KUNGL
TEKNISKA
HÖGSKOLAN

Ultra High Speed InP Heterojunction Bipolar Transistors

Mattias Dahlström

Stockholm 2003

Doctoral Dissertation
Royal Institute of Technology
Department of Microelectronics and Information Technology

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknisk doktorsexamen måndagen den 26 maj 2003 kl 10.00 i sal C2, Electrum Kungl Tekniska Högskolan, Isafjordsvägen 22, Kista.

ISBN 91-7283-496-X
TRITA-TRITA-MVT Report 2003:2
ISSN ISSN 0348-4467
ISRN ISRN KTH/MVT/FR–03/2–SE

Printed by Universitetsservice AB, Stockholm 2003

Abstract

This thesis deals with the development of high speed InP mesa HBT's with power gain cut–off frequencies up to and above 300 GHz, with high current density and low collector discharging times.

Key developments are Pd–based base ohmics yielding base contact resistances as low as 10 $\Omega\mu\text{m}^2$, base–collector grades to enable to use of InP in the collector, and an increase in the maximum current density through collector design and thermal optimization. HBT's with a linear doping gradient in the base are for the first time reported and compared to HBT's with a bandgap graded base. The effect of degenerate base doping is simulated, as well as the base transit time.

Key results include a DHBT with a 215 nm thick collector and an $f_\tau = 280$ GHz, and f_{max}=400 GHz. This represents the highest f_{max} reported for a mesa HBT. Results also include a DHBT with a 150 nm thick collector and an $f_\tau = 300$ GHz, and f_{max}=280 GHz. The maximum operating current density has been increased to above 10 mAμm while maintaining f_τ and $f_{max} \geq 200$ GHz.

A mesa DHBT process with and as much yield and simplicity as possible has been developed, while maintaining or pushing world–class performance.

ISBN 91-7283-496-X • TRITA-TRITA-MVT Report 2003:2 • ISSN ISSN 0348-4467 • ISRN ISRN KTH/MVT/FR–03/2–SE

Acknowledgements

For my work I am indebted to many people, both at KTH and at UCSB. In Stockholm I want to thank .... Lars Thylén, Eilert Berglind, Patrik Evaldsson, Urban Eriksson and Urban Westergren. Also my thanks to Robert Lewén, Stefan Irmscher and all others! Especially my *innebandy* pals!

Moving over to UCSB was a wonderful opportunity and I thank Prof. Thylén and Prof. Rodwell for making it possible. I also want to grab the opportunity thank *Johan och Karin Engbloms Stipendiefond* and *L.M. Ericsson Stipendiefond*'s generous contributions for my voyage over to expensive Santa Barbara! Working for Prof. Rodwell has been a pleasure and a privilege. Usually depth of knowledge is inversely coupled with width of knowledge, but I have failed to find this error in him ...

Most of my time at UCSB was spent in the clean room and life would have been a lot more miserable if not for Jack, Brian, Bob, Mike, Neil and Luiz. Thanks for their never ending dedication despite often grumpy and unknowledgeable graduate students ! Thanks are also due to other members of the Rodwell Empire....they are very dedicated bunch and they deserve the best, Zachary, PK, Miguel, Dennis, Navin, P-diddi, Yun-Wei, Young-Min, Sangmin, Jong-Uk, Heng-Kuang and Christoph. Prof. Harrison from University of Nottingham have his share in helping me, as well as being a really nice guy! I also want to express a heartfelt thank to Dr. Amy Liu and her coworkers at IQE Inc for their dedicated efforts and great wafers. I also want to thank Donato who helped cure homesickness and clean-room ennui — it is so nice to speak Swedish!

Finally I want to express my love and gratitude to my wife Virginia. Siempre!

Contents

List of Figures

List of Tables

List of acronyms

III-V Group III - Group V semiconductors (refering to the periodic system)

BCB A plastic dielectrica

BJT Bipolar Junction Transistor

DOS Density Of States

HBT Heterojunction Bipolar Transistor

MBE Molecular Beam Epitaxy

MOVPE Metal-Organic Vapor Phase Epitaxy

RIE Reactive Ion Etching

SIMS Secondary Ion Mass Spectroscopy

CPW Coplanar Waveguides

PECVD Plasma Enhanced Chemical Vapour Deposition

Till mina föräldrar

Chapter 1

Introduction to InP Heterojunction Bipolar Transistors

1.1 Overall goal of the work at UCSB: narrow mesa HBT

Development of analog and digital ICs operating at 80-160 GHz clock frequencies requires improved transistor performance and manufacturabilty. Based upon analyzes of emitter-coupled logic (ECL) gate delay [1, 2], target specifications for 160 Gb/s optical transmission include > 3 V breakdown, > 450 GHz f_τ and f_{max} , maximum emitter current density $J_e \geq 7$ mA/μm^2 at V_{cb}=0 V, and low base-collector capacitance charging time ($C_{cb}/I_e \leq 0.3$ ps/V). Improved transistor bandwidth can be obtained by simultaneously reducing the collector depletion thickness, the collector and emitter junction widths, the emitter contact resistivity, and, in mesa HBT's, the base sheet and contact resistivity [1].

The goal of this work was to demonstrate a conventional emitter up HBT technology with performance approaching them of transferred substrate HBT's. The underlying reason for this was to improve the manufacturabilty of high frequency transistors with transit time f_τ and power gain cut off frequency f_{max} of more than 300 Ghz.

Transfered substrate HBT's suffer from complications involved in removing the InP substrate without damaging the collector region. The problems are more severe for InP collector transistors than for devices with InGaAs collector since the InP can be etched by the selective substrate removal etch. For reaching this the following needs to be achieved:

- Very low base resistance
- Very good base alignment
- Narrow base-collector mesa
- High current density

The reason behind this is that the base-collector capacitance must be kept to a minimum just as is achieved in transferred subststrate HBT's and therefore the base contact must be as narrow as possible. But in order to keep the total RC delay as small as possible the base resistance must also be as low as possible given the constraints. The base resistance is composed of two parts, the intrinsic resistance and the contact resistance (equation 2.25). The intrinsic resistance is minimized by doping the base region as high as possible and by keeping the emitter and base-emitter spacing narrow. Above $5 \cdot 10^{19}$ cm^{-3} carbon has to be used instead of beryllium or zinc as the dopant. The contact resistance is inversely proportional to the square root of the doping under idealized conditions and is thus minimized by increasing the doping. The correct choice of contact metal and annealing procedure (page 33) is also very important.

To keep the base–collector capacitance C_{cb} as small as possible the base contact width should be on the order the base contact transfer length, $0.15 - 0.4$ μm. This

makes the necessary base alignment tolerance is on the order of 0.1-0.2 μm which is a demanding task in the average university clean room.

One thing that cannot be overlooked is the importance of thermal conductivity. The transfered substrate devices are sensitive to this since the heat generated in the device is removed through the emitter region which poses a high thermal resistance since it is narrow (1-0.3 μm) and often made of ternary alloys such as InGaAs or InAlAs. InP has a much higher thermal conductivity than InGaAs or InAlAs. By contrast a narrow mesa HBT can — especially if the collector region is InP and the amount of InGaAs or InAlAs in the subcollector and buffer regions are kept to a minimum — tolerate a higher current density which should result in a higher frequency of operation.

1.2 Overview of transistor technology

Due to their respective advantages, III-V Heterojunction Bipolar Transistors (HBT's) and Si/SiGe HBT's are primarily used in high-speed digital and mixed-signal applications. The principal advantages of III-V InP-based HBT's is superior bandwidth and breakdown. The main factors contributing to this is emitter whose bandgap energy is much larger than that of the base, such as InP with E_g=1.35 eV for emitter and $In_{47}Ga_{53}As$ with E_g=0.76 eV for base. This allows the base doping to be increased to the limits of incorporation in growth (10^{20} cm^{-3}), and results in very low base sheet resistance and high Early voltage. High electron velocities are a second significant advantage of III-V HBT's, which also allows a trade-off for thicker regions with better breakdown voltage. Best reported results of InP-based HBT's include 351 Ghz f_τ [3], simultaneous 329 Ghz f_τ and f_{max} [3,5], and 300 GHz f_τ and f_{max} for GaAsSb HBT [4]. Meanwhile Si/SiGe HBT's have obtained 210 Ghz f_τ [8] and 285 GHz f_{max} [7] for an integration scale several orders of magnitude larger. Despite the advantages of III-V HBT's provided by superior material properties, Si/SiGe HBT's remain highly competitive. The high bandwidths of Si/SiGe HBT's arise from aggressive submicron scaling, made possible through polysilicon contacts, making the metal-semiconductor contacts much larger than the intrinsic transistor. In devices with a 0.12 μm base-emitter junction, 207 GHz f_τ and 285 GHz f_{max} have been obtained [7]. Self-aligned polysilicon contacts reduce both the parasitic collector-base capacitance and the base resistance. In marked contrast to the aggressive submicron scaling and aggressive parasitic reduction employed in Si/SiGe HBT's, III-V HBT's are typically fabricated with 1-2 μm emitter junction widths. Current densities are also much lower in III-V transistors despite similar thermal conductivity, and contribute strongly to improved circuit performance. Further submicron scaling be needed to improve the bandwidth of III-V heterojunction transistors and is critical to their continued success.

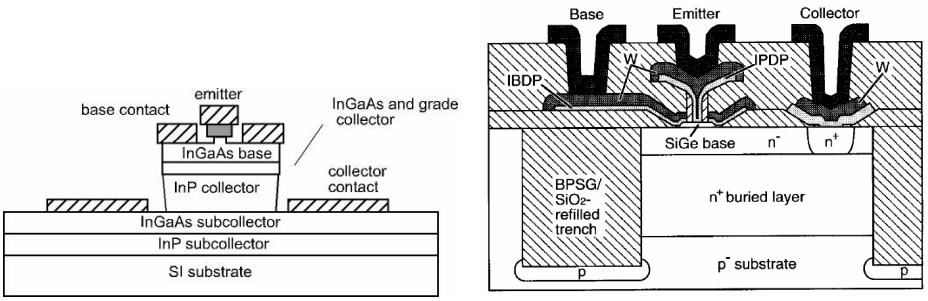

Figure 1.1. Mesa InP HBT cross section (left). SiGe HBT cross section (right)

1.3 The InP transistor

A HBT is composed of three main regions: the emitter, the base and the collector. Simply put the emitter sends out electrons, the base modulates that current and the collector collects them all. The key point is that a small variation in base current is translated to a larger collector current. The ratio is referred to as the gain of the device and is usually 20-200. What sets a HBT apart from a Bipolar Junction Transistor (BJT) are the heterojunctions (section 2.2), [12], which permits very high base doping. The high base doping permits a number of advantages such as low base resistance and thinner bases, which results in higher device speed and gain. Figure 1.3 shows an InP/InGaAs/InP HBT with abrupt emitter base junction, figure 1.4 shows an InP/InGaAs/InP HBT with graded emitter base junction, and figure 1.5 shows an InP/GaAsSb/InP HBT with abrupt emitter base junction. These represent the main types of DHBT's available. Figure 1.2 illustrates the different regions in a HBT and the denominations. In this work W_e, W_c indicate *widths* or horizontal dimensions, and T_c, T_b indicate *thickness* or vertical dimensions.

A typical layer structure is shown in Table 1.1.

1.3.1 Criteria for high speed devices

To achieve a mesa HBT with simultaneously high f_τ and f_{max} suited for high speed circuits the factors involved need to be identified [6]. The current–gain cutoff frequency f_τ,

$$\frac{1}{2\pi f_\tau} = \tau_b + \tau_c + \frac{kT}{qI_c}\,(C_{je} + C_{cb}) + (R_{ex} + R_c)C_{cb}, \tag{1.1}$$

where R_{ex} and R_c are the parasitic emitter and collector resistances. R_{ex} and R_c are discussed in chapter 3.7 and are on the order of 4 Ohms each for a device with a 0.7×8 μm emitter. C_{cb} is the collector junction capacitance, and I_c the collector

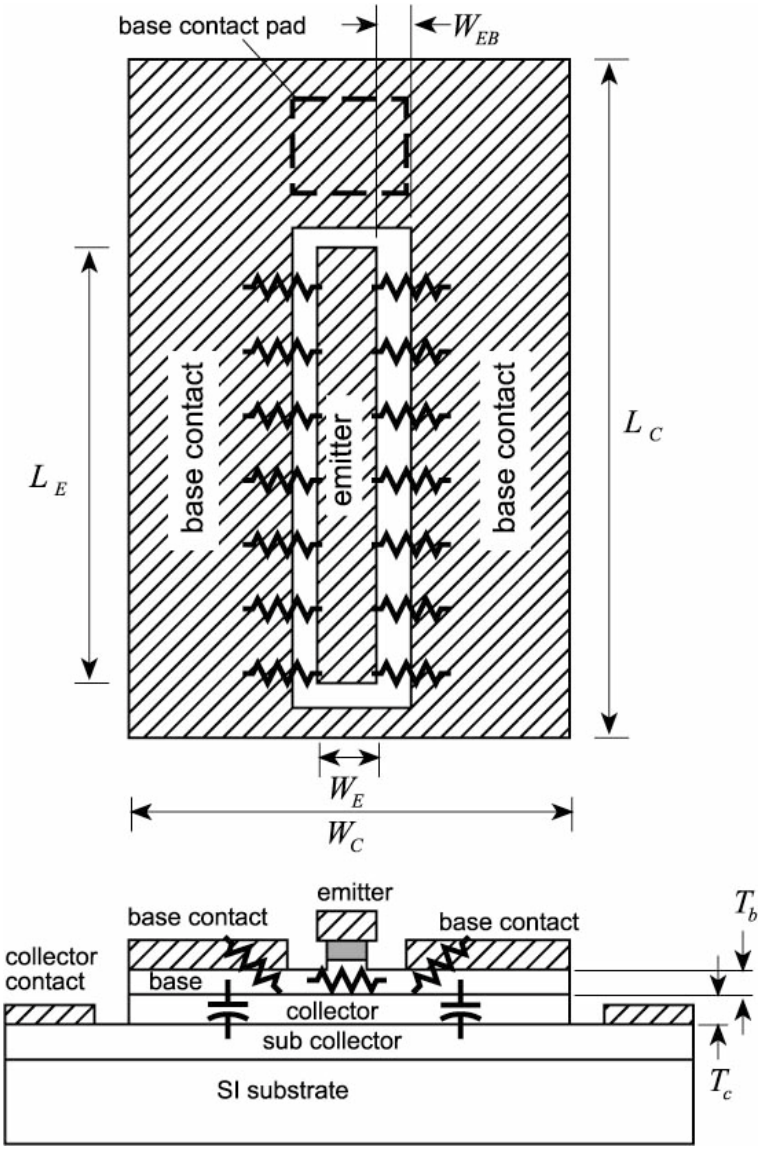

Figure 1.2. Plan and cross-section of a typical mesa HBT. The emitter-base junction has width W_e, length L_e and area $A_e = L_e W_e$, while the collector-base junction has width W_c, length L_c and area $A_c = L_c W_c$

