III-V Bismides as a new Heterojunction Material System for Electronic Devices

by

Zefram Dael Marks

B.S., Santa Clara University, 2009

M.S., University of Colorado Boulder, 2011

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy Electrical Engineering

Department of Electrical, Computer, and Energy Engineering

2014

This thesis entitled:
III-V Bismides as a new Heterojunction Material System for Electronic Devices
written by Zefram Dael Marks
has been approved for the Department of Electrical, Computer, and Energy Engineering

_____
Bart Van Zeghbroeck

_____
Garret Moddel

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Marks, Zefram Dael (Ph.D., Electrical Engineering)

III-V Bismides as a new Heterojunction Material System for Electronic Devices

Thesis directed by Prof. Bart Van Zeghbroeck

Incorporating bismuth into epitaxially grown GaAs layers produces the alloy $GaAs_{1-x}Bi_x$. This new material system shifts the band gap of GaAs down by approximately 88 meV/Bi%, while maintaining a small lattice mismatch of less than 0.25 % for a 200 meV to 300 meV band gap shift. This material has many potential applications in optical and electron devices. In this work the material is studied for use in device applications, specifically heterojunction bipolar transistors with a narrow band gap GaAsBi base layer. The performance of this device is simulated to find its maximum potential gain and frequency of operation in the X-band at devices sizes of 0.5 µm for < 2.5 % bismuth alloying. P-N and HBT devices are fabricated to characterize material quality and HBT performance. Loss mechanisms are studied to improve future devices in the GaAsBi material system.

Dedication

To my fiancé, Hope Sheffield.

Acknowledgements

Graduating with a Ph.D. and writing a thesis seemed like a far away and impossibly difficult task when I arrived at CU Boulder almost five years ago. It still seems like that now, but I actually have a finished document and people are calling me “doctor.” I’d first like to thank my advisor, Bart Van Zeghbroeck. Bart took me on as a research assistant when I first entered the program, allowing me to start researching right away. He has always been there to share his expertise, provide direct support and feedback in research, and to ensure continuous funding to support my research. His work ethic and attitude to life is something I hope to emulate in my own life. ‘Obrigado’ for our NNIN iWSG trip to Brasil!

I want to thank my committee members both for their support at my comprehensive exam and final defense, and for their guidance through my graduate school career. Kirstin Alberi, who has always been open and helpful collaborator at NREL,. Won Park, whose classes have proved very rewarding and for being there for any questions I might have. To Minhyea Lee for serving on my committee. And to Garret Moddel, whose classes and conversations have always challenged me to think deeper than I otherwise would have done.

I also want to give a huge thank-you to my collaborators and friends at the National Renewable Energy Laboratory. Kirstin Alberi again for her help, accessibility, and openness. To Aaron Ptak and Daniel Beaton for growing sample after sample for me throughout this work. I hope you have learned as much as I have from all those (and future!) material growths. I’d also like to thank Angelo Mascarenhas for making all of my NREL collaborations possible. I would also like to thank Ryan

Lewis and Tom Tiedje at the University of British Columbia for their material growth and bismide expertise.

Grad school would not have been the same without my fellow grad students and colleagues. Thank you to my group-mates Josh Ford, Tzu-Min Ou, and Ian Haygood, researcher Tomoko Borsa whose has been there since the beginning, other to my fellow grad students for being there to talk and hang-out, and for sharing this experience with me. Thanks as well to Ryan Brow and Jan Van Zeghbroeck for their non-stop support of the Colorado Nanofabrication Laboratory and everyone's continued research. They have always being the line between a functional lab and broken equipment chaos.

Finally, I want to thank all my friends and family for their support. Most of all to my fiancé, Hope Sheffield. She has been there non-stop to help me during good times and bad, and support me in finishing this work. Your love has helped me through this more than anything else.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

1.1 III-V Compound Semiconductors for Electronic Devices

III-V compound semiconductors consist of combinations of equal amounts of elements from the column groups IIIA (boron family) and VA (nitrogen family), as in figure 1.1. Compounds made of just two elements are called binary (e.g. GaAs), of three are called ternary (e.g. $Al_xGa_{1-x}As$), and of four elements are called quaternary (e.g. $Ga_xIn_{1-x}As_yP_{1-y}$) where $0 \leq x, y \leq 1$. The huge range of different compounds and their intermediaries that can be made from these elements makes it possible to select compounds from a wide parameter space of band gaps, lattice spacings, and other material properties. One of the biggest advantage of III-V semiconductors is the ability to grow crystalline thin films of heterogeneous materials on top of each other, a process called **epitaxy**. The abrupt material change at these interfaces are called **heterojunctions**. This allows the band gap of a semiconductor to vary spatially across length scales of nanometers to micrometers. This band gap engineering technique can be used to control the flow of electrons and holes in a device to a degree not possible in homogeneous materials. For example, a thin layer of a narrower band gap semiconductor sandwiched between wider band gap semiconductors can make a quantum well to confine carriers. Another example is regular alternating layers of different band gaps to form super-lattices where the carrier wave-function undergoes Bragg diffraction and behaves in novel ways inside the composite material.

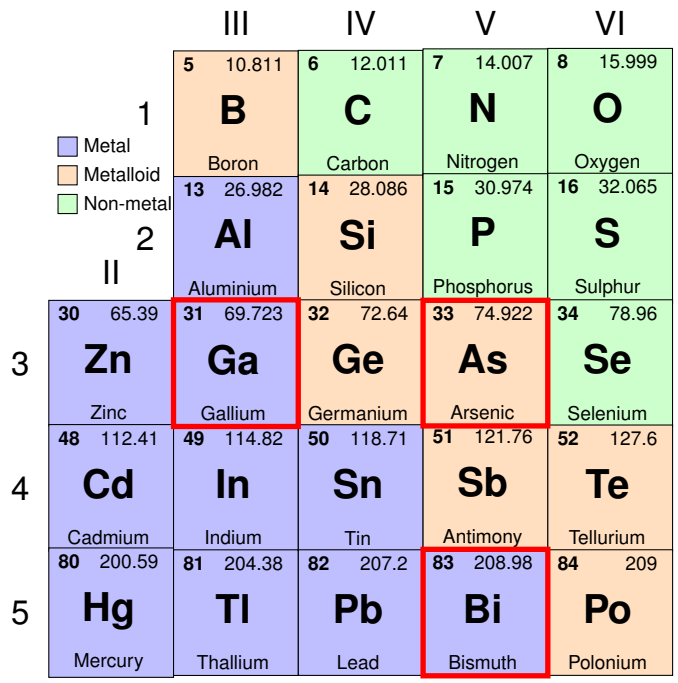


Figure 1.1: Group II through VI of the periodic table of the elements. Ga, As, and Bi highlighted.[1]

III-V compound semiconductors also have several advantages over elemental semiconductors such as silicon. Some of these semiconductors have extremely high carrier mobility (14,000 $\mathrm{cm}^2/(\mathrm{V\,s})$ for undoped InGaAs [3]) and transistors with cut-off frequencies of ~ 1 THz have been reported [4]. Many III-V semiconductors have a direct band gap. That is, in the bandstructure the valence band maximum (VBM) and conduction band minimum (CBM) exist at the same point in momentum-space, usually where $\vec{k} = 0$, called the Γ-point. In contrast, silicon has an indirect band gap, where the VBM and CBM are at different points in momentum-space. This means that in a direct band gap semiconductor an electron and hole recombining across the band gap does not need a change in momentum, just a change in energy. Thus direct band gap semiconductors have a much higher probability of producing a photon (which has very low momentum) from an electron-hole recombination. In silicon the probability is much lower since the electron and hole would need an additional interaction with a large momentum (quasi)-particle, such as a phonon. Thus a larger fraction of the carrier

[1] This figure derived from work licensed under the LaTeX Project Public License (`www.latex-project.org/lppl.txt`). Original from `TeXample.net` by Ivan Griffin.

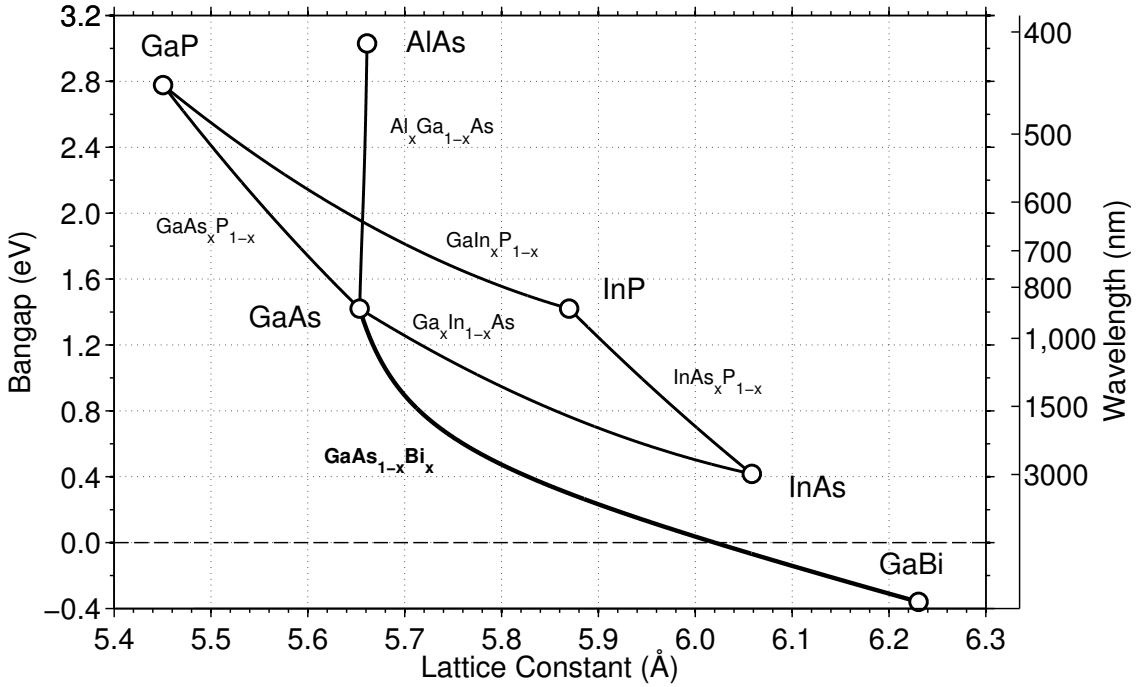


Figure 1.2: Band gap versus lattice-constant for various III-V compound semiconductors, including the GaAs/GaAsBi system [5].

recombination in III-V semiconductors is due to radiative transitions (ΔE in the form of a photon) as opposed to non-radiative transitions (ΔE in the form of phonon lattice vibrations). For this reason many optoelectronic devices such as lasers are made using III-V compound semiconductors.

1.1.1 Bismide as a III-V ternary compound semiconductor

Incorporating dilute amounts of substitutional bismuth into the epitaxial growth of GaAs forms an alloy between GaAs and the theoretically predicted semi-metal GaBi to create $GaAs_{1-x}Bi_x$ [6]. The growth and characterization of this material system has attracted study in recent years due to the alloy's novel electronic, optical, and spin properties [1, 7, 2, 8]. The growth of dilute bismides uses atomic concentrations typically between 0 % to 5 %, and up to 12 % have been reported [9, 10]. Small atomic concentrations of bismuth cause large band gap shifts in the electronic structure of GaAs. This is analogous to dilute $GaAs_{1-x}N_x$ alloys, which also have a large band gap shift for small amounts of nitrogen. In GaAsBi alloys small amounts of bismuth incorporation causes spin-orbital

splitting only in the valence band, causing the VBM to move up in energy and decrease the band gap [11]. Whereas in GaAsN the same effect causes the conduction band energy to decrease, which reduces the total band gap. The incorporation of either dilute nitrogen or bismuth has been reported experimentally to reduce the electron and hole mobility, respectively [11].

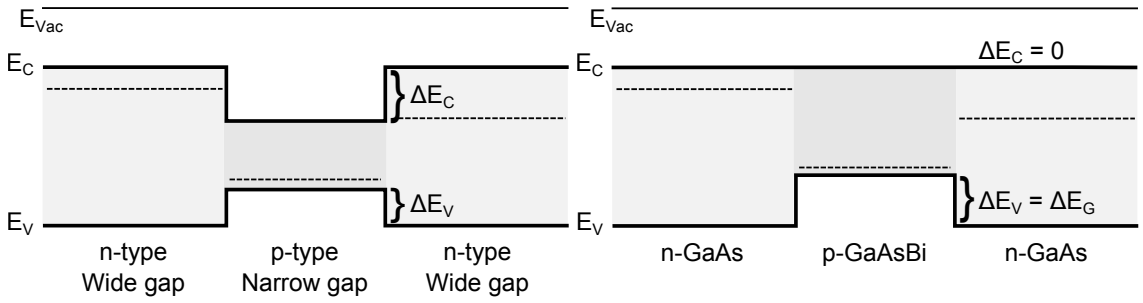
Figure 1.2 shows the band gap versus lattice-constant parameter space of many III-V compound semiconductors, including $GaAs_{1-x}Bi_x$. GaAsBi alloys open up a new area of this parameter space with band gaps from 1.4 eV to 0.8 eV or smaller with a small mismatch ($< 1\,\%$) to GaAs. This band gap range is also accessible with InGaAs, but only for compositions lattice-matched to InP substrates. At low bismuth concentrations GaAsBi will cause a linear decrease in the band gap of GaAs of approximately 88 meV/Bi% [12].

1.1.2 The GaAs/GaAsBi material system

Because only dilute amounts of bismuth are needed for heterostructure electronic devices, most of the properties of GaAsBi are expected to be similar to GaAs, except for some key properties such as the band gap. Bismuth atomic concentrations only need to be around 1 % to 5 % to achieve band gap changes of ∼100–400 meV necessary for most electronic devices. The lattice-constant mismatch to GaAs at these concentrations is between 0.1–0.5 %—an acceptable range for typical epitaxial layer thicknesses.

Because bismuth tends to surface segregate at typical GaAs growth temperatures (∼ 450 °C to 600 °C) the growth of GaAsBi films must be done at temperatures below 350 °C to get reasonable bismuth incorporation [13]. GaAs itself will grow at these low[2] temperatures, but the quality will suffer mostly from a reduction in the minority carrier lifetime and mobility from crystalline defects. This reduction in carrier transport has been observed in molecular beam epitaxy (MBE) grown GaAsBi films. The carrier lifetime is known to be reduced, but the exact amount depends on the specific process conditions, as well as post-growth annealing. Nearly all the reported research on

[2] This is called LT-GaAs, or low-temperature gallium arsenide. This material is sometimes used in high speed photoconductive switches where the short lifetime can be advantageous.



(a) Typical (type I) HBT

(b) GaAs/GaAsBi HBT

Figure 1.3: Idealized flat band diagram of both typical (type I) and GaAs/GaAsBi showing band discontinuities.

GaAsBi shows a reduction of the hole mobility in GaAsBi films compared to GaAs [14, 15]. The electron mobility however does not appear to be nearly as affected—especially at the low bismuth concentrations used in electronic devices [16]. However there is evidence that the minority carrier lifetime and hole mobility reduction are not intrinsic to the bismuth incorporation but just a byproduct of the low-temperature growth, and that certain growth conditions and annealing steps could reduce or prevent this reduction [17].

An important feature of the GaAs/GaAsBi material system is the unique property of the heterojunction band discontinuity. In most material systems the heterojunction between a wide and narrow band gap semiconductor has both a valence and conduction band discontinuity, as shown in figure 1.3a. This causes an energy barrier for carriers traveling from the narrow to wide band gap semiconductor. This barrier is desirable in a HBT for one type of carrier but a disadvantage for the other type. In the case of an N-P-N HBT the hole current from the base into the emitter should be suppressed, so the energy barrier at the heterojunction is desirable. For the electron current, the barrier from the wide gap emitter to the narrow gap base will not impede the flow of electrons. But this same barrier for the electrons diffusing from the base into the collector could reduce the total current, and thus the gain. The GaAs/GaAsBi material system has the advantageous property that the band gap distortion from the bismuth is entirely in the valence band. That is, there is only a discontinuity in the valence band ($\Delta E_{\mathrm{C}} = 0$, $\Delta E_{\mathrm{G}} = \Delta E_{\mathrm{V}}$) in a GaAs/GaAsBi HBT, as shown

in figure 1.3b [18]. This property of the GaAs/GaAsBi system and its consequences for electron devices will be explored more in chapter 2.

1.2 Applications of III-V Compound Semiconductors

1.2.1 Heterojunction bipolar transistors

A major III-V compound semiconductor device is the heterojunction bipolar transistor (HBT). Like bipolar junction transistors(bipolar junction transistors), HBTs consist of two P-N junctions that share a common region to form three layers: The emitter, base, and collector. Using heterojunctions to improve bipolar transistors was originally postulated by Shockley [19], then studied in detail by Kroemer, who derived the gain of a HBT [20, 21]. HBTs are operated by applying a bias between the base-emitter P-N junction, which causes electrons to be injected into the base and diffuse towards the collector junction. This allows the emitter-collector current to be modulated by the bias voltage across the base-emitter junction. Changing the base-emitter junction voltage will induce a hole flux into the emitter and an electron flux into the base. This hole current, along with other currents from effects such as recombination, makes up the total base current while only the electrons diffuse towards the collector resulting in a collector current. The ratio of total collector to base current is called β, which is the DC current gain of the transistor and an important figure of merit. The difference between HBTs and bipolar junction transistors is that while bipolar junction transistors use multiple P-N junctions (N-P-N or P-N-P layers) [3] with different doping levels, HBTs additionally use junctions of dissimilar materials to affect the flow of carriers. This gives additional control over the carrier flow through band gap engineering of the heterojunction layers. In the case of an N-P-N HBT the n-type emitter has a larger band gap material than the base.

The gain of a HBT, ignoring recombination and transport through the base, is given by [21]

$$\beta = \frac{N_{\mathrm{E}}}{N_{\mathrm{B}}} \frac{D_{\mathrm{n,B}}}{D_{\mathrm{p,E}}} \frac{w'_{\mathrm{E}}}{w'_{\mathrm{B}}} \exp\left(\frac{\Delta E_{\mathrm{G}}}{k_{\mathrm{B}} T}\right) \tag{1.1}$$

[3] All HBTs and bipolar junction transistors in this text will implicitly be assumed to be N-P-N, unless otherwise noted. For a P-N-P transistor the electron and holes, as well as the sign of the currents and voltages, will be switched.

where N_E and N_B are the doping concentrations in the emitter and base, $D_{n,B}$ and $D_{p,E}$ are the minority carrier diffusion coefficients for electrons in the base and holes in the emitter, and w'_B and w'_E are the effective widths of the base and emitter respectively. The currents in a HBT are due to the injected minority carrier concentrations across the junction, which are proportional to the thermal equilibrium doping concentration and intrinsic concentration:

$$n_0 p_0 = n_i^2 \rightarrow n_{p0} = \frac{n_i^2}{N_A} \text{ and } p_{n0} = \frac{n_i^2}{N_D} \tag{1.2}$$

Reducing the band gap will increase the intrinsic carrier concentration n_i and increase the amount of injected minority carriers. The emitter efficiency (the ratio of electron current to total current through the emitter), which is part of the overall gain, depends on the minority carrier concentration and thus the doping levels and intrinsic carrier concentration in the base and emitter. The smaller the hole current component compared to the electron current the larger the emitter efficiency and gain will be. By changing the band gap in the base compared to the emitter the intrinsic carrier density, and therefore the minority carrier density will change and affect the ratio of electron to hole current independently of the doping. This allows a greater degree of freedom in device design since the gain can be controlled through the doping as well as the band gap. As seen in equation (1.1), a change in band gap or temperature between the base and emitter will exponentially affect the gain, while the gain will only change linearly with doping. This means that the base can be doped much higher than the emitter for the same amount of gain. A higher doped base has lower resistance, and will be able to operate at higher frequencies. More information about the operation of bipolar junction transistors and HBTs will be covered in chapter 3.

In digital logic, the dominant technology is the silicon complementary metal-oxide-semiconductor field effect transistor (MOSFET). The insulated gate as well as the complementary structures will produce (ideally) zero steady-state current, which is best for logic operations. For analog electronics such as radio frequency (RF) signal amplification for RADAR and wireless communication there are many more material systems and devices to choose from, including discrete or integrated III-V HBTs [22, 23, 24, 25].

The first material system used to make a HBT was GaAs/AlGaAs, which consists of an $Al_xGa_{1-x}As$ epitaxial layer grown on a GaAs substrate. These first GaAs/GaAsBi heterojunctions were grown using liquid phase epitaxy (LPE) in 1963 [26], and the first demonstrated transistor in this system was shown in 1972 [27]. The lattice mismatch between GaAs and AlGaAs is less than 0.15 % across the entire composition range [28] (see figure 1.2) so high quality heterojunctions can be made. This system uses a wide band gap AlGaAs emitter and a GaAs base, collector, and substrate. In general, the HBT active device layers are epitaxially grown on a wafer such as GaAs, with different materials of similar lattice-constant and doping as planar stacks. The entire active device layer is typically no more than a few micrometers thick, with individual layers as small as tens of nanometers. More recently, high performance HBTs have been fabricated in the InP/GaInAs and GaAs/GaInP [29] material system, with record performance of up to 1.1 THz demonstrated using a InGaAs base [30]. Although the high mobility of InGaAs allows for a much faster transistor, an InP substrate must be used in order to lattice-match the epitaxial layers to the substrate. Fabrication on this substrate increases production costs due to the smaller and more expensive wafer as well as a decreased yield from the brittle nature of InP compared to GaAs [31]. In addition to this increased cost, the InP/GaInAs system has ∼40 % of its total band gap discontinuity in the conduction band, for the lattice-matched composition of $Ga_{0.47}In_{0.53}As$ [32]. The GaAs/GaInP system also has a significant conduction band discontinuity of ∼30% of the total energy gap [23]. While for N-P-N HBTs the band gap discontinuity in the valence band is advantageous for the gain, the discontinuity in the conduction band blocks the flow of electrons through the base into the larger band gap collector. This barrier can be mitigated with techniques such as graded or super-lattice layers. The GaAs/GaAsBi system does not have this complication since for dilute amounts of bismuth the band gap discontinuity is almost entirely in the valence band [11].

Chapter 2

Bismide Material System Background

2.1 History of Dilute Nitrides and Bismides

One of the first applications and the original motivation for studying dilute III-V compounds of nitrogen and bismuth was to find a new material for multi-junction solar cells. The typical high-efficiency multi-junction solar cell is epitaxially grown on a germanium wafer, due to its comparatively low cost and close lattice match to GaAs and other III-V compounds [33]. Layers of III-V semiconductor with progressively wider band gaps, but similar lattice spacing, are grown on top of each other to absorb shorter wavelength light before passing it down to the next layer. Because the top layers limit the spectrum seen by the lower layers, there is an optimum energy gap for each semiconductor layer for a given number of unique material layers [34]. However in practice exact structure is limited by what materials can be grown together with low lattice mismatch. For $N = 3$, the usual configuration is to use a stack of GaInP, GaAs (or InGaAs), and Ge [35, 36, 33]. This configuration is not the most optimal: the 0.67 eV band gap of the Ge substrate is too small compared to the other materials in the stack [37]. A third junction with a band gap closer to ~ 1 eV has a higher theoretical efficiency. Going to $N = 4$ and using a material with a band gap closer to ~ 1 eV as the third junction and a substrate of Ge as a forth junction has the potential to be a much more efficient solar cell [38]. A new material was needed to fill this role: One whose band gap was ~ 1 eV and with a low lattice mismatch to GaAs.

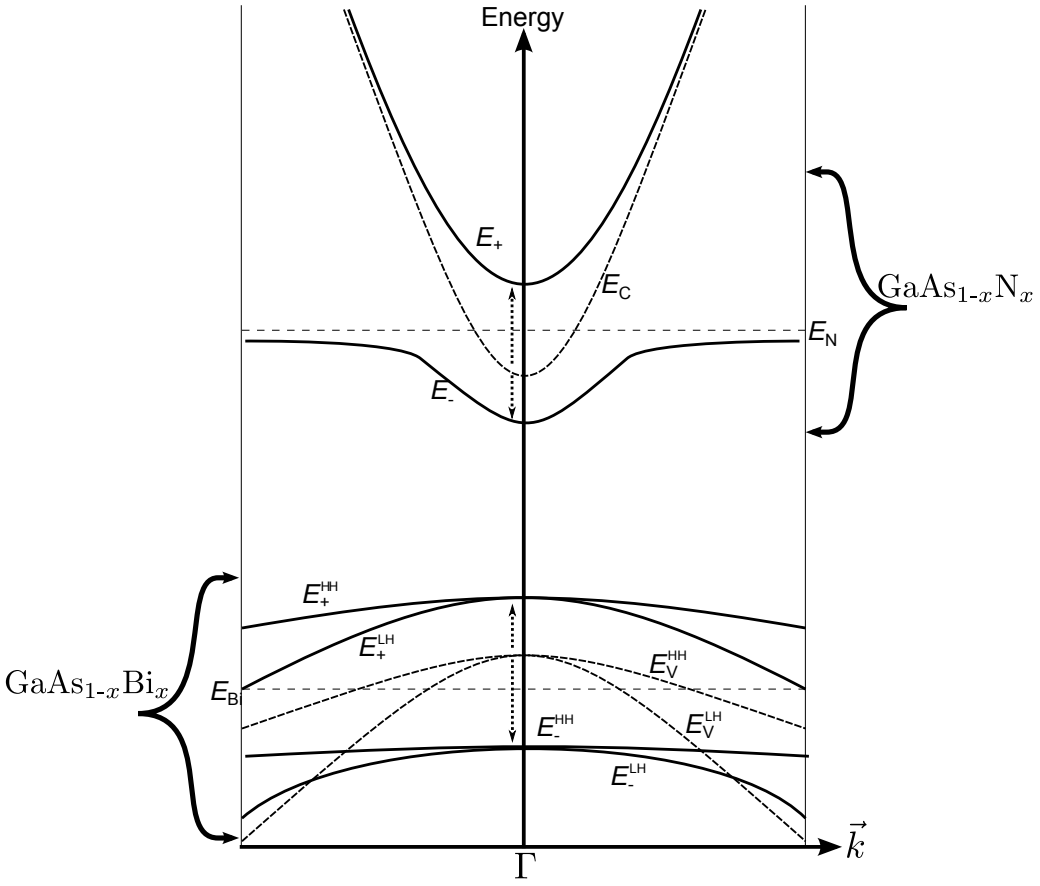


Figure 2.1: Band structure anti-crossing in GaAs valence (heavy and light hole) and conduction band from highly electronegative impurities Bi and N. The band gap narrows in both materials with $\Delta E_{\mathrm{G}} \approx \Delta E_{\mathrm{V}}$ for GaAsBi and $\Delta E_{\mathrm{G}} \approx \Delta E_{\mathrm{C}}$ for GaAsN. The effective mass of carriers will increase in both systems. The split orbit valence band not is shown. [39, 40].

GaAsN was one of the first in the class of novel materials called highly mismatched alloys (HMAs) studied for four-junction solar cells [40]. Adding dilute amounts of nitrogen to GaAs introduces a resonant nitrogen state above the conduction band that perturbs the bandstructure to cause a splitting of the conduction band energy level. A schematic of this process is shown in the top part of figure 2.1. This energy level splitting causes the conduction band to move closer to the valence band (lowering the CBM) and reduces the total band gap by $\sim$ 150–200 meV/N% [41]. Adding dilute amounts of nitrogen to GaAs primarily affects the conduction band—both in its band offset and carrier transport properties. The band gap discontinuity in a GaAsN heterojunction is almost entirely in the conduction band ($\Delta E_{\mathrm{V}} \sim 20$ meV/N% [41]). The addition of nitrogen into the

crystal structures results in a decrease in electron mobility, while the hole mobility remains largely the same [42].

Research into HMAs also included the heaviest group V element: Bismuth. Like GaAsN, and shown on the bottom of figure 2.1, the isolated bismuth impurity causes a splitting of the valence bands (heavy, light, and spin-orbit) which pushes the VBM up to reduce the total band gap. Analogously to GaAsN and as mentioned in section 1.1.2 almost all the perturbation from the incorporation of the bismuth occurs in the valence band and primarily affects the hole transport properties. The interaction with the bismuth impurities takes place in the valence band, so that the electron transport is largely unaffected at low bismuth concentrations [16]. Because of the similar physical origin of the band bending in GaAsBi compared to GaAsN the conduction band discontinuity ΔE_C is suspected to be nearly zero for low bismuth concentrations, but direct empirical evidence of this is scarce. Since the electron transport in GaAsBi films is close to that of GaAs, high performance devices are possible in this system. The simulated device potential for GaAs/GaAsBi HBTs is shown in [43], and will be explored further in chapter 3.

Another advance in HMA dilute nitrides and bismides is the $GaAs_{1-x-y}N_yBi_x$ quaternary material system. Adding the small nitrogen atom to GaAs reduces the lattice-constant, and adding the large bismuth atom will increase it. By combining both of these elements in the correct concentrations a perfectly lattice matched, narrow band gap epitaxial film can be grown on GaAs [44]. This material has especially great potential to make quantum wells for optoelectronic devices. This would make possible active optoelectronic devices like lasers that operate in the near to mid-IR wavelength region to be built on GaAs substrates, as opposed to the more expensive InP. While the single band discontinuity of either dilute nitrides or bismides is advantageous for many electronic devices, in optoelectronic devices such as lasers having a quantum well for both electrons and holes is desirable. GaAsNBi can be used to make perfectly lattice-matched quantum wells in GaAs, but at the expense of both the electron and hole mobility [44]. A possible advantage of GaAsBi, GaAsN, or GaAsNBi all have for optoelectronics is their lower sensitivity to temperature for the band gap ($-0.15\,\mathrm{meV/K}$ [45]) and refractive index ($2.47 \times 10^{-4}/\mathrm{K}$ [46]), both $\sim 1/3$ that of GaAs. In addition, both GaAsBi

and GaAsNBi have large spin-orbital splitting in the valence band due to the bismuth, which at sufficiently high concentrations could be used to suppress Auger recombination in lasers [44, 47].

2.2 Epitaxial Growth and Material Properties of GaAsBi

Molecular beam epitaxy (MBE) is used to grow the majority of GaAsBi reported in the literature, and consequently is the growth technique used to get most of the published material parameters to date. MBE is an ultra-high vacuum epitaxial process where different elemental materials or gaseous sources are heated up and evaporated onto an existing seed crystal (the substrate). For example, to grow GaAs epitaxially pure elemental sources of both gallium and arsenic are heated until they evaporate inside the ultra-high vacuum onto the GaAs seed substrate. These molecular beams will condense onto the heated substrate and either re-evaporate off, move along the surface, or adhere to the existing crystal structure to grow crystalline layers. Different impurities can also be flown in during the growth, such as beryllium or carbon as a p-type dopant, or silicon as an n-type dopant. Different group III or group V elemental species (as in figure 1.1) and dopants will flow in at specific partial pressures while the substrate is maintained at specific temperatures to grow the desired material layer. Heterojunctions can be made by changing the composition of the gaseous elements during the growth. If the crystalline structure of the epitaxial layer is too different from the substrate seed layer the epi-film will not grow as an ideal planar layer. The lattice-constant of each epitaxial layer must be close to that of the seed layer or the film will "relax" after a certain **critical thickness** and form dislocations in the crystal structure causing deep, or mid-gap energy states which are strong carrier recombination centers—and will prevent most devices from operating properly. Mismatched "metamorphic" thin films can be grown as long as they are under this critical thickness, which is a function of the percent lattice mismatch between the material layers.

Bismide films with bismuth concentrations of more than 12 % have been grown [9]. The growth temperature of GaAsBi must be lower than that of GaAs by about 200 °C to 300 °C or the bismuth atoms will desorb off the surface and not incorporate, with higher bismuth concentrations

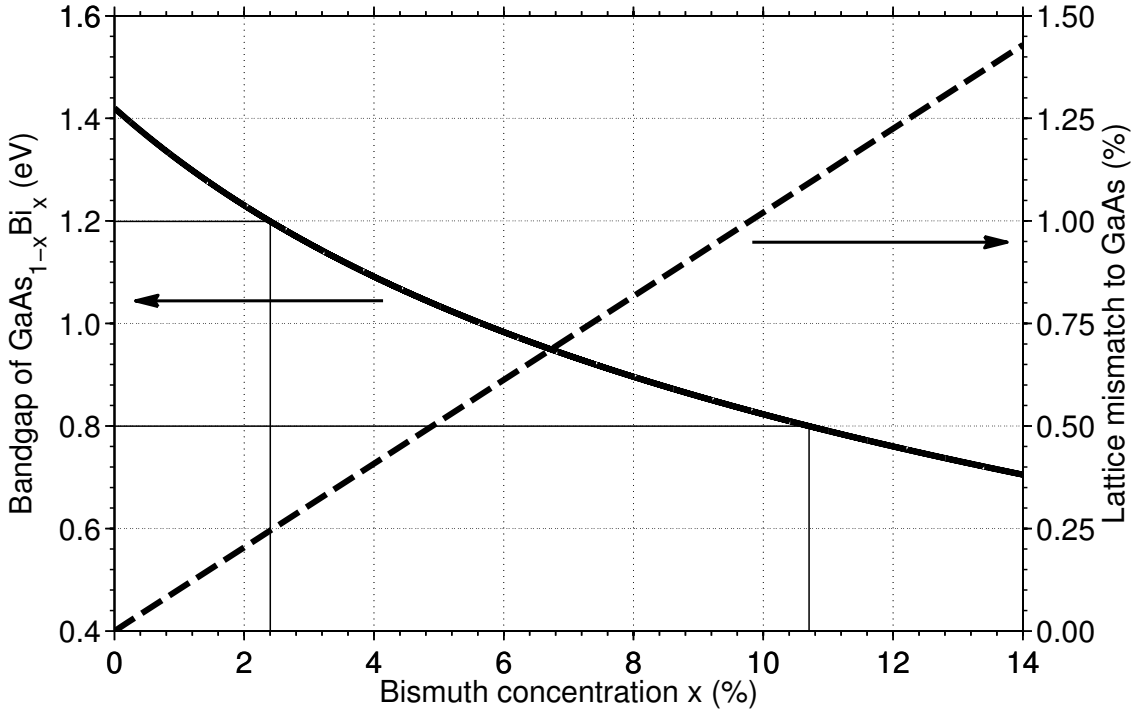


Figure 2.2: Band gap of GaAsBi and percent lattice mismatch to GaAs as a function of Bi concentration [1, 2]. The two points of interest $E_G = 1.2\,\text{eV}$ (HBTs) and $E_G = 0.8\,\text{eV}$ (telecom optoelectronics) where $x_{Bi} = 2.4\,\%$ and $x_{Bi} = 10.7\,\%$ with percent mismatch of $0.25\,\%$ and $1.1\,\%$ respectively are noted with straight lines.

requiring lower growth temperatures [9]. This reduction in temperature can introduce crystal defects which affect the electrical properties independent of the bismuth's perturbation of the bandstructure. This could be responsible for some of the deteriorated transport properties (especially for electrons) and decreased minority carrier lifetime observed in bismide materials in the literature. These reduced material properties might not be an intrinsic property of GaAsBi, but just a consequence of the growth conditions. There could exist a growth parameter window that minimize these effects [17].

GaAsBi films can be grown on GaAs despite the small lattice mismatch because only a small atomic percentage of bismuth is needed to get a large band bowing (band gap change), as seen in figure 1.2. Figure 2.2 shows the band gap and percent lattice mismatch between $GaAs_{1-x}Bi_x$ films and a GaAs substrate as a function of bismuth atomic concentration. For $E_G = 1.2\,\text{eV}$, which is the band gap target for electronic devices like HBTs, a bismuth atomic concentration of $x_{Bi} = 2.4\,\%$ is needed, with a percent lattice mismatch of only $0.25\,\%$. The formula for the composition dependent

lattice-constant for $GaAs_{1-x}Bi_x$ is found via Vegard's law and the calculated lattice-constant of GaBi to be [2]

$$a_0^{\mathrm{GaAsBi}} = (1-x)a_0^{\mathrm{GaAs}} + xa_0^{\mathrm{GaBi}} = 5.6535 + 0.577x \quad (\text{Å}) \tag{2.1}$$

and the fitted empirical band gap with compositions dependent band bowing $b(x) = \alpha/(1+\beta x)$ is [1]

$$\begin{aligned} E_{\mathrm{G}}^{\mathrm{GaAsBi}} &= xE_{\mathrm{G}}^{\mathrm{GaBi}} + (1-x)E_{\mathrm{G}}^{\mathrm{GaAs}} - b(x)x(1-x) \\ &= -0.36x + 1.42(1-x) - \frac{9.5}{1+10.4x}x(1-x) \quad (\text{eV}) \end{aligned} \tag{2.2}$$

where in both equations $0 \leq x \leq 1$. In any epitaxial thin film the mismatch in lattice-constant will limit the maximum thickness of the layer before it undergoes relaxation and defects form. Although there are many models for the critical film thickness of epitaxial layers involving energy minimization [48, 49], most real-world films differ in their actual thickness before dislocations appear due to lack of defect nucleation sites or non-equilibrium growth conditions. GaAsBi seems to exhibit few dislocations for moderate mismatch ($< 1\,\%$), with low-defect layers grown up to several hundred nanometers thick [50].

At a bismuth concentration of 2.5 % ($\Delta E_{\mathrm{G}} \approx 230\,\mathrm{meV}$) the lattice mismatch to GaAs is 0.25 %, as shown in figure 2.2. Because of the dilute amount of bismuth, and the low percent mismatch to GaAs, a model for the critical thickness h_{C} using energy balance is most accurate, as developed in [49]. Using this model the critical layer thickness of a GaAsBi epitaxial film on GaAs at $x = 2.5\,\%$ is $h_{\mathrm{C}} \sim 3\,\mu\mathrm{m} \gg w_{\mathrm{B}}$. This analysis is supported from results in the literature. Crystalline films $> 600\,\mathrm{nm}$ thick have been grown at moderate bismuth concentrations [50]. Even large mismatch films of $> 10\,\%$ bismuth concentration have been successfully grown with thicknesses between 100 nm to 300 nm [9].

The hole mobility of GaAsBi will decrease with increasing bismuth concentration. The exact mechanism of this reduction is not entirety known, but both the modification of the valence band from the energy level splitting (figure 2.1) and scattering off isolated Bi and Bi-Bi complexes are thought to be the primary causes [51]. P-type mobility is most affected by bismuth incorporation,

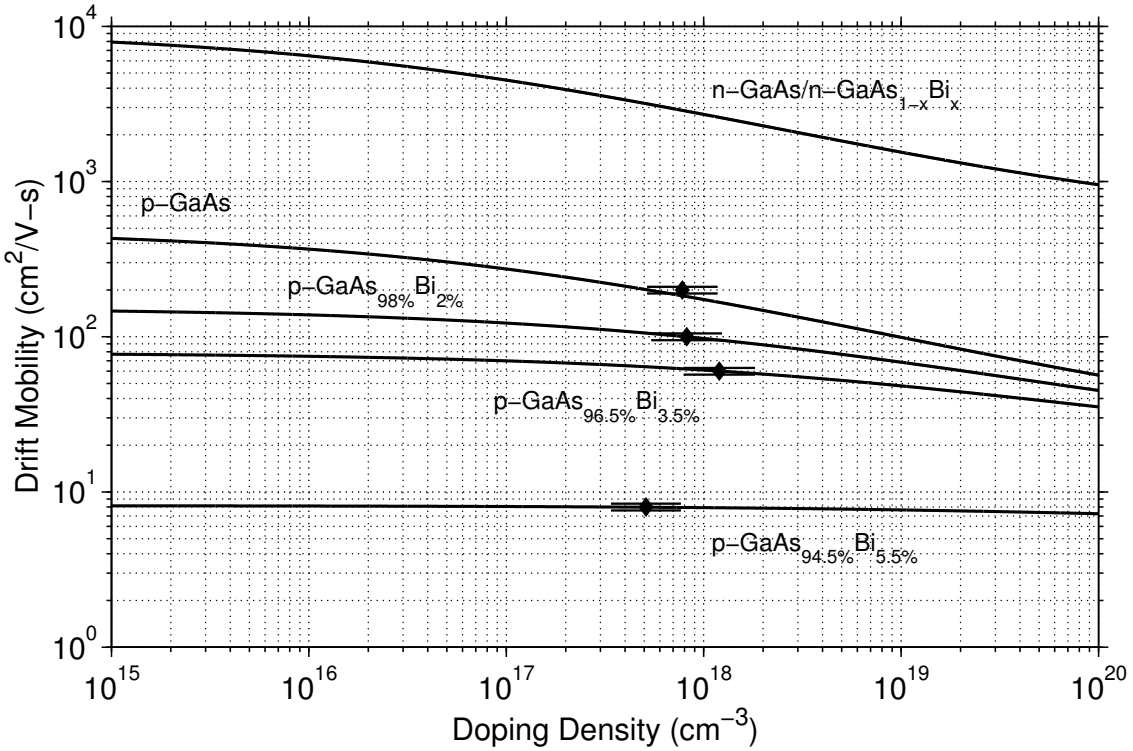


Figure 2.3: Empirical model for electron and hole drift mobility versus impurity doping concentration for n-type and p-type GaAs [3] and n-type and p-type GaAsBi at various Bi concentrations. Black markers are experimental values used for fitting [8].

which is consistent with the primary electronic bandstructure modification occurring in the valence band and the suspected Bi-complex energy levels being a few hundred meVs above the valence band [52]. N-type electron mobility is only slightly affected by the incorporation of Bi, especially at higher concentrations, but this may just be an artifact of the lower growth temperature of the substrate [17].

