb T2SL absorber layer, the wide band gap superlattice barrier layer, and a T2SL top n-type contact layer. The T2SL absorber structure used was 49 Å InAs/45 Å InAs$_{0.81}$Sb$_{0.19}$, resulting in a superlattice bandgap of 5.5 µm. The bottom contact was created using 950 nm (101 periods) of the same T2SL structure as the absorber, doped n-type using Si with doping concentrations of $1 \times 10^{18}$ cm$^{-3}$. Similarly, the top contact was formed with 96 nm (10 periods) of the absorber T2SL structure followed by 30 nm of bulk InAs. The top contact was uniformly doped n-type at $1 \times 10^{18}$ cm$^{-3}$, also using Si. The barrier layer consisted of 20 periods of a 29.3 Å InAs / 21.7 Å AlGaSb superlattice. Top and bottom contacts were formed by etching mesas down to the bottom contact layer using a standard lithography and chemical wet etch process, followed by deposition of Ti/Pt/Au metal contacts via a traditional metallization and lift-off technique.

We set out to verify the optical response of the T2SL absorber region of the device, along with the capability of detecting the targeted wavelengths of light with the nBn structure. Two optical methods were used, both taken with the sample cooled to 77 K with liquid nitrogen: first, we obtained an uncalibrated optical responsivity spectrum from a fabricated device using the FTIR methodology described in Chapter 3. For this measurement, we used a CaF$_2$ beam splitter with the globar source to target the MWIR cutoff wavelength. Second, we measured the PL spectra of the sample to verify the band gap of T2SL. The PL measurement setup and methodology were as follows: the sample was mounted in a cryostat behind a ZnSe window and excited using a 980 nm diode laser
pulsed at 45 kHz and incident at −45° through a quartz window. The excited PL was collected at +45°, collimated via a Ge lens (which blocks the pump laser wavelength) and fed to a Bruker V80v FTIR. The response of the internal FTIR MCT detector was taken to an external lock-in amplifier tuned to the laser pulse frequency, and returned to the FTIR for processing. The motion of the moving mirror in the FTIR is controlled such that the full lock-in signal can be retrieved for each calculated position of the interferogram. This mode of operation is known as amplitude modulation step-scan mode, and allows for a spectrum to be measured for the excited PL even as the result of a pulsed excitation. The results of both measurements are shown together in Figure 24. From both the FTIR and PL measurements, the band gap of the T2SL was estimated to be approximately 4.8 µm, in good agreement with the provided estimate.
Figure 24. Uncalibrated responsivity (left axis) and PL (right axis) spectra of the nBn detector taken via Fourier transform spectroscopy. The raw responsivity data (points) are shown along with a smoothed average (solid line), and the PL spectrum (dashed line) is overlaid for comparison. Measurements were taken at 77 K with a 0.35 V bias applied to the detector for the responsivity spectrum.

Minority Carrier Lifetime Characterization

Following the verification of the device optical response, we set out to measure the minority carrier lifetime of the T2SL. TRPL measurements were carried out at Sandia National Laboratories using a 1064 nm pump laser with a 4 ns pulse width and 1 kHz repetition rate. The sample was housed in a cryostat with a closed-cycle helium exchange system for temperature control and excited with a range of optical powers while cooled to 16 K. The resulting emission was collected with a parabolic mirror and focused onto a MCT detector with a nanosecond response time. The voltage from the detector was recorded with a high-speed oscilloscope to obtain time-series data of the PL intensity.
synchronized to the pulses of the pump laser with a trigger signal. The pump power incident on the sample was measured by blocking the beam with a Thorlabs optical power meter. Multiple pump powers were used to obtain TRPL data: 3, 9.6, 50, and 245 µW. The data for each pump power value are plotted on a log scale with arbitrary units (since the responsivity of the detector, as well as the collection efficiency of the PL optics, were unknown) in Figure 25.

Figure 25. Time resolved photoluminescence data obtained from the InAs/InAsSb T2SL nBn detector sample at 16 K, for a range of optical pumping powers. The fit used to extract carrier lifetime (dashed line) from the 3 µW injection curve is shown for comparison.

For the data from the higher pump pulse powers, there is a distinctly faster decay rate of the PL intensity at the beginning of the transients (< 0.2 µs). This indicates that the initial excess carrier concentration exceeds the low-level injection conditions necessary for single exponential decay, and that radiative and Auger recombination processes contribute to the total recombination rate. To extract the minority carrier lifetime, we used a linear regression of the logarithmic data for 3 µW pump power for t > 0.2 µs. The fit is shown
as a dashed line in Figure 25. Based on this result, we estimated a hole lifetime of 200 ns. Our results fall short compared to comparable devices published in the literature, e.g. that of Olson et al. [53], but based on other publications we believe that the results fall within reasonable values for a MWIR T2SL (Connelly et al. [57] reported TRPL lifetimes as low as 100 ns at low temperatures for Ga-free T2SLs).

