GaAsP/InGaP Heterojunction Bipolar Transistors for III-V on Si Microelectronics

by

Christopher Heidelberger

Bachelor of Science in Materials Science and Engineering
Cornell University, 2012

Submitted to the Department of Materials Science and Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Materials Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

Signature of Author: _____
Department of Materials Science and Engineering
August 31, 2017

Certified by: _____
Eugene A. Fitzgerald
Merton C. Flemings-SMA Professor of Materials Science and Engineering
Thesis Supervisor

Accepted by: _____
Donald R. Sadoway
John F. Elliot Professor of Materials Chemistry
Chair, Departmental Committee on Graduate Students

GaAsP/InGaP Heterojunction Bipolar Transistors for III-V on Si Microelectronics

by

Christopher Heidelberger

Submitted to the Department of Materials Science and Engineering
on August 31, 2017 in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Materials Science and Engineering

Abstract

GaAs-based transistors are capable of operating at high frequency with low noise, and are produced in large volumes for a wide range of applications including microwave frequency ICs for input/output in mobile devices. However, Si CMOS still holds an advantage for digital logic due to wide market penetration resulting in decades of development and lower cost. Monolithic integration of III-V analog circuity and Si CMOS gives circuit designers the best of both materials. In addition, by substituting $GaAs_xP_{1-x}$ ($0.8 < x < 1$) for GaAs as an active material, we can take advantage of its higher breakdown voltage and reduced lattice mismatch with Si. In this thesis, we study GaAsP/InGaP heterojunction bipolar transistors (HBTs) grown via MOCVD as a test-bed for III-V microelectronics integration with Si.

Epitaxial challenges involving growth of GaAsP/InGaP HBT structures on Si substrates were addressed. Heavy C p-type doping of GaAsP via MOCVD, necessary for the HBT base region, was studied. Growth rate, composition, and hole concentration dependence on C precursor ($CBrCl_3$) input was investigated, yielding GaAsP films with hole concentrations in excess of 2×10^{19} cm^{-3}. $GaAs_{0.825}P$ was grown on Si substrates via a SiGe graded buffer with a threading dislocation density of 3.7×10^6 cm^{-2} measured by PV-TEM and EBIC. This density is appropriate for fabrication of minority-carrier devices such as HBTs.

GaAsP/InGaP HBTs were fabricated on both GaAs and Si substrates with a range of defect densities to measure the effect on DC performance and prove the feasibility of GaAsP transistor growth on Si. Models for the effect of threading dislocation density and misfit dislocation density (in the active device layers) on current gain were developed. A GaAsP/InGaP HBT grown on Si was demonstrated with a current gain as high as 158. Changes in $GaAs_xP_{1-x}$ composition from $0.825 < x < 1$ did not have a significant effect on current gain. Collector current was determined not to be controlled by thermionic emission of electrons from the emitter into the base, contrary to prior reports. In addition, GaAsP was shown to support a higher breakdown voltage than GaAs, consistent with modeling.

Thesis Supervisor: Eugene A. Fitzgerald
Title: Merton C. Flemings-SMA Professor of Materials Science and Engineering

Table of Contents

List of Figures

List of Tables

Acknowledgements

> You know, if you take everything I've accomplished in my entire life and condense it down into one day, it looks decent!
>
> –George Costanza

When you watch someone's PhD defense, it's easy to get the impression that they did their experiments in a well-planned, linear sequence. But any doctoral student will learn that this is far from the truth. Scientific research, and more broadly, innovation, are highly non-linear processes that are difficult, if not impossible, to predict from the outset (Gene can tell you more about this). Dead ends and roadblocks are common. One of my biggest lessons during my PhD was learning how to be persistent in the face of this high level of uncertainty. As I underwent this journey, I was lucky to be surrounded by amazing people who helped me at every step.

The person who deserves my first and foremost thanks is my advisor, Prof. Eugene Fitzgerald. Gene gave me a great deal of latitude in shaping my project with just the right amount of guidance and support along the way. His level of enthusiasm and positivity always had me excited to come to group meeting to discuss my latest results. Gene's approach to advising has taught me to think for myself as a scientist and I am truly grateful for that. I would also like to thank my thesis committee members Prof. Silvija Gradečak and Prof. Lionel Kimerling for providing helpful feedback over the years.

I would like to acknowledge the Center for Energy Efficient Electronics Science (E^3S, NSF Award No. 0939514) for financially supporting my thesis work, as well as for creating an exciting cross-institutional environment for collaboration and learning. In addition, thanks to IBM for supporting me as a PhD Fellow.

The members of the Fitzgerald Group share a common bond in running and maintaining our MOCVD reactor. The reactor is a substantial, yet delicate multi-component system requiring a special blend of scientific understanding, practical know-how, patience, persistence, and humor to keep operational. Because of this, our group has developed a rare level of comradery and friendship. Mayank Bulsara served as almost a second advisor for my first couple of years and was always available to help with lab issues. Prithu Sharma welcomed me into the group as my first office-mate. Adam Jandl taught me most of what I know about managing and fixing the reactor. Kunal Mukherjee had an infectious curiosity and was one of my foremost technical advisors. Tim Milakovich taught me all about growing GaAsP on SiGe, UHVCVD maintenance, and lifting weights. I will always remember trips to Berkeley with Ryan Iutzi, which always culminated with our tradition of cigars and cocktails with Gene. Roger Jia always selflessly helped others in the group. As the youngest and final two students in the Fitzgerald Group, Rushabh Shah and I were especially collaborative. Rushabh has a physics background and I have always appreciated his unique perspective.

I am lucky to have been part of a number of fruitful collaborations with other groups from MIT and around the world. I would particularly like to thank Prof. Jeehwan Kim for including me as a collaborator as he started up his group at MIT, and to his students Yunjo Kim, Samuel Cruz,

and Scott Tan for eagerly helping with MOCVD maintenance. Collaborations that nucleated through E^3S have included those with Prof. Ming Wu and Seth Fortuna at UC Berkeley, as well as with Prof. Jesús del Alamo and Xin Zhao at MIT. I have also collaborated with David Kohen and Lee Kwang Hong from SMART. I am grateful for the opportunity to work with these wonderful scientists.

Many others at MIT have helped me in various capacities. Jamie Teherani and Winston Chern from Prof. Dimitri Antoniadis' group assisted with electrical characterization, along with Corentin Monmeyran and Danhao Ma from the Kimerling Group. I made extensive use of both CMSE and MTL, and I wish to thank the helpful and dedicated staff of both of these invaluable user facilities. In addition, many thanks to Mark Beals and the rest of MPC for supporting the SEL. Thank you to Angelita Mireles and Elissa Haverty from the DMSE academic office for your assistance.

Throughout my entire life, my family has been incredibly supportive of me. My parents have given me everything that I need to succeed and much more. Mom, Dad, and Leigh, thank you for your unconditional love and generosity. Also, thanks to all of the friends I've made at MIT and in the Boston area. You have made this a great five years and I look forward to more. In particular, I want to thank my former roommates at The Palace, Scott Grindy, Steve Dacek, and Mike Campion, for all of the family dinners, parties, Game of Thrones viewings, and generally great times. Lastly, thanks to Jamie for your love, kindness, and encouragement (and for putting up with so many long conversations about materials science). I can't wait to share many more adventures with you.

Chapter 1: Introduction

1.1. Motivation

High-gain, low-noise transistors for analog radio frequency (RF) signal processing have many established and emerging applications. These include broadband radio communications; high data-rate fiber systems; concealed weapon detection; passive imaging systems capable of "seeing" through rain, snow, and fog; and environmental, atmospheric, and pollution monitoring [1]. Perhaps the most prominent today is broadband radio communications, which includes input/output from now-ubiquitous smartphones and other mobile devices. Transistors for RF circuitry have fundamentally different requirements compared to transistors used for digital data processing—such as silicon CMOS—which have benefitted from Moore's Law scaling over the past several decades. While Si CMOS has moved to small sizes and low voltages to increase speed and decrease power, RF devices are required to handle relatively high voltages and are less constrained by their footprint. With higher electron mobilities, band gaps, and critical breakdown fields, III-V semiconductors allow RF devices to operate at higher temperatures, higher voltages, and with lower on-state resistances than Si [2], [3].

One attractive prospect is the integration of analog III-V-based circuitry and digital Si CMOS monolithically on the same chip. This delivers the benefits of III-V materials for RF signal processing while still leveraging well-developed and cost-effective Si CMOS for data processing [4]. This level of integration is possible by growing the III-V films epitaxially on Si substrates. While this approach has been viewed as problematic due to the large lattice mismatch and inherent incompatibility of Si and most III-V compounds [5], recent advances in metamorphic epitaxy have

made the growth of device-quality III-V films on Si a reality [6], [7]. Further benefits of monolithic integration of III-V compounds with Si include its superior mechanical strength, increased size, and lower cost compared to III-Vs and Ge [5], as well as energy savings from reduced chip-to-chip communication. In addition, a monolithic III-V on Si platform offers the possibility of layer transfer of the III-V active region onto another Si handle wafer allowing reuse of the original substrate and buffer layers. This is easier than doing the same from a III-V substrate (e.g. GaAs) because it eliminates the thermal expansion mismatch between the donor and recipient substrates [3].

1.2. Challenges of III-V semiconductor growth on Si substrates

Growth of III-V semiconductors (particularly III-P, III-As, and III-Sb) on Si has several difficulties. The foremost issue is the lattice mismatch between these III-V semiconductors and Si. The lattice constants of GaAs and InP are highlighted in Figure 1.1 along with Si. These two lattice constants are very popular for growth of III-V heterostructure devices [8]. Their lattice mismatch (defined as the percent difference in lattice constant) with Si is 4% and 7%, respectively. While these numbers may sound small, if GaAs or InP are grown directly on Si, they develop very large strain energies which must relax via plastic deformation of the film. This results in high defect density and poor film morphology.

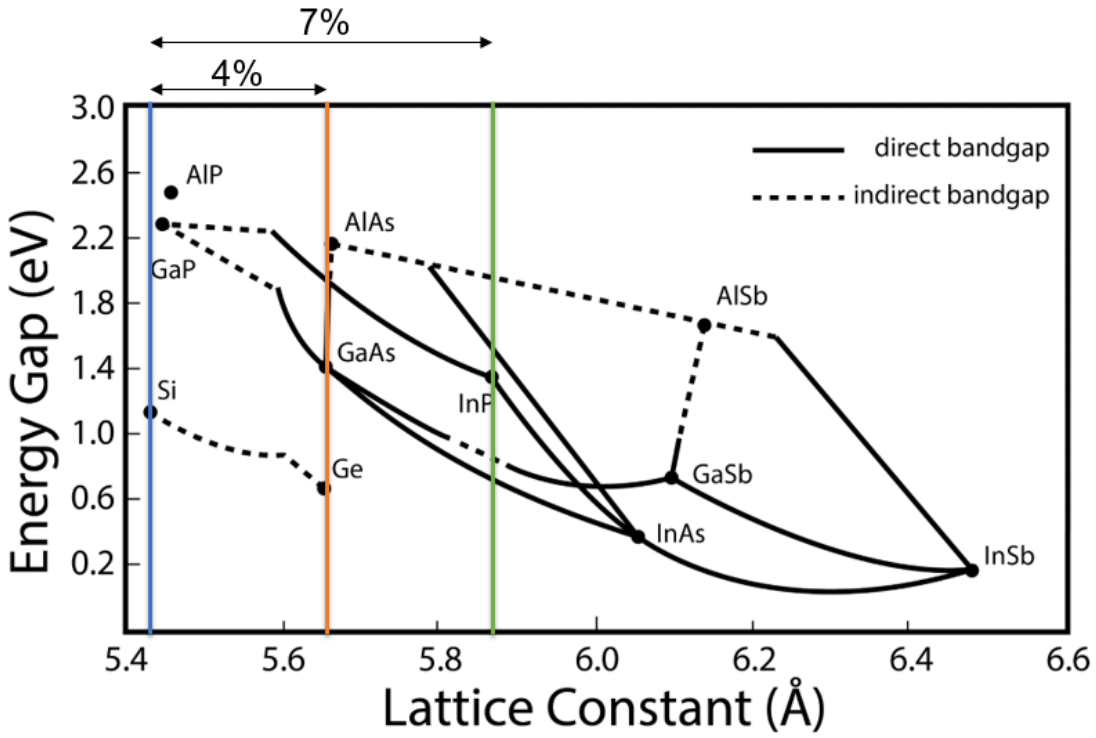

Figure 1.1: Band gap vs lattice constant of Si, Ge, and various III-V semiconductors. Binary or pseudobinary alloys are represented as solid or dashed lines, while end members are represented by dots. Lattice mismatch between Si–GaAs and Si–InP highlighted.

The concept of a critical thickness, above which a lattice-mismatched film will relax via plastic deformation, was developed by Mathews and Blakeslee [9]. A detailed review of this topic and expansions to include kinetic considerations was published by Fitzgerald, and will not be discussed further here [10]. However, it is important to note that the critical thickness for a film with 1% lattice mismatch is typically less than 10 nm [11]. Therefore, any useful III-V devices grown on Si cannot be grown in a pseudomorphically strained state.

Figure 1.2 demonstrates how the strain energy of a film grown beyond its critical thickness is relieved by the formation of misfit dislocations at the film-substrate interface. These misfit dislocations themselves are not generally a problem, because they reside well below the active III-V device layers. However, because a dislocation cannot end inside of a crystal, each misfit dislocation will end in two threading segments that penetrate through the surface of the film (Figure 1.3). These threading dislocations will therefore will interact with the active device layers. Deleterious effects of threading dislocations include reduced carrier lifetime [12], [13] and

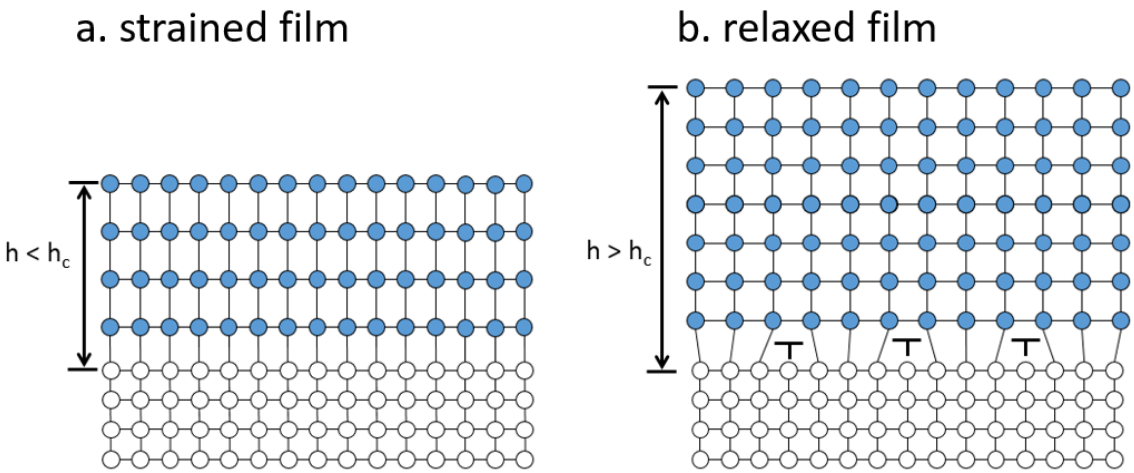

Figure 1.2: Ball-and-stick diagram of a mismatched epitaxial film. In (a), the film is below the critical thickness and is therefore pseudomorphically strained to the substrate. In (b), the film is above the critical thickness and has relaxed through the formation of misfit dislocations along the film-substrate interface.

mobility [14], as well as increased leakage current [15]. A quantitative understanding of these interactions is critical to growing devices on mismatched substrates.

Another issue with growth of III-V semiconductors on Si or Ge is the heterovalent III-V on group IV interface. At this interface, there is a break in crystal symmetry between the diamond cubic structure of Si/Ge and the zinc blende structure of the III-V semiconductors in Figure 1.1. The III-V lattice forms domains with one of two different orientations, where the only difference is that the group III and group V atoms are swapped [16]. This is also equivalent to a 90° rotation and translation. At the intersection of these domains is what is known as an anti-phase boundary (APB). These defects are expected to act as strong scattering and recombination centers, as well as causing local energy band bending in their vicinity, similarly to dislocations [17]–[19].

It has been demonstrated that by growing the III-V semiconductor on an intentionally offcut Si or Ge substrate, the formation of APBs is suppressed [20]–[22]. Figure 1.4a shows a schematic of a III-V/IV interface where an APB has formed at a step edge in the group IV material. By using a (100) substrate that is offcut by 6° towards the [111] direction and including the proper annealing step, a lower-energy double-step structure is formed on the group IV surface, thereby preserving the III-V sublattice orientation between neighboring terraces. This allows the III-V

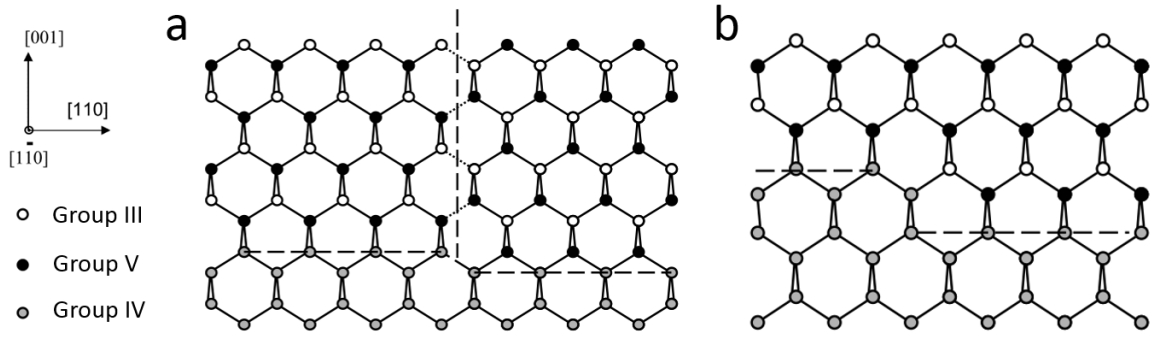

Figure 1.4: Ball and stick model of III-V/IV epitaxial interface. In (a), a single step in the group IV material leads to an APB in the III-V. In (b), a double step in the group IV material yields APB-free III-V. (Reproduced from Ting, 1999 [98]*.)*

material to form in a single domain with no APBs (Figure 1.4b). Specific details of this procedure are discussed in Section 4.2.

1.3. Strain relief schemes for lattice-mismatched epitaxy

Direct growth of GaAs on Si has been reported to yield a threading dislocation density (ρ_{TD}) as high as 10^{10} cm^{-3} [7]. Increased GaAs thickness decreases ρ_{TD}—none the less, GaAs at practical thicknesses possesses $\rho_{TD} > 10^8$ cm^{-2}, a value that renders devices with poor performance and low reliability. Therefore, if GaAs or InP is grown directly on Si, the resultant dislocation density will be in a range that prohibits fabrication of any useful electronic devices. However, through the use of one or more strain relief schemes, we are able to obtain a dislocation density at an acceptable value. In this section, we will discuss the general theory behind two of the schemes used in this thesis.

1.3.1. Compositionally graded buffers

A structure known as a compositionally graded buffer, or just graded buffer, can be used to achieve large levels of lattice mismatch with substantially reduced threading dislocation densities [7], [23]. For any lattice-mismatched film to relax completely, a certain total length of misfit dislocations must be present at the interface. This length can be achieved by one of two ways, both illustrated in Figure 1.5. A preexisting threading dislocation can glide along the mismatched interface, creating new misfit length and increasing the relaxation of the film (Figure 1.5a). Alternatively, a dislocation loop can nucleate from the surface of the film, creating new misfit length but also creating two new threading dislocations (Figure 1.5b). The former of these is preferred, because it does not increase the threading dislocation density of the film. However,

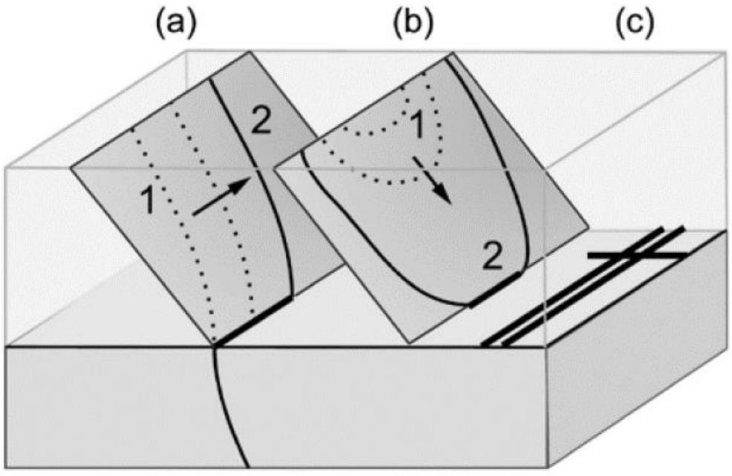

Figure 1.5: (a) A preexisting threading dislocation glides to create new misfit length. (b) A new dislocation half-loop is nucleated to create new misfit length. (c) A typical misfit dislocation network. (Reproduced from Mukherjee, 2014 [11]*)*

for an interface with a large lattice mismatch (e.g. GaAs grown directly on Si), the rate of process (b) is very high and creates a large number of threading dislocations.

An alternate method would be to introduce a small amount of strain at a time, layer by layer. In the first layer, process (b) would still be present, but only a relatively small number of threading dislocations would be nucleated because of the small mismatch at this point. Then, in subsequent layers, those threading dislocations are "reused" through process (a) to create the necessary misfit dislocation length for relaxation to take place. This is illustrated in Figure 1.6. Leitz, et al. have studied the kinetics of relaxation by dislocation glide in graded buffers, and have determined that the threading dislocation density (ρ_{TD}) necessary, in equilibrium, to relieve strain without further nucleation is [24]:

$$\rho_{TD} = \frac{2R_g R_{gr} \exp\left(\frac{E_a}{kT}\right)}{b v_0 Y^m \varepsilon_{eff}^m}, \qquad 1.1$$

where R_g is the growth rate, R_{gr} is the grade rate (in units of lattice mismatch per thickness), E_a is the activation energy for dislocation glide, b is the magnitude of the Burgers vector, Y is the

Young's modulus of the film, ε_{eff} is the residual strain in the top-most layer of the film, and m and v_0 are fitting parameters. The equilibrium value of ρ_{TD} is linearly dependent on grade rate, so a slower grade rate should yield a higher-quality film. However, slower grade rate means a thicker film and longer growth time, which is costly. In addition, ρ_{TD} has an Arrhenius temperature dependence. Thus, higher growth temperatures are generally preferred for graded buffer growth absent other factors.

High-quality $GaAs_{0.7}P_{0.3}$ has been grown on Si via a SiGe graded buffer with a threading dislocation density near 10^6 cm^{-2} [22]. GaAsP photovoltaic cells have been demonstrated on this platform and have performance comparable to devices grown on GaAs substrates [25]. Thus, SiGe graded buffers are likely to provide a viable platform for the integration of III-V microelectronic devices with Si. Fully relaxed Ge has been demonstrated on Si via a SiGe graded buffer with a threading dislocation density of 2.1×10^6 cm^{-2} [26]. One drawback to using graded buffers is that they tend to be relatively thick epitaxial structures, which can substantially increase processing costs. One approach to circumventing this cost would be to use a layer transfer technique to

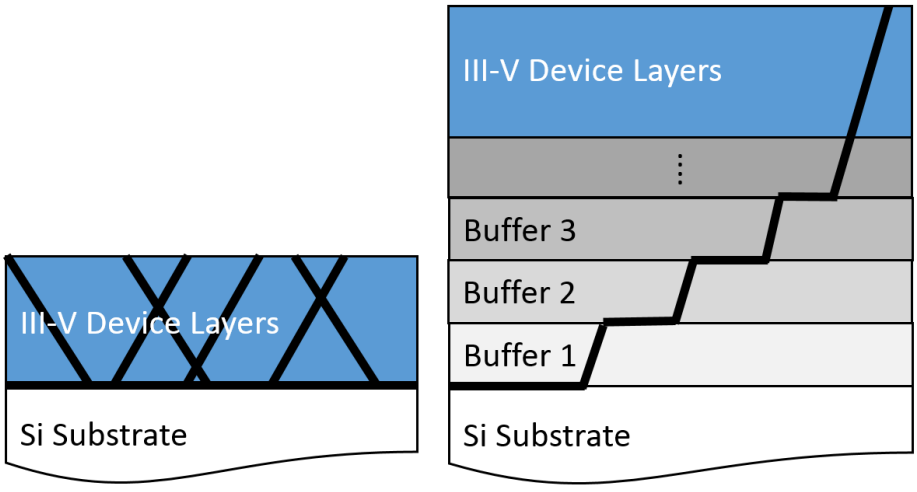

Figure 1.6: Direct growth of mismatched III-V films on a Si substrate vs growth via a graded buffer.

remove the device layers and then reuse the graded buffer and substrate to grow more device layers.

1.3.2. Direct "two-step" Ge growth on Si

Here, we will discuss a method for growing Ge on Si with a reasonably low dislocation density, but without requiring a thick SiGe graded buffer. Known as "two-step Ge," this technique involves direct growth of Ge on Si in a two-step process with a subsequent cyclic annealing step to reduce dislocation density. This process is capable of generating Ge/Si structures with a threading dislocation density of around 2×10^7 cm^{-2} [27], [28]. First, a very thin (~30 nm) Ge film is deposited at relative low temperature (350–400 °C). The low temperature is chosen so that the Ge film is conformal and does not form 3-dimensional islands due to the high strain (i.e. Volmer-Weber growth mode). Second, a thicker Ge film (1 µm) is grown at higher temperature (600 °C). The films are then subjected to a cyclic thermal anneal, which due to the difference in coefficient of thermal expansion between the Ge film and Si substrate, causes dislocations to glide back and forth and annihilate. This reduces the threading dislocation density from an initial value of around 1×10^9 cm^{-2} to the final value mentioned above [27], [29].

Benefits of the two-step Ge process as compared to the SiGe graded buffer include a much thinner epitaxial structure (~1 µm vs ~10 µm) and not having to calibrate SiGe growth across the entire range of composition. However, it cannot achieve as low of a threading dislocation density for Ge on Si as the SiGe graded buffer (~2×10^6 cm^{-2} [26]). In addition, it cannot be used to access intermediate lattice constants between those of Si and Ge.

1.4. GaAsP/InGaP heterojunction bipolar transistors as a test-bed for III-V on Si integration

In this thesis, we choose to use the GaAsP/InGaP heterojunction bipolar transistor (HBT) as a "test-bed" of sorts to explore the possibility of using III-V on Si as a platform for microelectronics fabrication. HBTs have a unique set of materials requirements as compared to photonic devices and photovoltaic cells. GaAsP/InGaP high-electron-mobility transistors (HEMTs) were also considered, and will be discussed further in Section 3.2. In this section, we will talk about why the GaAsP/InGaP pairing was chosen instead of more conventional materials, such as GaAs/AlGaAs, GaAs/InGaP, or InGaAs/InAlAs.

The III-As and III-P semiconductor system offers the possibility of fabricating structures where adjacent layers are made of different materials with different band gaps. Known as band structure engineering, this can allow for vast improvement of device performance over devices made from a single semiconductor. This approach is used widely in almost every type of device made from III-V semiconductors, including transistors, light-emitting diodes, lasers, photovoltaic cells, photodetectors, and more.

An HBT structure requires materials with two different bandgaps: a higher bandgap for the emitter layer, and a lower bandgap for the base and collector layers. These layers must be lattice-matched or close to lattice-matched to each other so that defects are not nucleated in the active region of the device. The higher bandgap of the emitter (more specifically, the valence band offset between the emitter and base) is necessary to prevent reverse injection of holes from the base into the emitter. The different current components of an HBT and its band structure requirements will be discussed further in Section 3.1. In this work, we will refer to the materials used in an HBT

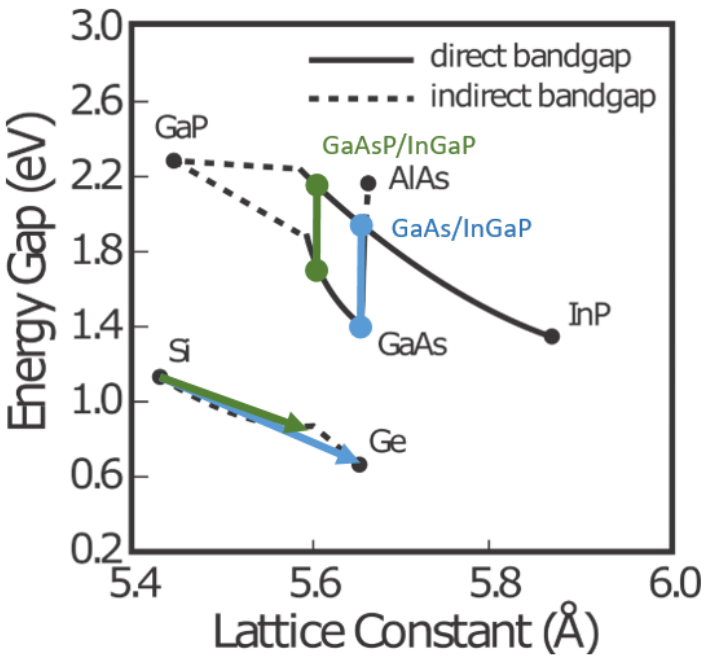

Figure 1.7: Lattice constant and band gap of GaAsP/InGaP vs GaAs/InGaP heterostructures. GaAsP/InGaP offers higher band gaps (higher breakdown voltage) and a closer lattice constant to Si.

either as [emitter material]/[base and collector material] (e.g. GaAs/InGaP) , or by just the base and collector material (e.g. GaAs) where the emitter material is implied.

GaAs is used commonly as a base/collector material for HBTs found widely in the marketplace (especially for cell phone power amplifiers), with InGaP used as the emitter [30]–[36]. Compared to InP-based structures, GaAs-based structures are more readily integratable with Si because they have lower lattice mismatch. InGaP is generally a better choice for the emitter material than AlGaAs because of its lower conduction band offset with GaAs, which results in improved injection efficiency [30], [34]. In addition, InGaP avoids oxygen-related defects associated with AlGaAs layers and InGaP/GaAs has superior wet-etching selectivity to AlGaAs/GaAs heterostructures.

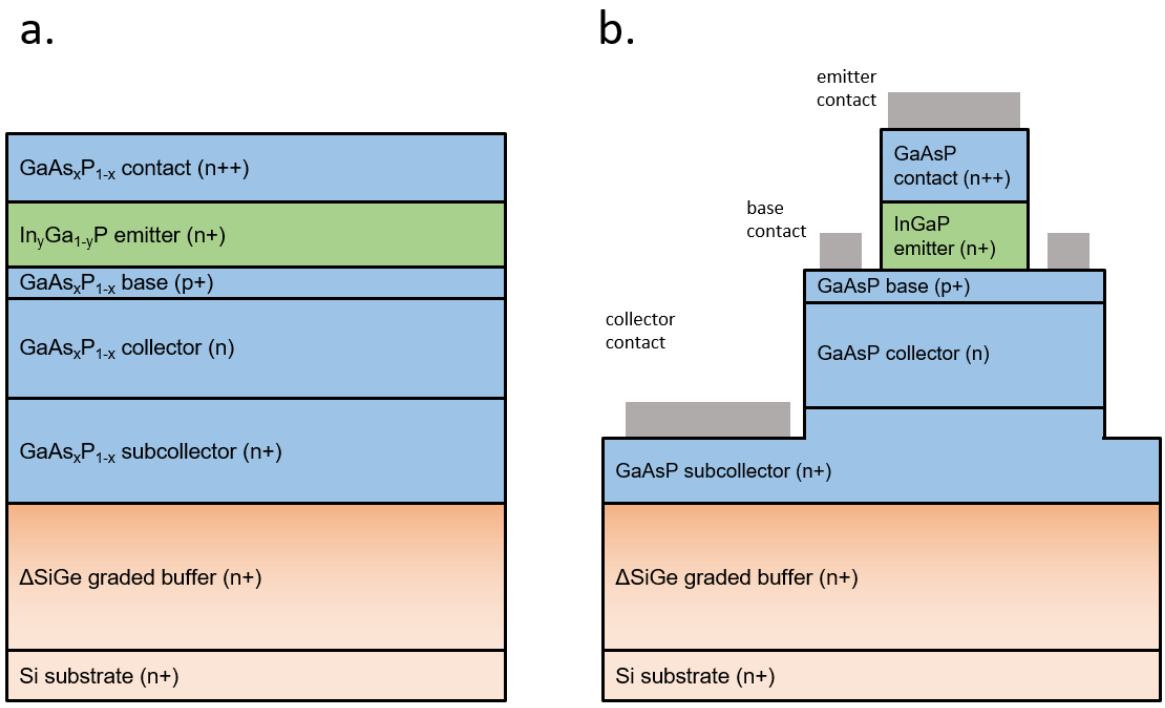

Figure 1.8: Schematic of (a) the GaAsP/InGaP HBT epitaxial structure to be studied in this thesis and (b) the fabricated HBT device.

The blue markings on Figure 1.7 denote the GaAs/InGaP pairing along with the SiGe graded buffer that could be used to access it from the Si lattice constant. What if, instead of using GaAs as the base and collector material, we used $GaAs_xP_{1-x}$ with $0.8 < x < 1$? $In_yGa_{1-y}P$ can still be used for the emitter, except with y adjusted so that it remains lattice-matched with the $GaAs_xP_{1-x}$ base/collector. This arrangement is shown in green in Figure 1.7. A SiGe graded buffer is still used to bridge the lattice mismatch between the GaAsP/InGaP device layers and the substrate. A generalized schematic of this epitaxial structure is shown in Figure 1.8. At these compositions of GaAsP, there is not a significant decrease in electron mobility from that of GaAs [37]. The use of GaAsP instead of GaAs has several advantages. The higher band gap of GaAsP allows for a higher breakdown voltage. Also, with decreasing As content, the lattice constant of GaAsP decreases, getting closer to that of Si. The reduced lattice mismatch simplifies growth on Si of the III-V device layers with sufficiently low defect density. Lastly, the coefficient of thermal

expansion (CTE) of GaAsP is lower than that of GaAs and closer to that of Si, reducing the likelihood of the III-V epi-layers cracking due to CTE mismatch between the films and the Si substrate [38]. In this thesis, we seek to verify these claims, as well as gain an understanding of the effect of dislocations on HBT performance.

Chapter 2: Methods

In this chapter, we will discuss the experimental methods used in this thesis for epitaxial growth, film characterization, and device fabrication in general terms. Specific details of each experiment are included in subsequent chapters.

2.1. Band structure simulation

Band structure simulation of various heterostructures was performed using Nextnano. Nextnano is a commercially-available software tool which can solve the Poisson, Schrodinger, drift, and diffusion equations to calculate band structure, carrier densities, and I-V behavior of a given semiconductor structure [39]. It includes a database of necessary materials parameters for many different semiconductors, including Si, Ge, and all of the III-Vs (including ternaries and quaternaries). One feature is the ability to define the strain states of each layer with the band gaps calculated via the deformation potentials. It is reasonably user-friendly.

In this work, all of the Nextnano simulations were conducted in one dimension. Certain parameters, such as the band gap of InGaP and the band alignment between GaAs(P) and InGaP, are dependent on growth parameters such as temperature and V/III ratio. These dependencies were not taken into account in the simulations presented here. Although the effects are estimated to be small, it should be kept in mind that the simulation is only as good as its inputs.

2.2. Metalorganic chemical vapor deposition (MOCVD) growth

2.2.1. Reactor design

Most of the epitaxial samples (including all of the III-V samples) discussed in this work were grown by metalorganic chemical vapor deposition (MOCVD). MOCVD is chosen over molecular beam epitaxy (MBE) because of superior throughput and cost, and the fact that it is widely used in industry, making this work more transferrable. Our reactor is a Thomas Swan/Aixtron close-coupled showerhead flip-top design, manufactured c. 2004, pictured in Figure 2.1. The available precursors are dimethylzinc (DMZn), B_2H_6 (1% in H2), trimethylaluminum (TMAl), trimethylgallium (TMGa), $CBrCl_3$, SiH_4, Si_2H_6 (0.1% in H_2), GeH_4 (15% in H_2), PH_3, AsH_3, trimethylantimony (TMSb), and diethyltelluride (DETe).