Figure 2.3 shows the mobility of electrons and holes in GaAsBi at various bismuth concentrations. These mobilities were calculated from the mobility of bulk GaAs using the model

$$\frac{1}{\mu_{\mathrm{p}}(T,N,x)} = \frac{1}{\mu_{\mathrm{p}}^{\mathrm{GaAs}}(T,N)} + \frac{1}{C_{\mathrm{Bi}}(x)} \tag{2.3}$$

where $\mu_{\mathrm{p}}^{\mathrm{GaAs}}(T,N)$ is the mobility of bulk GaAs as a function of temperature and doping concentration and $C_{\mathrm{Bi}}(x)$ is the effect of the Bi alloying on the mobility as a function of Bi atomic concentration x. $C_{\mathrm{Bi}}(x)$ is assumed to be independent of temperature and dopant concentration for

dilute Bi concentrations. The coefficient $C_{\mathrm{Bi}}(x)$ is due to the scattering from single bismuth atoms distributed randomly in the GaAs host lattice. From kinetic theory the scattering from single Bi atom sites is [8]

$$C_{\mathrm{Bi}} = \frac{q}{m_{\mathrm{h}}^{*} v_{\mathrm{th}} \sigma_{\mathrm{Bi}} n_{\mathrm{Bi}}} \tag{2.4}$$

where q is the electron charge, v_{th} is the thermal velocity of holes, σ_{Bi} is the scattering cross-section of Bi atoms, and n_{Bi} is the density of Bi atoms. From [53] the scattering cross-section due to an isolated atom in a dilute alloy is

$$\sigma_{\mathrm{Bi}} = \frac{\pi}{4} \left(\frac{m_{\mathrm{h}}^{*}}{2\pi\hbar^{2}} \right)^{2} \left[\frac{dE_{\mathrm{G}}}{dx} \right]^{2} a_{0}^{6} \tag{2.5}$$

where a_0 is the lattice-constant of unalloyed GaAs and $\frac{dE_{\mathrm{G}}}{dx}$ is the change in band gap with increased Bi concentration, which is approximately constant for low ($x < 5\,\%$) Bi contractions. With the assumption that σ_{Bi} is constant for a given alloy concentration and n_{Bi} is linearly proportional to x then fitting the data from [8]

$$\frac{1}{C_{\mathrm{Bi}}} \approx (0.315{,}98x - 3.736 \times 10^{-4}) \propto x \tag{2.6}$$

which are some of the best hole mobility data for GaAsBi found in the literature, to produce the mobility curves seen in figure 2.3. The hole mobility of bulk GaAs is from the empirical formula [3]

$$\mu_{\mathrm{p}}^{\mathrm{GaAs}}(T, N) = 20 + \frac{491.5 \times \left(\frac{300}{T} \right)^{2.2} - 20}{1 + \left(\frac{N}{1.48 \times 10^{17} \left(\frac{T}{300} \right)^{3}} \right)^{0.38}} \tag{2.7}$$

2.3 Comparison of GaAsBi to Other III-V Compound Semiconductors

Table 2.1 lists the material properties of several binary and ternary III-V compound semiconductors, the same that are shown in figure 1.2. GaBi data is from [54], while GaAs, InAs, and InP data is from [55], and AlAs and GaP data is from [56]. The GaAsBi band gap expression is from [1] and the lattice constant expression is from [2]. The AlGaAs and InGaAs composition expressions are from [55]. $\mathrm{In}_{x}\mathrm{Ga}_{1-x}\mathrm{As}$ with $x = 0.53$ is the composition lattice-matched to InP.

Table 2.1: Material properties of various III-V material systems. All mobilities are for intrinsic materials. "lh" = light holes. "—" indicates no data found in literature.

	GaAs	**GaBi**[4]	**AlAs**	**InAs**	**GaP**	**InP**
$E_G^{300\,K}$ $\left(\text{eV}\right)$	1.424	−1.45	2.36	0.354	2.26	1.34
a_0 $\left(\text{Å}\right)$	5.653	6.324	5.661	6.058	5.451	5.869
$\mu_n^{300\,K}$ $\left(\frac{\text{cm}^2}{\text{V}\cdot\text{s}}\right)$	8500	—	180	33,000	110	4600
$\mu_p^{300\,K}$ $\left(\frac{\text{cm}^2}{\text{V}\cdot\text{s}}\right)$	400	—	100	450	75	150
m_n^*/m_0	0.063	—	0.11	0.022	0.82	0.077
$m_p^{*\text{lh}}/m_0$	0.076	—	0.22	0.026	0.60	0.12
ϵ_S/ϵ_0	12.9	—	10.1	15.15	11.1	12.56

(a) Binary materials [54, 55, 56].

	$GaAs_{1-x}Bi_x$	**$Al_xGa_{1-x}As$**	**$In_xGa_{1-x}As$**
$E_G^{300\,K}$ $\left(\text{eV}\right)$	$-0.36x + 1.42(1-x)$ $-\frac{9.5}{1+10.4x}x(1-x)$	$\begin{cases} 1.424 + 1.247x & (x < 0.45) \\ 1.9 + 0.125x + 0.143x^2 & (x > 0.45) \end{cases}$	0.75 ($x = 0.53$)
a_0 $\left(\text{Å}\right)$	$5.653 + 0.557x$	$5.653 + 0.0078x$	$6.058 - 0.405x$
$\mu_n^{300\,K}$ $\left(\frac{\text{cm}^2}{\text{V}\cdot\text{s}}\right)$	—	$\begin{cases} 8000 - 22{,}000x + 10{,}000x^2 & (x < 0.45) \\ -255 + 1160x - 720x^2 & (x > 0.45) \end{cases}$	13,800 ($x = 0.53$)
$\mu_p^{300\,K}$ $\left(\frac{\text{cm}^2}{\text{V}\cdot\text{s}}\right)$	Eqs. (2.3) to (2.7)	$370 - 970x + 740x^2$	—

(b) Ternary materials [1, 2, 55].

[4] GaBi has never been grown and its properties are not well know and have only been obtained from simulations [54]. GaBi is suspected to be an indirect, negative band gap semi-metal.

Chapter 3

Modeling and Simulation

3.1 Heterojunction Bipolar Transistor Background

The first solid-state transistor was made at Bell Laboratories in 1947 [57]. This first transistor used gold point-contacts on germanium to create a bipolar junction. Later designs used diffused "sandwich" junctions to form the P-N-P or N-P-N junctions to allow bipolar transistor action. The BJT was the primary transistor used in production applications for transistor-transistor logic (TTL) circuits, power switches, and amplifiers until the MOSFET became prominent for logic circuits. In 1982 Kroemer proposed that crystal growth technology had advanced enough to make devices using **heterojunctions**, which are the physical and electrical junction between two semiconducting materials with different electronic properties, primarily their band gap [21].

3.1.1 Bipolar junction transistor operation

From a circuits perspective bipolar junction transistors operate by modulating the current through the **collector** and **emitter** terminals with the current through the **base** terminal. In a regular BJT there are two P-N junctions which share a common region. In "forward active" transistor mode one junction is reverse biased (the base-collector junction) and the other is forward biased (the base-emitter junction). When the base-emitter junction is forward biased both electrons and holes will diffuse across the junction. In an N-P-N BJT holes will flow into the emitter from the base

($J_{p,B}$), and electrons into the base from the emitter ($J_{n,E}$). Ignoring recombination in the base and depletion regions and other non-ideal effects $J_B \sim J_{p,B}$ and $J_E \sim J_{n,E} \sim J_{n,C}$, as shown in figure 3.1. Transistor action occurs because the device is designed so that the base-emitter hole current is much smaller than the electron current $J_{p,B} \ll J_{n,E}$, and therefore $\beta \triangleq J_C/J_B$ is large and the device has current gain.

To achieve this asymmetry between $J_{p,B}$ and $J_{n,E}$ which is needed for device gain consider the basic "short channel" ($w'_{p/n} \ll L_{p/n}$) P-N diode equation (Ch. 2 of [56]):

$$J_p = q\frac{D_p p_{n0}}{w'_n}\left[\exp\left(\frac{V}{k_B T}\right) - 1\right] = q\frac{D_p n_i^2}{w'_n N_D}\left[\exp\left(\frac{V}{k_B T}\right) - 1\right] \tag{3.1}$$

$$J_n = q\frac{D_n n_{p0}}{w'_p}\left[\exp\left(\frac{V}{k_B T}\right) - 1\right] = q\frac{D_n n_i^2}{w'_p N_A}\left[\exp\left(\frac{V}{k_B T}\right) - 1\right] \tag{3.2}$$

where $w'_{p/n}$ is the effective widths of the quasi-neutral p-type and n-type regions (layer width minus depletion region width), n_i is the intrinsic carrier concentration, $D_{p/n}$ is the minority carrier diffusivity, and $n_{p0} = n_i^2/N_A$ is the minority electron concentrations in the p-type region and $p_{n0} = n_i^2/N_D$ is the minority hole concentrations in the n-type region. The two main design parameters in BJT device design are the layer width and the doping concentration. Since the gain of a BJT is ideally $\beta = J_C/J_B \approx J_n/J_p$ for the base-emitter P-N junction the gain is:

$$\beta = \frac{D_n}{D_p}\frac{w'_n}{w'_p}\frac{N_D}{N_A} = \frac{D_B}{D_E}\frac{w'_E}{w'_B}\frac{N_E}{N_B} \tag{3.3}$$

So to get gain in a BJT generally the base must be doped much lower than the emitter. This is because there is a much wider range of control over the doping concentration compared to layer width or the diffusion constant, which is a material property. Equation (3.3) also demonstrates why N-P-N are preferred over P-N-P devices. The critical material property for both the gain and the transit time is the minority carrier diffusion constant (or mobility) in the base. In almost all materials the n-type mobility is larger than the p-type mobility. So a N-P-N device can benefit from the larger n-type minority carrier mobility.

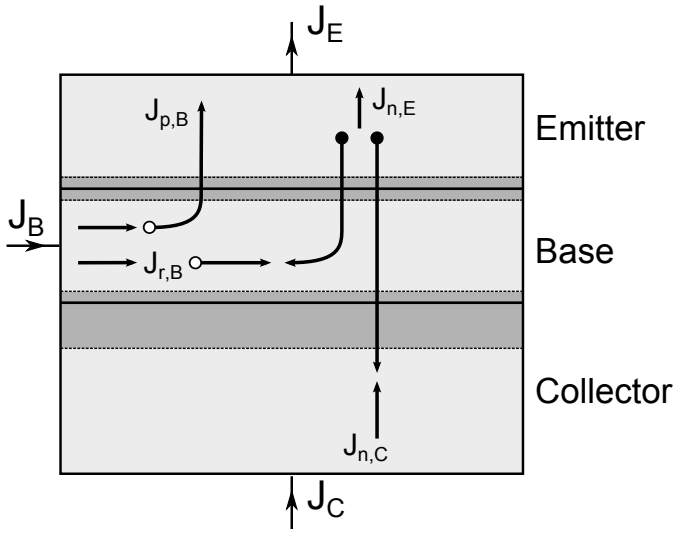


Figure 3.1: Simplified diagram of hole and electron currents in a BJT or HBT.

3.1.2 Heterojunction bipolar transistor operation

There is a trade-off in homojunction bipolar junction transistors between gain and speed. The resistor–capacitor (RC) time-constant of a BJT is greatly affected by the access resistance to the base. This base resistance should be as low as possible for maximum RF frequency operation. To get a lower base resistance the layer must be doped to a higher carrier concentration or the layer grown to a larger thickness, or both. However, increasing the doping will lower the maximum operating voltage of the transistor by reducing the depletion region thickness, and increasing either the doping or layer thickness will directly reduce the gain, as in equation (3.3). An increase in the base doping will reduce the base-emitter depletion region into the base quasi-neutral region, and reduce how high of a reverse bias can be applied to the base-collector junction before the electric field exceeds the breakdown field in the material. A base doping that is too low is also undesirable, since the base resistance will increase (lowering the maximum frequency of operation) and the base-emitter depletion region will extend more into the base quasi-neutral region, reducing how high of a reverse bias can be applied to the base-collector junction before the quasi-neutral region disappears—leading to an effect called punch-through. The few free parameters available in BJT design limit the achievable performance.

Heterojunction materials provide a desirable additional degree-of-freedom. At its most basic level a HBT has a heterojunction between the base and emitter. The base should have a narrower band gap than the emitter. Looking back at equations (3.1) and (3.2) in a heterojunction the intrinsic carrier concentration n_i will be different between the two junctions. n_i is the amount of thermally excited carriers across the band gap in an undoped semiconductor. At 0 K all the energy bands below the VBM will be filled, and everything above will be empty. At any finite temperature some free carriers will be excited. The amount of free carriers depends on the band gap, and therefore will be different in each layer of a heterojunction device.

The intrinsic carrier concentration is the integral of the semiconductor density of states [56]

$$g_\mathrm{C}(E) = \frac{\sqrt{2}}{\pi^2}\frac{m_\mathrm{e}^{*3/2}}{\hbar^3}\sqrt{E - E_\mathrm{C}} \qquad \left(E \geq E_\mathrm{C}\right) \tag{3.4}$$

$$g_\mathrm{V}(E) = \frac{\sqrt{2}}{\pi^2}\frac{m_\mathrm{h}^{*3/2}}{\hbar^3}\sqrt{E_\mathrm{V} - E} \qquad \left(E \leq E_\mathrm{V}\right) \tag{3.5}$$

with the energy range restrictions because $g_\mathrm{C}(E) = g_\mathrm{V}(E) = 0$ for $E_\mathrm{V} \leq E \leq E_\mathrm{C}$, and the Fermi-Dirac function

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_\mathrm{F}}{k_\mathrm{B}T}\right)} \tag{3.6}$$

The free electron and hole concentrations are then

$$n_0 = \int_{E_\mathrm{C}}^{\infty} g_\mathrm{C}(E) f(E)\,\mathrm{d}E \tag{3.7}$$

$$p_0 = \int_{-\infty}^{E_\mathrm{V}} g_\mathrm{V}(E)\left[1 - f(E)\right]\,\mathrm{d}E \tag{3.8}$$

The integrals in equations (3.7) and (3.8) do not have a general analytic solution. For the general case we can define the complete Fermi-Dirac integral function as [58, 59, 60]

$$\mathscr{F}_j(\eta_\mathrm{F}) \equiv \frac{1}{\Gamma(j+1)}\int_0^{\infty} \frac{\varepsilon^j}{1 + \exp\left(\varepsilon - \eta_\mathrm{F}\right)}\,\mathrm{d}\varepsilon \tag{3.9}$$

where $\eta_\mathrm{F} = \left(E_\mathrm{F} - E_\mathrm{C}\right)/k_\mathrm{B}T = \left(E_\mathrm{V} - E_\mathrm{F}\right)/k_\mathrm{B}T$ for electrons or holes, and $j = \mathrm{D}/2 - 1$ where D is the dimension of the semiconductor system. For 3D $j = 1/2$, and equation (3.9) is called the

"1/2 order Fermi-Dirac integral." With this definition the carrier concentration is found with

$$n = N_{\mathrm{DOS}} \mathscr{F}_{1/2}(\eta_{\mathrm{F}}) \tag{3.10}$$

where N_{DOS} is called the "effective density of states" and is

$$N_{\mathrm{C/V}} = 2 \left(\frac{m^*_{\mathrm{e/h}} k_{\mathrm{B}} T}{2\pi\hbar^2} \right)^{3/2} \tag{3.11}$$

for electrons or holes. For non-degenerate semiconductors (Fermi level is more than $\sim 3k_{\mathrm{B}}T$ inside the band gap away from the conduction or valence band edges) and using the equilibrium mass-action law of $n_{\mathrm{i}}^2 = n_0 p_0$ equation (3.10) can be approximated as

$$n_{\mathrm{i}} = \sqrt{N_{\mathrm{C}} N_{\mathrm{V}}}\, \mathrm{e}^{-E_{\mathrm{G}}/2k_{\mathrm{B}}T} \tag{3.12}$$

which is exponentially dependent on the band gap, In a heterojunction device n_{i} will have different values in different layers, and most importantly will affect the amount of injected minority carriers in each layer.

To achieve higher HBT gain, either the injected minority carrier concentration into the emitter p_{n0} must decrease or the minority carrier injection into the base n_{p0} must increase, as in equations (3.1) and (3.2). In homojunction devices the only option was to decrease the base doping relative to the

Table 3.1: Transport factors in HBT operation

Transport factor	**Symbol**	**Definition**	**Equation**
Emitter Efficiency	γ_{E}	$J_{\mathrm{n,E}}/J_{\mathrm{E}}$	$\left[1 + \frac{D_{\mathrm{E}} N_{\mathrm{B}} w'_{\mathrm{B}}}{D_{\mathrm{B}} N_{\mathrm{E}} w'_{\mathrm{E}}} \exp\left(\frac{-\Delta E_{\mathrm{G}}}{kT}\right)\right]^{-1}$
Base Transport	α_{T}	$J_{\mathrm{n,C}}/J_{\mathrm{n,E}}$	$\mathrm{sech}\left(\frac{w'_{\mathrm{B}}}{L_{\mathrm{B}}}\right) \approx 1 - \frac{1}{2}\left(\frac{w'_{\mathrm{B}}}{L_{\mathrm{B}}}\right)^2$
Recombination	δ_{R}	$\left(J_{\mathrm{E}} - J_{\mathrm{r,d}}\right)/J_{\mathrm{E}}$	—
Total Transport Factor	α	$J_{\mathrm{C}}/J_{\mathrm{E}}$	$\gamma_{\mathrm{E}}\alpha_{\mathrm{T}}\delta_{\mathrm{R}}$
Current Gain	β	$J_{\mathrm{C}}/J_{\mathrm{B}}$	$\alpha/(1-\alpha)$

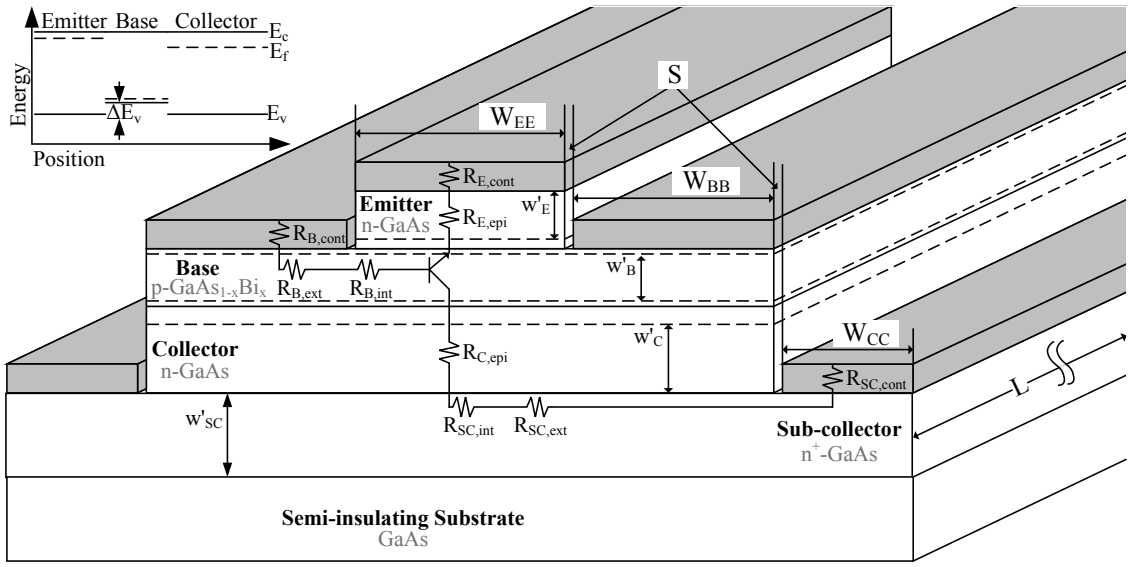


Figure 3.2: Schematic cross-section of double-mesa n-GaAs/p-GaAsBi/n-GaAs DHBT on a semi-insulating GaAs substrate with circuit diagram showing resistive current paths. Inset: Flat-band diagram of the DHBT.

emitter doping, since $n_{p0} = {}^{n_{i,B}^2}/_{N_B}$. But now including the result from equation (3.12), the minority carrier injection into the base is increased compared to the injection into the emitter. Now the current gain becomes

$$\beta = \frac{D_B}{D_E}\frac{w'_E}{w'_B}\frac{N_E}{N_B}\left(\frac{n_{i,B}^2}{n_{i,E}^2}\right) = \beta_0 \exp\left(\frac{E_G^{Emitter} - E_G^{Base}}{k_B T}\right) = \beta_0 \exp\left(\frac{\Delta E_G}{k_B T}\right) \tag{3.13}$$

which is the same as equation (1.1). Now there is much more freedom to dope the base higher and still maintain the gain, as well as reduce punch-through effects. Of course this exponential increase in gain cannot go on indefinitely. In a real transistor the total base current is made up of multiple currents, such as the base recombination current in figure 3.1. The gain derived in equation (3.13) is just one element of total HBT current called the emitter efficiency. Even if the hole current was reduced to zero, there would still be recombination as the electrons diffused across the base quasi-neutral region to the collector. The various transport factors in a HBT are shown in table 3.1.

Table 3.2: Material and device properties used in performance simulation.

Parameter	**Symbol**	**Value**	**Unit**
Bismuth concentration in the base	x	2.5	%
Emitter doping density	N_{E}	10^{18}	cm^{-3}
Base doping density	N_{B}	10^{19}	cm^{-3}
Base Gummel number	GN_{B}	1.5×10^{15}	cm^{-2}
Collector doping density	N_{C}	10^{16}	cm^{-3}
Sub-collector doping density	N_{SC}	10^{18}	cm^{-3}
Emitter thickness	w_{E}	300	nm
Base thickness	w_{B}	150	nm
Collector thickness	w_{C}	1000	nm
Sub-collector thickness	w_{SC}	500	nm
Device length	L	5	µm
Emitter stripe width	W_{EE}	0.5	µm
Base metal contact width	W_{BB}	0.5	µm
Collector metal contact width	W_{CC}	3	µm
Alignment width	S	0.1	µm
Contact resistivity [61]	$\rho_{\mathrm{Cont.}}$	10^{-6}	$\Omega \cdot \mathrm{cm}^2$
Base electron effective mass [11]	m_{e}^*	0.103	m_{e}
Base hole effective mass [11]	m_{h}^*	0.847	m_{e}

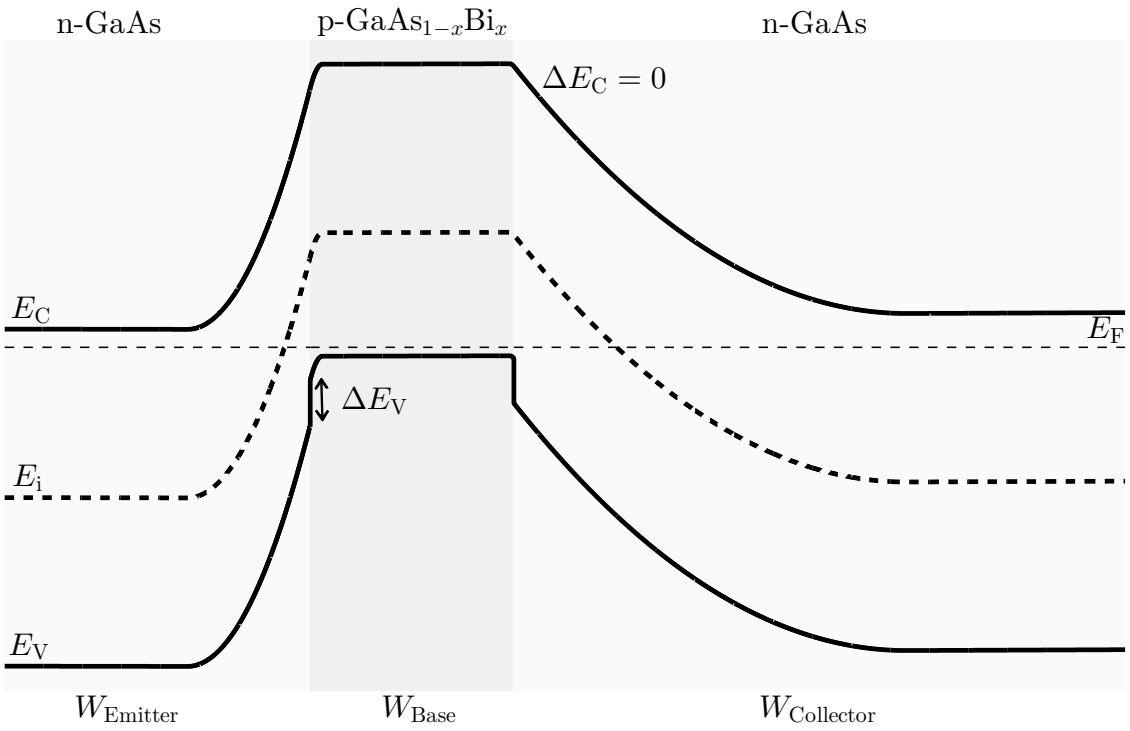
3.2 HBT Modeling

The device structure to model and simulate is a double-mesa vertical HBT with an n-type GaAs emitter, a p-type GaAsBi base, and an n-type GaAs collector and sub-collector grown on a semi-insulating GaAs substrate. A schematic of the device is shown in figure 3.2. The energy band diagrams of the model device under both zero and active bias are shown in figure 3.3a and figure 3.3b, respectively. Variable names and descriptions, as well as nominal values used in the simulations are listed in table 3.2. Material parameters for GaAsBi have been taken from the literature [8, 3, 53, 11]. When no material parameters were available typical values for GaAs were used. The device dimensions and doping levels were chosen to maximize performance while staying reasonable for fabrication. The resistivity of the layers were calculated from the doping, effective thickness, and mobility calculated in figure 2.3. In this model the alignment width S is set to 0.1 µm, with the assumption of a self-aligned HBT. Heating effects are not included in the model. Overall the parameter space of the HBT is explored to maximize the operating frequency and find good performance at reasonable device dimensions. This chapter will explore the details of the simulation model and its results. The specific implementation and architecture of the MATLAB code used to compute these results are provided in appendix B under an open-source license.

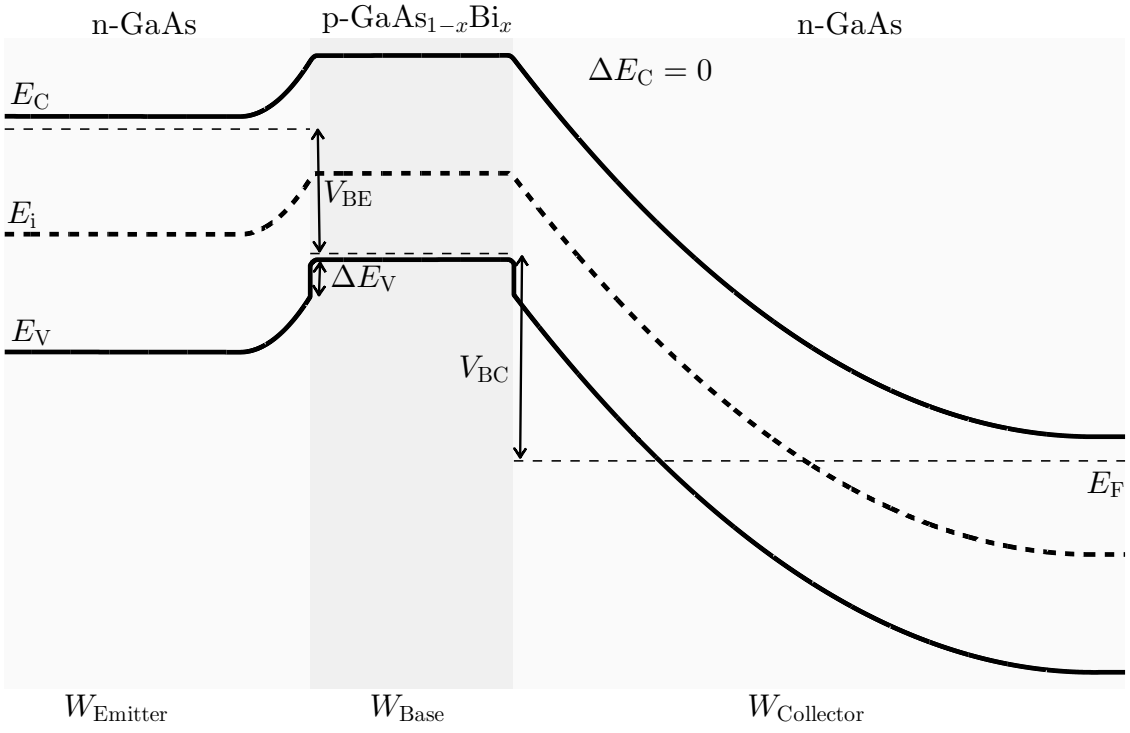
3.2.1 Drift–diffusion model

At each P-N junction in the HBT the equilibrium depletion region width is computed using the full-depletion approximation. That is, the carrier density is assumed to be zero inside the depletion region (fully depleted) and the charge is zero outside the depletion region (the quasi-neutral region). This approximation is actually very accurate compared to the full electrostatic solution to the Poisson equation because of the high doping concentrations used in this transistor model. This depletion analysis is used to find the effective quasi-neutral region width $w'_{\mathrm{p/n}} = w_{\mathrm{p/n}} - x_{\mathrm{p/n}}$, the junction capacitance, and the built-in voltage.

The electron and hole currents at each junction in figure 3.2 are computed from the general



(a) Zero bias. $V_{BE} = V_{BC} = 0$



(b) Active bias. $V_{BE} > 0$ and $V_{BC} < 0$

Figure 3.3: Energy band diagram of an n-p-n GaAs/GaAsBi HBT

solution of the analytic drift–diffusion differential equation for P-N junctions [62]. These are solved by considering the boundary conditions of the minority carrier concentrations at each side of the junction. No assumptions are made about the relative width of the junction versus the diffusion length (so called "short diode" and "long diode" approximation). Instead the generalized hyperbolic equations are used [56]. Starting with

$$J = J_{\mathrm{S}} \left(\mathrm{e}^{V_{\mathrm{A}}/V_{\mathrm{T}}} - 1 \right) \tag{3.14}$$

where the current density J is a function of the thermal voltage $V_{\mathrm{T}} = k_{\mathrm{B}}T/q$ and the applied voltage bias V_{A} between the p-type an n-type sides of the junction. J_{S} is the diode saturation current density and is

$$J_{\mathrm{S,p}} = q\frac{D_{\mathrm{n}}n_{\mathrm{p0}}}{L_{\mathrm{n}}}\coth\left(\frac{w'_{\mathrm{p}}}{L_{\mathrm{n}}}\right) \qquad \text{(Hole current)} \tag{3.15}$$

$$J_{\mathrm{S,n}} = q\frac{D_{\mathrm{p}}p_{\mathrm{n0}}}{L_{\mathrm{p}}}\coth\left(\frac{w'_{\mathrm{n}}}{L_{\mathrm{p}}}\right) \qquad \text{(Electron current)} \tag{3.16}$$

for either holes from the p-type to n-type side or electrons from the n-type to p-type side. $L_{\mathrm{p/n}}$ is the minority carrier diffusion length $L_{\mathrm{p/n}} = \sqrt{D_{\mathrm{p/n}}\tau_{\mathrm{p/n}}}$. In the limit $L_{\mathrm{p/n}} \gg w'_{\mathrm{p/n}}$ then equations (3.15) and (3.16) becomes the same as equations (3.1) and (3.2). The equilibrium minority carrier concentrations p_{n0}, n_{p0} are found through the mass-action law from equation (1.2). The n-type collector current is a special case. Here the boundary conditions have to include the depletion region interface on both the emitter and collector side of the base. From figure 3.4 the electron collector current is $J_{\mathrm{n,C}} = J_{\mathrm{n,E}} - J_{\mathrm{r,B}}$. The recombination current in the base $J_{\mathrm{r,B}}$ can be computed from the recombination of the integrated injected minority carrier density throughout the base divided by the carrier lifetime. Another way is to solve the drift–diffusion equation directly for these boundary conditions, which (as in chapter 5 of [56]) at the base-collector interface is

$$J_{\mathrm{n,C}} = q\frac{D_{\mathrm{n,B}}n_{\mathrm{i}}^2}{L_{\mathrm{n,B}}N_{\mathrm{B}}}\operatorname{csch}\left(\frac{w'_{\mathrm{B}}}{L_{\mathrm{n,B}}}\right) \tag{3.17}$$

The carriers that are injected into the base will diffuse and recombine until they reach the base-collector junction. At that point the reverse biased junction will sweep the carriers into the collector,

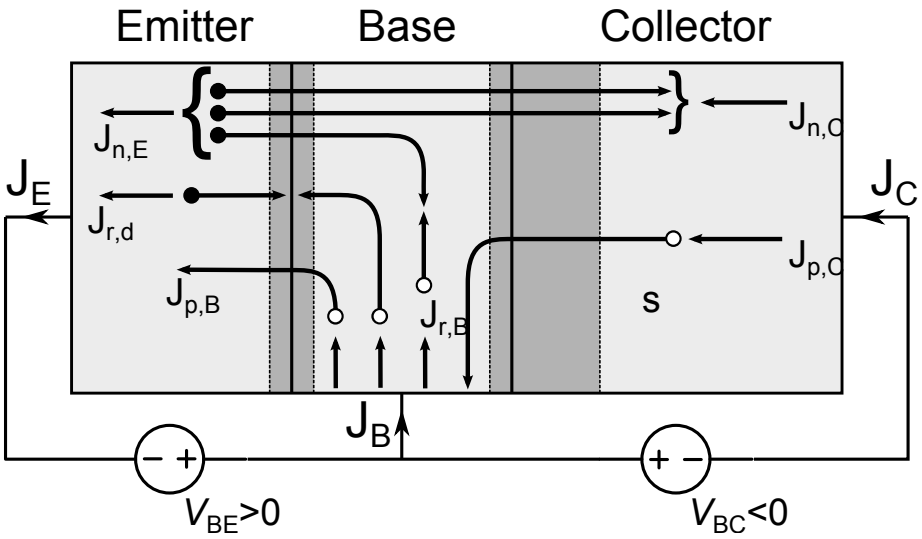


Figure 3.4: Full diagram of hole, electron, and recombination currents in a BJT or HBT.

contributing to the total collector current. The current components in figure 3.4 are summarized as:

$$J_{\mathrm{E}} = J_{\mathrm{n,E}} + J_{\mathrm{p,B}} + J_{\mathrm{r,d}} \tag{3.18}$$

$$J_{\mathrm{B}} = J_{\mathrm{p,B}} + J_{\mathrm{r,B}} + J_{\mathrm{r,d}} - J_{\mathrm{p,C}} \tag{3.19}$$

$$J_{\mathrm{C}} = \left(J_{\mathrm{n,E}} - J_{\mathrm{r,B}}\right) + J_{\mathrm{p,C}} \tag{3.20}$$

where $\left(J_{\mathrm{n,E}} - J_{\mathrm{r,B}}\right)$ is computed from equation (3.17). $J_{\mathrm{r,d}}$ is found from Shockley Reed Hall (SRH) recombination in the depletion region, approximated using an effective width factor of 0.1 (chapter 5 of [63]).

In a real P-N junction's current–voltage (I-V) curve the forward bias current doesn't increase indefinitely. There will be a roll-off from exponential increase to linear increase for any circuit with a non-zero resistance. Once the diode current is high enough, the limiting factor becomes the linear resistor. This effect can be included by defining $V_{\mathrm{A}} \rightarrow V_{\mathrm{A}}^{\mathrm{int}}$ as the internal voltage across the diode and the total voltage drop of both the diode and resistor as:

$$V_{\mathrm{A}}^{\mathrm{ext}} = V_{\mathrm{A}}^{\mathrm{int}} + IR \tag{3.21}$$

where R is the total resistance of the circuit and $I = AJ$ is the total current, calculated from the current density in equation (3.14) and the junction area A. This turns equation (3.14) into a

transcendental equation that cannot be solved analytically. In the simulation model the unnecessary transcendental calculation is avoided by computing the currents using the internal voltage $V_{\mathrm{A}}^{\mathrm{int}}$ up to a fraction of the built-in voltage and then plotting the results against the external voltage in equation (3.21). Unless otherwise stated all voltages in later figures will display the external voltages V_{BE} and V_{BC} from figure 3.4. This method improves computation time greatly but results in non-equally spaced data-points at higher current levels.

3.2.2 Current crowding

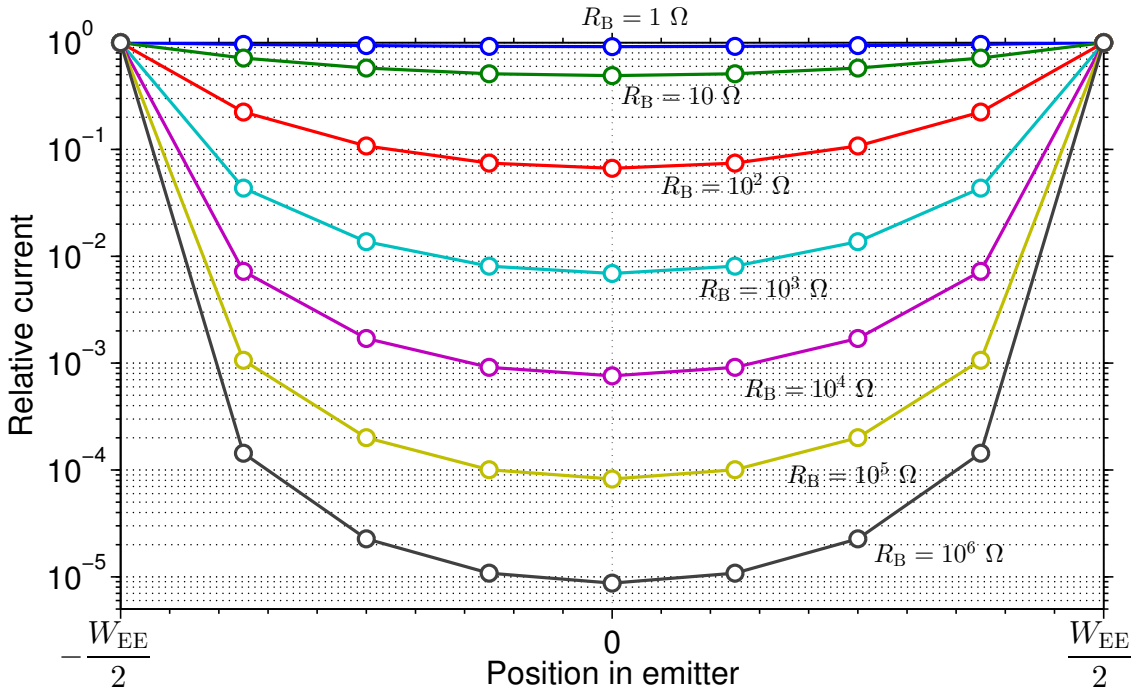


Figure 3.5: Current crowding effect in the base-emitter junction of a HBT as a function of base resistance. Generated from a distributed transistor SPICE simulation. Currents are normalized relative to the incoming current at the emitter mesa edge.