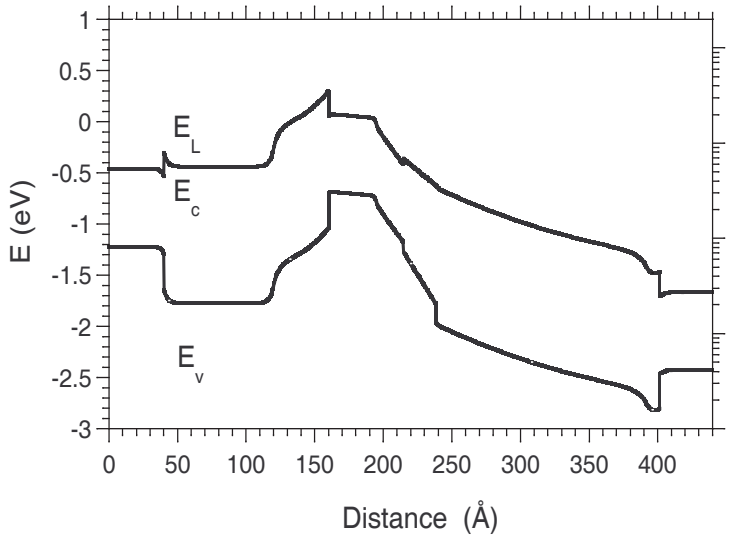

Figure 1.3. InP/InGaAs/InP HBT band diagram, with abrupt base–emitter, V_{ce}=1.3 V and V_{be}=0.8 V

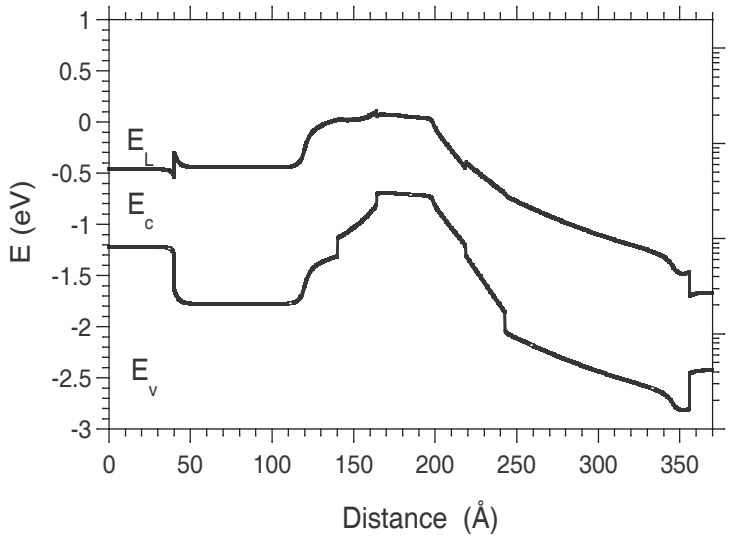

Figure 1.4. InP/InGaAs/InP HBT band diagram, with graded base–emitter, V_{ce}=1.3 V and V_{be}=0.8 V

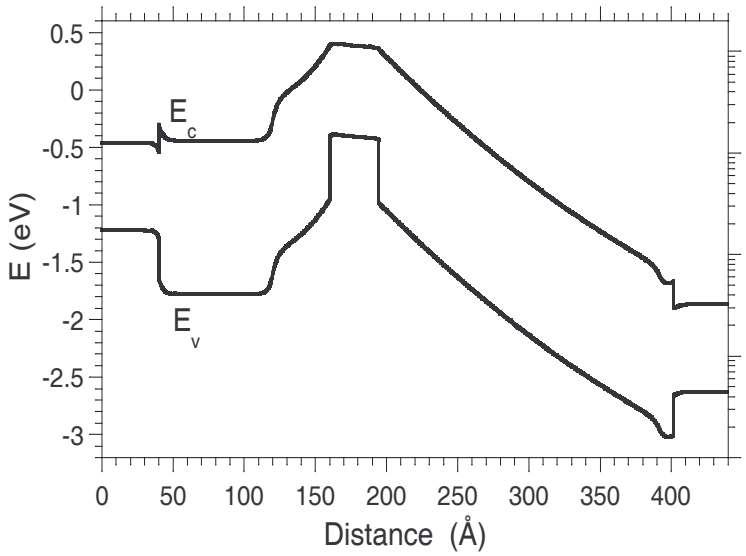

Figure 1.5. InP/GaAsSb/InP HBT band diagram, V_{ce}=1.5 V and $V_{be} = 0.5$ V

Table 1.1. DHBT layer structure

Material	Doping(cm^{-3})	Thickness(nm)
n-InGaAs	$3 \cdot 10^{19}$	80
n-InP	$3 \cdot 10^{19}$	90
n-InP	$8 \cdot 10^{17}$	10
n-InP	$3 \cdot 10^{17}$	30
p-InGaAs	$8 \rightarrow 5 \cdot 10^{19}$	30
n-InGaAs	$2 \cdot 10^{16}$	20
n-InAlGaAs	$2 \cdot 10^{16}$	24
n-InP	$3 \cdot 10^{18}$	3
n-InP	$2 \cdot 10^{16}$	170
n-InP	$1.5 \cdot 10^{19}$	50
n-InGaAs	$2 \cdot 10^{19}$	25
n-InP	$3 \cdot 10^{19}$	200
SI-InP	UID	

current. τ_b and τ_c are the base and collector transit times. They are on the order of 180 fs and 400 fs for our DHBT's. C_{je} is the emitter–base junction capacitance. Naturally, the achieve a good f_τ close attention has to be paid to all terms in (1.1). Compared to a transfered substrate HBT (T.S.) the collector junction capacitance C_{cb} is much higher in a mesa HBT unless the size of the base contact is kept to a minimum. Regardless of the value of f_τ, transistors cannot provide power gain at frequencies above f_{max} and a good design should pay attention to both. Independent of f_τ, f_{max} defines the maximum usable frequency of a transistor in either narrowband reactively-tuned or broadband distributed circuits [13]. In more general circuits, all transistor parasitics play a significant role. The f_τ and f_{max} of a transistor are then cited to give a first-order summary of the device transit delays and of the magnitude of its dominant parasitics. C_{cb}/I_c - the ratio of collector capacitance to collector current (discharging time) and the breakdown voltage also play a critical role. In an HBT with base resistance R_{bb} and collector capacitance C_{cb} the power-gain cutoff frequency is approximately:

$$f_{max} \simeq (f_\tau / 8\pi R_{bb} C_{cbi})^{1/2} \tag{1.2}$$

The base-collector junction is a distributed network, and $R_{bb}C_{cbi}$ represents an effective, weighted time constant. It arises from the fact the current distribution is not homogenous over the base-collector mesa, most of the current is directly beneath the emitter [6, 14]. C_{cbi} is the *intrinsic* part of the base–collector capacitance, and $C_{cb} = C_{cbi} + C_{cbe}$ with C_{cbe} the extrinsic part of the base–collector capacitance. To answer the question whether a mesa HBT could reach performance similar to a transfered substrate HBT simulations were performed using a distributed mesh model, see section 3.1.

Chapter 2

Theory of the InP Heterojunction Bipolar Transistor

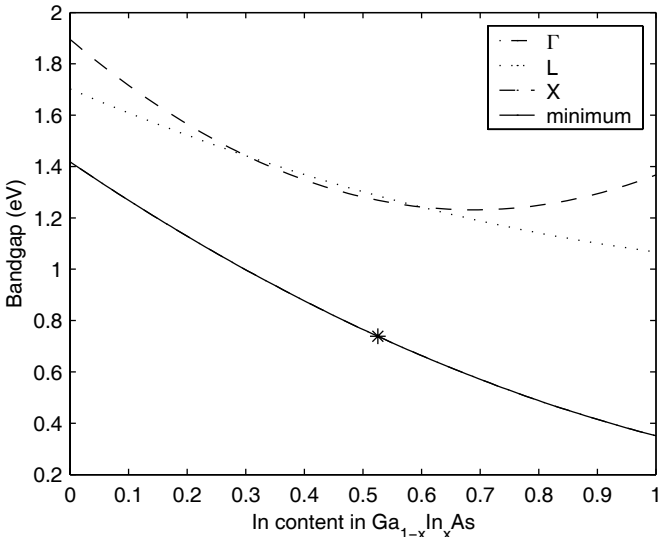

Figure 2.1. Energy of Γ, L and X-bands in $In_xGa_{1-x}As$. The star shows the InP–lattice matched condition.

2.1 The materials

2.1.1 Band structure of III-V materials

For calculation of base transit time and collector properties precise knowledge about the relevant materials is important. Materials are typically grown lattice matched on InP substrates, with the same lattice constant, 5.8 Å as InP. Material can be grown strained (not lattice matched) for a certain distance, but above a certain distance the material will relax and become polycrystalline. The distance is in practice larger than theoretically predicted and seems to be a function of growth parameters. Published data about especially band offsets but also bandgaps show considerable variation [17], and it is not clear which data to use. Data from different research groups are grouped together, but if the reason is due to measurement method or due to growth is not clear. One reason for changed band offsets and bandgaps is strain in the heterojunction or the interface type, as is known for the InAs-GaSb junction [17]. III-V semiconductors have three energy valleys for electrons, denoted Γ, L and X (figure 2.1) [18]. The Γ valley is typically the one lowest in energy. There are also three energy valleys for the holes, heavy hole , light hole and the split-off band. The heavy hole band and light hole band are typically very close to each other. The separation between the lowest electron band and the highest hole band define the bandgap E_g. Energy bands represent modes of propagation, how an ensemble of electrons move through a crystal. If an electron

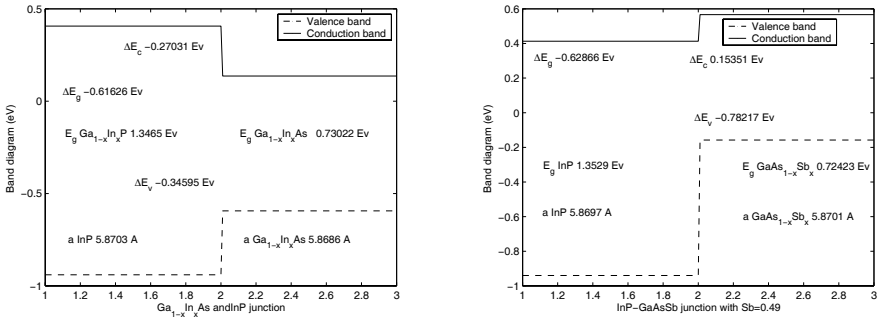

Figure 2.2. Band line up of lattice matched InGaAs–InP and GaAsSb–InP

become energetic enough to aquire an energy large than $\Delta E_{\Gamma-L}$, the separation between the Γ band and the L band, the electron can jump over to the L-band, so called $\Gamma - \mathrm{L}$ scattering. This results in a slowing down of the electron with an energy at least equal to the energy difference [1]. This is an important mechanism for the collector transit time in a HBT. If electrons never get enough energy to make the jump to the next band they can travel at a substantial velocity. This is important for the region next to the base in a HBT: the electron velocity can be very high there, which is what is desired, while $\Gamma - \mathrm{L}$ scattering can drastically lower it [19]. Figure 2.2 shows the band line–up between InP/InGaAs and InP/GaAsSb, for the lattice matched, low doped case. When the two bandgaps ΔE_g are different the difference is split up between the conduction band and the valence band, ΔE_c and ΔE_v. The ratio is very important for a (npn) HBT: the holes should be confined to the base and thus the valence band offset should be large. The electrons should easily travel through the base and into the collector, and the conduction band offset should thus the small or even negative. Section 2.5.4 discusses methods for eliminating the effective conduction band offset in the base–collector junction that otherwise would hinder electrons from the leaving the base.

2.1.2 Doping of semiconductors

The background doping in common semiconductors when grown with MBE or MOCVD is in the $\approx 10^{15}$ cm^{-3} range, and it is generally n-type. To achieve p-type doping an acceptor has be incorporated and to achieve n–type doping a donator has to be incorporated into the lattice. An acceptor is an doping atom that can accept an extra electron and a donor is an doping atom that can donate an extra electron. The situation is made more complicated in composite semiconductors such as InP compared to Si since the doping type achieved will depend on which lattice position the dopant atom occupies. One example is carbon doping of

[1] It's like paying to change lanes on the freeway only you change into the slower lane!

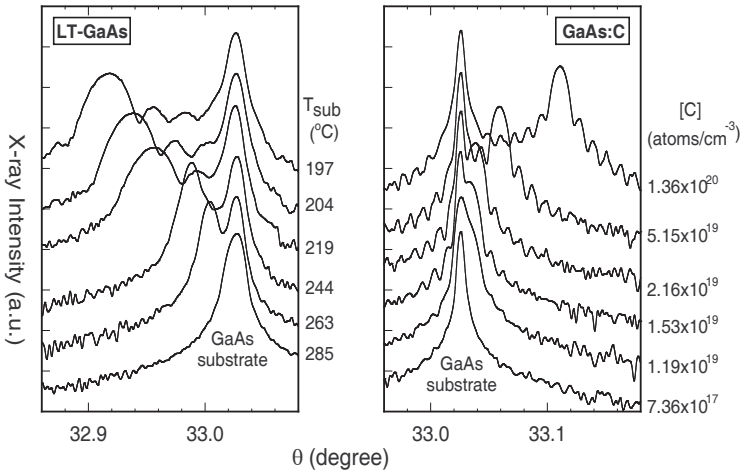

Figure 2.3. Lattice distortion from carbon doping in GaAs. X–ray from [76]

InGaAs and InP: in InGaAs and GaAs the carbon atom typically occupies a group V position (As) and can receive an electron: an acceptor. In InP the situation is reversed: carbon occupies a group III position (In) and functions as a donor [22]. The dopant we use for n–type is Si, in both InP and InGaAs. It has low diffusivity though we have observed the doping junction can be found 4.5 nm away from the heterojunction [23] when using Si doping concentration of $1 \cdot 10^{19}$ cm^{-3} or more. Doping InGaAs higher than $\approx 3 \cdot 10^{19}$ cm^{-3} leads to increasingly poor surface morphology [24]. This puts a limit to the practical doping density in the subcollector region in the HBT layers grown upon it: a too high doping will lead to poor material quality in the HBT layers above it.