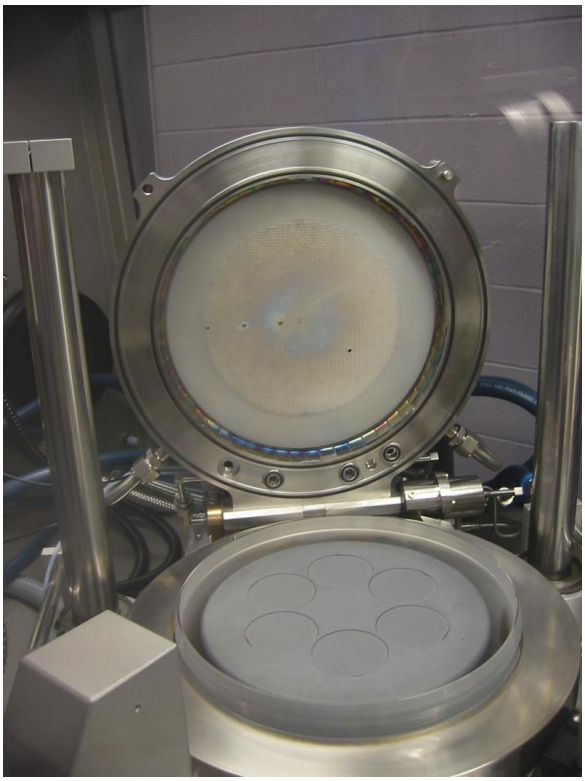

Figure 2.1: Picture of the close-coupled showerhead MOCVD used in this thesis.

A heavily simplified schematic of the gas layout is shown in Figure 2.2. Precursors are dosed using mass flow controllers and are heavily diluted and carried into the reaction chamber with either a N_2 or H_2 carrier gas. The carrier gas flow was held constant at 20 slm for all process runs. Almost all of the III-V material grown in this work uses a N_2 carrier gas. While H_2 is more common in literature, N_2 has higher purity on this system and its use alleviates the requirement of monitoring and changing of H_2 cylinders during process runs. Each precursor can be individually flowed either into the reactor or into a bypass line called the vent line. By flowing each precursor to the vent line for at least 60 s before flowing into the reactor, the flow rates are allowed to stabilize, reducing any non-uniformities at the beginnings of layers. The reactor includes both feed-back and look-ahead based flow stabilization systems, which further improve flow stability and abruptness at interfaces. The close-coupled showerhead design of the reactor optimizes gas flow geometry and provides excellent growth rate and compositional uniformity across the wafers. The showerhead features a dual-plenum design, where group III and group V sources are not

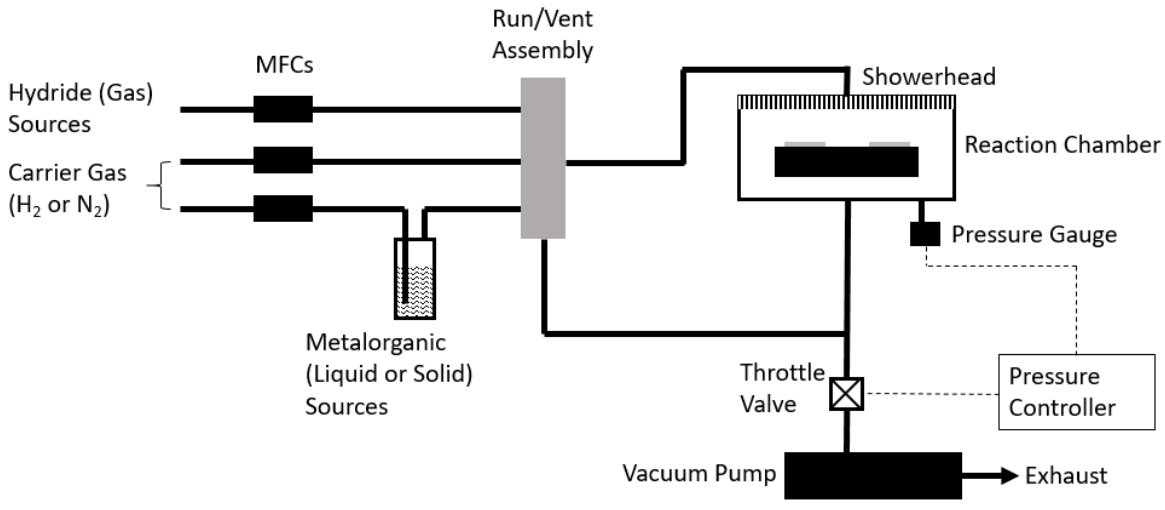

Figure 2.2: Simplified schematic of the MOCVD reactor used in this thesis.

allowed to mix until directly above the wafer surface. This reduces the opportunity for upstream reactions and adduct formation between the precursors.

Wafers sit on a SiC-coated graphite susceptor in the reaction chamber. The susceptor is heated from underneath by a graphite heater. The reactor is cold-walled, with the stainless steel reactor body being continuously water-cooled to 50 °C. The graphite heater has three concentric zones. The relative power output of each heater zone can be adjusted such that the thermal profile across the top of the susceptor is constant. The temperature of the top of the susceptor is calibrated using an optical pyrometer, which itself is calibrated using a blackbody calibration furnace. The pyrometer can measure in three locations spaced radially across the susceptor, one corresponding to each heater zone. This way, the three heater zones can be balanced, in addition to calibrating the average temperature to the desired setpoint. This is done for each temperature setpoint once every several months in a process we call "temp. balancing." Temp. balances were completed under controlled conditions of Si-coated glassware (no other growth) and an uncoated 4" susceptor, and were done separately for H_2 and N_2 ambients. All MOCVD growth temperatures quoted in this work are the susceptor surface temperatures calibrated in this manner. The minimum growth temperature is 400 °C and the maximum is 825 °C.

The typical temperature error corrected during a temp. balance is less than 3 °C. Therefore, the run-to-run variation in temperature is of this magnitude or less. However, there is another source of systematic growth temperature variation: the difference between the susceptor surface temperature and the wafer surface temperature. We estimate that the wafer surface temperature is around 20–40 °C cooler than the susceptor surface temperature. This varies due to substrate material and thickness. This effect must be accounted for when growing materials that are particularly sensitive to temperature variation. For example, GaAsP grown on a 650 μm Si

substrate will have a different composition than GaAsP grown on a 300 μm GaAs substrate under the same nominal growth conditions. This is because the surface temperature of the GaAs wafer is estimated to be around 15 °C higher than that of the Si wafer.

The reaction byproducts are pumped from the reactor by an Ebara ESA25-D dry pump. They are then passed through a dry scrubber to remove any toxic or reactive species. For the past 5 years, we have used Matheson Ultimasorb as a scrubber medium with excellent results.

Our reactor has the somewhat unique capability of growing Si and Ge alongside III-V semiconductors. This allows for growth of III-Vs on newly-grown group IV layers without removal of the samples from the reactor chamber (or even cooling down from growth temperature). There are a few modifications made to the reactor to enable this. The primary modification is the addition of a quartz chamber liner. The liner is in two parts: the "J-liner", which covers the outer walls, and the "ceiling", which covers the showerhead and has the same pattern of holes for the precursors to pass through. After Si or Ge growth, the reactor walls and ceiling are covered in Si and Ge which then incorporates into subsequent III-V films. The level of Si or Ge autodoping in GaAs grown under these conditions can be as high as 1×10^{18} cm^{-3}, much too high for any device with sensitive doping requirements. However, by removing the quartz chamber liner, etching off the deposits in acid, and recoating with GaAs, the level of Si and Ge autodoping can be reduced down to less than 1×10^{16} cm^{-3} (as measured by SIMS).

One important consideration for MOCVD growth of In-containing compounds is the bubbler design for the In precursor. The most common In precursor and what we use here, TMIn, is a solid at room temperature. A standard bubbler with solid TMIn is prone to "channeling" effects, where the carrier gas forms wide channels through the solid source material in the bubbler and no longer saturates with the full vapor pressure of TMIn. This causes a reduction in the

effective flow rate of TMIn into the reactor. This historically has been addressed by measuring the TMIn concentration of the gas coming out of the bubbler, and adjusting the flow rate to compensate for any change in concentration with a closed-loop algorithm. This method is prone to stability issues and relies on the accuracy of the TMIn concentration measurement. Alternatively, we use a "Hiperquad" TMIn bubbler manufactured by AkzoNobel. This bubbler, which fits in the same water bath, has four consecutive TMIn chambers and promises very stable pickup rates through the first > 90% of the bubbler lifetime, eliminating the need to run the TMIn line in closed-loop control with the concentration measurement [40]. We have found that the Hiperquad bubbler does in fact yield good pickup rate stability and linearity of TMIn effective flow with programmed carrier flow rate.

2.2.2. Element incorporation and effect of growth temperature

Compositional control of ternary and quaternary III-V compounds as well as SiGe is a critical part of growing successful heteroepitaxial structures. Both band gap and lattice constant vary strongly with composition (Figure 1.1); therefore, composition must often be controlled to within 2 atomic percent or better. Lattice constant (and therefore composition, by Vegard's law) can be measured to well-within this specification using high-resolution x-ray diffraction (see Section 2.4.1). In order to understand the effects of different process parameters such as precursor flow rates, growth temperature, and total reactor pressure, it is important to have at least a rudimentary understanding of the underlying reactions taking place during MOCVD. There are many different steps occurring in series, including diffusion of the precursor through the boundary layer, cracking of the precursor, physisorption and chemisorption of the adatom onto the surface, desorption of byproducts, and diffusion of byproducts away from the film surface. However, for

a particular set of growth conditions, this can often be simplified to a single rate-limiting step that determines both the growth rate and film composition.

Typical MOCVD growth temperatures vary from 550 °C to 750 °C. At lower temperatures, growth rate is kinetically limited by one or more of the physical processes described above [41]. In addition, film quality suffers due to poor adatom mobility. At higher temperatures, the growth rate decreases due to reduced thermodynamic driving force as well as parasitic upstream reactions of the precursors. Intermediate growth temperatures are generally selected at which the growth rate is mass-transport limited.

For mixed group-III pseudobinary alloys such as AlGaAs, InGaP, and InGaAs, compositional control is relatively straightforward. At typical III-V growth temperatures such as 650 °C, the growth is mass-transport limited [42]. This means that the rate-limiting step is diffusion of the precursors through the gas boundary layer above the surface of the wafer. Equation 2.1 shows the dependence of growth rate (G_r) on growth temperature (T_G) and gas velocity (v) [43].

$$G_r \propto T_G^{1.8} v^{0.5} \tag{2.1}$$

Because of the relatively weak dependence on T_G, the composition and growth rate of mixed group-III alloys in this temperature regime is almost independent of temperature. However, because of the dependence on v, reactor geometry and gas dynamics must be considered carefully to ensure a uniform boundary layer thickness.

At lower growth temperatures, the decomposition (cracking) of certain group-III precursors begins to slow to the point where they now limit incorporation. For instance, at 500 °C, slower TMGa cracking begins to limit Ga incorporation into InGaAs, such that the film will

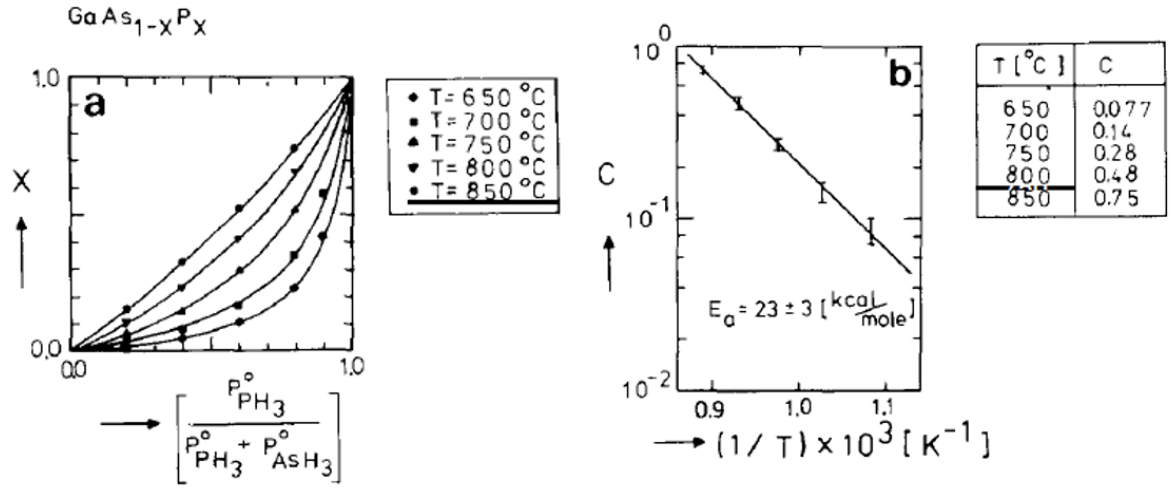

Figure 2.3: (a) GaAsP composition vs PH_3 fraction in gas phase. (b) Proportionality constant C vs temperature in Arrhenius plot. Reproduced from Smeets, 1987 [44].

be Ga-deficient if using the same precursor flows as are used 650 °C. When growth becomes reaction-limited, growth rate takes on an Arrhenius dependence [43]:

$$G_r \propto e^{-\Delta H/kT_G} \qquad 2.2$$

Because of the strong exponential dependence on T_G, composition and growth rate at these lower temperatures is a bit trickier to pin down and rely on precise and repeatable temperature calibration.

Mixed group-V ternary or quaternary compounds, such as GaAsP or InGaAsP, are more difficult to calibrate than mixed group-IIIs. In general, the growth rate is still determined by the mass transport of the group III element(s). The ratio of the group V elements depends on the ratio of As to P in the gas phase. This ratio is controlled by the ratio of AsH_3 to PH_3 in the gas phase, and also by the cracking rate of both precursors. At typical growth temperatures near 650 °C, the AsH_3 cracks almost fully but the PH_3 cracking is kinetically limited. Because of this, the As to P ratio in the film has an exponential dependence on temperature. For $GaAs_xP_{1-x}$, the As to P ratio can be written as [44]:

$$\frac{1-x}{x} = C\frac{P_{PH_3}}{P_{AsH_3}}, \tag{2.3}$$

where P_{AsH_3} and P_{PH_3} are the partial pressures of AsH_3 and PH_3, respectively, and C is a proportionality constant with the form described below:

$$C = C_0 \exp\left(-\frac{E_a}{kT}\right) \tag{2.4}$$

Figure 2.3a shows the P fraction in the solid phase plotted against the PH_3 fraction in the gas phase, demonstrating the hyperbolic dependence of Equation 2.3 across a range of growth temperatures from 650–850 °C. In Figure 2.3b, the proportionality factor C is plotted Arrheniusly, yielding an activation energy (E_a) of 23 ± 3 kcal/mol (1.0 ± 0.13 eV/atom). We have reproduced this data on our reactor for temperatures ranging from 600–725 °C, with E_a = 1.06 eV/atom. This value of E_a is somewhat lower than that reported for the decomposition of PH_3 in the presence of Si and glass. It is hypothesized that the presence of GaAs(P) lowers the activation energy for PH_3 cracking [44].

Si, Ge, and SiGe growth in the MOCVD is in the reaction-limited regime for the entire range of growth temperatures. This means that the growth rate and composition is governed by Equation 2.2 and is therefore highly temperature-dependent.

For both mixed group-III and mixed group-V III-V semiconductors as well as for SiGe, other effects can effect growth rate and composition as well. One such effect is pseudomorphic strain, which will not be considered further here because nearly all layers in this thesis are grown in a relaxed or near-relaxed state. Another is the presence of dopants. This can cause changes in growth rate and composition both due to reactions with dopant atoms on the surface as well as gas-phase reactions with dopant precursors or their byproducts.

2.2.3. V/III ratio

V/III ratio is an important parameter for MOCVD growth. It is defined as the total amount (in mol/min) of group V flow divided by the total amount of group III flow. The group V flow rate is calculated by multiplying the MFC setpoint times the MFC gas correction factor (0.67 for AsH_3 and 0.76 for PH_3) times 4.5×10^{-5} to convert from sccm to mol/min. The group III flow is calculated by the following equation:

$$F_{TMX} = S_{MFC} \frac{P_{TMX}}{P_{Bubbler}} * 4.5 \times 10^{-5}, \quad 2.5$$

where F_{TMX} is the flow rate of the metalorganic precursor in mol/min, S_{MFC} is the setpoint of the MFC in sccm, P_{TMX} is the vapor pressure of the metalorganic at the bubbler temperature, and $P_{Bubbler}$ is the total pressure inside of the bubbler.

V/III ratios typically used in our reactor at MIT are generally on the order of 100 [45]. To be more economical, industry processes use V/III ratios that are somewhat lower, but still >> 1. During growth, there is an equilibrium between the III-V material and the group V element in the gas phase. Then, III-V growth is then controlled by the introduction of group III atoms. The same cannot be done in reverse because at growth temperature, the group III elemental phase is a liquid instead of a gas. Therefore, if a high enough group V overpressure is not maintained while at high temperature, liquid droplets of group III atoms will form on the surface as group V atoms desorb. This is why it is important to maintain a high V/III ratio during growth.

2.2.4. In situ wafer reflectance measurement (EpiTT)

Our MOCVD reactor is fitted with an in situ reflectance monitor, known as EpiTT, manufactured by LayTec. It has a white light source and a detector which measures light at 950 nm. By using a chopper on the light source, it can differentiate between the signal from reflected

light and from blackbody emission from the wafer. It can then calculate wafer temperature and reflectance. The wafer temperature measurement is useful for determining the extent of equilibration after a change in temperature, but is too inaccurate to be of any further use. The reflectance measurement has two distinct uses. First, it can measure approximate film thickness and growth rate. By looking at the number intensity oscillations of reflected light during growth, the film thickness can be calculated by this equation, which is derived from the equation for constructive or destructive interference:

$$d = \frac{m\lambda}{2n} \tag{2.6}$$

Here, d is the film thickness, m is the number of oscillations, λ is the wavelength of light (950 nm), and n is the refractive index of the film. For GaAs, $n \approx 3.3$, so the thickness of film grown per oscillation is about 140 nm. This is similar for all of the III-V materials grown in this thesis. Unfortunately, because of uncertainty in the n values, particularly for ternary semiconductors of arbitrary composition, this method is limited in its usefulness to just a zeroth-order estimation of thickness or growth rate.

Reflectance can also be used to give an indication of surface morphology. If the wafer surface roughens significantly during growth, the reflectance signal drops and eventually falls to near zero. This is useful for troubleshooting recipes where there is a catastrophic issue with the epitaxy quality. With knowledge of the time at which the reflectivity begins to fall, the particular interface or layer with the problem can be identified.

2.3. UHVCVD growth

The remainder of the epitaxial samples in this work that were not grown by MOCVD were grown by ultra-high-vacuum chemical vapor deposition (UHVCVD). These include SiGe graded

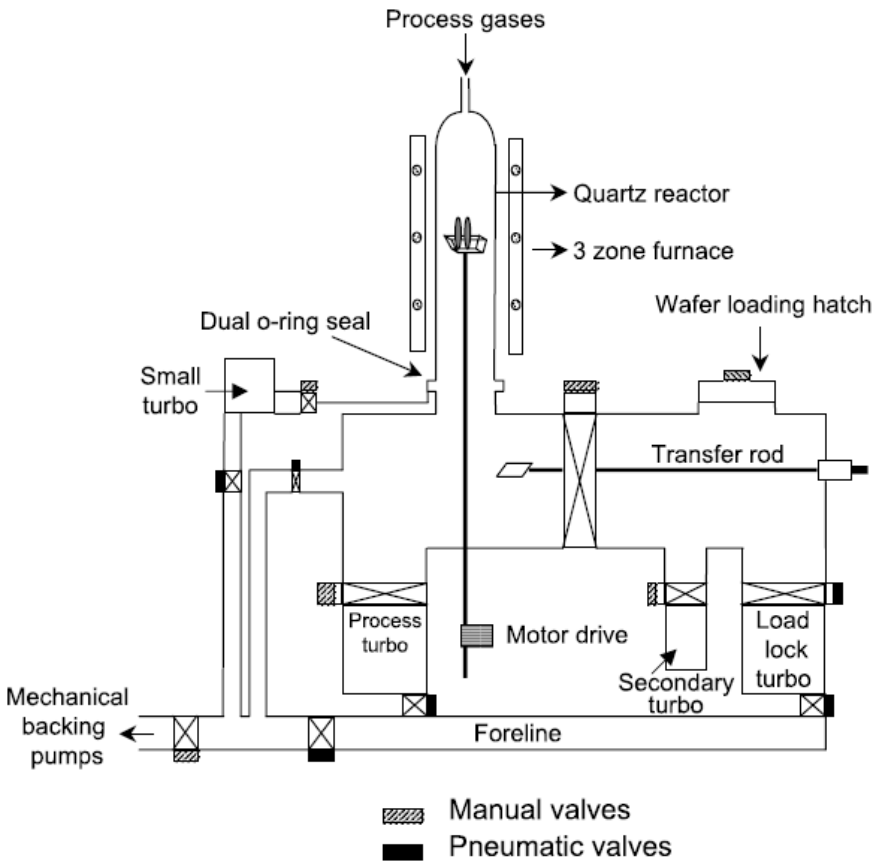

Figure 2.4: Schematic of the SiGe UHVCVD reactor (reproduced from Leitz, 2002 [24]*).*

buffers on Si substrates, which were used as a platform for further growth of SiGe and III-V semiconductors in the MOCVD (see Section 4.2). The UHVCVD reactor used in this work is a hot-walled, vertical tube reactor that was custom built for this purpose. A schematic of the reactor is shown in Figure 2.4. Up to ten 6" wafers are held in a quartz holder centered in a quartz reaction tube. A three-zone electric heater surrounds the chamber and heats it to growth temperatures ranging from 600 °C to 900 °C. Precursors, metered by MFCs, are injected directly into the top of the tube and reaction byproducts are pumped out the bottom via a turbomolecular pump, which is backed by a large roots-blower mechanical roughing pump.

The base pressure of the reaction chamber after a bakeout and Si chamber coating recipe is 1×10^{-9} Torr. The growth pressure is regulated by a manual throttle valve at the outlet of the

reaction chamber directly above the turbomolecular pump, and can vary from 5–30 mTorr. In this pressure and temperature regime, growth is limited by reaction kinetics and not by mass transport of the precursors. This allows for uniform batch processing of wafers without careful regard to the gas flow geometry. A second benefit is that it allows for epitaxial growth on both sides of every wafer if they are double-side polished. This allows for strain induced by coefficient of thermal expansion mismatch between the film and substrate to be cancelled out and wafer bow thereby dramatically reduced.

The precursors used for Si and Ge were SiH_4 and GeH_4, respectively. B and P dopants were provided by B_2H_6 (1% in H_2) and PH_3 (1% in H_2) for p-type and n-type doping, respectively. These dopant precursors were further diluted with Ar before injection into the reactor in two dilution stages. This provides a wide dynamic range allowing for doping levels ranging from 10^{15}–10^{19} cm^{-3}.

2.4. Blanket film characterization

2.4.1. X-ray diffraction (XRD)

High-resolution x-ray diffraction (XRD) allows for the precise measurement of in-plane and out-of-plane lattice constant as well as film tilt. From this information, relaxed lattice constant and strain state of the epitaxial films can be calculated. These measurements are critical to the growth of SiGe and pseudobinary III-V alloys. Diffraction measurements for this entire thesis were taken using a Bruker D8 high-resolution x-ray diffractometer with a Cu K-alpha-1 x-ray source. Reciprocal space maps (RSMs) were gathered using a 1-D Linxeye detector for faster data collection.

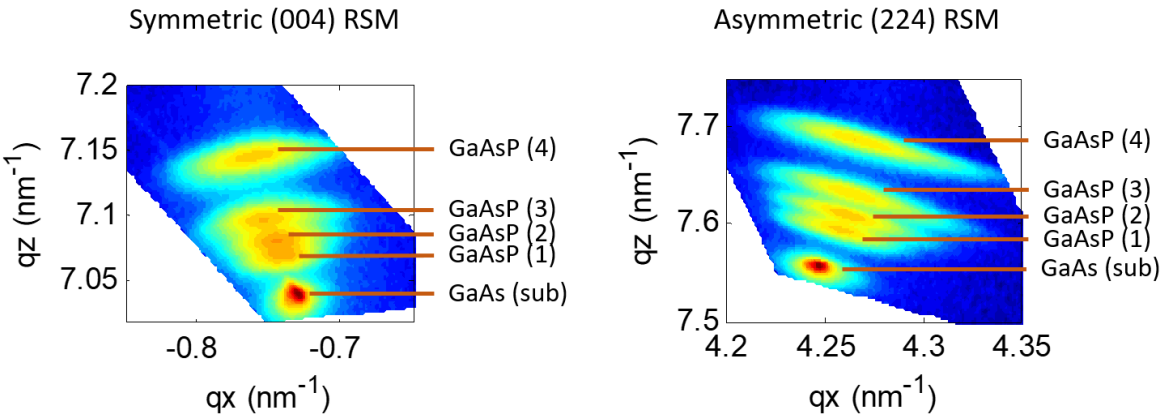

Figure 2.5: Example XRD data from a GaAsP calibration structure (four GaAsP layers of varying composition) on a GaAs substrate.

(004) and (224) RSMs were gathered from samples in order to provide enough data to calculate in-plane lattice constant ($a_{\parallel}$), out-of-plane lattice constant ($a_{\perp}$), and tilt (γ) of the film, according to the method of Roesener, et al. [46]. Figure 2.5 shows sample XRD data from a GaAsP calibration structure. The GaAs substrate peak and four GaAsP film peaks are indexed in each of the RSMs. From each reciprocal space map, the q_x and q_z coordinates of the substrate peak and each film peak were obtained by truncating the upper half of each peak and taking the centroid of that. To simplify the upcoming calculations, the coordinates of the film peaks are redefined in relation to the substrate peaks by $\Delta q_i^{(hh4)} = q_{i,L}^{(hh4)} - q_{i,S}^{(hh4)}$, where $i = x, z$; $h = 0, 2$; and subscripts L and S denote layer and substrate, respectively. Next, γ is calculated for each film layer:

$$\tan(\gamma) = \frac{\Delta q_x^{(004)}}{2\lambda/a_s - \left|\Delta q_z^{(004)}\right|}, \qquad 2.7$$

where λ is the x-ray wavelength (0.154056 nm). Once the tilts for each epitaxial layer are known, they are removed from the data by the following transformation:

$$\begin{pmatrix} q_x^{(hh4)} \\ q_x^{(hh4)} \end{pmatrix}_{corrected} = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} q_x^{(hh4)} \\ q_x^{(hh4)} \end{pmatrix}_{tilted} \quad (h = 0,2) \tag{2.8}$$

Now, using the corrected q_x and q_z values, $a_\parallel$ and $a_\perp$ for each film layer can be calculated as follows:

$$a_\parallel = a_S \times \left[1 - \frac{\Delta q_x^{(224)}}{\sqrt{2}\lambda/a_S + \Delta q_x^{(224)}}\right] \tag{2.9}$$

$$a_\perp = a_S \times \left[1 - \frac{\Delta q_z^{(004)}}{2\lambda/a_S + \Delta q_z^{(004)}}\right] \tag{2.10}$$

From $a_\parallel$ and $a_\perp$, the relaxed lattice constant of each film layer (a_R) can be calculated with knowledge of the Poisson ratio of the film (ν):

$$a_R = \frac{1-\nu}{1+\nu} a_\perp + \frac{2\nu}{1+\nu} a_\parallel \tag{2.11}$$

The value of ν was linearly interpolated from the binary endpoints. The composition of the film can then be calculated using Vegard's Law, which assumes that the relaxed lattice constant of the mixture is equal to a weighted average of the end members. For example, for $Si_{1-x}Ge_x$:

$$a_R = (1-x)a_{Si} + x a_{Ge} \tag{2.12}$$

It is also helpful to define a relaxation parameter R, which is a measure of how relaxed a mismatched film is. An R of 0 indicates a fully pseudomorphically strained film, while an R of 1 indicates a fully relaxed film.

$$R = \frac{a_\parallel - a_S}{a_R - a_S} \tag{2.13}$$

More details of this process, including figures, are available in Roesener, et al. [46].

For calibration of GaAsP, we commonly grew calibration structures with three or four GaAsP layers on either a GaAs substrate or a GaAsP "virtual substrate" on Si. For calibration of SiGe, several SiGe layers were grown on a Si substrate. However, for InGaP, we have found that

In/Ga incorporation is heavily dependent on strain state of the films, and that InGaP films do not relax as easily as GaAsP films. Therefore, InGaP layers were calibrated one at a time on a relaxed GaAsP buffer at approximately the same lattice constant.

2.4.2. Transmission electron microscopy (TEM)

Transmission electron microscopy (TEM) is an extremely useful characterization technique for epitaxial films. A high-energy collimated electron beam is passed through a thin specimen, and the resultant transmitted electrons are collected using magnetic lenses and then either displayed on a phosphor screen or imaged using a CCD. The wavelength of the energetic electrons is much smaller than the lattice spacing in the semiconductor crystal. Therefore, TEM is capable of measuring very small phenomena—much smaller than what can be measured in an optical microscope. Among the many things that it can measure are layer thicknesses from nanometers through tens of microns; film morphology; presence and density of extended defects such as dislocations, stacking faults, and anti-phase boundaries; strain state of films; and presence of precipitates.

Various types of interactions between the high-energy electrons and the sample can be observed. Inelastically scattered electrons can be discarded while transmitted electrons are used to create an image, similarly to how a transmission light microscope works. This image has what is known as mass-thickness contrast, because higher-Z or thicker regions will have higher inelastic scattering and will therefore appear darker. In addition, contrast can be obtained from elastically scattered, or diffracted, electrons [47]. To do this, the sample is tilted into an orientation that only allows for diffraction from one set of crystallographic planes, with Miller indices denoted $g = (hkl)$ (see Figure 2.6). Then, an image is collected using only the directly transmitted beam and

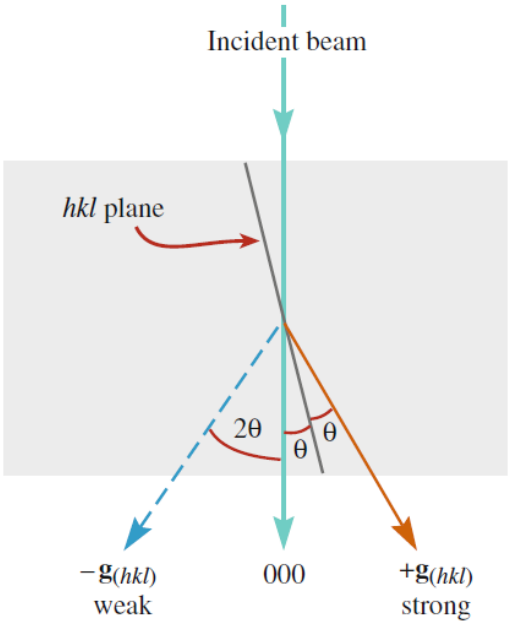

Figure 2.6: Two-beam diffraction condition, where sample is tilted to obtain a single strong diffraction spot, $g = (hkl)$ [47].

excluding the diffracted beam. This way, any perturbations in lattice spacing along the direction of g manifest as changes in contrast in the image. Images with g = (220) are useful for seeing dislocations and stacking faults in diamond cubic and zinc blende semiconductors. Images with g = (004) are useful for observing lattice strain in the [001] growth direction.

Samples can be prepared in either a cross section (XTEM) or plan-view (PVTEM) orientation. While PVTEM cannot determine which particular layer a defect is in with great accuracy, it is able to detect defects such as threading dislocations at a significantly lower density than XTEM. If no dislocations are visible near the surface in XTEM, that only implies a dislocation density less than ~10^8 cm^{-2} [10]. If no dislocations are visible in a given 5000x PVTEM image, that implies dislocation density less than ~5×10^6 cm^{-2}. However, by taking many PVTEM images from different areas of the sample, dislocation densities as low as ~5×10^5 cm^{-2} can be measured with sufficient accuracy.

TEM specimens were prepared from epitaxial samples by grinding/polishing the sample into a ~10 μm thick membrane by hand, followed by Ar-ion milling in a Fischione 1010 ion mill until the sample has reached electron transparency. This process is described in detail in Appendix A. Images were obtained using a JEOL 2011 microscope with a LaB_6 filament, an accelerating voltage of 200 kV, and a double-tilt holder.

2.4.3. Scanning electron microscopy (SEM)

For certain samples, particularly those in the GaAsP C doping study described in Chapter 4, cross-section scanning electron microscopy (SEM) was used to measure film thickness. A Zeiss Merlin High-Resolution SEM in secondary electron (SE) mode was used to image the samples. Samples were cleaved along a {110} plane and imaged directly in cross-section. In SE mode, changes in doping level and polarity can cause large contrast variations due to changes in ionization energy [48]. This makes it easy to measure the layer thickness of the p-doped GaAsP layers above n-doped buffers with minimal sample preparation.

2.4.4. Etch pit density (EPD) measurements

Etch pit density (EPD) is a valuable chemical technique for measuring threading dislocation density (ρ_{TD}) in crystalline materials, particularly in mismatched epitaxial films. First, the sample is dipped in an etchant which preferentially etches near threading dislocations due to their strained cores [49]. This causes a pit to form around the dislocation, which can then be viewed either by optical microscopy or by SEM.

We used this technique extensively to measure ρ_{TD} for $Si_{1-x}Ge_x$ films with $x \leq 0.5$. The samples were dipped in a solution of 8 g CrO_3, 200 ml HF, and 250 ml H_2O, known as the

Schimmel etch, for 30 s, then rinsed in DI H_2O [50], [51]. The resultant etch pits were then counted in a differential interference contrast (DIC) microscope.

We have not been able to reliably replicate any EPD results for any ternary III-V semiconductors, such as GaAsP or InGaP. Therefore, we rely on different techniques for measuring ρ_{TD} in these materials, such as PVTEM and EBIC (see Section 2.6.2).

2.4.5. Photoluminescence (PL)

The photonic emission of the samples was probed using photoluminescence (PL) measurements. In a PL measurement, light is focused on the sample, creating a large number of electron-hole pairs. The excess minority carriers then recombine, causing emission of photons at an energy close to the band gap (E_g) of the material. The emitted light is then collected and analyzed.

A 514.5 nm (green) laser with around 500 mW of power was used for the excitation. The laser was passed through a chopper and a focusing lens before hitting the sample at an angle of around 60°. The emitted light was collected and focused into a computer-controlled monochromator (adjustable diffraction grating). The selected wavelength was then incident on a Si photodetector. The signal from the photodetector was amplified with a transimpedance amplifier, and then input to a lock-in amplifier. The lock-in amplifier was configured to measure signal only with the same frequency as the chopper, thereby removing other sources of noise, such as the room lights. By varying the spacing of the diffraction grating, a full emission spectrum can be measured. The entire setup was controlled using LabVIEW.

These measurements are useful for several reasons. First, the wavelength of the emission can be used to calculate E_g. Due to carriers recombining with non-zero momentum, the emission

peak is generally at higher energy than E_g. To account for this, we measured the offset for a sample with known E_g (i.e. GaAs) and applied that same offset to samples with uncertain E_g (i.e. $GaAs_xP_{1-x}$).

The intensity of the PL signal can give qualitative information about the radiative recombination rate in the sample as compared to the rated of other non-radiative recombination processes. A stronger PL signal signifies a higher radiative recombination rate or a lower non-radiative recombination rate. Comparing PL signal between similar samples can give a quick sense of the comparative quality of the epitaxial structures.

2.4.6. Secondary ion mass spectrometry (SIMS)

Dynamic secondary ion mass spectrometry (SIMS) is a technique for measuring elemental concentrations in a sample as a function of depth. An area of the sample is sputtered at a calibrated rate. After leaving the surface, the sputtered (secondary) ions are collected and the mass to charge ratio is measured. The concentration of each element can thereby be measured for each depth location in the sample. Concentrations as low as 10^{15} cm^{-3} can be detected. Because there is a large matrix effect (i.e. a large difference in signal from element A in a matrix of B vs A in a matrix of C), it is important to have good calibration standards for SIMS.

SIMS is most useful for measuring the concentration of dopant atoms or contaminants such as O or C as a function of depth in epitaxial samples. It can also be used to measure the atomic fraction of components in alloys such as SiGe or GaAsP.

All SIMS measurements in this work were performed by Evans Analytical Group.

2.4.7. Hall effect measurements

Hall effect measurements were taken using the van der Pauw geometry. From the measurements made using this technique, resistivity, sheet carrier density, mobility, and carrier type (p or n) can be calculated [52]. The measurements are taken on approximately 1 cm × 1 cm square pieces, cleaved by hand. Ohmic contacts are made to the four corners of the sample by melting on small drops of In metal on a hotplate set to 250 °C. It is important that the contacts be small in diameter relative to the width of the sample (i.e. less than 1/10 the sample width). Given that in our samples the contact size is less than 1 mm, we estimate an upper bound for the error in R_S of 1% and in n_s and μ of 5% [53]. A magnetic field of 0.412 T is provided by a permanent magnet.

2.5. Device fabrication

HBTs were fabricated from the epitaxial structures using standard contact lithography techniques, wet etching, and metal lift-off. All of the HBTs discussed in this work have an annular geometry where the emitter and base mesas are concentric circles. The emitter mesa diameters, which define the size of the emitter-base junction, range from 15 μm to 240 μm. Depictions of the major steps are shown in Figure 2.7.