**EBIC Characterization of nBn Detectors**

With data on the lifetime of the Ga-free T2SLs in hand, we separately obtained fits for the hole diffusion length and surface recombination velocity to diffusivity ratio using EBIC measurements of other devices fabricated with the same process. As before, etched devices with metal contacts deposited on them were cleaved to expose the nBn layers, and the sample was mounted with silver paste to a brass block to orient the exposed surface parallel to the SEM stage. Samples were cryogenically cooled with helium gas in a JEOL 7000F SEM, with wire-bonded contacts fed outside the chamber via vacuum-sealed cables within the sample exchange rod. Similar to previous measurements, the device current was measured with a Stanford Research Systems SR570 pre-amplifier, with the voltage signal fed to a DigiScan imaging circuit to correlate beam position with EBIC signal intensity. EBIC measurements were conducted at 6 K with the device at zero bias voltage for a variety of beam energies (5, 10, 15, 20, and 25 keV) using a beam current of 1 nA. A combination of low device contact resistance, current saturation of the pre-amplifier, and signal voltage threshold of the imaging circuit prevented the application of bias to the device. Nonetheless, we were able to obtain distinct signals from the device in the area local to the barrier layers. The experimental data is shown in Figure 26a. The data presented with
strong peaks located within the barrier region, with a fall-off in intensity as the beam position was swept through the n-type absorber region towards the substrate. Figure 26b shows the modeled curves we used to extract parameters from the data as a comparison.

Before discussing the results of the fitting, a discussion of the method for fitting the data with unique consideration of the nBn structure follows.

Figure 26. EBIC signals plotted versus beam position on the sample (right horizontal axis, peak signal corresponding to x=0, with positive x direction indicating the direction of growth) as well as beam energy in keV (left horizontal axis). (a) Experimental data collected at 6 K in an SEM for the cleaved nBn detector, with the sample oriented with the exposed side wall normal to the beam. (b) Theoretical simulation of the EBIC signal for the sample used to extract minority carrier diffusion length and surface recombination parameters.

In the nBn structure, the barrier region is designed to suppress SRH current mostly through minimization of the depletion region width, while stopping majority carrier flow when the device is subjected to bias voltage. In normal operation, the larger band gap of the barrier combined with its low thickness means that its influence on the optical response of the device is negligible. This axiom does not hold when considering the EBIC measurement, however. The energies of the electron beam are several orders of magnitude above the semiconductor band gaps, and their locations are precisely controlled to achieve
high imaging resolution, meaning that the generation of carriers in the barrier layer can be a significant influence on the EBIC current when the beam is scanned near it. At zero bias voltage, carriers generated by the inelastic scattering process in the barrier region can contribute more to the EBIC current than those of the absorber region. This is by virtue of electrons and holes being free to diffuse to the n-type regions from a higher potential. In contrast, if there is any offset in the valence band at the barrier, holes from the absorber region require thermal excitation to pass through as current. Figure 27 diagrams the concept behind our reasoning.

Figure 27. Schematic diagram of the ASU nBn photodetector. (Top) Labelled regions of the device. (Bottom) expanded superlattice regions with electron and hole minibands (not to scale). In the barrier region, generated electron hole pairs can fully contribute to device current. In the absorber region, electrons are blocked and holes may be impeded by a non-zero valence band offset.

To model the behavior of the EBIC signal, we separated the current contributions from the absorber and barrier regions into separate models. The current caused by minority carriers in the absorber T2SL use the same model for $\eta$ as described in Chapter 2. For the current caused by carriers in the barrier region, we assume that the collection probability $\phi$
is unity for a region of thickness comparable to the diffusion length and allow $\eta$ to derive solely from the normalized energy absorbed by the barrier material (Equation (39)). This energy is calculated by running a Monte Carlo simulation per Chapter 2, with an additional specification of recording the energy lost to scattering within the barrier material. The simulation is repeated for a series of points along the sample to construct a plot of absorbed energy vs. beam position. The simulated curves for 5, 10, 15, 20, and 25 keV beam energy are plotted in Figure 28.

![Figure 28](image.png)

**Figure 28.** Plot of modeled energy absorbed due to inelastic scattering in the barrier region of the InAs/InAsSb nBn detector, as a function of the beam position along the epitaxial axis. Normalized log scale plots are shown for the various beam energies used in the experiment.