The dopant used for p–type is Be, Zn or C. Zn has very high diffusivity and a solubility limit $\approx 4 \cdot 10^{19}$ cm^{-3}. Be diffuses somewhat (≈ 5 nm) and a solubility limit $\approx 5 \cdot 10^{19}$ cm^{-3}, and is very toxic. C shows no diffusivity and is a n–type dopant in InP which makes the p–n junction coincide with the heterojunction. The solubility limit is higher than $\approx 1 \cdot 10^{20}$ cm^{-3} in lattice matched InGaAs [20], and $\approx 4 \cdot 10^{20}$ cm^{-3} in GaAs [32]. The main problem with carbon is hydrogen passivation (appendix B.1). Further, the gain for carbon doped InP HBT's has been lower than for corresponding Be doped HBT's [36, 47]. The carbon atom is smaller than the As atom, and at high doping levels the contraction of the material is measurable (figure 2.3). For our latest DHBT we increase the In to Ga ratio to compensate for this since the In atom is large than the Ga atom. Doping introduces defects in the lattice, and the mobility decreases with increased doping level (figure 2.4).

One distinction needs to be made about mobilities: majority and minority mobility. Majority mobility is the situation when the majority carrier is of the same polarity as the dopant, i.e. electrons in n-InP (figure 2.4). Minority mobility is the

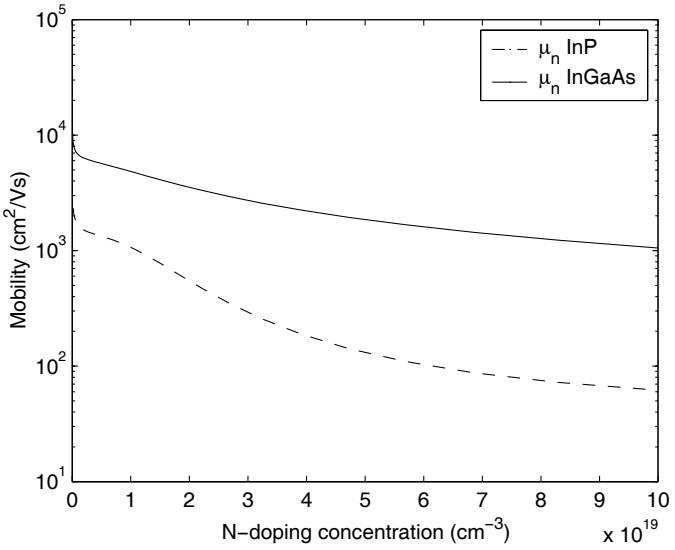

Figure 2.4. Electron majority mobility in n–InGaAs and n–InP as a function of emitter and collector doping level.

situation electrons encounter in the p–doped base. The minority mobility shows a small increase at very high doping densities due to screening (figure 2.5), [15]. The plotted data are from a compilation of published data.

Early experiments with carbon doped InP HBT's with base carbon doping showed lower gain and lower hole mobility than expected [64] coupled with high base resistance. The main reason is carbon passivation by hydrogen but there are also a number of other reasons.

Carbon doping in InGaAs is complicated by the fact that carbon is an amphoteric dopant in InAs, and the growth conditions need to be carefully adjusted to make sure carbon occupies the correct lattice position. In fact, doping of InGaAs might be thought of as doping of GaAs in an InAs lattice. Thus, in a base with varying degree of In to Ga ratio the carbon flux must be adjusted due to the different incorporation efficiency. Carbon is a weak n-type dopant in InP which in fact makes the crystallographic junction coincide with the electrical in an InP HBT. This makes it possible to achieve very highly doped regions with very abrupt p-n junctions [32, 34, 35].

The most severe problem with carbon doping of InGaAs is hydrogen passivation: hydrogen incorporated in the InGaAs material during growth or subsequent processing binds to the carbon atoms and negates the doping properties (appendix B.1). The carbon–hydrogen junction is by nature very strong and an annealing temperature of 400 degrees or more is needed to break the hydrogen carbon bond and cause the hydrogen to out-diffuse [14, 26]. Any hydrogen-passivated carbon

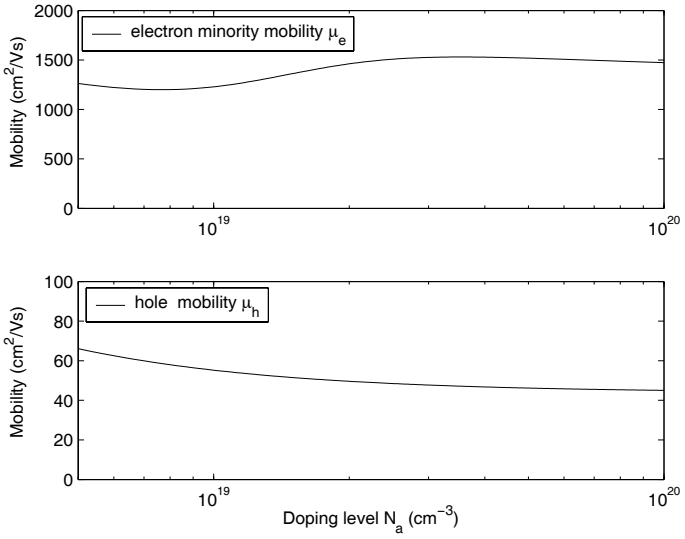

Figure 2.5. Electron minority mobility and hole majority mobility in p–InGaAs.

atoms still contribute to reducing the electron mobility in the base, and thus the situation can occur when a transistor has a low gain due to low base mobility as well as high base resistance due to low effective doping level in the base! Out diffusion of hydrogen through annealing is paramount for Metal-Organic Chemical Vapor Deposition (MOCVD) grown carbon doped layers due to the hydrogen containing precursor chemicals. For Molecular Beam Epitaxy (MBE) grown material the precursors do not contain hydrogen but incorporation of hydrogen can still occur during processing steps such as Chemical Vapor Deposition (CVD) of SiN or SiO.

Measurements show the bandgap shrinks for very high doping levels, so called BandGap Narrowing (BGN) [44, 45]. This effect is important in the base, where doping levels approach $\sim 1 \cdot 10^{20}$ cm^{-3}. From [45] a value of 2/3 is adopted for the ratio of bandgap reduction split between the conduction and the valence band: most of the band gap reduction will be in the valence band. (figure 2.1.2). For the highest base doping level used in this thesis, BGN shrinks the base bandgap with roughly 110 meV, and 70 meV of that is in the valence band.

2.1.3 Thermal properties

The thermal conductivity of several III–V material is shown in figure 2.7 [15]. The thermal conductivity of alloy materials such as InGaAs ($\approx$ 5 W/Km) and InAlAs ($\approx$ 10 W/Km) is much lower than the thermal conductivity of binary materials such as InP ($\approx$ 68 W/Km)or GaAs ($\approx$ 46 W/Km). The thermal conductivity

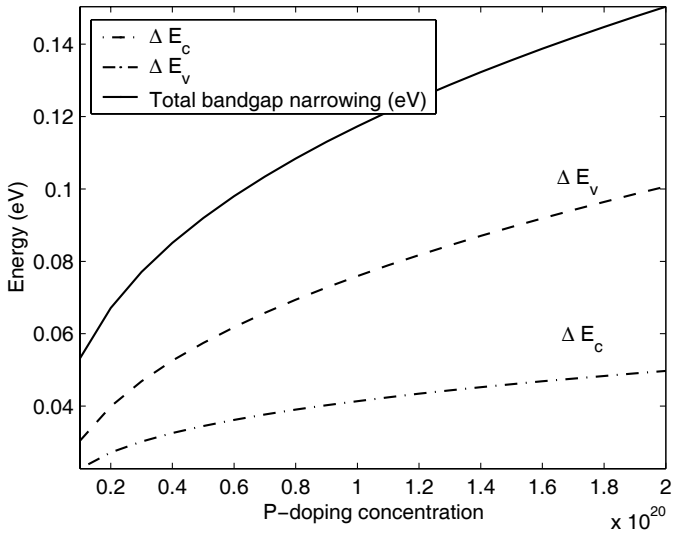

Figure 2.6. Bandgap narrowing (BGN) in InGaAs.

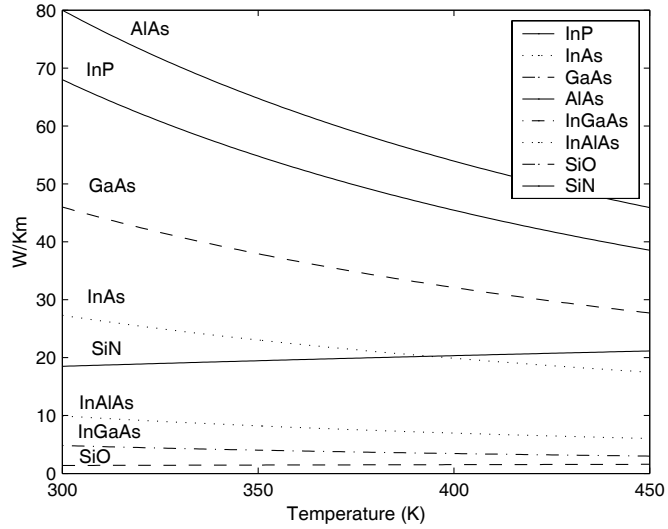

Figure 2.7. Thermal conductivity of common materials

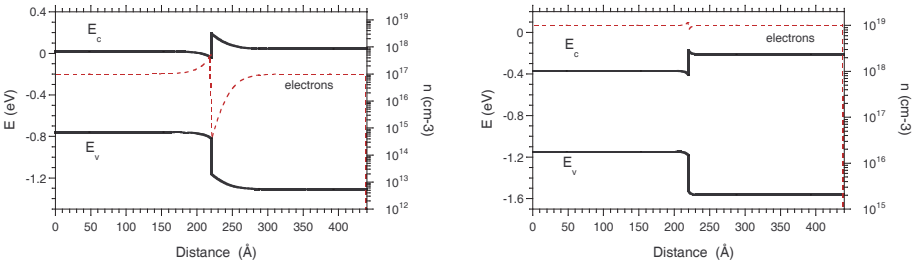

Figure 2.8. InGaAs/InP N-N junctions at different doping levels

is temperature dependant and increases slowly with temperature. The thermal conductivity of highly doped semiconductors is reported to be up to 30 % lower. However under certain conditions the carriers in a highly doped semiconductors can contribute to the thermal conductivity [18].

2.2 Heterojunctions

2.2.1 The isotype junction

The isotype junction represent a junction between two different materials but with the same type of doping. An example is the emitter region in a HBT, where the emitter cap is InGaAs and the emitter is InP (figure 2.8).

In the early DHBT designs the emitter region contained a grade between InGaAs and InP to smooth out the conduction band spikes, like the ones shown in figure 2.8. However, when the doping level is very high – $1 \cdot 10^{19}$ cm^{-3} or higher, the simulated band profile and carrier concentration shown in the right part of figure 2.8 suggests no grade is necessary. Measured emitter resistances are lower for devices without the grade, suggesting removing it did not make things worse at least.

2.2.2 P-n junctions

The governing equation for semiconductor materials is Poisson's equation, which describes the shape of the potential as a function of charge distribution.

$$\nabla \mathcal{E} = \frac{1}{\varepsilon_r} \left(q N_c(x) \right) \tag{2.1}$$

where $qN_c(x)$ represent the charge in the region is the governing relation for semiconductor junctions. The emitter–base and the base–collector junction have a depletion depth, over which the electric field is changing. Solving 2.1 with the boundary conditions that the electric field in $0, \infty = 0$ gives the following relation for the

depletion width:

$$x_d = \sqrt{\frac{2\varepsilon_r \left(V_{applied} + \phi_{bi}\right)}{qN_{e,c}}} \tag{2.2}$$

where the simplifying assumption that $N_{base} \gg N_e, N_c$ has been used. With $N_{base} \approx 4{-}8{\cdot}10^{19}$ cm^{-3} and $N_{emitter} \approx 5{\cdot}10^{17}$ cm^{-3} and $N_{collector} \approx 2{\cdot}10^{16}$ cm^{-3} that holds true. The built in voltage ϕ_{bi} can be defined as the difference between the conduction band in the first and second material, and can be calculated as:

$$\phi_{bi} = \frac{E_{g,b} + \Delta E_c - \Phi_p - \Phi_n}{q} \tag{2.3}$$

$E_{g,b}$ is the bandgap in the base, typically InGaAs with $E_g \approx 0.76$ eV, ΔE_c is the conduction band offset to InP, around 0.26 eV. Φ_p is the Fermi level position in the base from the valence band edge, and Φ_n is the Fermi level position in the collector (emitter) from the conduction band edge. When the doping levels are higher than the respective density-of-state (DOS) in the conduction or valence band Boltzmann statistics cannot be used. One good approximative method is the the Selberherr approximation [25] or numerical calculation of the full Fermi-Dirac statistics cans be used to calculate Φ_n or Φ_p. In effect, this only applies to the base Fermi level. For non–degenerate regions the electron and hole Fermi levels are described by:

$$\Phi_n = E_c - E_f = k_B T \ln \frac{N_c}{N_n} \tag{2.4}$$

$$\Phi_p = E_f - E_v = k_B T \ln \frac{N_v}{N_p} \tag{2.5}$$

, where N_c and N_v are the conduction and valence density of states. $In_{47}Ga_{53}As$ have $N_c = 2.48 \cdot 10^{17}$ cm^{-3}, $N_v = 4.70 \cdot 10^{18}$ cm^{-3}, and InP has $N_c = 5.38 \cdot 10^{17}$ cm^{-3}. In regions where the composition or temperature changes N_c and N_v should be calculated from

$$N_i = 2\left(\frac{2\pi k_B T}{h^2}\right)^{3/2} m_i^{*3/2} \tag{2.6}$$

where m_i are the electron and hole effective masses.