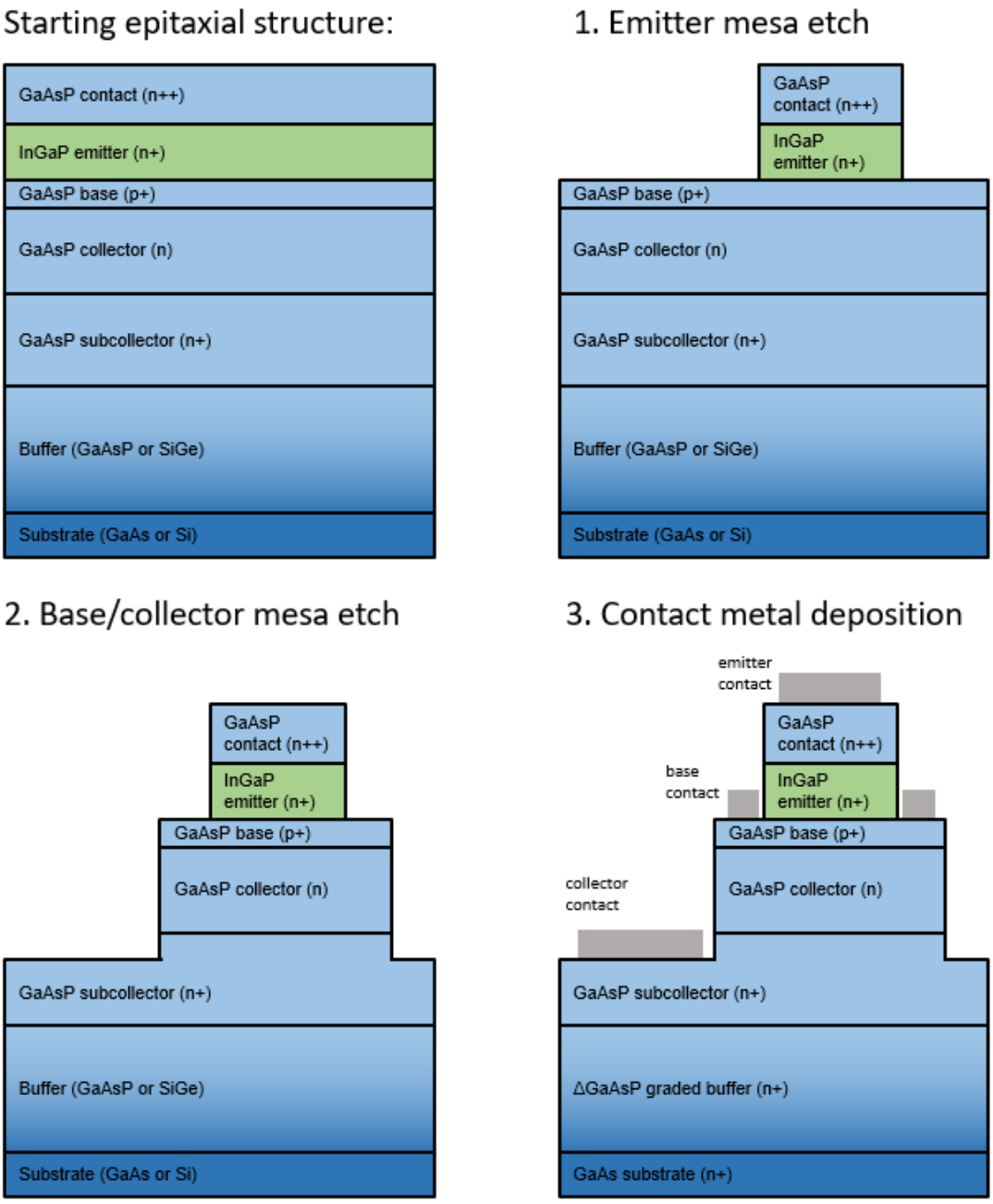

Figure 2.7: Cross-sectional schematic of HBT device fabrication steps.

First, the emitter mesa was etched using a two-step process. The GaAsP contact layer was removed using H_2SO_4:H_2O_2:H_2O = 1:1:10; then, the InGaP emitter was removed using HCl:H_3PO_4 = 1:1. The H_2SO_4 etchant has been shown to not etch InGaP and the HCl etchant not to etch GaAs [54], [55]. Even with x = 0.825, the HCl etchant is still highly selective against $GaAs_xP_{1-x}$ and

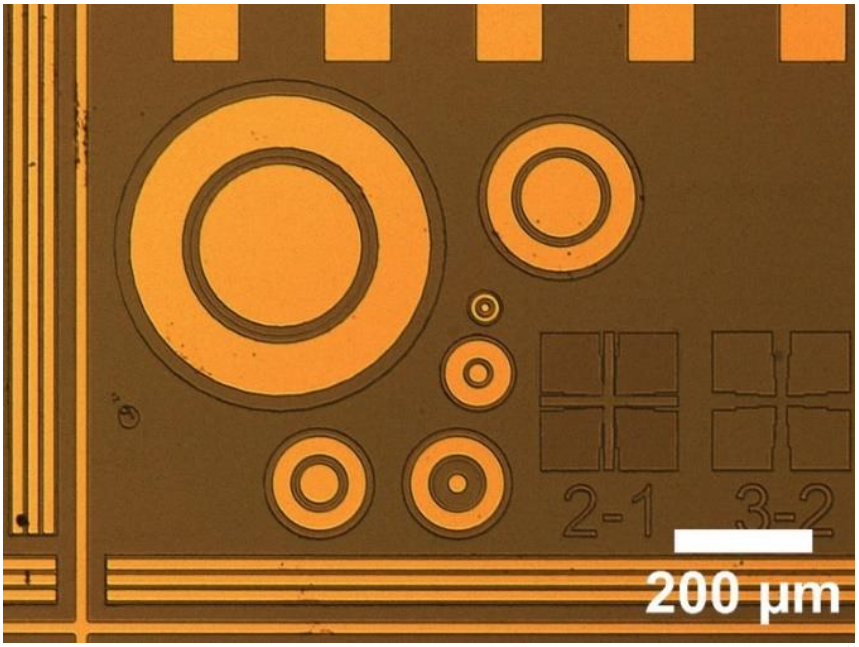

Figure 2.8: Image of annular HBTs of various sizes after fabrication is complete.

therefore does not etch into the base region. Next, the base/collector mesa was then patterned using the H_2SO_4 etchant, timed to stop in the subcollector layer.

After patterning, the mesa sidewalls were passivated. The samples were dipped in H_2SO_4:H_2O = 1:10 for 60 s to remove any native oxide and then a 10 nm Al_2O_3 passivation layer was immediately deposited by atomic layer deposition (ALD) at 200 °C.

After developing the photoresist, but before depositing the metal, the Al_2O_3 passivation was removed with 7:1 buffered oxide etch in the areas underneath the contacts. Non-alloyed emitter, base, and collector contacts were then deposited simultaneously by e-beam evaporation of Ti/Pt/Au (5 nm/40 nm/120 nm; base pressure of 1×10^{-6} Torr) and lifted off in acetone.

All device fabrication was completed at the Microsystems Technology Laboratory (MTL) at MIT. A detailed list of the process steps is given in Appendix B. An image of the finished devices is shown in Figure 2.8.

2.6. Device characterization

2.6.1. DC measurements

Devices were electrically tested at 300 K using a Keysight B1500A semiconductor parameter analyzer. Several types of measurements were taken. Output characteristics were measured by sweeping the collector-emitter voltage (V_{CE}) while maintaining constant base current (I_B), and measuring collector current (I_C). These generally do not yield much information on material quality, and will therefore not be discussed further in this thesis. Gummel plots were measured by holding V_{BC} constant while varying V_{BE}, and measuring I_B and I_C. Breakdown voltage was also measured, the details of which are in Section 5.4.3.

2.6.2. Electron-beam-induced current (EBIC)

Electron-beam-induced current (EBIC) is a technique for measuring defects in semiconductors by probing local minority carrier recombination rates using an SEM. A focused electron beam is rastered across the sample just as in normal SE mode. However, rather than generating an image by collecting secondary electrons, the image is generated by measuring the current collected by the built-in electric field of a diode or other heterostructure. Figure 2.9 shows a schematic of the setup used in this work to measure HBTs. The high-energy electron beam creates many electron-hole pairs, which are then collected by the HBT, which has an applied bias of V_A. This current is amplified and transformed into an image. In areas near defects such as dislocations, the recombination rate is increased, and therefore the collected current is smaller. This creates a dark area in the image that is recognizable as a threading dislocation if it is a dot and a misfit dislocation if it is a line.

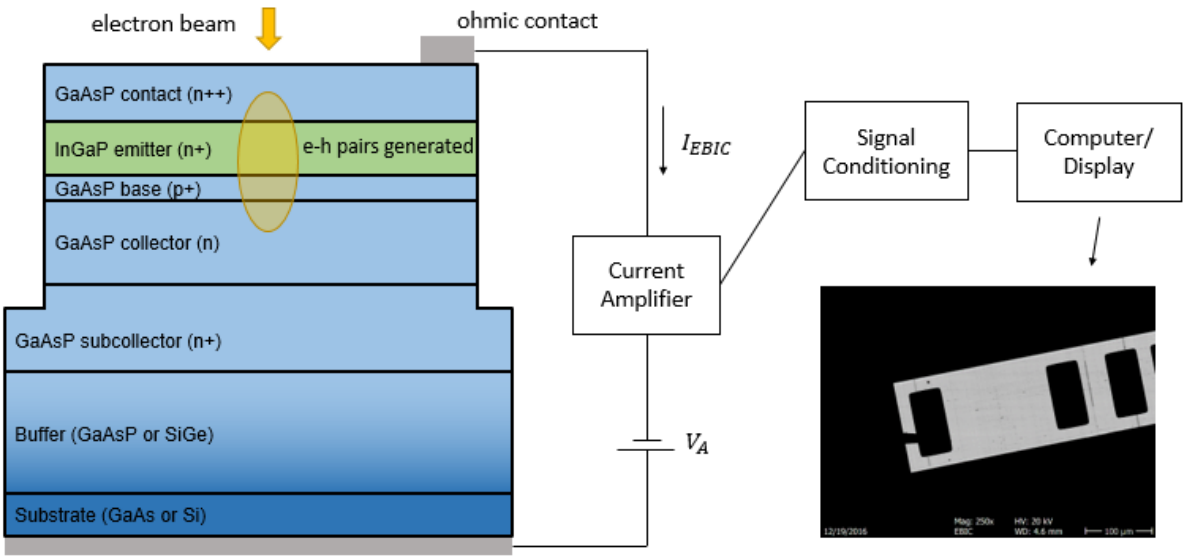

Figure 2.9: Schematic of the EBIC setup for measuring HBTs.

EBIC measurements were taken using an FEI Helios Nanolab 600 dual-beam SEM/FIB. This instrument is particularly useful because of its Omniprobe, a computer-controlled micromanipulator which can be used to probe multiple different devices without breaking vacuum in the SEM. An accelerating voltage of 15 kV and a beam current of 0.34 nA were used. The applied voltage V_A was set to between 0 V and 1 V. The gain, current offset, and contrast parameters in the EBIC acquisition software were adjusted for each image to create an image with proper brightness and contrast, with clearly visible defects.

EBIC is a useful technique for measuring defects for several reasons. First, it detects defects by probing the increase in minority carrier recombination rate in the surrounding area. This is the same effect that makes the defects detrimental to device performance in the first place. In addition, large areas can be measured quickly, allowing for low defect densities to be observed. Finally, as long as HBTs are being fabricated anyway to measure device performance, there is no additional time-consuming sample prep for EBIC as there would be for PVTEM.

Chapter 3: Comparison of transistor architectures for GaAsP/InGaP integration with Si

3.1. Heterojunction bipolar transistors (HBTs)

3.1.1. Device structure

A heterojunction bipolar transistor (HBT) is a variation of a bipolar junction transistor (BJT) where the emitter is made from a different material than the base and collector. A simplified

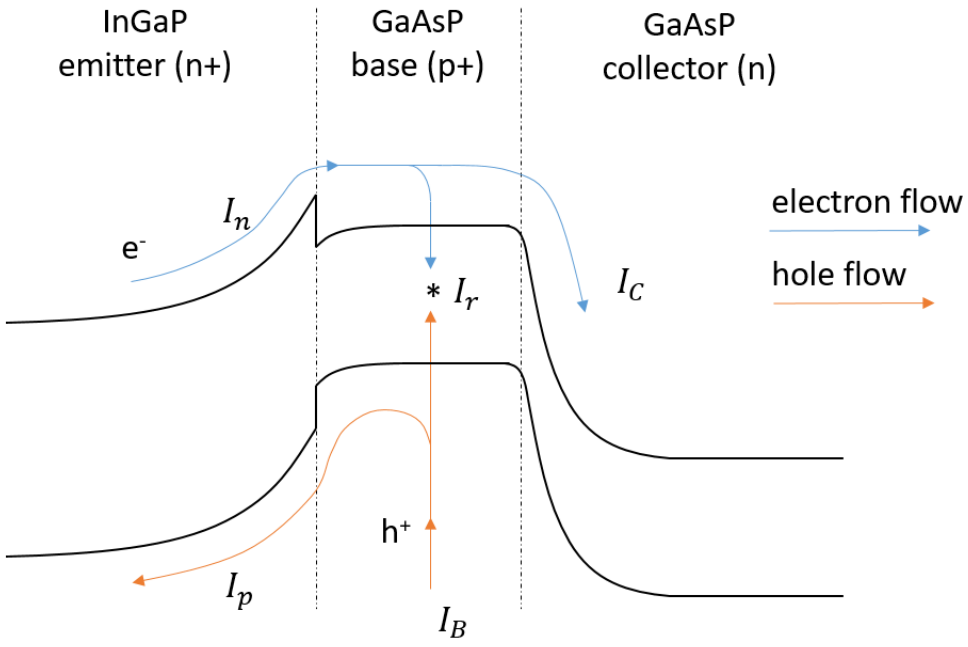

Figure 3.1: Band diagram of a GaAsP/InGaP HBT in FAR with important current components labeled. Adapted from Yuan, 1999 [57].

cross-section of the GaAsP/InGaP HBTs studied in this thesis is shown in Figure 1.8b. A band diagram of this same structure is illustrated in Figure 3.1.

Here, we will discuss the basic operation of an HBT in the forward-active regime (FAR). In FAR, the emitter-base junction is in forward bias and the base-collector junction is in reverse bias. A current sent into the base is amplified and output from the collector (while the emitter is grounded). During this process, electrons are injected from the emitter into the base region (I_n). Most of these electrons diffuse across the base region and into the collector region where they are extracted as collector current (I_C). A small minority recombine with majority holes in the base (I_r), contributing to base current (I_B). Another contribution to the base current is reverse injection of holes from the base into the emitter (I_p). The valence-band offset at the emitter-base interface of an HBT acts as a barrier to I_p that is not present in a BJT. This means that unlike in a BJT, an HBT can have a very high base doping. Benefits of this include lower base contact resistance, lower intrinsic and extrinsic base series resistances, and a thinner base, all of which are good for high frequency performance [56].

3.1.2. Critical DC parameters

The equations below govern the relationships between the different current components in Figure 3.1:

$$I_E = I_n + I_p \tag{3.1}$$

$$I_B = I_p + I_r \tag{3.2}$$

$$I_C = I_n - I_r \tag{3.3}$$

The common emitter current gain (β) is defined as:

$$\beta = \frac{I_C}{I_B} \tag{3.4}$$

The current gain is an important parameter to maximize and should be on the order of 100 for most device applications. In order to maximize β for a given I_C, we must attempt to minimize I_B. In an HBT, I_p is very small due to the valence band offset, and is therefore generally negligible compared to I_r. For this to be true, the valence band offset must be at least several kT. I_r has several components [56], [57]:

1. Recombination in the bulk quasi-neutral region of the base. This recombination can be due to the heavy doping in the base, or other defects in the base region such as threading dislocations. This is accompanied by a base current ideality factor close to 1.
2. Recombination in the space-charge region of the emitter-base junction (SCR). This is usually due to contamination, point defects, or other defects present at the emitter-base interface. This is accompanied by a base current ideality factor close to 2.
3. Recombination at the surface (edge) of the emitter-base mesa. If this is the primary recombination mechanism, I_B will scale with device perimeter rather than with device area. This can be reduced or eliminated by proper sidewall passivation of the device.

We will seek to minimize all of these components during HBT fabrication.

A second important characteristic of an HBT is the breakdown voltage. This is the reverse-bias voltage across the base-collector junction at which avalanche breakdown will occur. For circuit designers, this determines the maximum voltage that the device should be subjected to in normal operation. The breakdown voltage is dependent on materials properties, most importantly, the band gap of the base/collector regions. A higher band gap results in a higher breakdown voltage. This is one of the main benefits of using GaAsP as the base/collector material instead of GaAs, as discussed in Section 1.4.

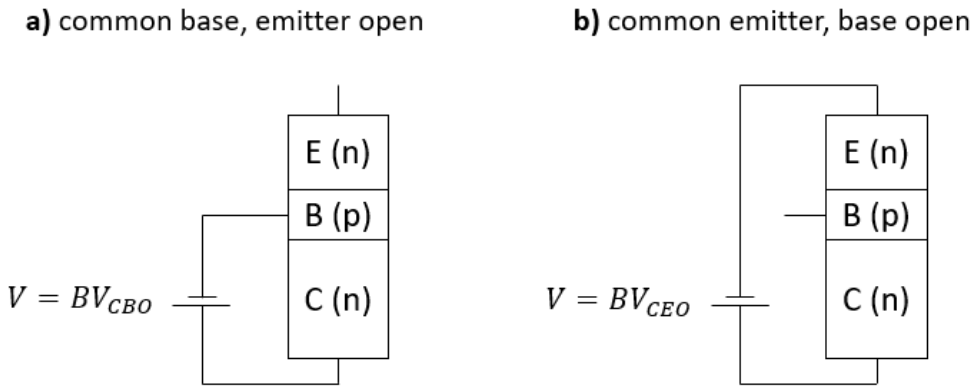

Figure 3.2: Two important cases HBT breakdown: (a) BV_{CBO} and (b) BV_{CEO}.

The breakdown voltage of an HBT varies depending on how the emitter and base are biased with respect to each other. This results in a continuous range of V_{BC} at which breakdown can occur, depending on the value of V_{BE}. This behavior can be characterized satisfactorily by two distinct corner cases, which describe the bounds of the entire breakdown voltage envelope. The first, known as "common base, emitter open" (notated as BV_{CBO}), is the collector voltage at breakdown where the base is grounded and the emitter is left floating (see Figure 3.2a). This is equivalent to the breakdown of a p-n diode made up of the p-type base and the n-type emitter. The second case, known as "common emitter, base open" (notated as BV_{CEO}), is the collector voltage at breakdown where the emitter is grounded and the base is left floating. At first glance, it may appear as if this is equivalent to BV_{CBO} because the base-collector diode is in reverse bias and supports the entire voltage difference. However, any leakage current that passes through the base-collector junction is then amplified by the emitter-base junction, causing β times that leakage current of to be injected into the base. This amplification of the leakage current causes breakdown to occur more readily in the common emitter case than in the common base case. Therefore, in general, $BV_{CBO} > BV_{CEO}$. Most device operation occurs somewhere in between these two cases [58].

For HBTs in this work, we focused mostly on measuring and modeling BV_{CBO}. This is because it's easier to model and it depends more directly on materials parameters and less on β. With knowledge of BV_{CBO} and β, BV_{CEO} can be estimated.

3.2. High-electron-mobility transistors (HEMTs)

GaAsP/InGaP high-electron-mobility transistors (HEMTs) were also considered alongside GaAsP/InGaP HBTs during this work as candidates for integration onto Si substrates. HEMTs are a type of heterojunction-based field-effect transistor that takes advantage of a technique called modulation doping in order to achieve very high carrier mobilities. In a modulation-doped heterostructure, a low-bandgap channel is left undoped and grown adjacent to a heavily doped, high-bandgap barrier layer. The high doping in the barrier causes band bending, which allows a

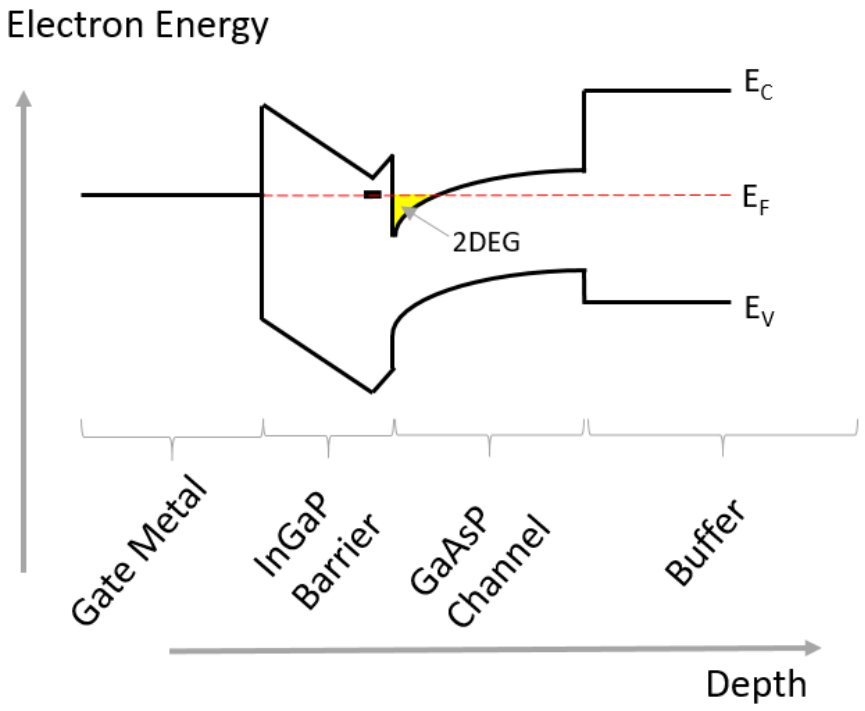

Figure 3.3: Band diagram of a GaAsP/InGaP HEMT. Modulation doping in the InGaP barrier causes a 2DEG to form in the GaAsP channel.

2D electron gas (2DEG) to form in the channel. This way, the free electrons associated with the heavy doping in the barrier are instead located in the undoped channel, allowing them to have a much higher mobility. A band diagram illustrating this concept for a proposed GaAsP/InGaP HEMT is shown in Figure 3.3.

HEMTs are majority carrier devices, and are therefore less sensitive to threading dislocations than minority carrier devices such as HBTs. A high carrier concentration in the 2DEG shields the charged dislocation cores, further reducing the effect of dislocations on electron mobility [14]. This makes them strong candidates for integration onto Si substrates. However,

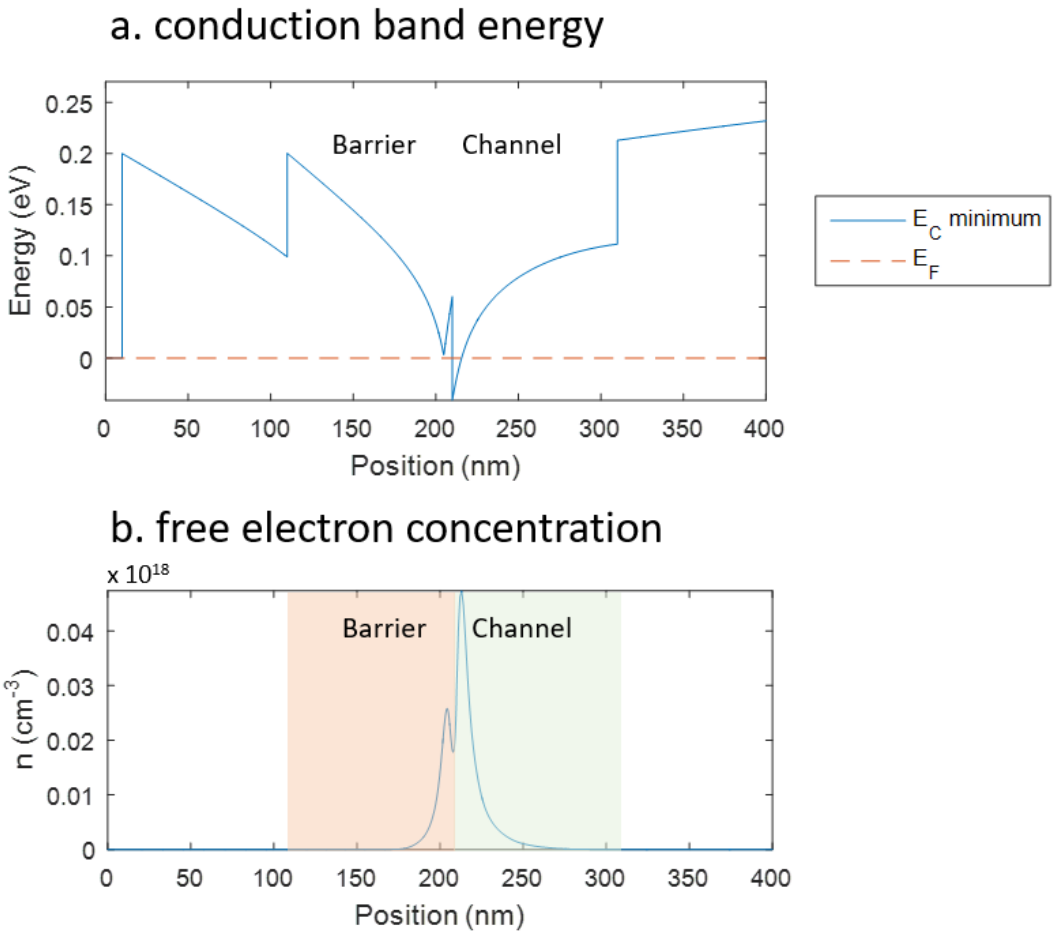

Figure 3.4: Example Nextnano calculations for a GaAsP/InGaP HEMT structure as a function of depth: (a) conduction band energy and (b) free electron concentration.

before embarking on GaAsP/InGaP HEMT device fabrication, we decided to simulate parameters of GaAsP/InGaP HEMTs as a function of GaAsP composition to predict its effect on device performance.

The $n_s\mu$ product (channel conductance) is an important figure of merit for HEMT performance. Low channel conductance is critical for high transconductance, which in turn is an important factor for good AC performance (high cut-off frequency) [1]. We must therefore understand how this product varies in order to predict HEMT performance as a function of GaAsP channel composition.

Sheet carrier concentration (n_s) is approximately equal to the delta doping concentration in the barrier layer. However, electron mobility (μ) is a weighted average of the electrons conducting through the high-mobility channel and the low-mobility barrier. We expect the mobility of the InGaP barrier to be around 10 times lower than the GaAsP channel [59]. We can estimate the ratio of electrons in the channel to electrons in the barrier ($n_{channel}/n_{barrier}$) by modeling the GaAsP/InGaP HEMT band structure as a function of composition. A 1D simulation (Nextnano) of electron energy vs device depth is shown in Figure 3.4a for an example GaAsP/InGaP HEMT. Next, electron concentration was calculated as a function of depth (Figure 3.4b). We then integrated this curve in both the channel and barrier regions to find the ratio of electrons in the channel to the ratio of electrons in the barrier. This entire procedure was repeated for various lattice constants/compositions and the results plotted in Figure 3.5.

As the As content of the GaAsP channel decreases, $n_{channel}/n_{barrier}$ also decreases. This is a problem for GaAsP/InGaP HEMTs, because this means that the $n_s\mu$ product will be significantly lower than that of a GaAs/InGaP HEMT. The exact value of $n_{channel}/n_{barrier}$ depends strongly on the conduction band offset between the GaAsP channel and the InGaP emitter.

As this value is only approximately known as a function of GaAsP composition [60], these values of $n_{channel}/n_{barrier}$ may not be exact. However, the trend is expected to hold. Because of this trend, we did not allocate further resources to study GaAsP/InGaP HEMTs, but instead decided to focus on HBTs, which have less stringent restrictions on band alignment.

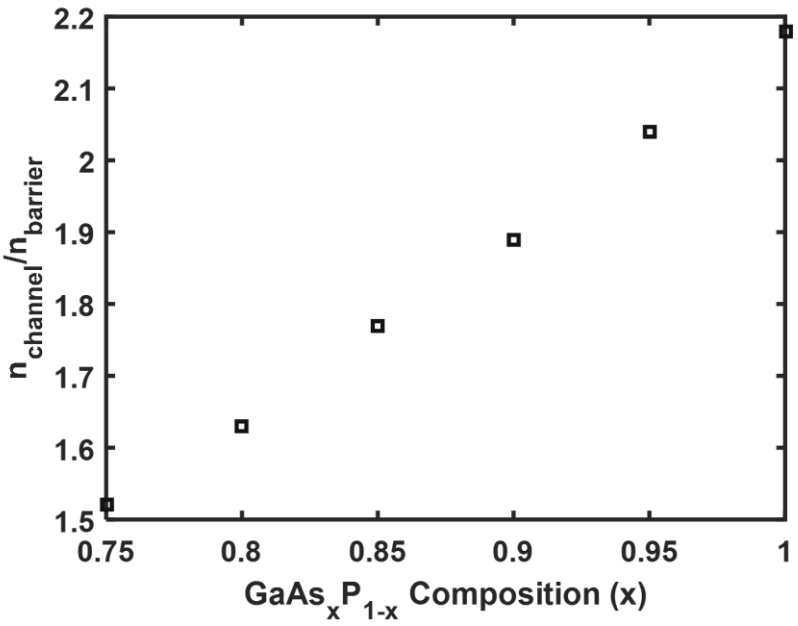

Figure 3.5: Ratio of free electrons in the channel to in the barrier of GaAsP/InGaP HEMTs vs composition (simulated using Nextnano).

Chapter 4: Epitaxial challenges for GaAsP/InGaP HBTs grown on GaAs and Si substrates

4.1. GaAsP tensile graded buffers on GaAs substrates

In order to study the effect of $GaAs_xP_{1-x}$ composition on HBT performance while minimizing the effect of defects, we grew GaAsP HBT structures on GaAs substrates via a tensile $GaAs_xP_{1-x}$ graded buffer (see Chapter 5). This dramatically reduces the defect density compared with GaAsP devices grown on Si due to a greatly reduced lattice mismatch as well as the lack of a III-V on group IV interface. The tensile GaAsP graded buffers require special consideration compared with compressive SiGe graded buffers. Compressive strain is generally chosen for metamorphic structures because sufficient tensile strain will cause 60° dislocations to dissociate into two partial dislocations with a stacking fault in between [61], [62]. These stacking faults act as barriers to dislocation glide, causing a rapid increase in threading defect density. Therefore, for tensile graded buffers, a much slower grade rate must be used than for compressive graded buffers to obtain the same threading dislocation density (0.2% strain/μm compared to 1–2% strain/μm) [63]. In addition, at high strain rates, so-called "faceted trenches" can form in tensile GaAsP graded buffers rather than deformation by plastic deformation [63], [64]. These crack-like trenches preclude device fabrication in their vicinity. A grade rate of 0.2% strain/μm is sufficient to reduce the density of faceted trenches to acceptable levels.

4.2. GaAsP "virtual substrates" on Si and GaAsP regrowth

For GaAsP/InGaP HBTs grown on Si substrates, our main approach was to integrate via a SiGe virtual substrate. Because of the many steps involved in growing GaAsP on Si using this method, we first made GaAsP "virtual substrates" on a 150 mm Si wafers, which we then cleaved into ~2 cm square pieces for growth of various HBT structures. That way, many different HBT structures could be grown on one GaAsP virtual substrate.

Figure 4.1 shows the process flow for the creation of GaAsP virtual substrates used in this work. This process was developed by Sharma, et al. from our research group [25], [65]–[67]. The GaAsP virtual substrates were grown starting with 150 mm (100) Si wafers with a 6° offcut towards the nearest [111] direction. This 6° offcut is important for growth of single-domain III-V material in later steps. A SiGe graded buffer was grown via UHVCVD to $SiGe_{0.5}$ and a thick relaxed $SiGe_{0.5}$ layer grown on top. At this point, the wafer has developed significant surface

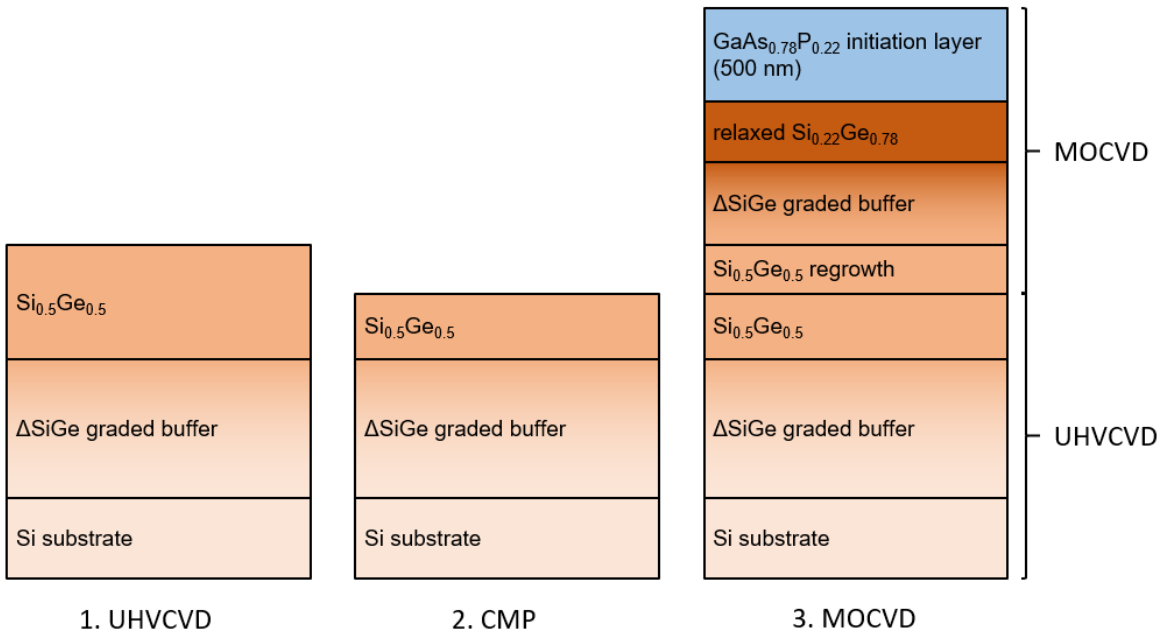

Figure 4.1: Schematic of the process flow for GaAsP/Si integration.

roughness, known as “crosshatch,” due to local differences in growth rate due to the strain fields surrounding misfit dislocations [68]. If grading were continued past $SiGe_{0.5}$, this surface roughness might become too large and cause the nucleation of further threading dislocations [26]. To avoid this, at this point the wafers are planarized via chemical mechanical polishing (CMP). CMP was completed externally by Entrepix, Inc. This process removes some of the $SiGe_{0.5}$ cap layer but reduces the surface roughness dramatically. Then, the wafers are cleaned and loaded into the MOCVD reactor. Here, $SiGe_{0.5}$ is regrown on the existing surface, and then graded to $SiGe_{0.78}$ and

ended with a relaxed $SiGe_{0.78}$ cap. Finally, a $GaAs_{0.78}P$, which is lattice-matched to the $SiGe_{0.78}$, is initiated on the surface. This initiation step is highly sensitive to residual strain [22]. A highly-detailed explanation of this entire process is available in Chapter 3 of Milakovich, 2015 [67].