The base resistance in all simulated and fabricated GaAs/GaAsBi HBTs is large due to the low p-type carrier mobility in GaAsBi. An effect that can occur in HBTs with large access resistance is called **current crowding**. Because of the high resistance there is a voltage drop over the P-N junction between the edge and the middle of the emitter mesa. Because the base-emitter P-N junction

current is exponentially dependent on the junction voltage there can be a significant difference in the current densities between the mesa edge and middle. This violates the assumption used in most device simulations that the junction current density is constant across the emitter stripe width. The current density at the edge becomes much larger than in the middle—making the problem of surface recombination worse since more carriers are closer to surface recombination centers.

Figure 3.5 is a SPICE simulation of the current crowding effect using a distributed transistor model with nine sub-transistors. The currents are normalized to the current density at the mesa edge and simulated for various base resistances. Even for a moderate base resistance of 1 kΩ the current density at the center of the junction width is $< 1\,\%$ of the density at the edge. This means that the majority of the base-emitter current is likely concentrated at the emitter mesa edge. Since the surface recombination rate is similar to SRH recombination in that it is a function of the minority carrier concentration, the surface recombination current might be a significant effect in a fabricated device.

In the simulation model the current crowding effect is included by calculating an effective area for the base-emitter junction $A'_{\mathrm{BE}} = \left(\frac{\sin Z \cos Z}{Z}\right) A_{\mathrm{BE}}$. The factor Z is found from the transcendental equation [56, 64]

$$Z \tan(Z) = \frac{J_{\mathrm{B}} W_{\mathrm{BB}}^2 \rho_{\mathrm{B}}}{8 w'_{\mathrm{B}} V_{\mathrm{T}}} \tag{3.22}$$

where the numerical solution is fitted to a polynomial spline to speed up computation time during the simulation.

3.2.3 Time-constant RF model

The circuit model used to compute the parasitic resistances and capacitances are shown in figure 3.2. The junction capacitance of the base-emitter and base-collector junctions are $C_{\mathrm{j,BE}}$ and $C_{\mathrm{j,BC}}$, respectively. These capacitances are calculated from the depletion region width at each junction as

$$C_{\mathrm{j}} = \frac{\epsilon_{\mathrm{s}} \epsilon_0}{x_{\mathrm{d}(V_{\mathrm{A}})}} \tag{3.23}$$

where $x_{d(V_A)}$ is the depletion region width from the full depletion approximation as a function of the junction voltage bias V_A. The resistances are:

$$\begin{aligned} R_E &= R_{E,Epi} + R_{E,Contact} \\ &= \frac{\rho_E w'_E}{A_{BE}} + \frac{\rho_{Cont.}}{W_{EE} L} \end{aligned} \tag{3.24}$$

$$\begin{aligned} R_B &= R_{B,int} + \frac{1}{2} R_{B,ext} + \frac{1}{2} R_{B,Contact} \\ &= \frac{\rho_B A_{BE}}{12 L^2 w'_B} + \frac{1}{2} \frac{\rho_B S}{L w'_B} + \frac{1}{2} \sqrt{\frac{\rho_B \rho_{Cont.}}{w'_B L^2}} \coth\left(\sqrt{\frac{W_{BB}^2 \rho_B}{w'_B \rho_{Cont.}}} \right) \end{aligned} \tag{3.25}$$

$$\begin{aligned} R_C &= R_{C,Epi} + R_{SC,int} + \frac{1}{2} R_{SC,ext} + \frac{1}{2} R_{SC,contact} \\ &= \frac{\rho_C w'_C}{A_{BC}} + \frac{\rho_{SC} A_{BC}}{12 L^2 w'_{SC}} + \frac{1}{2} \frac{\rho_{SC} S}{L w'_{SC}} + \frac{1}{2} \sqrt{\frac{\rho_{SC} \rho_{Cont.}}{w'_{SC} L^2}} \coth\left(\sqrt{\frac{W_{CC}^2 \rho_{SC}}{w'_{SC} \rho_{Cont.}}} \right) \end{aligned} \tag{3.26}$$

and the junction areas are:

$$A_{BE} = L \cdot W_{EE} \tag{3.27}$$

$$A_{BC} = L \cdot \left(W_{EE} + 2 W_{BB} + 4 S \right) \tag{3.28}$$

with all variables defined in table 3.2 and figure 3.2. R_{int} is the intrinsic spreading resistance of the current going from vertical to lateral transport, R_{ext} is the extrinsic resistance across the alignment gap S, and $R_{Contact}$ is the contact resistance between the metal contact and the semiconductor [56]. The factor of $1/2$ is due to the dual-sided contacts.

RF performance of the HBT is calculated using a transit-time and RC time-constant model. The total HBT time-constant is the sum of the individual transit and RC times in the system. This total time-constant τ (and maximum transit frequency f_T) is [56]:

$$\frac{1}{2\pi f_T} = \tau = \tau_{E,RC} + \tau_{B,Transit} + \tau_{C,Transit} + \tau_{C,RC} \tag{3.29}$$

where the individual elements are:

$$\tau_{\mathrm{E,RC}} = \frac{V_{\mathrm{T}}}{J_{\mathrm{E}}}\left(C_{\mathrm{j,BE}} + C_{\mathrm{j,BC}}\right) \tag{3.30}$$

$$\tau_{\mathrm{B,Transit}} = \frac{w'^{2}_{\mathrm{B}}}{2D_{\mathrm{n,B}}} \tag{3.31}$$

$$\tau_{\mathrm{C,Transit}} = \frac{x_{\mathrm{d,BC}}}{2v_{\mathrm{Sat.}}} \tag{3.32}$$

$$\tau_{\mathrm{C,RC}} = \left(R_{\mathrm{E}} + R_{\mathrm{C}}\right) C_{\mathrm{j,BC}} A_{\mathrm{BC}} \tag{3.33}$$

where $\tau_{\mathrm{E,RC}}$ and $\tau_{\mathrm{C,RC}}$ are the RC charging time constants of the emitter and collector junctions. $\tau_{\mathrm{B,Transit}}$ is the carrier transit-time for diffusion across the quasi-neutral base region and $\tau_{\mathrm{C,Transit}}$ is the transit-time of carrier drift across the base-collector depletion region. $v_{\mathrm{Sat.}}$ is the high-field saturation velocity in GaAs, with a value of 7×10^{6} cm/s. The base-collector junction is assumed to be under high reverse bias, so any carriers that make it to the junction are swept out at the saturation velocity. The maximum oscillation frequency (where power gain is equal to one) is the geometric mean of the transit-frequency, f_{T}, and the base-collector RC charging time:

$$f_{\mathrm{Max}} = \sqrt{\frac{f_{\mathrm{T}}}{8\pi R_{\mathrm{B}} C_{\mathrm{j,BC}}}} \tag{3.34}$$

3.2.4 Multi-level partial dopant ionization model

An important effect observed in GaAsBi with implications for hole transport, and therefore device design, is the presence of energy states within the band gap near the VBM. A continuum of energy states from the VBM up to about 90 meV are suspected to be from bismuth induced spatially localized states near the valence band, likely from neutrally charged clusters of Bi atoms. The density of these states in one paper was found to be $\sim 10^{17}\,\mathrm{cm}^{-3}$, or about $\sim 2 \times 10^{-4}$ of the incorporated bismuth atoms [51]. The rest of the bismuth states are de-localized and form the modified valence band of GaAsBi. Deep-level states greater than 200 meV above the VBM are also found in MBE grown epi-layers at densities of $\sim 10^{15}\,\mathrm{cm}^{-3}$ [52]. These deep level states are at energy levels similar to LT-GaAs, so they are suspected to be from arsenic and bismuth incorporation into gallium

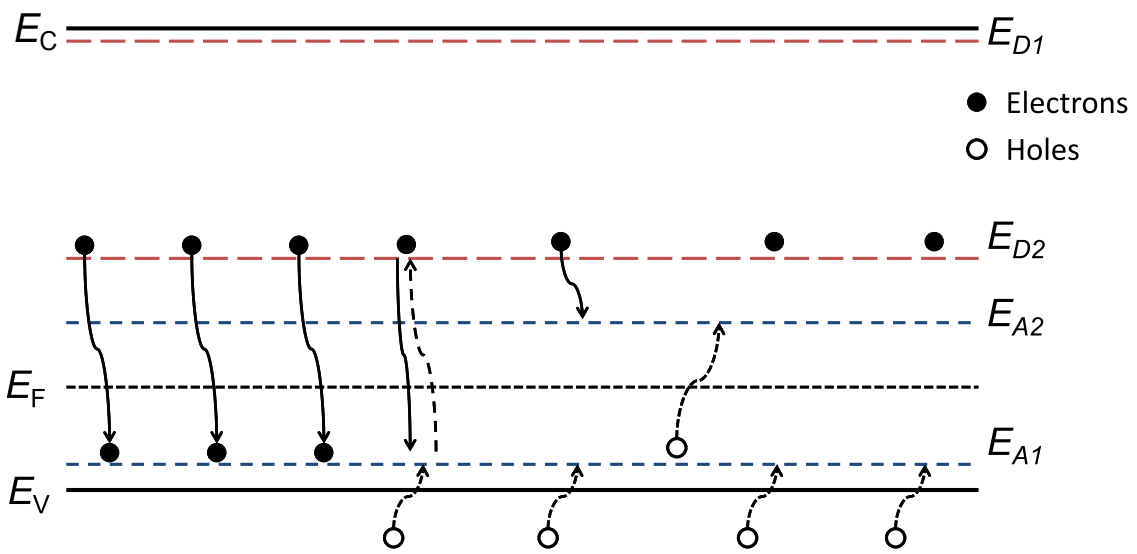
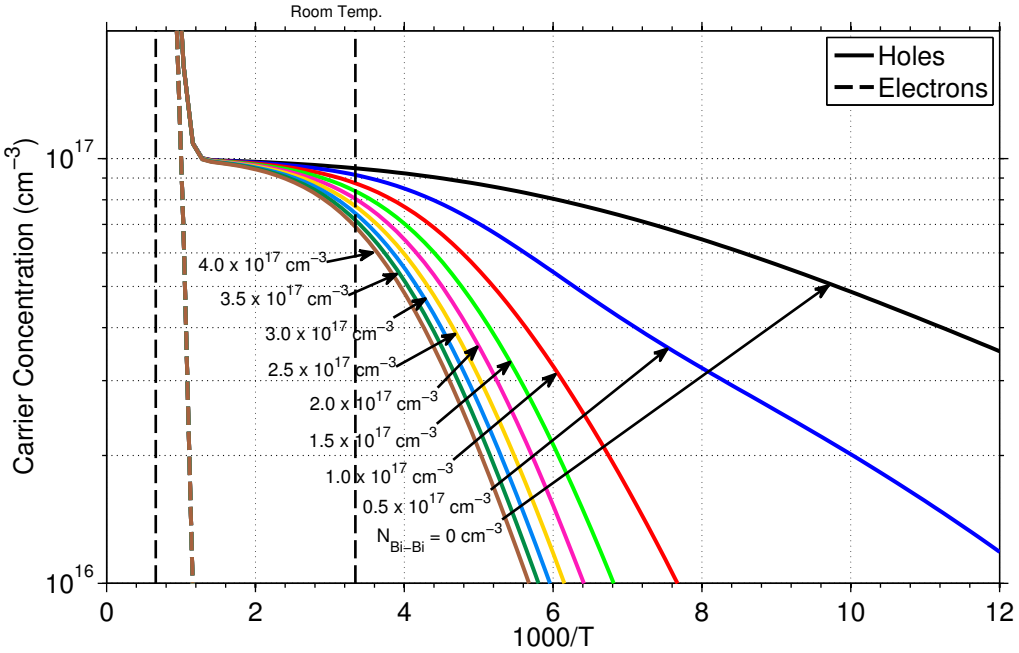


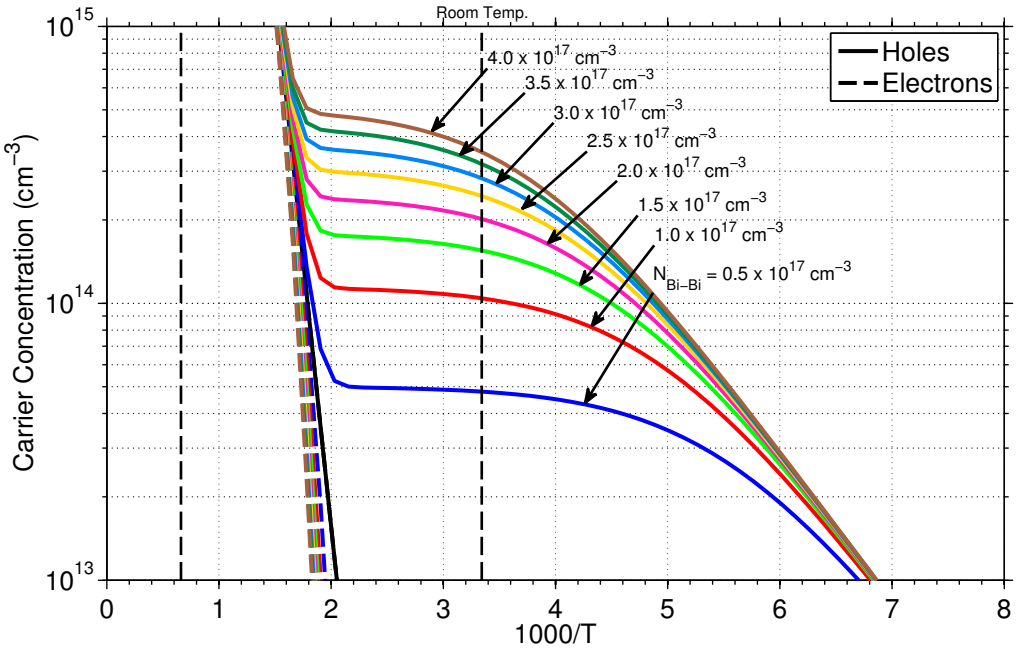
Figure 3.6: Band diagram of the multi-energy defect state model in p-type GaAsBi. Includes both shallow and deep donors (E_{D1}, E_{D2}) and acceptors (E_{A1}, E_{A2}).

antisites (As_{Ga} and Bi_{Ga}) from the low temperature growth. GaAsBi has been observed to be p-type doped even in nominally undoped samples, with stronger p-type doping observed at higher bismuth incorporation ($\sim 10^{14}$ cm^{-3} at 2 %$_{Bi}$ to $\sim 10^{17}$ cm^{-3} at 10 %$_{Bi}$) [65]. It has also been observed that bismuth incorporation can compensate p-type dopants in extrinsic GaAsBi and reduce the free hole concentration [15], again attributed to the Bi-Bi clusters. These seemingly contradictorily effects can be understood by considering the localized Bi clusters to occupy energy states around 90 meV above the VBM. The Bi-clusters act mostly as donor-like states with concentrations of around $\sim 10^{17}$ cm^{-3} for bismuth incorporation at 2.5 %$_{Bi}$, which compensate the shallow acceptors in GaAsBi. In low doped GaAsBi the Bi-clusters are too deep (nearly the entire band gap) to ionize electrons to the conduction band. Instead a small fraction of these impurities act as acceptors and cause a small p-type doping, which at higher bismuth concentrations can become significant in intrinsic samples.

It is important to have a model of these multi-energy state defects when designing and testing GaAsBi devices since the free carrier concentration will be affected differently than for GaAs. Later in chapter 5 the model will be compared to Arrhenius plots of the conductivity versus temperature measurements of actual samples. The model uses charge neutrality to find the Fermi level (and



(a) Free carrier concentration versus inverse temperature (Arrhenius) plot using multi-level ionization model with compensation. $N_{\text{Bi-Bi}}$ is amount of Bi-clusters in GaAsBi corresponding to the range of $\sim 0.5\,\%_{\text{Bi}}$ to $5\,\%_{\text{Bi}}$. Shallow doping set at $N_{\text{A1}} = 10^{17}\,\text{cm}^{-3}$.



(b) Arrhenius plot using multi-level ionization model showing p-type doping. Nominally undoped ($N_{\text{A1}} \sim 0$) with $N_{\text{A2}} = 10^{-3} \times N_{\text{Bi-Bi}}$.

Figure 3.7: Arrhenius plots using multi-level ionization model

therefore the free carrier concentration via equation (3.10)) in an equilibrium semiconductor in the presence of multiple impurity levels. The Fermi level is a unique single-value that describes the state of the system. Starting with charge neutrality:

$$\rho = q\left(p_0 - n_0 + N_{\mathrm{D}}^{+} - N_{\mathrm{A}}^{-}\right) = 0 \tag{3.35}$$

where the ionized impurities are each split up into shallow and deep levels. Expanded using the Fermi-Dirac distribution from equation (3.6) and the integral carrier distributing from equation (3.10) this becomes

$$\begin{aligned} N_{\mathrm{V}}\mathscr{F}_{\mathrm{p}}(T, E_{\mathrm{F}}) &+ \frac{N_{\mathrm{D1}}}{1 + g_{\mathrm{n}} \exp\left(\frac{E_{\mathrm{F}} - E_{\mathrm{D1}}}{k_{\mathrm{B}}T}\right)} + \frac{N_{\mathrm{D2}}}{1 + g_{\mathrm{n}} \exp\left(\frac{E_{\mathrm{F}} - E_{\mathrm{D2}}}{k_{\mathrm{B}}T}\right)} \\ = N_{\mathrm{C}}\mathscr{F}_{\mathrm{n}}(T, E_{\mathrm{F}}) &+ \frac{N_{\mathrm{A1}}}{1 + g_{\mathrm{p}} \exp\left(\frac{E_{\mathrm{A1}} - E_{\mathrm{F}}}{k_{\mathrm{B}}T}\right)} + \frac{N_{\mathrm{A2}}}{1 + g_{\mathrm{p}} \exp\left(\frac{E_{\mathrm{A2}} - E_{\mathrm{F}}}{k_{\mathrm{B}}T}\right)} \end{aligned} \tag{3.36}$$

where $g_{\mathrm{n}} = 2$ and $g_{\mathrm{p}} = 4$ are the degeneracy factors for donors and acceptors. (E_{D1}, N_{D1}) and (E_{D2}, N_{D2}) are the energy levels and impurity concentration of the shallow and deep donors while (E_{A1}, N_{A1}) and (E_{A2}, N_{A2}) are for the shallow and deep acceptors, respectively. The free variable in equation (3.36) is E_{F}, and there exists a single unique solution to the transcendental equation that can be found numerically. From the Fermi level the free carrier concentration can be calculated from equation (3.10).

In the GaAsBi system (E_{D1}, N_{D1}) are zero for p-type material. (E_{D2}, N_{D2}), and (E_{A2}, N_{A2}) are the deep level donors and acceptors for the Bi-cluster induced states. (E_{A1}, N_{A1}) are from the shallow acceptor dopants, which is carbon in this system at $E_{\mathrm{A1}} = 26\,\mathrm{meV}$ [56]. Examples of this model are shown in figures 3.7a and 3.7b. Here the deep energy levels are set to $E_{\mathrm{D2}} = 90\,\mathrm{meV}$ and $E_{\mathrm{A2}} = 27\,\mathrm{meV}$ relative to the valence band edge. The Bi-induced cluster concentration is swept with $N_{\mathrm{D2}} = 1 \times N_{\mathrm{Bi\text{-}Bi}}$ and $N_{\mathrm{A2}} = 10^{-3} \times N_{\mathrm{Bi\text{-}Bi}}$. This accounts for the qualitative behavior of GaAsBi seen in the literature. At low doping levels bismuth alloying will induce p-type carriers [65], while at high doping the bismuth will cause compensation [51]. The model shown in figure 3.7 shows this same qualitative behavior.

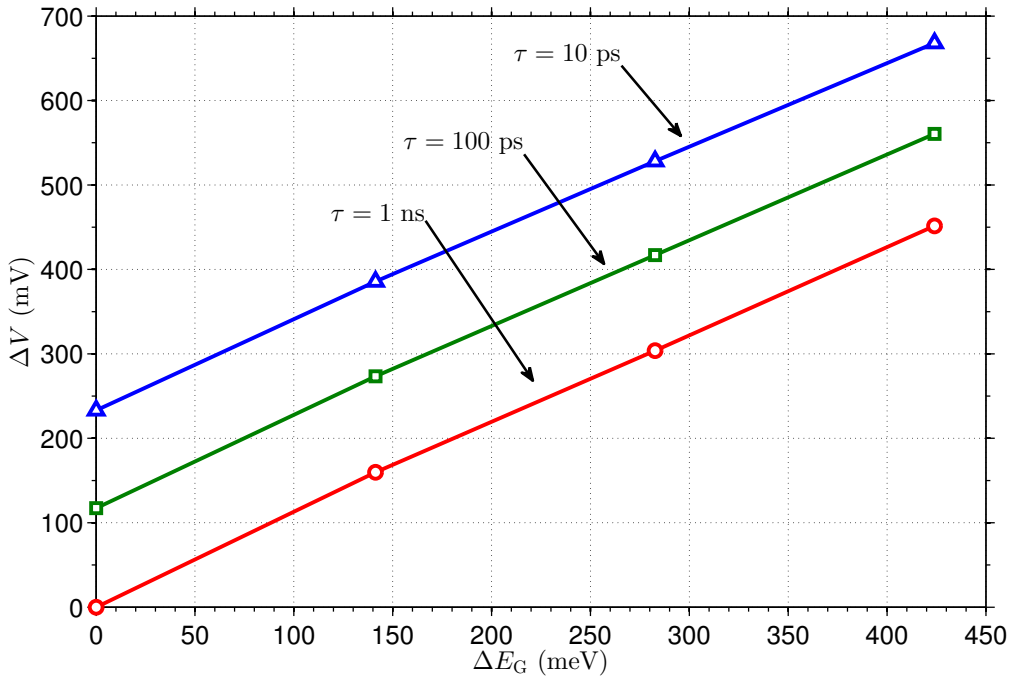


Figure 3.8: Simulated turn-on voltage shift as a function of band gap shift for different life-times. Structure is a GaAsBi/GaAs P-N diode. The lifetime in the n-GaAs layer is 1 ns. Both lower band gaps and reduced life-times in the p-GaAsBi layer reduces the the diode turn-of voltage.

3.3 GaAs/GaAsBi HBT Simulation Results

3.3.1 DC simulation results

In order to both verify the analytic model, and to assist in analyzing fabricated devices it is helpful to determine to what extent different material parameters affect basic devices, such as P-N junctions. Because of the reduction in band gap from the bismuth alloying, the turn-on voltage (voltage required to achieve some current level in a P-N junction) should shift to lower voltages at reduced band gaps. Using the numerical simulation program `PC1D` [66] this voltage shift is found as both a function of band gap and of life-time, as shown in figure 3.8. This result can be used with measured devices to confirm both the presence of a band gap reduction, and as an approximation of the life-time difference between homojunction and heterojunction diodes.

Nominal parameters used for the device simulation are found in table 3.2. The first device simulation is the Gummel plot of $J_C(V_{BE})$ and $J_B(V_{BE})$, as well as the current gain β, shown in

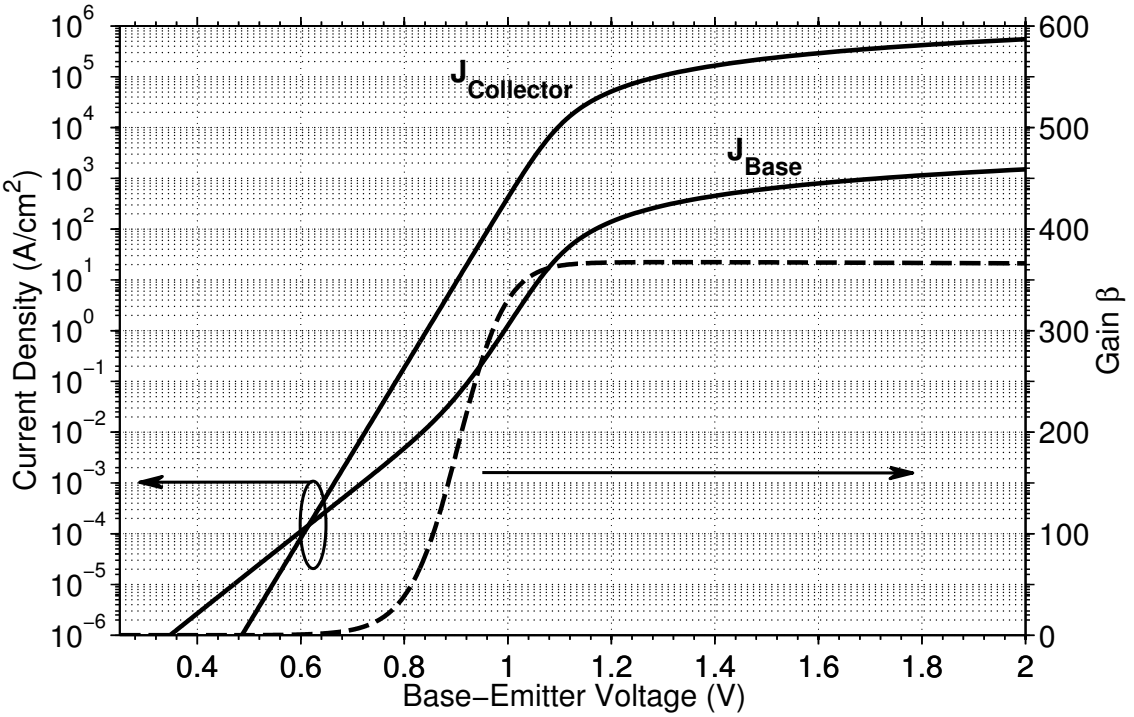


Figure 3.9: Simulated Gummel plot of base and emitter current density and gain β of GaAs/GaAsBi HBT device with $x = 2.5\,\%_{\mathrm{Bi}}$ at maximum reverse bias before breakdown of $V_{\mathrm{BC}} = -22\,\mathrm{V}$.

figure 3.9. The maximum gain β versus the bismuth fraction x in the base is shown in figure 3.10 for various base thicknesses w_{B}. In figure 3.9 at low forward voltage bias the base current J_{B} has an ideality factor of $n \sim 2$ from SRH recombination in the depletion region being the dominant current mechanism. At larger forward bias both the collector J_{C} and base J_{B} currents have ideality factors of $n \sim 1$ as minority carrier injection and recombination in the base becomes dominant. At even larger bias the series resistance in the HBT becomes the dominant voltage drop, causing the current to roll off and the gain to saturate. This is from the additional voltage drop across the ohmic series resistance (equations (3.24) to (3.26)). The internal gain will still increase for higher voltages across the junction, but additional applied voltage at high current levels will cause more voltage drop across the resistor. High-injection effects (i.e. when the injected minority carrier concentration becomes comparable to the doping level) such as the diode current slope or transit time reduction are not needed in this model. The base and emitter layers are highly doped, so other effects such as the series resistance occurs well before the onset of high-injection at high bias in a HBT.

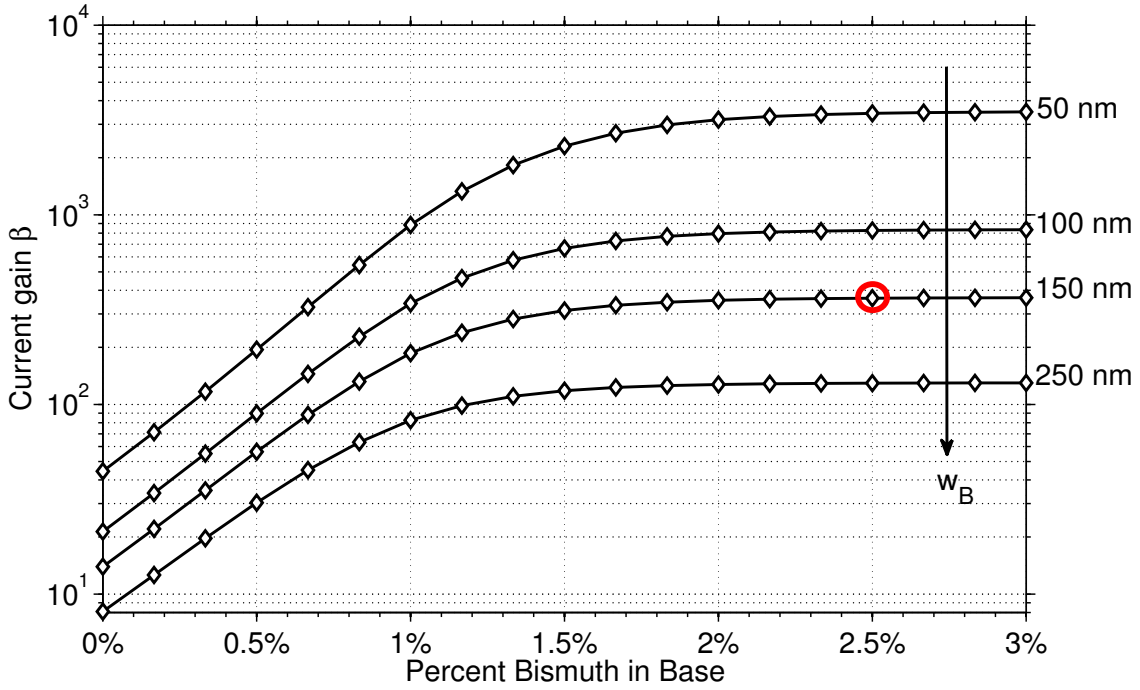


Figure 3.10: Simulated gain β versus bismuth concentration in the base, for $J_E > 10^5$ A/cm^2. The bismuth fraction and base layer thickness used for simulations in this chapter is circled.

Figure 3.10 shows the maximum of the current gain for $J_E > 10^5$ A/cm^2 as a function of Bi concentrations x in the base over various base layer thicknesses w_B. This emitter current density was chosen because it is the point where the gain saturates, but before significant heating effects might occur. The absolute gain of an actual HBT device will likely be lower than shown in figure 3.10 due to carrier recombination effects on the surface or larger than calculated base-emitter depletion region recombination from interface states or compensation during growth. However the general trends will remain the same. Because the GaAs/GaAsBi system does not have a band gap discontinuity in the conduction band, electrons injected into the base are unimpeded to travel across the base-collector junction. As the Bi concentration is increased, the band gap of the base layer will decrease by ~ 88 meV/Bi%, which will increase the emitter electron current relative to the base hole current and increase the gain. At low Bi concentration the HBT current gain is limited by the emitter efficiency. In this region the gain will improve exponentially with the reduction in base band gap ΔE_G. At higher Bi concentrations $J_{p,B}$ becomes negligible compared to the recombination currents $J_{r,d}$ and

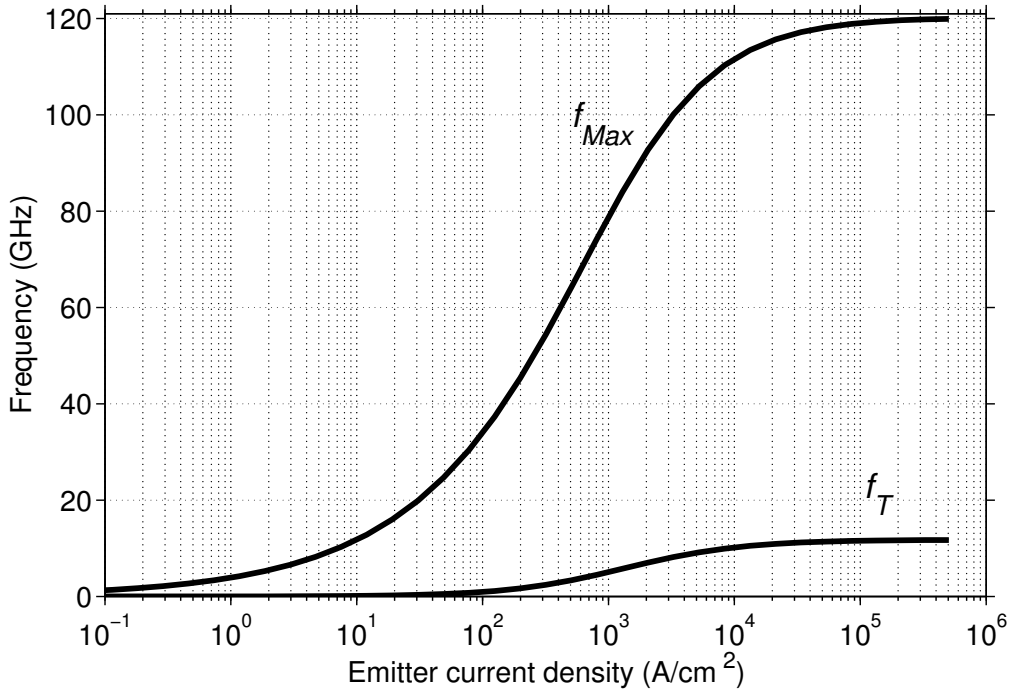


Figure 3.11: f_{T} and f_{Max} versusJ_{E} based on transit-time model. $V_{\mathrm{BC}} = -15\,\mathrm{V}$ with $w_{\mathrm{B}} = 150\,\mathrm{nm}$, $N_{\mathrm{B}} = 10^{19}\,\mathrm{cm}^{-3}$ and $N_{\mathrm{C}} = 10^{16}\,\mathrm{cm}^{-3}$

$J_{\mathrm{r,B}}$, so that the gain becomes limited by the transport-factor α_{T} as the emitter efficiency and gain saturate. The maximum gain begins to saturate for bismuth concentrations of $x > 2\,\%$ to $3\,\%$. In this range the gain of the device is recombination, rather than hole injection current dependent. This means that the gain will be less sensitive to variations in the band gap with temperature. Bismuth concentrations greater than $\sim 3\,\%$ are unnecessary in HBT devices since the mobility and base resistance will decrease with increased bismuth incorporation, as in figure 2.3, without improving the gain. As in table 3.2, all simulations will use this value of $x = 2.5\,\%$ ($\Delta E_{\mathrm{G}} \approx 230\,\mathrm{meV}$) for the bismuth concentration, unless explicitly stated.

3.3.2 RF simulation results

The maximum frequency of operation versus emitter current density for the parameters specified in table 3.2 is shown in figure 3.11. Here the base-collector reverse voltage bias is $V_{\mathrm{BC}} = -15\,\mathrm{V}$. The maximum frequency increases with the current density primarily because the RC charging

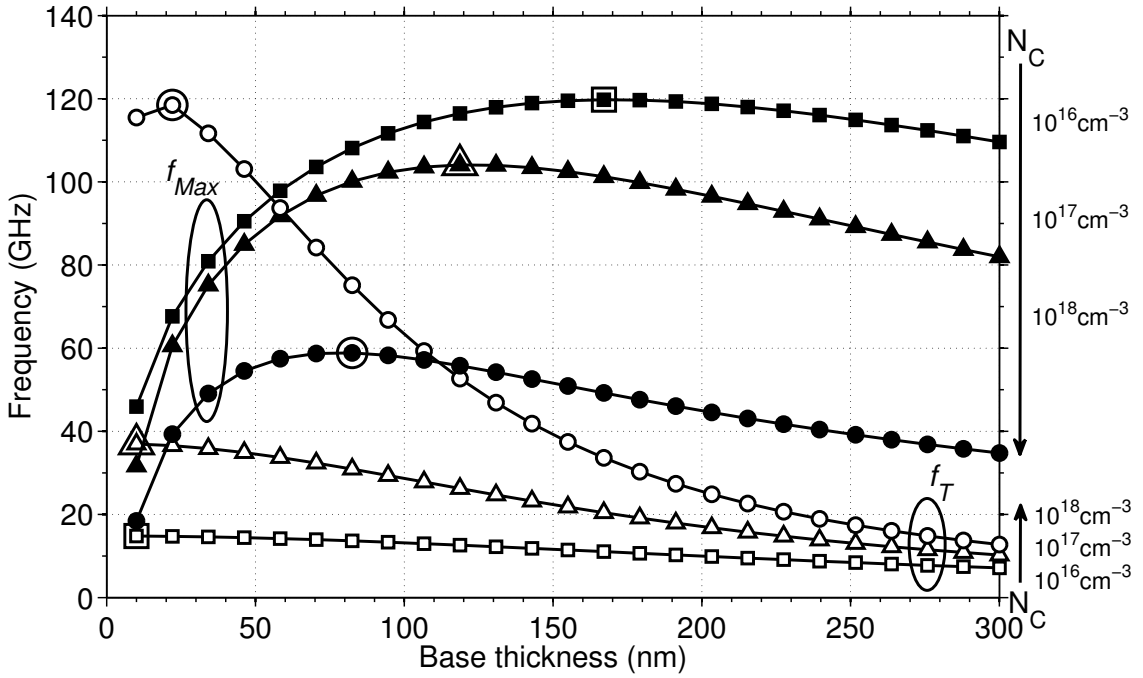


Figure 3.12: f_{T} and f_{Max} versus base thickness w_{B} at constant base doping, $N_{\mathrm{B}} = 10^{19}\,\mathrm{cm}^{-3}$, for collector doping of $N_{\mathrm{C}} = 10^{16}\,\mathrm{cm}^{-3}$, $10^{17}\,\mathrm{cm}^{-3}$, and $10^{18}\,\mathrm{cm}^{-3}$. Maximum f_{Max} of 120 GHz and f_{T} of 13 GHz for $w_{\mathrm{B}} = 150\,\mathrm{nm}$ and $N_{\mathrm{C}} = 10^{16}\,\mathrm{cm}^{-3}$ at maximum current density.

time-constant of the emitter-base junction decreases due to the decrease in differential resistance $V_{\mathrm{T}}/J_{\mathrm{E}}$, as in equation (3.30). f_{Max} is the geometric mean of the transit-frequency, f_{T}, and the base-collector RC charging time, as in equation (3.34). The high f_{Max} relative to f_{T} is because of the low base-collector capacitance $C_{\mathrm{j,BC}}$ despite the high base resistivity caused by the low hole mobility in GaAsBi. The main factor limiting f_{T} is the collector transit-time (equation (3.32)), which is a function of the base-collector depletion region thickness. This collector transit time can be reduced to increase f_{T} by increasing the collector doping N_{C} or decreasing the base-collector reverse voltage bias V_{BC}, both of which will reduce the base-collector depletion region thickness. However both of these parameters will increase the base-collector junction capacitance $C_{\mathrm{j,BC}}$, and decrease f_{Max}. There is a trade-off between f_{T} and f_{Max} that can be adjusted to optimize RF performance at a specific operating frequency. The design target for these simulations will be a HBT capable of X-band RF amplification (i.e. 6–12 dB of gain at 10 GHz), which would require a $f_{\mathrm{Max}} > 100\,\mathrm{GHz}$ and $f_{\mathrm{T}} > \sqrt{f_{\mathrm{Max}}} \approx 30\,\mathrm{GHz}$.

Figure 3.12 shows how the maximum frequency of the HBT is affected by the key variable of the base thickness w_B. The base thickness is varied while the doping remains constant at $N_B = 10^{19}\,\mathrm{cm}^{-3}$. f_T and f_{Max} are found for $J_E > 10^5\,\mathrm{A/cm}^2$ and $V_{BC} = -15\,\mathrm{V}$. f_T increases for smaller base thickness due to the decrease in base transit-time, as in equation (3.31). f_{Max} reaches a maximum of 120 GHz at $w_B \approx 150\,\mathrm{nm}$ and decreases again for thicker base layers due to the reduction in f_T. Increasing N_C reduces the base-collector depletion layer thickness. This increases f_T from the reduction in collector transit-time and decrease f_{Max} from the increased base-collector capacitance. For the target X-band operation at 10 GHz with a gain of ~ 10 a $f_{Max} \geq 100\,\mathrm{GHz}$ and $f_T \geq 30\,\mathrm{GHz}$ is needed. In figure 3.12 this occurs for $w_B = 100\,\mathrm{nm}$ and $N_C = 10^{17}\,\mathrm{cm}^{-3}$.