For doping densities above the intrinsic carrier concentration Boltzmann statistics is not correct to use, and the Joyce-Dixon or Selberherr approximations are often used or full Fermi-Dirac statistics. At degenerate doping levels the variation of the hole Fermi level Φ_p with doping level is strong, as shown in figure 2.2.2. Above $N_c = 10^{19}$ cm^{-3} the deviation from Boltzmann statistics is considerable. For a InGaAs/InP junction with p-doped InGaAs ($6 \cdot 10^{19}$ cm^{-3}) and n-InP doped at ($2 \cdot 10^{16}$ cm^{-3}) the junction built-in voltage ϕ_{bi} becomes 0.95 V. For the same junction with n-InP doped at ($5 \cdot 10^{17}$ cm^{-3}) ϕ_{bi} becomes 1.04 V. In the calculation above bandgap narrowing was calculated from [44, 45] and the InGaAs bandgap

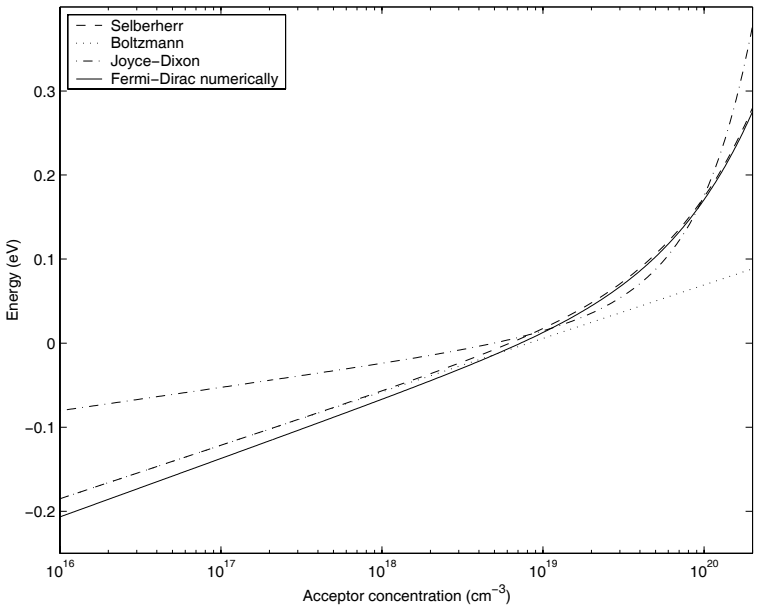

Figure 2.9. The hole Fermi level in InGaAs calculated with various approximations and Fermi Dirac statistics

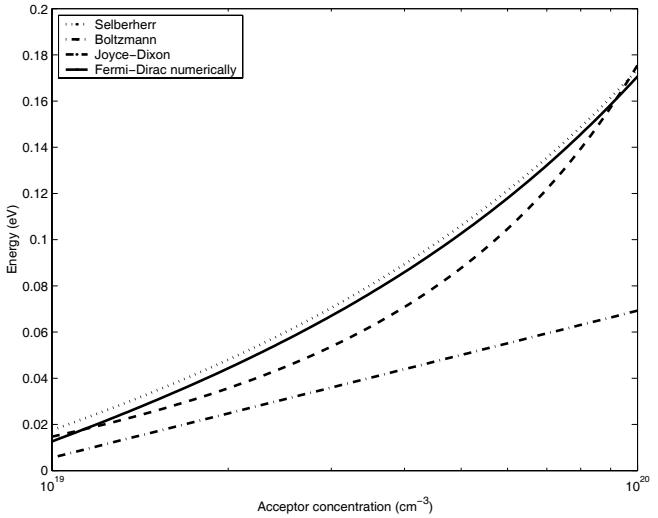

Figure 2.10. The hole Fermi level in InGaAs at base–like doping concentrations

reduced to 0.64 eV from 0.73 eV. The bandgap itself also has a weak temperature dependence, at room temperature $In_{53}Ga_{47}As$ has $E_g = 0.74$ eV, and at 100 °C $E_g = 0.71$ eV.

Since the built in voltage changes with base doping level (figures 2.3, 2.5) a HBT with a base doping of $N_c = 4 \cdot 10^{19}$ cm^{-3} has a smaller built in voltage than a HBT with a base doping of $N_c = 8 \cdot 10^{19}$ cm^{-3}. Judging from figure 2.2.2 $\Phi_p \approx 0.75$ eV in the first case and 0.83 eV in the second case. All else being equal this translates into a corresponding difference in base-emitter diode turn-on voltage. A device turns on when the applied bias is close to the built–in bias so that electrons can be injected into the base. More exact is the condition that the electron concentration at the emitter end of the base $n_b(0)$ is given by $n_b(0) \sim \exp(q(\phi_{bi} - V_{be})/k_bT)$, and $n_b(0)$ becomes significant when $\phi_{bi} - V_{be}$ is on the order of k_bT. k_bT at room temperature is 26 meV, meaning $V_{be} \approx \phi_{bi} - (0.05)$ V. For HBT's with graded base–emitter junctions, as verified by a diode ideality factor close to 1, we measure a turn-on voltage of ≈ 0.75 V for base doping of $N_c = 4 \cdot 10^{19}$ cm^{-3} while the turn-on voltage for $N_c = 8 \cdot 10^{19}$ cm^{-3} (DHBT-20) is ≈ 0.83 V. HBT's with abrupt emitter-base junctions have even larger turn-on voltage, closer to ≈ 0.87 V (DHBT-17,-18,-19).

The InP-GaAsSb junction has a much smaller built-in bias than InP-InGaAs, with a reported turn-on voltage of ≈ 0.4 V [4,66]. The reason for this can be derived from (2.3). Compared to InP-InGaAs, the conduction band step ΔE_c is negative with a reported value of $0.05-0.18$ eV [4,16,17]. The conduction band offset for InP-InGaAs is around 0.26 eV and positive. The difference in built-in voltage becomes, all else being equal, $\Delta E_c(\text{InP} - \text{InGaAs}) - \Delta E_c(\text{InP} - \text{GaAsSb}) = 0.26 - (-0.15) = -.41$ eV. Thus, we would expect the turn–on voltage for an InP-GaAsSb HBT to be -0.41 eV lower than the turn–on voltage for a InP-InGaAs HBT. The reported turn-on voltage for GaAsSb based transistors is ~ 0.4 eV, which is what is predicted.

The high doping in the base also means that the un–depleted base width does not vary appreciably with bias, as it would in a BJT with a low–doped base. This translates in a high Early voltage. The Early voltage $V_A$is defined as:

$$\frac{I_c}{V_A} = \frac{\partial I_c}{\partial V_{ce}}|_{V_{be}} \tag{2.7}$$

Flat curves give a high Early voltage and a current that increases with V_{cb} gives a low voltage. Our latest devices have an Early voltage of 10 V or more. The Early voltage at high current densities is hard to evaluate since thermal and high–injection effects change the apparent voltage.

A heterojunction can be made graded or abrupt. Figure 2.11 show how the base–emitter junction looks an InP-InGaAs HBT that is abrupt or graded. There are several implications of the choice between a graded and an abrupt junction. An abrupt junction is signified by the following:

- Higher turn–on voltage
- Simpler layer design

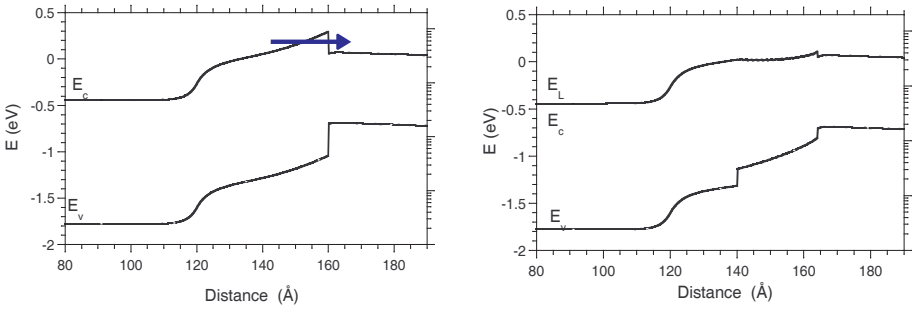

Figure 2.11. Abrupt (left) and graded (right) emitter-base junctions

- Sensitive for base dopant diffusion
- Exposed base surface
- Lower threshold for hole–back injection
- Ideality factor $n \to 2$

A graded junction is signified by the following:

- Lower turn–on voltage
- Need a non–electron trapping layer design
- Provides a barrier to base dopant diffusion
- Possibly protected base surface
- Higher threshold for hole–back injection
- Ideality factor $n \to 1$

An abrupt junction can provide higher gain since electrons are injected into the base with a minimum energy $\Delta E_{barrier}$, symbolized by the arrow in figure 2.11. It is an open debate how much this contributes to the current gain and base transit time of a HBT. For the latest HBT's we observe 30 % higher gain for submicron abrupt devices, and 50 % higher gain for large area devices.

We suspect that base–emitter surface leakage is an important effect lowering the gain, and the gain difference can be due to that as well.

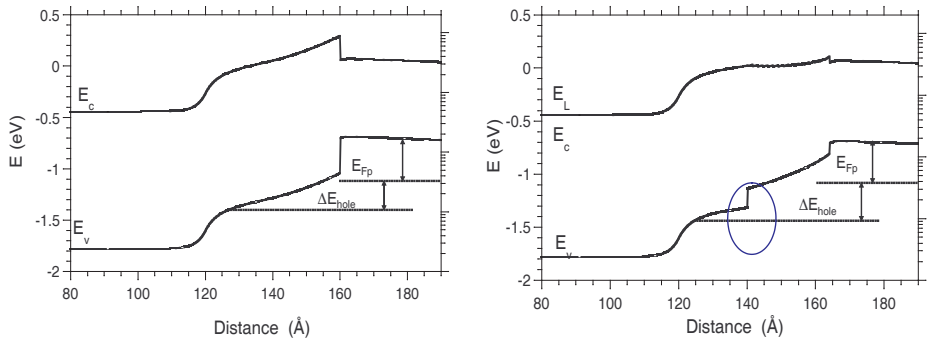

Figure 2.12. Difference in hole back–injection threshold for abrupt and graded emitter-base junctions

2.3 The HBT base region

2.3.1 Theoretical background

Based on the work of Kroemer [31] the electron and hole currents in the base of a HBT can be expressed as :

$$J_n = q\mu_n \frac{d}{dx}\left(\frac{\Phi_n}{q}\right) = \mu_n \frac{d}{dx}\Phi_n \tag{2.8}$$

$$J_p = q\mu_p \frac{d}{dx}\left(\frac{\Phi_p}{q}\right) = \mu_p \frac{d}{dx}\Phi_p \approx 0 \tag{2.9}$$

The hole current should be very close to zero in a proper HBT. If equation 2.8 is integrated over the base, with the boundary condition that the electron concentration of the emitter side is determined by the applied bias in base–emitter diode, $n(0) \sim \exp(qV_{be}/kT)$, we get the following:

$$J_n = -q\frac{\exp(qV_{be}/kT)}{\int_0^{T_b}[p/D_n n_i^2]\,dx} \tag{2.10}$$

The base transit time is the sum of the electron concentration in the base divided by the electron current;

$$\tau_b = q\frac{\int_0^{T_b} n(x)dx}{J_n} \tag{2.11}$$

Equation 2.11 forms the basis for the calculation of base transit time in section 2.3.4, where τ_b is expressed in a more general form.

2.3.2 General expressions for the base

The base of a modern InP HBT is 20-70 nm thick and doped above $1 \cdot 10^{19}$ cm^{-3} with beryllium, zinc or carbon. Due to the diffusivity of zinc and beryllium they are not used for doping levels higher than $5 \cdot 10^{19}$ cm^{-3}. From the viewpoint of a designer of a high speed HBT the base ought to be very thin ($\leq$ 40 nm) for decreased base transit time and increased gain, while at the same time the base resistance must be kept to a minimum. The base transit time can be expressed as [31]

$$\tau_b = T_b^2/D_e, b \tag{2.12}$$

where T_b is the base thickness and $D_{e,b}$ is the electron diffusivity [2] in the base.

[2] The diffusivity D needs to be calculated with Fermi–Dirac statistics if the carrier concentration is degenerate. In the base the hole concentration is but not the electron concentration [33]

The gain can be expressed as

$$\beta = \frac{2L_b^2}{T_b^2} \tag{2.13}$$

where $L_{e,b}$ is the base transfer length, $L_b = \sqrt{D_{e,b}\tau_{e,b}}$. $\tau_{e,b}$ is the electron lifetime in the base, inversely proportional to the doping level $N_{a,b}$ of the base (figures 3.8 and 3.9). The gain β can also be expressed as $\beta = \tau_n/\tau_b$, and can be thought of as the less time spent in the base the less time for recombination which would subtract from the gain. To achieve a HBT with high f_{max} it is imperative that the base sheet resistance R_s and the base contact resistance ρ_c are as small as possible. The base sheet resistance is

$$\rho_s \cdot \frac{1}{T_b} = \frac{1}{q\mu_{h,b}N_{a,b}} \cdot \frac{1}{T_b} \tag{2.14}$$

where $\mu_{h,b}$ is the hole mobility in the base, $N_{a,b}$ is the acceptor concentration of the base, and T_b is the base thickness (figure 3.7).

The contact resistance is proportional to the inverse square of the base doping level, $\rho_p \approx 1/N_{a,b}$ but is also heavily dependant on processing and metalization. A sheet resistance of 400-1000 $\Omega/\Box$ and a base contact resistance 1-500 $\Omega - \mu\text{m}^2$ is typical of a modern InP HBT.

For recent devices not suffering from hydrogen passivation, the hole mobility in the base is usually 50-55 $\mu\text{m}^2/\text{Vs}$ for carbon concentrations around $6{\cdot}10^{19}$ cm^{-3} [99], (table 5.7). Despite these good mobilities HBTs with carbon doped bases often show DC gain lower than expected [36]. As an example for two HBT's with the same base doping the beryllium doped HBT had a gain of 50, while the carbon doped device had a gain of 45, [28].

Oka [32] presents a theory of interest for highly doped bases in GaAs HBTs. For a very heavily doped base - with a doping level of the order of $1 \cdot 10^{20}$ cm^{-3} the Fermi level in the base moves deeply into the valence band due to strong degeneracy (figure 2.2.2). This results in an effective lowering of the valence band barrier for hole back-injection from the base to the emitter. This valence band barrier is the very essence of a HBT compared to BJT, that permits the base doping to be increased by a factor $\sim \exp(\Delta E_v/k_bT)$. In a BJT the current gain is limited by this hole back–injection. The results is a an increase in the reverse hole current, and a corresponding reduction of the gain of the device. At a certain doping level the Fermi level will move so deep into the base valence band that the hole barrier will be reduced to zero, and the gain of the device will be zero. Shown in figure 2.3.2 is the difference between the base Fermi level and the InP valence band. As the difference approaches zero the reverse hole current will increase and the gain of the device decrease. When the barrier becomes zero the HBT will have no gain.

For a full calculation of this one needs to incorporate the bandgap narrowing (BGN) effect in the base, and it's distribution between conduction and valence bands [44, 45] (figure 2.1.2). For the InP-InGaAs material system the onset of the collapse of current gain will happen at a base doping level of $2 \cdot 10^{20}$ cm^{-3} for an abrupt base-emitter junction.