The entire SiGe graded buffer of a GaAsP virtual substrate is shown in a series of interlaced cross-section TEM images in Figure 4.2. This complexity of this structure and the sensitivity of its different interfaces necessitates tight process control. Composition of the various layers were calibrated with XRD or SIMS before growth of the entire structure. The ability to grow individual

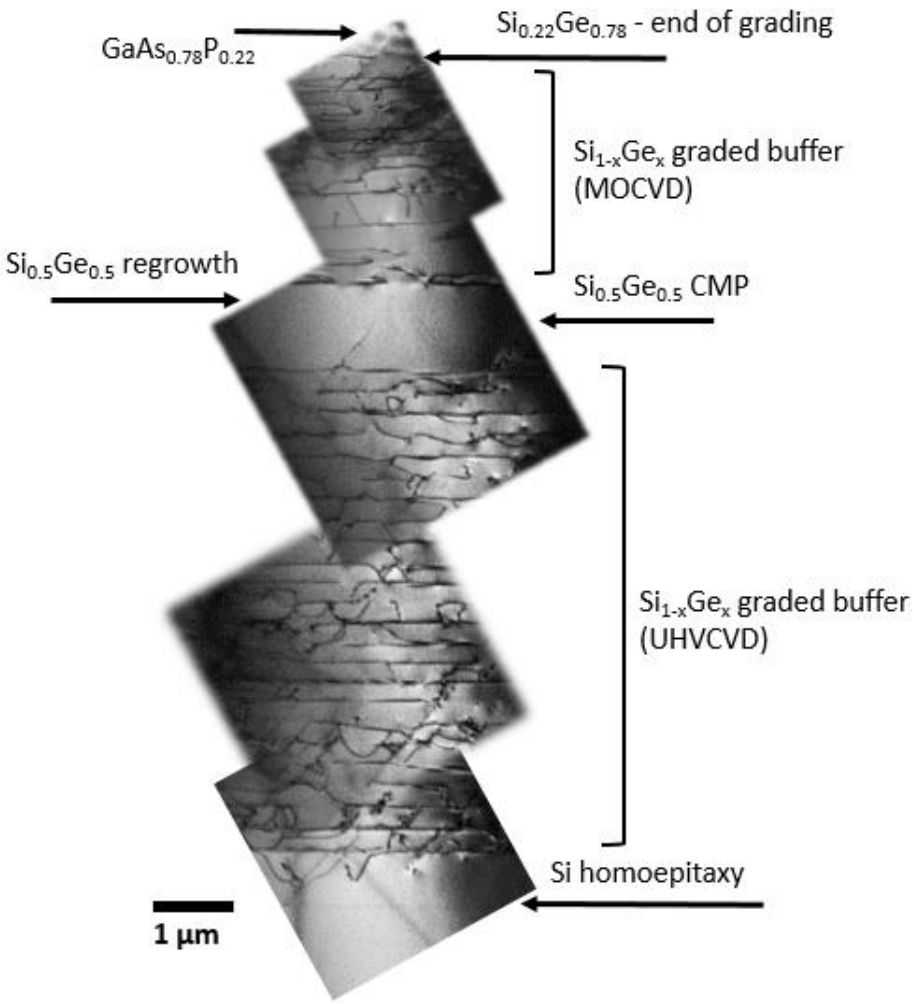

Figure 4.2: Cross-section TEM of SiGe graded buffers from GaAsP virtual substrate structure. g = (220) and Acc V. = 200 kV. (Courtesy of Tim Milakovich.)

HBT structures on pieces of the GaAsP virtual substrate was highly beneficial because it avoided having to grow the thick graded buffers each time.

A plan-view TEM image of the GaAsP virtual substrate containing one threading dislocation is shown in Figure 4.3. Many such images were taken from random areas of the sample to count enough dislocations to determine ρ_{TD}, which was $(3.7 \pm 0.7) \times 10^6$ cm^{-2}. The error in this value was determined using Poisson statistics.

Immediately before regrowth on the pieces of GaAsP virtual substrate, they were rinsed sequentially in acetone, methanol, isopropanol, and then dried with N_2. The final $GaAs_xP_{1-x}$ composition (x = 0.825) was reached via a short compressive GaAsP graded buffer with a grade rate of 1% strain/um. Based on EBIC data shown in Chapter 6, this does not significantly increase ρ_{TD}.

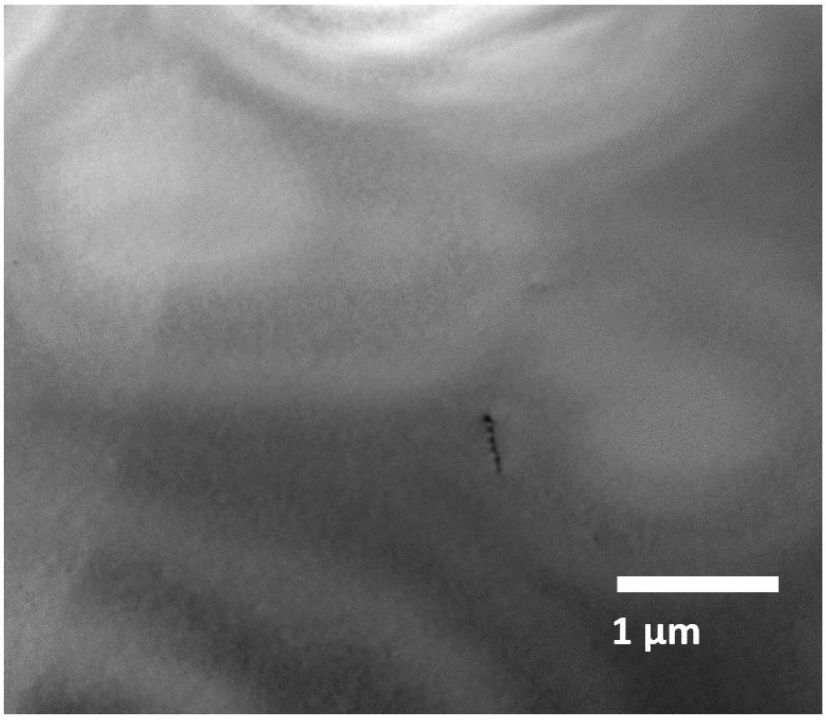

Figure 4.3: Plan-view TEM image of GaAsP virtual substrate showing a threading dislocation. g = (220) and Acc. V. = 200 kV.

4.3. Heavy carbon doping of GaAsP

Heavy p-type doping of GaAsP in excess of 1×10^{19} cm^{-3} is required for the base regions of HBTs. This type of doping is also useful in other device elements, such as Ohmic contact layers and tunnel junctions. Carbon has an advantage over group II p-type dopants (e.g. Zn) in its lower diffusivity [69] and higher solid solubility in GaAs [70]. There are several ways of doping with C in MOCVD. Various hydrocarbon sources have been used, including the methyl group from TMGa, but these generally require low growth temperatures and/or low V/III ratios and therefore result in a high density of deep-level traps and bad surface morphology [71], [72].

$CBrCl_3$ and other halomethanes allow higher growth temperatures and V/III ratios so high-quality C-doped films can be grown [73]. However, they create reactive byproducts which can have two adverse effects on $GaAs_xP_{1-x}$ growth: substantial growth rate reduction [74] and composition shift (change in x) [75]. Control over growth rate and composition is critical for device fabrication; therefore, it is important that we are able to understand the effect of C doping by $CBrCl_3$ on these parameters. In this paper, we report the thicknesses, compositions, and electrical properties of various C-doped GaAsP films.

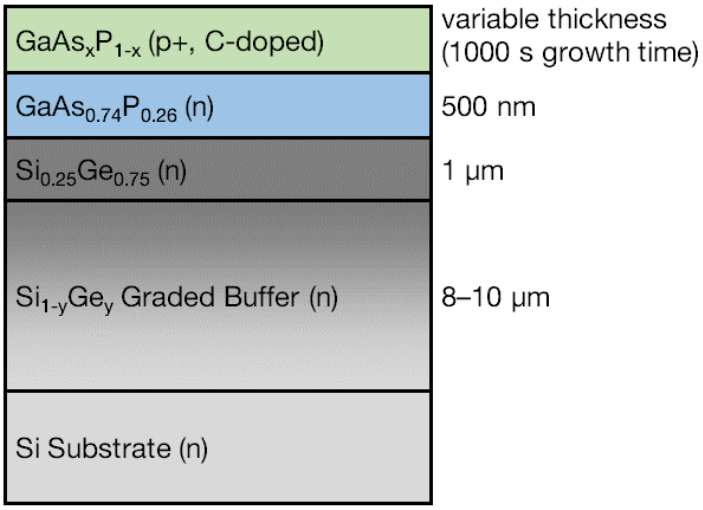

Figure 4.4: Cross-section schematic of the epitaxial samples for GaAsP C-doping experiments.

4.3.1. Details of experimental samples

Samples were grown with a structure shown in Figure 4.4. Si substrates were used for all GaAsP samples because tensile GaAsP films grown on GaAs substrates had a tendency to crack, yielding inaccurate Hall effect measurements. N-type (100) Si substrates with a 6° offcut were used. The offcut is necessary to obtain single-domain GaAsP [76]. $Si_{0.5}Ge_{0.5}$ was grown via a $Si_{1-y}Ge_y$ graded buffer using a vertical-tube UHVCVD reactor. The wafers were then chemical-mechanically polished to remove surface roughness [26]. Next, in the MOCVD reactor, a $Si_{1-y}Ge_y$ graded buffer was grown from y = 0.5 to y = 0.75, upon which an n-type $GaAs_{0.74}P_{0.26}$ film was initiated at 725 °C and grown to 500 nm. Then, under a mixed AsH_3 and PH_3 overpressure, the growth temperature was lowered to the final growth temperature (600 °C or 650 °C). A C-doped GaAsP film was then deposited for 1000 s with a TMGa flow of 132 µmol/min (50 sccm, T_{source} = 5 °C, P_{source} = 1500 Torr) and a V/III ratio of about 100. The C-doped GaAs and GaP films were grown directly on their respective n-type substrates rather than on Si.

Table 4.1 shows a summary of the process conditions for all of the C-doped GaAsP films, grown at a variety of temperatures, $CBrCl_3$ flow rates, and AsH_3 fractions. AsH_3 fraction, the chosen metric of As precursor in the gas phase, is defined as:

$$AsH_3 \text{ Fraction} = \frac{P_{AsH_3}}{P_{AsH_3} + P_{PH_3}}, \tag{4.1}$$

where P_{AsH_3} and P_{PH_3} are the input partial pressures of AsH_3 and PH_3, respectively. A $CBrCl_3$ flow rate of 59 μmol/min corresponds with a carrier flow of 40 sccm, T_{source} of 26 °C, and P_{source} of 1200 Torr.

Sample A1 was grown at a temperature of 650 °C, while all other samples were grown at 600 °C. Samples A2–A6 vary AsH_3 fraction, and therefore the solid $GaAs_xP_{1-x}$ composition, while keeping a constant $CBrCl_3$ flow. Samples A7, A3, and A8 vary the $CBrCl_3$ flow with a constant AsH_3 fraction. Sample A9 has multiple GaAsP layers with varying AsH_3 fractions and no $CBrCl_3$ flow. The AsH_3 fractions of 0.18–0.37 were chosen because at a growth temperature of 600 °C, they result in GaAsP compositions which span most of the direct band gap range [77]. The V/III ratio was kept constant while varying AsH_3 fraction.

Table 4.1: List of GaAsP samples grown for C doping study.

Sample ID	Temperature (°C)	$CBrCl_3$ Flow (μmol/min)	AsH_3 Fraction
A1	650	118	0.53
A2	600	59	0
A3	600	59	0.18
A4	600	59	0.27
A5	600	59	0.37
A6	600	59	1
A7	600	15	0.18
A8	600	118	0.18
A9	600	0	0.18, 0.27, 0.37

4.3.2. Effect of growth temperature on C-doped GaAsP

Sample A1, grown at 650 °C, yielded a hole concentration of 9×10^{17} cm^{-3}. This is too low for the applications discussed on page 69 . Growth temperatures of 650 °C and higher are preferred for MOCVD growth of GaAsP. This is due to increased concentrations of impurities at lower temperatures as well as decreased PH_3 cracking efficiency [78], [79]. However, it is known that a decreased growth temperature of 600 °C or lower produces the highest hole concentrations in C-doped GaAs [73], [74]. Sample A8, grown with the same $CBrCl_3$ flow rate as Sample A1 but at 600 °C, yielded a hole concentration of 2×10^{19} cm^{-3}. This confirms that the active C doping of GaAsP also increases with decreasing growth temperature. The rest of this study focuses on a growth temperature of 600 °C because growth temperatures less than 600 °C are undesirable for the reasons mentioned above.

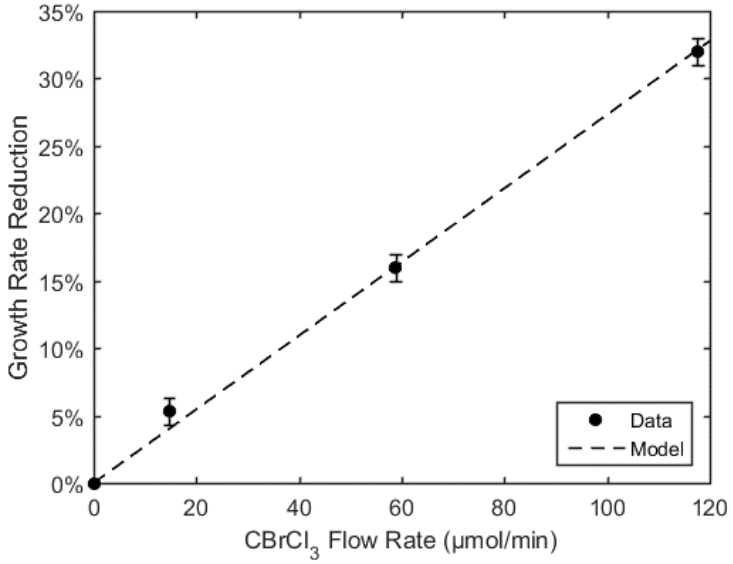

Figure 4.5: Reduction in GaAsP growth rate as a function of $CBrCl_3$ flow rate.

4.3.3. GaAsP growth rate reduction from $CBrCl_3$ precursor

The growth rate of GaAsP was decreased by the introduction of $CBrCl_3$. Figure 4.5 shows the growth rate reduction for a single AsH_3 fraction (0.18) as a function of $CBrCl_3$ flow. Lee et al. model the growth rate reduction of GaAs by CCl_4 as:

$$r = k_0[\mathrm{CX_4}][\mathrm{V/III}]^{-0.5} \exp\left(-\frac{E_a}{kT}\right), \qquad 4.2$$

where $[\mathrm{CX_4}]$ is the concentration of C precursor (CCl_4 in Lee et al.), E_a is the activation energy of formation of GaCl from Ga and Cl, and k_0 is a constant dependent on other factors [74]. The proportional dependence on C precursor concentration agrees with our data for GaAsP, suggesting that the growth rate reduction is caused by reaction of either Cl or Br byproducts with Ga or Ga precursor.

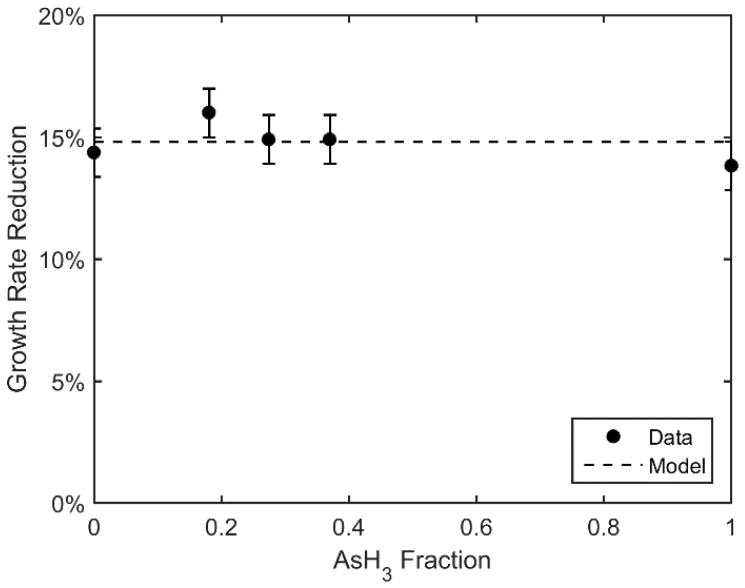

Figure 4.6: Reduction in GaAsP growth rate as a function of AsH_3 fraction.

Growth rate reduction does not change appreciably with AsH_3 fraction at a constant $CBrCl_3$ flow rate, as shown in Figure 4.6. This is further evidence that the growth rate reduction is limited by the reaction of Br or Cl byproducts with Ga and that group V species are not involved.

4.3.4. GaAsP composition shift due to $CBrCl_3$

Figure 4.7a shows the $GaAs_xP_{1-x}$ composition (x) with a constant AsH_3 fraction of 0.2 and a varying $CBrCl_3$ flow. The introduction of $CBrCl_3$ increases the As fraction in the solid phase (x). This trend was also observed by Tateno et al. for InGaAsP doped by CBr_4 [75]. The shift in x is approximately the same for all $CBrCl_3$ flows tested (15–118 µmol/min).

The incorporation of high amounts of C is expected to have an effect on the GaAsP lattice constant due to Vegard's Law [80]. However, this shift would be opposite in sign to the shift observed here, that is, towards smaller rather than larger lattice constants. In addition, the

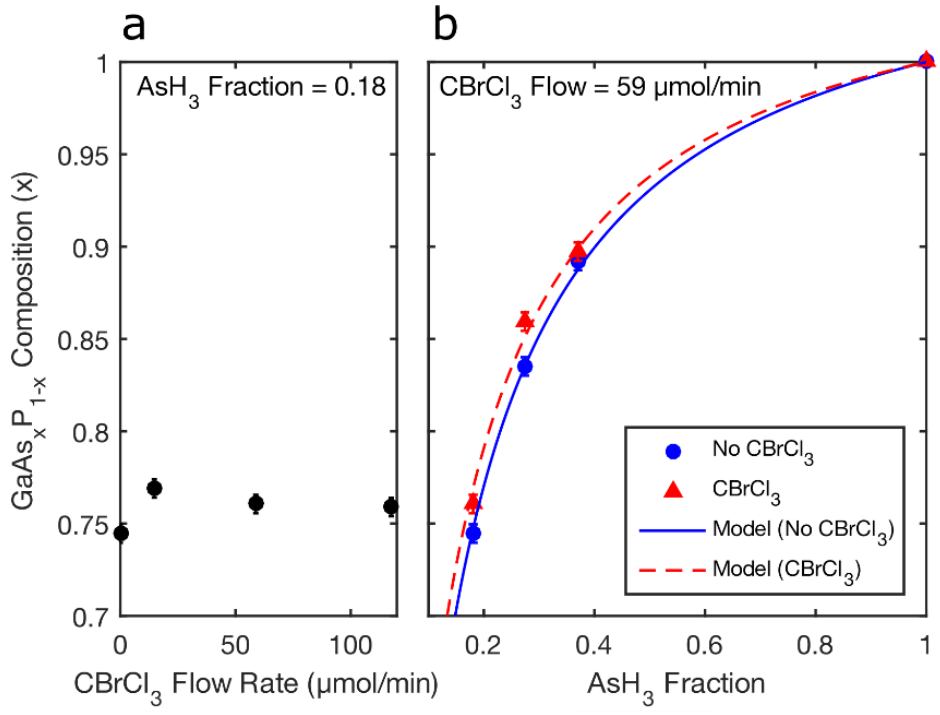

Figure 4.7: GaAsP composition (As fraction) versus (a) $CBrCl_3$ flow rate and (b) AsH_3 fraction.

magnitude of the shift due to C incorporation is expected to be a small fraction (< 10%) of the observed shift, so it is ignored here.

$GaAs_xP_{1-x}$ composition (x) versus AsH_3 fraction is shown in Figure 4.7b for both no $CBrCl_3$ and for a $CBrCl_3$ flow of 59 µmol/min. Smeets et al. model the relationship between the input partial pressures of AsH_3 and PH_3 and x as:

$$\frac{1-x}{x} = C\frac{P_{PH_3}}{P_{AsH_3}}, \tag{4.3}$$

where C is a fitting constant with an Arrhenius dependence and an activation energy close to that of the PH_3 cracking energy [44]. For the case of no $CBrCl_3$, this equation fits the data well with C equal to 0.075. For a $CBrCl_3$ flow of 59 µmol/min, the data is well-fit with an adjusted C of 0.066. Equation 4.3 with both values for C is plotted in Figure 4.7b.

We have two hypotheses as to what could cause the increase in x in the films grown with $CBrCl_3$ present. $CBrCl_3$ could be slowing the PH_3 cracking or otherwise consuming a certain

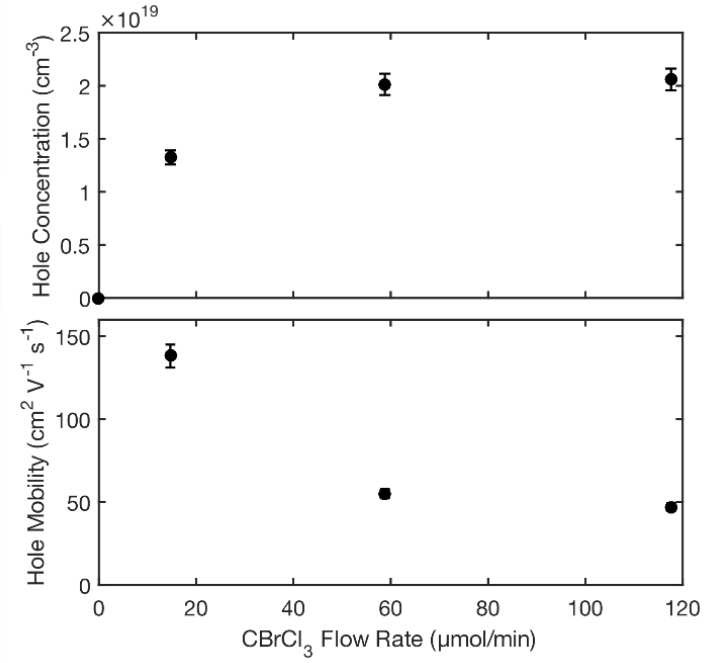

Figure 4.8: Hole concentration and mobility versus $CBrCl_3$ flow rate with a fixed AsH_3 fraction (0.18).

fraction of PH_3, reducing the amount of elemental P available to enter the film. This could be due to formation of PCl_3, PBr_3, or similar species. Alternatively, tensile strain in the GaAsP film due to incorporation of C could be causing an increase in As incorporation into the film in compensation.

4.3.5. Carrier concentration and mobility of C-doped GaAsP

Room-temperature hole concentration and mobility versus $CBrCl_3$ flow rate with a fixed AsH_3 fraction (0.18) are shown in Figure 4.8. A saturation of hole concentration at 2×10^{19} cm^{-3} is observed at high $CBrCl_3$ flow rates. The C concentration of Sample 8A (AsH_3 fraction of 0.18 and $CBrCl_3$ flow rate of 118 µmol/min) was measured by SIMS to be 4×10^{19} cm^{-3} ± 1×10^{19} cm^{-3}, corresponding to a dopant activation near 50%. The error in C concentration is large because of the unknown matrix effects of GaAsP compared to available GaAs standards. The dopant

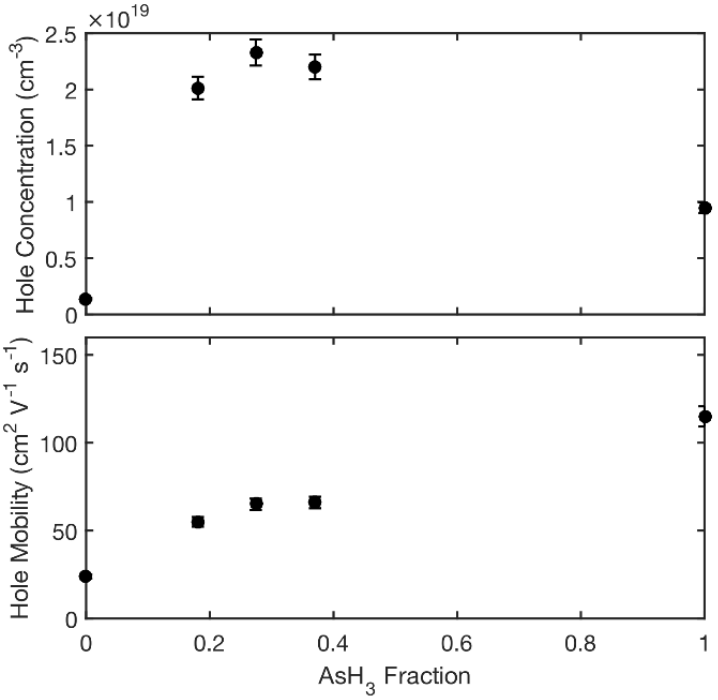

Figure 4.9: Hole concentration and mobility versus AsH_3 fraction with a fixed $CBrCl_3$ flow rate (59 µmol/min).

activation level of near 50% suggests that the saturation in hole concentration is due to a combination of multiple factors: a limitation in incorporation of C into the GaAsP film along with a reduction in dopant activation.

Figure 4.9 shows the room-temperature hole concentration and mobility as a function of AsH_3 fraction with a constant $CBrCl_3$ flow rate of 59 μmol/min. Hole concentrations for the three GaAsP samples are similar, while that of the GaAs and GaP samples are lower. Hole mobility increases almost linearly with increasing AsH_3 fraction. This behavior could be due to a decrease in hole effective mass with increasing As concentration in combination with changing amounts of dopant scattering from incorporated C.

Chapter 5: GaAsP/InGaP HBT performance with low defect densities

In this chapter, we will discuss the performance of GaAsP/InGaP HBTs grown on GaAs substrates with low defect densities. GaAs substrates are chosen instead of Si so that we can study the effects of $GaAs_xP_{1-x}$ composition on the HBTs absent the effects of the defects present from heavily lattice-mismatched III-V growth on Si. Although there are still some threading dislocations present from the GaAsP graded buffer, these have a much lower density and have minimal effect on HBT performance. The effect of dislocation density on HBT performance is discussed in depth in Chapter 6.

$GaAs_xP_{1-x}$ with $0.8 < x < 1$ will be used for the base and collector layers, and $In_yGa_{1-y}P$ for the emitter with y adjusted to keep the emitter lattice-matched to the base and collector. We focus on these GaAsP compositions because the electron mobility does not decrease strongly from that of GaAs as they do for lower values of x [37]. We will show that HBT current gain does not deviate strongly with GaAsP composition. Physical mechanisms governing base current, and therefore current gain, will be investigated. Collector current will be modeled as a function of emitter-base voltage. In addition, the breakdown voltage of GaAsP will be confirmed to increase over that of GaAs, consistent with modeling.

Table 5.1: Target epitaxial layer structure for GaAsP/InGaP HBTs on GaAs substrates. Four samples were grown with x = 0.825, 0.873, 0.941, and 1. y was chosen such that the InGaP layer is lattice-matched to the surrounding GaAsP layers.

Layer	**Material**	**Thickness (nm)**	**Growth Temperature (°C)**	**Polarity**	**Doping (cm^{-3})**
Contact	$GaAs_xP_{1-x}$	100	650	*n*	2×10^{19} Si
Emitter Cap	$In_yGa_{1-y}P$	25	650	*n*	1×10^{18} Si
Emitter	$In_yGa_{1-y}P$	50	650	*n*	6×10^{17} Si
Spacer	$GaAs_xP_{1-x}$	3	650	-	UID
Base	$GaAs_xP_{1-x}$	90	600	*p*	7×10^{17} C
Collector	$GaAs_xP_{1-x}$	500	650	*n*	1×10^{17} Si
Sub-collector	$GaAs_xP_{1-x}$	500	650	*n*	5×10^{18} Si
Graded Buffer	ΔGaAsP	0–4000 (depending on final composition)	725	*n*	5×10^{18} Si
Initiation	GaAs	100	725	*n*	5×10^{18} Si
Substrate	GaAs	-	-	*n*	2×10^{18} Si

5.1. Details of experimental samples

Four HBT structures were grown with $GaAs_xP_{1-x}$ compositions of x = 0.825, 0.873, 0.941, and 1. Table 5.1 shows the generalized epitaxial structure. All samples were grown on n+ (100) GaAs substrates, with a 6° offcut towards the nearest <111>B direction. This particular offcut was chosen because GaAsP or GaAs grown at temperatures above 600 °C on (100) SiGe or Ge, respectively, with a 6° offcut towards the nearest <111> direction will adopt this orientation [21]. A tensile $GaAs_xP_{1-x}$ compositionally graded buffer with a grade rate of 0.2% strain/μm was used to reach the desired lattice constant for the final device layers. This resulted in a buffer thickness ranging from 0 nm (x = 1) to 4 μm (x = 0.825). More details of the tensile GaAsP graded buffer are discussed in Section 4.1. The $In_yGa_{1-y}P$ emitter layer was grown with a composition lattice-

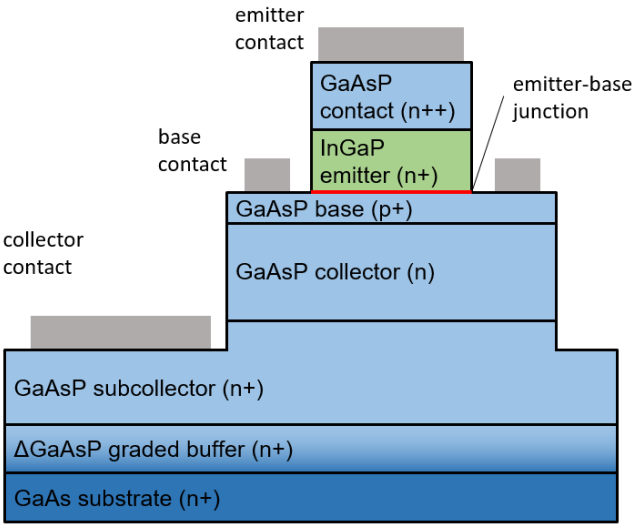

Figure 5.1: Cross-sectional schematic of GaAsP/InGaP HBT on GaAs substrate.

matched to the GaAsP layers directly above and below. A 5 s purge step, holding group V precursor flow rates constant from the previous layer, was implemented while switching from GaAsP to InGaP and from InGaP to GaAsP. The graded buffers were grown with a substrate temperature of 725 °C to increase dislocation glide velocities and therefore relaxation of the films [81]. Device layers were then grown at 650 °C, except for the GaAsP base layer, which was grown at 600 °C to increase the incorporation of the C dopant [82]. All temperature ramps were executed with a group V overpressure (mixed AsH_3 and PH_3) but with no group III precursor flow.

Figure 5.1 shows a cross-sectional schematic of the devices after fabrication is complete. Details of the fabrication process are discussed in Section 2.5.

5.2. Epitaxial film characterization

XRD was performed on each sample to measure the GaAsP lattice constant and composition. The compositions measured by XRD are those listed in Section 5.1. The GaAsP device layers are nearly fully relaxed, with a maximum residual strain of only 0.06%.

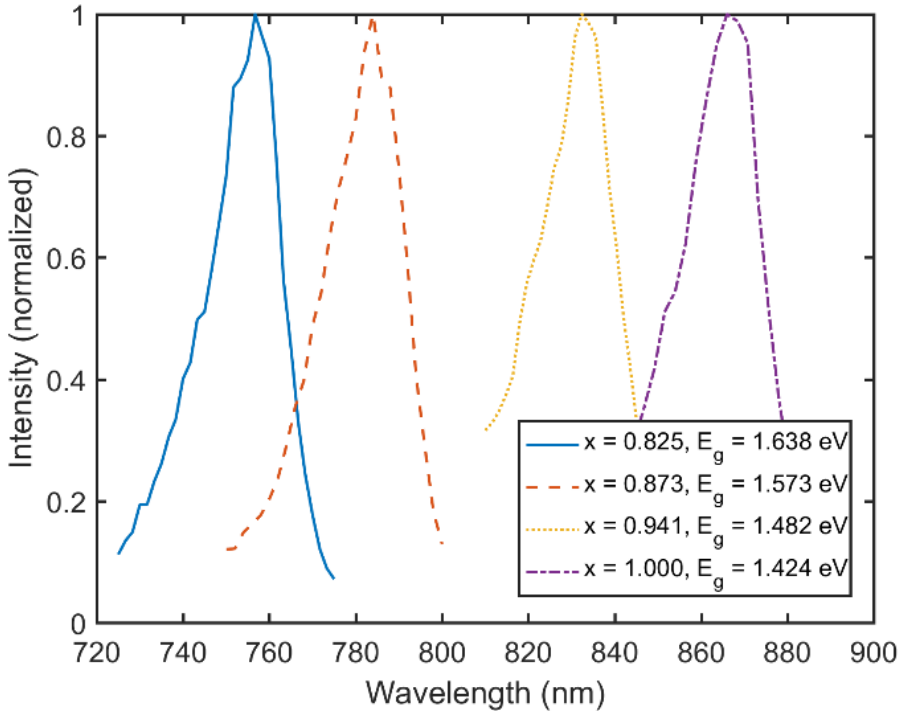

Figure 5.2: Photoluminescence spectra from the GaAsP base/collector layers.

PL spectra from the GaAsP base/collector layers are shown in Figure 5.2. Band gap (E_g) for each sample was calculated by shifting the corresponding photon energy for each peak by −7 meV, the amount necessary for the GaAs peak to coincide with 1.424 eV. This shift can be attributed to electron-hole pairs recombining with non-zero momentum. E_g values from the GaAsP samples correspond well with what would be predicted by the XRD composition data, varying by less than 10 meV [83].

Figure 5.3a shows a cross-section TEM image of the sample with x = 0.825. While misfit dislocations are visible in the GaAsP graded buffer region (not shown), there are no defects visible in the active device layers. The InGaP/GaAsP emitter-base interface is sharp. Figure 5.3b shows a transmission electron diffraction (TED) pattern from the InGaP emitter along the (110) orientation. The superspots indicate the presence of Cu-Pt ordering.

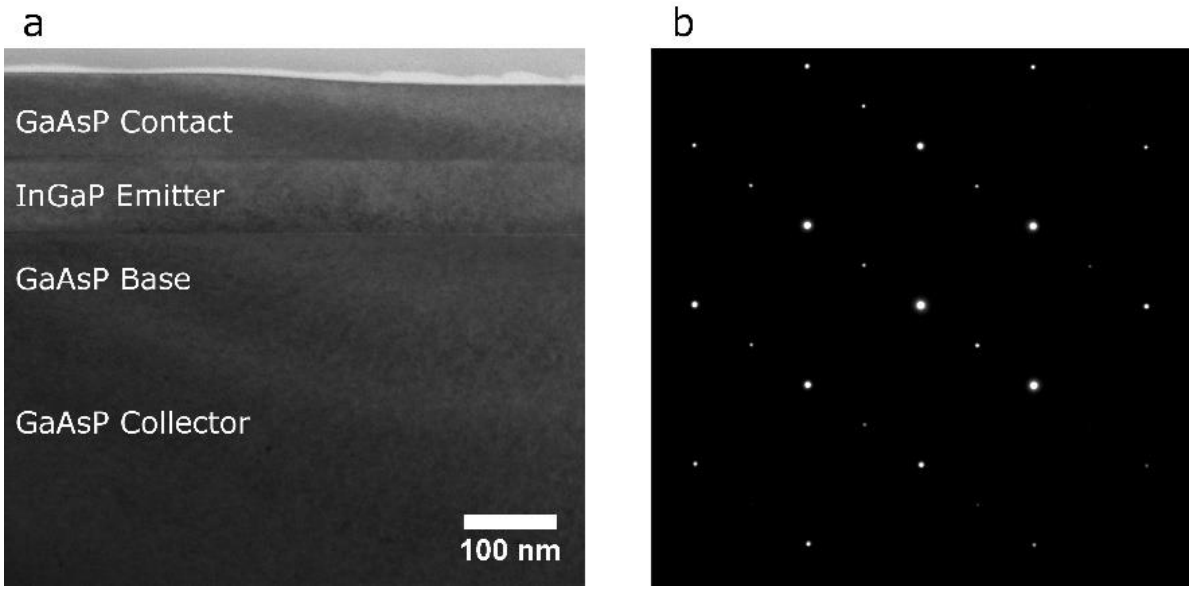

Figure 5.3: (a) Cross-section TEM image of GaAsP/InGaP HBT with x = 0.825 (Acc. Voltage = 200 kV). (b) (110) TED pattern from InGaP emitter area with superspots indicating Cu-Pt ordering.

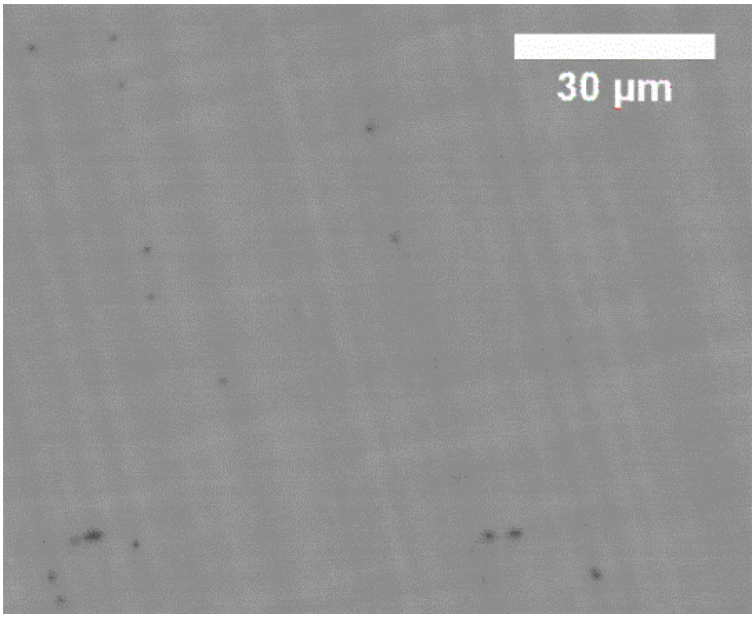

Figure 5.4: EBIC image of base-collector junction of the $GaAs_{0.825}P$ device under 0 V bias. Black spots are caused by recombination at threading dislocations originating from the underlying GaAsP graded buffer.