To fit to the experimental EBIC data for the nBn sample, the barrier absorption simulation data were added to the simulated EBIC current due to generation in the absorption region for the same beam position. One such fit is shown in Figure 29 for the
EBIC data at 15 keV. Since the relative barrier height at the valence band was not experimentally known, a weighted sum was used to combine the models. The method of determining this weighting factor was to match the simulated current’s inflection point at \( \sim 0.1 \mu m \) to that of the experimental data. The behavior of the data in the absorber region, along with the variation of the data with respect to beam energy, were used to obtain consistent fits to all of the experimental data, as shown in Figure 26.

![Graph showing modeled EBIC response](image)

Figure 29. Breakdown of the separate contributions to modeled EBIC response at 15 keV in the InAs/InAsSb nBn detector sample. The diffusion current due to minority carriers in the absorption region (dashed line) and the current due to carriers generated in the barrier region (dash-dot line) are combined in a weighted sum to form the total EBIC current (solid line). Shown for comparison is the experimental data (circles).

**Results and Discussion**

From the fit to the data, we were able to extract a hole diffusion length in the absorber region of 750 nm and a surface recombination to diffusivity ratio of \(10^6 \text{ cm}^{-1}\) at
the measurement temperature, 6 K. The diffusivity was calculated using the estimated minority carrier lifetime from the TRPL measurement: our estimation for the vertical hole diffusivity was $3 \times 10^{-2}$ cm$^2$/s. Based on the surface recombination to diffusivity ratio, we estimate a surface recombination velocity of $3 \times 10^4$ cm/s. The modeled EBIC response becomes increasingly less sensitive to the surface recombination parameter as $S/D$ grows larger in magnitude, so this value is an estimate of the lower bound for the true surface recombination velocity. Nonetheless, the large surface recombination velocity of our result compared to previous modeling done on comparable InAs/GaSb T2SL p-n junction photodiodes [33] suggests a strong contribution to the effective carrier lifetime by recombination at the surface of the InAs/InAsSb T2SL. To mitigate this, device fabrication methods which avoid etching below the barrier layers should be used to optimize device performance since effective methods of passivating the InAs/InAsSb T2SL surface have yet to be developed. Fortunately, the high band gap barrier region can act as a passivating layer if device arrays are fabricated without etching beneath it, but a weakness of the cross-sectional EBIC measurement procedure we have used is that it requires the entire vertical structure to be exposed. Using Einstein’s relation a vertical hole mobility of 60 cm$^2$/Vs is estimated in the Ga-free T2SL. This value is an order of magnitude larger than the vertical hole transport reported recently for InAs/GaSb T2SLs by Olson et al. [58], suggesting that the issue of hopping transport for holes in our Ga-free T2SL is less severe than for holes in an InAs/GaSb T2SL.

Perhaps more important than the quantities determined in this work is the demonstration of a relatively straightforward method of directly calculating the vertical
mobility in the superlattice. We have shown in our work that a rigorous application of EBIC analysis of a T2SL device with standard fabrication methods, when supplemented with time-resolved photoluminescence, allows for a direct measurement of the vertical mobility with very little ambiguity. While the nature of vertical transport is crucial to the performance of T2SL detectors, efforts to characterize it have been relatively sparse. Obtaining the in-plane mobility is fairly simple and requires only a Hall measurement to be performed on a sample, while recent efforts to study the vertical mobility have been restricted to theoretical discussion [59] and more complex magnetic field experiments reliant on specific sample layer structures and geometry [50]. Our work has the advantage of being applicable to any device that can be used as a photodetector, with relatively little special preparation methods (the only extra steps required post-fabrication are cleaving and mounting the sample).
CHAPTER 5

DEEP LEVEL TRANSIENT SPECTROSCOPY FOR INVESTIGATION OF DEFECT STATES IN TYPE-II SUPERLATTICES

Introduction

A key component in the effort to understand and improve the quality of mid- to longwave IR type-II superlattices is the ability to quantify the effects of defects that populate the crystal lattice. Since it has already been established that SRH recombination limits the performance of T2SL devices, the effort to reduce its impact requires understanding the types of defects that are common to each material system and what role they play in the T2SL carrier dynamics. Transmission electron microscopy, in addition to qualitative EBIC imaging as discussed previously, has long been used to observe defect superstructures in a sample. Imaging techniques also exist for assessing lattice quality at the subatomic level; recently Kim et al. [60] demonstrated the use of atom probe tomography to analyze the composition of InAs/GaSb SLs and quantified the interfacial mixing of anions and cations.

So far, the discussion of this work has involved measurements of the overall recombination lifetime $\tau_{SRH}$. This chapter details our efforts to investigate the capture and emission of carriers from distinct trap states in the T2SL band gap. The focus of this discussion is a technique known as deep level transient spectroscopy (DLTS), a technique first developed by Lang [61] in 1974 at Bell Laboratories. The following discussion details
the fundamentals of the technique, and summarizes our preliminary efforts thus far to apply the system to IR detector T2SL materials.