Figure 3.13 is the optimized performance of a GaAs/GaAsBi HBT, given the design constraints and simulation assumptions. f_T and f_{Max} at maximum current and V_{BC} are shown versus the emitter

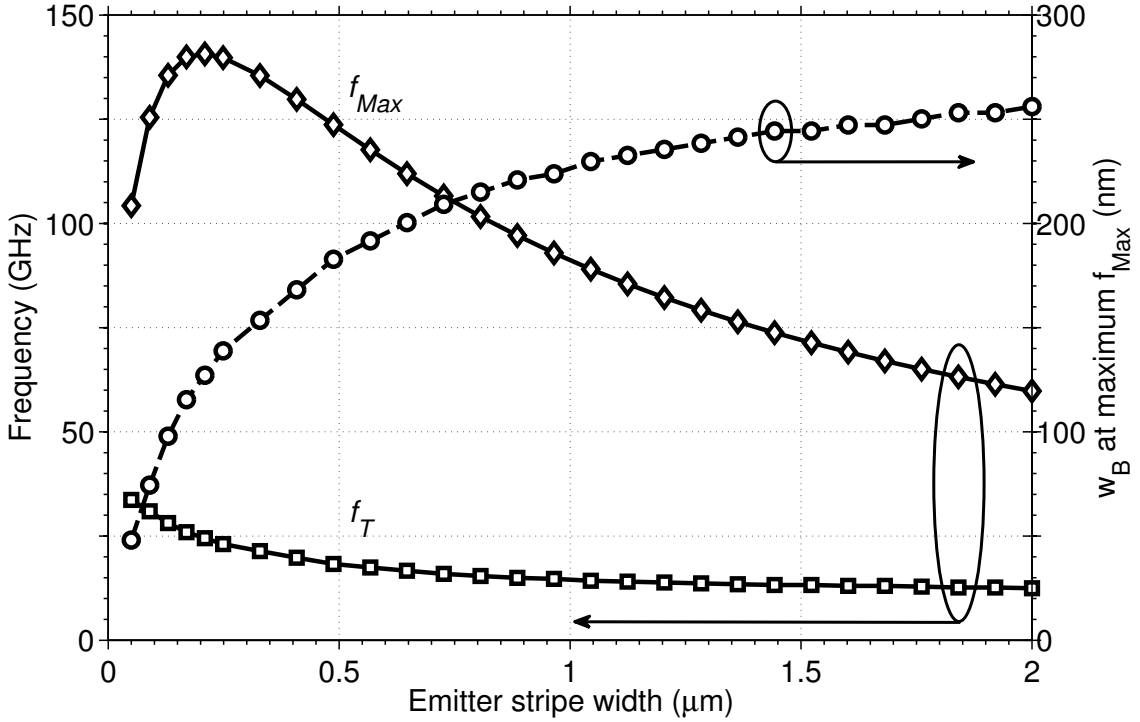


Figure 3.13: f_T and f_{Max} with $N_C = 10^{16}\,\mathrm{cm}^{-3}$ at maximum voltage and current density versus W_{EE} with base thickness w_B of maximum f_{Max} at each point. All other device parameters are scaled with W_{EE} for the same ratios as on table 3.2. As the emitter stripe width is varied the other device dimensions are varied at the ratios of $2w_E$, $8w_C$, $4w_{SC}$, W_{BB}, $6W_{CC}$, and $10L$.

stripe width W_{EE}. At each frequency the base thickness for maximum f_{Max} is calculated and shown in the left axis of figure 3.13. All other device parameters are scaled with W_{EE} for the same ratios as on table 3.2. As the emitter stripe width is varied the other device dimensions are varied at the ratios of $2w_E$, $8w_C$, $4w_{SC}$, W_{BB}, $6W_{CC}$, and $10L$. All of these results show that a GaAs/GaAsBi HBT is capable of good X-band RF performance at reasonable device dimensions and can be tuned to fit the needs of specific applications.

Chapter 4

Fabrication

4.1 Mask Design

There have been a total of five generations of HBT mask designs for this project. In this thesis the focus will be on three different mask-sets: **1)** The 2-step P-N junction mask-set for fabricating test structures on P-N junction samples, shown in figure A.4. **2)** The six-step HBT generation four mask-set, used to fabricate many of the devices in this thesis and shown in figure A.2. And **3)** the six-step HBT generation five mask-set, the latest mask-set shown in figure A.3. A process sequence of the full six-step HBT process is shown in figure A.1. All of these masks and process procedures are shown in appendix A. In order of processing, the steps consists of an emitter etch to reveal the base layer, a base etch to reveal the subcollector layer, an n-type emitter and collector ohmic contact metalization, a p-type base ohmic contact metalization, a blanket nitride encapsulation followed by a via etch through the SiN_x layer, and a wiring metalization to make the electrical test pads. A rapid thermal anneal (RTA) must be performed one or two times for the ohmic metalization. This can be done either before or after the nitride encapsulation, but generally before the wiring metalization. Each of these processes will be explored in more depth in this chapter. Specific process procedures are detailed in appendix A. Test structures common to all the mask-sets include P-N junctions for I-V

and capacitive–voltage (C-V) testing, transfer length measurement (TLM)[5] test structure to measure the sheet and contact resistance, various sized HBT structures, and alignment marks. Details of the TLM measurement technique can be found in [67, 68, 69].

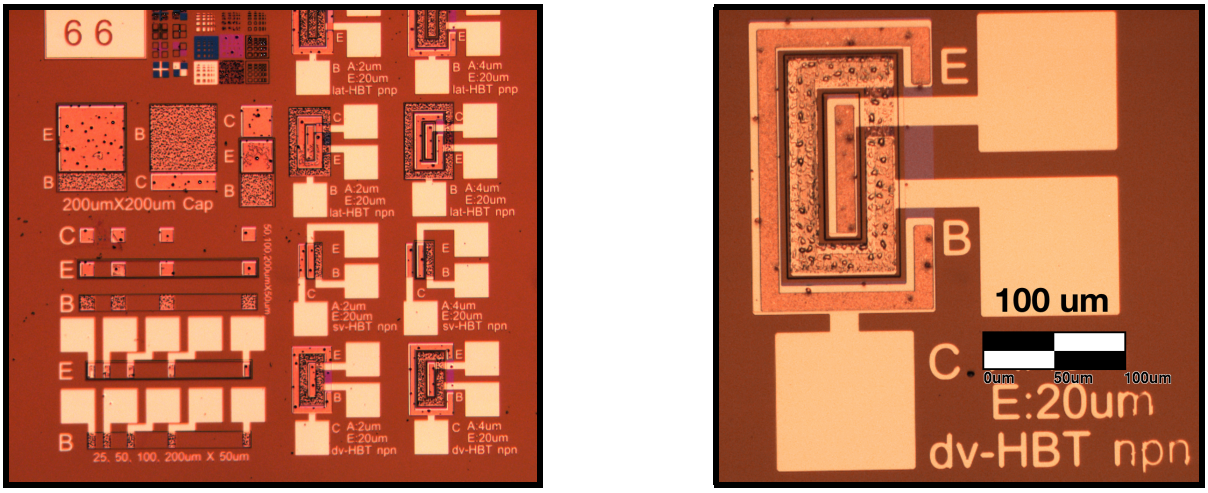
The P-N junction mask-set (shown in figure A.4) has two processing steps: one etch and one metalization. In addition there is a non-lithographic back contact metalization. This mask-set is designed for samples with a single P-N junction usually consisting of n-GaAs and p-GaAsBi epi-layers grown on a conducting n^+-GaAs substrate. The first process step etches down through the p-type layer into the n-type layer. The exact etch depth does not need to be well controlled since any etch greater than the p-type layer thickness is sufficient. Then the front is metalized with a p-type ohmic contact, and the back is metalized with a blanket n-type ohmic metalization. The sheet and contact resistance in the p-type layer is measured with a TLM structure. Various sizes of both square and circular mesa contacts are used to perform I-V and C-V testing on the P-N junction. There are lateral P-N-P HBT structures to help study minority transport in the material, but in a typical P-N junction sample the p-type layer is on top, making the "base" of this lateral HBT the n^+-GaAs substrate—which makes for a poor transistor and does not provide any information about the material properties of GaAsBi.

The generation four HBT requires six process steps: two etches (emitter and base mesa etch), two metalizations (n-type emitter and collector ohmic, and p-type base ohmic), a via etch through the SiN_x encapsulation layer and a Ti/Au wiring metalization for electrical probe contacts. The mask layout showing all structures is shown in figure A.2, and a micrograph and scanning electron microscope (SEM) image of a finished unit cell is shown in figure 4.1a and figure 4.2a. In this process the first etch is the most critical. The emitter layer must be etched away without removing too much of the base layer. Typically the etch needs to remove $\sim 300\,\text{nm}$ of the emitter layer and stop at the $\sim 150\,\text{nm}$ thick base. Details of this etch process are provided in section 4.2. The second

[5] There are a few different interpretations of what the acronym "TLM" means. In the context of semiconductor measurement they all mean the same thing, but some share their name with techniques in other fields. One is "transmission line method" or "transmission line measurement", a term borrowed from RF work. Another is "transfer length model" or "transfer length method" or even "transfer length measurement." In this thesis I will use "transfer length measurement" to denote the semiconductor measurement technique to extract sheet and contact resistance from a thin film.

etch through the base into the subcollector is less critical since the collector and subcollector layers are much thicker ($\geq$ 300 nm for each). The mask-set has a total of five TLM test structures: two isolated mesas for the n-type GaAs emitter layer and the p-type GaAsBi base layer with via contacts, two mesas for the emitter and base with pads directly to the ohmic metal, and a set of pads on the n-type GaAs subcollector. The subcollector TLM contacts are more for verifying the linearity of the contacts than for quantitative measurements since they are not isolated from the substrate. In addition the mask-set has two large area base-emitter and base-collector diodes for C-V testing and a large area N-P-N HBT structure for device testing before the last two process steps. The main structures are the eight HBTs on the right side of figure 4.1a. There are two sets of lithographic alignments, one with 2 μm alignment spacing and the other with 4 μm alignment spacing. In each of these sets there is a both a single and a double sided contacted N-P-N vertical, via-contacted HBT device. A micrograph of an individual double-sided HBT after all six mask steps is shown in figure 4.1b. A SEM image of a similar device after four mask steps is shown in figure 4.2b. There are also two lateral HBTs with an N-P-N version using the emitter-base-emitter layers and a P-N-P version using the base-collector-base layers. These are to test minority transport in different material layers. All the via-contacted HBT structures have a 20 μm $\times$ 100 μm emitter contact area.

The generation five mask-set (shown in figure A.3) is also a full six-step HBT process with two etches, two ohmic metalizations, a via etch and a wiring metalization. Like generation four this mask-set has two via-contacted TLM structures to the emitter and base layer and a base-emitter and base-collector C-V lateral diode. It also has a meandering line resistor to measure the wiring metal resistivity and a vertical, horizontal, and diagonal step continuity test structure. The actual transistor devices on this mask-set are a set of double-sided via-contacted HBTs in eight different sizes with different emitter strip widths, lengths, and alignments. The purpose of the different sized HBTs is to be able to extract the effect of surface recombination on the transistor gain. In a HBT there might be recombination at the semiconductor surface in addition to recombination in the bulk. The surface recombination is proportional to the emitter mesa perimeter, where all other currents are proportional to the emitter mesa area. By varying the dimensions the ratio of perimeter to area



(a) Unit cell

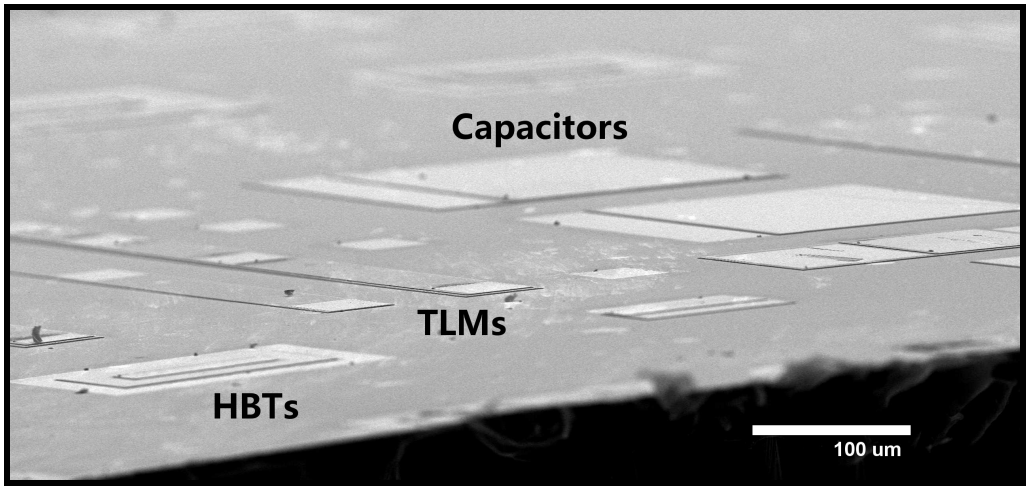
(b) N-P-N dv-HBT

Figure 4.1: Micrographs of finished six-step HBT mask (Generation 4)

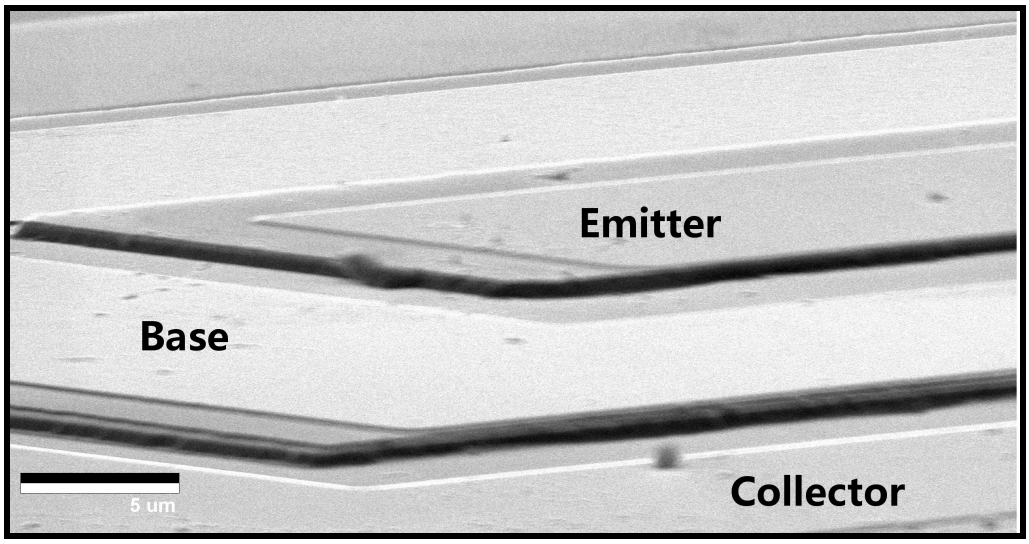
will change, and so will the proportion of the surface to bulk recombination current. Plotting the inverse gain $1/\beta$ versus the perimeter to area ratio, P/A, results in a straight line whose slope is proportional to the saturation current density of the surface recombination current. This additional information is helpful in understanding which non-ideal effect might be affecting the HBT device. Details of the measurement technique can be found in [70].

4.2 Mesa Etching

All the processed samples in this thesis require some sort of etch step. Since the material is MBE grown as planar stacks in order to make contact with the buried layers the material must be etched away, as in steps L1 and L2 in figure A.1. This forms mesas of material which are then metalized to form a working device. Both wet etching and dry etching processes were explored. Wet etching can have very different etch rates between different materials depending on their surface chemistry. This effect can be exploited to create an "etch stop" layer that will automatically stop a wet etch at a certain layer. However, if two materials have very similar etch rates, like GaAs and GaAsBi, then process control can be difficult since the exact etch rate depends on various factors



(a) Wide view of gen. 4 unit cell



(b) Double-side contacted HBT

Figure 4.2: SEM image of generation 4 HBT mask-set after four process steps

such as concentration, temperature, agitation, etc. Reactive ion etching (RIE) typically has much less selectivity between different materials, but much better process control. Both processes are established techniques in GaAs device fabrication, but there are some differences when etching GaAsBi. A typical III-V semiconductor RIE process uses a chlorine based gas, such as Cl_2 or BCl_3 to attack the surface of the semiconductor, chemically react with the elements of the compound, and carry away the reaction products [71]. The reaction products of the chemical interaction must

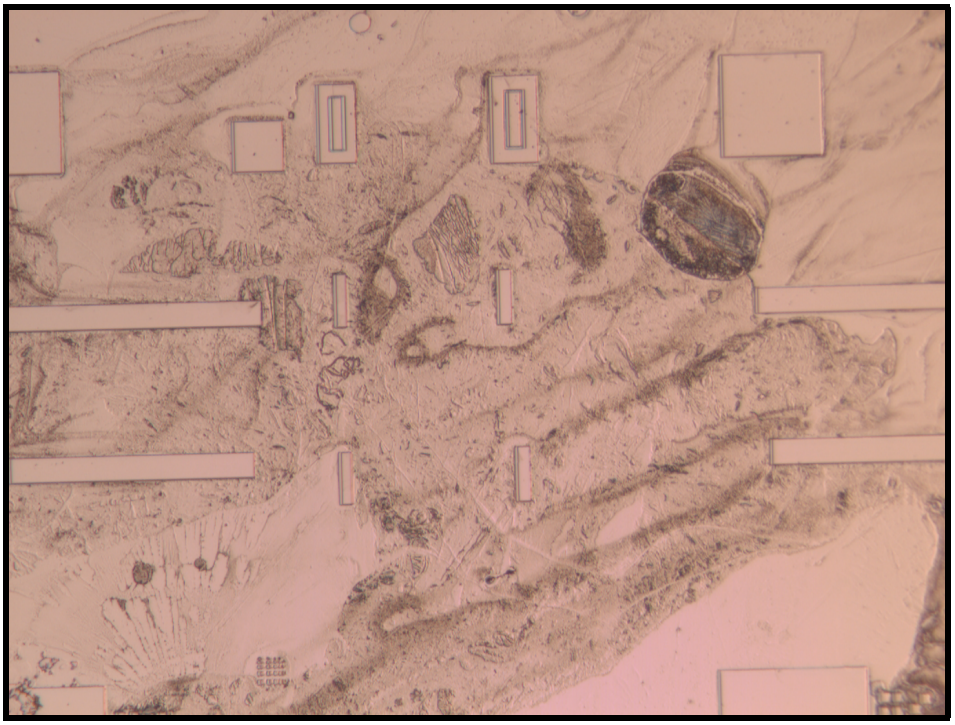
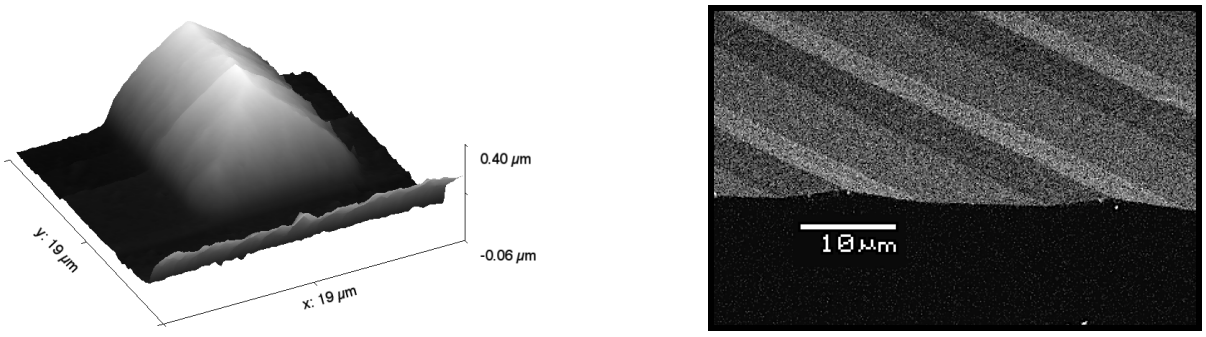


Figure 4.3: Micrograph of failed GaAsBi RIE using a BCl_3 etch gas

be gaseous or they will remain on the surface and slow down or even stop further etching. While this process works well for GaAs, the etching of GaAsBi presents some problems. The reaction products $AsCl_3$ and $GaCl_3$ are gaseous at the low pressures in a RIE system and will be carried out of the reaction chamber by the vacuum system. However the bismuth byproduct $BiCl_3$ of the process gas does not appear to be gaseous, but will redistribute back onto the surface, as shown in the micrograph in figure 4.3. Chlorine gas based etching will not work with GaAsBi, but there might exist another process gas that produces a gaseous byproduct with bismuth. RIE processes using HBr are also used to etch III-V semiconductors. The reaction product $BiBr_3$ appears to have a gaseous phase at the pressures and temperatures of the RIE process [72].

Because of this limitation of RIE for etching GaAsBi films all the samples in this thesis were fabricated using wet etching. The main components in III-V semiconductor wet etch chemistry are an acid and an oxidizer. A chemical such as H_2O_2 will oxidize the semiconductor surface to form As_2O_3, Ga_2O_3, or some other oxide compound. The acid will then remove the soluble oxide from



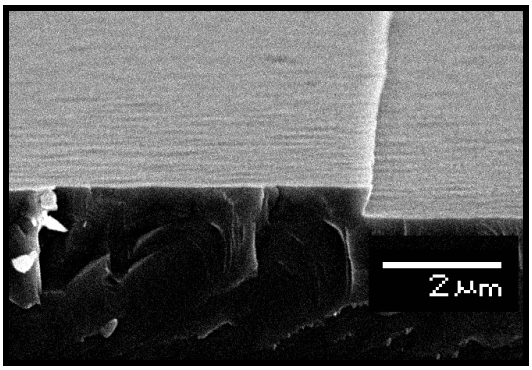
(a) AFM image of undercut etched bar

(b) SEM image of undercut etched bar

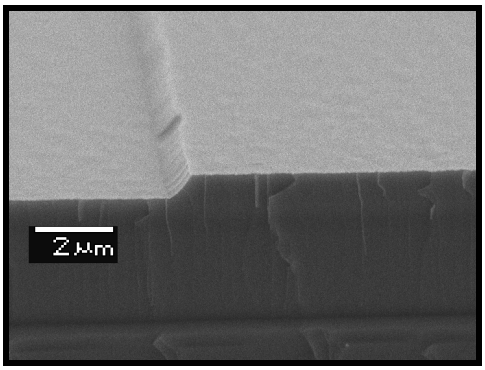
Figure 4.4: SEM and AFM images of wet etching of GaAsBi using H_2SO_4:H_2O_2:H_2O showing severe undercutting.

the surface and the oxidizer will once again react with the newly exposed surface [73]. Almost any acid and oxidizer combination will attack GaAs. A good wet etchant will have a reproducible etch rate and leave a clean, smooth surface. The first wet etchant tried on GaAsBi in this work was the H_2SO_4:H_2O_2:H_2O system. This is a standard etching solution in GaAs processing, with a controllable and reproducible etch rate by varying the ratios of acid, oxidizer, and water. In processing however this system caused severe undercutting of the photoresist which laterally etched away at the features, seen in the micrograph in figures 4.4a and 4.4b. Because of this lack of process control on feature size different etching chemicals were investigated. The $C_6H_8O_{7(aq)}$:H_2O_2 system [74] (citric acid and hydrogen peroxide) produced very well defined, crystallographic etch profiles, as seen in figure 4.5a and the 90° rotation in figure 4.5b. The citric acid is mixed from anhydrous powder into deionized (DI) water at 1 g:1 mL. This aqueous acid is then mixed with hydrogen peroxide at a ratio of 6:1 $C_6H_8O_{7(aq)}$: H_2O_2. This is the etchant used to fabricate all the bismide samples in this work.

A problem with wet etching in general that needs to be addressed is the etch rate variability of the solution. The first emitter etch in the six-step process is the most critical. In a typical case $\sim$ 300 nm to 350 nm of GaAs emitter must be etched away to expose $\sim$ 100 nm to 150 nm of the GaAsBi base. The more the solution etches into the base the higher the base resistivity. Etching too



(a) SEM image of etched ledge



(b) SEM image of etched ledge, 90° rotation from figure 4.5a

Figure 4.5: SEM images of wet etched GaAsBi using $C_6H_8O_{7(aq)}$:H_2O_2 showing crystallographic plane etching

far could cause the thinned base layer to become fully depletion from surface depletion. The layer thicknesses and composition are known from MBE growth calibrations and X-ray diffraction (XRD) measurements, therefore to etch to the correct thickness the etch rate must be well know. Even a small variation in the etch rate could cause the base to be completely removed. Slight changes in the exact ratio of mixed solutions could vary between different etchant solution batches, causing the rate to change. Even a known etch rate could change over weeks, days, or even hours as the volatile, unbuffered H_2O_2 changes in concentration. In order to address this processing problem a partial etch process was developed. Because the photoresist film is not attacked at all by the citric acid based solution the difference in photoresist film thickness, as measured by a profilometer, before and after the etch shows the amount of semiconductor material etched away. A partial etch can therefore be used to obtain the etch rate at the time of processing. Based on this the sample could be returned to the etchant to remove the desired amount of material. Typically only one or two profilometer measurements after the initial measurement are needed per etch step. This technique proved to be very robust for etching thin layers using only lab-mixed, non-commercial etching solutions.

The partial etch process is performed by first measuring the photoresist thickness after the lithography step to produce the etch mask. Initial etching of the sample is done to approximately

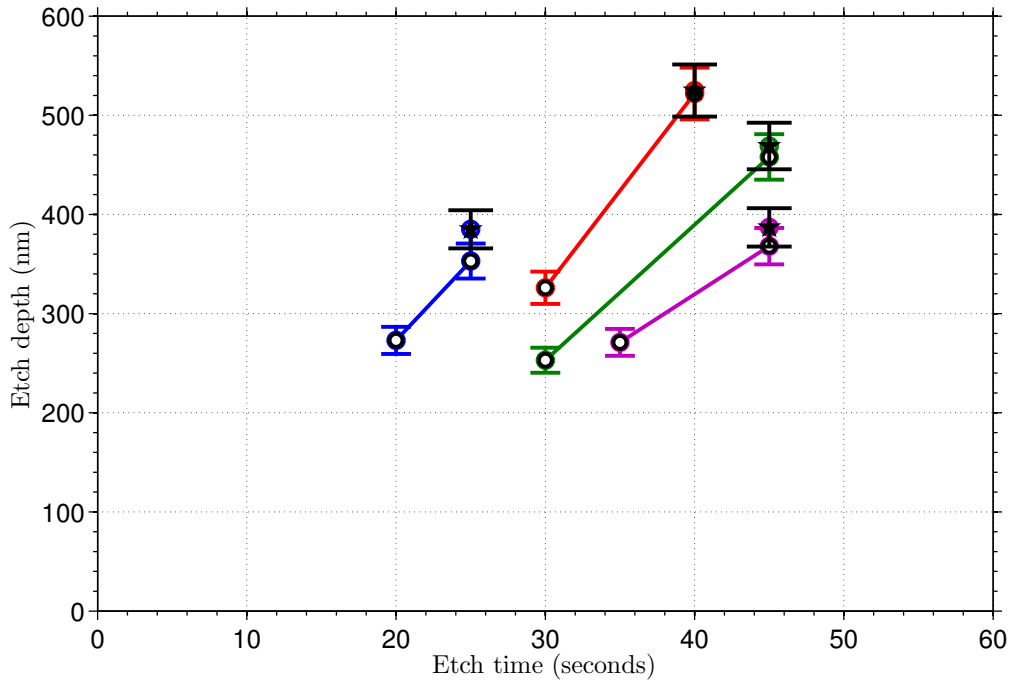
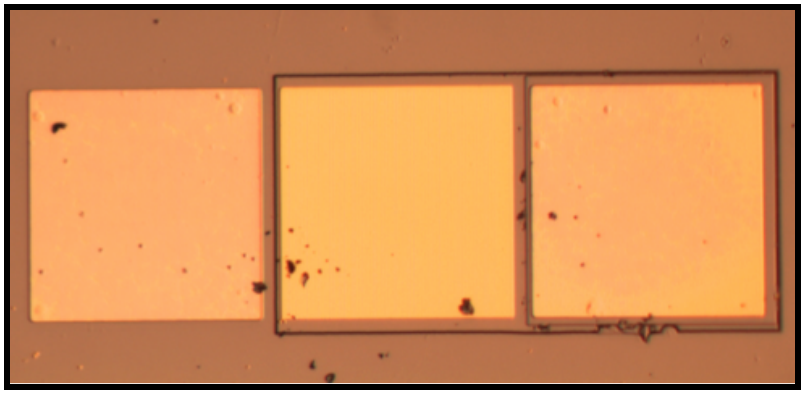
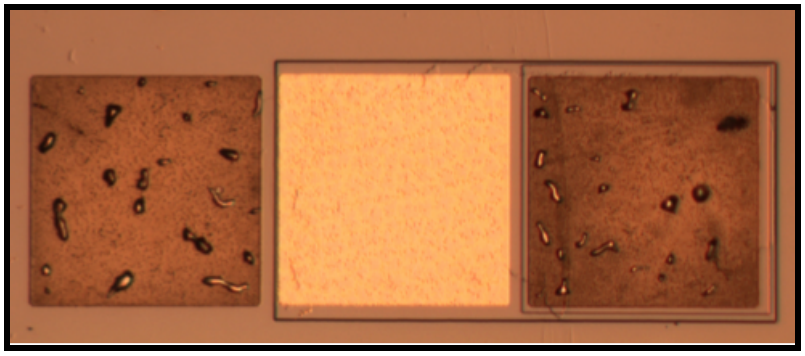


Figure 4.6: Etch depth versus time for $C_6H_8O_{7(aq)}$:H_2O_2 etchant system using partial etch technique. Datapoints marked as "∘" were measured during the etch process using the difference in photoresist thickness, while "★" datapoints are final etch measurements after photoresist removal.

half of the desired depth, based on previously observed etch rates. Before every etch the sample is cleaned in a 1:1 solution of HCl:H_2O to remove any surface oxide that could affect the etch rate. After the initial etch the photoresist thickness is measured again and the difference is assumed to be the etch depth. From this the instantaneous etch rate is found and used to compute the remaining time necessary to complete the etch. Once the photoresist thickness indicates that the etch is at the desired depth the photoresist is removed and the true etch depth of the sample is measured. Figure 4.6 shows the experimental etch depth versus time using the partial etch technique for various P-N and HBT GaAs/GaAsBi samples. Data points marked as "∘" were measured during the etch process using the difference in photoresist thickness, while "★" data points are final etch measurements after photoresist removal. Based on multiple thicknesses measured on the sample the profilometer measurement error was found to be $\sim 5\,\%$. The etch rate across all samples, with different etch solutions all nominally at 6:1 $C_6H_8O_{7(aq)}$:H_2O_2, is 413 ± 131 nm/min. The variability of the etch rate is $\sim 32\,\%$. With the partial etch technique the relative error in etch depth after PR removal is



(a) Ohmic metal contacts as deposited (unannealed)



(b) Ohmic metal contacts after RTA

Figure 4.7: Change in metal surface texture after rapid thermal annealing. Similar to the process described in [73]

only $\sim 10\,\%$, a factor of three improvement in accuracy.

4.3 Ohmic Contacts

The L3 and L4 metalization steps from figure A.1 in this project are to make ohmic contact between the metal layer and the semiconductor. At a metal-semiconductor interface in thermal equilibrium the Fermi levels in each material will align, causing the semiconductor's band to bend. If the work function of the metal and semiconductor are nearly the same there will be no energy barrier for carriers and the I-V response of the junction will be linear. For most metals and mid-range band gap semiconductors (band gaps of a few eV) the band alignment always creates an energy barrier for

transport, forming a rectifying Schottky junction. In this regime the conventional approach is to use a tunnel junction. By doping the semiconductor to a very high level ($> 10^{18}\,\mathrm{cm}^{-3}$) the barrier width can be made thin enough for carriers to tunnel through—forming a linear junction. In order to not limit what doping levels can be used in the HBT layer structure the area immediately underneath the contact is usually doped to these high levels, independent of the doping of the rest of the film [75]. This can be done many ways, including by thermal diffusion (difficult in GaAs due to As out-gassing) or ion implantation. The easier and more common way is to deposit metal contacts alloyed with a dopant and to anneal the metal post-deposition. This will dope an area directly under the contact but leave the doping level of the rest of the layer the same.

For n-type layers one of the oldest alloy metals for GaAs n-type ohmic contacts is AuGe [76, 77, 78]. Gold is a commonly used metal in III-V semiconductor because of its good electrical and chemical properties, while germanium is an n-type dopant in GaAs. The AuGe alloy has a eutectic melting point of 363 °C at the composition $\mathrm{Au}_{88\mathrm{wt.\%}}\mathrm{Ge}_{12\mathrm{wt.\%}}$ [79]. This low melting temperature facilitates the alloying of the metal and semiconductor junction and the diffusion of the Ge dopant into the semiconductor. To improve the alloying even more by changing the surface wetting properties a layer of Ni can be deposited as well onto the AuGe stack. A typical n-type GaAs contact might use $\sim 100\,\mathrm{nm}$ of AuGe and $\sim 20\,\mathrm{nm}$ of Ni. The contact then needs to be annealed to diffuse the dopants into the semiconductor. RTA is used to give the alloy sufficient energy to dope the semiconductor, but restricts the total time and thermal budget to avoid too much diffusion of the Ge or other dopants already in the layer. A typical RTA process anneals the metal at $\sim 350\,°\mathrm{C}$ to $450\,°\mathrm{C}$ for 0.5 min to 2 min. A micrograph of the metal surface both before and after annealing is shown in figure 4.7. The process window for this anneal is very wide, with many times and temperatures reported in the literature that produce a successful junction. Contact resistivities as low as $10^{-6}\,\Omega \cdot \mathrm{cm}^2$ have been reported [61].

P-type contacts also utilize alloyed metals to form ohmic contacts. All the p-type ohmic contacts in this work were made using the AuZn alloy system. Like AuGe, the AuZn alloy uses gold as a conductor and zinc as a p-type dopant for III-V contacts [80]. The composition typically

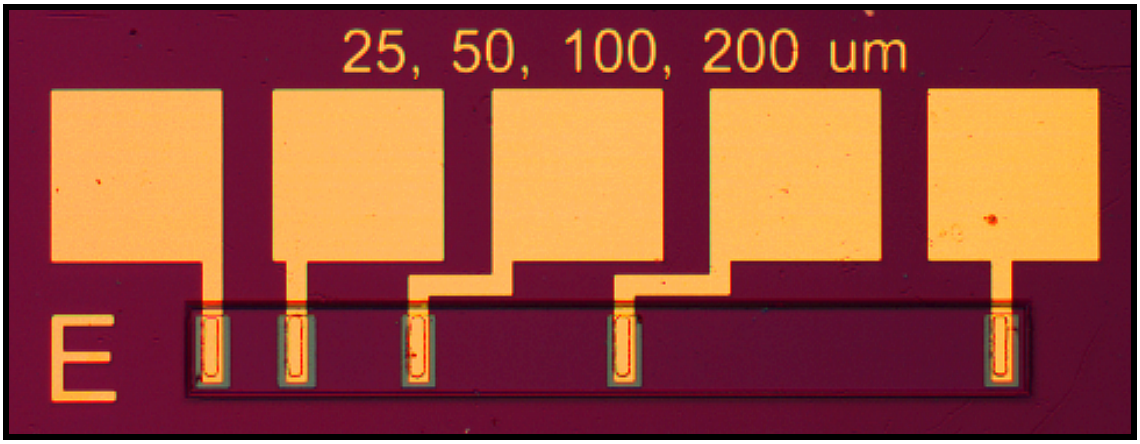


Figure 4.8: Micrograph of TLM via structure between the wiring metal (Ti/Au) and ohmic metal (AuGe/AuZn)

used is $Au_{90wt.\%}Zn_{10wt.\%}$. Although the melting point of this composition is much higher than the annealing temperature used to alloy the contacts, it is sufficient to diffuse the zinc dopants into the semiconductor to form an ohmic contact [81]. This alloy is either deposited as alternating gold and zinc layers, or evaporated using an entire charge of already alloyed material. The latter is used to make all the ohmic contacts in this work. In both the AuGe and AuZn contact systems the entire charge of the alloyed material is evaporated. Because of the different vapor pressures of the metals in the alloy, unless the entire amount is evaporated at once the composition deposited might not be the same as the original alloy. Also for both ohmic metal alloys a lift-off process is used to pattern the contacts. This avoids having to perform a metal etch on the sample, which could adversely affect the non-contacted semiconductor surface.

4.4 Encapsulation with Silicon Nitride and Wiring Metalization

To electrically contact a fabricated devices inside probe station metal pads are needed that are large enough to land a probe onto consistently. This requires a contact metal pad of at least 50 µm×50 µm, and preferably 100 µm×100 µm. So as not to limit the minimum size of the fabricated HBTs a wiring metalization step is added to separate the ohmic contact size from the metal-to-probe

contact pad. To do this there must be an insulating layer between the ohmic metal and semiconductor and the wiring metal layer, with etched via holes to electrically connect the layers. In addition to being an insulating layer for top-level metalization, depositing a dielectric such as silicon nitride has the potential to passivate the semiconductor and reduce surface states, or at least encapsulate the surface to prevent further oxidation after cleaning [82, 83]. Oxidation of III-V semiconductors produces surface states which causes an additional recombination current and reduces the HBT device gain [83].

Plasma enhanced chemical vapor deposition (PECVD) is used to deposit the silicon nitride film. The stoichiometric thin film composition Si_3N_4 has high residual compressive stress and can buckle under thermal cycling [84]. To counteract this a non-stoichiometric, low-stress PECVD recipe is used that incorporates hydrogen during deposition, resulting in a lower density but still insulating nitride thin film SiN_x [85]. The deposition is performed at 250 °C, below any temperature that could affect the semiconductor or metal contacts. A layer of SiN_x between 100 nm to 300 nm thick is deposited—thick enough to insulate properly but not so think that the metal step coverage is affected. The Ti/Au wiring metal is deposited through thermal evaporation, the same process as the ohmic metals. Because of its low reactivity, the adhesion of gold to most surfaces is poor. A layer of highly reactive titanium is deposited first to promote adhesion. Typically the wiring metalization consists of $\sim$ 10 nm to 20 nm of Ti and $\sim$ 80 nm to 120 nm of Au. A micrograph of this two-layer structure is shown in figure 4.8. Here the Ti/Au contact pads sit on top of a PECVD grown silicon nitride film with etched via holes to contact the ohmic metal on the TLM mesa.

Chapter 5

Characterization of Material and Devices

5.1 Material and Process Characterization

This chapter presents the measurement and characterization of fabricated devices in the GaAs/GaAsBi material system. Chapter 4 and appendix A covers how devices are fabricated and tested, while this chapter will cover the results and analysis of those measurements. More than 15 different complete process runs have been done with this material system, with almost as many different material growths from three different MBE systems. The different material origins will be referenced to as:

- P-N and HBT device structures grown from a MBE system at the National Renewable Energy Labortory (NREL)[6] , $x \sim 2\%$ to 3% **(Samples "A")**
- HBT device structures grown from a MBE system at the University of British Columbia[7] , $x \sim 0.95\%$ **(Samples "B")**
- P-N and HBT device structures grown from a new, dedicated bismide MBE system at NREL[8] **(Samples "C")**

with specific material properties detailed as needed.