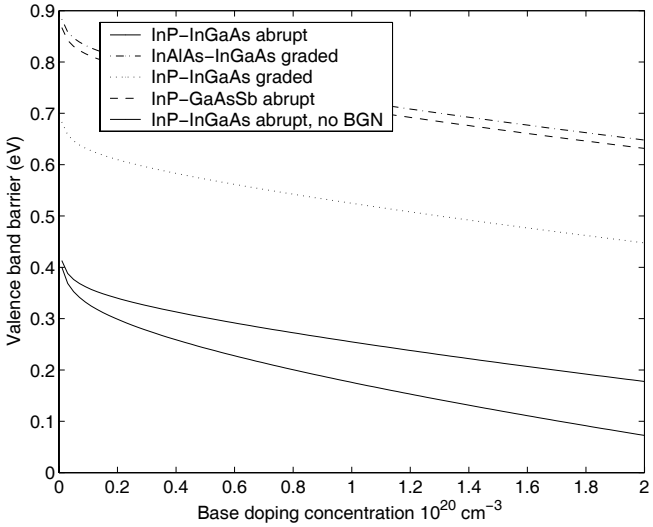

Figure 2.13. The hole barrier in an InP InGaAs junction taking degenerate doping and BGN in effect

For a graded base-emitter junction more or all of the total bandgap difference will be transferred to the valence band, and a higher doping level is possible, as illustrated in figure 2.12. Using GaAsSb/InP no grade is necessary since the conduction band difference is small and *negative*. The valence band difference is higher than InGaAs/InP (0.78 eV compared to 0.34 eV) and the critical doping density for an abrupt junction will occur at a higher base doping density. The key advantage of using GaAsSb is the band lineup, not only for the conduction band in terms of base–collector grade, but also this large valence band barrier.

2.3.3 Base grading

To reduce base transit time, and through $\beta = \tau_n/\tau_b$, where τ_n is the electron lifetime in the base, increase gain, the base can be graded as proposed by Kroemer [31]. The grading creates a quasi electric field that sweeps electron across the base. This can be achieved through either bandgap grading or doping grading. In bandgap grading the material composition is changed throughout the base, as in InGaAs where the In composition is changed from 53 % at the base-collector interface to 44 % at the emitter-base interface [1, 39, 47, 60, 64], which is the method used at UCSB [39,60]. Another approach is to have a quaternary base, e. g. incorporate Al in a GaAsSb base, in which case only a small amount of Al is necessary to create an electric field. Modern SiGe HBT have an increasing Ge content in the base, about 5-10 %.

Careful attention has to be paid to the strain and impurity effects these grading causes, but since the base is already heavily doped, no large effects are visible in hole or electron mobility in well-calibrated growths . The other approach to create a base grade is to change the base doping level through the base, heavy doping at the emitter-base interface and lower doping at the base-collector interface [41–43, 99]. Si-based BJTs often employ an exponentially decreasing base doping to create a linear electric field. However, at degenerate doping levels (above $\approx 1 \cdot 10^{19}$ cm^{-3}) a linear change in doping level is enough to create a strong electric field [99]. In this way a strong electric field can be created without paying a large penalty in base resistance, which otherwise is the big drawback with exponentially varying doping. Figure 2.2.2 shows how strong the variation of Fermi energy is at very high doping levels. This fact is often overlooked in device design or device simulators that rely on Boltzmann statistics or other simplifying assumptions.

Another way of looking at the base grading is that it permits a thicker base than would otherwise be possible, due to the advantages it creates in base gain and base transit time, and the base resistance can be decreased. Alternatively the base doping can be made higher than otherwise, benefiting the base contact resistance.

The built-in field from bandgap grading is, if we for the moment neglect the differences in mobilities and intrinsic carrier concentrations [31]:

$$\mathcal{E} = \frac{\Delta E_g}{T_b} \tag{2.15}$$

A doping grade also introduces a built-in field [43] (neglecting band gap narrowing effects):

$$\mathcal{E} = \frac{d\Phi_p}{dx} = -\frac{kT}{q}\frac{F_{1/2}(N_a(x))}{F_{-1/2}(N_a(x))} \tag{2.16}$$

If the doping levels are non-degenerate and the doping roll-off is exponential, as in $N_a(x) = N_a(0)\exp(-x/\tau_0)$, where τ_0 is a measure of the doping width, equation 2.16 simplifies to:

$$\mathcal{E} = \frac{kT}{q}\frac{1}{N_a}\frac{dN_a}{dx} = -\frac{kT}{q}\frac{1}{\tau_0} \tag{2.17}$$

That the field is independent of N_a in (2.17) might look surprising at first, but the field corresponds to the gradient of the dopant, which is constant for an exponential profile (see figure 2.2.2). This changes at degenerate doping densities (as the Fermi level leaves the line defined by the Boltzmann expression).

2.3.4 Calculation of base grade and base transit time

The base transit time in a HBT is given by Kroemers double integral [31]:

$$\tau_b = \int_0^{T_b} \frac{n_i^2(x)}{N_a(x)} \left(\int_0^{T_b} \frac{N_a(z)}{n_i(z)^2 D_{n,b}(z)} dz \right) dx \tag{2.18}$$

N_a is the acceptor dopant concentration in the base, n_i is the intrinsic carrier concentration, that can also be used to model bandgap variations. $D_{n,b}$ is the diffusivity of the electrons in the base, and is also the term with the strongest temperature dependence through Einsteins relation: $D_{n,b} = k_B T/q\mu_{n,b}$, where $\mu_{n,b}$ is the electron mobility in the base. Kroemers integral should be put in relation to the expression for base transit time in the constant structure case

$$\tau_b = \frac{T_b^2}{2D_{n,b}} \tag{2.19}$$

Equation 2.18 simplifies into (2.19) when N_a, n_i and $D_{n,b}$ are all constant. For small base widths a correction term needs to be added to equation 2.18 due to the finite exit velocity for electron on the collector side of the base [31]. The reason is the breakdown of the boundary condition that the electron concentration at the collector end of the base should be zero – it is not and the concentration of electrons there is dependant on which velocity they might be moved away with.

$$\tau_{exit} = \frac{1}{v_{sat}} \int_0^{T_b} \frac{N_a(T_b) n_i^2(x)}{N_a(x) n_i^2(T_b)} dx \tag{2.20}$$

When all terms are constant equation 2.20 simplifies to $\tau_{exit} = T_b/v_{sat}$ The total base transit time is the sum of τ_b and τ_{exit}.

$$\tau_{b,tot} = \tau_b + \tau_{exit} \tag{2.21}$$

To calculate the base transit time for different base structures a general derivation of the base transit time was made, relying on numerical calculation based on a model of relevant parameters. All parameters such as mobility, diffusivity, temperature, material composition and temperature was fitted using fits of the type $f(x) = e^{Az+B}$. The fit is good when the parameter is monotonic over the base. Parameters that are strongly Gaussian in shape such as an implanted dopant profile or the dopant profile of a highly diffusive material such as zinc are not of interest for these calculations, that are intended for epitaxial bases with non-diffusive dopants such as carbon. Using $N_a = e^{N_{a1}z+N_{a2}}$, $n_i^2 = e^{N_{i1}z+N_{i2}}$, $D_{n,b} = e^{D_1 z+D_2}$ (2.18) turns into:

$$\frac{e^{-D_1 T_b - D_2} N_{a1} - e^{-D_1 T_b - D_2} D_1 - e^{-D_1 T_b - D_2} N_{i1} + N_{i1} e^{-D_2} - N_{a1} e^{-D_2} + \beta}{(-N_{a1} + D_1 + N_{i1}) D_1 \ (N_{i1} - N_{a1})} \tag{2.22}$$

where $\beta = D_1 e^{-N_{i1} T_b - D_1 T_b - D_2 + N_{a1} T_b}$.

And the exit term τ_{exit} becomes.

$$\frac{N(T_b) T_b \left(e^{N_{i1} T_b + N_{i2} - N_{a1} T_b - N_{a2}} - e^{N_{i2} - N_{a2}}\right)}{2 v_{exit} T_b \ (N_{i1} - N_{a1})} \tag{2.23}$$

The model for material parameters was taken from [15–17]. The thermal gradient over the base is taken from [21] and represents a significant correction to the

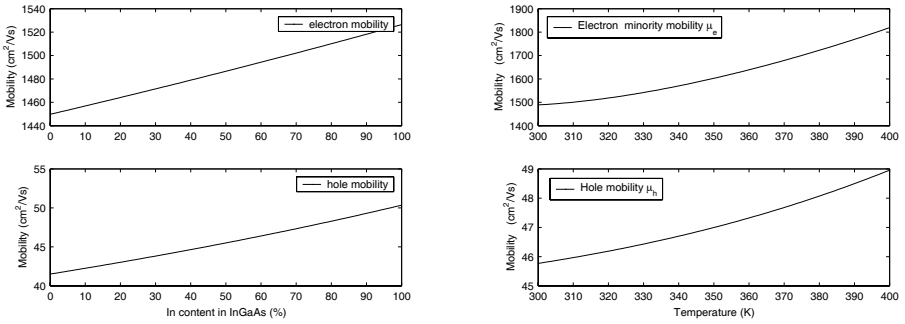

Figure 2.14. Mobility in InGaAs as a function of lattice composition and temperature

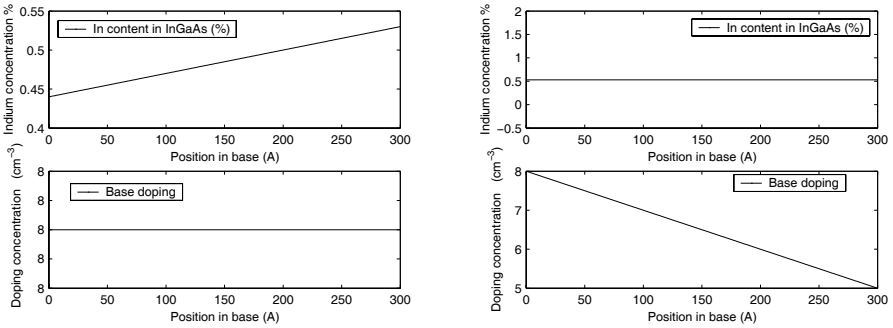

Figure 2.15. Simulation setup for a bandgap graded base (left) and a doping graded base (right)

calculation due to the strong dependence of diffusivity on temperature. The calculation is done using Fermi-Dirac statistics. Shown in figure 2.14 is the mobility in InGaAs as a function of In to Ga ratio, and as a function of lattice temperature for the lattice matched case. The calculated electron minority mobility in InGaAs and majority mobility in InP as a function of base and collector doping level is shown in figure 2.4. The simulation setup regarding base lattice composition and doping is shown in figure 2.15. The resulting conduction band profile in the base is shown in figure 2.16 where also the conduction band profile would bandgap narrowing not be used is shown. The resulting base electron minority mobility and hole majority mobility in the base is shown in figure 2.17 as a function of position.

The resulting electric field in the base is shown in figure 2.18 where also the resulting electric field would degenerate statistics not be used is shown. The figure shows the field is significantly weaker for non–degenerate statistics.

This is the key observation when it comes to highly doped bases and resulting quasi fields.

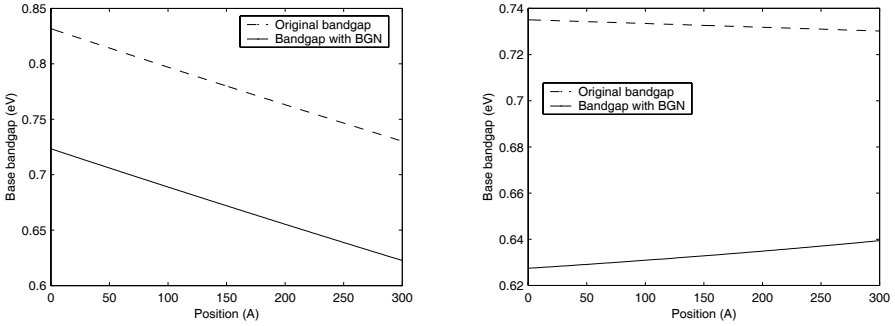

Figure 2.16. Effective base bandgap for a bandgap graded base (left) and a doping graded base (right)

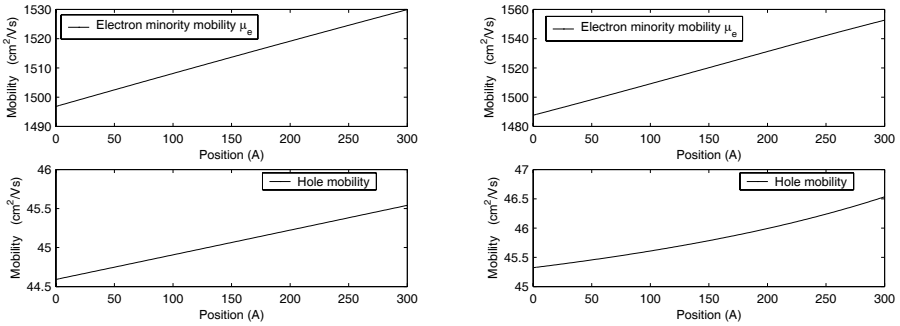

Figure 2.17. Electron and hole mobility as a function of position for a bandgap graded base (left) and a doping graded base (right)

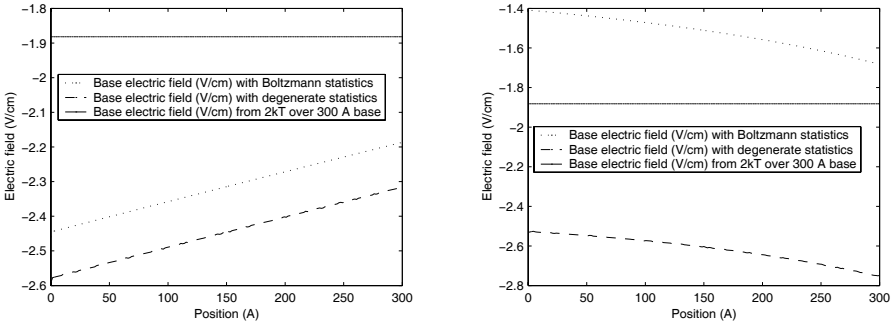

Figure 2.18. Resulting base electric field for a bandgap graded base (left) and a doping graded base (right)

The internal base transit time has a $\sim T_b^2$ dependence and the exit term a $\sim T_b$ dependence, and at very thin base widths the relative importance of the exit term is increased. In the calculation an exit velocity of $3 \cdot 10^5$ m/s was assumed [1], which it is a typical saturation velocity in InGaAs, that forms the top part of the collector. Under conditions that form a conduction band barrier in the base-collector interface – either a badly designed or grown base–collector structure or under strong injection conditions so the Kirk effect forms it is possible that the exit velocity term increases. However, to substantiate this we would need to extract τ_b from τ_{bc} but the author is unaware of a method for this. To control the validity of the base transit time calculation routine the gain difference ($\sim 1/\tau_b$) was calculated for a structure reported by Ohkubo *et al.* [47]. This structure has a base thick enough (100 nm) that the influence of ballistic electrons injected from the base–emitter interface is non–dominant. The structure has an abrupt base-emitter junction and a 100 nm thick base, either bandgap graded or ungraded. The grade was achieved by changing the In composition in the InGaAs base from 42 to 49 %. The difference in calculated base transit time using (2.22,2.23) would predict a gain difference of 1.73, very close to the reported gain difference of 1.70.