Principles of DLTS

To understand the mechanics of DLTS we re-visit the carrier dynamics of deep level traps, which were touched upon in Chapter 1. The DLTS measurement involves the isolation of occupied trap states in a region with no free carriers, and so the details of the capture and emission process are important here and rely on the trap level with respect to the band edge. Consider a trap with energy $E_T$ that is occupied by an electron within the depletion region of a p-n diode. Absent a hole population to allow capture of a hole to complete the recombination process, the electron will eventually be emitted back to the conduction band by means of thermal excitation. The emission rate for this process is given as [61]

$$e_T = v_n \sigma_n N_C \exp \left( \frac{- (E_C - E_T)}{k_B T} \right) = v_n \sigma_n N_C \exp \left( \frac{- \Delta E_C}{k_B T} \right) \quad (64)$$

Provided that the trap state has a net charge either when it is occupied or empty, the emission process will affect a change in the overall capacitance of the device. In the simplest case, if all traps of a single type are filled and then the depletion region is expanded across them at time $t = 0$, the capacitance changes according to the emission rate of the traps:

$$C(t) = \Delta C \times \exp \left( - \frac{t}{\tau_T} \right) \quad (65)$$

where $\Delta C$ is the change in capacitance due to the filled traps, and $\tau_T$ is the emission lifetime, equal to the inverse of $e_T$. The dependence of the capacitance transient on the
emission rate means that its behavior can be “tuned” by altering the temperature of the device under test. For illustration, Figure 30 demonstrates this concept using exponential functions with an amplitude of unity. In the extreme cases, the temperature will be so high that the emission rate is near instantaneous and the capacitance instantly relaxes, or so low that the capacitance does not appear to change.

![Figure 30. Demonstration plot of the DLTS rate window capture process. Main figure: a series of simple exponential plots (solid lines) representing capacitance, with varying lifetimes $\tau$. The difference in the capacitance between 1 and 2 sec (dashed lines) is calculated for each transient. Inset: plot of the rate window differential as a function of emission rate. Circles correspond to curves in the main plot.]

The goal is then to observe the capacitance over a rate window given by two times, $t_1$ and $t_2$, so as to find the transition point between the two extremes such that the value $|\Delta C| = |C(t_2) - C(t_1)|$ is maximized. The emission rate at which this occurs can be found by taking the derivative of $\Delta C$ and finding the value of $\tau$ that makes the result zero:

$$\frac{\partial \Delta C}{\partial \tau_T} = -\Delta C \left[ \frac{t_2}{\tau_T^2} \exp\left( -\frac{t_2}{\tau_T} \right) - \frac{t_1}{\tau_T^2} \exp\left( -\frac{t_1}{\tau_T} \right) \right].$$  (66)
\[
\tau_{T,\text{max}} = \frac{t_2 - t_1}{\ln\left(\frac{t_2}{t_1}\right)}.
\]

(67)

To examine the temperatures at which this maximum occurs, it is helpful to rearrange Equation (64):

\[
\ln\left(\frac{T^2}{e_T}\right) = \ln\left(\frac{T^2}{\nu_n \sigma_n N_C}\right) + \frac{\Delta E_C}{k_B T}.
\]

(68)

Equations (67) and (68) are the foundation of the DLTS method: if the transient \(C(t)\) is recorded for a range of temperatures, a wide variety of rate windows can be selected. The peak in \(\Delta C\) over temperature thus can be paired with the associated emission rate in Equation (67), and from a group of such pairings Equation (68) can be used to deduce the trap energy level \(\Delta E_C\) and the capture cross section \(\sigma_n\) (the temperature dependence of \(\nu_n N_C\) is \(T^2\)—see Equations (27) and (9)—thus only the effective mass is needed to calculate \(\sigma_n\)).

**Experimental Method**

Our system for DLTS measurements is a commercial system obtained from Semetrol, LLC. Sample housing and cooling is performed in a Janis cryostat with a closed-cycle helium gas exchange system providing constant cooling down to \(~16\) K. A combination of diode temperature sensors and diode heating elements is used for temperature control, coordinated by a Cryo-Con 32B temperature controller. A Boonton 7200 capacitance meter is used to measure the alternating current (AC) capacitance of the sample at a probe frequency of 1 MHz through co-axial cable connections to the sample through the cryostat housing. Capacitance values from the meter are captured and stored
via an analog-to-digital conversion unit that is connected to the instrumentation computer, which also controls the sample temperature and bias voltages using proprietary LabVIEW software communicating via IEEE-488 standard connections to the hardware.