[6] Grown by Aaron Ptak and Dan Beaton
[7] Grown by Ryan Lewis
[8] Grown by Kirstin Alberi and Dan Beaton

Table 5.1: HBT and material structures for samples A, B, and C

	Sample A	**Sample B**	**Sample C**
Emitter Cap (n^+-GaAs)	10^{18} cm^{-3}, 50 nm	1.5×10^{18} cm^{-3}, 61 nm	$> 10^{18}$ cm^{-3}, 10 nm
Emitter (n-GaAs)	10^{17} cm^{-3}, 300 nm	2×10^{17} cm^{-3}, 274 nm	$\sim 10^{17}$ cm^{-3}, 300 nm
Base (p-$GaAs_{1-x}Bi_x$)	10^{18} cm^{-3}, 300 nm, 2.3 % Bi	3×10^{18} cm^{-3}, 144 nm, 0.95 % Bi	$> 10^{18}$ cm^{-3}, 150 nm, $\sim$ 1 % Bi
Collector (n-GaAs)	2×10^{16} cm^{-3}, 300 nm	10^{17} cm^{-3}, 320 nm	$> 10^{17}$ cm^{-3}, 150 nm
Sub-collector (n^+-GaAs)	10^{18} cm^{-3}, 300 nm	1.5×10^{18} cm^{-3}, 290 nm	$> 10^{18}$ cm^{-3}, > 300 nm
Substrate	S.I. GaAs	S.I. GaAs	n^+-GaAs

5.1.1 Rapid thermal annealing of contacts

As discussed in section 4.3, rapid thermal anneal (RTA) is used to alloy the contacts to form ohmic junctions. There is a wide range of anneal conditions used in the literature, with almost every lab having their own unique recipe [86, 87, 61, 73]. Most of the anneal recipes use a temperature of ~ 450 °C for times ranging anywhere from 15 s to 180 s for RTA [86] to longer than 5 min to 12 min for furnace annealing [87]. Annealing at higher temperatures can increase the contact resistivity rather than reduce it, especially in alloyed contacts. Annealing AuGe at temperatures much larger than the eutectic temperature (363 °C [79]) will cause the alloy metal to segregate, forming germanium droplets, or cause unwanted diffusion of dopants too far into the semiconductor. Non-alloyed ohmic contacts, such as silicon-palladium-tungsten based refractory metals, or multi-metal stacks using platinum diffusion barriers, are sometimes used to increase device reliability [88].

Figure 5.1 shows the two-point I-V characteristics of an ohmic junction before and after annealing. This is the typical behavior of these junctions, going from a small non-linear current to a much larger linear current. Note the different scales for both I-V curves. To find the correct annealing

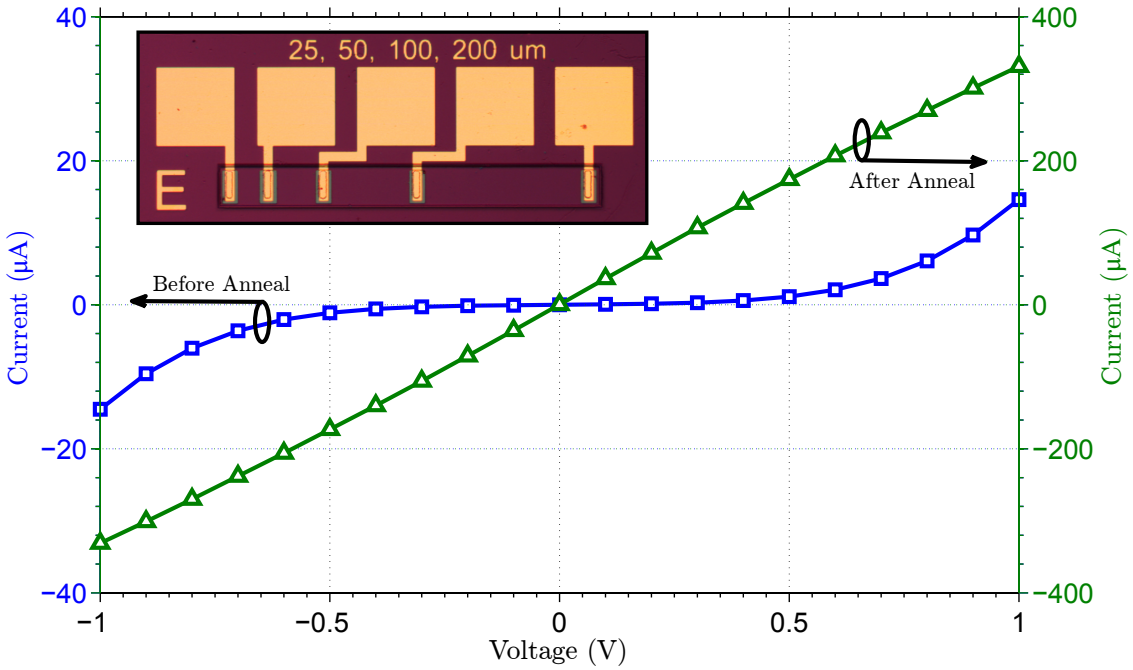


Figure 5.1: Two-point I-V of typical GaAs ohmic contacts, before and after RTA. Left-axis is 1/10 the scale of the right-axis.

time and temperature the two-point I-V is measured before and after each anneal. The anneal at which the junction I-V becomes linear is the correct process condition. As in GaAs, annealing GaAsBi contacts for longer or higher temperatures than this optimal could cause the P-N junction to short, either from dopant diffusion or metal spiking through the junction [73]. The RTA process conditions were found by starting with typical annealing conditions for GaAs and verifying the linearity. The annealing properties of GaAsBi are assumed to be similar to GaAs. The RTA process conditions used for fabricating devices in this work is a ramp to 350 °C and hold for 30 s, followed by ramp to 425 °C and hold for 30 s, then an unaided cool-down. All annealing is done in an argon atmosphere inside a vacuum chamber pumped to the mTorr range. Contact resistivities of $7 \times 10^{-5}\ \Omega \cdot \mathrm{cm}^2$ for p-type contacts and $3 \times 10^{-5}\ \Omega \cdot \mathrm{cm}^2$ for n-type contacts have been achieved in fabricated devices.

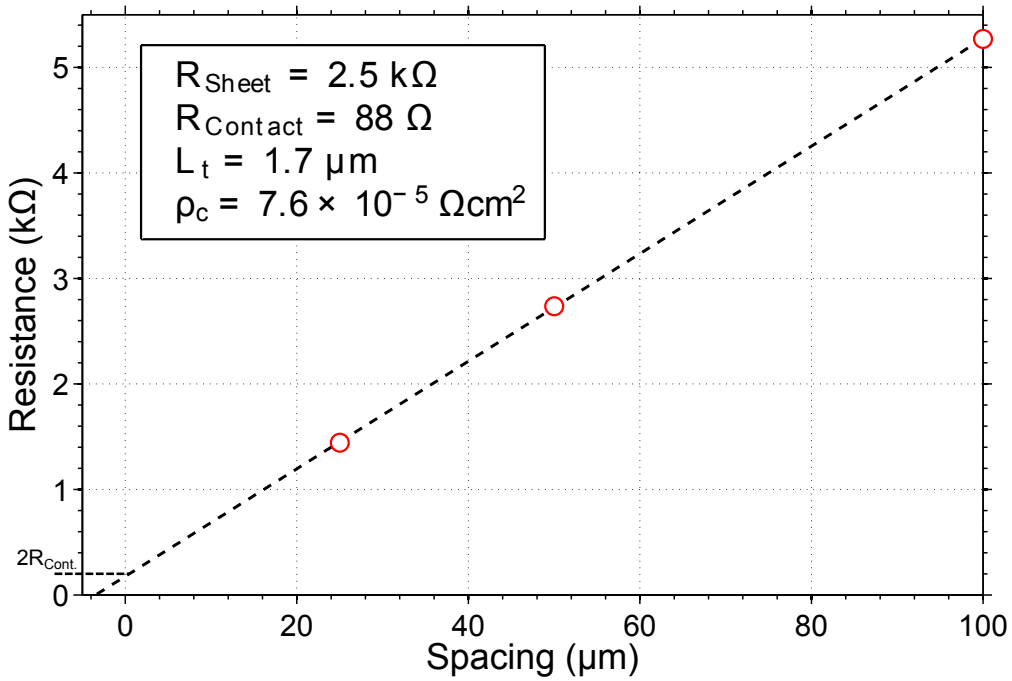


Figure 5.2: TLM measurement of GaAsBi base epi-layer. Sample B.

5.1.2 Resistivity, doping, and mobility

Two measurement techniques are used to characterize the resistivity, doping, and contact resistance of the material system. The transfer length measurement (TLM) technique is used to extract the sheet resistance and contact resistivity using the TLM structure discussed in section 4.1 and detailed in [69]. The mobility can then be inferred from these measurements. Figure 5.2 shows a typical TLM measurement of the GaAsBi base layer in sample B. L_T is the transfer length, a measure of how far into the junction before the voltage drop is $1/e$ of the initial value. A typical junction has contacts wider than this value, and beyond that the contact resistance does not become any smaller.

Capacitive–voltage (C-V) measurements are used to find the doping level and built-in voltage of a P-N junction. This technique works by modulating the depletion layer width and junction capacitance versus the applied voltage. The doping concentration can be extracted using the full-depletion model. With the independent measurement of the doping, the layer mobility can then be

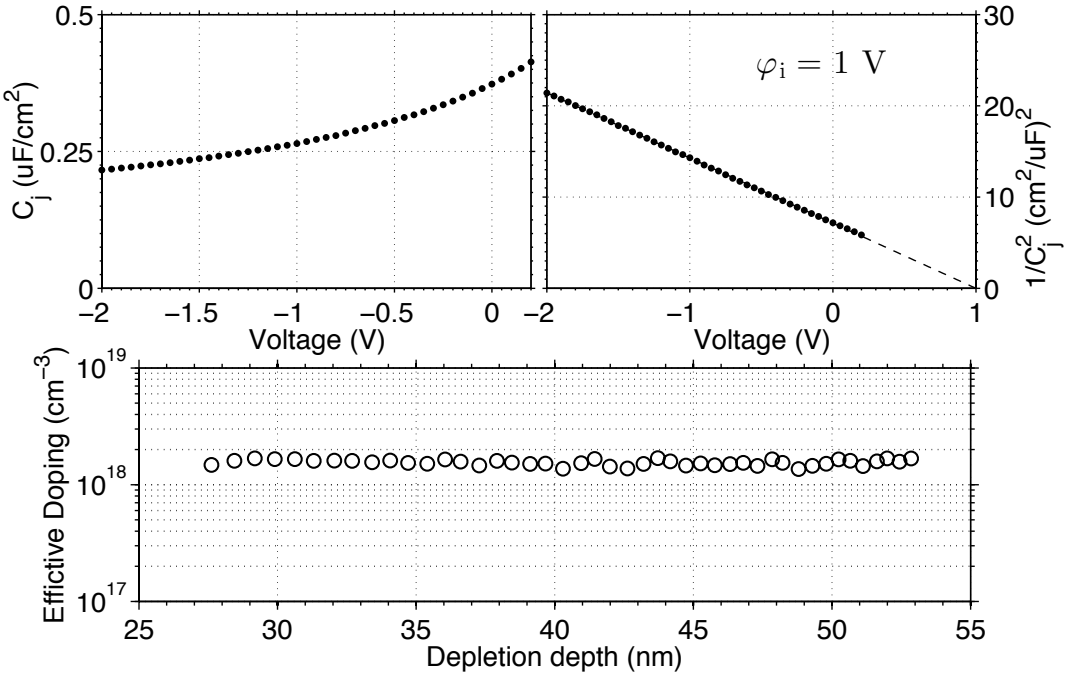


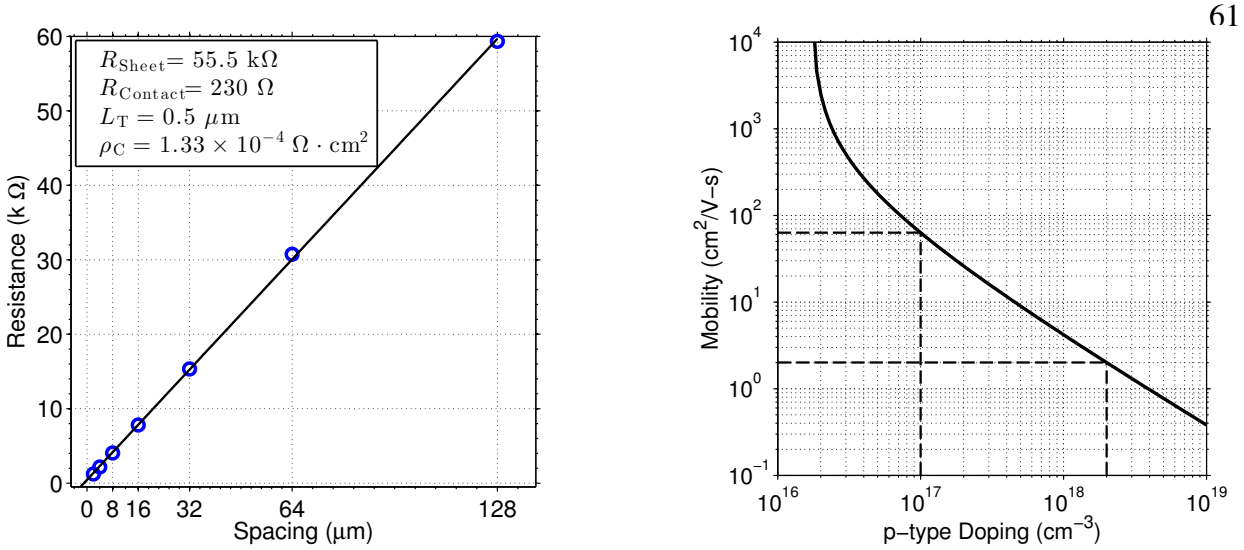
Figure 5.3: C-V measurement of GaAs/GaAsBi P-N junction showing a dopant concentration of $\sim 2 \times 10^{18}\,\mathrm{cm}^{-3}$, Sample A

inferred using

$$R_{\mathrm{Sheet}} = \frac{1}{q\mu N w'} \tag{5.1}$$

where the doping N is from the C-V measurement, R_{Sheet} is from the TLM, and the layer thickness w' is the junction effective width $w'_{\mathrm{p/n}} = w_{\mathrm{p/n}} - x_{\mathrm{p/n}}$, where $w_{\mathrm{p/n}}$ known from the growth and $x_{\mathrm{p/n}}$ is the depletion region width on either side. The C-V measurement of sample A, as well as the extracted doping concentration, is shown in figure 5.3.

There is an additional complication with these measurements when trying to infer the values of interest. The electrically active area of the junction is the effective width $w'_{\mathrm{p/n}}$, which is the junction width minus the depletion region width. For a highly doped junction this value is close to the layer thickness, but at lower doping levels a significant portion of the junction might be depleted from either junction or surface depletion. This introduces ambiguity into the inferred measurement as either the doping or the mobility or both could be lower than expected. Figure 5.4 shows the TLM plot for a GaAsBi layer from sample A. For the layer resistance found from the TLM in figure 5.4a,



(a) TLM

(b) Doping versus mobility for measured resistivity

Figure 5.4: Doping vs. mobility for measured TLM resistivity, Sample A

figure 5.4b shows the mobility versus doping concentration from equation (5.1), where the effective width is calculated as a function of doping. For this specific sample either the mobility is much lower than expected ($\sim 2\ \frac{\text{cm}^2}{\text{V s}}$ for the expected doping level of $2 \times 10^{18}\ \text{cm}^{-3}$), or the doping is much lower than expected ($\sim 10^{17}\ \text{cm}^{-3}$ for the expected mobility of $60\ \frac{\text{cm}^2}{\text{V s}}$ from figure 2.3). At first it seemed that the mobility was worse than expected in these samples, but the free carrier concentration might actually be much less than expected because of the compensation effect of localized bismuth states mentioned in section 3.2.4. The C-V measurements are performed under high reverse bias, so all the immobile dopants will be in the extended depletion region and will be ionized. This means that the measured dopant concentration and the actual free carrier concentration are different. This explains the inconsistency of an expected doping of $\sim 10^{17}\ \text{cm}^{-3}$ versus the C-V measured doping of $\sim 10^{18}\ \text{cm}^{-3}$. The conductivity is too low for the measured doping for expected mobility values. But the C-V test is actually measuring the total impurity concentration, not the free carrier concentration. An actual lower free carrier concentration (that contributes to conductivity) would be consistent with the expected mobility from the literature, as in figure 2.3. Figure 5.4b shows this trade-off. The

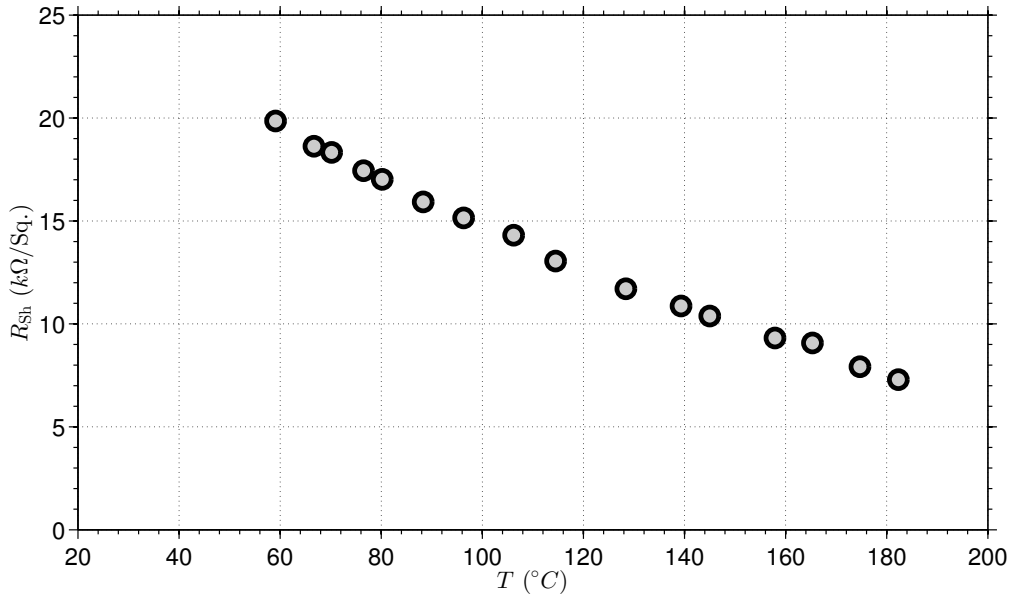


Figure 5.5: Four-point sheet resistance versus temperature of a GaAsBi/GaAs P-N heterojunction (Sample A)

actual values must lay between the indicated values.

5.1.3 Incomplete ionization and compensation

As introduced in section 3.2.4, there are numerous reports in the literature of energy states within the band gap near the VBM that are unique to GaAsBi. To experimentally verify the existence of these states, and to understand the possible compensation occurring that the measurements in section 5.1.2 suggested, the GaAsBi samples need to be measured at different temperatures. Like in figure 3.7a, the parameter of interest is the free carrier concentration as a function of temperature. This will identify the activation energy of any states—whether from shallow dopants or localized bismuth states. As a proxy for the carrier concentration the resistivity of the p-type bismuth layer was measured over various temperatures using a Peltier heater set-up. The resistivity was measured using a four-point probe technique on the TLM structure to remove the effect of any contact resistance.

Figure 5.5 is the measurement of the sheet resistance of the TLM structure versus temperature of a GaAsBi/GaAs P-N junction of sample A. There is a clear dependence of the resistivity on temperature. The resistivity is a function of both the free carrier concentration and the mobility, as

seen in equation (5.1). The dopant activation was modeled in equation (3.36) using the Fermi-Dirac distribution. The free carrier concentration will have an exponential dependence on inverse temperature for large activation energies ($E_A > k_B$ T), while the mobility will have a power dependence on temperature (e.g. $T^{-1.5}$). The exact power dependences of the mobility on temperature is difficult to find from first principles and will change with bismuth concentration. It is usually fitted from empirical Hall effect data. It is expected that the mobility will decrease with temperature, while the free carrier concentration will increase, if there is partial ionization. Because of the exponential increase in conductivity shown in figure 5.5, the effect of the temperature dependence of the mobility will be ignored and the conductivity will be assumed to be proportional to the free carrier concentration $G(T) \propto N(T)$. This assumption is justified because of the much stronger exponential dependence of the free carrier concentration with temperature—especially for the large activation energies expected to be seen. Additionally, just as the band gap dependence on temperature is reduced with the addition of bismuth, the mobility becomes less dependent on temperature at higher alloy concentration as the bismuth becomes the dominant scattering mechanism.

Taking the data from figure 5.5 and plotting the logarithm of the conductivity versus inverse temperature produces the Arrhenius plot in figure 5.6. There is a clear exponential dependence of the conductivity, and therefore the carrier concentration on temperature. The activation energy is fitted to the function $G(T) \propto N(T) \propto \exp\left(-E_A/k_B T\right)$. Fitting this curve results in an apparent activation energy of $E_A = 98.0 \pm 8.1$ meV, which is consistent with the values reported in the literature for GaAsBi [51, 52].

5.2 DC Device Testing

5.2.1 P-N junction testing

Because of the narrower band gap of bismide the P-N junction I-V of a GaAs/GaAsBi heterojunction will require a lower “turn-on” voltage to produce the same current density as an equivalent homojunction. The turn-on voltage is defined as the forward bias voltage required to induce a certain

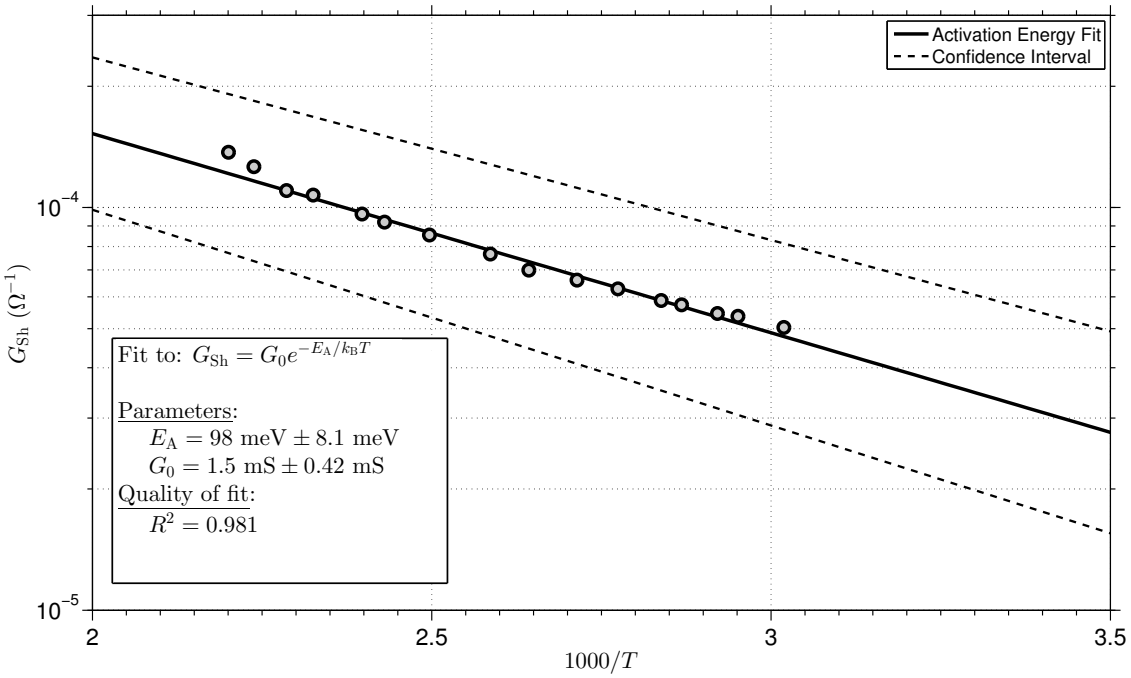
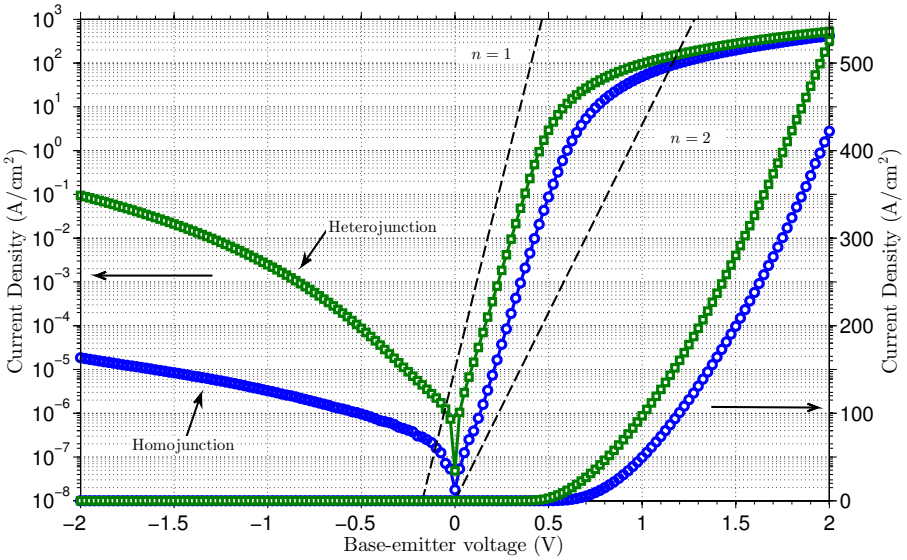
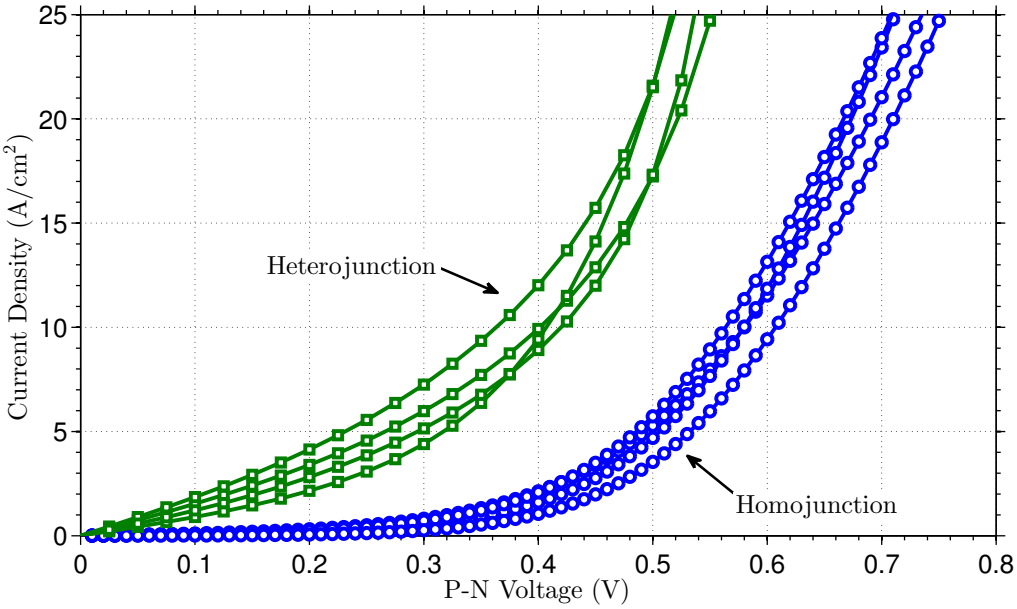


Figure 5.6: Arrhenius plot of four-point sheet conductivity of a GaAs/GaAsBi P-N heterojunction to extract activation energy (Sample A). Extracted activation energy is $E_\mathrm{A} = 98.0 \pm 8.1$ meV.

current across a P-N diode. The turn-on voltage shift is the difference in voltage between two diodes required to produce the same current density. Sample sets "B" and "C" have both homojunction and heterojunction diodes. The base-emitter P-N junction I-V of GaAs/GaAsBi heterojunctions and GaAs homojunctions for **sample B** are shown in figure 5.7a. In sample B both junction diodes have an ideality factor between $n = 1$ and $n = 2$, indicating ideal drift–diffusion current through the junction with a depletion region modulation effect, and possibly surface recombination. An ideality factor of $n = 2$ would indicate that the current is dominated by SRH recombination current in the depletion region. The two most notable differences between the heterojunction and homojunction diodes of sample B are the reverse bias leakage and the turn-on voltage. The reverse bias leakage current is larger in the GaAs/GaAsBi heterojunction than in the GaAs homojunction. Both of the reverse bias currents are from generation of carriers in the depletion region through SRH generation (inverse of forward bias recombination). The larger current in the GaAs/GaAsBi junction is consistent with the higher intrinsic and minority carrier concentration in GaAsBi from the narrower band



(a) Base-emitter junction diode, sample B. $\Delta V_{\text{Turn-on}} \approx 125\,\text{mV}$



(b) P-N junction diode, sample C. $\Delta V_{\text{Turn-on}} = 174.0 \pm 20.4\,\text{mV}$ at $10\,\text{A/cm}^2$

Figure 5.7: Homojunction and heterojunction P-N diode I-V plots.

gap. Under forward voltage bias there is a clear shift towards lower voltage for the same current level for the heterojunction versus the homojunction. This turn-on voltage shift of ~ 125 meV is the same sign and approximately equal to the expected band gap shift of ~ 100 meV for 0.95 % bismuth in sample B, plus any differences in life-time, mobility, or free carrier concentration. The voltage shift is also affected by any difference in doping, thickness, or mobility between the homojunction and heterojunction materials, but aside from the mobility those factors are approximately the same between the two junctions. The current roll-off at high forward bias is from the series resistance voltage drop. Both of the junctions in sample B have a semi-insulating (SI) substrate, so the current is lateral through the epitaxial base layer.

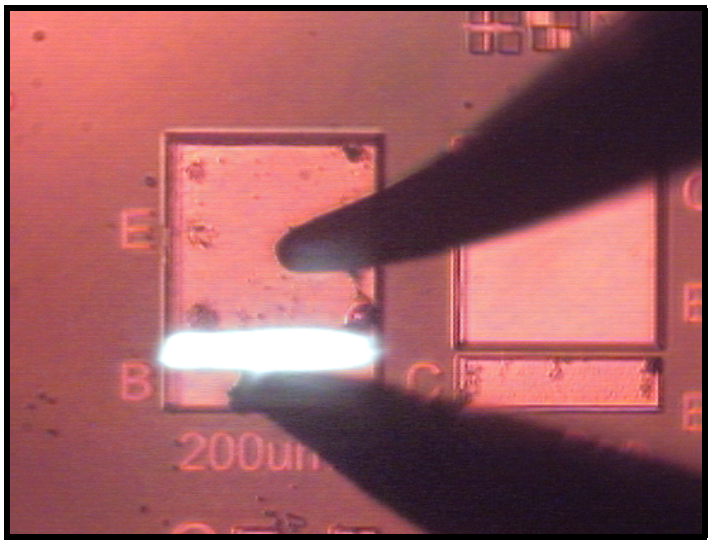


Figure 5.8: Light emission from a forward biased GaAsBi/GaAs P-N heterojunction (Sample B). Images with microscope light on and off overlaid.

The P-N junction I-V of GaAs/GaAsBi heterojunctions and GaAs homojunctions for **sample C** are shown in figure 5.7b. Both are P-N junction diodes with n^+-GaAs substrates and p-type GaAs or GaAsBi layers on top. All currents are vertical through these samples. For this sample set I-V measurements were taken on diodes with different contact area. All currents are normalized to the junction area. Like sample B, the heterojunction in sample C has a lower turn-on voltage than the homojunction, as well as increased reverse bias leakage. The difference in turn-on voltage between

the junctions is $\Delta V_{\text{Turn-on}} = 174.0 \pm 20.4\,\text{mV}$ at $10\,\text{A/cm}^2$, again consistent with the band gap shift of GaAsBi in sample C.

Figure 5.8 shows a micrograph of light emission from a forward biased GaAsBi/GaAs P-N heterojunction of sample B. Images with the microscope light on and off are overlaid for clarity. The emission of light is from radiative recombination in the depletion region of the forward biased P-N junction. This light emission can only of come from minority carriers. The emission wavelength or power output was not determined. Confirmation of minority transport is a necessary condition for HBT operation, even if other effects are dominant.

5.2.2 Heterojunction bipolar transistor testing

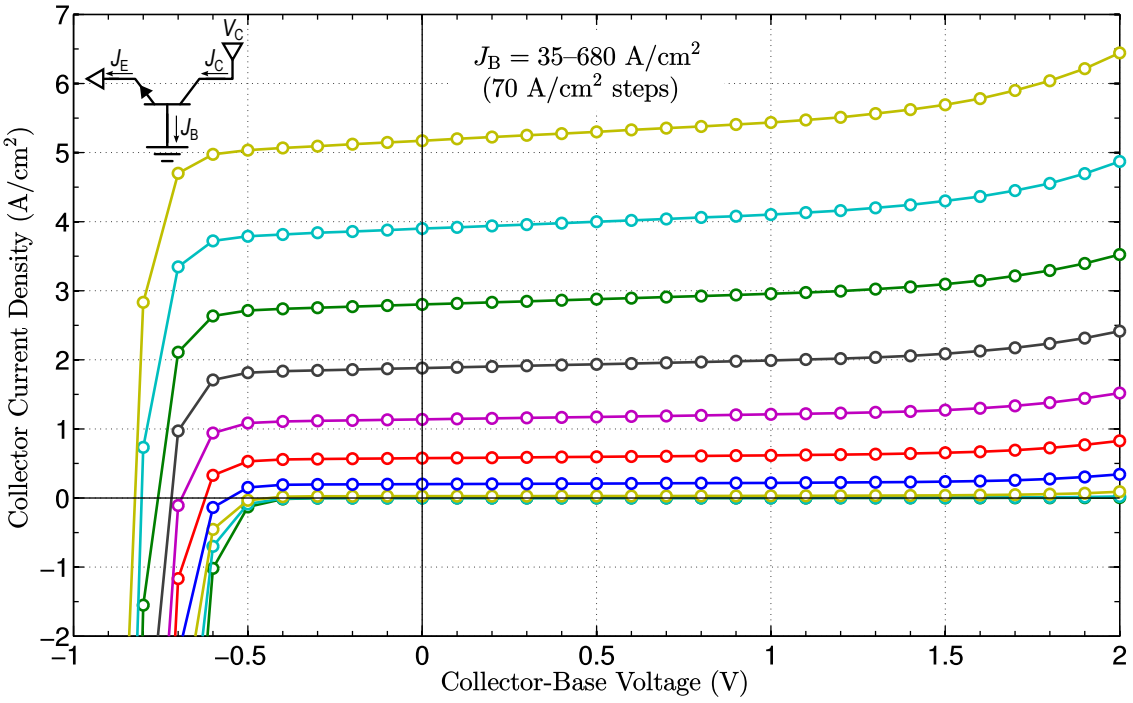
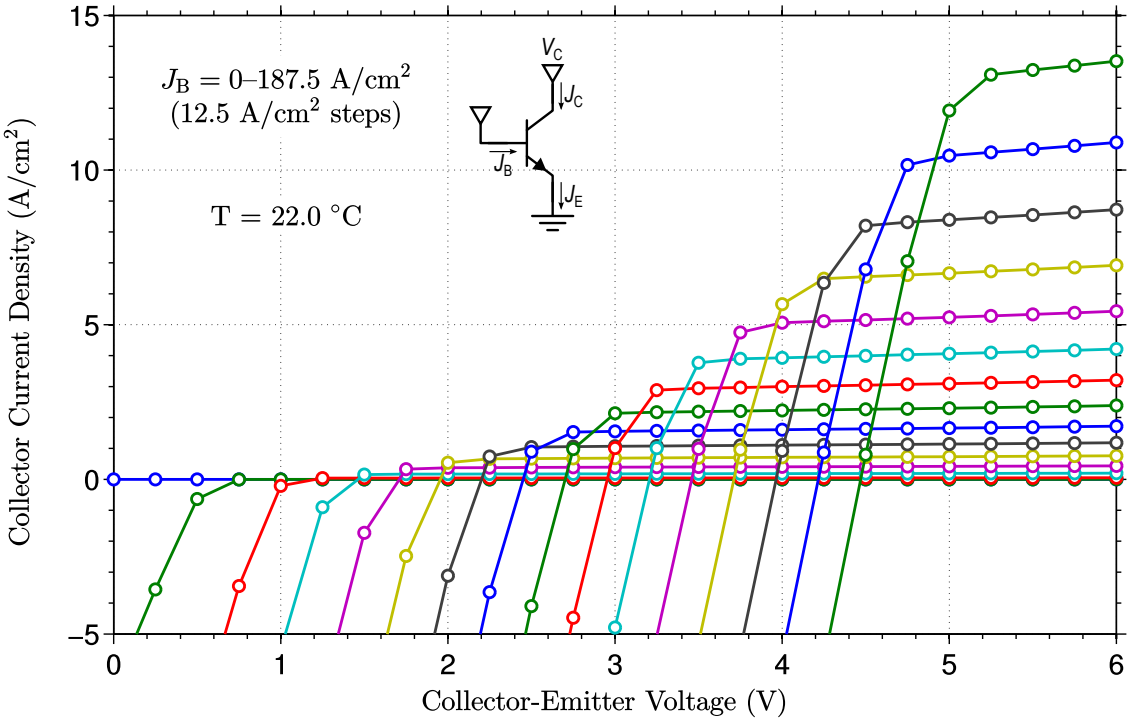
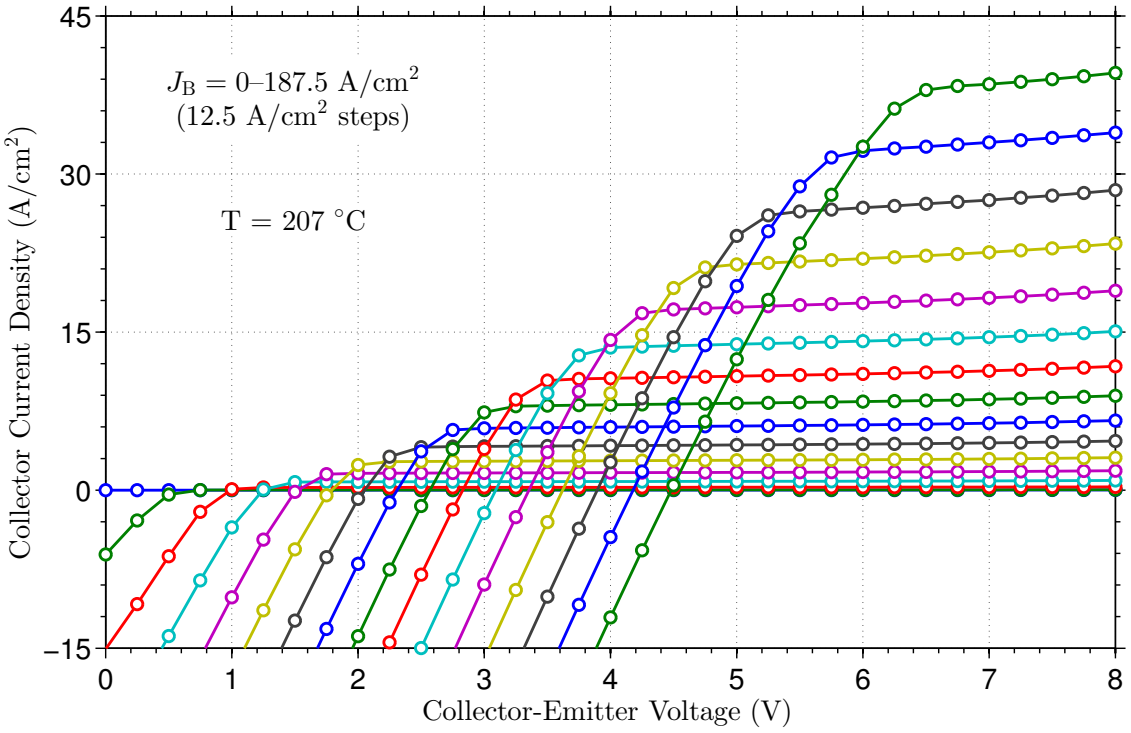


Figure 5.9: Measured HBT common-base I-V plot (Sample B). Diode area is $20\,\mu\text{m} \times 100\,\mu\text{m}$.