Figure 5.4 shows an EBIC image of the base-collector diode of the structure with x = 0.825 (upper contact and emitter layers etched off). Threading dislocations originating in the GaAsP graded buffer layer extend to the surface of the sample, passing through the base-collector junction. The minority carrier lifetime is suppressed near these dislocations, resulting in a reduction in collected current and a dark spot in the EBIC image. The threading dislocation density (ρ_{TD}) in this sample is $(1.5 \pm 0.4) \times 10^5$ cm^{-2}. This dislocation density only has a small effect on the electron minority carrier lifetime in GaAs, such that current gain in an HBT with a narrow base width should not be affected [84]. The other samples with $x > 0.825$ have less lattice mismatch with the GaAs substrate. Therefore, this is likely an upper bound of the TDD for all of the samples discussed here.

5.3. Issues with preliminary HBTs

My first attempts at fabricating HBTs were done with a lattice-matched GaAs/InGaP device on a GaAs substrate. The goal here was to fabricate a repeatable GaAs/InGaP control device against which GaAsP/InGaP devices on both GaAs and Si substrates could be compared. A DC current gain of at least 100 was desired for the control device. That way, even small increases in base current in the GaAsP/InGaP devices caused by experimental variables such as device composition, substrate, and strain relief scheme would not be overwhelmed by other base

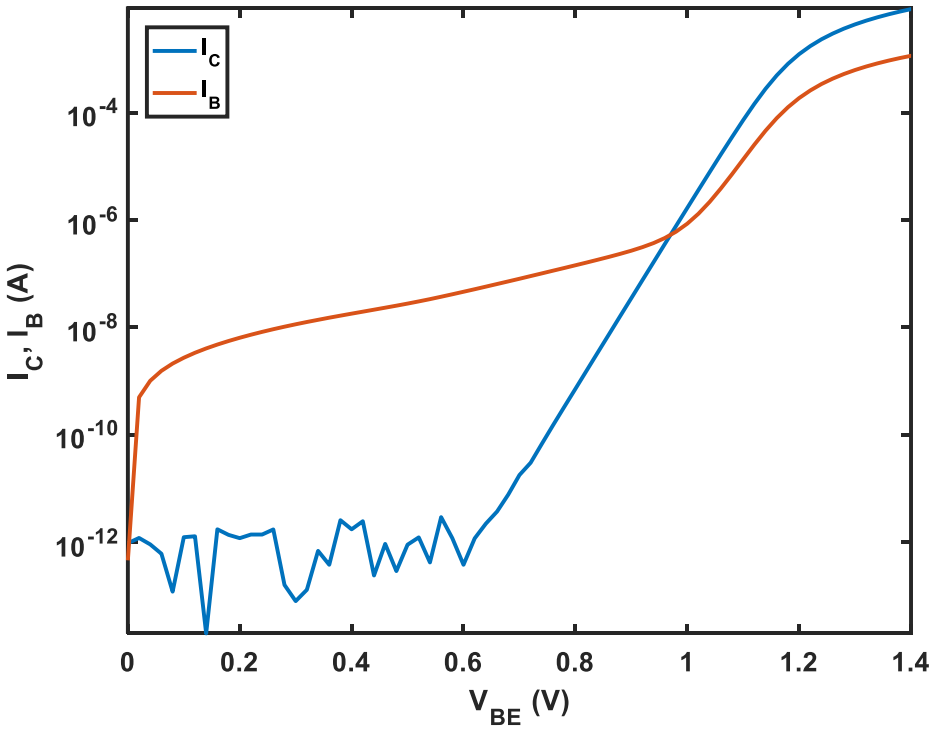

Figure 5.5: Gummel plot for initial GaAs/InGaP control HBT with poor device performance.

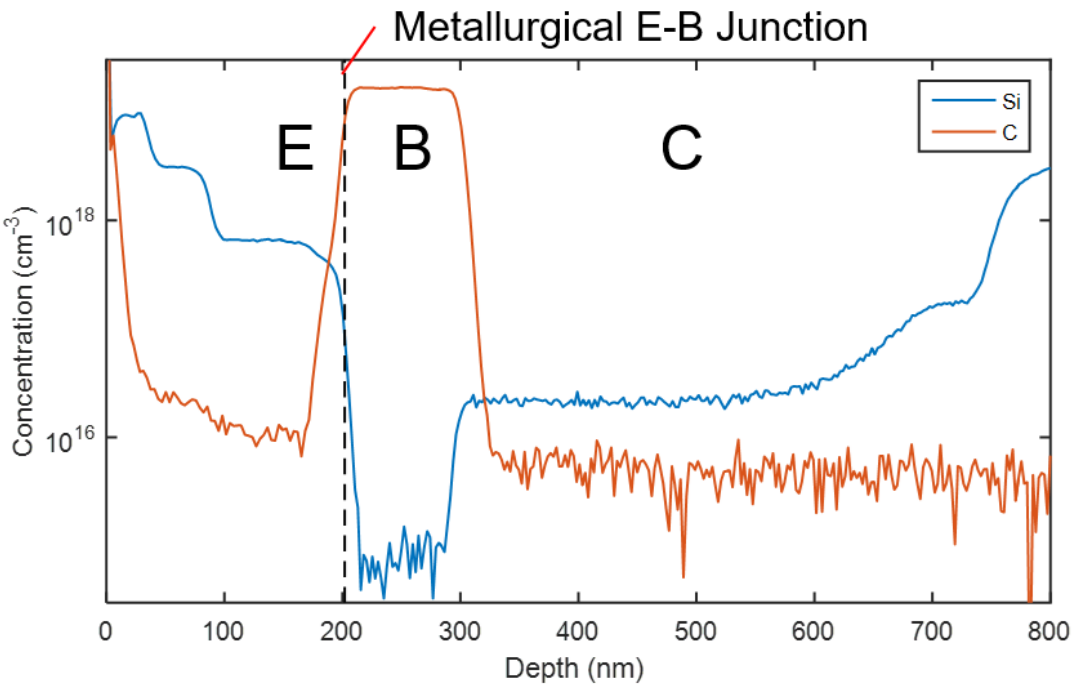

Figure 5.6: SIMS profiles of Si and C dopants in initial GaAs/InGaP control HBTs.

current components effecting even the control device. Unfortunately, the initial attempts to fabricate control devices did not meet this standard for several reasons, to be discussed here.

The initial epitaxial structure for the GaAs/InGaP control HBT was very similar to that of the finalized devices in Table 5.1. Three major differences in the epitaxial structure were: (1) no unintentionally-doped setback layer between the base and emitter layers, (2) higher p-type C doping in the base region (1.7×10^{19} cm^{-3}), and (3) a lattice-mismatched InGaAs emitter cap layer. A fourth difference between the initial HBTs and the finalized devices were that the initial devices did not have sidewall passivation. I will address each of these changes and why they were made.

A Gummel plot for a representative initial HBT (60 μm diameter) is shown in Figure 5.5. The maximum current gain measured was approximately 10, about a factor of 10 lower than the goal of 100. In addition, the range of voltages with current gain > 1 is severely limited by a base leakage current that dominated at V_{BE} < 1 V.

SIMS was used to check the doping levels in the various device layers (Figure 5.6). While the doping levels were within an acceptable range from of the target levels, the heavy C doping of the base had extended significantly into the InGaP emitter. We suspected that this may have a deleterious effect on the valence band offset between the GaAs base and InGaP emitter, thereby allowing for reverse-injection of holes from the base into the emitter. To investigate further, the band structure of the emitter-base junction was simulated using Nextnano with both abrupt doping and the doping profile measured using SIMS. The simulated band structure is shown in Figure 5.7. For the case of abrupt doping, there is a sharp valence band offset of about 400 meV. However, for the case of the actual doping profile, the effective valance band offset has been reduced dramatically to around 200 meV. Because reverse injection of holes from the base into the emitter depends exponentially on the valence band offset, this could be the cause of the increased base current and therefore the higher current gain. This simulation is dependent on the accuracy of the Nextnano input parameters (i.e. band gaps and band offsets of GaAs and InGaP).

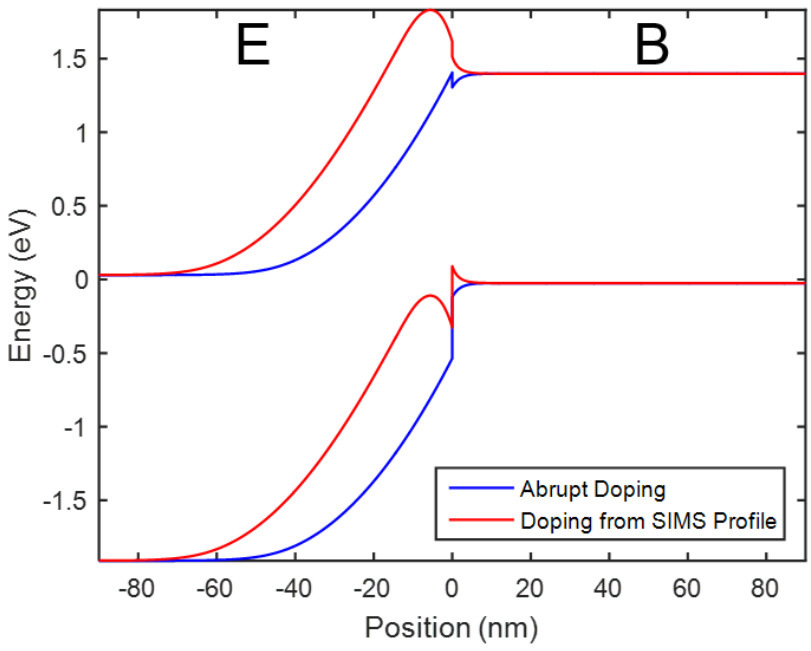

Figure 5.7: Simulated band structure of emitter-base junction with different doping profiles.

These parameters have some dependence on growth parameters. However, despite this possible source of inaccuracy, the trend of reduced effective valence band offset still holds.

In order to reduce the effect of the extension of the heavy C base doping into the emitter, a base doping setback was inserted. This is a 3 nm unintentionally doped GaAs layer in between the base and collector, grown at 650 °C. For devices with a GaAsP base and collector, the base doping setback was comprised of GaAsP with the same composition.

In addition to the base doping setback, the base doping level was decreased to further increase the current gain. A high base doping increases base current due to recombination of minority electrons during transit of the base. This can be due to increased recombination due to impurity scattering and/or increased Auger recombination. We found that decreasing the base doping from 1.7×10^{19} cm^{-3} to 7×10^{17} cm^{-3} was sufficient to increase current gain at $V_{BE} > 1$ to greater than 100. In order to reduce the base C doping by this amount, the $CBrCl_3$ pickup flow rate was reduced from 40 sccm to 7 sccm, the minimum stable flow for that particular MFC. In addition, to further reduce the mass flow of $CBrCl_3$ into the reactor, the $CBrCl_3$ bubbler temperature was reduced from 26 °C to 5 °C. This caused an estimated factor of four reduction in

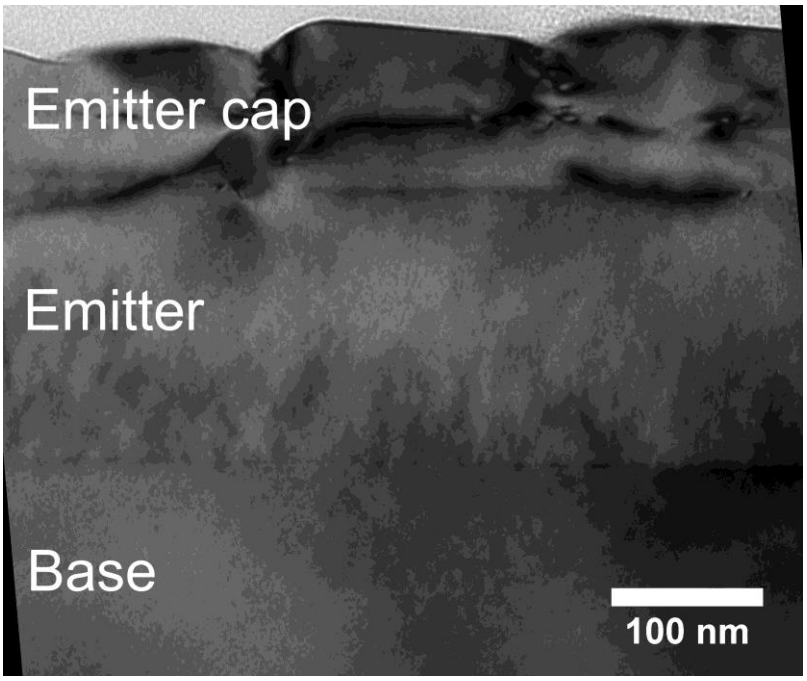

Figure 5.8: TEM image of GaAsP/InGaP HBT with InGaAs emitter cap (Acc. V = 200 kV, g = (220)).

vapor pressure in the source based on limited vapor pressure data available from the precursor vendor.

The initial GaAs/InGaP control HBT structure had a lattice-mismatched n++ InGaAs emitter cap. This was included to reduce the emitter contact resistance. This did not seem to introduce any problems with the GaAs/InGaP HBT. However, for early GaAsP/InGaP HBTs it caused dislocations to "punch down" into the InGaP emitter region. This can be seen on the left-hand side of the TEM image in Figure 5.8. The punch down was likely to be more prominent for GaAsP with lower As content because of the greater lattice mismatch between the emitter and emitter cap layers. Because of this, it was decided to leave out the InGaAs emitter cap layer and only use a n++ GaAs or GaAsP emitter cap lattice-matched to the underlying device layers. This change is at the expense of increased contact resistance, especially for the GaAsP devices which have higher band gaps. However, even with the increased series resistance, DC current gain in the

desired range is still measureable. If AC measurements were to be attempted, contact resistance would play a more important role and this issue would need to be addressed [56].

The final issue that was addressed was the high base leakage current at lower voltages. The base leakage current can be seen for values of V_{BE} between 0 and 1 V with an ideality factor much greater than 2 (see Figure 5.5). This leakage current does not have a considerable effect at V_{BE} > 1 V, but completely masks the base current due to bulk recombination at lower voltages. This severely limits the device's range of exponential behavior, restricting analysis of the DC characteristics. If base current at V_{BE} = 0.4 V is plotted for different device circumferences (Figure 5.9), it is clear that it varies linearly with circumference and is therefore due to recombination or leakage at the sidewalls. Therefore, it was deemed important that suitable mesa sidewall passivation be introduced. We decided to use a 10 nm Al_2O_3 layer as sidewall passivation, the details of which are explained in Section 2.5. All further devices that are discussed have this sidewall passivation, which severely reduces the base leakage current at low voltages.

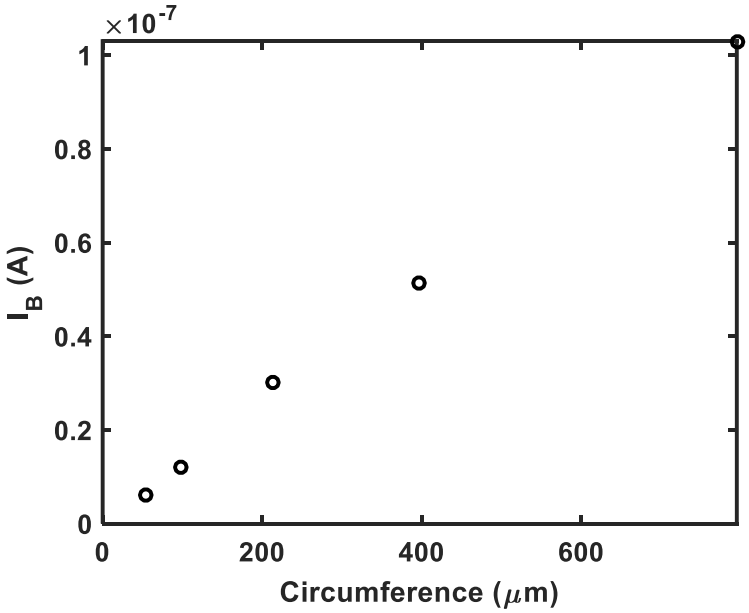

Figure 5.9: Base current vs device circumference for GaAs/InGaP HBTs without surface passivation (V_{BE} = 0.4 V).

5.4. DC characterization

5.4.1. Gummel plots and DC current gain

Figure 5.10 shows Gummel plots—collector current (I_C) and base current (I_B) plotted as a function of the base-emitter voltage (V_{BE})—for the $GaAs_xP_{1-x}/In_yGa_{1-y}P$ HBTs of four different compositions (x = 1, 0.941, 0.873, 0.825). The diameter of the emitter-base junction (d, highlighted in Figure 5.1) is 60 μm and the base-collector voltage (V_{BC}) is 0 V. The ideality factor (n) of the collector current is close to 1 for all GaAsP compositions, ranging from 1.02 to 1.08. I_B has two exponential regimes: $n \approx 1.8$ for I_B less than 10^{-9} A and $n \approx 1.4$ for I_B greater than 10^{-9} A.

Roll-off of both I_C and I_B at high V_{BE} is caused by series resistance at the emitter ohmic contact. This series resistance worsens as x decreases because of the increased band gap of the

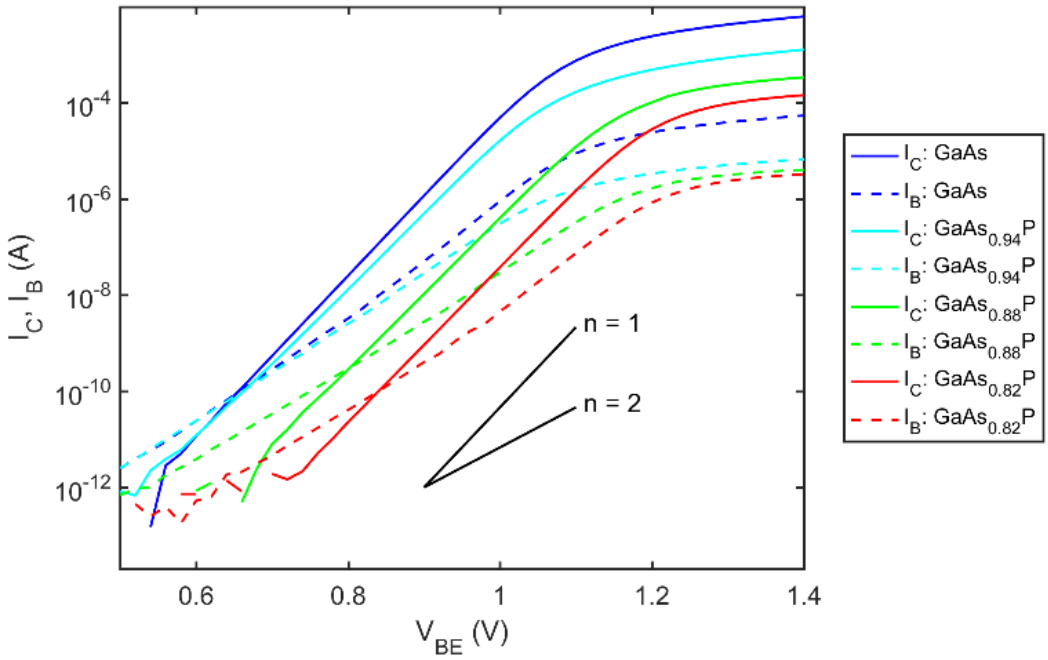

Figure 5.10: Gummel plots for GaAsP/InGaP HBTs of four different compositions. Measurement was taken with V_{BC} = 0 V.

$GaAs_xP_{1-x}$ emitter contact layer. This could be mitigated by adding a mismatched n++ GaAs or InGaAs contact layer to all of the devices. In addition, contact resistance could be improved by using a separate alloyed Ge/Ni/Au contact to the n-type emitter and collector [85].

Figure 5.11a shows I_C and I_B for a range of device sizes at V_{BE} = 0.930 V and V_{BC} = 0 V. At this biasing condition, I_C scales with area (d^2), indicating that the entire emitter-base junction is uniformly injecting electrons into the base. The collector current density (J_C) for all device sizes is 8×10^{-5} A·cm^{-2}. There are two different dependences of I_B on device diameter. For diameters smaller than 60 µm, I_B scales with $d^{1.3}$, while for those larger than 60 µm, it scales with d^2. Figure 5.11b shows the same as Figure 5.11a but at V_{BE} = 1.134 V and J_C = 0.13 A·cm^{-2}. At this current density, both I_C and I_B scale with d^2 across the entire range of device sizes.

A DC current gain (β) of over 100 was measured for all GaAsP compositions at relatively high collector current densities (J_C), greater than about 20 A·cm^{-2}. The only exception was the device with x = 0.825, where J_C could not reach this value due to series resistance and β only reached 60. However, β is reduced at lower current densities. The main sources of I_B—which

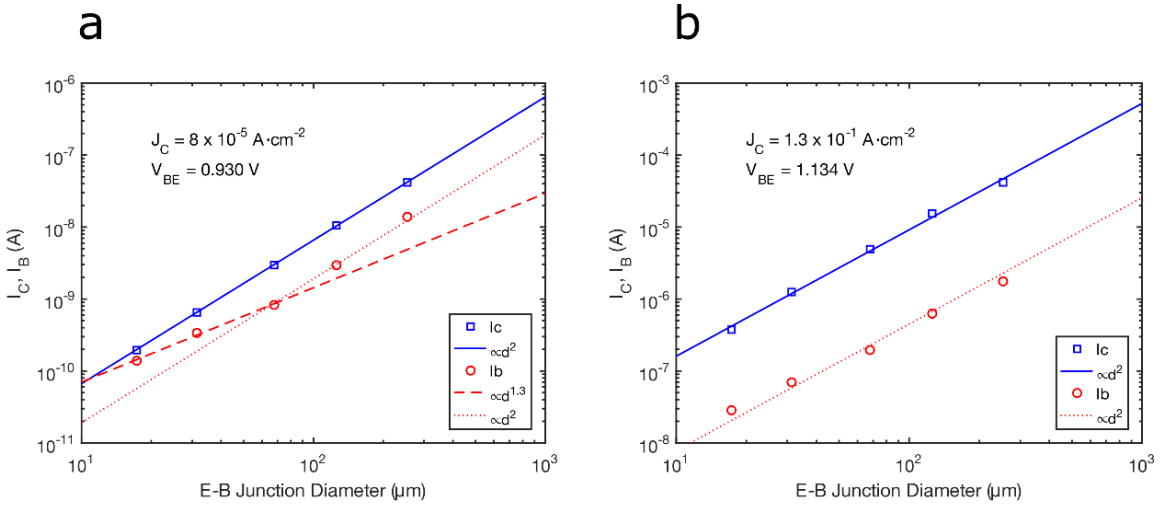

Figure 5.11: Collector current and base current vs device size (emitter-base junction diameter) for $GaAs_{0.825}P$ HBT. Plotted at low (a) and high (b) current densities. Trend lines are plotted with approximate slopes noted in the legend.

limit β at lower current densities—can be ascertained by observing how I_B varies with active layer composition, with V_{BE}, and with emitter-base junction area (A_E).

For a given I_C, β does not change appreciably with GaAsP composition. With decreasing x, the valence band offset between $GaAs_xP_{1-x}$ and lattice-matched InGaP is expected to decrease [60]. This indicates that backwards injection of holes from the base into the emitter does not contribute significantly to I_B, because this injection current would increase exponentially with decreasing valence band offset, thereby reducing β.

At smaller device sizes (d < 60 μm) and the lower current densities shown in Figure 5.11a, I_B is dominated by a process occurring at the perimeter of the emitter-base junction. Because I_B varies exponentially with V_{BE} in this regime with n close to 2 (Figure 5.10), we can conclude that it is likely space-charge region (SCR) recombination at the emitter-base interface occurring close to the perimeter of the device. This could be due to imperfect sidewall passivation near the emitter-base interface, or defects in the emitter base interface formed near the perimeter of the mesa during device processing. To improve performance at small device sizes, particularly for those smaller than what was fabricated for this study, it will be critical to identify and address this source of I_B.

At larger device sizes and larger currents, I_B is proportional to the area of the emitter-base junction (Figure 5.11). $n \approx 1.4$ in this regime for all GaAsP compositions (Figure 5.10), which indicates a combination of SCR recombination and quasi-neutral region (QNR) recombination during electron transit across the base. The fact that SCR recombination still occurs even though no defects are visible in cross-section TEM of the emitter-base interface suggests that DC characteristics of GaAsP/InGaP HBTs are a more sensitive measure of interface quality than TEM of the GaAsP/InGaP interface. The GaAs/InGaP devices (x = 1) exhibit similar I_B behavior as the x = 0.825 devices shown in Figure 5.11. This is further evidence that the recombination

mechanisms governing I_B are not related to dislocations, but rather to interfacial defects between the InGaP emitter and GaAs(P) base. Future work may involve optimization of the growth conditions of this interface to reduce the rate of SCR.

5.4.2. Collector current behavior

The collector current of GaAs/InGaP HBTs has been previously modelled as thermionic emission of electrons from the emitter into the base, due to the sharp peak in the conduction band created by the abrupt emitter-base junction [34]. This process is illustrated in Figure 5.12a. These injected electrons diffuse across the base, combining only slightly with majority holes, into the collector. Therefore, collector current can be written as:

$$I_C = A_E A^* T^2 \exp\left(-\frac{E_A}{kT}\right), \quad 5.1$$

where A_E is the area of the emitter-base junction, $A^* = 4\pi q m_e^* k^2/h^3$ is the effective Richardson constant for thermionic emission, and E_A is the activation energy for injected electrons—the difference between the Fermi level in the emitter and the top of the conduction band peak at the emitter-base interface. Equation 5.1 can also be extended to the GaAsP/InGaP system. As in Kobayashi, et al., the conduction band offset (ΔE_C) can be written in terms of known quantities:

$$\Delta E_C = nkT\ln\left(\frac{A^* T^2}{I_C}\right) + (1-n)\delta_1 + \delta_2 + qV_{BE} - E_{g,B}, \quad 5.2$$

where $E_{g,B}$ is the band gap of the base, V_{BE} is the base-emitter voltage, δ_1 is the energy difference between the Fermi level and conduction band edge in the neutral emitter, and δ_2 is the energy difference between the Fermi level and the valence band edge in the neutral base. n, the collector current ideality factor, can be measured directly from the I_C vs V_{BE} curve. δ_1 and δ_2 are not ignored in Equation 5.2 because unlike in Kobayashi, et al., the dopant concentrations of the

emitter and base are such that they are non-negligible . $E_{g,B}$ values can be obtained from the PL data in Figure 5.2 as described in Section 3.1. m_e^* values for InGaP can be linearly interpolated between $0.15m_0$ for GaP (Γ valley) and $0.08m_0$ for InP, and m_h^* values for GaAsP can be linearly interpolated between $0.79m_0$ for GaP and $0.51m_0$ for GaAs. Therefore, all values on the right-hand side of Equation 5.2 are known.

If Equation 5.2 is used to calculate ΔE_C for the GaAsP/InGaP HBTs presented here, negative values ranging from –20 to –40 meV are yielded. This is true for any I_C - V_{BE} pair from the $n \approx 1$ regime. Of course, a negative or even a sufficiently small positive value of ΔE_C means that the above model of thermionic emission does not apply. Kobayashi, et al. report a ΔE_C of 30 meV for the InGaP/GaAs heterojunction according to the thermionic emission model. The discrepancy between their result and ours (for the GaAs/InGaP device) is due to one of two possibilities. First, the InGaP emitter in this work was grown at a temperature of 650 °C, while in Kobayashi, et al., it was grown at 700 °C. Evidence of Cu-Pt ordering in the emitter regions of the device can be seen in Figure 5.3. Cu-Pt ordering occurs in InGaP grown at temperatures

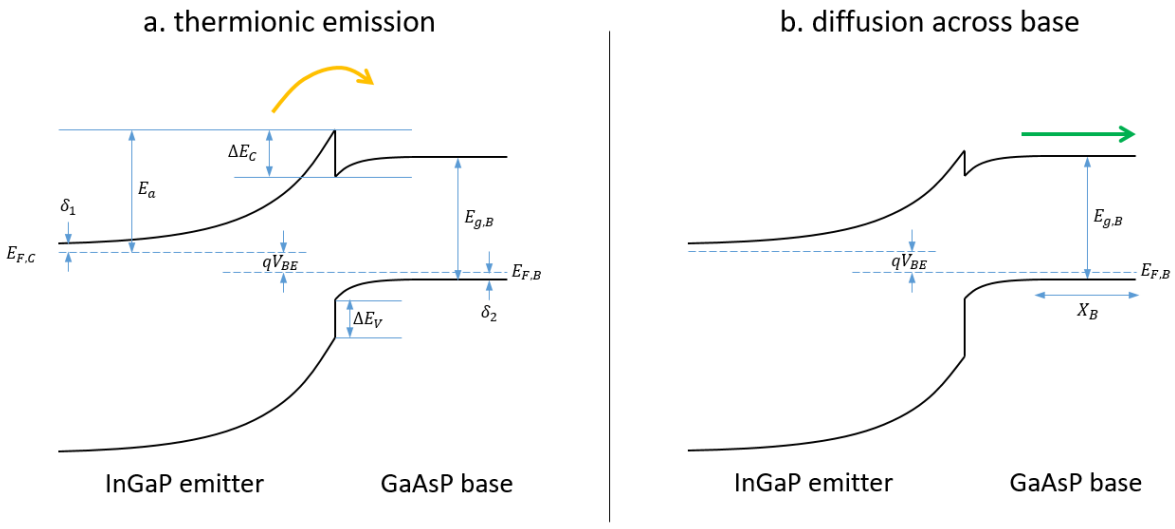

Figure 5.12: Possible limiting processes for I_C: (a) thermionic emission over the conduction band at the emitter-base junction and (b) diffusion across the quasi-neutral base region.

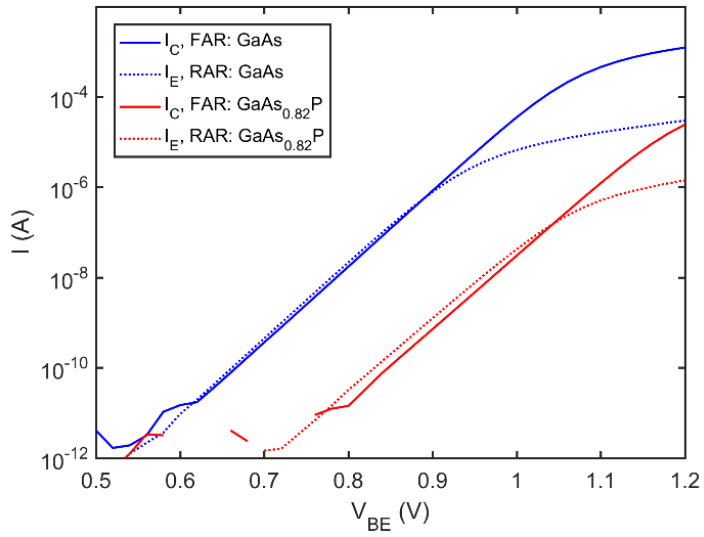

Figure 5.13: Collector current in the forward-active regime (FAR) and emitter current in the reverse-active regime (RAR) for GaAs/InGaP and $GaAs_{0.825}P$/InGaP HBTs. The emitter-base junction diameter is 60 µm.

between around 550 °C and 750 °C with a corresponding reduction in band gap that is greatest near 650 °C [86]. The shift in band gap would therefore be greater for our GaAs/InGaP sample than for Kobayashi, et al., resulting in a smaller conduction band offset. Second, because our samples have a lower base doping than those in Kobayashi, et al., the depletion width of the emitter-base junction extends further into the base. This reduces the prominence of the conduction band peak at the emitter-base interface.

In either case, it can be concluded that thermionic emission from the InGaP emitter into the GaAsP base due to the conduction band offset does not significantly affect I_C. This can be confirmed by comparing I_C in the forward-active regime (FAR) with the emitter current (I_E) in the reverse-active regime (RAR). In RAR, I_E is a measure of electrons injected from the collector into the base. Because the collector and base are the same material, there is no conduction band offset at the base-collector junction. In Figure 5.13, I_C in FAR and I_E in RAR are compared for GaAs/InGaP and $GaAs_{0.825}P$/InGaP HBTs, both with emitter-base junction diameters of 60 µm. The ideality factors for both FAR and RAR are n = 1.02 for the GaAs/InGaP device and n = 1.06

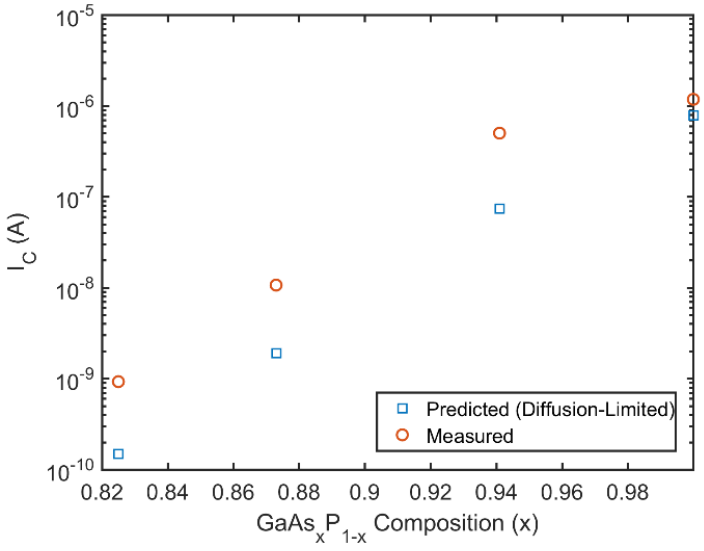

Figure 5.14: Predicted and measured collector current for V_{BE} = 0.9 V and V_{BC} = 0 V. Emitter-base junction diameter is 60 µm.

for the GaAs$_{0.825}$P/InGaP device. This demonstrates that in both modes of operation, carrier transport is limited by diffusion across the base layer, not by the conduction band offset. The ideality factor would be expected to be significantly higher in FAR than in RAR if it was limited by thermionic emission across the emitter-base junction [57].

Therefore, I_C can be modeled solely by diffusion of electrons across the quasi-neutral base layer, which can be written as:

$$I_C = A_E \frac{qD_{n,B}}{X_B} \frac{n_{i,B}^2}{N_B} \exp\left(\frac{qV_{BE}}{kT}\right), \tag{5.3}$$

where $D_{n,B}$ is the diffusivity of electrons in the base, X_B is the quasi-neutral base thickness, $n_{i,B}$ is the intrinsic carrier concentration in the base, N_B is the p-type doping level in the base [56]. This process is illustrated in Figure 5.12b. Figure 5.14 shows I_C predicted by this diffusion process along with the measured I_C for the GaAsP HBTs of different composition, all at V_{BE} = 0.9 V and V_{BC} = 0 V. For the GaAs device (x = 1), the measured I_C is well-predicted by this model to within the expected error. However, for all of the GaAsP devices (x < 1), the measured I_C is about 10

times higher than what is predicted. The origin of this behavior is unknown and under further investigation, but a higher I_C for a given V_{BE} is beneficial because it yields a higher transconductance.

5.4.3. Breakdown voltage

Breakdown voltage in the common base, emitter open configuration (BV_{CBO}) was measured for the HBTs with varying GaAsP compositions. This was done by sweeping V_{BC} from 0 V to −20 V in 0.1 V steps, while leaving the emitter terminal floating. Figure 5.15 shows the I-V trace from a single example breakdown voltage measurement. BV_{CBO} is defined here as the point of highest curvature, which is determined by finding the zero of the third derivative of a 5-point moving average.

BV_{CBO} for the GaAsP HBTs of varying composition are plotted in Figure 5.16. The data shows a general trend of decreasing breakdown voltage with decreasing x. The increasing band gap of $GaAs_xP_{1-x}$ with lower x should yield devices with higher breakdown voltage, so this is

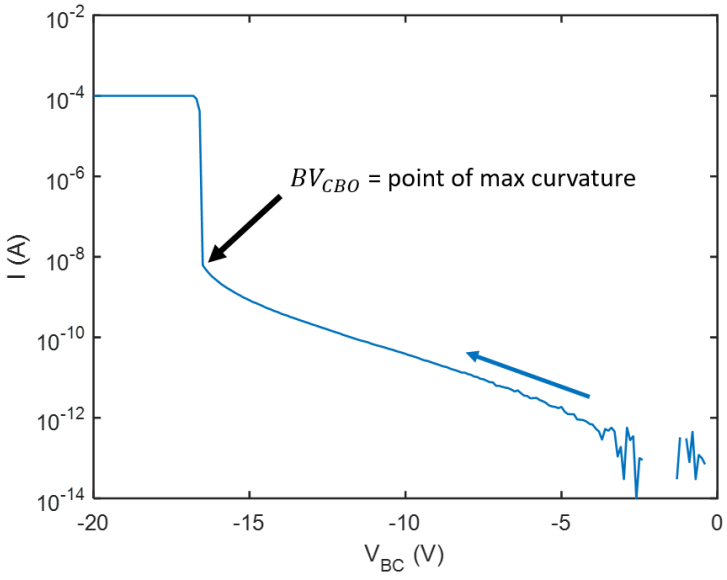

Figure 5.15: I-V curve from an example breakdown voltage measurement.

unexpected. However, by looking at the underlying physics governing device breakdown, it is possible to understand the origin of this behavior.