Capacitance transients are generated by pulsing an applied voltage to the device, called the filling pulse, at a consistent frequency with a short duty cycle. The filling pulse is followed by a near-instantaneous shift to a reverse bias, known as the measurement bias. During each cycle, the capacitance signal is recorded for a fixed time after the voltage is reverted to the measurement level, and the data is saved. The process repeats several hundred times and the transients are averaged together. In this manner, a transient is obtained for many temperatures within a range that is stepped through by the temperature controller between measurements. The values of $\Delta C$ for a selected rate window are plotted for each temperature, smoothed and interpolated to estimate the peak temperatures, each of which is tabulated along with the corresponding value of $e$. By plotting $\ln(T_{max}^2/e)$ vs. $1/k_B T_{max}$, these pairs will ideally form lines whose slopes correspond to $\Delta E_T$, in units of eV.

**Sample Design**

Our goal was to study recombination centers in T2SL structures by designing devices specifically for the capacitance measurement. Existing T2SL devices available to us had high dark current densities for relatively small bias voltages, which interferes with the DLTS measurement in several ways. First, the capacitance meter relies on a measurement of the AC capacitance of the device by applying a 1 MHz, 15 mV signal over the direct current (DC) voltage. The corresponding AC response due to capacitance will be
completely out of phase with the probe signal, since the impedance of capacitance is purely imaginary. If the direct response current, which includes the leakage current in the device, is too large, it makes the isolation and retrieval of the out-of-phase capacitance current difficult and significantly degrades the sensitivity of the measurement. This effect has long hampered the use of DLTS for narrow band semiconductor structures, which tend to have larger leakage currents than their wide band gap counterparts.

Second, the presence of a significant flow of carriers in the device alters the emission dynamics of the traps after the voltage is pulsed. Recall that the formulation of the rate window analysis assumes that there are no carriers in the conduction band. With a leakage current in the device, this assumption is not valid since a carrier flux will exist throughout the device, and there will be carriers available to be captured by empty traps. If the current density is significant, the measured energy levels will then be inaccurate if the current is not properly accounted for.

To counter this effect, we set out to design a T2SL device specifically for the DLTS measurement. Our design incorporates an InAs/GaSb T2SL adjacent to a bulk GaSb p+-n junction, reducing dark current by having a larger band gap in the majority of the depletion region while having the T2SL close enough to be within the depletion region in reverse bias. Figure 31 shows the layer design of the sample, along with a control design that does not have a T2SL region.
Figure 31. Sample layer structures for DLTS measurement. The control sample (a) is a bulk GaSb p⁺-n diode. The T2SL sample (b) has the same dimensions and similar doping profile, but a 200 nm InAs/GaSb T2SL with a band gap of ~4.5 µm is inserted in the n-type region.

The T2SL region was designed using 8-band k·p simulation, in the same fashion as the samples in Chapter 3. The target cut-off wavelength was 5 µm. The SL consists of 24 Å of GaSb and 26 Å of InAs layers, repeated 40 times for a total length of 200 nm. The T2SL is incorporated into the n-type region of a GaSb p⁺-n junction with a p-type concentration of $1 \times 10^{17}$ cm⁻³ and an n-type concentration of $1 \times 10^{16}$ cm⁻³. The T2SL separation from the junction was set to 500 nm, based on estimations of the behavior of the depletion region width as a function of reverse bias. The sample designs were grown via MBE by the IQE Corporation. A particular detail of the growth of the sample worth noting is that the n-type regions of the design were grown by extending the Te-doped p-type substrate, and using compensation doping with implanted Si to reduce the concentration.
Ultimately, the impurity concentration using this method is greater and includes more impurity types compared to altering the concentration of Te.

Using the PL measurement procedure from Chapter 4, the samples were tested to determine the efficacy of the T2SL band structure design. Figure 32 shows the PL spectra for the control sample and the T2SL taken at a sample temperature of 77 K. The data roughly confirms the design, although the true cut-off wavelength of the T2SL appears to be closer to 4.5 µm. We note that the control sample PL spectrum shows a weak emission peak at around 2 µm, which matches well with the band gap of GaSb.