The measurements in figures 5.9, 5.10, and 5.11 are the first results of a fabricated HBT in the GaAs/GaAsBi material system. The HBT structures of **sample B** were made using the generation four mask-set (see figure A.2) in the six-step fabrication process in figure A.1. The gain of the



(a) Measured at 22.0 °C



(b) Measured at 207 °C

Figure 5.10: Measured HBT common-emitter I-V plot at low and high temperature (Sample B). Diode area is 20 µm × 100 µm.

HBT device is less than one ($\beta < 1$) likely from additional recombination mechanisms such as surface recombination or interface states in the base-emitter heterojunction. Despite the low gain the device does demonstrate minority carrier transport and transistor-like operation. Figure 5.9 is the common-base measurement of the sample B HBT device. Here the collector current density is plotted while the base terminal is grounded and the collector-base voltage is swept for various emitter current densities. Most of the injected emitter-base current is due to the afore mentioned recombination currents, as well as unintentional leakage due to non-uniform injection. The current shown in figure 5.9 and the collector current in figure 5.11 is the fraction of injected electrons that make it through the base into the collector. Figure 5.8 showed that there is minority transport occurring in the device. The collector current in figure 5.11 is the fraction of this minority carrier current that reaches the base-collector junction. Here the small minority transport fraction means that the gain and transport factor are roughly equal to $\beta \sim \alpha \sim 0.74\,\%$.

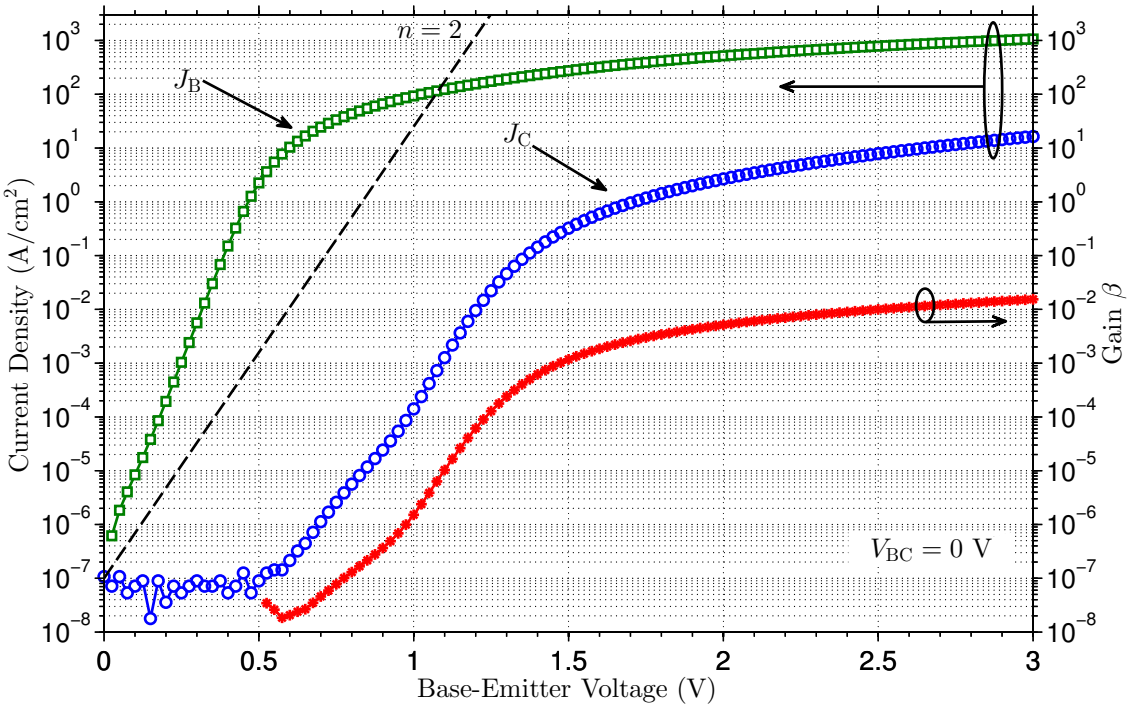


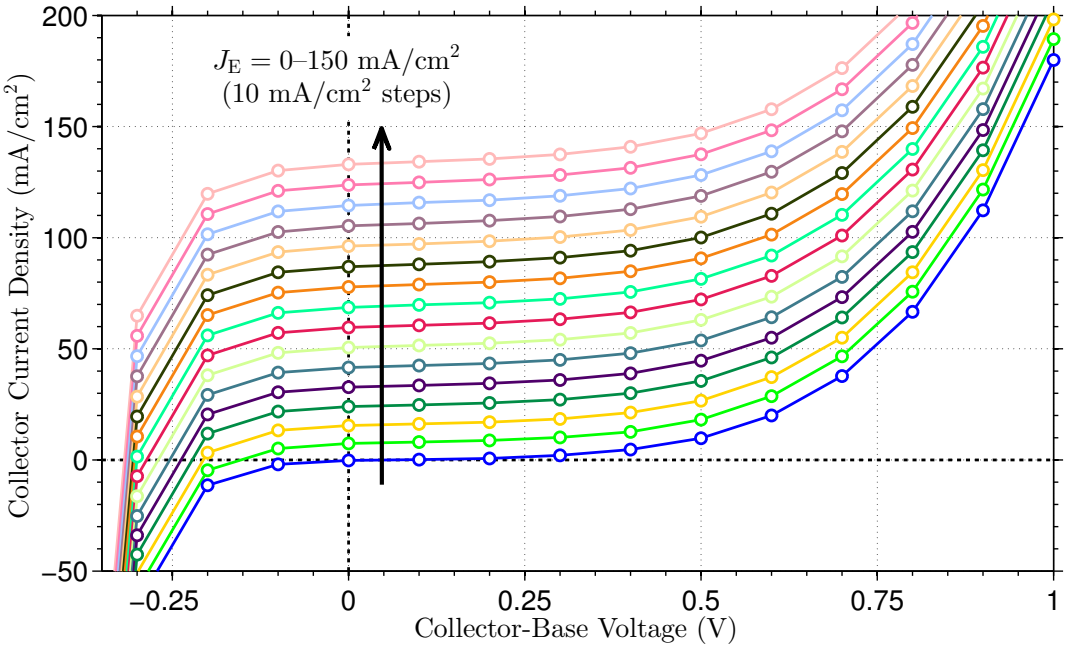
Figure 5.11: Measured HBT Gummel I-V plot (Sample B). Area is 20 µm × 100 µm.

The common-emitter configuration measurement for the sample B HBT at two different

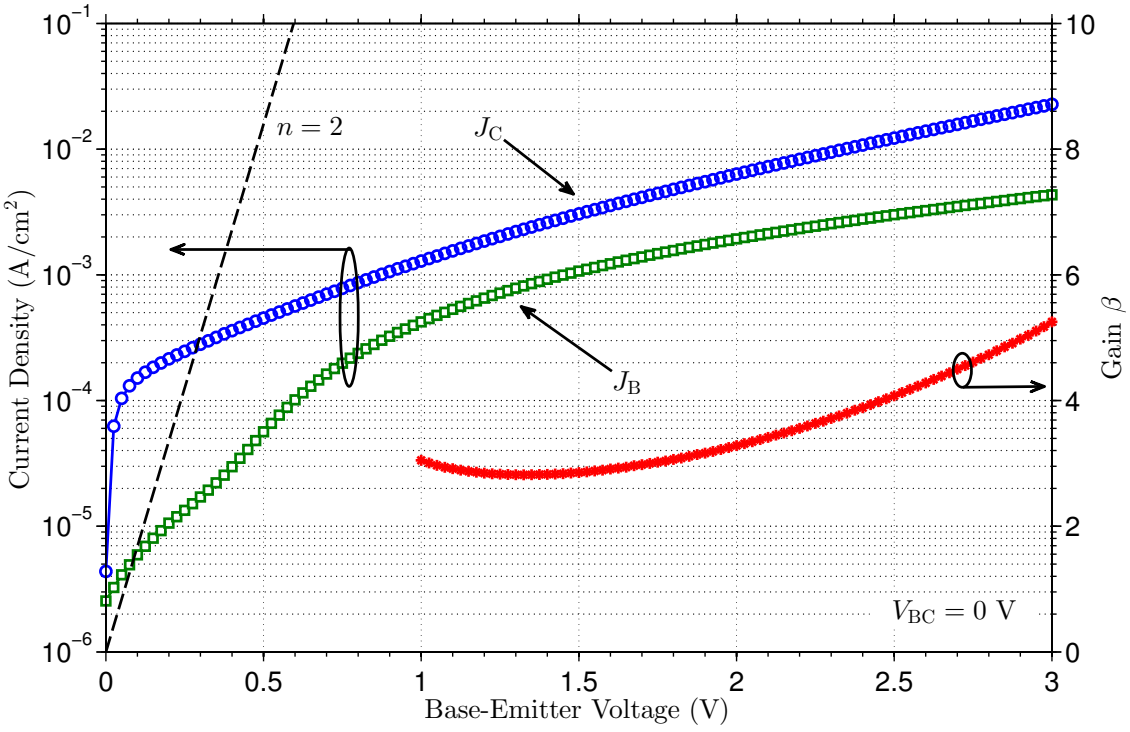
temperatures is shown in figure 5.10. In this setup the collector current is again shown while the emitter terminal is grounded and the voltage on the collector is swept for various base current densities. Once again most of the applied base current does not go towards minority carrier injection but is due to depletion region or surface recombination currents. However the fraction of the current from successful base minority transport demonstrates typical HBT behavior. The forward voltage shift in the collector current is caused by the high lateral base series resistance. The lowest current line in figure 5.10 is the zero base-current sweep and shows the amount of collector-emitter leakage occurring in the transistor. The gain of the HBT is larger for the higher temperature measurement. The change with temperature of most material parameters will reduce the gain at higher temperatures, but the device in figure 5.9 shows the opposite behavior. If the device is limited by the base transit factor (as in table 3.1) rather than emitter efficiency then the dominant temperature effect will be the linear increase in diffusion constant ($D_{\mathrm{n}} = k_{\mathrm{B}} T \mu_{\mathrm{n}}$), increasing the gain.

Figure 5.11 is the Gummel plot of the base and collector current at $V_{\mathrm{BC}} = 0$. For $\beta > 1$ the collector current should be larger than the base current. In the sample B device the base current is much larger than the collector current and increases much earlier. The fraction of collector to base current is the gain of the HBT. Here the effect of reverse bias leakage current from the collector into the base is minimized since the base-collector junction is biased at zero volts. However some leakage can still occur due to non-uniform injection caused by the large series resistance. The main non-ideal effects are then the recombination current in the base-emitter depletion region and at the surface, as well as the recombination in the quasi-neutral base layer. All the previous P-N junction I-V measurements (such as in figure 5.7a) shows that each junction is good individually, but it is the poor minority transport that causes the low gain—either from greater than expected recombination or a much lower minority carrier lifetime in the base.

The device made from the latest material growth in **sample C** is shown in figure 5.12. The junctions in the sample C device, especially the base-collector junction, exhibit a lot of leakage. This leakage prevents good common-emitter measurements. However both common-base and Gummel measurements can be made, and for the first time a GaAs/GaAsBi HBT has been measured with an



(a) Common-base measurement



(b) Gummel plot

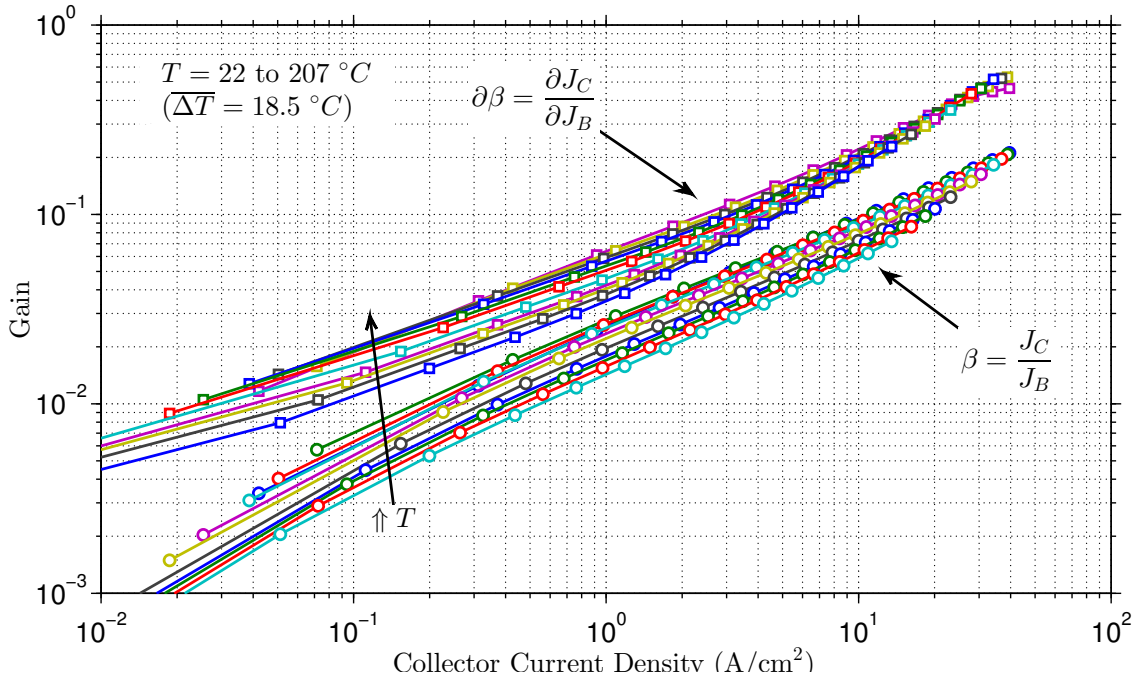
Figure 5.12: I-V measurement of 100 µm × 100 µm HBT device on sample C

apparent gain of $\beta > 1$. Both the sample B and sample C devices have similar epitaxial structures. The main differences, besides being grown at different times in different systems, is that sample C likely has a larger bismuth concentration.

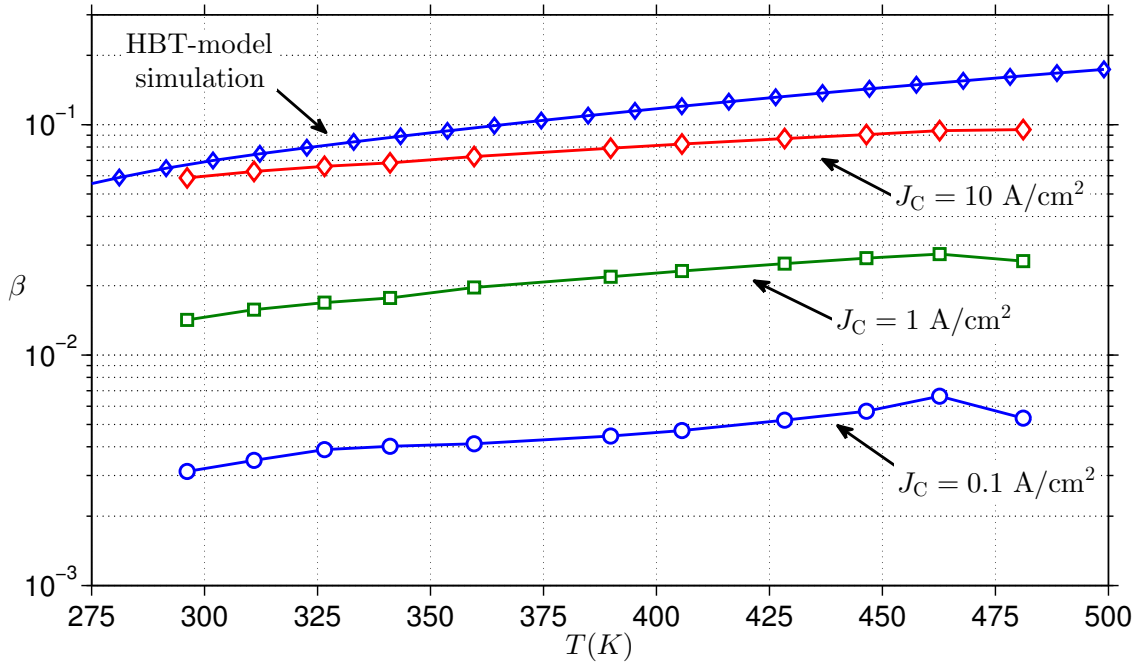
5.2.3 Minority carrier transport

To understand what is happening in these two samples the gain as a function of current density and temperature can be analyzed. Starting with **sample B**, the HBT current gain can be directly extracted from the common-emitter plot in figure 5.10. The gain is plotted against collector current density at various temperatures in figure 5.13a. Both the standard ratio gain and the differential gain are displayed. The differential gain will remove the effect of leakage current from the reverse biased base-collector junction. The gain of the HBT increases for both higher collector current densities and at higher temperatures, reaching a maximum of $\beta \sim 10^{-1}$. This is a strong indication that the base transport α_{T} is dominating the gain. Nearly every parameter of the HBT will reduce the gain at higher temperatures. However the transport factor is inversely proportional to the diffusion length squared, and therefore the diffusion constant ($L_{\mathrm{n}}^2 = \tau_{\mathrm{n}} D_{\mathrm{n}}$). Since the diffusion constant is directly proportional to the temperature $D_{\mathrm{n}} = k_{\mathrm{B}} T \mu_{\mathrm{n}}$ the diffusion length, and therefore the gain will increase with temperature.

Assuming that the emitter efficiency γ_{E} is close to one, and that all the carrier loss is the result of minority carrier recombination in the base, then from the transport factor equation in table 3.1 and using $D_{\mathrm{B}} \sim 1\,\mathrm{cm}^2/\mathrm{s}$ and $w'_{\mathrm{B}} = 144\,\mathrm{nm}$ for sample B then $\tau_{\mathrm{B}} < 18\,\mathrm{ps}$. This is the lifetime for the maximum gain at high emitter current levels. In this range the assumption to use the simple diffusion model in table 3.1 is reasonable. The derived lifetime value will be even smaller for other measurements done at lower current levels, like in figure 5.9. Even in severely compensated material this kind of minority carrier lifetime is unreasonably small. The minority carrier lifetime in GaAsBi has been measured to be $> 1\,\mathrm{ns}$ at low bismuth concentrations (such as the 0.95 % bismuth concentration in sample B), and still $> 100\,\mathrm{ps}$ at greater than 6 % bismuth concentration [14]. Even in MBE grown LT-GaAs the lifetime does not enter the sub-picosecond range until



(a) Regular and differential measured gain for Sample B HBT at different temperatures



(b) Gain versus temperature at different J_C, with HBT-model simulation of sample B

Figure 5.13: Measured HBT common-emitter I-V plot at low and high temperature (Sample B)

growth temperatures drop below 200 °C [89], much lower than the growth temperature of sample B.

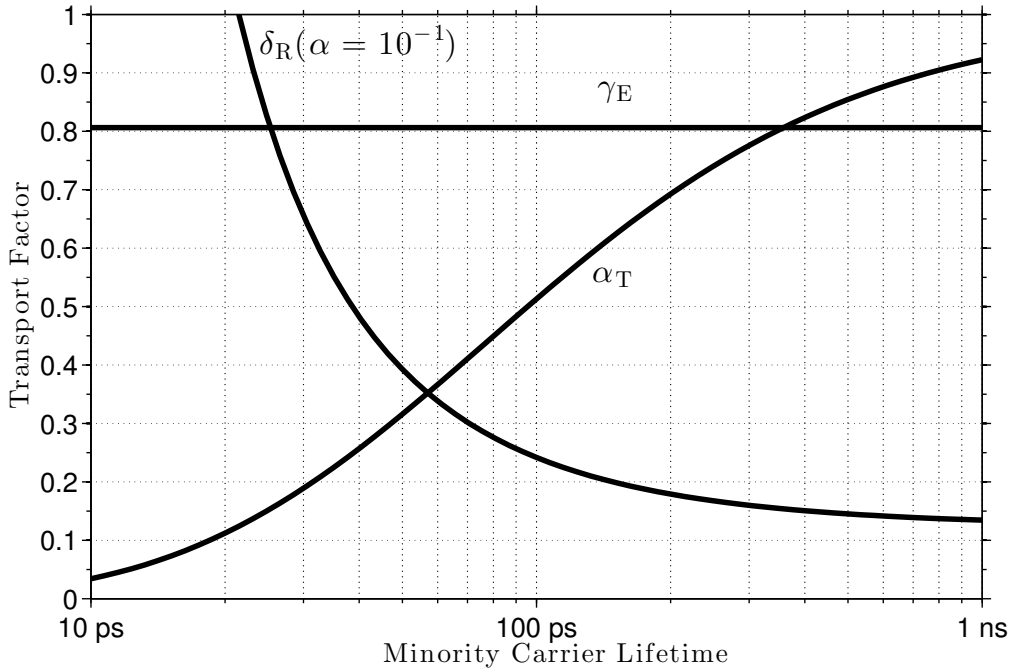
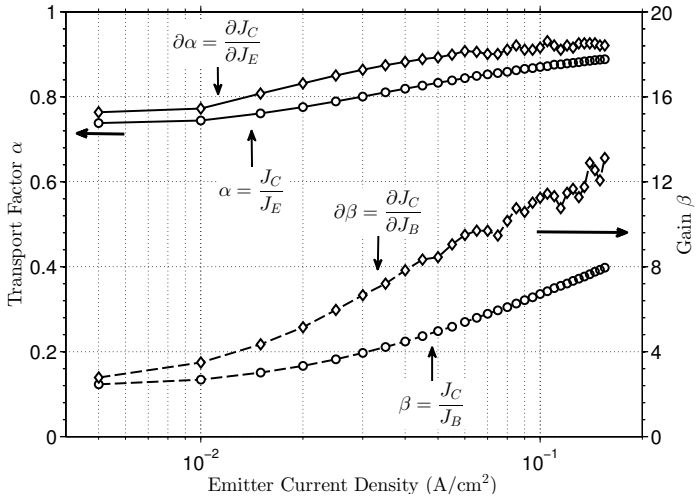


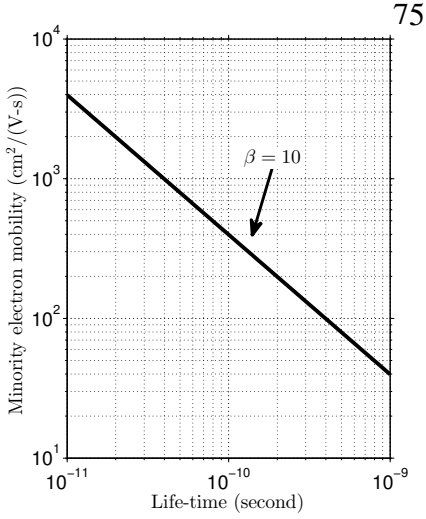
Figure 5.14: Transport factors for sample B device as a function of minority carrier lifetime in the base. The recombination factor δ_R calculated from γ_E and α_T as a function of lifetime and assuming a constant total transport factor of $\alpha = 10^{-1}$

If minority carrier recombination in the base of the sample B device is not the cause of the low gain, then the likely cause is excessive recombination current in the base-emitter depletion region or surface recombination. From table 3.1 the total transport factor is $\alpha = \gamma_E \alpha_T \delta_R$. Using the known layer properties for sample B the different transport factors as a function of carrier lifetime in the base is shown in figure 5.14. The recombination factor δ_R is computed using a constant total transport factor $\alpha = 10^{-1}$. This gives a lower limit to the lifetime to achieve the gain seen at high current levels in figure 5.13. The lifetime cannot be less than 20 ps or the gain seen would not be possible even with $\delta_R \to 1$. For more typical lifetimes of > 100 ps recombination is the dominant mechanism limiting the transistor gain. Either the lifetime in the base is much shorter than previously seen or expected in GaAsBi, or recombination is dominating the transport in the sample B HBT. Further proof of this will be explored in section 5.2.4.

In **sample C** the gain and transport factor as a function of emitter current density is plotted in



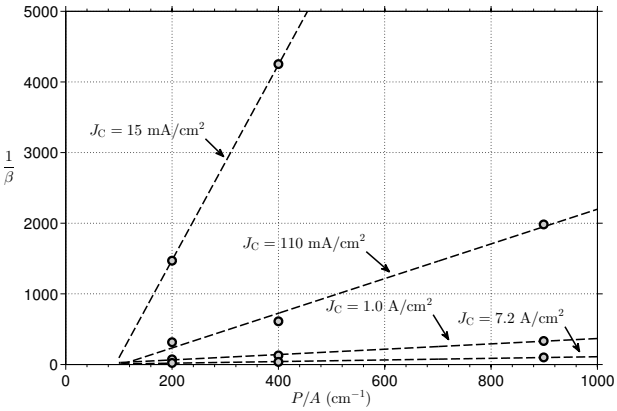
(a) Extracted gain and transport factor



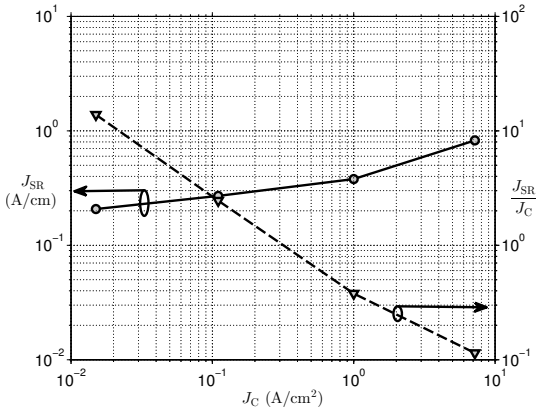
(b) Mobility versus life-time

Figure 5.15: Gain analysis of 20 µm × 100 µm sample C HBT. Calculated assuming constant gain ($\beta = 10$, $\alpha = 0.9$), no recombination ($\delta_R \to 1$), and a constant emitter efficiency.

figure 5.15a. In the measurement of figure 5.12a both the collector and base currents are directly measured for a given emitter current. This allows a direct measurement of both the gain and transport factor. Once again both the ratio and differential form of these parameters is calculated. Since the sample C device has a higher bismuth concentration the emitter efficiency is even less of a limiting factor than for sample B. Ignoring for now the effect of recombination ($\delta_R \to 1$), the gain seen in the sample C device can be explained entirely on the base transport, α_T. Both the electron lifetime in the base, and the diffusion constant (which is a function of the minority carrier mobility) influence the base transport factor. For a gain of $\beta = 10$ the possible values of minority electron lifetime and carrier mobility in the base of sample C is shown in figure 5.15b. The device either has a good lifetime and poor mobility ($\tau \geq 1$ ns, $\mu_n \leq 40\,\mathrm{cm^2/(V\,s)}$), poor (but not unreasonable) lifetime and good mobility ($\tau \geq 20$ ps, $\mu_n \leq 2000\,\mathrm{cm^2/(V\,s)}$), or some value between these two extremes. An electron mobility of $2000\,\mathrm{cm^2/(V\,s)}$ is the expected value for pure GaAs and dilute GaAsBi from figure 2.3.



(a) P/A versus $1/\beta$ to extract surface recombination current



(b) J_C versus J_{SR} and surface recombination current fraction

Figure 5.16: Extracted surface recombination current, Sample B

5.2.4 Surface recombination

As mentioned in section 4.1 transistors with different area to perimeter ratios can be used to estimate the proportion of surface recombination current. The technique is described in more detail in [90, 70]. The technique assumes a simplified current model for the base and collector, with $J_C = J_{n,C}$ and $J_B = J_{p,B} + J_{r,d} + J_{SR}$ where J_{SR} is the surface recombination current density and the total surface current is proportional to the device perimeter. Using the definition of the gain β and taking the inverse

$$\frac{1}{\beta} = \frac{AJ_{p,B} + AJ_{r,B} + PJ_{SR}}{AJ_{n,C}} = \frac{1}{\beta_0} + \frac{P}{A}\frac{J_{SR}}{J_{n,C}} \tag{5.2}$$

This means that plotting $1/\beta$ versus the device perimeter to area ratio P/A will show a linear line whose slope is the ratio of surface recombination linear current density to the collector current density. This plot is shown in figure 5.16a. On mask-set four for sample B there are three different sized transistor structures: the via-contacted HBTs, the large area HBT, and the very large capacitor test structure. The gain in all these devices is $\beta < 1$, so $1/\beta$ is very large. This makes finding the y-intercept, and thus the intrinsic gain β_0 difficult, but the slope and thus the surface recombination fraction will still be an accurate estimate.

Figure 5.16b is the extracted linear current density J_{SR} and surface recombination fraction

J_{SR}/J_C versus the collector current J_C. At low current levels the surface recombination is almost 10 times larger than the collector current. At higher current levels this fraction drops to 0.1. This is consistent with the gain improving at higher current levels seen in figure 5.13. This is a simplified model, and there is likely additional recombination occurring that is reducing the gain. This analysis does point to surface recombination being one of the primary factors in the reduced device gain of sample B.

Chapter 6

Conclusion

6.1 Summary

In this work the first HBT made using a GaAs/GaAsBi heterojunction has been simulated, fabricated, and tested. The history and development of the dilute bismide GaAsBi III-V material system, as well as its latest material property data has been reviewed from the literature. Empirical models for the mobility and lattice-constant change as a function of bismuth concentration have been developed. The feasibility and potential of the material system for use in electrical and optoelectronic semiconductor devices was discussed, and the material was compared to other similar III-V compounds.

A detailed model of a HBT utilizing the specific properties of GaAsBi was developed and implemented to understand the performance potential of this new material for device applications. The gain and frequency potential of a device with the GaAsBi material was found, as well as the design trade-offs necessary for certain applications. In addition to simulating the operation of a generic transistor, the model incorporated details specific to the GaAs/GaAsBi material system. A multi-level ionization model was developed to better understand the mid-band gap states seen in GaAsBi epitaxial layers in the literature and in the work itself.

A full micro-fabrication process was developed to produce actual HBTs using epitaxially grown bismide material. Many iterations of process design and testing were performed to produce a

working transistor using a six-step process. Specific fabrication challenges for this material were addressed, including finding a suitable etch solution and process technique to adapt to variable etch rates. A variety of test structures were used to test the conductivity, doping concentration, minority carrier transport, and thermal effects such as partial ionization.

Measurements of fabricated GaAs/GaAsBi devices were taken, such as I-V measurements of diodes and transistors, C-V measurements of P-N junctions, and TLM and thermal measurements of conductivity. Key parameters were extracted from these measurements such as doping concentration and conductivity, while other parameters were inferred from these such as mobility, mid-gap state activation energy, lifetime and transport measurements, and surface recombination effects. Overall this contributes to the characterization of the first HBT made using the new III-V bismide material system.

6.2 Conclusions and Discussion

This work has been the first investigation of GaAsBi as a novel material for HBT electronic devices. The simulation, processing, and characterization undergone in this work will be useful for future electronic and optoelectronic device applications utilizing dilute bismides. The material review and simulation have demonstrated the potential of this material for research and commercial device applications. The framework developed here will be useful for future projects using this material—whether for novel research devices or for commercial products.

The potential performance for HBT devices using the material was investigated in chapter 3. By using measured material parameters, and a simulation model taking into account the unique properties of GaAsBi, it was shown that high performance devices can be made using this material. Transistors using GaAs/GaAsBi heterojunction have several advantages that make them competitive among other III-V based devices. The biggest advantage is that bismide-based epitaxial layers open up the narrow band gap region to devices using a GaAs substrate. The small lattice mismatch and large band gap shift allows band gaps down to the telecom range of 0.8 eV (1.55 μm wavelength)

grown on a GaAs substrates. Compared to InGaAs grown on InP substrates, bismide devices can leverage the large wafer size and mature processing development of GaAs, as well as avoiding the need for potentially dangerous phosphorus-based process gases. The simulations in chapter 3 show that low-power GaAs/GaAsBi HBTs are capable of X-band operation at reasonable (0.5 μm) device dimensions.

The process development of a prototype GaAs/GaAsBi HBT has addressed several hurdles necessary to build a working device. The lack of wet etch selectivity between GaAs and GaAsBi necessitated the development of the partial etch technique outlined in chapter 4. This wet etch process is sufficient for the initial device dimensions, but smaller devices will require dry etch techniques. The standard III-V RIE recipes was shown to not work with GaAsBi layers, but this could be solved by using different gas mixtures.

The first prototype GaAs/GaAsBi P-N diodes and HBT were demonstrated in chapter 5. Good P-N junction diodes and GaAsBi layer mobility have been shown. Minority carrier transport has also been demonstrated in fabricated HBT devices. The minority carrier transport is small—even for the thick base prototype device. The reason for this poor transport is not fully understood, but is strongly suspected to by due to decreased carrier lifetime from the additional mid-gap states near the valance band in GaAsBi or from large surface or depletion region recombination currents. Direct experimental evidence of these mid-gap states have been found with energies of ~ 100 meV above the valance band, which is consistent with the values reported in the literature. This unique effect in GaAsBi thin films has been seen in samples grown in many MBE systems at multiple institutions. It is still not know if this is an intrinsic effect in bismide materials or if there are certain growth conditions that can minimize its impact. In addition the proportion of the base current due to surface recombination was found to be very large, especially for all but the highest current densities. Neither of these factors fully explains the low transport, but improving the intrinsic material lifetime as well as reducing the surface recombination would greatly improve fabricated bismide devices.

6.3 Future Work

There are many more experiments that can be done to further understand the GaAsBi material system. Some of these experiments require specific material structures, or specialized equipment to perform. Some of the properties that were previously inferred can be measured directly. Hall effect measurements can be used to get the device mobility directly, which could help resolve the ambiguity between doping and mobility seen in section 5.1.2. A thermal Hall effect could also get more information about the temperature dependence of the mobility—which would make the thermal conductivity extraction of the mid-gap state activation energy more accurate. Right now the mobility is assumed to be a weak function of temperature. This is a justified assumption since the scattering from the bismuth is the dominant mechanism, but especially at low alloy concentrations the additional data would be helpful.

Other useful material property measurements include secondary-ion mass spectrometry (SIMS) to find the exact concentration of both bismuth and dopant impurities. Comparing this to the measured free carrier concentration the effects of dopant compensation, partial ionization, and even unintentional doping from localized bismuth states could be further understood. Independent measurements of the minority carrier lifetime in GaAsBi could aid understanding of the minority carrier transport in a transistor. Photoluminescence (PL) can be used to find this lifetime, but might be difficult to measure for very short lifetimes. However, even knowing an upper-bound on the lifetime would be very useful.

A key feature of GaAsBi mentioned throughout the literature is the lack of a band discontinuity in the conduction band for dilute bismuth GaAs/GaAsBi heterojunctions. This effect has been assumed for the material because of its analogous properties to dilute nitride materials. At this time however no direct measurement of the discontinuity has been performed. Using a specific device structure of an iso-type heterojunction of GaAs/GaAsBi near either a P-N or Schottky junction could be used to measure this value. Using C-V measurements of the reverse biased junction to extend the depletion region over the iso-type hetero-interface the band discontinuity can be inferred

from established measurement techniques [91, 92, 93, 94]. This measurement could be especially effective if the doping was found from independent measurements of the impurity concentration from SIMS or inferred from conductivity and Hall effect measurements.

Now that a first concept device has been made in this material other device designs can be explored for specific applications. Simulations of GaAs/GaAsBi HBTs in this work indicate the potential for X-band RF amplifiers using this material. Once the unideal effects in the material are further understood and addressed it would just be a matter of scaling to achieve high-performance RF GaAs/GaAsBi transistors. From the simulations in chapter 3 the device dimensions necessary to reach the target frequency are reasonable, and could be done in a small lab using e-beam lithography. At these smaller dimensions it might be necessary to develop a dry RIE process for etching the transistor mesa.

Overall there is a lot a potential left to explore in devices made using GaAsBi. There is great potential in optoelectronic devices such as lasers. The reduced dependence of the material properties on temperature, such as band gap and mobility, makes optical sources a promising application. In addition, a band gap in the telecom range of 0.8 eV, or 1.5 μm can be produced on a GaAs substrate—potential reducing the material and manufacturing cost of telecom lasers. Fully integrated photonic circuits can also be made on a GaAs substrate—with large band gap materials like GaAs or AlGaAs used as waveguides and GaAsBi for sources and detectors. This work will serve as the first investigation of the device applications of GaAsBi—hopefully the first of many.

Bibliography

[1] X. Lu, D. A. Beaton, R. B. Lewis, T. Tiedje, and Y. Zhang, "Composition dependence of photoluminescence of $GaAs_{1-x}Bi_x$ alloys," Applied Physics Letters, vol. 95, no. 4, 2009.

[2] Y. Takehara, M. Yoshimoto, W. Huang, J. Saraie, K. Oe, A. Chayahara, and Y. Horino, "Lattice distortion of GaAsBi alloy grown on GaAs by molecular beam epitaxy," Japanese Journal of Applied Physics, vol. 45, pp. 67–69, Jan. 2006.

[3] M. Sotoodeh, A. Khalid, and A. Rezazadeh, "Empirical low-field mobility model for III-V compounds applicable in device simulation codes," Journal of Applied Physics, vol. 87, no. 6, pp. 2890–2900, 2000.

[4] V. Jain, E. Lobisser, A. Baraskar, B. J. Thibeault, M. J. Rodwell, Z. Griffith, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and W. Liu, "InGaAs/InP DHBTs in a dry-etched refractory metal emitter process demonstrating simultaneous $f_t/f_{max} \sim 430/800$ GHz," IEEE Electron Device Letters, vol. 32, no. 1, pp. 24–26, 2011.

[5] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," Journal of Applied Physics, vol. 89, no. 11, pp. 5815–5875, 2001.

[6] H. Achour, S. Louhibi, B. Amrani, A. Tebboune, and N. Sekkal, "Structural and electronic properties of GaAsBi," Superlattices and Microstructures, vol. 44, no. 2, pp. 223–229, 2008.

[7] D. Madouri, A. Boukra, A. Zaoui, and M. Ferhat, "Bismuth alloying in GaAs: a first-principles study," Computational Materials Science, vol. 43, pp. 818–822, Oct. 2008.

[8] D. A. Beaton, R. B. Lewis, M. Masnadi-Shirazi, and T. Tiedje, "Temperature dependence of hole mobility in $GaAs_{1-x}Bi_x$ alloys," Journal of Applied Physics, vol. 108, p. 083708, Oct. 2010.

[9] A. J. Ptak, R. France, D. A. Beaton, K. Alberi, J. Simon, A. Mascarenhas, and C. S. Jiang, "Kinetically limited growth of GaAsBi by molecular-beam epitaxy," Journal of Crystal Growth, vol. 338, no. 1, pp. 107–110, 2012.

[10] M. Yoshimoto, S. Murata, A. Chayahara, Y. Horino, J. Saraie, and K. Oe, "Metastable GaAsBi alloy grown by molecular beam epitaxy," Japanese Journal of Applied Physics, vol. 42, p. 1235, Oct. 2003.

[11] G. Pettinari, A. Polimeni, M. Capizzi, J. H. Blokland, P. C. M. Christianen, J. C. Maan, E. C. Young, and T. Tiedje, "Influence of bismuth incorporation on the valence and conduction band edges of $GaAs_{1-x}Bi_x$," Applied Physics Letters, vol. 92, no. 26, p. 262105, 2008.

[12] S. Francoeur, M.-J. Seong, A. Mascarenhas, S. Tixier, M. Adamcyk, and T. Tiedje, "Band gap of $GaAs_{1-x}Bi_x$, 0<x<3.6%," Applied Physics Letters, vol. 82, p. 3874, June 2003.

[13] X. Lu, D. A. Beaton, R. B. Lewis, T. Tiedje, and M. B. Whitwick, "Effect of molecular beam epitaxy growth conditions on the Bi content of $GaAs_{1-x}Bi_x$," Applied Physics Letters, vol. 92, no. 19, pp. 192110–192110, 2008.

[14] S. Nargelas, K. Jarašiūnas, K. Bertulis, and V. Pačebutas, "Hole diffusivity in GaAsBi alloys measured by a picosecond transient grating technique," Applied Physics Letters, vol. 98, no. 8, p. 082115, 2011.

[15] R. N. Kini, A. J. Ptak, B. Fluegel, R. France, R. C. Reedy, and A. Mascarenhas, "Effect of Bi alloying on the hole transport in the dilute bismide alloy GaAsBi," Physical Review B, vol. 83, no. 7, p. 075307, 2011.