The conclusion from this is, that for bases 100 nm thick the calculation of transit time holds true. What can offset the calculated base transit time from the real transit time is the influence of energetic electrons injected into the base by means of an abrupt base–emitter heterojunction. However, the evidence here is that the influence of the abrupt junctions is the same for both types of base. The base transit time is not fully described by the expressions for τ_b and τ_{exit}. The conduction band difference in an abrupt base presents a barrier that electrons need a certain energy to $\Delta E_{barrier}$ to traverse. The electrons are then entering the base with an velocity close to:

$$\sqrt{\frac{2\Delta E_{barrier}}{m_e^*}} \tag{2.24}$$

The electrons quickly undergo scattering with the lattice and impurities such as dopants and it is dependant on average length before scattering (L_d) how large the influence of the ballistic injection will be. From a base dominated by ballistic transport an $\sim T_b$ dependence on base thickness is expected for base transit time, and indirectly, gain. It is possible that for very thin bases the influence of the launcher of the abrupt junction becomes dominant over any grades in the base. If a majority of the electron injected travel through the base without undergoing scattering the influence of the abrupt junction - the launcher - will dominate. Teissier *et al.* [49] report the scattering relaxation length L_d to be 18 nm at room temperature for a base doping level of $1 - 2 \cdot 10^{19}$ cm^{-3}. The scattering relaxation length is reported to be rather constant over the entire 10^{19} cm^{-3} doping range [48]. Using $\eta = e^{(-T_b/L_b)}$, where η represents the proportion of electrons in the base not undergoing scattering events and retaining the entry velocity $\Delta E_{barrier}$. η becomes 19-25 % for 25 and 30 nm bases, respectively. For the structure reported by Ohkubo *et al.* [47] and investigated above, the corresponding value would be 4 %. Ito [35] find that for

bases thicker than 20 nm with abrupt emitters, the base transit time $\tau_b \sim 1/W_b^2$, and for bases thinner than 20 nm $\tau_b \sim 1/W_b$. The dependence for thicker bases is what one would expect for diffusive transport (equation 2.19).

For thinner bases the $\tau_b \sim 1/W_b$ dependence can be explained both by ballistic transport from the launcher and by the exit term in the base transit time (equation 2.20).

The effective energy by which the electrons are launched into the base is however not equal to the entire InP-InGaAs conduction band offset. Because of the high doping levels involved tunneling will substantially lower the launching threshold, from $\approx \Delta E_c = 0.26$ eV to $\Delta E_{barrier} \approx= 0.12$ eV [48] and the electron energy retained at the base-collector interface $\approx \Delta E_{c,\text{effective}} e^{(-T_b/2L_b)}$ which is on the order 50-60 meV, similar in energy to the grade. Comparing devices with abrupt and graded base–emitter junctions, with the same base doping, the graded devices show a turn-on voltage ≈ 60 meV lower.

Calculation of base resistance

In an actual HBT is the base resistance R_{bb} is composed of the sum of contact resistance R_c, base-emitter gap resistance R_{gap}, and spreading resistance under the emitter R_{spread}. With base sheet resistance ρ_s, and specific (vertical) contact access resistance ρ_c, we have

$$\begin{aligned} R_{bb} &= R_{b,cont} + R_{gap} + R_{spread} \\ R_{b,cont} &= \sqrt{\rho_s \rho_c}/2L_e \\ R_{gap} &= \rho_s W_{eb}/2L_e \\ R_{spread} &= \rho_s W_e/12L_e. \end{aligned} \tag{2.25}$$

ρ_s is the sheet resistance in the base, $\rho_s = (q\mu_h N_a T_b)^{-1}$. The base sheet resistance does not have to be the same in the different regions. Specifically, we observe evidence ρ_s is higher in the base–emitter gap region than under the emitter. L_e is the emitter length. W_e is the emitter width, W_{eb} is the base–emitter gap between the emitter semiconductor and the base metal, W_b is the base metal width, that only comes in if the base contact resistivity is high, in which case the expression for $R_{b,cont}$ is changed to:

$$R_{b,cont} = \frac{\sqrt{\rho_s \rho_c}}{2L_e} \coth\left(W_b \sqrt{\frac{\rho_s}{\rho_c}}\right) \tag{2.26}$$

An important observation is the strong $1/W_e$ dependence: narrow emitters decrease the base resistance – an important argument for increased device scaling. For calculation of expected device performance we can use the ρ_s calculated in the same sequence as the base transit time or measurements on similar structures. For

ρ_c we rely on measured values from earlier samples with the same doping level. In early devices the R_{bb} values were very high, and great attention was paid to minimizing it. The $R_{b,cont}$ term was large due to the often large contact resistance to p-InGaAs, and led to an investigation of base contact metalization schemes (page 33).

Of the different terms in equation 2.25 the $R_{b,cont}$ term and the R_{gap} are most influenced by processing. The gap term R_{gap} is dependant on the amount of undercut created in the base–emitter junction, W_{eb}. We typically observe an undercut on the order of 50-100 nm using SEM. We typically aim to create a very small undercut to reduce the R_{gap} term. The choice of etching procedure and emitter cap thickness and composition play an important role in this.

In devices with thin bases (less than 40 nm) and with abrupt emitter-base junctions the sheet resistance ρ_s can be substantially different between the *intrinsic* region beneath the emitter-base junction and the *extrinsic* region between the base contact and the emitter-base junction. The resistance is always higher in the extrinsic region, and (section 5.3.1, table 5.5) a combination of wet etching and surface depletion seems to be the reason. As the base becomes progressively thinner the surface depletion cannot be ignored despite the very high doping levels. The p-side depletion depth in an abrupt junction can be expressed as

$$x_d = \sqrt{\frac{2\varepsilon_r \left(V_m + V_{bi}\right)}{qN_a}} \tag{2.27}$$

For an InGaAs base doped to $6{\cdot}10^{19}$ cm^{-3} the value of x_d is around 4 nm, depending on the exact position of the surface pinning potential. From TLM measurements we find that in average 5.1 nm of the base was missing for abrupt base–emitter junctions where the base surface is exposed. For graded base–emitter junctions where the grade is kept in place as a ledge the average distance is 1.7 nm.

2.4 Contacts

2.4.1 Overview of semiconductor metal contacts

A metal-semiconductor junction can be described as a metal with a certain work function Φ_m in contact with a semiconductor. In equilibrium the semiconductor Fermi level will line up with the metal Fermi–level, and the band offset is determined by the difference between the metal work function and the semiconductor band lineup. The line-up is described as surface pinning, and is a function not only of metal and semiconductor but also of processing. Normally this creates a depletion region and a barrier that the carriers (electron or holes) need to tunnel through. The lowest contact resistance is achieved when the barrier for carrier transport is as thin as possible, and this occurs if the line-up pins the in or near the conduction(valence) band for electron (hole) transport, the doping level is high, the semiconductor has a narrow bandgap and there is no interfacial oxide.

The semiconductor with the most potential for a good contact is n-type InAs where the line-up is such the surface pinning occurs in the conduction band, and no depletion region is thus formed. There is evidence that the best p-type contact would occur for GaSb due to similar pinning in or close to the valence band, but reported data for device contacts has not been quite so favorable, possibly due to processing problems with GaSb. Several theoretical models have been proposed for contacts but the actual correlation to device experiments has been poor, and instead the following conclusions can be made:

- Contact resistance is inversely proportional to the tunneling probability
- Electron tunnel more easily than holes, and smaller effective mass is better
- The tunneling probability is proportional to the root of the doping level $\sqrt{N}$.
- The tunneling probability is inversely proportional to the bandgap, favoring low bandgap materials such as InGaAs or InAs over InP or InAlAs
- Any intermediate oxide will reduce the tunneling probability

The ideal contact should thus contain no oxide, the band line-up should be favorable, the semiconductor should be doped as high as possible, and the bandgap should be as small as possible. Only to the top 10-20 nm of the semiconductor need to fulfill these properties. From a device viewpoint the contacts need to be thermally stable up to the maximum processing and operating temperature, usually 350 and 200 °C respectively. Further, for contacts to thin doped layers such as the base in a HBT, the contacts metallurgical reaction depth must be limited to avoid punch-through or reliability problems. Finally, the contact metal itself should provide low electrical and thermal resistance and be compatible with the processing environment.

2.4.2 Semiconductor metal reactions

There exist many metalization schemes for III-V semiconductor fabrication. Au, AuGe, Sn, Ti, TiW, Zn, Pt are among those used, often in different sequences. They are usually divided into two main groups: *alloyed* and *non-alloyed* contacts.

In *alloyed* contacts an alloy is formed between the metals and the semiconductor during an annealing step. The reaction usually progresses to a substantial depth and is limited by annealing conditions or when all the metal has reacted. AuGe is the most common of these reactions. They are not suitable for HBT fabrication with the possible exception of the collector contact. The reason is the deep reaction - often of the order of 200 nm - that makes it impossible to control it for emitter and base contacts.

Non-alloyed contacts only have a very limited surface reaction - sometimes on the order of the thickness of the surface oxides present, on the order of 1-2 nm. Tungsten is a prime example of a *non-alloyed* contact metal. Ti-based contacts

such as TiPtAu are also included in this group, even though the surface reaction of Ti penetrates the oxide. While the W-based contacts can be annealed at temperatures high enough to cause the III-V material to disintegrate the Ti-based contacts start reacting at 350-400 °C [51]. A note of caution must be addressed concerning metalization schemes encountered even in recent literature and it is the use of gold-based contacts without a gold diffusion barrier. A TiAu or CrAu contact will not survive an anneal higher than 200 °C without the gold penetrating the Ti or Cr layer and diffuse deep into the semiconductor. A barrier layer such as Pt or Pd must be used [53]. This has resulted in TiPtAu or TiPdAu contacts, often used for emitter and base contacts. Such a contact can easily provide a contact resistance to n-InGaAs in the low $\approx 10^{-7}$ Ωcm^{-2} range. High n-type doping level and good surface preparation is as always necessary.

Good p-type contacts have been more difficult to obtain with contact resistances a magnitude or more higher. This is expected because of the larger effective mass of holes. PtTiPtAu or PdTiPtAu contact have been shown to provide better contact resistance [52, 55] to p-type material, and the reason for this has been ascribed to the metal work function of Pt and Pd that provides a better band-lineup to the valence band. A more pragmatic explanation lies in the metal–semiconductor reactions that take place and change the semiconductor into a semi–metal.

Work in GaAs HBTs reported different semi-metallic phases being formed between the Pt or Pd and the GaAs [54–56] or InP [57]. Immediately after deposition phases such as PdGa are formed, and upon annealing several different phases are formed. They are usually of the form Pd/Pt-group III-group V, as in PdGaAs_x, where x is less than one. Different compounds are usually observed, with different ratio of the constituents. The depth of this reaction is on the order of 2-10 nm. The metallic phase is usually lacking in arsenic, that moves upwards in the structure until it reaches the Ti-interface where it remains. The Pd or Pt layer above the Ti is necessary as an Au diffusion block [53].

Chor *et al.* [50, 51] reported very good contact resistances to p-InGaAs using PdTiPdAu contacts and the p-type contacts done in this thesis are based on that. The depth of reaction of a PdTiPdAu or PtTiPtAu contact can be controlled by the thickness of Pd or Pt. In our experience the depth of reaction is roughly 4 times the Pd thickness, less for Pt. The depth of the Pd-diffusion can be judged by the reverse leakage current in the base–collector diode. For a 25 nm thick base that accidently received 8 nm Pd instead of 3 nm the leakage increased by an order of magnitude after annealing at 300 °C.

Palladium and platinum are both in the same group in the periodic system and have similar chemical properties. Palladium has a lower melting point than platinum (1552°C compared to 1769°C) and evaporates quite a bit easier (vapor pressure of 26 μmHg compared to 0.17 μmHg [51]). Since all metal deposition for contacts in this work was done be electron beam deposition this means less power can be used during the deposition, which results in less out-gassing and possible heat-related problems with the photoresist. TiPtAu was therefore exchanged

to TiPdAu and PdTiPdAu. For that reason part of the work of this thesis was dedicated to the development of good Pd-based p-contacts.

2.4.3 The base contact

In our base contacts we use Pd/Ti/Pd/Au of 3/15/15/70 nm thickness for placement directly on the base. For DHBT-20 we employed a 10 nm thick Pd layer annealed through the grade down to the base. The lowermost Pd reacts with the InGaAs base, and the Ti layer above stops out-diffusion of As, and prevents the above lying Pd from reacting. If not present, the upper Pd layer would continue to move into the base and create a base–collector short. The purpose of the second Pd is to prevent Au diffusion into the base. Gold will react quickly with InGaAs at rather low temperatures [53] and create a base–collector short. The thickness of the Ti and the second Pd layer are chosen empirically to be as thin as possible and still be good diffusion barriers. No specific study has taken place on the minimum thickness needed. The gold thickness is decided by the need to avoid emitter–base shorts. The contact is self–aligned, meaning the undercut emitter–base junction creates a gap in the base metal, that otherwise would cover it entirely (figure 1.1). However, if the base metal thickness is more than the InP thickness in the emitter there is a risk the base metal touches the highly doped InGaAs emitter cap layer or the emitter metal, resulting into en emitter–base short. This thin layer of gold has a resistance ≈ 50 % higher than expected (figure 5.16). The overall base metal resistance is about 0.5 $\Omega/\square$, which is not negligible.