Breakdown in electrical devices can occur through several effects, but the most common is impact ionization, which will be discussed here. In the presence of an electric field, electrons travelling in a semiconductor will randomly collide with the lattice. If the electron travels far enough in between collisions, it will have gained enough kinetic energy to create an electron-hole pair upon its next collision. The minimum kinetic energy required for this to occur, known as the impact ionization threshold energy (E_{ii}), scales with about $1.5E_g$ [58]. The rate of this impact ionization process (α, number of ionization events per unit length that an electron travels) goes as:

$$\alpha \cong \frac{q|\mathcal{E}|}{E_{ii}} \exp\left(-\frac{E_{ii}}{q|\mathcal{E}|l_c}\right), \tag{5.4}$$

where $\mathcal{E}$ is the electric field and l_c is the mean free path of an electron between collisions [58]. If the electric field in the device is high enough, this process creates an unchecked increase in current known as avalanche breakdown. A multiplication factor M can be defined as the ratio between the total current passing through the device and the current that started the multiplication process. In the case of a p-n junction (such as the base-collector junction of an HBT), this can be described as:

$$M = \frac{1}{1 - \int_{-x_p}^{x_n} \alpha_e \exp\left(\int_x^{x_n} (\alpha_h - \alpha_e) dx\right) dx}, \tag{5.5}$$

where x_n and x_p are the depletion widths on the n and p sides of the junction, respectively [58]. α_h and α_e are both functions of x because they both depend on the electric field, which varies depending on the location in the space charge region. When M approaches infinity (the denominator equals 0), avalanche breakdown occurs. For an HBT biased across the base and collector with the emitter open, the voltage at which this occurs is called BV_{CBO}.

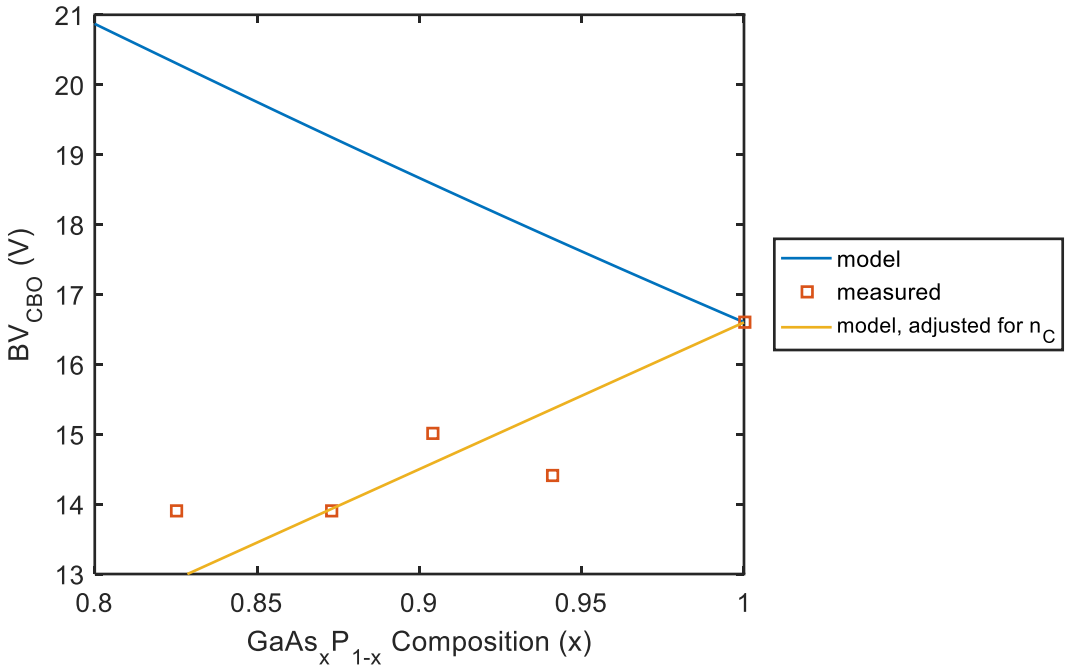

Figure 5.16: Breakdown voltage of GaAsP HBTs of various compositions. The initial model predicted increasing BV_{CBO} with decreasing As content, contrary to the data. The model accounting for changes in collector doping follows the trend of the data.

For an HBT, there are two aspects of the device structure that significantly affect the breakdown voltage: the collector doping and the collector width. Both of these parameters can potentially affect the collector depletion width x_n, which drastically effects the value of M and therefore the voltage at which breakdown occurs. In simplified terms, when the collector doping is high, the depletion width will be lower, and therefore the average electric field across the junction will be higher for a given biasing condition. This causes breakdown to occur at a lower voltage. Likewise, a shorter collector width has the same effect on the average electric field and breakdown voltage.

The depletion widths x_n and x_p are dependent on the voltage, as is the electric field $\mathcal{E}$. Therefore, a closed-form solution for BV_{CBO} cannot be written. However, x_n and x_p can be found

using Gauss' law for any given biasing condition. Therefore, the value of BV_{CBO} can be found using an iterative numerical approach.

BV_{CBO} is predicted using the equations above as a function of $GaAs_xP_{1-x}$ composition for the target GaAsP/InGaP HBT structure presented in Table 5.1. This model was implemented using a MATLAB script, which is shown in full in Appendix C. l_c from Equation 5.4 is unknown, so it is chosen such that the model predicts the GaAs breakdown exactly and assumed to remain constant for the GaAsP HBTs. This is a reasonable assumption because the electron mobility of $GaAs_xP_{1-x}$ in this range of compositions ($x > 0.8$) has been shown to be approximately constant [37]. The predicted BV_{CBO} is shown in blue in Figure 5.16. An HBT with base/collector regions of $GaAs_{0.825}P$ should have a ~20% higher BV_{CBO} than a GaAs HBT.

The measured data does not agree with this model. Instead of the predicted 20% increase in breakdown voltage by replacing GaAs with $GaAs_{0.825}P$, there is a ~20% decrease in BV_{CBO}. This could be due to one of two effects: a difference in collector thickness or doping in the GaAsP devices leading to poorer electrostatics for breakdown, or an otherwise increased rate of impact ionization (α) for GaAsP due to intrinsic materials properties such as decrease in E_{ii}. We believed that the most likely of these issues was the collector doping level. GaAsP growth rate does not change as a function of composition, and E_{ii} should decrease with increasing band gap.

The collector doping of the GaAs HBT structure as well as the $GaAs_{0.825}P$ HBT structure were measured using SIMS. While the GaAs HBT had a collector doping of 1.2×10^{17} cm^{-3}, which was acceptably close to the target value of 1×10^{17} cm^{-3}, the $GaAs_{0.825}P$ HBT had a collector doping of 3.5×10^{17} cm^{-3}, well over the target. The collector layers of all of the GaAs(P) HBT structures in this study were grown with the same Si precursor flow, under the assumption that the incorporation of Si was limited solely by the cracking of the Si_2H_6 and that changes in GaAsP

composition would not affect Si incorporation. Clearly, this assumption was wrong. The increase in Si doping with decreasing As fraction causes a shorter depletion width for a given V_{BC}, resulting in higher electric fields across the base-collector junction. This causes an increase in impact ionization leading to a lower BV_{CBO}.

Figure 5.16 shows the same model as above, but adjusted for the changing collector doping levels measured by SIMS, in yellow. With this adjustment, the model follows the data well, suggesting that the model is valid for these devices. This means that if the GaAsP HBTs structures had been grown with the same collector doping as the GaAs HBT, then they should have increased breakdown voltages following the blue curve.

In order to verify that GaAsP HBTs are capable of yielding higher breakdown characteristics than GaAs HBTs—specifically, that E_{ii} of GaAsP is at least $1.5E_g$—we fabricated two devices with the exact same structure and doping, but one using GaAs and the other using $GaAs_{0.825}P$. Rather than growing and fabricating full HBTs, only the base-collector junction was

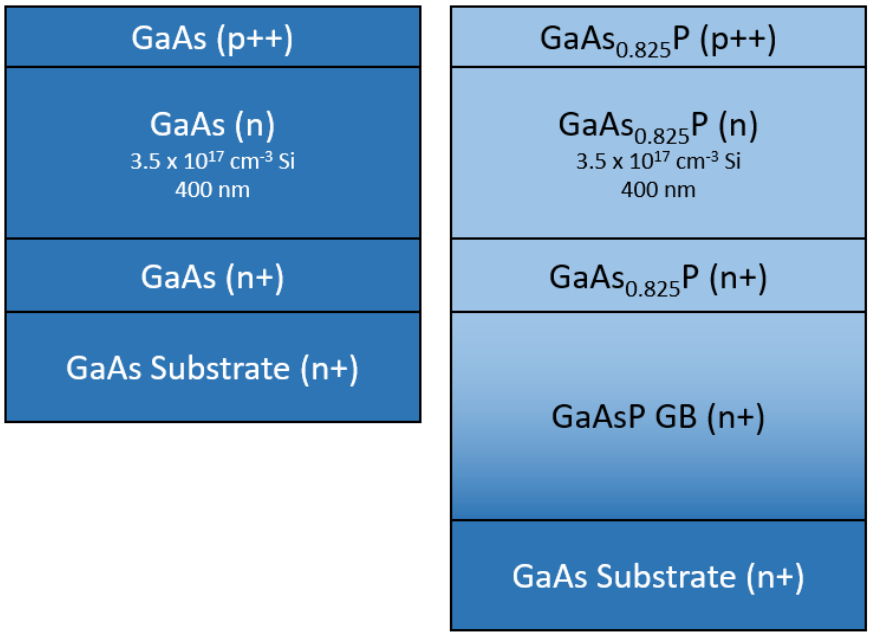

Figure 5.17: Schematic of GaAs(P) diode epitaxial structures for measuring breakdown voltage.

grown, leaving out the InGaP emitter. These diodes should behave identically to the base-collector junction of the full HBT. The epitaxial structures of the diodes are shown in Figure 5.17. For the GaAs diode structure, the Si_2H_6 flow was raised proportionally to increase the Si doping to the same level as that of the $GaAs_{0.285}P$. The diodes were fabricated with a process similar to that used for the HBT fabrication, described in Section 2.5.

The breakdown voltages of the GaAs and $GaAs_{0.825}P$ diodes are plotted in Figure 5.18. BV_{CBO} for the $GaAs_{0.825}P$ device is 15.7 V, increased from 11.5 V for the GaAs device. Plotted alongside is the breakdown voltage according to the model described above. The breakdown voltage of the $GaAs_{0.825}P$ diode is even higher than that predicted by the model. This shows with certainty that E_{ii} for GaAsP is equal to at least $1.5E_g$. An increase in breakdown voltage of this magnitude is helpful for the design of high-power circuitry.

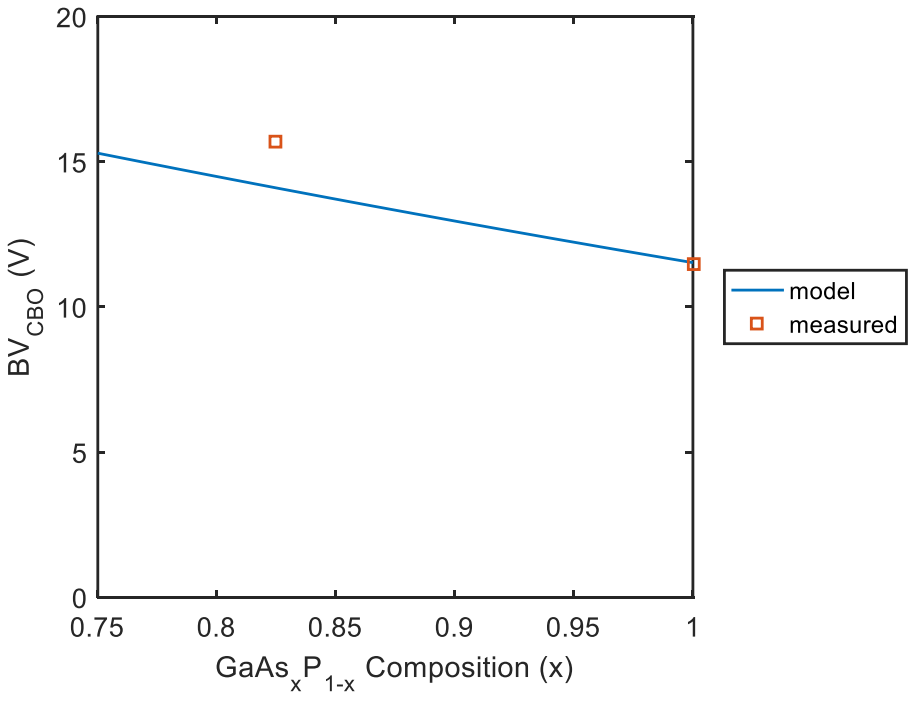

Figure 5.18: Breakdown voltage of GaAs(P) diodes with controlled thickness and doping.

Chapter 6: GaAsP/InGaP HBTs on Si substrates: effect of dislocations on HBT performance

For any semiconductor device, it is important to quantify and understand the deleterious effects of any crystalline defects that may be present. These defects could be isolated to one or more of the active layers—for example, the heavy p-type base doping in an HBT or interface traps at the emitter-base junction. Alternatively, they could be extended defects originating from layers below the active region, such as dislocations, stacking faults, or anti-phase boundaries. These extended defects are particularly important in structures grown on lattice-mismatched substrates. In this chapter, we develop an understanding the effects of crystalline defects, particularly those originating from III-V growth on Si substrates, on HBT performance.

6.1. Details of experimental samples

Many different attempts were made at growing and fabricating GaAs(P)-based HBTs on Si substrates. Both strain relaxation schemes described in Section 1.3 (SiGe graded buffer and 2-step Ge) were employed. In addition, a few other parameters were varied to engineer an HBT with lower defect density and better performance. The structure of the active device layers grown in this section are identical to that of the HBTs from Chapter 5 (i.e. the top seven rows of Table 5.1), except for the differences discussed below. Just as in Chapter 5, all of the active layers of a given HBT structure were targeted to be lattice-matched to each other. The composition of the $In_yGa_{1-y}P$ emitter was chosen such that it would be lattice-matched to the $GaAs_xP_{1-x}$ layers above and below. Table 6.1 lists all of the samples that will be considered in this section. The black text shows

Table 6.1: List of samples for studying effect of defect density on HBT performance with varying structural parameters. Measured dislocation densities (EBIC) and current gains for 60μm diameter HBTs are shown in green.

Sample #	$GaAs_xP_{1-x}$ Comp. (x)	Base Dopant	Substrate	Strain Relaxation Scheme	Notes	ρ_{TD} (cm^{-2})	ρ_{MD} (cm^{-1})	β at I_C = 2 x 10^{-4} A	β at I_C = 1 x 10^{-1} A
B1	1	C	GaAs	n/a		< 6000	< 1	70	263
B2	0.825	C	GaAs	GaAsP GB		1.5×10^5	52	50	-
B3	0.825	C	GaAs	GaAsP GB, fast grade	defect density purposefully increased from B2	9.7×10^5	119	41	-
B4	1	C	GaAs	GaAsP GB then jump back to GaAs	defect density purposefully increased from B1	1.7×10^6	335	23	-
B5	0.825	C	Si	SiGe GB		2.7×10^6	1880	1.6	-
B6	0.825	Zn	Si	SiGe GB		1.6×10^6	594	16	-
B7	0.825	Zn	Si	SiGe GB	intentionally mismatched InGaP emitter	3.4×10^6	897	9.2	-
B8	0.825	Zn	Si	SiGe GB	thick subcollector layer	3.7×10^6	< 1	-	158
B9	1	C	Si	2-step Ge		2.2×10^7	< 1	-	60

important details about each epitaxial structure, while the green text denotes experimental results, which will be discussed in Section 6.2.

Either C or Zn was used as the p-type dopant for the base region. We first used C because it of its lower diffusivity and capability of producing higher doping levels than Zn. Details of the use of C for heavy p doping of GaAsP are discussed extensively in Section 4.3. However, later samples were grown with Zn base doping. Use of Zn allows the entire device to be grown at one temperature (650 °C) rather than having to cool to 600 °C for the C-doped base region. This enables better control of lattice matching between the layers by reducing the number of discrete compositional calibrations from three (InGaP emitter at 650 °C; GaAsP base at 600 °C; GaAsP collector at 650 °C) to two (InGaP emitter at 650 °C; GaAsP base/collector at 650 °C). It also

eliminates the need for long growth pauses (2-3 minutes each) for temperature ramping and stabilization before and after the growth of the base layer. Contamination can accumulate at the surface during these growth pauses, so it is good to avoid them if possible.

Samples B1 through B4 were all grown on GaAs substrates. Sample B1 is a fully lattice-matched GaAs HBT control sample on a GaAs substrate, and is the same GaAs sample used in Chapter 5. Sample B2 is a $GaAs_{0.825}P$ HBT grown on a GaAs substrate via a slow tensile GaAsP graded buffer (0.2% strain/μm grade rate), and is the same as the $GaAs_{0.825}P$ sample from Chapter 5. Sample B3 is the same as B2, except with a faster grade rate (0.8% strain/μm) in the GaAsP graded buffer. This was done to intentionally increase the defect density for the purposes of understanding the impact of dislocations on device performance. Sample B4 also has an intentionally raised defect density for the same reason. Here, we graded from GaAs to $GaAs_{0.88}P$ with a 0.2% strain/μm grade rate, immediately jumped back to GaAs, and then grew a GaAs HBT structure, well-exceeding the critical thickness. This abrupt introduction of lattice mismatch and growth above the critical thickness introduced nucleation of a high number of dislocations.

Samples B5 through B9 were all grown on Si substrates. B5–B8 are all $GaAs_{0.825}P$ HBTs grown on GaAsP "virtual substrates." These utilize SiGe graded buffers to accommodate the lattice mismatch between the GaAsP device layers and the Si substrate. Details of the GaAsP virtual substrate growth and regrowth on the GaAsP virtual substrate can be found in Section 4.2. B5 uses C base doping just like B1-B4. Sample B6 uses Zn base doping for the reasons mentioned earlier. Sample B7 is the same as B6 but has an intentionally mismatched InGaP emitter layer (compositional shift of 2% In, corresponding to 0.15% strain). Sample B8 is also the same as B6, but includes a thicker subcollector layer grown at 650 °C (1100 nm). This allows for more complete relaxation of the $GaAs_{0.825}P$ subcollector before growth of the active device layers.

Sample B9 is a GaAs HBT grown on 2-step Ge on a Si substrate. More information about 2-step Ge is located in Section 1.3. The 2-step Ge on Si sample for this study was provided by Kwang Lee from the Singapore-MIT Alliance for Research and Technology (SMART). Details of the growth of the 2-step Ge can be found in Lee, et al. [87].

HBTs were fabricated in an annular geometry from the epitaxial structures using standard lithography and wet-etching techniques, as described in in Section 2.5. 60 μm diameter devices were electrically tested at 300 K using a Keysight B1500 semiconductor parameter analyzer.

6.2. Results

6.2.1. Epitaxial film characterization

EBIC was used to measure the threading dislocation density (ρ_{TD}) and misfit dislocation density in the active region (ρ_{MD}) for each sample. Representative EBIC images from B1, B6, and B9 are shown in Figure 6.1 (a), (b), and (c), respectively. In B1, there are no defects visible. In B6, there are both threading dislocations (black dots) and misfit dislocations (black lines) visible.

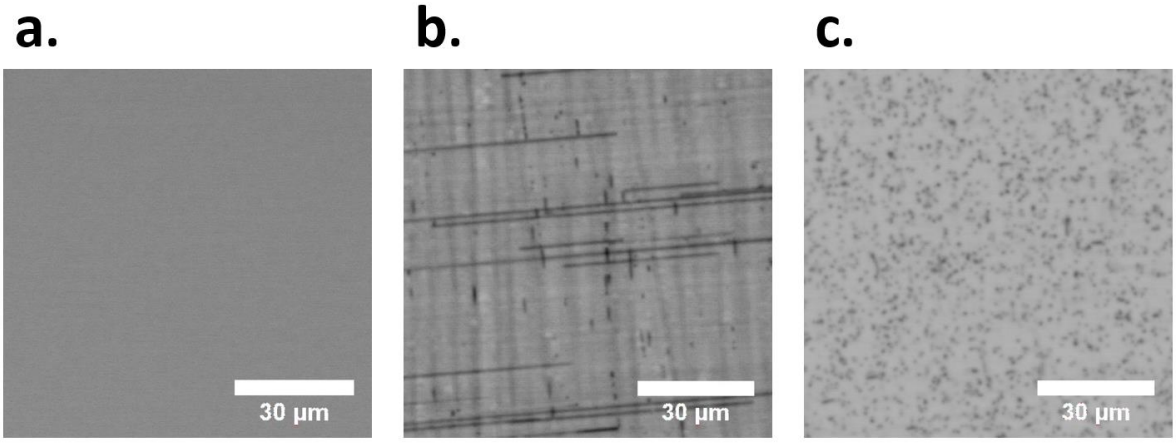

Figure 6.1: EBIC images from (a) B1, (b) B6, and (c) B9. Threading dislocations are visible in (b) and (c) and misfit dislocations are visible in (b).

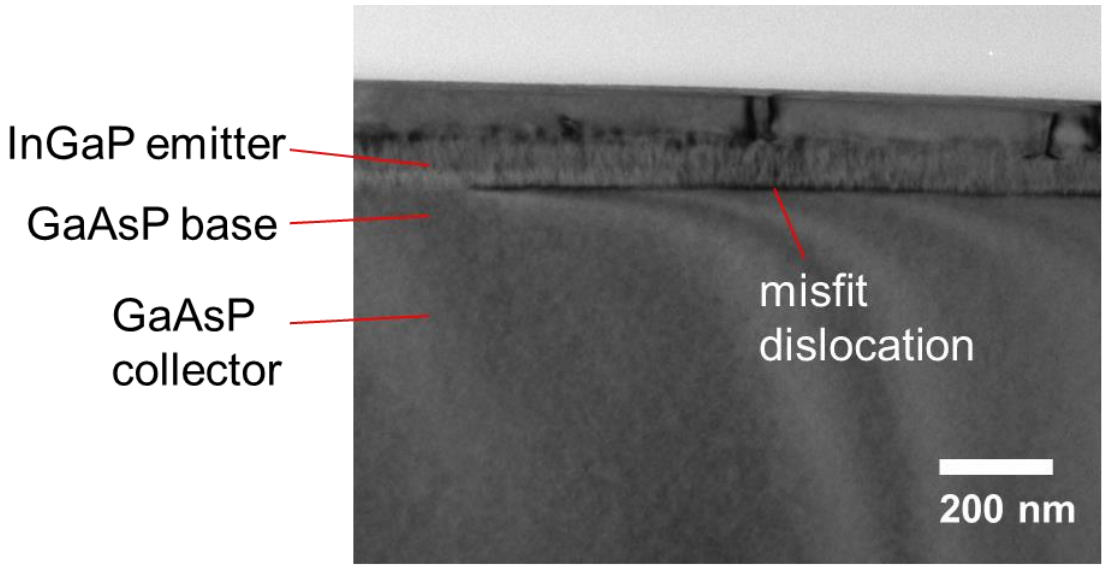

Figure 6.2: XTEM image of Sample B5, showing a misfit dislocation at the emitter-base interface. g = (220) and Acc. V = 200 kV.

While all of the lattice-mismatched samples have misfit dislocations in the buffer layers in order to accommodate their relaxation, the misfits detected by EBIC are in the active regions and therefor have a negative impact on device performance. The grey vertical features in B6 are due to surface roughness, known as "cross-hatch", induced during growth by strain fields surrounding underlying misfit dislocations [88]. In B9, the black dots are threading dislocations. To calculate the threading dislocation density (ρ_{TD}), the total number of threading dislocations in an image are divided by the area. To calculate the misfit dislocation density (ρ_{MD}), the total misfit dislocation length is divided by the image area. ρ_{TD} and ρ_{MD} are listed for all of the samples in Table 6.1.

XTEM of Sample B5 was used to measure film morphology, observe any dislocations in the film, and verify layer thickness (see Figure 6.2). While the thicknesses are on target, there is a misfit dislocation visible at the emitter-base interface. This is consistent with the high ρ_{MD} measured with EBIC for this sample.

6.2.2. DC characterization

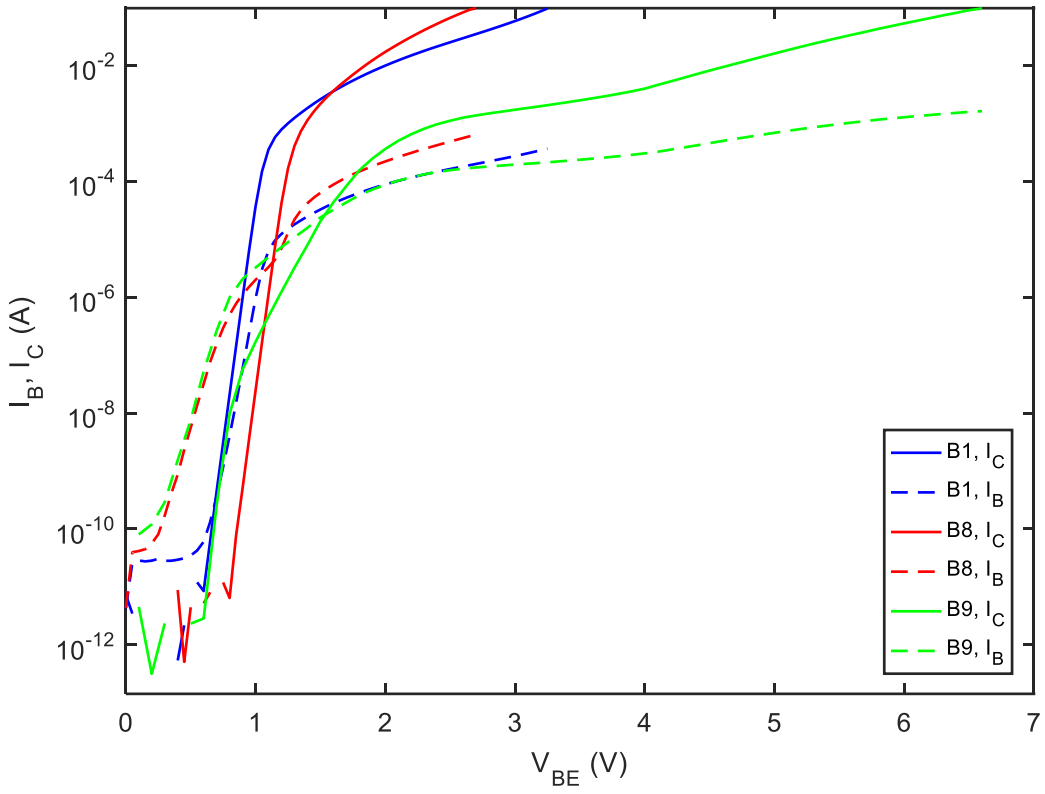

Figure 6.3: Gummel plots for Samples B1, B8, and B9. B9 exhibits a high series resistance.

Gummel characteristics were measured for all nine samples on devices with an emitter-base junction width of 60 µm. Gummel plots for Samples B1, B8, and B9 are plotted in Figure 6.3. Sample B1, a GaAs HBT on a GaAs substrate, serves as a control device with no measured dislocations. Sample B8 is the GaAsP sample on Si with the lowest defect density. Sample B9 is the GaAs sample on Si with the lowest defect density, and was also grown without the use of graded buffers. Samples B2–B4 are not shown here because they are on GaAs substrates, and B5–B7 are not shown because although they are on Si substrates, their performance is worse than that of B8.

The series resistance for Sample B9 is much higher than for B1 and B8. This is why when compared to B1, V_{BE} must be higher to yield the same I_C for I_C > ~10^{-7} A. The series resistance for B9 is higher because the Ge buffer is undoped, so the collector current is forced to travel laterally tens of microns through the relatively thin GaAs subcollector. The GaAs subcollector could not be made thicker because if the III-V layers are too thick, the film cracks due to CTE mismatch between the films and the substrate. The high series resistance exhibited in this case is not expected to impact performance in real applications because the contacts would be made much nearer to the device.

At low V_{BE} of less than 1–1.5 V, a significant leakage component is observed in I_B for the samples on Si substrates. This reduces β in this voltage range. It is unlikely that this leakage is from generation currents from trap states associated with threading dislocations near the emitter-base interface, as these currents have been shown to be many orders of magnitude lower than what is observed here [15]. It is therefore unclear what the mechanism of this leakage current is. However, because it is only present at lower V_{BE}, β can still be compared between devices at high V_{BE} without its effect.

The differences in series resistance between samples makes direct comparison of device performance (i.e. β) difficult from looking at the Gummel plot. To make this comparison easier, we plot β as a function of I_C in Figure 6.4. In this figure, we can get a better idea of the relative performance of B1, B8, and B9. The maximum current gain for B8 is 158, which is the highest current gain that we were able to achieve for a GaAs(P) HBT grown on Si. The maximum current gain for B9 is 60, which is the highest that we achieved for a III-V device grown on a Si substrate without the use of a graded buffer.

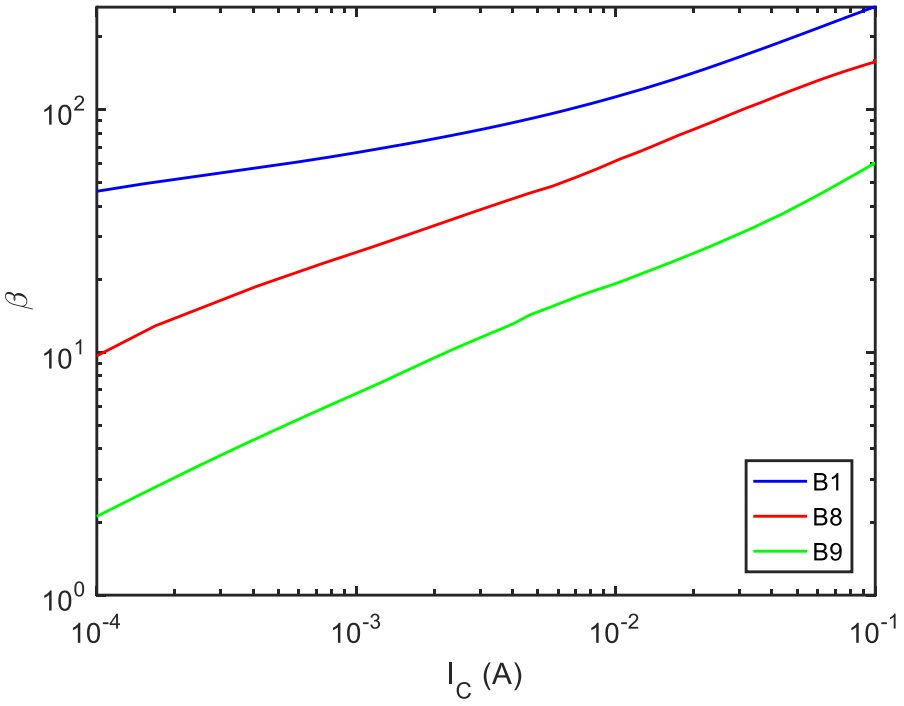

Figure 6.4: Current gain vs collector current for Samples B1, B8, and B9.

In order to quantitatively compare current gain for all of the samples, we list the current gain at singular values of I_C in Table 6.1. These values will be used in the next section for modelling the current gain as a function of defect density.

6.3. Discussion

6.3.1. Effect of threading dislocations on current gain

Threading dislocations are necessarily generated when lattice-mismatched layers are allowed to relax, as discussed in Section 1.2. Therefore, it is important to understand the effect of threading dislocations on device performance. Here, we will focus on how threading dislocation

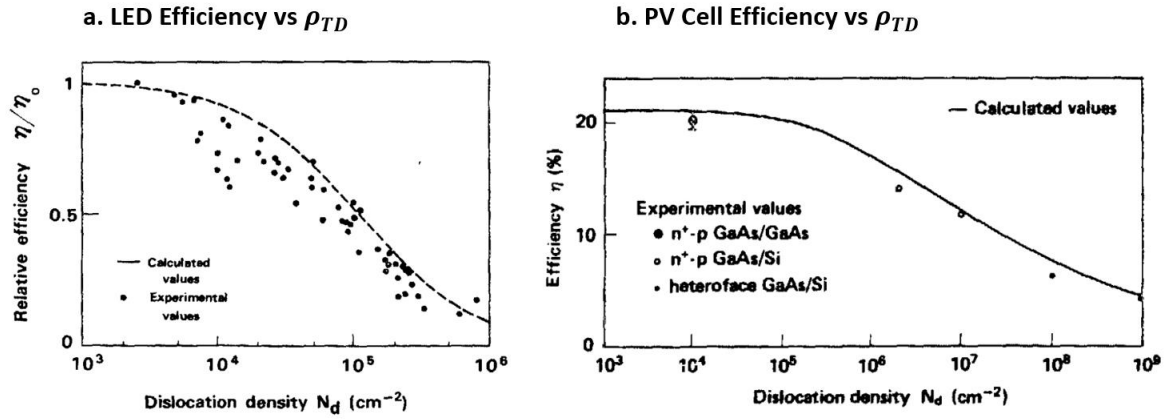

Figure 6.5: Model for the effect of threading dislocation density (called N_d here) on (a) GaAs/AlGaAs LED efficiency and (b) GaAs PV cell efficiency (reproduced, with permission, from Yamaguchi, et al., 1986 [89]*).*

density (ρ_{TD}) affects the common emitter current gain (β), as this is a critical device parameter that is highly sensitive to enhanced minority carrier recombination.

The effect of misfit dislocations in the active device layers will be discussed later in Section 6.3.2. However, it is important for now to note that misfit dislocations cause an comparatively larger reduction in β than threading dislocations. Therefore, in order to isolate the effect of ρ_{TD} on β, in this section we will consider only devices without any misfit dislocations detected in the active region. Of the samples described in Table 6.1, only samples B1, B8, and B9 will be considered here because they have negligible misfit dislocation densities (ρ_{MD}) as determined by EBIC measurements.

The effect of threading dislocations on GaAs- or GaAsP-based HBT performance can be estimated by a model for minority carrier lifetime vs threading dislocation density developed by Roedel, et al. for describing the efficiency of GaAs/AlGaAs LEDs grown on lattice-mismatched substrates [12]. This model was further developed by Yamaguchi, et al. for describing the efficiency and open-circuit voltage of GaAs photovoltaic cells in the presence of threading dislocations [13], [89]. In Figure 6.5, Yamaguchi's model is plotted alongside efficiency data for

Roedel's LEDs as well as for GaAs photovoltaic cells over a wide range of ρ_{TD}. The model and data are in reasonably good agreement, varying by a factor of 2–3 in ρ_{TD}. This model has been verified by others at different values of ρ_{TD} [90]. We can use this model as a basis for describing the effect of threading dislocations on HBT current gain.

The model starts by calculating the average distance that a given electron must diffuse in the p-type base to reach a threading dislocation at a given threading dislocation density ρ_{TD}:

$$L_{TD} = \frac{2}{\pi^{1.5}\rho_{TD}^{0.5}} \tag{6.1}$$

The details of the derivation of this equation can be found in Section II-A of Yamaguchi's 1996 paper. From this, we can calculate the overall electron diffusion length in the base region L:

$$L = \left(\frac{1}{L_0^2} + \frac{1}{L_{TD}^2}\right)^{-0.5}, \tag{6.2}$$

where L_0 is the diffusion length due to all other recombination mechanisms. Major contributors to L_0 in our HBTs include recombination in the quasi-neutral base region due to the heavy p-type doping, recombination at the emitter-base junction, and recombination at the device sidewalls [91]. From here, we calculate the minority carrier lifetime in the base in the presence of threading dislocations ($\tau_{n,TD}$):

$$\tau_{n,TD} = \frac{L^2}{D_{nB}} = \frac{1}{D_{nB}}\left(\frac{1}{L_0^2} + \frac{\pi^3\rho_{TD}}{4}\right)^{-1}, \tag{6.3}$$

where D_{nB} is the diffusivity of electrons in the base region.