![Figure 32. PL spectra of the DLTS samples obtained via step-scan FTIR. Solid line: control sample, a bulk GaSb p-n junction. Dashed line: T2SL sample, with a 4.5 µm T2SL layer embedded in the control sample design.](image-url)
Results

Samples were fabricated with a standard wet etch process using photolithography to create square mesas. Both the control and T2SL samples were etched down to a depth of 1.3 µm with a citric and phosphoric acid solution to expose the bottom lightly n-type GaSb, and top and bottom contacts were formed by evaporating Ti/Pt/Au with a photoresist lift-off. For capacitance measurements, the samples were mounted in an integrated circuit testing chip carrier with gold pads for wire bonding. For devices with a total electrical area of $9 \times 10^{-4}$ cm$^2$, we measured the capacitance as a function of bias applied to the devices across a temperature range of 20 to 100 K. The junction capacitance of a p-n junction as a function of applied bias voltage $V$ is given as

$$C_j(V) = \sqrt{\frac{A^2 q \epsilon_s}{2(\phi_i - V) \left( \frac{N_a N_d}{N_a + N_d} \right)}},$$

where $A$ is the device area, $\epsilon_s$ is the semiconductor permittivity, $\phi_i$ is the potential drop across the junction at zero bias, $N_a$ is the acceptor doping concentration in the n-type region, and $N_d$ is the donor doping concentration in the p-type region. Rearranging and taking the derivative with respect to applied voltage, this becomes

$$\frac{d(1/C_j^2)}{dV} = \frac{2}{A^2 q \epsilon_s \left( \frac{N_a + N_d}{N_a N_d} \right)},$$

and thus by fitting to the slope of $1/C_j^2$ the apparent carrier concentration can be found:

For $N_a \gg N_d$, $(N_a + N_d/N_a N_d)$ can be approximated as $N_d^{-1}$. Figure 33 shows the capacitance-voltage results at 80 K for both the control and T2SL samples.
Figure 33. (Left axes) Plots of the inverse squared capacitance (solid lines) as a function of bias voltage for the control and T2SL samples for DLTS, taken at 80 K. (Right axes) Scatter plots of the apparent carrier concentration, based on the diode capacitance function, as a function of bias voltage.
The results for the control sample are as expected for a plain p-n junction. The shape of the $1/C^2$ curve appears mostly linear, in accordance with Equation (69). The apparent carrier density values were calculated by taking nearest-neighbor numerical derivatives at each voltage and using Equation (70) with a permittivity of $15.2\epsilon_0$, where $\epsilon_0$ is the permittivity of vacuum. The device electrical area of the samples was $400 \times 400 \mu\text{m}^2$. Assuming $N_A = 1 \times 10^{17} \text{ cm}^{-3}$, based on the apparent carrier density from the control sample data we estimate that $N_D \approx 5 \times 10^{16} \text{ cm}^{-1}$. The abrupt shift in response at low voltages in the C-V data implies a competing capacitance in series with the GaSb junction, likely a Schottky contact with the n-type region. We used Ti/Pt/Au contacts for both contacts, which is known to be ohmic with p-type GaSb [62] but likely forms a potential barrier when applied to n-type GaSb.

The data from the T2SL sample, on the other hand, behaves differently. As seen in Figure 33, there appear to be three major regions in the capacitance curve. First, at lower voltages, after what is presumed to be an inflection due to a Schottky contact, the carrier density appears to be around $2 \times 10^{16} \text{ cm}^{-3}$, as the depletion region is confined to the bulk GaSb n-type region above the T2SL. This is followed by a change in the slope of $1/C^2$ and the apparent carrier density as the depletion region begins to overlap with the T2SL. Based on the apparent density calculated within this voltage regime we estimate an electron density of $\sim 3 \times 10^{15} \text{ cm}^{-3}$ within the T2SL. At higher voltages ($V < -2$), the apparent carrier density appears to rise back to bulk levels, implying that the T2SL is fully depleted at this point onwards.
DLTS measurements were performed on both samples within a temperature range of 16 to 150 K. The parameters of the voltage pulses were tested at 80 K to verify stability in the capacitance transient and the equilibrium capacitance at the measurement voltage. Spectra were obtained using a filling voltage of 0 V and a measurement voltage of -2 V, with a pulse width of 50 ms and a capacitance polling frequency of 10 kHz. Transients were recorded for 25 ms after the switch in the measurement voltage and averaged 200 times for every 1 K in the temperature range. Figure 34 shows a selection of the results plotted such that the apparent equilibrium capacitance $C(\infty)$ is subtracted from each transient. In both samples, the capacitance transients rise up to the equilibrium value, suggesting that the trapped carriers are majority carriers. Given the disparity in dopant concentrations, this means that the most likely cause is the emission of trapped electrons in the n-type region.

![Figure 34. Capacitance over time for the control and T2SL DLTS samples across a range of temperatures. Capacitance data (z axis) at each temperature are shifted such that the steady state value is 0, for ease of comparison.](image)
From the data in Figure 34, sampling of $\Delta C$ vs. $T$ for various values of $t_1$ and $t_2$ was used to create so-called rate window plots. Based on the appearance of the capacitance transients for the various temperatures, we chose rate window ranges individually suited to the control and T2SL samples. For the control sample, we chose to observe the capacitance within a range of 1 to 15 ms, the temporal range in which the control sample, at low temperatures, shows the majority of its change in capacitance. In contrast, the T2SL sample was observed with a rate window extending from 0.2 msec to 3 msec, which allowed us to see more clearly changes in the upper half of the temperature range. Example rate window plots within these ranges are shown in Figure 35.