2.5 The collector

2.5.1 Overview of the collector

The collector is probably the HBT region that is changed the most to suit the designers goal: a thicker collector directly translates into higher breakdown for power devices, and a thinner collector results in higher f_t but increased C_{cb}. The doping can be increased to increase the current tolerance, or be minimized to achieve full depletion over a wider range of base–collector bias. Two separate types of InP HBTs exist: SHBT's and DHBT's. SHBT's (Single Heterojunction Bipolar Transistor) use the same material, e.g. InGaAs in the base and the collector. DHBT's (Double Heterojunction Bipolar Transistor) use a wide bandgap material such as InP in the collector.

2.5.2 Collector design

SHBTs have achieved very impressive device performance [61] and are comparably easier to design and grow. The collector is of InGaAs and in the growth all that is needed is to change the dopant sources to grow the collector to base region and probably lower the temperature in the base growth. The main problem of InGaAs is its low breakdown. A 200 nm thick collector can only take around 2.2 V before the current starts to avalanche, whereas the same thickness for a composite InP collector can take up to 7.5 V [99]. This prohibits using even thinner collectors, that otherwise would have permitted higher current densities (see section 3.4). A second problem is the low thermal conductivity of InGaAs (figure 2.7), which is around 1/15 of InP's thermal conductivity. At high current densities heat management becomes a critical issue.

2.5.3 Base collector grade

To use an InP collector and an InGaAs base the conduction band discontinuity between InGaAs and InP must be removed or else device performance will be severely degraded. The discontinuity will increase collector transit time and electron–hole recombination due to electron trapping. The value of the discontinuity to InP is 0.26 eV for low doped $In_{53}Ga_{47}As$. The very high doping levels used here might increase the discontinuity by 29 meV [44, 45]. Three main approaches exist to deal with the discontinuity:

- InGaAsP grade with intermediate bandgap
- InAlAs/InGaAs super lattice grade with intermediate effective bandgap
- InGaAs setback layer

The first approach is particularly suited for MOCVD growth, and the type of quaternaries are often grown for InP lasers. For example, Kurishima *et al.* employed

three InGaAsP layers [14, 26, 35, 64], the layer closest to the base is often InGaAs ($E_g \approx 0.76$ eV, and then staggered to approach the bandgap of InP ($E_g = 1.35$ eV).

The second approach is easier to accomplish with MBE and is used at UCSB and it is to use a chirped super lattice of thin layers of InGaAs/InAlAs. Attempts have been made to use InGaAs and InP [65] but due to the inter diffusivity of arsenic and phosphorous the lattice quality becomes very bad. An all–arsenic grade with InAlAs and InGaAs does not have this problem. Design of the grade is critical and part of this thesis is designated to that.

The third approach is to keep InGaAs next to the base, and switch over to InP further down in the collector. If the potential drop from the conduction band in the base down to the onset of InP is in excess of ΔE_c much of the current blocking effect of the conduction band offset will be removed. The fastest HBT's reported use method 1 or 2 but combined with a InGaAs setback.

2.5.4 The grade

A conduction or valence band offset can be neutralized by a grading region with a length T_{grade} followed by δ-dopings on either side [9, 70]. The necessary doping to balance out the offset is given by [46]:

$$N_\delta T_\delta = \frac{\varepsilon_r \Delta E_c}{q^2 T_{\text{grade}}} \tag{2.28}$$

For a 24 nm thick grade N_δ becomes $9 \cdot 10^{10}$ cm^{-2}, or for a δ-doping thickness of 3 nm, $3 \cdot 10^{18}$ cm^{-3}.

If the grade is a super lattice — that is very thin layers of e.g. InAlAs and InGaAs sandwiched together — there is a risk of creating quantum wells in the InGaAs surrounded by InAlAs (wider–bandgap) barriers. The lattice is usually constructed with a fixed period width T_{period}, with varying widths of InGaAs and InAlAs. Next to the InGaAs base the majority of the period is InGaAs, and in the other end, a majority is InAlAs. The final composition depends on the desired effective bandgap. One needs to make sure the super–lattice does not act like quantum wells so there should be no allowed quantum states. The solution for the energy Eigenstates for Schrödinger's equation for an well surrounded by infinite barriers is :

$$E_n = \frac{\pi^2 \hbar^2 n^2}{2 m_e T_{SL}} \quad n = 1, 2, 3 \ldots \tag{2.29}$$

Where T_{SL} is the thickness of the widest InGaAs layers, m_e is the electron effective mass. If the energy for the lowest Eigenstate (n=1) is less than the conduction band offset between InAlAs and InGaAs ($\Delta E_c = 0.47$ eV) a quantum well is formed and trapping can occur (figure 2.19). Solving 2.29 shows that the maximum allowed InGaAs layer width is 2.5 nm for hole trapping and in the range 3.2 nm for electron trapping. The answer depends on which effective mass to use. In reality equation 2.29 relies on a number of simplifications; the barriers are not infinite, and

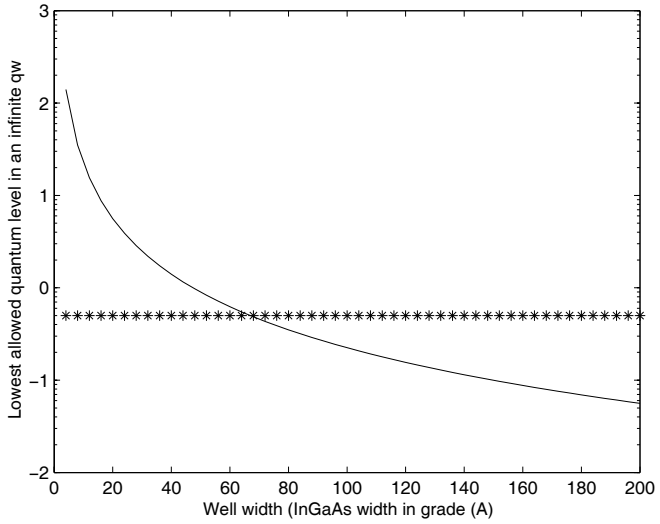

Figure 2.19. Minimum allowed quantum well width for electron trapping in InGaAs/InAlAs

the conduction band offset and effective mass are not well defined inside the super lattice. However, Nguyen *et al.* [67] reported evidence of quantum well trapping for T_{SL}=2.5 nm, while T_{SL}=1.5 nm showed no such effects. Our latest grade designs use T_{SL}=1.5 nm.

In constructing such grades one needs to keep in mind the interatomic distance of InGaAs/InP. The distance between a group III atom and a group IV atom is 0.25 nm, and defining layers in a super lattice thinner than that is pointless. This puts a limit on the range of bandgaps achievable with a certain super lattice period (this assumes the island formation during growth results in mono–layer islands much larger than a base–collector junction, otherwise if the islands are much smaller than the base–collector junction super lattice layers much less than one monolayer can be used as they would average. The first is true as far as I know from MOCVD).

Our latest HBT's have an InGaAs setback layer between the base and the grade (table A.3) and the models for bandgap narrowing (BGN) indicate a small conduction band offset also there. A base doping level of $6 \cdot 10^{19}$ cm^{-3} gives a conduction band offset of 36 meV [44]. Any such discontinuity can be offset by changing the lattice composition of the base to Ga-rich [3].

A special mentioning must be made about GaAsSb base HBT's. In these HBT's, pioneered by C. Bolognesi *et al.* the base is of $GaAs_{50}Sb_{50}$, that has a *negative* conduction band discontinuity to InP of 0.05-0.18 eV [4, 66], (figure 1.5). No base-collector grade is necessary.

2.5.5 The collector transit time

The collector transit time can be expressed as [1, 19, 30] :

$$\tau_c = \frac{1}{T_c}\int_0^{T_c} \frac{T_c - x}{v(x)} dx \tag{2.30}$$

In the case where the collector velocity can be assumed to be a constant the equation 2.30 simplifies into

$$\tau_c = \frac{T_c}{2v_c} \tag{2.31}$$

The important conclusion from (2.30) is that the region next to the base is most important for the collector transit time. The velocity in that region is also important for the base transit time in the expression for the base exit time (2.23).

To investigate the importance of high collector velocity near the base lets use a model where the collector velocity is a two-step function, v_1 in the region $x = [0 \dots T_1]$ and v_2 in the region $x = [T_1 \dots T_c]$. A more complete treatment has been made by Ishibashi *et al.* [19]. (2.30) becomes

$$\tau_c = \frac{1}{T_c}\int_0^{T_1} \frac{T_c - x}{v_1} dx + \frac{1}{T_c}\int_{T_1}^{T_c} \frac{T_c - x}{v_2} dx \tag{2.32}$$

Solving the equation gives:

$$\tau_c = \frac{1}{T_c}\left[\frac{2T_cT_1 - T_1^2}{2v_1} + \frac{(T_c - T_1)^2}{2v_2}\right] \tag{2.33}$$

To show two possible scenarios let first $v_1 = 5v_0$, $v_2 = 3v_0$ and $T_1 = T_c/3$, as when the region next to the base is high-mobility InGaAs and the rest is InP with somewhat lower mobility. For the second scenario – to illustrate a HBT near Kirk threshold or with a badly designed base-collector grade – let $v_1 = 1v_0$, $v_2 = 3v_0$ and $T_1 = T_c/3$. The geometric mean for the first case is $3.66v_0$ and $2.33v_0$ for the second. Using these numbers in equation 2.33 we get $v_{eff} = 3.86v_0$ and $v_{eff} = 1.42v_0$. This means that the collector transit time in the high-medium case is 105 % higher than the average velocity would give, and in the second case 61 % of the average velocity. With $v_0 = 1 \cdot 10^5$ m/s this corresponds to 260 fs and 704 fs for a 200 nm thick collector for the first and second case. Obviously a low mobility region next to the base has a large influence on collector transit time. This shows the need for properly designed base–collector grades.

2.5.6 Maximum current density

Under high current densities the sheer amount of injected charge starts to put a limit on the operation of the HBT. As the current density in the collector increases, this injected charge screens the original doping, and changes the shape of the band

structure. At a sufficiently high current density the collector will act as if it is p-doped, and the base will move out into the collector, making the base transit time larger. For DHBT's the effects can be even worse as the conduction band offset will move above the base conduction band, and create an electron trap, with very detrimental results for device performance, see figure 2.21. The way the current density effects the band structure can be estimated through Poisson's equation (equation 2.34) for the collector, with J_e/qv_c as the injected charge.

$$\frac{d\mathcal{E}}{dx} = \frac{1}{\varepsilon_r}\left(qN_c - \frac{J(x)}{v_c}\right) \tag{2.34}$$

Where N_c is the collector doping level, typically $\sim 2 \cdot 10^{16}$ cm^{-3}, $J(x)$ is the current density: J_e in the region directly beneath the base and changing to a position dependant value $J(x)$ further down in the collector due to current spreading. Integrating (2.34) twice yields the potential drop over the base-collector junction $\Phi_{bi} + V_{\text{applied}}$. Φ_{bi} is the junction built–in voltage (section 2.2.2) and V_{applied} is the applied voltage, $\approx V_{bc} = V_{ce} - V_{be}$. Resistive losses in the base and collector will reduce the applied voltage, $V_{\text{applied}} = V_{bc} - R_{bb}I_c/\beta - R_{subc}I_c$. R_{bb} and R_{subc} are the base resistance and subcollector access resistances. However, for time being the resistive losses will be ignored as their magnitude is rather small. Assuming a current of 20 mA, a DC gain of 24, 12 Ω base resistance and 4 Ω collector resistance, the potential drop from these resistive losses become 0.08 V, small compared to typical bias potentials of 1-2 V. v_c is the effective electron velocity in the collector, in the order of $3 \cdot 10^5$ m/s, [1]. Integrating 2.34 gives:

$$\int d\mathcal{E} = \frac{1}{\varepsilon_r}\int_0^x \left(qN_c - \frac{J(x)}{v_c}\right) dx \tag{2.35}$$

The second integration gives the total potential drop over the junction $\Phi_{bi}+V_{\text{applied}}$. The Kirk current threshold is reached when the electric field at the base–collector junction is zero, $dE(0)/dx = 0$, so the start of the conduction band in the collector is flat, and an electron barrier is formed. That boundary value is used in the next integration.

$$\Phi_{bi} + V_{\text{applied}} = \int \mathcal{E}(x) = -\frac{1}{\varepsilon_r}\int_0^{T_c}\int_0^x qN_c - \frac{J(x)}{v_c} dx dx \tag{2.36}$$

when $J(x)$ and N_c are constant it turns into the familiar formula:

$$\Phi_{bi} + V_{\text{applied}} = -\frac{qN_cT_c^2}{2\varepsilon_r} + \frac{J_cT_c^2}{2\varepsilon_r v_c} \tag{2.37}$$

Equation 2.37 can be rearranged into:

$$J_c = qN_cv_c + \frac{(\Phi_{bi} + V_{\text{applied}})2\varepsilon_r v_c}{T_c^2} \tag{2.38}$$

if we insert typical values into (2.38), $N_c = 2 \cdot 10^{16}$ cm^{-3}, $v_c = 3 \cdot 10^5$ cm/s, $\Phi_{bi} = 1.01$V, $V_{\text{applied}} = 0.5$ V and $T_c = 200$ nm we get $J_{Kirk} = 3.5$ mA/μm^2

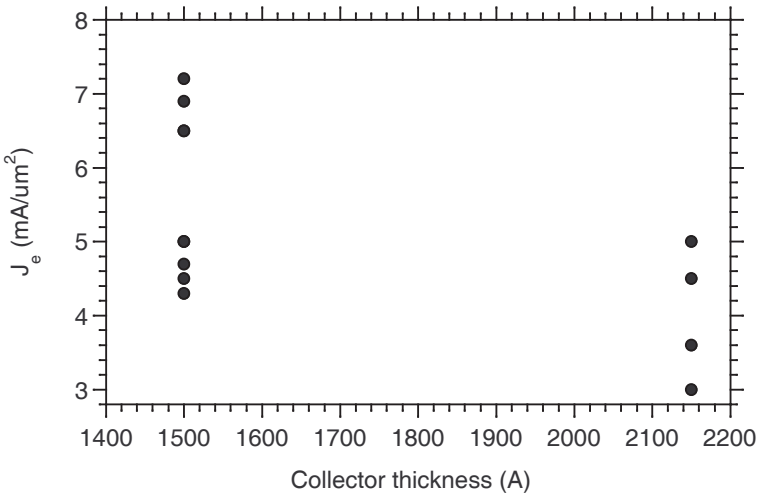

Figure 2.20. Extracted Kirk current density from capacitance data. The different J_{Kirk} for 150 nm correspond to narrow emitters (top) and wide (bottom)

for $T_c = 200$ nm and $J_{Kirk} = 5.5$ mA/μm^2 for $T_c = 150$ nm. In this example the collector doping N_c was the same for both collector thicknesses, while in fact, according to equation 2.40 a thinner collector can sustain a higher doping level, and thus an even higher current.