[16] R. N. Kini, L. Bhusal, A. J. Ptak, R. France, and A. Mascarenhas, "Electron hall mobility in GaAsBi," Journal of Applied Physics, vol. 106, no. 4, pp. 043705–043705, 2009.

[17] K. Kado, T. Fuyuki, K. Yamada, K. Oe, and M. Yoshimoto, "High hole mobility in $GaAs_{1-x}Bi_x$ alloys," Japanese Journal of Applied Physics, vol. 51, p. 040204, Mar. 2012.

[18] K. Alberi, J. Wu, W. Walukiewicz, K. M. Yu, O. D. Dubon, S. P. Watkins, C. X. Wang, X. Liu, Y.-J. Cho, and J. Furdyna, "Valence-band anticrossing in mismatched III-V semiconductor alloys," Phys. Rev. B, vol. 75, p. 045203, Jan. 2007.

[19] W. Shockley, "Circuit element utilizing semiconductive material," Sept. 1951. US Patent 2,569,347.

[20] H. Kroemer, "Theory of a wide-gap emitter for transistors," Proceedings of the IRE, vol. 45, no. 11, pp. 1535–1537, 1957.

[21] H. Kroemer, "Heterostructure bipolar transistors and integrated circuits," Proceedings of the IEEE, vol. 70, no. 1, pp. 13–25, 1982.

[22] P. Asbeck, F. Chang, K. Wang, G. Sullivan, and D. Cheung, "GaAs-based heterojunction bipolar transistors for very high performance electronic circuits," Proceedings of the IEEE, vol. 81, no. 12, pp. 1909–1726, 1993.

[23] H. Kroemer, "Heterostructure bipolar transistors: What should we build?," Journal of Vacuum Science Technology B: Microelectronics and Nanometer Structures, vol. 1, pp. 126–130, Apr. 1983.

[24] F. Schwierz and J. J. Liou, "RF transistors: Recent developments and roadmap toward terahertz applications," Solid State Electronics, vol. 51, no. 8, pp. 1079–1091, 2007.

[25] B. Bayraktaroglu, "GaAs HBTs for microwave integrated circuits," in Proceedings of the IEEE, pp. 1762–1785, 1993.

[26] H. Nelson, "Epitaxial growth from the liquid state and its application to the fabrication of tunnel and laser diodes," RCA Review, vol. 24, no. 4, pp. 603–615, 1963.

[27] W. Dumke, J. Woodall, and V. Rideout, "GaAs-GaAlAs heterojunction transistor for high frequency operation," Solid-State Electronics, vol. 15, pp. 1339–1343, Dec. 1972.

[28] M. Ettenberg and R. J. Paff, "Thermal expansion of AlAs," Journal of Applied Physics, vol. 41, no. 10, pp. 3926–3927, 1970.

[29] M.-C. Tu, H.-Y. Ueng, and Y.-C. Wang, "Performance of high-reliability and high-linearity InGaP/GaAs HBT PAs for wireless communication," IEEE Transactions on Electron Devices, vol. 57, pp. 188–194, Jan. 2010.

[30] M. Urteaga, M. Seo, J. Hacker, Z. Griffith, A. Young, R. Pierson, P. Rowell, A. Skalare, V. Jain, E. Lobisser, and M. Rodwell, "InP HBTs for THz frequency integrated circuits," in 2011 and 23rd International Conference on Indium Phosphide and Related Materials, pp. 1–4, May 2011.

[31] W. Hoke, P. Lemonias, J. Mosca, P. Lyman, A. Torabi, P. Marsh, R. McTaggart, S. Lardizabal, and K. Hetzler, "Molecular beam epitaxial growth and device performance of metamorphic high electron mobility transistor structures fabricated on GaAs substrates," Journal of Vacuum Science & Technology B, vol. 17, no. 3, pp. 1131–1135, 1999.

[32] D. V. Lang, M. B. Panish, F. Capasso, J. Allam, R. A. Hamm, A. M. Sergent, and W. T. Tsang, "Measurement of heterojunction band offsets by admittance spectroscopy: InP/$Ga_{0.47}In_{0.53}As$," Applied Physics Letters, vol. 50, no. 12, p. 736, 1987.

[33] M. Yamaguchi, T. Takamoto, K. Araki, and N. Ekins-Daukes, "Multi-junction III-V solar cells: current status and future potential," Solar Energy, vol. 79, no. 1, pp. 78 – 85, 2005.

[34] C. H. Henry, "Limiting efficiencies of ideal single and multiple energy gap terrestrial solar cells," Journal of Applied Physics, vol. 51, no. 8, pp. 4494–4500, 1980.

[35] R. King, D. Law, K. Edmondson, C. Fetzer, G. Kinsey, H. Yoon, R. Sherif, and N. Karam, "40% efficient metamorphic GaInP/GaInAs/Ge multijunction solar cells," Applied Physics Letters, vol. 90, no. 18, pp. 183516–183516, 2007.

[36] W. Guter, J. Schone, S. P. Philipps, M. Steiner, G. Siefer, A. Wekkeli, E. Welser, E. Oliva, A. W. Bett, and F. Dimroth, "Current-matched triple-junction solar cell reaching 41.1% conversion efficiency under concentrated sunlight," Applied Physics Letters, vol. 94, no. 22, pp. –, 2009.

[37] S. Kurtz, D. Myers, and J. Olson, "Projected performance of three- and four-junction devices using GaAs and GaInP," in 2007 Photovoltaic Specialists Conference, pp. 875–878, Sept. 1997.

[38] D. Friedman, J. Geisz, S. Kurtz, and J. Olson, "1-eV solar cells with GaInNAs active layer," Journal of Crystal Growth, vol. 195, no. 1-4, pp. 409 – 415, 1998.

[39] K. Alberi, O. Dubon, W. Walukiewicz, K. Yu, K. Bertulis, and A. Krotkus, "Valence band anticrossing in $GaBi_xAs_{1-x}$," Applied Physics Letters, vol. 91, pp. 051909–051909–3, July 2007.

[40] W. Walukiewicz, K. Alberi, J. Wu, W. Shan, K. Yu, and I. Ager, J.W., "Electronic band structure of highly mismatched semiconductor alloys," in Dilute III-V Nitride Semiconductors and Material Systems (A. Erol, ed.), vol. 105 of Materials Science, pp. 65–89, Springer Berlin Heidelberg, 2008.

[41] T. Kitatani, M. Kondow, T. Kikawa, Y. Yazawa, M. Okai, and K. Uomi, "Analysis of band offset in GaNAs/GaAs by X-Ray photoelectron spectroscopy," Japanese Journal of Applied Physics, vol. 38, p. 5003, Sept. 1999.

[42] A. Hashimoto, T. Yamaguchi, T. Suzuki, and A. Yamamoto, "Hall electron mobility versus N spatial distribution in III-V-N systems," Journal of Crystal Growth, vol. 278, pp. 532–537, May 2005.

[43] Z. Marks, I. Haygood, and B. Van Zeghbroeck, "Analysis of the GaAs/GaAsBi material system for heterojunction bipolar transistors," IEEE Transactions on Electron Devices, vol. 60, pp. 200–205, Jan. 2013.

[44] S. J. Sweeney and S. R. Jin, "Bismide-nitride alloys: Promising for efficient light emitting devices in the near- and mid-infrared," Journal of Applied Physics, vol. 113, no. 4, 2013.

[45] J. Yoshida, T. Kita, O. Wada, and K. Oe, "Temperature dependence of $GaAs_{1-x}Bi_x$ band gap studied by photoreflectance spectroscopy," Japanese Journal of Applied Physics, vol. 42, no. 2R, p. 371, 2003.

[46] K. Yamashita, M. Yoshimoto, and K. Oe, "Temperature-insensitive refractive index of GaAsBi alloy for laser diode in WDM optical communication," physica status solidi (c), vol. 3, no. 3, pp. 693–696, 2006.

[47] S. Sweeney, I. Marko, S. Jin, K. Hild, Z. Batool, N. Hossain, and T. Hosea, "Bismide-based photonic devices for near- and mid-infrared applications," in Bismuth-Containing Compounds (H. Li and Z. M. Wang, eds.), vol. 186 of Springer Series in Materials Science, pp. 29–53, Springer New York, 2013.

[48] J. Matthews and A. Blakeslee, "Defects in epitaxial multilayers: I. misfit dislocations," Journal of Crystal Growth, vol. 27, no. 0, pp. 118 – 125, 1974.

[49] R. People and J. Bean, "Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures," Applied Physics Letters, vol. 47, no. 3, pp. 322–324, 1985.

[50] R. France, C. S. Jiang, and A. J. Ptak, "In situ strain relaxation comparison between GaAsBi and GaInAs grown by molecular-beam epitaxy," Applied Physics Letters, vol. 98, no. 10, p. 101908, 2011.

[51] M. Yoshimoto, M. Itoh, Y. Tominaga, and K. Oe, "Quantitative estimation of density of Bi-induced localized states in $GaAs_{1-x}Bi_x$ grown by molecular beam epitaxy," Journal of Crystal Growth, vol. 378, pp. 73–76, Sept. 2013.

[52] M. Yoshimoto and T. Fuyuki, "Localized states in GaAsBi and GaAs/GaAsBi heterostructures," in Bismuth-Containing Compounds (H. Li and Z. M. Wang, eds.), vol. 186 of Springer Series in Materials Science, pp. 201–224, Springer New York, 2013.

[53] S. Fahy, A. Lindsay, and E. O'Reilly, "Intrinsic limits on electron mobility in disordered dilute nitride semiconductor alloys," IEE Proceedings—Optoelectronics, vol. 151, pp. 352–356, Oct. 2004.

[54] A. Janotti, S.-H. Wei, and S. B. Zhang, "Theoretical study of the effects of isovalent coalloying of bi and n in gaas," Phys. Rev. B, vol. 65, p. 115203, Feb. 2002.

[55] M. Shur, Physics of Semiconductor Devices. Prentice-Hall series in solid state physical electronics, Prentice Hall, 1990.

[56] S. M. Sze and K. N. Kwok, Physics of Semiconductor Devices. Hoboken, NJ: Wiley-Interscience, third ed., 2007.

[57] J. Bardeen and W. H. Brattain, "The transistor, a semi-conductor triode," Physical Review, vol. 74, no. 2, pp. 230–231, 1948.

[58] J. S. Blakemore, "Semiconducting and other major properties of gallium arsenide," Journal of Applied Physics, vol. 53, no. 10, pp. R123–R181, 1982.

[59] R. B. Dingle, "The fermi-dirac integrals $\mathcal{F}_p(\eta) = (p!)^{-1} \int_0^\infty \varepsilon^p (e^{\varepsilon-\eta}+1)^{-1} d\varepsilon$," Applied Scientific Research, Section B, vol. 6, pp. 225–239, Dec. 1957.

[60] R. Kim and M. Lundstrom, "Notes on fermi-dirac integrals," ArXiv e-prints, Nov. 2008.

[61] A. Baca, F. Ren, J. Zolper, R. Briggs, and S. Pearton, "A survey of ohmic contacts to III-V compound semiconductors," Thin Solid Films, vol. 308, pp. 599–606, 1997.

[62] B. Van Zeghbroeck, "Principles of semiconductor devices." `http://ecee.colorado.edu/~bart/book/`, 2004. [Online; accessed May-2014].

[63] R. Muller, T. Kamins, and M. Chan, Device Electronics for Integrated Circuits. John Wiley & Sons, 2003.

[64] V. Fournier, J. Dangla, and C. Dubon-Chevallier, "Investigation of emitter current crowding effect in heterojunction bipolar transistors," Electronics Letters, vol. 29, pp. 1799–1800, Sept. 1993.

[65] G. Pettinari, A. Patanè, A. Polimeni, M. Capizzi, X. Lu, and T. Tiedje, "Bi-induced p-type conductivity in nominally undoped Ga(AsBi)," Applied Physics Letters, vol. 100, no. 9, p. 092109, 2012.

[66] D. Clugston and P. Basore, "PC1D version 5: 32-bit solar cell modeling on personal computers," in Photovoltaic Specialists Conference, 1997., Conference Record of the Twenty-Sixth IEEE, pp. 207–210, Sep 1997.

[67] G. S. Marlow and M. B. Das, "The effects of contact size and non-zero metal resistance on the determination of specific contact resistance," Solid State Electronics, vol. 25, pp. 91–94, Feb. 1982.

[68] G. K. Reeves and H. B. Harrison, "Obtaining the specific contact resistance from transmission line model measurements," IEEE Electron Device Letters, vol. 3, no. 5, pp. 111–113, 1982.

[69] E. F. Chor and J. Lerdworatawee, "Quasi-two-dimensional transmission line model (QTD-TLM) for planar ohmic contact studies," IEEE Transactions on Electron Devices, vol. 49, no. 1, pp. 105–111, 2002.

[70] P. E. Dodd, T. B. Stellwag, M. R. Melloch, and M. S. Lundstrom, "Surface and perimeter recombination in GaAs diodes: an experimental and theoretical investigation," IEEE Transactions on Electron Devices, vol. 38, no. 6, pp. 1253–1261, 1991.

[71] S. Campbell, The Science and Engineering of Microelectronic Fabrication. Oxford series in electrical and computer engineering, Oxford University Press, second ed., 2001.

[72] J. Kim and S. Blairs, "Sublimation study of $BiBr_3$," Journal of Materials Science, vol. 30, no. 10, pp. 2559–2565, 1995.

[73] R. Williams, Modern GaAs processing methods. Artech House, 1990.

[74] G. DeSalvo, R. Kaspi, and C. Bozada, "Citric-acid etching of $GaAs_{1-x}Sb_x$, $Al_{0.5}Ga_{0.5}Sb$, and InAs for heterostructure device fabrication," Journal of The Electrochemical Society, vol. 141, no. 12, pp. 3526–3531, 1994.

[75] V. L. Rideout, "A review of the theory and technology for ohmic contacts to group III-V compound semiconductors," Solid State Electronics, vol. 18, pp. 541–550, June 1975.

[76] R. H. Cox and H. Strack, "Ohmic contacts for GaAs devices," Solid State Electronics, vol. 10, pp. 1213–1218, Dec. 1967.

[77] M. N. Yoder, "Ohmic contacts in GaAs," Solid State Electronics, vol. 23, no. 2, pp. 117–119, 1980.

[78] N. Braslau, "Ohmic contacts to GaAs," Thin Solid Films, vol. 104, pp. 291–297, June 1983.

[79] R. Elliott and F. Shunk, "The Au-Ge system (gold-germanium)," Bulletin of Alloy Phase Diagrams, vol. 1, no. 2, pp. 51–54, 1980.

[80] T. Sanada and O. Wada, "Ohmic contacts to p-GaAs with Au/Zn/Au structure," Japanese Journal of Applied Physics, vol. 19, pp. L491–L494, Aug. 1980.

[81] H. Okamoto and T. Massalski, "The Au-Zn (gold-zinc) system," Bulletin of Alloy Phase Diagrams, vol. 10, no. 1, pp. 59–69, 1989.

[82] E. Chang, G. T. Cibuzar, and K. P. Pande, "Passivation of GaAs FETs with PECVD silicon nitride films of different stress states," IEEE Transactions on Electron Devices, vol. 35, no. 9, pp. 1412–1418, 1988.

[83] Y. Hirota, K. Sugii, and Y. Homma, "Cleaning effects of running deionized water on a GaAs surface," Journal of The Electrochemical Society, vol. 138, no. 3, pp. 799–802, 1991.

[84] J. A. Taylor, "The mechanical properties and microstructure of plasma enhanced chemical vapor deposited silicon nitride thin films," Journal of Vacuum Science & Technology A, vol. 9, no. 4, pp. 2464–2468, 1991.

[85] H. Huang, K. Winchester, A. Suvorova, B. Lawn, Y. Liu, X. Hu, J. Dell, and L. Faraone, "Effect of deposition conditions on mechanical properties of low-temperature PECVD silicon nitride films," Materials Science and Engineering: A, vol. 435-436, pp. 453–459, 2006.

[86] A. Piotrowska, A. Guivarc'h, and G. Pelous, "Ohmic contacts to III-V compound semiconductors: A review of fabrication techniques," Solid State Electronics, vol. 26, no. 3, pp. 179–197, 1983.

[87] Y. Lu, T. S. Kalkur, and C. P. de Araujo, "Rapid thermal alloyed ohmic contacts to p-type GaAs," Journal of The Electrochemical Society, vol. 136, no. 10, pp. 3123–3129, 1989.

[88] T. C. Shen, G. B. Gao, and H. Morko, "Recent developments in ohmic contacts for III-V compound semiconductors," Journal of Vacuum Science & Technology B, vol. 10, no. 5, pp. 2113–2132, 1992.

[89] S. Gupta, M. Y. Frankel, J. A. Valdmanis, J. F. Whitaker, G. A. Mourou, F. W. Smith, and A. R. Calawa, "Subpicosecond carrier lifetime in GaAs grown by molecular beam epitaxy at low temperatures," Applied Physics Letters, vol. 59, no. 25, pp. 3276–3278, 1991.

[90] Y. F. Yang, C. C. Hsu, and E. S. Yang, "Surface recombination current in InGaP/GaAs heterostructure-emitter bipolar transistors," IEEE Transactions on Electron Devices, vol. 41, pp. 643–647, May 1994.

[91] H. Kroemer, W.-Y. Chien, J. S. Harris, and D. D. Edwall, "Measurement of isotype heterojunction barriers by C-V profiling," Applied Physics Letters, vol. 36, no. 4, p. 295, 1980.

[92] M. O. Watanabe, J. Yoshida, M. Mashita, T. Nakanisi, and A. Hojo, "Band discontinuity for GaAs/AlGaAs heterojunction determined by C-V profiling technique," Journal of Applied Physics, vol. 57, no. 12, pp. 5340–5344, 1985.

[93] D. Arnold, A. Ketterson, T. Henderson, J. Klem, and H. Morkoc, "Determination of the valence-band discontinuity between GaAs and (Al,Ga)As by the use of p^+-GaAs-(Al,Ga)As-p^--GaAs capacitors," Applied Physics Letters, vol. 45, pp. 1237–1239, Dec. 1984.

[94] H. Okumura, S. Misawa, and S. Yoshida, "Reliability of the band discontinuity determination by capacitance-voltage method: Relation of the interface charge density and the trap concentration near the interface," Surface Science, vol. 174, pp. 324–330, Aug. 1986.

Appendix A

Fabrication and processing

A.1 Processing procedure steps

The basic steps and processes used in the device fabrication are as follows:

Spinner clean

Place sample on pneumatic spinner. Wet sample with acetone before spinning. Spray acetone for $\sim$ 10 s, followed by $\sim$ 10 s spray of isopropyl alcohol. Dry while still spinning with N_2 gun.

Positive resist

Pre-dry sample on a $>$ 120 °C vacuum hotplate for 2 min. Deposit Clariant AZ-P4210 positive resist with pneumatic spinner at 6000 RPM for 30 s. Pre-bake sample and resist on a 100 °C vacuum hotplate for 90 s. After exposure, immersion develop with agitation in 1:4 AZ400K:H_2O_2 for $\sim$ 60 s.

Negative resist

No pre-dry required. Deposit Futurrex NR71-1500PY (lift-off) or NR9-1000P (etching) negative resist with pneumatic spinner at 4000 RPM for 40 s. Pre-bake sample and resist on a 150 °C vacuum hotplate for 60 s. After exposure, post-bake sample on a 100 °C vacuum

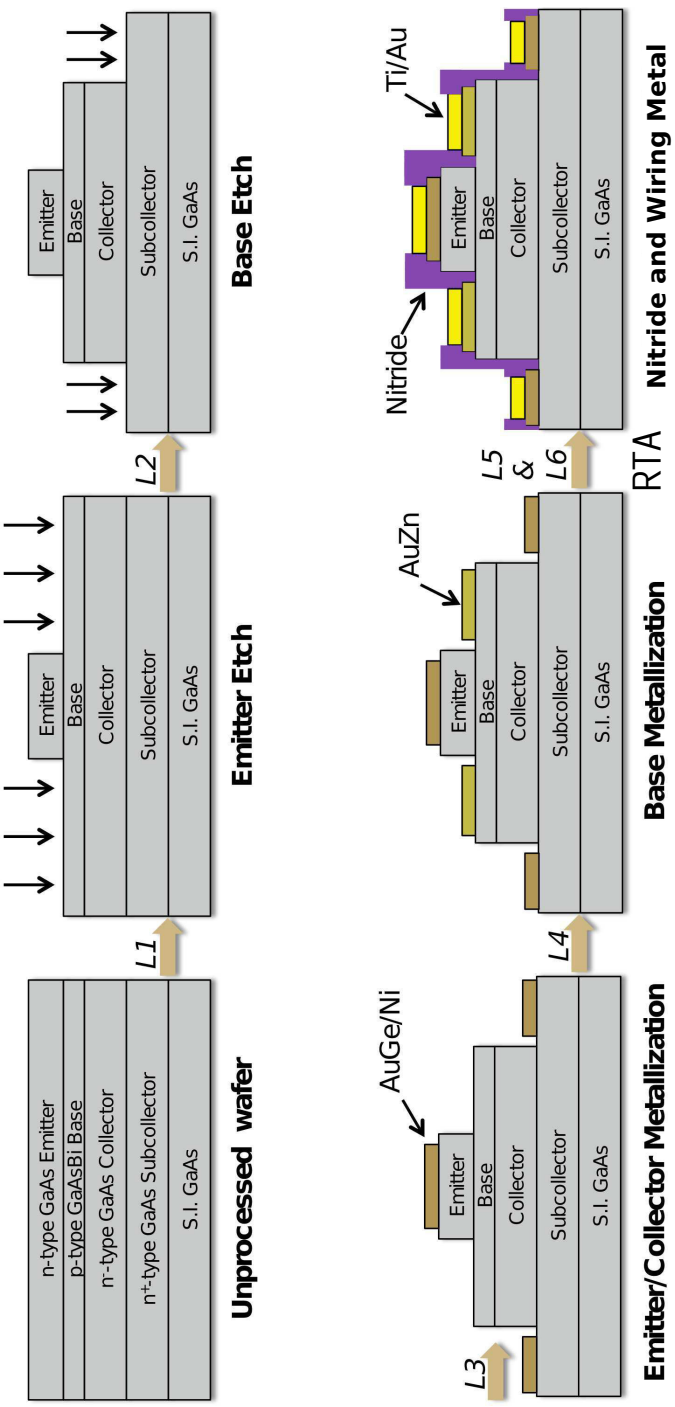


Figure A.1: Six-step process cartoon for N-P-N GaAs/GaAsBi HBT

hotplate for 60 s. Pneumatic spinner development with spray of 10 s DI water, 10 s RD6 developer, and 10 s DI water rinse spray.

Etch Process

See section 4.2 for details about etchant mixture. Before etch clean the semiconductor surface with a 2 min soak in 1:1 HCl:H_2O solution. Transfer sample from cleaning solution directly into etchant (5:1 $C_6H_8O_7$:H_2O_2) for desired etch time. Soak in DI water rinse tank, starting from dirtiest and moving to cleanest tank.

Table A.1: HBT fabrication process procedure

Step	Process	Instructions
L0	Sample Prep	**Spinner clean** and 2 min soak in 1:1 HCl:H_2O to remove surface oxide
L1a	Emitter Etch Lithography	**Positive resist** process. Expose with L1 mask on MJB3/MJB4 for 8 s to 12 s. Possible descum in O_2 RIE for 30 s
L1b	Emitter Etch	**Etch process** to etch through emitter cap and emitter layers to expose base. Use partial etch process from section 4.2.
L2a	Base Etch Lithography	**Positive resist** process. Expose with L2 mask on MJB3/MJB4 for 8–12 s. Possible descum in O_2 RIE for 30 s
L2b	Base Etch	**Etch process** to etch through base and collector layers to expose sub-collector.

Step	Process	Instructions
L3a	Emitter/Collector Metalization Lithography	**Negative resist** (lift-off) process. Expose with L3 mask on MJB3/MJB4 for 80 120s. Possible descum in O_2 RIE for 30 s
L3b	Emitter/Collector Metal Evaporation	Remove oxide with 1 min soak in 1:1 HCl:H_2O solution. Evaporate 80–100 nm of AuGe (use entire charge), and ∼ 20 nm of Ni.
L3c	Emitter/Collector Metal lift-off	Soak in acetone for > 1 min, sonicate for 30 s then **spinner clean**.
L4a	Base Metalization Lithography	**Negative resist** (lift-off) process. Expose with L4 mask on MJB3/MJB4 for 80 120s. Possible descum in O_2 RIE for 30 s
L4b	Base Metal Evaporation	Remove oxide with 1 min soak in 1:1 HCl:H_2O solution. Evaporate 80–100 nm of AuZn (use entire charge).
L4c	Base Metal lift-off	Soak in acetone for > 1 min, sonicate for 30 s then **spinner clean**.
L4d	Rapid Thermal Annealing	See section 4.3. RTA under Ar atmosphere for 30 s at 350 °C followed immediately by 15–30 s at 415–450 °C
L5a	Nitride Deposition	**Spinner clean** and remove oxide with 1 min soak in 1:1 HCl:H_2O solution. Deposit 200–400 nm of silicon nitride in PECVD system (`LDN50.PRC`). Approximately 30 min for a 225 nm SiN_x thin-film.

Step	Process	Instructions
L5b	Nitride Etch Lithography	**Negative resist** (lift-off) process. Expose with L5 mask on MJB3/MJB4 for 8 12s. Possible descum in O_2 RIE for 30 s
L5c	Nitride Etch	Remove nitride using buffered oxide etchant (BOE), rate $\sim 180\,\mathrm{nm/min}$)
L6a	Wiring Metalization Lithography	**Negative resist** (lift-off) process. Expose with L6 mask on MJB3/MJB4 for 80 120s. Possible descum in O_2 RIE for 30 s
L6b	Wiring Metal Evaporation	Remove oxide with 1 min soak in 1:1 HCl:H_2O solution. Evaporate 20 nm Ti adhesion layer, followed by 100–150 nm of Au.
L6c	Wiring Metal lift-off	Soak in acetone for $>$ 1 min, sonicate for 30 s then **spinner clean**.

A.2 Mask layout

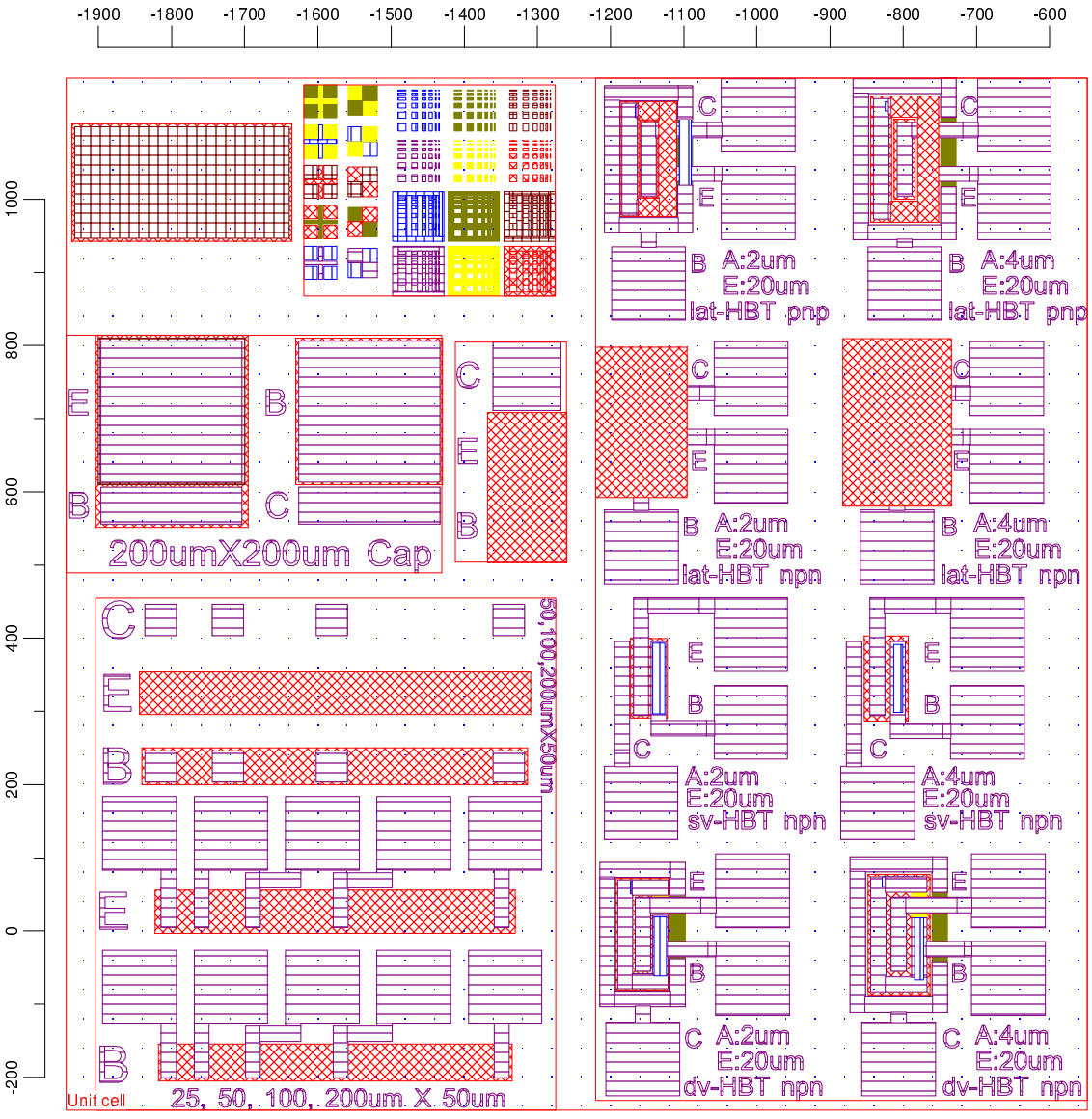


Figure A.2: Six-step HBT structure mask design (Generation 4)

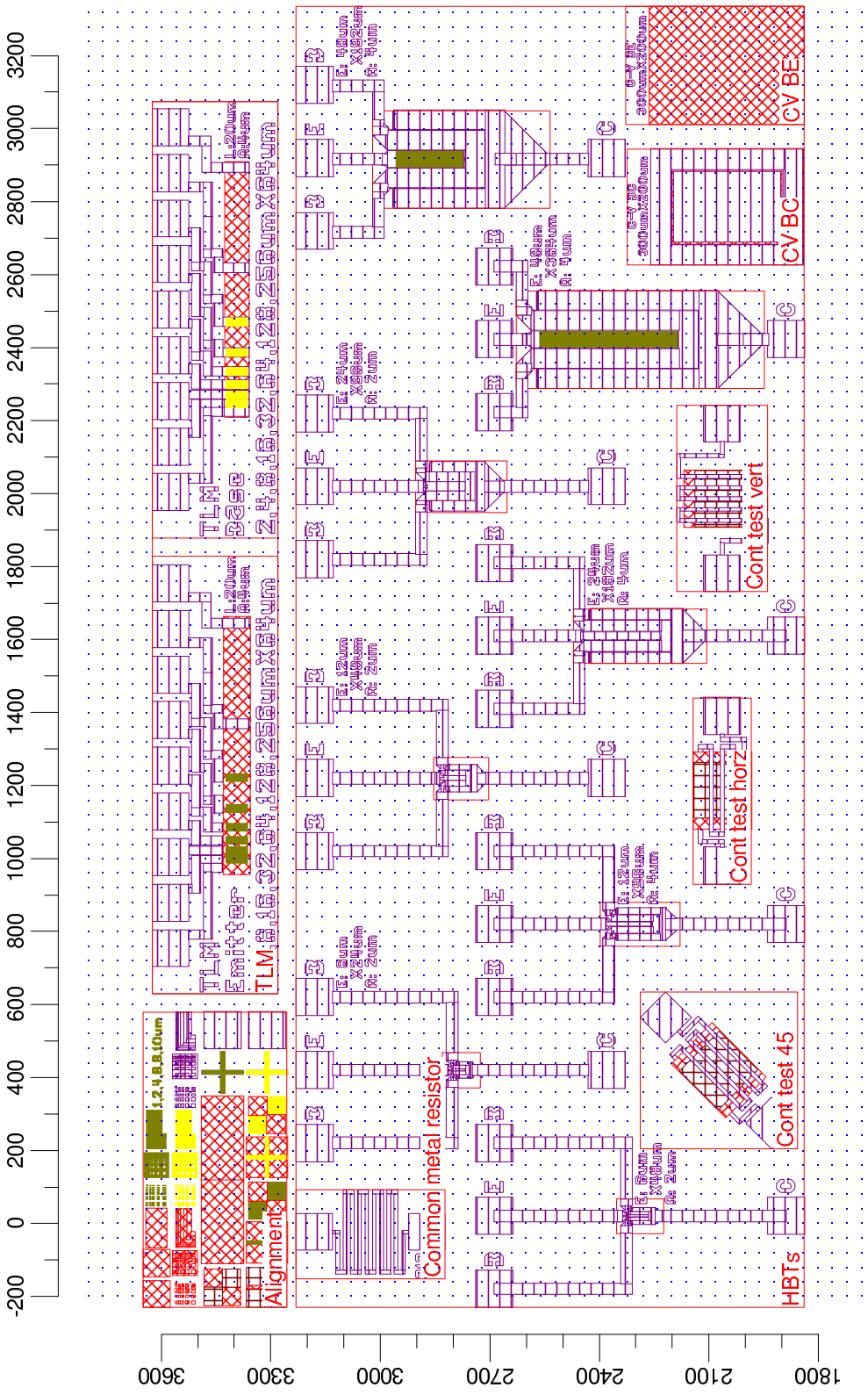


Figure A.3: Six-step HBT structure mask design (Generation 5)

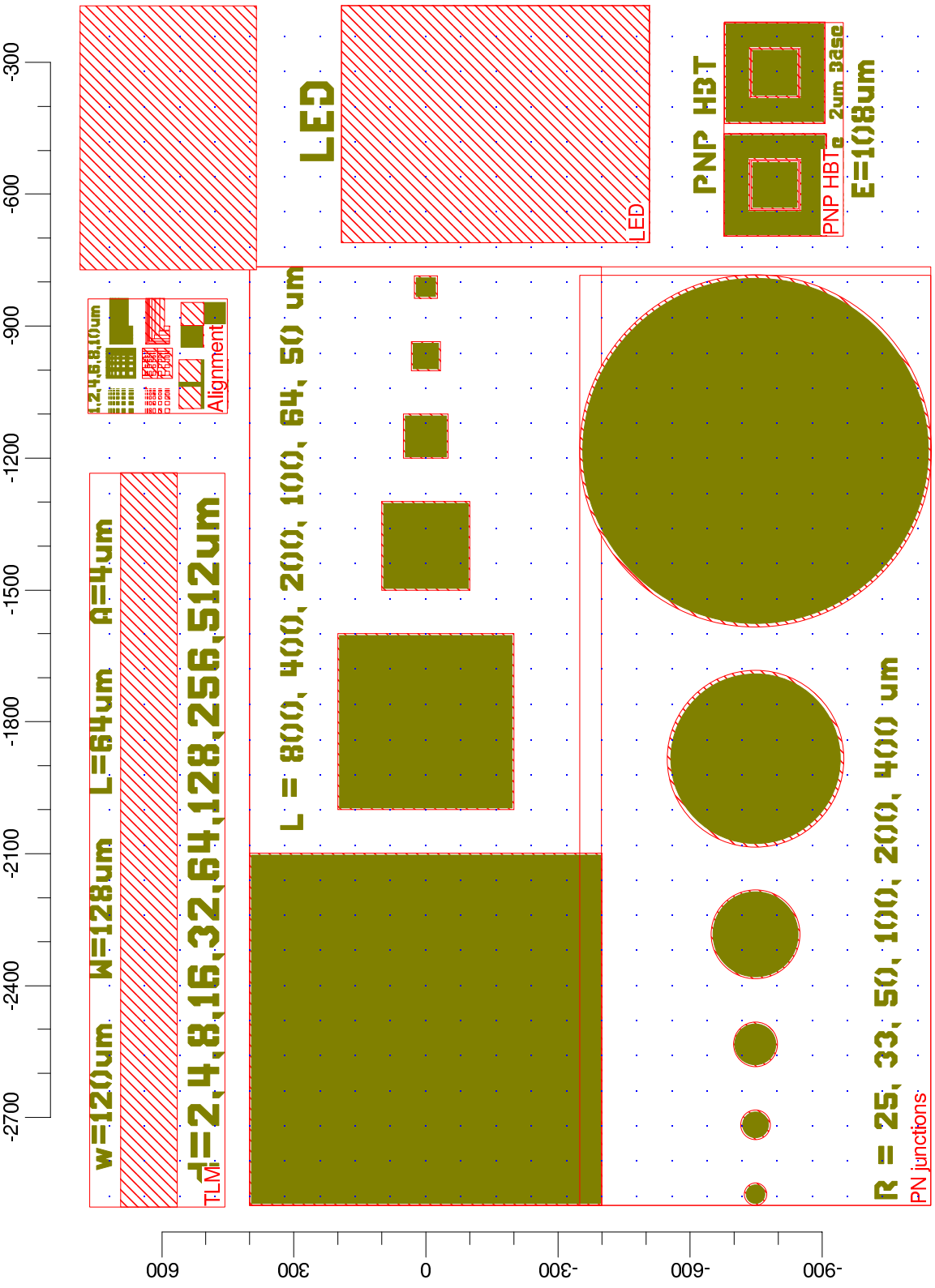


Figure A.4: Two-step P-N structure mask design

Appendix B

Simulation and analysis code documentation

B.1 HBT Simulation Code

The functions here are the core algorithms for both the HBT simulated from chapter 3 and the multi-level ionization code. Additional wrapper functions are used to compute and plot the results in this thesis. Full code is avaliable upon request at zefram.marks@gmail.com or online at www.zeframmarks.com.