![Rate window plots for control and T2SL samples](image)

**Figure 35.** Rate window plots extracted from the DLTS data represented in Figure 34. The data plotted is the change in capacitance over the rate window $\Delta C$ normalized by the apparent equilibrium capacitance value $C(\infty)$. The data for the control sample (left) were collected over a range of rate emission windows from $5.2 \times 10^{-3}$ to $7.5 \times 10^{-3}$ sec$^{-1}$. The T2SL data were taken over a range of $1.4 \times 10^{-3}$ to $3.7 \times 10^{-3}$ sec$^{-1}$.

The local maxima from the rate window data was iteratively collected to create Arrhenius plots for extracting energy levels. Both sets of data were fit using Gaussian functions to extract the peak locations. In the case of the T2SL data, a shoulder in the data
at higher temperature (~125 K) can be seen, so we extracted data by fitting two separate Gaussian functions added together. The calculated peak temperatures, along with the corresponding emission rates (Equation (67)), are plotted in the semi-log format of Equation (68) in Figure 36, with $\ln(T_{\text{max}}^2/e)$ as the y-axis and $1/k_B T$ as the x-axis.

Figure 36. Arrhenius plots of rate window results from the DLTS measurement. The rate windows for various emission values (see Figure 35) were analyzed to extract trap signatures. For the control sample (left), peak values were calculated via interpolation of the data. For the T2SL sample, the data was fitted to a two-term Gaussian expression and the resulting peak centers were analyzed. The peak data (crosses) for the control and T2SL sample are plotted along with linear fits used to extract the activation energies (various lines).

The results of the fits to the data are collected in Table 1. The trap energies were taken as the slope of a linear fit to the data. In the case of the T2SL sample, the data from distinct Gaussians were treated with separate linear fits, with a preference for data from higher emission rates where the separate Gaussians are more distinguishable. To calculate the capture cross section $\sigma_n$, the y-intercept values from the linear fits were used in conjunction with Equation (68) and the electron effective mass in bulk GaSb ($0.04 \times m_0$...
We chose to use this effective mass given the similarity of the calculated energy levels of each signature.

Table 1. Calculated trap signature results from the DLTS analysis.

<table>
<thead>
<tr>
<th>Signature</th>
<th>$\Delta E$ (meV)</th>
<th>$\sigma_n$ (GaSb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>50</td>
<td>$1.6 \times 10^{-16}$ cm$^2$</td>
</tr>
<tr>
<td>T2SL (1)</td>
<td>46</td>
<td>$4.2 \times 10^{-20}$ cm$^2$</td>
</tr>
<tr>
<td>T2SL (2)</td>
<td>49</td>
<td>$2.3 \times 10^{-19}$ cm$^2$</td>
</tr>
</tbody>
</table>

All three measured signatures have an energy of 50 meV within the bounds of 95% confidence. Based on this finding, and the similarity to published calculations of point defect levels in bulk GaSb by Hakala et al. [64], we conjecture that the observed trends are due to the emissions from defects in the bulk GaSb, both in the control sample and the T2SL sample. A large reduction in capture cross section and a split into two separate signatures with an order of magnitude separation in cross section is likely due to the presence of the T2SL layer, which likely introduces additional defects throughout the device structure. Further investigation is thus necessary to demonstrate the observation of carrier emission from defects in the T2SL, where electrons reside in the InAs layers far beneath the GaSb conduction band.

There is currently no published study of experimentally determined defect levels with trap concentration and capture cross section in a T2SL as of this writing. This is despite the fact that the understanding of specific impurity and defect sources and their impact on the carrier transport are crucial to optimizing the growth and fabrication of such devices. Our work so far has demonstrated T2SL devices suitable for capacitance-voltage measurements and shown that DLTS measurements are capable of resolving trap levels as
shallow as 50 meV. While we have not yet conclusively observed deep levels within the T2SL itself, our findings are important in establishing the feasibility of the DLTS measurement technique for the type-II broken band gap alignment system.
CHAPTER 6
CONCLUSION

Summary

In conclusion, we have investigated the transport properties of InAs/GaSb and InAs/InAsSb T2SL structures for detection of light in the mid- to long-wavelength IR. Our study has allowed us to characterize the minority carrier diffusion length, surface recombination velocity, diffusivity, and carrier mobility along the vertical axis in a variety of T2SL photodetector structures. We have also begun the investigation of defect energy level characterization in a mid-wavelength InAs/GaSb T2SL structure using the DLTS technique.