Now, to calculate the current gain in the presence of threading dislocations (β_{TD}), we find the ratio of electrons which diffuse all the way across the base to those which recombine in the base. This is simply equal to $\tau_{n,TD}$ divided by the base transit time (τ_B) [56]:

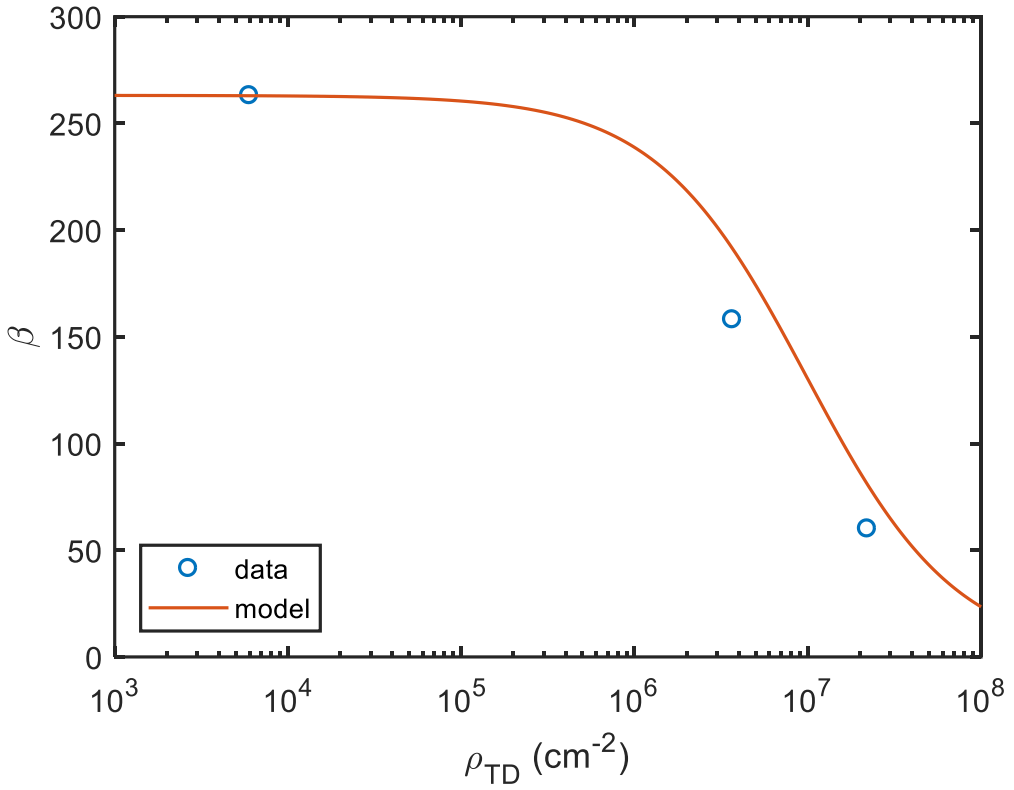

Figure 6.6: Effect of threading dislocation density on GaAs(P)/InGaP HBT current gain. In order to normalize for GaAsP composition and series resistance, β is taken when $I_C = 8 \times 10^{-1}$ A.

$$\beta_{TD} = \frac{\tau_{n,TD}}{\tau_B} \tag{6.4}$$

τ_B is the average time for an electron to transit the base region, and is a function of the quasi-neutral base width (x_B) and of $D_{n,B}$ [56]:

$$\tau_B = \frac{x_B^2}{2D_{nB}} \tag{6.5}$$

By combining equations 6.3–6.5, we can now find β_{MD} as a function of ρ_{TD} with only one unknown parameter (L_0). For our purposes, we will set L_0 such that the current gain in the case of no threading dislocations is predicted correctly. It should be noted that β_{MD} has no dependence on $D_{n,B}$.

β vs ρ_{TD} data for Samples B1, B8, and B9 is plotted in Figure 6.6, along with the model discussed above. The model slightly overpredicts β—the ρ_{TD} for a given value β is shifted by about a factor of 2. This error is similar in direction and magnitude to the errors exhibited in the model from Yamaguchi, et al. [13]. One source of error could be the estimation of the base thickness x_B (in this case assumed to be 100 nm). A larger base thickness (caused by diffusion of the base dopant into the emitter or collector) would cause a higher τ_B and therefore a lower β. Another source of error is our approximation that recombination only occurs at the dislocation core and not at any distance away from the dislocation. A final source of error is the estimation of the average distance between dislocations, which following Yamaguchi, et al., we set as $2/(\pi\rho_{TD})^{0.5}$ [89]. However, Roedel, et al. simply uses $1/\rho_{TD}^{0.5}$ for this distance [12], which yields a β vs ρ_{TD} model in closer alignment to our data.

6.3.2. Effect of misfit dislocations on current gain

Misfit dislocations in the active region of an HBT have a particularly detrimental effect on performance. While a threading dislocation pierces through the active layers near-perpendicularly on its way to the upper surface, a misfit dislocation can run along the active region for considerable distances. Therefore, for the same overall dislocation density (total dislocation length divided by volume), misfit dislocations have the potential to cause much greater carrier recombination in the base region. In any relaxed lattice-mismatched epitaxial structure, misfit dislocations are necessary to accommodate the plastic deformation of the mismatched layers. However, in general, it is desirable to ensure that these misfits are located sufficiently below (or above) the active layers such that they do not interfere with device performance.

In all but one of the GaAsP HBTs grown on Si substrates for this thesis, misfit dislocations were observed at one or both of the base-collector and emitter-base interfaces. These were seen in the EBIC images of Samples B2–B7 (e.g. in Figure 6.1b). Misfit dislocation density (ρ_{MD}) for these samples is listed in Table 6.1. We can see that the changes made in S6 and S8 (Zn base doping and a thick subcollector layer, respectively) caused improvements in ρ_{MD}, while the intentionally mismatched emitter layer of S7 caused a worsening of ρ_{MD}.

By considering samples B1–B7, we see a wide range of ρ_{MD} measured by EBIC. In addition, for B2–B7, the ρ_{TD} is low enough that it should not significantly affect β according to the model developed in the previous section. In other words, threading dislocations do not make a considerable contribution to minority carrier recombination in the base compared with recombination caused by misfit dislocations. In order to compare β from HBTs with different GaAsP compositions without seeing the effects of differing band gaps, β should be taken with constant I_C rather than constant V_{BE} [91]. In Figure 6.7, β is plotted at $I_C = 2 \times 10^{-4}$ A as a function of ρ_{MD}. We pick this current because it is in a regime where both I_C and I_B vary exponentially with V_{BE} for all of the devices. The first data point, with $\beta = 70$, was from Sample B1, which had no observable misfit dislocations. This was assigned an upper bound ρ_{MD} of 1 cm^{-1} based on the

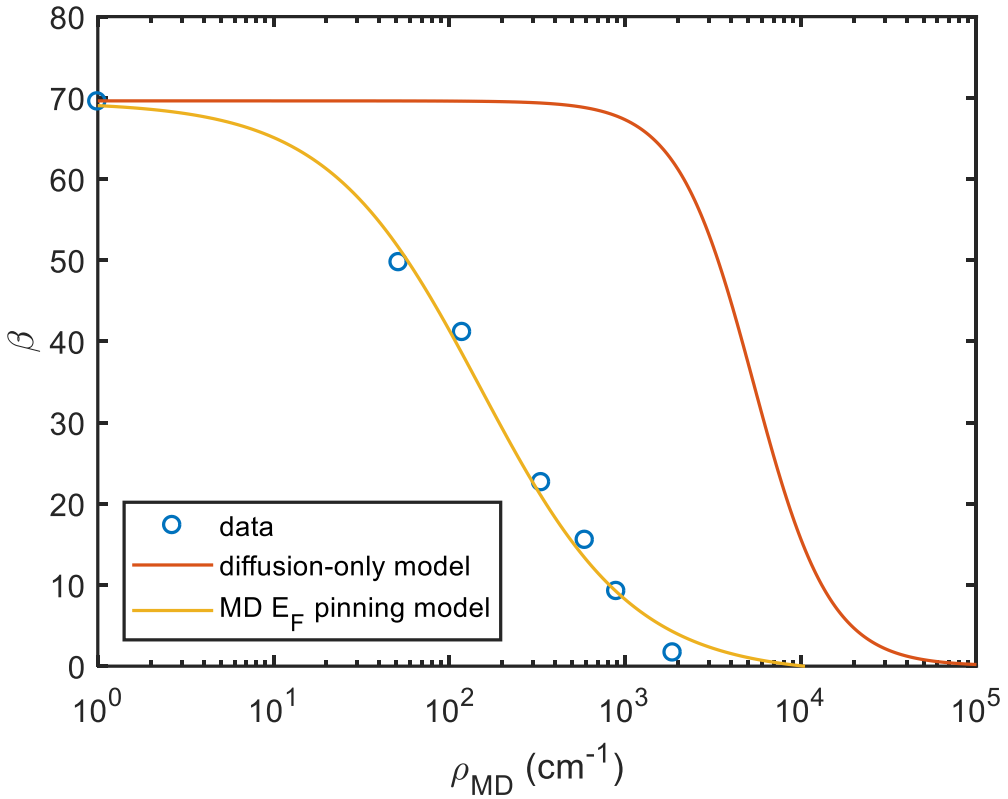

Figure 6.7: Effect of misfit dislocation density on GaAs(P)/InGaP HBT current gain. In order to normalize for GaAsP composition and series resistance, β is taken where I_C = 2 × 10^{-4} A.

area of the sample measured by EBIC. As expected, β decreases monotonically with increasing ρ_{MD}.

The model for the effect of threading dislocations on minority carrier lifetime described in Section 6.3.1 can be extended to apply to misfit dislocations as well. We will refer to this model as the diffusion-only model, because the additional component of I_B due to misfit dislocations is controlled by diffusion of electrons in the quasi-neutral base region. A schematic of this process is shown in Figure 6.8a. Electrons are injected from the emitter into the quasi-neutral base uniformly across the emitter-base junction, unaffected by the presence of dislocations. Similarly to the model for β_{TD} from Section 6.3.1, these electrons can then either diffuse to the collector, where they are captured as I_C, or to a misfit dislocation, where they recombine and contribute to

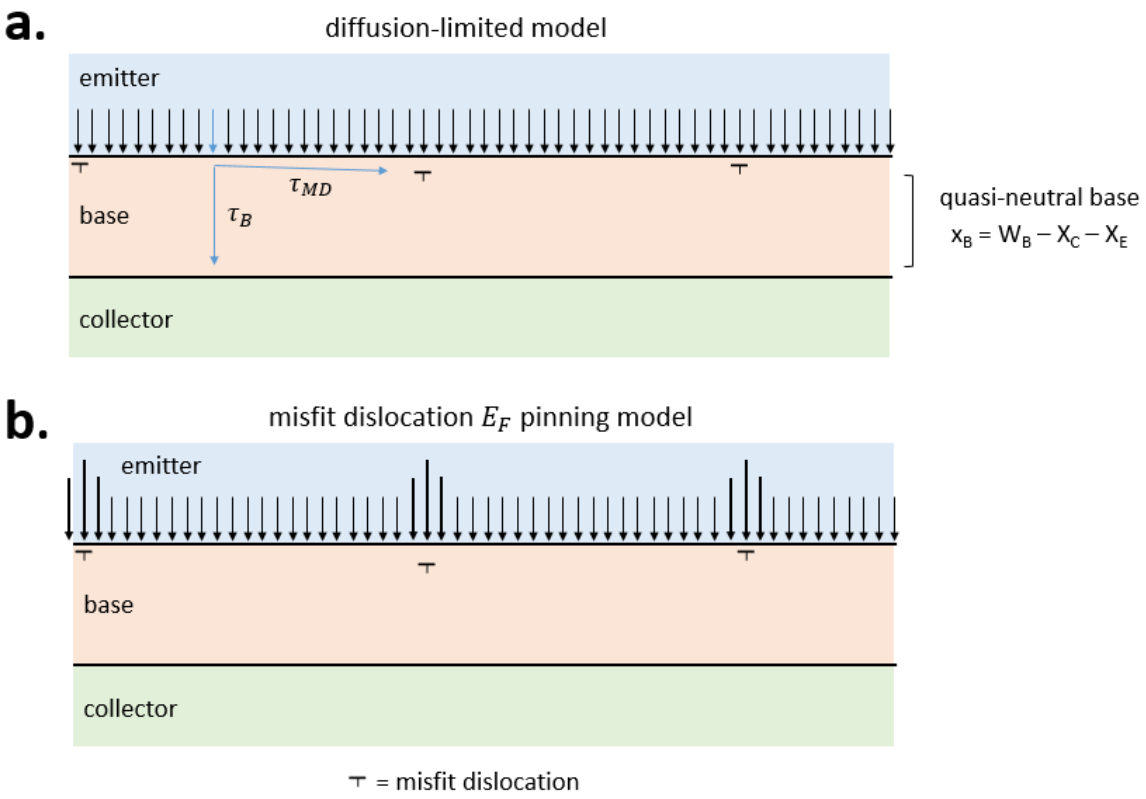

Figure 6.8: Schematics of two models for describing the effect of misfit dislocations on current gain: (a) diffusion-limited model and (b) misfit dislocation E_F pinning model.

I_B. The ratio of electrons which make it across the base to those which recombine is therefore equal to the minority carrier lifetime in the presence of misfit dislocations ($\tau_{n,MD}$) divided by the base transit time (τ_B). This ratio can be defined as the current gain in the presence of misfit dislocations (β_{MD}):

$$\beta_{MD} = \frac{\tau_{n,MD}}{\tau_B} \tag{6.6}$$

τ_B is the same as in the model for β_{TD} and is defined in Equation 6.5.

We can estimate $\tau_{n,MD}$ based on Equation 6.3 for $\tau_{n,TD}$. In the case of threading dislocations, the average distance between two dislocations is $2/(\pi\rho_{TD})^{0.5}$ [13]. However, for the case of misfit dislocations, the average distance between two misfits is simply $1/\rho_{MD}$. Therefore, by substitution into Equation 6.3, we derive an equation for $\tau_{n,MD}$ of:

$$\tau_{n,MD} = \frac{1}{D_{nB}}\left(\frac{1}{L_0^2} + \pi^2\rho_{MD}^2\right)^{-1} \quad 6.7$$

Like in Section 6.3.1, we will set L_0 such that the current gain for the HBT with no misfit dislocations is predicted correctly.

β_{MD} as predicted by the diffusion-only model is plotted in orange in Figure 6.7. This model fails in multiple respects. First, it overestimates the misfit dislocation density at which the current gain begins to fall by two orders of magnitude. Some of this error may be attributed to an error in the conversion of the model from accounting for threading dislocations to misfit dislocations, or more specifically, an error in the average distance an electron must travel to encounter a misfit dislocation and recombine. However, this distance should be on the order of $1/\rho_{MD}$, so the model should not be off by more than one order of magnitude.

A second failure of the diffusion-only model is that is predicts the wrong slope for β vs ρ_{MD} in the section where β is rapidly dropping. This stems from the fact that in this model, for higher values of ρ_{MD}, β is proportional to ρ_{MD}^{-2}. However, the data follows more closely with ρ_{MD}^{-1}. This suggests that the diffusion-only mechanism that governs the effect of threading dislocations on β does not also apply to misfit dislocations in the active region of the HBT.

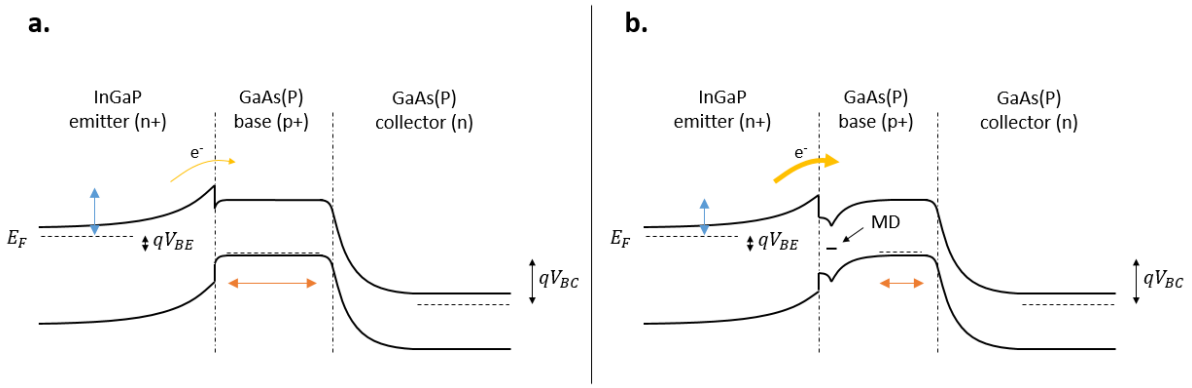

Figure 6.9: Schematic of an HBT band diagram (a) from a normal area and (b) near a misfit dislocation. The band bending around the dislocation in (b) causes a reduced barrier to electrons flowing from the emitter to the base (blue) and a thinner quasi-neutral base region (orange).

In place of the diffusion-only model, we hypothesize another model that we call the misfit dislocation (MD) E_F pinning model. Instead of being injected uniformly across the emitter-base interface, electrons are injected preferentially near the locations of misfit dislocations. Dislocations have mid-gap energy states that cause local pinning of the Fermi level (E_F) [92]. This causes band bending surrounding the dislocations [93], which in turn causes turn-on of the emitter-base junction at lower V_{BE} in these areas compared to areas far from dislocations. This effect is illustrated qualitatively in Figure 6.9. The band bending surrounding the dislocation in Figure 6.9b causes a reduction in the energy barrier for electrons moving from the emitter to the base (blue arrows). It also causes the quasi-neutral base thickness to decrease (orange arrows). Both of these effects can increase the rate of electron injection from the emitter into the base [91]. The extra electron current caused by this effect is injected in close proximity to the misfit dislocations, which act as recombination centers, and therefore largely recombines with majority holes in the base. This process is illustrated in Figure 6.8b.

In order to describe this phenomenon, we define $I_{B,MD}$ as the extra base current that is caused by E_F pinning near misfit dislocations. $I_{B,MD}$ is proportional to the total length of misfit dislocations in the area of the emitter-base junction:

$$I_{B,MD} = CI_{C,0}\rho_{MD} \tag{6.8}$$

The arbitrary proportionality constant C accounts for the area of the emitter-base junction and the amount of extra current per unit length of misfit dislocation. $I_{C,0}$ is the collector current in the absence of misfit dislocations. From this, we can calculate the overall β with misfit dislocations:

$$\beta_{MD} = \frac{I_C}{I_B} = \frac{I_{C,0} - I_{B,MD}}{I_{B,0} + I_{B,MD}}, \tag{6.9}$$

where $I_{B,0}$ is the base current in the absence of misfit dislocations. By dividing all terms in the numerator and denominator of Equation 6.9 by $I_{C,0}$, it can be written as a function of β_0 (current gain without misfit dislocations):

$$\beta_{MD} = (1 - C\rho_{MD})\left(\frac{1}{\beta_0} + C\rho_{MD}\right)^{-1} \tag{6.10}$$

In Figure 6.7, β_{MD} according to this model is plotted in yellow. β_0 is 70, taken from sample B1 which has no misfit dislocations. Adjusting C simply translates the curve horizontally without any other distortion. It is chosen such that the curve fits the data. The model is approximately proportional to ρ_{MD}^{-1} for higher values of ρ_{MD}, and therefore fits the data well. We therefore believe that the MD E_F pinning mechanism is the main factor determining the reduction in β due to misfit dislocations into the active layers.

It is interesting that the E_F pinning model applies to misfit dislocations while the diffusion-only model applies to threads. The increase in I_B associated with increased dislocation density in the diffusion-only model is inversely proportional to the spacing between dislocations, whether they are misfits or threads. However, the increase in I_B associated with the E_F pinning model is proportional to the total area of the emitter-base interface that is intersected by dislocations. Because the intersection of misfit dislocations with this interface is 1-dimensional while the intersection of threading dislocations with this interface is 0-dimensional, misfit dislocations have a much larger contribution given the same average spacing (see Figure 6.10). This causes the E_F pinning mechanism to be significantly weaker for threading dislocations than for misfit dislocations.

By combining the applicable models for the effects of threading and misfit dislocations on current gain and normalizing by β_0, we can create a contour map showing the relative effects of threads and misfits in a sample containing both. The map is shown in Figure 6.11. It is important

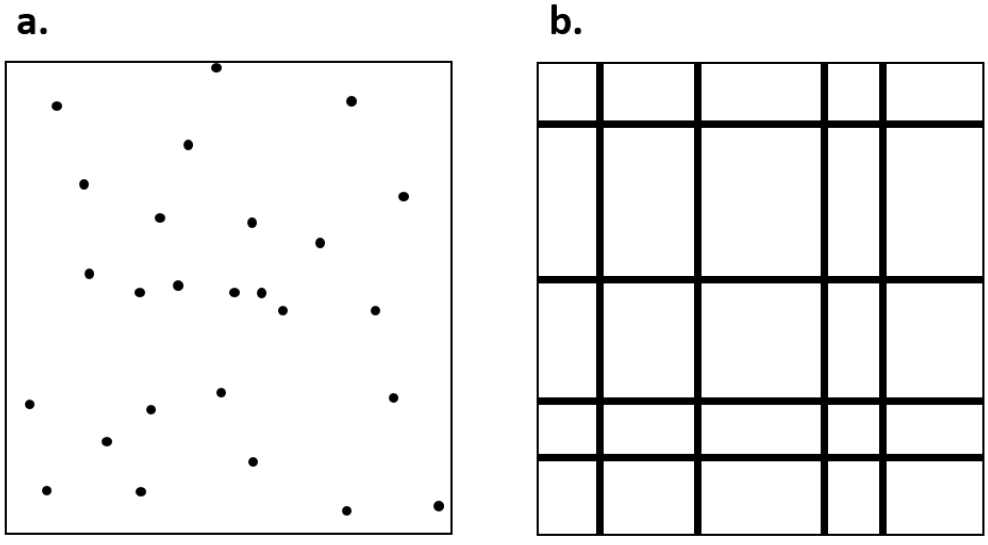

Figure 6.10: Intersection of (a) threading dislocations and (b) misfit dislocations with the emitter-base interface. For densities with a similar average spacing, misfit dislocations intersect a much larger effective area.

to note that this map is only valid for HBTs of this design, with the same base width and materials parameters governing recombination by other mechanisms.

Clearly, misfit dislocations in the active region have a pronounced effect on current gain. Even with an average spacing as large as 1 mm, misfit dislocations result in a 20% drop in β. This spacing of misfit dislocation is far too large to be observed via TEM. Therefore EBIC, which has a very large sampling area, is required for observing misfit dislocations at this density.

The following equation describes the average spacing between misfit dislocations (S) below a relaxed lattice-mismatched film:

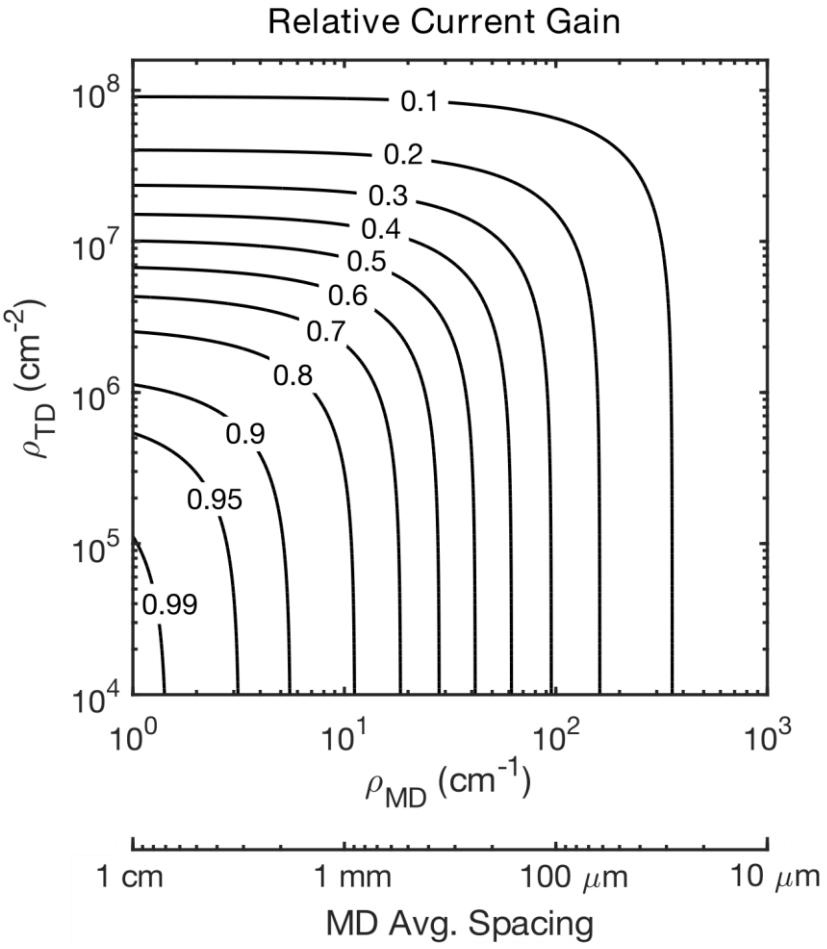

Figure 6.11: Map of relative current gain (β/β_0) as a function of misfit dislocation density and threading dislocation density.

$$S = \frac{b_{eff}}{\delta}, \tag{6.11}$$

where b_{eff} is the effective Burger's vector (in-plane component of $\boldsymbol{b}$ in the direction of spacing S) and δ is the plastic strain [10]. A 1 mm average misfit spacing corresponds with an extremely small plastic strain ($\delta = 1.4 \times 10^{-7}$). This amount of relaxation is too small to detect using XRD. Therefore, the base and emitter films must not be allowed to relax via plastic deformation at all during growth if β is to be maintained.

Chapter 7: Conclusions

7.1. Summary of results

In this thesis, we have explored a route towards monolithic integration of III-V microelectronics onto a Si substrate. GaAsP/InGaP HBT structures were grown on Si with defect densities in a range that allowed for fabrication of working transistors. This involved the growth of several complex and sensitive structures, including: a SiGe graded buffer from Si to $SiGe_{0.50}$ grown via UHVCVD, regrowth of $SiGe_{0.50}/SiGe_{0.50}$ after CMP of the initial graded buffer, growth of the heterovalent GaAsP on SiGe interface, and finally growth of a GaAsP/InGaP HBT structure. We studied the C doping of GaAsP for the heavily p-type base region of the HBT and characterized the interactions between C doping level and GaAsP composition. C-doped GaAsP with hole concentrations greater than 2×10^{19} cm^{-3} were demonstrated across a range of compositions using a $CBrCl_3$ flow rate of 59 μmol/min. At higher flow rates, the C incorporation and hole concentration saturated.

GaAsP/InGaP HBTs were grown on GaAs substrates in order to study the effect of GaAsP composition independent of any defects that might be present from growing the structures on Si substrates. $GaAs_xP_{1-x}$ active layers with x ranging from 1 to 0.825 were demonstrated. GaAsP composition was shown to have minimal effect on current gain in this range. For low current densities and smaller device sizes, I_B is dominated by SCR recombination at the perimeter of the device. For higher current densities and larger device sizes, I_B is dominated by a combination of QNR and SCR recombination across the entire area of the emitter-base junction. I_C was shown not to be limited by thermionic emission over a barrier formed by the conduction band

discontinuity between the InGaP emitter and the GaAsP base. Modeling I_C as diffusion of electrons across the quasi-neutral base agrees well with measurements for the GaAs/InGaP HBT, but underestimates measured I_C for all GaAsP/InGaP devices by almost 10 times. Therefore, GaAsP/InGaP HBTs have a higher transconductance than what was predicted based solely on known materials parameters. Breakdown voltage was measured for GaAsP HBTs and p-n junctions as a function of composition. GaAsP breakdown voltage was confirmed to be higher than that of GaAs, with an impact ionization energy equal to at least $1.5E_g$.

Finally, GaAsP/InGaP HBTs were grown on Si substrates via SiGe graded buffers with varying epitaxial structures and were fabricated and tested. Threading and misfit dislocation densities were measured by EBIC. Understanding the exact nature of defects in III-V HBTs on Si substrates is necessary to engineer successful devices. The differences in base dopant, InGaP composition, subcollector thickness, and grade rate yielded samples with a wide range of threading and misfit dislocation densities. The effect of threading dislocation density on current gain was modeled based on an earlier model developed for describing the effect of threading dislocation density on LED and photovoltaic cell efficiency. A different model was developed to describe the effect of misfit dislocations in the active region on current gain involving Fermi level pinning near dislocations, which causes an increase in injected current in their vicinity. By using Zn base doping instead of C and by inserting a thick GaAsP subcollector layer between the graded buffer and device layers, misfit dislocations were eliminated from the active device layers. This allowed for the demonstration of a $GaAs_{0.825}P$/InGaP HBT on Si with a current gain of 158. This current gain is high enough for use in many applications, demonstrating our GaAsP on Si growth via SiGe graded buffer as a viable route towards III-V microelectronics integration with Si.

7.2. Recommendations for future work

We have identified several areas as candidates for future work. The first three involve optimization of the GaAsP HBT structure as grown on Si. The fourth involves measuring AC electrical characteristics of the GaAsP HBTs on Si. Lastly, the fifth involves measuring materials properties of GaAsP/InGaP heterojunctions.

1. In order to use the GaAsP/InGaP devices presented in this thesis in any kind of application, the contact resistances will have to be improved. First, a heavily doped InGaAs cap layer can be added to the emitter. This way, the contact will be made to a semiconductor with a lower E_g, which should reduce the contact resistance. This was omitted from the devices in this work for simplicity, and to reduce the likelihood of dislocations from the mismatched InGaAs cap from punching down into the InGaP emitter. However, future lattice-mismatch engineering can eliminate misfit injection from lower-barrier contacts such as InGaAs. The emitter contact resistance can also be improved by using a different contact type for the n-type emitter contact, such as an alloyed Ge/Ni/Au contact [85]. In addition, contact resistance of the base contact must be reduced. This might be done by using a different metal stack for the contact or by exploring different pre-treatments of the semiconductor surface before metal deposition such as ashing, UV-ozone exposure, or wet chemical treatment.
2. The GaAsP HBTs grown on Si exhibit a leakage component of I_B at low V_{BE}. This significantly reduces the current gain at low current densities. We do not know the mechanism of this leakage current as of now. One possibility is E_F pinning from threading dislocations, made stronger by dopant segregation in the heavily doped base region. A

better understanding of what physically causes the leakage component would be a good first step towards trying to eliminate it.

3. GaAs HBTs on GaAs substrates grown with heavy C doping (i.e. greater than 1×10^{19} cm^{-3}) had very low current gains of 10 or less. This forced us to make all further HBTs either with relatively low C base doping, or with Zn doping, to get acceptably high current gains. Having a GaAs control device with high current gain was necessary in order to study the full effect of substrate and GaAsP composition on current gain. Based on SIMS data from our C-doped GaAs HBT structure, the C doping is almost fully activated. We also do not see any evidence of Cl or Br incorporation into the film as byproducts of the C precursor. C doping of greater than 1×10^{19} cm^{-2} has been shown many times in the literature to yield working HBTs with high gain [30]–[32]. Therefore, it seems that there is a specific issue with growth or processing that causes heavily C-doped GaAs to have an anomalously low carrier lifetime. Using a lower base doping was fine for our HBTs that were only subjected to DC measurements. However, AC performance improves greatly with higher base doping due to lower base series resistance. It would therefore be important to find out why heavy C doping causes low current gain in order to fabricate devices optimized for AC performance.
4. For HBTs to be useful, they must be able to operate well at high frequency. Measuring current gain and breakdown voltage are important first steps, but it will be necessary to measure the AC behavior of the GaAsP/InGaP HBTs integrated on Si substrates. Parameters such as f_T (cut-off frequency) and f_{max} could be effected by dislocations resulting from the mismatched epitaxy. These parameters were not measured in this work

because they require a highly optimized device geometry and fabrication flow in order to reduce parasitic resistances, inductances, and capacitances.

5. The band alignment of GaAsP with InGaP has implications for the engineering of any semiconductor device using these two materials. For HBTs, the valence band offset dictates the maximum current gain. For HEMTs, the conduction band offset sets the sheet carrier concentration of the 2-D electron gas, a critical parameter for device operation. Currently, the best estimates for the band offsets between GaAsP and InGaP are interpolations between the binary endpoints [60]. In addition, the band alignment of GaAs/$In_{0.5}Ga_{0.5}P$ has many conflicting values in the literature, likely because the band structure of InGaP changes with growth conditions due to ordering and phase separation [34], [94]–[97]. We would therefore recommend a careful study of GaAsP/InGaP band alignment as a function of composition and InGaP growth conditions.

References

[1] I. Thayne, “Advanced III-V HEMTs,” *III-Vs Rev.*, vol. 16, no. 8, pp. 48–51, 2003.

[2] J. Hudgins, G. Simin, E. Santi, and M. Khan, “An assessment of wide bandgap semiconductors for power devices,” *IEEE Trans. Power Electron.*, vol. 18, no. 3, pp. 907–914, 2003.

[3] K. H. Lee, S. Bao, E. Fitzgerald, and C. S. Tan, “Integration of III – V materials and Si-CMOS through double layer transfer process,” *Jpn. J. Appl. Phys.*, vol. 54, no. 30209, pp. 1–5, 2015.

[4] E. Fitzgerald, M. T. Bulsara, Y. Bai, C. Cheng, W. K. Liu, D. Lubyshev, J. Fastenau, Y. Wu, M. Urtega, W. Ha, J. Bergman, B. Briar, C. Drazek, N. Daval, F. Letertre, W. E. Hoke, J. R. LaRoche, K. J. Herrick, and T. E. Kazior, “Monolithic III-V/Si Integration,” *ECS Trans.*, vol. 19, no. 5, pp. 345–350, 2009.

[5] P. Chen, Y. Jing, S. S. Lau, D. Xu, L. Mawst, T. L. Alford, C. Paulson, and T. F. Kuech, “High crystalline-quality III-V layer transfer onto Si substrate,” *Appl. Phys. Lett.*, vol. 92, no. 9, p. 92107, 2008.

[6] E. A. Fitzgerald, P. Sharma, M. T. Bulsara, T. Milakovich, S. A. Ringel, A. J. Pitera, J. Hennessy, and A. C. Malonis, “Materials Integration for III-V/SiGe+CMOS Integrated Circuit Platforms,” in *The Electrochemical Society*, 2012.

[7] J. Carlin, S. Ringel, E. Fitzgerald, and M. Bulsara, “High quality GaAs qrowth by MBE on Si using GeSi buffers and prospects for space photovoltaics,” *Prog. Photovoltaics Res. Appl.*, vol. 8, pp. 323–332, 2000.

[8] W. Liu, *Fundamentals of III-V Devices*. John Wiley & Sons, 1999.

[9] J. Mathews and A. Blakeslee, “Defects in epitaxial multilayers,” *J. Cryst. Growth*, vol. 27,

pp. 118–125, 1974.

[10] E. A. Fitzgerald, “Dislocations in strained-layer epitaxy: theory, experiment, and applications,” *Mater. Sci. Reports*, vol. 7, no. 3, pp. 87–142, 1991.

[11] K. Mukherjee, “Lattice-Mismatched Epitaxy of AlInP and Characterization of its Microstructure and Luminescence,” MIT, 2014.

[12] R. J. Roedel, A. R. Von Neida, R. Caruso, and L. R. Dawson, “The Effect of Dislocations in GaAlAs:Si Light-Emitting Diodes,” *J. Electrochem. Soc.*, vol. 126, no. 15, p. 637, 1979.

[13] M. Yamaguchi and C. Amano, “Efficiency calculations of thin-film GaAs solar cells on Si substrates,” *J. Appl. Phys.*, vol. 58, no. 9, pp. 3601–3606, 1985.

[14] T. Ohori, S. Ohkubo, K. Kasai, and J. Komeno, “Effect of threading dislocations on mobility in selectively doped heterostructures grown on Si substrates,” *J. Appl. Phys.*, vol. 75, no. 7, p. 3681, 1994.

[15] L. M. Giovane, H. C. Luan, A. M. Agarwal, and L. C. Kimerling, “Correlation between leakage current density and threading dislocation density in SiGe p-i-n diodes grown on relaxed graded buffer layers,” *Appl. Phys. Lett.*, vol. 78, no. 4, pp. 541–543, 2001.

[16] N. H. Cho, B. C. De Cooman, C. B. Carter, R. Fletcher, and D. K. Wagner, “Antiphase boundaries in GaAs,” *Appl. Phys. Lett.*, vol. 47, no. 8, pp. 879–881, 1985.

[17] H. Kroemer, “Polar-on-nonpolar epitaxy,” *J. Cryst. Growth*, vol. 81, no. 1–4, pp. 193–204, 1987.

[18] D. B. Holt, “Antiphase boundaries in semiconducting compounds,” *J. Phys. Chem. Solids*, vol. 30, no. 6, pp. 1297–1308, 1969.

[19] P. M. Petroff, “Nucleation and growth of GaAs on Ge and the structure of antiphase

boundaries," *J. Vac. Sci. Technol.*, vol. 874, no. 4, pp. 874–877, 1986.