B.1.1 computeHBT.m

```
function [ HBT device materials ] = computehbt( materials, device )
%COMPUTEHBT This primary function will compute all
%           the HBT device values
%from the input structure.
%   [ HBT_solution ] = computehbt( materials, device )
%
%   materials   ==> Structure containing material data
%                   for each layer.
%   Values must be named E, B, and C. i.e. materials.E = GaAs.
%
%   device      ==>  Structure containing device parameters

% This work is licensed under the
% Creative Commons Attribution-ShareAlike
% 3.0 Unported License. To view a copy of this license, visit
```

```
% http://creativecommons.org/licenses/by-sa/3.0/ or send a letter to
% Creative Commons, 444 Castro Street, Suite 900, Mountain View,
% California, 94041, USA.

%% Physical Constants
kb = 1.38065e-23; % Boltzmann's contant, in J/K.
q = 1.602177e-19; % Electron charge, in Coulombs.
Vt = kb*device.T/q; % Threshold voltage, in volts.

%% Initialize material properties

materials = initializematerials(materials, device);

% Areas
device.Abc = device.L*(device.Wee + ...
    2*(device.Seb+device.Sbs+device.Wbb));
device.Abe = device.L*device.Wee;

%% General parameters

% Compute built-in voltage for heterojunction.
% From Bart's book, (4.3.56).
HBT.ViBE = (materials.dEc - materials.dEv)/2 + ...
            Vt*log((device.Nb+device.Nb_immobile) * device.Ne / ...
                  ( materials.E.ni * materials.B.ni) ) + Vt/2 * ...
               log( materials.E.Nv * materials.B.Nc / ...
                  ( materials.E.Nc * materials.B.Nv) );

HBT.ViBC = (materials.dEc - materials.dEv)/2 + ...
            Vt*log((device.Nb+device.Nb_immobile) * device.Nc / ...
                  ( materials.C.ni * materials.B.ni) ) + Vt/2 * ...
               log( materials.C.Nv * materials.B.Nc / ...
                  ( materials.C.Nc * materials.B.Nv) );

% Generate (internal) voltages to sweep.
if device.symmetricV
    [VBE VBC] = meshgrid( ...
    linspace(-device.ViBE_factor*HBT.ViBE, ...
              device.ViBE_factor*HBT.ViBE, device.Npt), ...
    linspace( device.ViBC_factor*HBT.ViBC, ...
             -device.ViBC_factor*HBT.ViBC, device.Npt) );
    VCE = VBE - VBC;
else
    [VBE VBC] = meshgrid( ...
    linspace(0,  device.ViBE_factor*HBT.ViBE, device.Npt), ...
    linspace(0, -device.ViBC_factor*HBT.ViBC, device.Npt) );
    VCE = VBE - VBC;
```

```
63 end
64
65 % Change signs if device is pnp.
66 if strcmpi(device.type, 'pnp')
67 VBE = -1*VBE;
68 VBC = -1*VBC;
69 VCE = VBE - VBC;
70 end
71
72 % Compute high injection voltage drop for each junction
73 HBT.hi.hiVe = Vt*log( ( sqrt( 1 + ( 4*materials.E.ni^2 * ...
74     exp(VBE/Vt)/device.Ne^2 ) ) - 1 ) / 2 + 1 );
75 HBT.hi.hiVbe = Vt*log( ( sqrt( 1 + ( 4*materials.B.ni^2 * ...
76     exp(VBE/Vt)/(device.Nb+device.Nb_immobile)^2 ) ) ...
77     - 1 ) / 2 + 1 );
78 HBT.hi.hiVbc = Vt*log( ( sqrt( 1 + ( 4*materials.B.ni^2 * ...
79     exp(VBC/Vt)/(device.Nb+device.Nb_immobile)^2 ) ) ...
80     - 1 ) / 2 + 1 );
81 HBT.hi.hiVc = Vt*log( ( sqrt( 1 + ( 4*materials.C.ni^2 * ...
82     exp(VBC/Vt)/device.Nc^2 ) ) - 1 ) / 2 + 1 );
83
84 % Find depletion region widths for each voltage.
85 HBT.xBE = real( sqrt( 2*materials.B.eps*materials.E.eps/q * ...
86     ((device.Nb+device.Nb_immobile) + device.Ne).^2 / ...
87    (((device.Nb+device.Nb_immobile)*device.Ne) * ...
88     ((device.Nb+device.Nb_immobile)*materials.B.eps + ...
89       device.Ne*materials.E.eps) ) .* ...
90    (HBT.ViBE - VBE + HBT.hi.hiVe + HBT.hi.hiVbe) ) );
91
92 HBT.xBC = real( sqrt( 2*materials.B.eps*materials.C.eps/q * ...
93     ((device.Nb+device.Nb_immobile) + device.Nc).^2 / ...
94    (((device.Nb+device.Nb_immobile)*device.Nc) * ...
95     ((device.Nb+device.Nb_immobile)*materials.B.eps + ...
96       device.Nc*materials.C.eps) ) .*...
97       (HBT.ViBC - VBC + HBT.hi.hiVc + HBT.hi.hiVbc) ) );
98
99 % Find effective quasi-neutral region width.
100 device.Wb_eff = device.Wb - ( HBT.xBE*device.Ne / ...
101     ((device.Nb+device.Nb_immobile)+device.Ne) + ...
102     HBT.xBC*device.Nc/((device.Nb+device.Nb_immobile)+device.Nc) );
103 device.Wb_eff_lat = device.Wb - ( HBT.xBC*device.Nc / ...
104     ((device.Nb+device.Nb_immobile)+device.Nc) + ...
105         device.Wb_overetch );
106 device.We_eff = device.We - HBT.xBE * ...
107     (device.Nb+device.Nb_immobile) / ...
108     ((device.Nb+device.Nb_immobile)+device.Ne);
109
```

```
% Short circuit if Xbc > Wc, Wc_eff = 0 HBT.xBC(
% device.Nb/(device.Nb+device.Nc)*HBT.xBC ≥ device.Wc ) =
% (device.Nb+device.Nc)/device.Nb*device.Wc;
device.Wc_eff = device.Wc - ...
    HBT.xBC*(device.Nb+device.Nb_immobile) / ...
    ((device.Nb+device.Nb_immobile)+device.Nc);
device.Wc_eff( device.Wc_eff < 0 ) = 0;
device.Wsc_eff = device.Wsc; % Ignore nn+ depletion region

% Find epi layer resistivity.
materials.E.n*materials.E.mu_n + ...
materials.E.p*materials.E.mu_p) ).^(-1); R.pB = ( q * (
materials.B.n*materials.B.mu_n + ...
materials.B.p*materials.B.mu_p) ).^(-1); R.pC = ( q * (
materials.C.n*materials.C.mu_n + ...
materials.C.p*materials.C.mu_p) ).^(-1); R.pSC = ( q * (
materials.SC.n*materials.SC.mu_n + ...
materials.SC.p*materials.SC.mu_p) ).^(-1);

% Check if physically valid
if device.hbtcheckflag
[ VBE VBC VCE ] = hbtcheck( VBE, VBC, VCE );
end

%% Execute simulation function

% Execute I-V function
[ IV, R, device ] = hbtiv(materials, device, R);

% % Excute RF function
[ RF ] = hbtrf(materials, device, R);

%% Output solutions

% Output last good values
IV.max_vbc = find( isnan( IV.Vbc ) == 0, 1, 'last');
IV.max_vbe = find( isnan( IV.Vbe ) == 0, 1, 'last');

HBT.IV = IV;
HBT.R = R;

HBT.RF = RF;

%% Function definitions

% Device RF
function [ RF ] = hbtrf(materials, device, R)
```

```
157
158 % Compute junction capacitances
159 C.CjBE = materials.E.eps*materials.B.eps ./ HBT.xBE * ...
160     ((device.Nb+device.Nb_immobile) + device.Nb) * ...
161     1/(materials.E.eps*device.Ne + materials.B.eps * ...
162         (device.Nb+device.Nb_immobile));
163 C.CjBC = materials.C.eps*materials.B.eps ./ HBT.xBC * ...
164     (device.Nc + (device.Nb+device.Nb_immobile)) * ...
165     1/(materials.C.eps*device.Nc + materials.B.eps * ...
166         (device.Nb+device.Nb_immobile));
167
168 HBT.C = C;
169
170 % Compute each time constant
171 RF.tauE = Vt./IV.JE.*(C.CjBE + C.CjBC); % Emitter charging time.
172 RF.tauB = device.Wb_eff.^2/(2*materials.B.D); % Base transit time.
173 RF.tauSC = HBT.xBC/2.*1./(device.Nc + ...
174     (device.Nb+device.Nb_immobile)) .* ...
175     ((device.Nb+device.Nb_immobile)/materials.C.vsat + ...
176     device.Nc/materials.B.vsat); % Base-collector transit time.
177 RF.tauC = (R.Re_total + R.Rc_total).*C.CjBC.*device.Abc_eff;
178
179 RF.tau_total = RF.tauE + RF.tauB + RF.tauSC + RF.tauC;
180
181 % Find ft and fmax
182 RF.ft = ( 2*pi * RF.tau_total ).^(-1);
183 RF.fmax = sqrt( RF.ft ./ ...
184     (8*pi*R.Rb_total.*C.CjBC.*device.Abc_eff) );
185
186 end
187
188 % Device I-V
189 function [ IV, R, device ] = hbtiv(materials, device, R)
190
191 % Calculate with no high injection, then if flag
192 % is true recalculate for high currents only.
193 JnEs = q*materials.B.D*materials.B.ni^2 / ...
194     (materials.B.L*device.Nb) .* ...
195     coth(device.Wb_eff/materials.B.L);
196 JnE = JnEs .* ( (exp(VBE/Vt) - 1) - ...
197     sech(device.Wb_eff/materials.B.L) .* ...
198     (exp(VBC/Vt) - 1) );
199
200 JpEs = q*materials.E.D*materials.E.ni^2 / ...
201     (materials.E.L*device.Ne) .* ...
202     coth(device.We_eff/materials.E.L);
203 JpE = JpEs .* ( exp(VBE/Vt) - 1 );
```

```
204
205 % Collector currents
206 JnCs = q*materials.B.D*materials.B.ni^2 / ...
207     (materials.B.L*device.Nb) .* ...
208     csch(device.Wb_eff/materials.B.L);
209 JnC = JnCs .* ( (exp(VBE/Vt) - 1) - ...
210     coth(device.Wb_eff/materials.B.L) .* ...
211     (exp(VBC/Vt) - 1) );
212
213 if device.highinject
214 % With high injection
215 hi_ind_p = exp(VBE/Vt) > 1e-3*((device.Ne/materials.E.ni/2).^2 + 1);
216 hi_ind_n = exp(VBE/Vt) > 1e-3*((device.Nb/materials.B.ni/2).^2 + 1);
217
218 % Emitter currents
219 JnEsBE = q*materials.B.D*device.Nb / (2*materials.B.L) .* ...
220     coth(device.Wb_eff(hi_ind_n)/materials.B.L);
221 JnE(hi_ind_n) = JnEsBE .* ( sqrt( 1 + ...
222     4*materials.B.ni^2*(exp(VBE(hi_ind_n)/Vt)- 1)/device.Nb^2 ) ...
223     - 1 ) - JnEs(hi_ind_n) .* ...
224     sech(device.Wb_eff(hi_ind_n)/materials.B.L) .* ...
225     (exp(VBC(hi_ind_n)/Vt) - 1);
226
227 JpEs_hi = q*materials.E.D*device.Ne / (2*materials.E.L) .* ...
228     coth(device.We_eff(hi_ind_p)/materials.E.L);
229 JpE(hi_ind_p) = JpEs_hi .* ( sqrt( 1 + ...
230     4*materials.E.ni^2*(exp(VBE(hi_ind_p)/Vt) - ...
231         1)/device.Ne^2 ) - 1 );
232
233 % Collector currents
234 JnCsBE = q*materials.B.D*device.Nb / (2*materials.B.L) .* ...
235      csch(device.Wb_eff(hi_ind_n)/materials.B.L);
236 JnC(hi_ind_n) = JnCsBE .* ( sqrt( 1 + 4*materials.B.ni^2 * ...
237     (exp(VBE(hi_ind_n)/Vt) - 1)/device.Nb^2 ) - 1 ) - ...
238     JnCs(hi_ind_n) .* ...
239     coth(device.Wb_eff(hi_ind_n)/materials.B.L) .* ...
240     (exp(VBC(hi_ind_n)/Vt) - 1);
241
242 end
243
244 JCOs =  q*( materials.B.D*materials.B.ni^2 / ...
245         (materials.B.L*device.Nb) .* ...
246         coth(device.Wb_eff/materials.B.L) + ...
247         materials.C.D*materials.C.ni^2 / ...
248         (materials.C.L*device.Nc) .* ...
249         coth((device.Wc_eff+device.Wsc_eff)/materials.C.L) );
250 JCO = -JCOs .* (exp(VBC/Vt) - 1);
```

```
JCO = JCO + device.IC_leak./device.Abc;

Jbbs =   q*( materials.E.ni^2*materials.E.b*HBT.xBE*device.Nb / ...
         ( device.Nb+device.Ne) + ...
         materials.B.ni^2*materials.B.b * ...
         HBT.xBE*device.Ne/(device.Nb+device.Ne) );
Jbb = Jbbs .* (exp(VBE/(2*Vt))-1); % Band-to-band

% Note: For SRH, assume n=p=ni and Et=Ei. Worst case, maximized
% recombination.
Jsrhs = q/2*( materials.E.ni/materials.E.tau_recomb*HBT.xBE * ...
    device.U_eff*device.Nb/(device.Nb+device.Ne) + ...
    materials.B.ni/materials.B.tau_recomb * ...
    HBT.xBE*device.U_eff * device.Ne/(device.Nb+device.Ne) );
Jsrh = Jsrhs .* (exp(VBE/(2*Vt))-1); % SRH

if device.surfacerecomb
Ls = 100e-7; % Surface diffusion length. Assume 100 nm.
% Emitter parameter of device, cm.
parameter = 2*device.L + 2*device.Wee;
nsurf = 2;

Isurf_factor = 1;

Isurf = Isurf_factor * parameter*q*materials.B.Srec * Ls * ...
        materials.B.ni * ( exp(VBE/(nsurf*Vt)) );
else
Isurf = zeros(size(VBE));
end

JrB = JnC .* ( JnE./(JnC+eps) - 1 );

% Record internal voltages
IV.VBE = VBE;
IV.VBC = VBC;
IV.VCE = VCE;

%%% Calculate resistances (with emitter current crowding) %%%

% Current crowding effective emitter width factor
if device.currentcrowding
% Use prefitted data in 'Effective_area_fit.mat'
Jb_crowd =  JpE + (Jbb + Jsrh) - JCO + ...
            JrB + Isurf./device.Abe;
% Note: kb in this function is J/K, NOT eV/K, as in physcost.m
X = reshape(abs(q * Jb_crowd * device.Wee.^2 * R.pB ./ ...
    (8 * device.Wb_eff * kb * device.T )), 1, device.Npt^2);
```

```
Ewidth_factor = reshape(exp(feval(device.effective_area_fit, ...
                log(X))), device.Npt, device.Npt);
device.Ewidth_factor = Ewidth_factor;
device.Wee_eff = device.Wee.*device.Ewidth_factor;
% Effective areas
device.Abc_eff = repmat(device.Abc, device.Npt, device.Npt);
device.Abe_eff = device.L*device.Wee_eff;
else
device.Ewidth_factor = ones(device.Npt,device.Npt);
device.Wee_eff = repmat(device.Wee,device.Npt,device.Npt);
% Effective areas
device.Abc_eff = repmat(device.Abc,device.Npt,device.Npt);
device.Abe_eff = repmat(device.Abe,device.Npt,device.Npt);
end

% Find device voltages and currents to output.
IV.JE = JnE + JpE + (Jbb + Jsrh) + Isurf./device.Abe_eff;
IV.JC = JnC + JCO;
IV.JB = JpE + (Jbb + Jsrh) - JCO + JrB + Isurf./device.Abe_eff;
% Total currents
IV.IE = IV.JE .* device.Abe_eff;
IV.IC = IV.JC .* device.Abe_eff;
IV.IB = IV.JB .* device.Abe_eff;

% Emitter series resistance
R.Re_epi = R.pE*device.We_eff./device.Abe_eff;
R.Re_contact = device.rcE./(device.Wee_eff*device.L);
R.Re_total = R.Re_epi + R.Re_contact; % RE = Re_epi + Re_contact

% Base series resistance
R.Rb_int = R.pB.*device.Wee_eff./(12*device.L.*device.Wb_eff);
R.Rb_ext = R.pB.*device.Seb./(device.L.*device.Wb_eff_lat);
R.Rb_contact = sqrt(R.pB./device.Wb_eff_lat.*device.rcB) ./ ...
    device.L.*coth(device.Wbb.*...
    sqrt(R.pB./(device.Wb_eff_lat.*device.rcB)));
% RB = Rb_int + 1/2*Rb_ext + 1/2*Rb_contact
R.Rb_total = R.Rb_int + 1/2*R.Rb_ext + 1/2*R.Rb_contact;

% Collector series resistance
R.Rc_epi = R.pC*device.Wc_eff./device.Abc_eff;
R.Rsc_int = R.pSC.*(device.Wee + ...
    2*(device.Seb+device.Sbs+device.Wbb)) ./ ...
    (12*device.L.*device.Wsc_eff);
R.Rsc_ext = R.pSC.*device.Sbc./(device.L.*device.Wsc_eff);
R.Rc_contact = sqrt(R.pSC./device.Wsc_eff.*device.rcC) ./ ...
    device.L.*coth(device.Wcc.*...
    sqrt(R.pSC./(device.Wsc_eff.*device.rcC)));
```

```
% RC = Rc_epi + Rsc_int + 1/2*Rsc_ext + 1/2*Rc_contact
R.Rc_total =    R.Rc_epi + R.Rsc_int + ...
                1/2*R.Rsc_ext + 1/2*R.Rc_contact;

% Calculate external voltages
if device.seriesR
IV.VBEext = VBE + device.Abe_eff.*( IV.JB.*R.Rb_total + ...
    IV.JE.*R.Re_total );
IV.VBCext = VBC + device.Abe_eff.*( IV.JB.*R.Rb_total - ...
    IV.JC.*R.Rc_total );
IV.VCEext = IV.VBEext - IV.VBCext;
else
IV.VBEext = VBE;
IV.VBCext = VBC;
IV.VCEext = IV.VBEext - IV.VBCext;
end

outofrangeVbe = abs(IV.VBEext) ≥ ( device.Vcutoff*abs(VBE) + 1 );
outofrangeVbc = abs(IV.VBCext) ≥ ( device.Vcutoff*abs(VBC) + 1 );

IV.outofrangeVbe = ones(device.Npt,device.Npt);
IV.outofrangeVbc = ones(device.Npt,device.Npt);
IV.outofrangeVbe( outofrangeVbe ) = nan;
IV.outofrangeVbc( outofrangeVbc ) = nan;

% Take value where other voltage is zero.
% 1 for 0 -- V, N/2 for -V -- V
if device.symmetricV
IV.Vbe = IV.VBEext(round(device.Npt/2),:).* ...
         IV.outofrangeVbe(round(device.Npt/2),:);
IV.Vbc = IV.VBCext(:,round(device.Npt/2)).* ...
         IV.outofrangeVbc(:,round(device.Npt/2));
else
IV.Vbe = IV.VBEext(1,:).*IV.outofrangeVbe(1,:);
IV.Vbc = IV.VBCext(:,1).*IV.outofrangeVbc(:,1);
end
IV.Vce = IV.Vbe' - IV.Vbc;

IV.JnE = JnE;
IV.JpE = JpE;
IV.JnC = JnC;
IV.JCO = JCO;
IV.Jsrh = Jsrh;
IV.Jbb = Jbb;
IV.JrB = JrB;
IV.Jsurf = Isurf./device.Abe_eff;

```

```
392 IV.gamma = JnE ./ IV.JE; % Emitter injection efficiency
393 IV.alphaT = JnC ./ JnE; % Base transport efficiency
394 IV.dr = ( IV.JE - (Jbb + Jsrh) ) ./ IV.JE;
395 IV.alpha = IV.JC ./ IV.JE; % Common-base current gain
396 IV.beta = IV.JC ./ IV.JB; % Common-emitter current gain
397
398 end
399
400 % Check for unphysical conditions.
401 function [ VBE VBC VCE ] = hbtcheck( VBE, VBC, VCE )
402 %HBTCHECK Checks conditions of HBT to make sure
403 % it is physically valid.
404 %Will put NaN in invalid V.
405 %   [ VBE VBC VCE ] = hbtcheck( VBE, VBC, VCE )
406 %       Will check for full depletion is base and collector and for
407 %       breakdown field in xBE and xBC.
408 end
409
410 % Compute and initialize material paramters
411 function [ materials ] = initializematerials( materials, device )
412 end
413 end
```

B.1.2 multi_ionization.m

```
1  % Compute fermi level and carrier concentration for
2  %   multi-energy level semiconductor systems with both
3  %   shallow and deep acceptors and donors.
4  % October, 2013. Zefram Marks
5
6  %% Physical Constants
7  physconst;
8  mats = materialdata();
9
10 %% Measured data
11     %%%%% String name of matfile of 4pt thermal fitting data %%%%%
12 fitdata.matname = 'thermal_data_GaAsBi.mat';
13 fitdata.importFlag = true; % Select which sheets get imported?
14 fitdata.Nguess = 1e18;
15
16 %% User Input
17
18 %%%%%%%%%%%%%%%%%%%%   USER PARAMETERS (START) %%%%%%%%%%%%%%%%%%%%
```

```
19
20      %%%%% Name of material from materialdat(), e.g. 'GaAsBi' %%%%%
21  inputstruct.material_name = 'GaAsBi';
22
23      %%%%% Reference deep levels to either valance band
24      %     or conduction band.
25  % Either 'val' or 'con' %%%%%
26  inputstruct.deepref = 'val';
27  % inputstruct.deepref = 'both';
28
29      %%%%% Alloy fraction, if applicable
30      %     (e.g. for GaAsBi) (0<x<1) %%%%%
31  inputstruct.x = 2/1e2; % Bi %, not really used right now.
32
33      %%%%% Number of points to compute over %%%%%
34  inputstruct.Npt = 1e2;
35
36      %%%%% [ T_low T_high ] temperature range to compute (K) %%%%%
37  inputstruct.Trange = ...
38      [ 77 C2K(mats.(inputstruct.material_name).meltpt) ];
39  % inputstruct.Trange = [ 50 300 ];
40  % inputstruct.Trange = C2K([0 200]);
41
42      %%%%% Temperature range type.
43      %     Either linear over 'T' or 'invT' %%%%%
44  % inputstruct.Ttype = 'invT';
45  inputstruct.Ttype = 'invT';
46
47
48      % Activation energies for deep/shallow acceptors/donors (eV) and
49      % Concentrations of deep/shallow acceptors/donors (cm^-3) %%%%%
50
51      % NOTE: Fit experimental thermal data to
52      %       Ln(p) = Ln(N0) - Ea/kb*(1/T),
53      %       with Ed2 = Ea and Na1-Nd2=N0
54
55      %%%%%%%%%% NOMINAL VALUES %%%%%%%%%%
56
57      %%%%% Shallow Donors %%%%%
58  inputstruct.Nd1 = 0; % cm^-3
59  inputstruct.Ed1 = 6e-3; % eV
60      %%%%% Deep Donors
61  inputstruct.Nd2 = 1e17; % cm^-3
62  inputstruct.Ed2 = 90e-3; % eV
63      %%%%% Shallow Acceptors
64  inputstruct.Na1 = 1e17; % cm^-3
65  inputstruct.Ea1 = 26e-3; % eV
```

```
    %%%%% Deep Acceptors
inputstruct.Na2 = 1e17; % cm^-3
inputstruct.Ea2 = 27e-3; % eV


    %%%%%%%%%% LOOP VARIABLES %%%%%%%%%%

loopvarname = 'Nd2';
loopvarNpt = 8;
loopvartype = 'linear';        % Use either 'log' or 'linear'
loopvarrange = [5e16 5e17];
% loopvarrange = [45 50]*1e-3;

includezeroFlag = true;     % Include zero element in range.

    %%%%% Cell array of string variable names (N or E) %%%%%
linkedvarnames = { 'Nd2', 'Na2' };
% linkedvarnames = { 'Na1' };
    %%%%% Cell array of scaling types %%%%%
    %%%%% 'diff' NOT IMPLEMENTED YET. IGNORE FOR NOW %%%%%
    %%%%% Multiplicative factor 'factor'
    % (linkedvar = linkedvarscales*loopvar)
    %%%%% Or arithmetic difference 'diff'
    % (loopvar - linkedvar = linkedvarscales).
% linkedvartypes = { 'factor' };
linkedvartypes = { 'factor', 'factor' };
    %%%%% Vector array of scaling values to loopvarname %%%%%
% linkedvarscales = { 1 };
linkedvarscales = { 1, 1e-3 };

    %%%%%%%%%% PLOT VARIABLES %%%%%%%%%%

    % Range to plot y-axis for p0 vs. invT
    DopePlotRange   =   [ 15 19 ];
    plotFlag        =   true;
    refreshfigFlag  =   true;

%%%%%%%%%%%%%%%%%%%%   USER PARAMETERS (END)   %%%%%%%%%%%%%%%%%%%%


%% Generate input and output common structures

mats = materialdata();

% Semiconductor Parameters
Nc = @(T) mats.(inputstruct.material_name).Nc(T);
Nv = @(T) mats.(inputstruct.material_name).Nv(T);
```

```
113 Eg = @(T, x) mats.(inputstruct.material_name).Eg(T, x);
114 Ei = @(T, x) Eg(T, x)/2 + 3/4*kb*T * ...
115     log(mats.(inputstruct.material_name).mh / ...
116         mats.(inputstruct.material_name).me);
117
118 % Test if loop variables are correct
119 if ¬isequal(size(linkedvarnames), size(linkedvartypes), ...
120             size(linkedvarscales) )
121     error('Loop cell arrays are not all the same size.')
122 end
123
124 switch lower(loopvartype)
125     case 'log'
126         loopvarval = logspace(log10(loopvarrange(1)), ...
127                               log10(loopvarrange(2)), loopvarNpt);
128     case 'linear'
129         loopvarval = linspace(loopvarrange(1), ...
130             loopvarrange(2), loopvarNpt);
131     otherwise
132         error('Variable loopvartype not reconized.')
133 end
134
135 % Prepend zero to sweep vector
136 if includezeroFlag
137     loopvarval = [ 0 loopvarval ];
138     loopvarNpt = loopvarNpt + 1;
139     cbarrange = [ 0 loopvarrange(2) ];
140 else
141     cbarrange = loopvarrange;
142 end
143
144 % Generate temperature range
145 switch lower(inputstruct.Ttype)
146     case 't'
147         % Temperature Input, K
148         outputstruct.Tk = linspace(inputstruct.Trange(1), ...
149             inputstruct.Trange(2), inputstruct.Npt);
150         outputstruct.T = K2C(outputstruct.Tk);
151         outputstruct.invT = 1000./outputstruct.Tk;
152     case 'invt'
153         % Temperature Input, 1000/T (K^-1)
154         outputstruct.invT = linspace(1e3./inputstruct.Trange(2), ...
155             1e3/inputstruct.Trange(1), inputstruct.Npt); % 1000/T
156         outputstruct.Tk = 1000./outputstruct.invT;
157         outputstruct.T = K2C(outputstruct.Tk);
158     otherwise
159         error('Temperature range type not understood. ...
```

```
160                 Use either T or invT.')
161 end
162
163 inputstruct.Nc = Nc(outputstruct.Tk);
164 inputstruct.Nv = Nv(outputstruct.Tk);
165
166 % Output calculated bands
167 outputstruct.bands.Eval = zeros(1,inputstruct.Npt);
168 outputstruct.bands.Econ = Eg(outputstruct.Tk, inputstruct.x);
169 outputstruct.bands.Ei = Ei(outputstruct.Tk, inputstruct.x);
170
171 % Preallocate cell structure arrays
172 inputstructcell = cell(1,loopvarNpt);
173 outputstructcell = cell(1,loopvarNpt);
174
175 % Put in nominal values to each structure in cell
176 [ inputstructcell{:} ] = deal(inputstruct);
177 [ outputstructcell{:} ] = deal(outputstruct);
178
179 % Prepare non-common factors in cell structure
180 parfor ii = 1:loopvarNpt
181
182     % Input loop variables
183     inputstructcell{ii}.(loopvarname) = loopvarval(ii);
184
185     for jj = 1:length(linkedvarnames)
186         switch lower(linkedvartypes{jj})
187         case 'factor'
188             inputstructcell{ii}.(linkedvarnames{jj}) = ...
189                 linkedvarscales{jj} .* loopvarval(ii);
190         case 'diff'
191             error('Difference scaling factor not implemented yet.')
192         otherwise
193             error('Variable linkedvartypes not reconized.')
194         end
195     end
196
197     % Reference deep levels to either valance band
198     % or conduction band
199     %   BOTH ==> FUNCTION ASSUMES THAT [EA1, EA2]>0 ARE REFERENCED
200     %            TO THE VALENCE BAND, AND THAT [ED1, ED2]>0 ARE
201     %            REFERENCED TO THE CONDUCTION BAND.
202     switch lower(inputstructcell{ii}.deepref)
203         case 'val'
204             outputstructcell{ii}.bands.Ed2 = ...
205                 Eg(outputstructcell{ii}.Tk, ...
206                 inputstructcell{ii}.x) - inputstructcell{ii}.Ed2;
```

```
207             outputstructcell{ii}.bands.Ea2 = ...
208                 repmat(inputstructcell{ii}.Ea2, 1, ...
209                         inputstructcell{ii}.Npt);
210         case 'con'
211             outputstructcell{ii}.bands.Ed2 = ...
212                 repmat(inputstructcell{ii}.Ed2, 1, ...
213                         inputstructcell{ii}.Npt);
214             outputstructcell{ii}.bands.Ea2 = ...
215                 Eg(outputstructcell{ii}.Tk, ...
216                 inputstructcell{ii}.x) - inputstructcell{ii}.Ea2;
217         case 'both'
218             outputstructcell{ii}.bands.Ed2 = ...
219                 repmat(inputstructcell{ii}.Ed2, 1, ...
220                         inputstructcell{ii}.Npt);
221             outputstructcell{ii}.bands.Ea2 = ...
222                 repmat(inputstructcell{ii}.Ea2, 1, ...
223                         inputstructcell{ii}.Npt);
224         otherwise
225             error('deepref variable not understood. ...
226                     Choose either val or con strings')
227     end
228     % Additional acceptor and donor levels
229     outputstructcell{ii}.bands.Ea1 = ...
230         repmat(inputstructcell{ii}.Ea1, 1, inputstructcell{ii}.Npt);
231     outputstructcell{ii}.bands.Ed1 = ...
232         repmat(inputstructcell{ii}.Ed1, 1, inputstructcell{ii}.Npt);
233
234 end
235
236 %% Compute ionized carrier concentration
237 parfor ii = 1:loopvarNpt
238
239     [ outputstructcell{ii} ] = ...
240         compute_ionization( inputstructcell{ii}, ...
241             outputstructcell{ii} );
242
243 end
244
245 %% Import measurement data from mat file (if exists)
246
247 % NOTE: Correct input structures are assumed to have fields
248 %       'Gsh', 'Tk',
249 % 'd', and 'source'. If the structure does not have these
250 % fields they will not be imported.
251
252     % If fitting structure exists, find doping amount
253 if fitdata.importFlag && exist(fitdata.matname, 'file') == 2
```

```
254
255     % Will load in all structures of the correct format.
256
257     % For existing data
258     necessary_fields_import = {'N_data', 'Tk', 'source'};
259     % For computation
260     necessary_fields = {'Gsh', 'Tk', 'd', 'x', 'source'};
261     toremove = [];
262
263     % Import mat file into workspace
264     fitdata.inputstruct = load(fitdata.matname);
265
266     % Find all field names and which ones are structures
267     fitdata.allstructnames = fieldnames(fitdata.inputstruct);
268     fitdata.allstructnames = fitdata.allstructnames( ...
269         structfun(@isstruct, fitdata.inputstruct) );
270
271     fitdata.source_names = cell(1,length(fitdata.allstructnames));
272
273
274     for ii = 1:length(fitdata.allstructnames)
275
276         field_check = isfield( fitdata.inputstruct.(...
277             fitdata.allstructnames{ii}), necessary_fields);
278
279         if all(isfield( fitdata.inputstruct.(...
280             fitdata.allstructnames{ii}), necessary_fields_import ))
281
282             % If no goodInd field, mark all data points as good
283             if ¬isfield(fitdata.inputstruct.(...
284                 fitdata.allstructnames{ii}), 'goodInd')
285
286                 fitdata.inputstruct.(...
287                     fitdata.allstructnames{ii}).goodInd = ...
288                     true(1, length( fitdata.inputstruct.(...
289                         fitdata.allstructnames{ii}).N_data ) );
290
291             end
292
293             fitdata.source_names{ii} = fitdata.inputstruct.(...
294                 fitdata.allstructnames{ii}).source;
295
296         elseif ¬all( field_check )
297
298             warning([ 'Structure "' fitdata.allstructnames{ii} ...
299                '" does not contain fields ''' ...
300                strjoin(necessary_fields(¬field_check), ...
```

```
301                     ''', ''') '''' ]);
302
303             toremove = [ toremove ii ];
304
305         else
306
307             % Calculate doping concentration
308             fitdata.inputstruct.(...
309                 fitdata.allstructnames{ii}).N_data = ...
310                 fitdata.inputstruct.(...
311                     fitdata.allstructnames{ii}).Gsh ./ ...
312                   ( q .* fitdata.inputstruct.(...
313                     fitdata.allstructnames{ii}).d .* ...
314                     mats.(inputstruct.material_name).mu_p( ...
315                     fitdata.inputstruct.(...
316                         fitdata.allstructnames{ii}).Tk, ...
317                     fitdata.Nguess, fitdata.inputstruct.(...
318                         fitdata.allstructnames{ii}).x ) );
319
320             % If no goodInd field, mark all data points as good
321             if ¬isfield(fitdata.inputstruct.(...
322                 fitdata.allstructnames{ii}), 'goodInd')
323
324                 fitdata.inputstruct.(...
325                     fitdata.allstructnames{ii}).goodInd = ...
326                     true(1, length( fitdata.inputstruct.(...
327                         fitdata.allstructnames{ii}).N_data ) );
328
329             end
330
331             fitdata.source_names{ii} = fitdata.inputstruct.(...
332                 fitdata.allstructnames{ii}).source;
333
334         end
335
336     end
337
338     % Remove structures without proper fields
339     fitdata.allstructnames(toremove) = [];
340     fitdata.source_names(toremove) = [];
341
342     % Ask user to choose which structures to import
343     toimport = listdlg('PromptString', ...
344         'Which fit data structures should be imported?', ...
345         'ListString', fitdata.source_names );
346
347     fitdata.allstructnames = fitdata.allstructnames( toimport );
```

```
348     fitdata.source_names = fitdata.source_names( toimport );
349 
350     Nfitdata = length(fitdata.allstructnames);
351 
352 else
353     Nfitdata = 0;
354     fitdata.plothandles = [];
355     fitdata.source_names = {};
356 end
```

Glossary of equation variable notations

A_0	Crystalline lattice constant
α	DC transport factor of transistor. Defined as $\beta \triangleq \frac{J_C}{J_E} = \gamma_E \alpha_T \delta_R$
α_T	Base transport factor. Defined as ratio of electron current into collector to electron current into base. $\alpha_T \triangleq \frac{J_{n,E} - J_{r,B}}{J_{n,E}}$
β	DC current gain of transistor. Defined as $\beta \triangleq \frac{J_C}{J_B}$
$C_{j,BC}$	Junction capacitance in the base-collector P-N junction.
$C_{j,BE}$	Junction capacitance in the base-emitter P-N junction.
δ_R	Depletion region recombination factor. Defined as ratio of electron and hole current across base-emitter junction to total emitter current. $\delta_R \triangleq \frac{J_{n,E} + J_{p,B}}{J_E}$
$D_{p/n}$	Diffusion coefficient of p-type or n-type minority carriers

E_A	Activation energy of mid-gap energy states.
E_G	Band gap energy. Energy difference between highest valance band and lowest conduction band level.
E_C	Energy level of lowest conduction band
ΔE_C	Band gap discontinuity in the conduction band. $\Delta E_C \triangleq E_C^{Wide} - E_C^{Narrow}$
ΔE_G	Total band gap discontinuity. $\Delta E_G = \Delta E_V + \Delta E_C$
ΔE_V	Band gap discontinuity in the valence band. $\Delta E_V \triangleq E_V^{Narrow} - E_V^{Wide}$
E_F	Fermi energy level. Reference energy for Fermi-Dirac function. Also called the "chemical potential."
E_V	Energy level of highest valence band
ϵ_0	Permittivity of free space
F_{Max}	Maximum frequency of a HBT to have power gain. Power gain is one at f_{Max}.
F_T	Maximum transit frequency of a HBT. Point when current gain of transistor is 1 ($\beta(f_T) = 1$)

γ_{E}	Emitter efficiency. Defined as ratio of electron current to total (electron and hole) current across the base-emitter junction. $\gamma_{\mathrm{E}} \triangleq \frac{J_{\mathrm{n,E}}}{J_{\mathrm{n,E}}+J_{\mathrm{p,B}}}$
H	Planck constant
$\hbar$	Reduced Planck constant, $\hbar = {}^{h}/_{2\pi}$
Type I Heterojunction	Straddling gap: The narrow band gap semiconductor is entirly within the band gap of the wide gap semiconductor. $\Delta E_{\mathrm{V}} > 0$, $\Delta E_{\mathrm{C}} > 0$
Type II Heterojunction	Staggered gap: The band gaps overlap, but the narrow gap semiconductor is either partial above or below the wide gap semiconductor. $\Delta E_{\mathrm{C}} \triangleq E_{\mathrm{C}}^{\mathrm{Wide}} - E_{\mathrm{C}}^{\mathrm{Narrow}}$
Type III Heterojunction	Broken-gap: The band gaps do not overlap at all.
J_{B}	Current density into base contact of transistor
J_{C}	Current density into collector contact of transistor
$J_{\mathrm{n,C}}$	Electron current density from minority carriers from the base into the collector. $J_{\mathrm{n,C}} = J_{\mathrm{n,E}} - J_{\mathrm{r,B}}$
$J_{\mathrm{n,E}}$	Electron current density from emitter into base and collector
J_{E}	Current density into emitter contact of transistor

$J_{\mathrm{p,C}}$	Hole current density from reversed biased collector into base
$J_{\mathrm{p,B}}$	Hole current density from base into the emitter
$J_{\mathrm{r,B}}$	Recombination current density of minority carriers in the base
$J_{\mathrm{r,d}}$	Recombination current density in base-emitter depletion region
J_{S}	Diode saturation current density.
J_{SR}	Surface recombination current linear density. Total surface current proportional to device perimeter.
K_{B}	Boltzmann constant
$L_{\mathrm{p/n}}$	Diffusion length of p-type or n-type minority carriers. $L_{\mathrm{p/n}} \triangleq \sqrt{D_{\mathrm{p/n}}\tau_{\mathrm{p/n}}}$
L_{T}	Transfer length in TLM measurement
$M^{*}_{\mathrm{n/p}}$	Effective mass in a semiconductor, as fraction of electron mass.
M_0	Fundamental electron rest mass.
$\mu_{\mathrm{n/p}}$	Electron or hole mobility in a semiconductor. Defined as $\mu \triangleq \frac{\lvert\vec{v}\rvert}{\lvert\vec{\varepsilon}\rvert}$, where $\vec{v}$ is the carrier velocity at low field strength and $\vec{\varepsilon}$ is the electric field.

$N_{A/D}$	Acceptor (p-type) or donor (n-type) dopant concentration
$N_{E/B/C}$	Dopent concentration in the Emitter, Base, or Collector.
N_C	Effective density of states of the conduction band
N_V	Effective density of states of the valence band
N_i	Intrinsic carrier concentration: Concentration of holes and electrons in an undoped semiconductor.
P_{n0}, n_{p0}	Equibrium minority carrier concentrations. Holes in the n-type region and electrons in the p-type region.
Q	Elementary charge (charge of an electron or proton)
R_B	Total resistance in the base layer.
R_C	Total resistance in the collector layer.
R_E	Total resistance in the emitter layer.
$\tau_{p/n}$	Lifetime of p-type or n-type minority carriers

V_{A}	Applied voltage bias across a P-N junction. Defined as $V_{\mathrm{A}} = V_{\mathrm{p}} - V_{\mathrm{n}}$.
$V_{\mathrm{A}}^{\mathrm{ext}}$	Externally applied voltage across a P-N junction. Includes the voltage drop from a non-zero resistance. $V_{\mathrm{A}}^{\mathrm{ext}} = V_{\mathrm{A}}^{\mathrm{int}} + IR$
V_{T}	Thermal voltage. $V_{\mathrm{T}} = \frac{1}{q} k_{\mathrm{B}} T = 25.86\,\mathrm{meV}$ at room temperature.
$W'_{\mathrm{p/n}}$	Effective width of the quasi-neutral p-type or n-type region ($w'_{\mathrm{p/n}} = w_{\mathrm{p/n}} - x_{\mathrm{p/n}}$)
X_{d}	P-N depletion region width. x_{p} and x_{n} represents the width of the depletion region on the p-type or n-type side.

Acronyms

BJT bipolar junction transistor

BOE buffered oxide etchant

C-V capacitive–voltage

CBM conduction band minimum

CMOS complementary metal-oxide-semiconductor

DI deionized

HBT heterojunction bipolar transistor

HMA highly mismatched alloy

I-V current–voltage

LPE liquid phase epitaxy

MBE molecular beam epitaxy

MOSFET metal-oxide-semiconductor field effect transistor

NREL National Renewable Energy Labortory

PECVD plasma enhanced chemical vapor deposition

PL photoluminescence

RC resistor–capacitor

RF radio frequency

RIE reactive ion etching

RTA rapid thermal anneal

SEM scanning electron microscope

SI semi-insulating

SIMS secondary-ion mass spectrometry

SRH Shockley Reed Hall

TLM transfer length measurement

TTL transistor-transistor logic

VBM valence band maximum

XRD X-ray diffraction