Using electron beam induced current, we measured the response of p-\(\pi\)-n photodiodes comprised of InAs/GaSb T2SLs to a localized excitation of carriers along the vertical growth direction. To do this, we fabricated mesa devices using a wet etch and cleaved the devices in order to expose the junction. Measurements were conducted with the electron beam of an SEM incident normal to the exposed edge. A new EBIC theoretical model was developed by altering the theory of Donolato, Bonard and Ganière [27], [28], [32] to account for the lightly p-type region in a three-region model of carrier collection probability. Using fits to the EBIC behavior at a beam energy of 15 keV, minority carrier lifetimes in the p- and n-type regions were extracted. We performed measurements on two near-identical devices, one with no interfacial treatment and another with an intentional InSb layer inserted between the InAs/GaSb superlattice layers. In addition to a
demonstrated shift in the cut-off wavelength, the use of InSb as a tool to manage strain was shown to improve the minority carrier diffusion lengths. By using values for the mobility and surface recombination velocity from published literature on similar structures, we observed a significant improvement in the diffusion length (from 500 nm to 900 nm), and thus the minority carrier lifetime (from 48.4 ns to 157 ns) of the holes in the n-region due to the interfacial treatment.

A comprehensive EBIC characterization was implemented in the course of studying an InAs/InAsSb superlattice based nBn photodetector. To successfully apply the EBIC technique to the novel device architecture, a Monte Carlo-based model of the absorbed energy within the barrier region was implemented to account for the disparity in current measured with the electron beam near the junction. Measurements and fits to multiple beam energies at 6 K allowed for determination of the hole lifetime in the absorber region and the surface recombination velocity to diffusivity ratio. When combined with measurement of the minority carrier lifetime via time-resolved photoluminescence, the vertical diffusivity and mobility were able to be determined. The hole diffusion length was measured as 750 nm and the surface recombination to diffusivity ratio was $10^6$ cm$^{-1}$. The minority carrier lifetime was measured as 200 ns, and the vertical hole diffusivity was $3 \times 10^{-2}$ cm$^2$/s. Thus we estimated a surface recombination velocity of $3 \times 10^4$ cm/s. The use of additional experimental techniques (TRPL) to determine the carrier lifetime allows the EBIC measurement to provide key information in determining the vertical transport properties of a detector. Together, these approaches give a full picture of the carrier dynamics and transport behavior of T2SL materials.
A device design for characterizing defects in low band gap T2SL structures was proposed. Using a bulk GaSb p⁺-n junction with a thin T2SL layer inserted in the n-type region close to the junction, the device leakage current can be reduced to allow for capacitance-voltage measurements. We successfully obtained capacitance vs. voltage data for the proposed design, and we compared the data to results from a control sample without T2SL, showing a clear difference in apparent carrier densities as the depletion region was swept through the T2SL region. We estimated an electron concentration of $3 \times 10^{15}$ cm⁻³ within the T2SL. Using DLTS measurements, a trap signature of 50 meV was observed in both the control and T2SL samples, with varying capture cross sections, likely due to bulk point defects in the GaSb. Further testing is necessary to determine whether defects within the T2SL can be isolated, or if a different device design is necessary.

**Future Work**

T2SL devices are rapidly improving, in terms of both design and overall material quality. As the carrier lifetimes demonstrated in these structures increase by orders of magnitude, the EBIC measurement technique is likely to encounter issues with localized concentrations of free carriers that exceed the low-level injection threshold. At such a point, the reciprocity theorem solution to the diffusion equation used by Donolato [28] may no longer be applicable, as carriers generated by the beam are screened by each other. Thus, a novel theoretical approach to the EBIC current in such a device may be necessary to keep pace with advances in growth technology. With this in mind, having demonstrated the ability to characterize the vertical mobility, the potential for study of the temperature-dependent vertical transport to better understand the underlying processes will be critical.
in improving detector performance in the future. As new devices are pushed towards room
temperature operation, the EBIC measurement will remain relevant as a means of studying
this behavior.

Further work on the DLTS characterization of the samples described in CHAPTER 5 can potentially yield interesting results. An investigation of the DLTS spectra as a function of different filling and measurement voltages can help to isolate the emission of electrons to the T2SL region. The use of etching to isolate mesas without exposing the T2SL, and ohmic contacts to the GaSb substrate with AuGeNi, can help improve the device performance and allow for better resolution of the capacitance transients. Additionally, the study of the response as a function of filling pulse width can allow for the characterization of trap concentrations within the sample [61].
REFERENCES