[20] Y. Li, G. Salviati, M. M. G. Bongers, L. Lazzarini, L. Nasi, and L. J. Giling, "On the formation of antiphase domains in the system of GaAs on Ge," *J. Cryst. Growth*, vol. 163, no. 3, pp. 195–202, 1996.

[21] S. M. Ting and E. A. Fitzgerald, "Metal-organic chemical vapor deposition of single domain GaAs on Ge/GeSi/Si and Ge substrates," *J. Appl. Phys.*, vol. 87, no. 5, p. 2618, 2000.

[22] P. Sharma, T. Milakovich, M. Bulsara, and E. Fitzgerald, "Controlling Epitaxial GaAsP/SiGe Heterovalent Interfaces," in *The Electrochemical Society*, 2012, vol. 4763, no. 2008, p. 4763.

[23] E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y.-J. Mii, and B. E. Weir, "Totally relaxed GeSi layers with low threading dislocation densities grown on Si substrates," *Appl. Phys. Lett.*, vol. 59, no. 7, p. 811, 1991.

[24] C. W. Leitz, "High mobility strained Si/SiGe heterostructure MOSFETs: channel engineering and virtual substrate optimization," 2002.

[25] T. Milakovich, R. Shah, S. Hadi, M. Bulsara, and A. Nayfeh, "Growth and characterization of GaAsP top cells for high efficiency III-V / Si tandem PV," in *Photovoltaic Specialist Conference (PVSC), 2015 IEEE 42nd*, 2015, pp. 1–4.

[26] M. T. Currie, S. B. Samavedam, T. A. Langdo, C. W. Leitz, and E. A. Fitzgerald, "Controlling threading dislocation densities in Ge on Si using graded SiGe layers and chemical-mechanical polishing," *Appl. Phys. Lett.*, vol. 72, no. 14, pp. 1718–1720, 1998.

[27] H.-C. Luan, D. R. Lim, K. K. Lee, K. M. Chen, J. G. Sandland, K. Wada, and L. C. Kimerling, "High-quality Ge epilayers on Si with low threading-dislocation densities,"

Appl. Phys. Lett., vol. 75, no. 19, p. 2909, 1999.

[28] J. M. Hartmann, A. Abbadie, A. M. Papon, P. Holliger, G. Rolland, T. Billon, J. M. Fédéli, M. Rouvière, L. Vivien, and S. Laval, "Reduced pressure-chemical vapor deposition of Ge thick layers on Si(001) for 1.3-1.55-μm photodetection," *J. Appl. Phys.*, vol. 95, no. 10, pp. 5905–5913, 2004.

[29] A. E. Romanov, W. Pompe, S. Mathis, G. E. Beltz, and J. S. Speck, "Threading dislocation reduction in strained layers," *J. Appl. Phys.*, vol. 85, no. 1, pp. 182–192, 1999.

[30] C. R. Abernathy, F. Ren, P. W. Wisk, S. J. Pearton, and R. Esagui, "Improved performance of carbon-doped GaAs base heterojunction bipolar transistors through the use of InGaP," *Appl. Phys. Lett.*, vol. 61, no. 9, p. 1092, 1992.

[31] W. Liu and S. K. Fan, "Near-Ideal I-V Characteristics of GaInP/GaAs Heterojunction Bipolar Transistors," *IEEE Electron Device Lett.*, vol. 13, no. 10, pp. 510–512, 1992.

[32] Y. S. Lin and J. J. Jiang, "Temperature dependence of current gain, ideality factor, and offset voltage of AlGaAs/GaAs and InGaP/GaAs HBTs," *IEEE Trans. Electron Devices*, vol. 56, no. 12, pp. 2945–2951, 2009.

[33] C. C. Hsu and E. S. Yang, "Surface Recombination Current in InGaP/GaAs Heterostructure-Emitter Bipolar Transistors," *IEEE Trans. Electron Devices*, vol. 41, no. 5, pp. 643–647, 1994.

[34] T. Kobayashi, K. Taira, F. Nakamura, and H. Kawai, "Band lineup for a GaInP/GaAs heterojunction measured by a high-gain Npn heterojunction bipolar transistor grown by metalorganic chemical vapor deposition," *J. Appl. Phys.*, vol. 65, no. 12, pp. 4898–4902, 1989.

[35] F. H. Chau, B. J. Lin, Y. Chen, M. Kretschmar, C. Lee, N. L. Wang, X. Sun, W. Ma, S.

Xu, and P. Hu, “Reliability Study of InGaP / GaAs HBT for 28V Operation,” pp. 191–194, 2006.

[36] S. G. Thomas, E. S. Johnson, C. Tracy, P. Maniar, X. Li, B. Roof, Q. Hartmann, and D. A. Ahmari, “Fabrication and Characterization of InGaP / GaAs Heterojunction Bipolar Transistors on GOI Substrates,” *IEEE Electron Device Lett.*, vol. 26, no. 7, pp. 438–440, 2005.

[37] J. J. Tietjen and L. R. Weisberg, “Electron Mobility in GaAsP Alloys,” *Appl. Phys. Lett.*, vol. 7, no. 10, pp. 261–263, 1965.

[38] V. K. Yang, M. Groenert, C. W. Leitz, A. J. Pitera, M. T. Currie, and E. A. Fitzgerald, “Crack formation in GaAs heteroepitaxial films on Si and SiGe virtual substrates,” *J. Appl. Phys.*, vol. 93, no. 7, pp. 3859–3865, 2003.

[39] S. Birner, T. Zibold, T. Andlauer, T. Kubis, M. Sabathil, A. Trellakis, and P. Vogl, “Nextnano: General purpose 3-D simulations,” *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2137–2142, 2007.

[40] C. L. Andre, N. El-Zein, and N. Tran, “Bubbler for constant vapor delivery of a solid chemical,” *J. Cryst. Growth*, vol. 298, no. SPEC. ISS, pp. 168–171, 2007.

[41] G. B. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practice*, Second Edi. San Diego, California: Academic Press, 1999.

[42] G. B. Stringfellow, “Thermodynamic aspects of OMVPE,” *J. Cryst. Growth*, vol. 70, no. 1–2, pp. 133–139, 1984.

[43] M. Ohring, *Materials Science of Thin Films*. Academic Press, 2002.

[44] E. T. J. M. Smeets, “Solid Composition of GaAsP Grown by Organometallic Vapour Phase Epitaxy,” *J. Cryst. Growth*, vol. 82, pp. 385–395, 1987.

[45] M. J. Mori and E. A. Fitzgerald, “Microstructure and luminescent properties of novel InGaP alloys on relaxed GaAsP substrates,” *J. Appl. Phys.*, vol. 105, no. 1, p. 13107, 2009.

[46] T. Roesener, V. Klinger, C. Weuffen, D. Lackner, and F. Dimroth, “Determination of heteroepitaxial layer relaxation at growth temperature from room temperature X-ray reciprocal space maps,” *J. Cryst. Growth*, vol. 368, pp. 21–28, Apr. 2013.

[47] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy*. Springer US, 2009.

[48] L. Li-lung and W. Min, “Investigation and Application of SEM dopant contrast on cross-section,” in *Proceedings of 12th IPFA*, 2005, no. 18, pp. 230–233.

[49] R. B. Heimann, *Crystals: Growth, Properties, and Applications - Silicon Chemical Etching*. Springer-Verlag, 1982.

[50] D. G. Schimmel, “Defect Etch for Silicon Evaluation,” *J. Electrochem. Soc.*, vol. 126, no. 3, p. 479, 1979.

[51] J. Lai, “Etch-Pit Density Survey of Germanium Silicon Graded Alloys on Silicon,” MIT, 1999.

[52] L. van der Pauw, “A method of measuring specific resistivity and Hall effect of discs of arbitrary shape,” *Philips Res. Reports*, vol. 13, no. 1, pp. 1–9, 1958.

[53] R. Chwang, B. J. Smith, and C. R. Crowell, “Contact Size Effects on the Van Der Pauw Method for Resistivity and Hall Coefficient Measurement,” *Solid. State. Electron.*, vol. 17, no. 12, pp. 1217–1227, 1974.

[54] W. S. Hobson, Y. K. Chen, and M. C. Wu, “InGaAs/AlGaAs ridge waveguide lasers utilizing an InGaP etch-stop layer,” *Semicond. Sci. Technol.*, vol. 7, pp. 1425–1427, 1992.

[55] M. Razeghi, F. Omnes, P. Maurel, Y. J. Chan, and D. Pavlidis, “Ga0.51In0.49P/GaxIn1-

xAs lattice-matched (x=1) and strained (x=0.85) two-dimensional electron gas field-effect transistors," *Semicond. Sci. Technol.*, vol. 6, pp. 103–107, 1991.

[56] W. Liu, *Handbook of III-V heterojunction bipolar transistors*. John Wiley & Sons, 1998.

[57] J. S. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors*. John Wiley & Sons, 1999.

[58] J. A. del Alamo, *Integrated Microelectronic Devices: Physics and Modeling*. Cambridge, MA, 2014.

[59] T. Shitara and K. Eberl, "Electronic properties of InGaP grown by solid-source molecular-beam epitaxy with a GaP decomposition source," *Appl. Phys. Lett.*, vol. 65, no. 3, pp. 356–358, 1994.

[60] S. Tiwari and D. J. Frank, "Empirical fit to band discontinuities and barrier heights in III–V alloy systems," *Appl. Phys. Lett.*, vol. 60, no. 5, p. 630, 1992.

[61] F. Cleton, B. Sieber, A. Lefebvre, A. Bensaada, R. A. Masut, J. M. Bonard, J. D. Ganiere, and M. Ambri, "Transmission electron microscopy and cathodoluminescence of tensile-strained GaxIn1-xP/InP heterostructures .1. Spatial variations of the tensile stress relaxation," *J. Appl. Phys.* , vol. 80, no. 2, pp. 827–836, 1996.

[62] F. Cleton, B. Sieber, A. Bensaada, R. A. Masut, J.M. Bonard, and J.D. Ganiere, "Transmission electron microscopy and cathodoluminescence of tensile- strained GaInP/InP heterostructures. II. On the origin of luminescence heterogeneities in tensile stress relaxed GaInP/InP x x heterostructures," *J. Appl. Phys.*, vol. 80, no. 2, pp. 837–845, 1996.

[63] M. J. Mori, S. T. Boles, and E. A. Fitzgerald, "Comparison of compressive and tensile relaxed composition-graded GaAsP and (Al)InGaP substrates," *J. Vac. Sci. Technol. A*

Vacuum, Surfaces, Film., vol. 28, no. 2, pp. 182–188, 2010.

[64] S. Tomasulo, K. N. Yaung, and M. L. Lee, "Metamorphic GaAsP and InGaP solar cells on GaAs," *IEEE J. Photovoltaics*, vol. 2, no. 1, pp. 56–61, 2012.

[65] P. Sharma, T. Milakovich, M. T. Bulsara, and E. A. Fitzgerald, "Controlling Epitaxial GaAsP/SiGe Heterovalent Interfaces," *Trans. E C S Soc. Electrochem.*, vol. 50, no. 9, pp. 333–337, 2012.

[66] P. Sharma, "Integration of GaAsP Alloys on SiGe Virtual Substrates for Si- based Dual-Junction Solar Cells," 2013.

[67] T. Milakovich, "Integration of GaAsP Alloys on Si for High-Efficiency III-V/Si PV," pp. 1–187, 2015.

[68] S. B. Samavedam and E. A. Fitzgerald, "Novel dislocation structure and surface morphology effects in relaxed Ge/Si-Ge(graded)/Si structures," *J. Appl. Phys.*, vol. 81, no. 7, p. 3108, 1997.

[69] B. T. Cunningham, L. J. Guido, J. E. Baker, J. S. M. Jr., N. H. Jr., and G. E. Stillman, "Carbon diffusion in undoped, n-type, and p-type GaAs," *Appl. Phys. Lett.*, vol. 55, no. 1989, pp. 687–689, 1989.

[70] M. Konagai, T. Yamada, T. Akatsuka, K. Saito, E. Tokumitsu, and K. Takahashi, "Metallic p-type GaAs and GaAlAs grown by metalorganic molecular beam epitaxy," *J. Cryst. Growth*, vol. 98, pp. 167–173, 1989.

[71] H. Zhu, Y. Adachi, and T. Ikoma, "Deep levels in MOCVD GaAs grown under different Ga/As mol fractions," *J. Cryst. Growth*, vol. 55, pp. 154–163, 1981.

[72] M. Kushibe, K. Eguchi, M. Funamizu, and Y. Ohba, "Heavy carbon doping in metalorganic chemical vapor deposition for GaAs using a low V/III ratio," *Appl. Phys.*

Lett., vol. 56, no. 13, p. 1248, 1990.

[73] B. T. Cunningham, M. A. Haase, M. J. McCollum, J. E. Baker, and G. E. Stillman, "Heavy carbon doping of metalorganic chemical vapor deposition grown GaAs using carbon tetrachloride," *Appl. Phys. Lett.*, vol. 54, no. 19, pp. 1905–1907, 1989.

[74] J.-S. Lee, I. Kim, B.-D. Choe, and W. G. Jeong, "Carbon doping and growth rate reduction by CCl4 during metalorganic chemical-vapor deposition of GaAs," *J. Appl. Phys.*, vol. 76, no. 9, pp. 5079–5084, 1994.

[75] K. Tateno and C. Amano, "Carbon doping and etching in GaInAsP on GaAs substrates using CBr4 by metalorganic chemical vapor deposition," *J. Electron. Mater.*, vol. 28, no. 1, pp. 63–68, 1999.

[76] Y. Bolkhovityanov and O. Pchelyakov, "GaAs epitaxy on Si substrates: modern status of research and engineering," *Physics-Uspekhi*, vol. 51, no. 5, pp. 437–456, 2008.

[77] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III–V compound semiconductors and their alloys," *J. Appl. Phys.*, vol. 89, no. 11, p. 5815, 2001.

[78] M. Kondo and T. Tanahashi, "Dependence of carbon incorporation on crystallographic orientation during metalorganic vapor phase epitaxy of GaAs and AlGaAs," *J. Cryst. Growth*, vol. 145, pp. 390–396, 1994.

[79] C. A. Larsen, N. I. Buchan, and G. B. Stringfellow, "Mass spectrometric studies of phosphine pyrolysis and OMVPE growth of InP," *J. Cryst. Growth*, vol. 85, pp. 148–153, 1987.

[80] M. C. Hanna, Z. H. Lu, and A. Majerfeld, "Very high carbon incorporation in metalorganic vapor phase epitaxy of heavily doped p-type GaAs," *Appl. Phys. Lett.*, vol. 58, no. 2, pp. 164–166, 1991.

[81] E. A. Fitzgerald, A. Y. Kim, M. T. Currie, T. A. Langdo, G. Taraschi, and M. T. Bulsara, "Dislocation dynamics in relaxed graded composition semiconductors," *Mater. Sci. Eng. B*, vol. B67, pp. 53–61, Dec. 1999.

[82] C. Heidelberger and E. A. Fitzgerald, "Heavy P-Type Carbon Doping of MOCVD GaAsP Using CBrCl3," *J. Cryst. Growth*, vol. 446, pp. 7–11, 2016.

[83] V. Swaminathan and A. T. Macrander, *Materials aspects of GaAs and InP based structures*. Prentice-Hall, Inc., 1991.

[84] C. L. Andre, M. Gonzalez, M. K. Hudait, E. A. Fitzgerald, J. A. Carlin, M. T. Currie, C. W. Leitz, T. A. Langdo, E. B. Clark, D. M. Wilt, and S. A. Ringel, "Impact of threading dislocations on both n/p and p/n single junction GaAs cells grown on Ge/SiGe/Si substrates," *Photovolt. Spec. Conf.*, pp. 1043–1046, 2002.

[85] R. A. Bruce and G. R. Piercy, "An improved Au-Ge-Ni ohmic contact to n-type GaAs," *Solid. State. Electron.*, vol. 30, no. 7, pp. 729–737, 1987.

[86] T. Suzuki, A. Gomyo, S. Iijima, K. Kobayashi, S. Kawata, I. Hino, and T. Yuasa, "Band-Gap Energy Anomaly and Sublattice Ordering in GaInP and AlGaInP Grown by Metalorganic Vapor Phase Epitaxy," *Jpn. J. Appl. Phys.*, vol. 27, no. 11, pp. 2098–2106, 1988.

[87] K. H. Lee, S. Bao, B. Wang, C. Wang, S. F. Yoon, J. Michel, E. A. Fitzgerald, and C. S. Tan, "Reduction of threading dislocation density in Ge/Si using a heavily As-doped Ge seed layer," *AIP Adv.*, vol. 6, no. 2, 2016.

[88] E. A. Fitzgerald, Y.-H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, "Relaxed GexSi1-x structures for III-V integration with Si and high mobility two-dimensional electron gases in Si," *J. Vac. Sci. Technol. B Microelectron.*

Nanom. Struct., vol. 10, no. 4, pp. 1807–1819, 1992.

[89] M. Yamaguchi, A. Yamamoto, and Y. Itoh, "Effect of dislocations on the efficiency of thin-film GaAs solar cells on Si substrates," *J. Appl. Phys.*, vol. 59, no. 5, pp. 1751–1753, 1986.

[90] S. A. Ringel, J. A. Carlin, C. L. Andre, M. K. Hudait, M. Gonzalez, D. M. Wilt, E. B. Clark, P. Jenkins, D. Scheiman, A. Allerman, E. A. Fitzgerald, and C. W. Leitz, "Single-junction InGaP/GaAs solar cells grown on Si substrates with SiGe buffer layers," *Prog. Photovoltaics Res. Appl.*, vol. 10, no. 6, pp. 417–426, Sep. 2002.

[91] C. Heidelberger and E. A. Fitzgerald, "GaAsP/InGaP heterojunction bipolar transistors grown by MOCVD," *J. Appl. Phys.*, vol. 121, no. 4, p. 45703, 2017.

[92] R. Jones, S. Öberg, and S. Marklund, "Structure and energy level calculations of dislocations in gallium arsenide," *Philos. Mag. Part B*, vol. 43, no. 5, pp. 839–852, 1981.

[93] D. B. Holt and B. G. Yacobi, *Extended Defects in Semiconductors*. Cambridge University Press, 2007.

[94] D. Biswas, N. Debbar, P. Bhattacharya, M. Razeghi, M. Defour, and F. Omnes, "Conduction- and valence-band offsets in GaAs/Ga0.51In0.49P single quantum wells grown by metalorganic chemical vapor deposition," *Appl. Phys. Lett.*, vol. 56, no. 9, pp. 833–835, 1990.

[95] J. J. O'Shea, C. M. Reaves, S. P. DenBaars, M. A. Chin, and V. Narayanamurti, "Conduction band offsets in ordered-GaInP/GaAs heterostructures studied by ballistic-electron-emission microscopy," *Appl. Phys. Lett.*, vol. 69, no. 20, p. 3022, 1996.

[96] A. Lindell, M. Pessa, A. Salokatve, F. Bernardini, R. M. Nieminen, and M. Paalanen, "Band offsets at the GaInP/GaAs heterojunction," *J. Appl. Phys.*, vol. 82, no. 7, p. 3374,

1997.

[97] S. L. Feng, J. Krynicki, V. Donchev, J. C. Bourgoin, M. Di Forte-Poisson, C. Brylinski, S. Delage, H. Blanck, and S. Alaya, “Band offset of GaAs-GaInP heterojunctions,” *Semicond. Sci. Technol.*, vol. 8, no. 12, pp. 2092–2096, 1993.

[98] S. M. Ting, “Monolithic Integration of III-V Semiconductor Materials and Devices with Silicon,” MIT, 1999.

Appendix A: TEM sample preparation

Imaging semiconductor samples by TEM requires that they be thinned to electron transparency, which occurs at a thickness on the order of 100 nm. This section describes the procedure used in this work to achieve this, which we refer to as the "hand-polishing method". This manual procedure involves two basic parts: first, thinning the specimen by grinding/polishing by hand to a thickness of around 10 μm, and second, using an ion mill to sputter a hole in the center of the sample. Surrounding this hole, the sample is wedge-shaped with a region of appropriate thickness for TEM analysis. This method contrasts with using focused ion beam (FIB) or dual-beam (SEM and FIB) instruments, which are more commonly used today. The hand-polishing method offers several advantages over FIB, listed here:

- Plan-view samples can be made easily.
- There is generally a larger imageable area. This is important especially for plan-view samples, when counting defects with relatively low densities.
- While the total time for preparing a sample is longer with the hand-polishing method than for FIB, the active time is shorter, because many of the steps involve long wait times during which other tasks can be accomplished. Efficiency is even higher when preparing multiple samples in parallel (e.g. two samples can be ground/polished simultaneously on the same grinding apparatus).
- Avoids scheduling issues with the CMSE dual-beam FIB tool (which is very heavily used).
- Avoids the cost of using the dual-beam FIB.

However, there are some advantages of FIB: namely, that a TEM sample can be produces from a specific area of the sample (e.g. from a particular defect seen in SEM or a device cross-section). In addition, samples prepared by FIB can have very good thickness uniformity, which is important for certain kinds of analysis such as HAADF STEM.

Here, the steps for the hand-polishing method are described in detail, first for preparing cross-section samples, then for plan-view samples.

Cross-section sample preparation

1. Cleave out several rectangular pieces of the wafer, about 4 mm by 2 mm. Using M-bond 610 adhesive, glue two of the pieces together face-to-face, making sure that they are clean and free of dust from the cleaving. It is important that the line of glue in between the samples is as thin as possible. Then glue a 4 mm by 2 mm Si piece to either side of the sandwich, also using the M-bond 610. Clamp the entire stack (Si/substrate/film/film/substrate/Si) together with reverse-action tweezers, and place on a hotplate to cure at 190 °C for at least 4 hours.
2. The grinding and polishing is done using a Gatan Model 623 disc grinder. This allows the sample to be polished very flat and at a controlled rate and final thickness. The Gatan model is better than its competitors because it has finer threading (250 μm per rotation rather than 500 μm per rotation). Mount the sample sideways onto the provided ceramic sample mount using QuickStick 135 mounting wax, by placing the mount on a 190 °C hotplate, applying the wax, and placing the sample in the molten wax. Allow to cool to room temperature. Grind and then polish one side of the sample, using these grits of SiC grinding paper in succession: 500, 1200, 4000. Do not thin the sample to less than 1 mm

at this time, just ensure that it is very smooth to the naked eye and perfectly flat. Continue grinding on the 4000 grit paper for at least a few minutes to ensure flatness. Polish for around 1 minute with DiaDuo 3 μm diamond slurry on a polishing pad, without advancing the sample out any further.

3. Place the ceramic sample mount on the hotplate and wait for the wax to melt. Flip the sample over with fine tweezers, pressing down so that the first polished side is flush against the ceramic mount. It is critical that the sample is mounted perfectly flat, or else it will be polished into a wedge shape.
4. Grind the second side of the sample with the 500 grit paper, advancing 200 um at a time, until it is around 100-200 um thick. Then, with the 1200 grit paper, grind until the sample is around 30 um thick, advancing only 10-20 microns at a time. When it is close to 30 um thick, the edges of the sample will begin to get chipped away (this is visible by eye) and the edges of the Si cladding layers will appear red in transmission (visible in a low-mag binocular microscope). Next, grind/polish the sample using the 4000 grit paper until is around 10-15 um thick, advancing only around 5 um at a time. With each advance of 5 um, make sure to grind for at least 1-2 minutes to ensure that the sample is still perfectly flat. Once most of the Si is red in transmission, or the entire edge of the sample is being worn away uniformly, the grinding is finished. Polish the sample again for 1 minute with the 3 um diamond slurry.
5. Attach a Cu TEM grid to the sample under a low-mag binocular microscope. For cross-section samples, use a grid with a 1 mm x 2 mm slot-shaped aperture. The grid can be attached to the sample with any quick-curing 2-part epoxy. Mix a small amount of the epoxy and apply to the polished sample in very small dots in a circular pattern to coincide

with where the grid will sit. Do not get any epoxy on the center of the sample where the milling and imaging will be done. To apply the epoxy, I use one of my eyelashes held using extra-fine-tipped tweezers as a mini paintbrush. However, it would also be possible to apply the epoxy using the tip of the extra-fine-tipped tweezers. Once the epoxy is on the sample, place the grid down so that the long axis of the slot is in line with the glue line made by the M-bond and push down lightly with the tweezers. Allow the epoxy to cure for 30 minutes or until hard.

6. Using the binocular microscope and a sharp razor blade, cut away the excess sample extending past the perimeter of the Cu. If the sample is thinned to 10 um, it should be easily cut by the razor blade and a moderate pressure. If it is impossible to cut the excess sample away, then it is probably too thick.
7. Place the ceramic sample holder with the sample attached into a beaker of acetone. The acetone will dissolve the mounting wax, freeing the sample from the ceramic holder, but it will not attack the M-bond or the epoxy. After dissolving for 30 min to 1 hour, pick up the sample from the holder using extra-fine-tipped tweezers. Rinse the sample in IPA, then gently dry with N2 or compressed air. Place the sample in a TEM grid holder.
8. The sample is ion milled in a Fischione Ar ion mill. The milling parameters depend on the substrate of the sample, and can be adjusted based on the level of mill damage observed in the TEM images and the time in which the samples finish milling:

Substrate	**Voltage**	**Current**
Si	4.0 kV	5.0 mA
GaAs or Ge	3.5 kV	4.5 mA
InP	3.2 kV	4.0 mA

The milling angle should be set to 15° and the rocking angle to 45°. The 180° rotation time should be 10 minutes. Mill the sample for 1 hour, then in 40-minute increments until a hole has just formed in the center of the sample, intersected by the glue line. For samples at the InP lattice constant, they should be cryo-milled at -100 °C to prevent amorphization of the In-containing compounds, but for all other samples room-temperature milling is sufficient.

9. At this point, the sample is ready to be imaged.

Plan-view sample preparation

1. Cleave out a rectangular piece of the substrate about 3 mm by 3 mm.
2. Attach the sample facedown to the ceramic sample holder with the mounting wax. While the wax is still hot, make sure that the sample is pressed firmly down to the holder such that the film is perfectly flat against the face of the holder.
3. Grind/polish the sample using the same procedure as cross-section step 4. For the step while grinding with the 4000 grit paper, the remaining sample thickness cannot be ascertained by the light coming through the sample because the III-V films are highly absorbent at all visible wavelengths. The sample is probably at the correct thickness when all edges of the sample are being worn away uniformly. In addition, when the sample is approaching the correct thickness, its edge can no longer be felt using your finger through a standard nitrile glove. Obviously, developing the judgement of when the sample is thin enough to continue takes practice.

4. Attach the TEM grid to the thinned sample using the same procedure as cross-section step 5. For plan view samples, a Cu grid with a 1 mm circular aperture should be used.
5. Once the epoxy has dried, remove the excess material from around the grid as in cross-section step 6.
6. Remove the sample from the ceramic sample holder as in cross-section step 7.
7. Mill a hole in the center of the sample using the Fischione ion mill. The sample should be loaded into the plan-view sample holder with the Cu grid on top. In this orientation, the upper surface of the film will be facing down and is protected from milling and redeposition of material by the holder. Mill using only the top gun, with 360° sample rotation. The same voltages and currents as in cross-section step 9 can be used. Mill for 90 minutes, then subsequently for 60-minute increments until a hole forms in the sample. Stop immediately once a small hole has formed.
8. At this point, the sample is ready to be imaged.

Appendix B: HBT fabrication process for MTL

Process name: GaAsP/InGaP HBT

Description: Fabricate a III-V HBT structure. All processes are red.

Starting Material: Epi-stack of InGaP/GaAsP/GaAs substrate or InGaP/GaAsP/SiGe/Si substrate, wafer pieces, grown in the Fitzgerald Group MOCVD in Building 13.

	Description	Device Processing	Lab	Machine
1	Degrease/clean	Acetone/IPA/Water/N2 blow dry	TRL	Photo-wet
2	Photolithography, Mask 1 (Emitter Mesa)	HMDS, recipe 1 SPR at 3000 rpm for 30s Soft-bake 90C 5 min Expose PR, 13 s Develop MA-CD-26, 80 s Hard bake 120C for 5 min	TRL	HMDS Coater Coater hotplate MA-6 Photo-wet Hotplate 1 or 2
3	Etch emitter mesa	1:1:10 H2SO4:H2O2:H2O, 60 s followed by 1:1 HCl:H2PO4, 60 s	TRL	Acidhood
4	Resist strip	Acetone in sonicator 20 min, Methanol rinse, IPA rinse, N2 dry	TRL	Acidhood
5	Inspection (optional)	Profilometry SEM	TRL ICL	Dektak semZeiss
6	Photolithography, Mask 2 (Base/Collector Mesa)	HMDS, recipe 1 SPR at 3000 rpm for 30s Soft-bake 90C 5 min Expose PR, 13 s Develop MA-CD-26, 80 s Hard bake 120C for 5 min	TRL	HMDS Coater Coater hotplate MA-6 Photo-wet Hotplate 1 or 2
7	Etch base mesa	1:1:10 H2SO4:H2O2:H2O, variable time Or, 1:8:80 H2SO4:H2O2:H2O, variable time	TRL	Acidhood
8	Resist strip	Acetone in sonicator 20 min, Methanol rinse, IPA rinse, N2 dry	TRL	Acidhood
9	Inspection (optional)	Profilometry	TRL	Dektak

		SEM	ICL	semZeiss
10	Pre-ALD clean (immediately before ALD)	1:10 H2SO4:H2O, 60 s	TRL	Acidhood
11	Deposit Oxide Passivation	Deposit 10 nm Al2O3, 200C	ICL	ALD (Cambridge Nanotech)
12	Photolithography, Mask 3 (Metal/oxide via)	HMDS, recipe 1 AZ5214 at 1500 rpm 30s Soft-bake, 90 C, 5 min Expose PR, 9 s Post-exposure bake, 115 C, 2 min Flood expose, 90 s Develop AZ422, 2 min	TRL	HMDS Coater Coater hotplate MA-6 Hotplate 1 or 2 MA-6 Photo-wet
13	Oxide Etch (immediately before metal dep)	7:1 BOE, 3 min	TRL	Acidhood
14	Deposit contact metal	50 A Ti/400 A Pt/1200 A Au	TRL	ebeamFP
15	Lift-off	Acetone soak 30 min (up to overnight), sonicator for 10 s, acetone rinse, methanol rinse, IPA rinse, N2 dry	TRL	Photo-wet
16	Inspection (optional)	Profilometry SEM	TRL ICL	Dektak semZeiss

Appendix C: MATLAB code for breakdown voltage modeling

Included here is the MATLAB code used for predicting breakdown voltage of HBTs and diode structures. For more details on the equations used, see Section 5.4.3 of the main text.

```
% Calculate and plot one-sided p+ - n diode BV as a function of GaAsP composition.

% Equations sourced from J. del Alamo, "Integrated Microelectronic Devices: Physics
% and Modeling" and from pp. 292-300 of W. Liu, "Handbook of III-V Heterojunction
% Bipolar Transistors"

clear all

% constants
p.q = 1.6e-19; % electron charge
p.eps_0 = 8.854e-14; % vacuum permittivity (F/cm)
p.kT = 0.025; % kT @ RT (eV)
p.m0 = 5.69e-16; % electron mass (eV.s^2/cm^2)

% materials parameters
p.EgGaAs = 1.42; % GaAs gamma point (eV)
p.EgGaP = 2.78; % GaP gamma point (eV)
p.CgGaAsP = -.21; % bowing parameter for GaAsP
p.Eiifactor = 1.7/1.42; % ratio of Eii to Eg
p.eps_rGaAs = 12.9; % dielectric constant of GaAs
p.eps_rGaP = 11.1; % dielectric constant of GaP
p.lc = 13.38e-7; % MFP between electron collisions (cm)
p.niGaAs = 2.1e6; % intrinsic carrier concentration of GaAs (cm^-3)
p.niGaP = 2; % intrinsic carrier concentration of GaP (cm^-3)
p.vd = 1e7; % electron drift velocity (cm/s)
p.mu = 3200; % electron mobility in collector (cm^2/V.s)
p.meff = .063; % electron effective mass/m0

% device parameters
p.N_c = 1.2e17; % collector doping (cm^-3)
p.N_b = 8e17; % base doping (cm^-3)
p.W_c = 500e-7; % collector width (cm)
p.beta = 30; % device current gain

% calculation-related
p.a = 1000; % number of bins for integration of alpha
rangeCBO = [15]; % starting point for BV_CBO
rangeCEO = [5]; % starting point for BV_CEO
xAs = linspace(.65,1,100); % composition range (As fraction)

%%%%%%%%%

% Find BV_CBO -- solve equation for M->inf
for i = 1:length(xAs)
    eqn = @(V) CBO_eqn(V,xAs(i),p);
    BV_CBO(i) = fzero(eqn,rangeCBO);
```

```
end

% Find BV_CEO -- solve equation for M->1+1/beta
for i = 1:length(xAs)
    eqn = @(V) CEO_eqn(V,xAs(i),p);
    BV_CEO(i) = fzero(eqn,rangeCEO);
end

% plot calculated data
figure(1)
plot(xAs,BV_CBO,'b-',xAs,BV_CEO,'r-')
xlabel('GaAs_xP_{1-x} As Fraction')
ylabel('BV (V)')
legend('BV_{CBO}','BV_{CEO}')
axis([0.75 1 0 inf])

%%%%%%%%%

function [a] = CBO_eqn(V,xAs,p)
%equation to solve for BV_CBO calculation

% calculate Eg
Eg = p.EgGaAs*xAs + p.EgGaP*(1-xAs) + p.CgGaAsP*xAs*(1-xAs);

% calculate Eii
Eii = Eg*p.Eiifactor;

% calculate ni (using geometric mean)
ni = (p.niGaAs^xAs)*(p.niGaP^(1-xAs));

% calculate built-in voltage
phi_bi = p.kT*log(p.N_b*p.N_c/ni^2);

% calculate dielectric constant
eps_r = p.eps_rGaAs*xAs + p.eps_rGaP*(1-xAs);

% calculate space charge region width
x_scr = sqrt(2.*p.eps_0.*eps_r.*(V+phi_bi)./(p.q.*p.N_c));

% perform integral to calculate RHS of BV_CBO eqn to be set to 0
a = 1 -
trapz(linspace(0,min(p.W_c,x_scr),p.a),alpha(linspace(0,min(p.W_c,x_scr),p.a),V,...
xAs,p));

end

%%%%%%%%%

function [a] = CEO_eqn(V,xAs,p)
%equation to solve for BV_CEO calculation

% calculate Eg
Eg = p.EgGaAs*xAs + p.EgGaP*(1-xAs) + p.CgGaAsP*xAs*(1-xAs);

% calculate Eii
Eii = Eg*p.Eiifactor;

% calculate ni (using geometric mean)
ni = (p.niGaAs^xAs)*(p.niGaP^(1-xAs));
```

```
% calculate built-in voltage
phi_bi = p.kT*log(p.N_b*p.N_c/ni^2);

% calculate dielectric constant
eps_r = p.eps_rGaAs*xAs + p.eps_rGaP*(1-xAs);

% calculate space charge region width
x_scr = sqrt(2.*p.eps_0.*eps_r.*(V+phi_bi)./(p.q.*p.N_c));

% perform integral to calculate RHS of BV_CEO equation to be set to 0
a = 1/(1 - trapz(linspace(0,min(x_scr,p.W_c),p.a),...
alpha(linspace(0,min(x_scr,p.W_c),p.a),V,xAs,p)))-(1+1/p.beta);

end

%%%%%%%%%

function [alph,x_scr] = alpha(x,V,xAs,p)
%ALPHA calculate impact ionization coefficients at vector of points x and applied
%voltage V with As fraction xAs.

% calculate Eg
Eg = p.EgGaAs*xAs + p.EgGaP*(1-xAs) + p.CgGaAsP*xAs*(1-xAs);

% calculate Eii
Eii = Eg*p.Eiifactor;

% calculate ni (using geometric mean)
ni = (p.niGaAs^xAs)*(p.niGaP^(1-xAs));

% calculate built-in voltage
phi_bi = p.kT*log(p.N_b*p.N_c/ni^2);

% calculate dielectric constant
eps_r = p.eps_rGaAs*xAs + p.eps_rGaP*(1-xAs);

% max electric field and depletion width
if 1/(2*p.eps_0*eps_r)*p.q*p.N_c*p.W_c^2 >= V + phi_bi
    Emax = sqrt(2*p.q*p.N_c*(V+phi_bi)/(p.eps_0*eps_r));
else
    Emax = 1/p.W_c*(V+phi_bi-1/(2*p.eps_0*eps_r)*p.q*p.N_c*p.W_c^2)...
    +1/(p.eps_0*eps_r)*p.q*p.N_c*p.W_c;
end

% electric field at point x
E = Emax - p.q*p.N_c/(p.eps_0*eps_r).*x;

% calculate impact ionization factor (assume same for h+ and e-) at points x
alph = abs(E)./Eii.*exp(-Eii./(abs(E).*p.lc));

end
```