From equation 2.38 one might be led to just increase the collector doping to increase the Kirk current threshold. One does not have full liberty to do that since at a chosen bias the collector must be depleted. If not the collector capacitance will increase and the HBT's usability decrease. For most digital circuits the chosen bias is $V_{cb} = V_{be} \rightarrow V_{bc} = 0$. This criteria can be expressed as:

$$[\Phi_{bi} + V_{\text{applied}}] = \frac{qN_cT_c^2}{2\varepsilon} + \frac{qN_\delta T_\delta T_{\text{setback}}}{\varepsilon} \tag{2.39}$$

Rearranged to yield the permissable collector doping as a function of collector thickness for our standard grade and setback:

$$N_c \leq \frac{2\varepsilon_r\,[\Phi_{bi} + V_{bc}]}{qT_c^2} - \frac{2N_\delta T_\delta T_{\text{setback}}}{T_c^2} \tag{2.40}$$

The maximum permissable collector doping has a $\sim 1/T_c^2$ dependence just like J_c.

The important conclusion from (2.38) is the Kirk current's $1/T_c^2$ dependence. This is a primary driving cause to decreasing the collector thickness.

Experimental evidence for this can be seen in figure 2.20, where the plotted current density is where the base–collector capacitance starts to increase. The different data for the 215 nm thick collector correspond to different bias V_{applied}

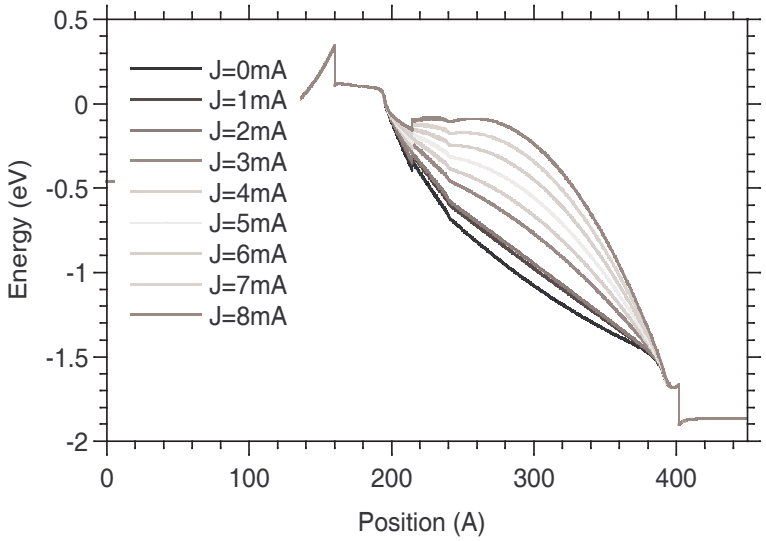

Figure 2.21. DHBT-17 base collector conduction band profile as a function of current density

that will change the critical current density. A stronger applied bias (V_{applied}) leads to an increase in J_c as can be seen in (2.38). The different Kirk current densities correspond to different device size, as well as different bias conditions: narrower emitter and stronger bias gives a larger Kirk threshold.

Figure 2.21 shows a numerical simulation of a DHBT base–collector junction with different current densities. The collector doping level N_c is reduced by J/qv_c, eventually making the collector space charge density negative when the current is high enough. The effective collector space charge density can be expressed as $\rho = qN_c - J/qv_c$. The collector saturation velocity v_c is assumed to be $5 \cdot 10^5$ m/s for InGaAs and $3 \cdot 10^5$ m/s for InP and the grade.

However, the electron current in a collector will have a lateral spread L_d, that can be approximated as $L_d = \sqrt{D_e \tau_c}$ where τ_c is the collector transit time, $\tau_c \approx T_c/2v_c$ and D_e is the electron diffusivity in the collector [63]. The value of the diffusivity is hard to estimate. Zampardi *et. al.* gives it as 200 cm^2/Vs for GaAs, and I apply an *ad–hoc* value of 300 cm^2/Vs for InP. For a 300 nm InP collector L_d becomes 0.12 μm, and smaller for thinner collectors. However, this is only a coarse estimate. The problem is dynamic: as the region beneath the emitter approaches the Kirk threshold the band structure in that region will change due to the massive amount of negative charge. The negative charge will act as a p-type doping $N_J \approx J_e/qv_c$ and move the conduction band up. This will in its turn effect the effective electron velocity in that region and also induce lateral fields which drive the electrons laterally away from the emitter stripe — and that will effect the

lateral spread. If we assume that the effective velocity goes down as we approach the Kirk threshold (a reasonable assumption in a HBT where a conduction band barrier can form, see figure 2.21), then the diffusion length L_d goes up because the collector transit time goes up, and possibly the diffusivity too - which would lower the effective Kirk current density! Thus, the expressions derived in here, can only serve as a rough estimate when Kirk threshold effects are *near*. It is possible that the very flat f_t vs J_c characteristics observed is due to this "soft" response.

For our HBTs the emitter is usually rather narrow compared to its length, with lengths of 6-8 μm, whereas widths are 0.5-1 μm for most of our devices, lending itself to approximate the emitter stripe as one-dimensional, with current spreading perpendicular to the length. If we let W_e be the width of the emitter, and L_e the length of it, we can similar to Zampardi *et. al.* let L_d denote the spread, causing the effective area for the collector current to spread as

$$(W_e + \frac{2L_d x}{T_c})L_e$$

where the factor of 2 arises from the current spreading on both sides of the emitter. Current spreading from the ends is neglected, and this formula is thus not expected to hold true when the length and width of the emitter become comparable.

If we solve Poisson's equation (2.36) with

$$J(x) = \frac{J_e A_e}{(W_e + \frac{2L_d x}{T_c})L_e}$$

we get the following double integral to solve:

$$\frac{1}{\varepsilon_r v_{sat}} \int_0^{T_c} \int_0^x J(x) dx^2 = \frac{1}{\varepsilon_r v_{sat}} \int_0^{T_c} \int_0^x \frac{J_e A_e}{W_e + \frac{2L_d x}{T_c})L_e} dx^2 \tag{2.41}$$

Using $A_e = W_e L_e$, we get:

$$J W_e \left(\frac{T_c}{2L_d}\right)^2 \left(W_e \ln\left(1 + \frac{2L_d}{W_e}\right) - 2\,L_d + 2 \ln\left(1 + \frac{2L_d}{W_e}\right) L_d\right) \tag{2.42}$$

As a control of the calculation let $L_d \to 0$ and Taylor series expansion gives us when taking to the second order,

$$\ln(1 + \frac{2L_d}{W_e}) \to \frac{2L_d}{W_e} - \frac{1}{2}\left(\frac{2L_d}{W_e}\right)^2$$

which inserted into (2.42) gives:

$$J_e \frac{T_c^2}{2} \tag{2.43}$$

Which is the answer as 2.38 in the constant case.

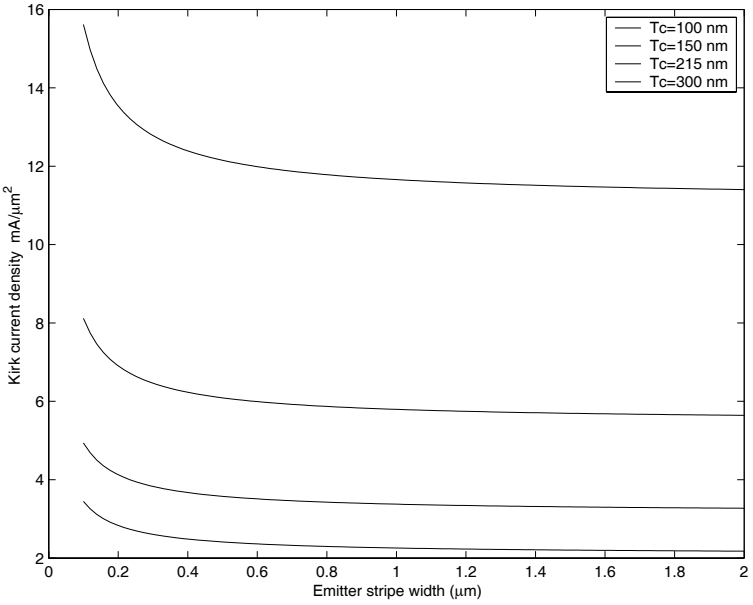

Figure 2.22. Calculated Kirk current density as a function of emitter stripe width and collector thickness

Combining (2.42) and (2.36):

$$\Phi_{bi} + V_{\text{applied}} + \frac{qN_cT_c^2}{2\varepsilon_r} = \frac{J_eL_eW_e}{\varepsilon_r v_c L_e}\frac{T_c^2}{2L_d^2}\left(W_e\ln(1+\frac{2L_d}{W_e}) + 2L_d\ln(1+\frac{2L_d}{W_e}) - 2L_d\right) \tag{2.44}$$

Using

$$\Gamma = \frac{W_e}{2L_d^2}\left(W_e\ln(1+\frac{2L_d}{W_e}) + 2L_d\ln(1+\frac{2L_d}{W_e}) - 2L_d\right) \tag{2.45}$$

we get

$$J_e = v_c\left[qN_c + \frac{2\varepsilon_r}{T_c^2}\left(\Phi_{bi} + V_{\text{applied}}\right)\right]\frac{1}{\Gamma} \tag{2.46}$$

A series expansion with $L_d \to 0$ of Γ goes to 1 as it should.

If we insert typical values into (2.44), $N_c = 2\cdot 10^{16}\text{cm}^{-3}$, $v_c = 3\cdot 10^5$ cm/s, $\Phi_{bi} = 1.01$ V, $V_{\text{applied}} = 0.5$ V and $T_c = 100 - 300$ nm we get the results shown in figure 2.22. The calculation use the simplifying assumption that the collector doping level is constant.

Experimental evidence comes from measurements of HBT's with different emitter width. The current at which the base–collector capacitance starts to increase (closely related to the Kirk current threshold) is shown in figure 2.23. The data shows how the current density where C_{cb} starts to increase varies with the emitter

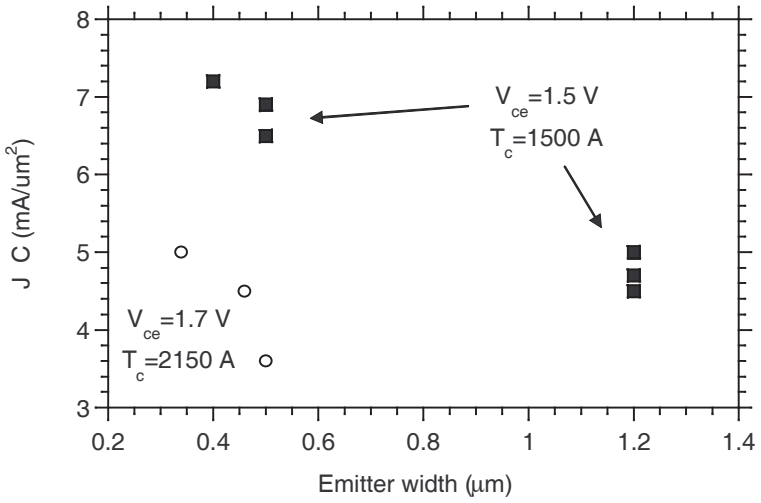

Figure 2.23. Measured current density J where C_{cb} starts to increase as a function of emitter stripe width

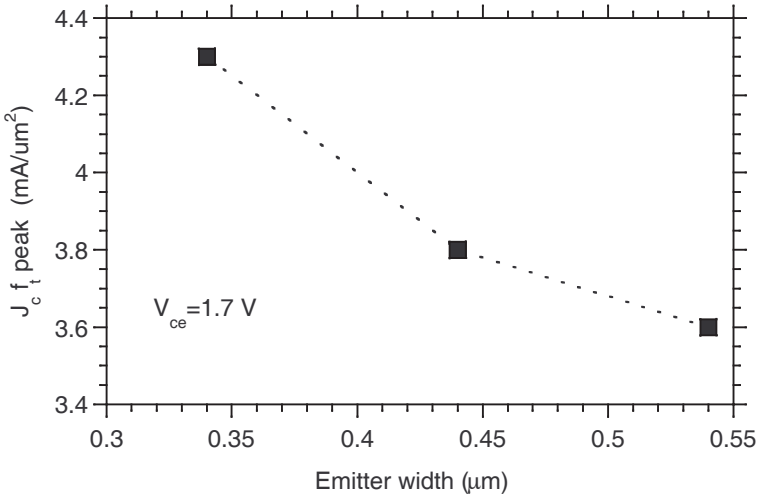

Figure 2.24. Current density for maximum f_τ as a function of emitter junction width. $T_c = 217$ nm

width in a fashion consistent with equation 2.46. Another evidence is shown in figure 2.24 where the current density for maximum f_τ is shown as a function of the emitter width. Once again the result is consistent with current spreading in the collector region in that narrow emitter devices show higher f_τ.

One issue is the collector saturation velocity v_c which is used in the calculations to predict the collector transit time and the charge storage. It is currently set to $3 \cdot 10^5$ cm/s but that value is slightly arbitrary, as is the electron diffusion length $L_d = \sqrt{D_e \tau_c}$ where τ_c is the collector transit time defined as $T_c/2v_c$, as previously discussed on page 43. The exact value of L_d determines the amount of lateral diffusion. We use a value of L_d of $0.12\,\mu m$ for a 300 nm collector. GaAs or Si are estimated to have higher value of L_d than InP due to their lower saturation velocity, the electrons simply have more time to diffuse sidewards before they are collected at the subcollector. As the collector current density approaches the Kirk current density and the conduction band starts to block electron transport through the base–collector region the transit time will increase – and L_d will increase through $L_d = \sqrt{D_e \tau_c}$ – which will lower the current density. Under conditions that increase the Γ-L scattering – such as an increased V_{bc} – the collector transit time will increase, and L_d will once again increase. This should not be confused with the capacitance cancelation effect since lateral diffusion of electrons will not explain the observed decrease in C_{cb} [72, 96].

One conclusion one can draw is that it would be beneficial to let the base contact extend slightly more than L_d, $\approx 0.12\,\mu m$ on each side of the emitter base junction. This value is close to contact transfer length L_c, the natural minimum base contact width. Further, equation (2.46) and the presented data are strong arguments for continued emitter width scaling, as the current spreading allows a larger current before reaching the Kirk threshold than the collector doping would suggest.

Conduction band profile with grade and δ-doping

The inclusion of the conduction band offset and the δ-doping modifies Poisson's equation (2.34). The Kirk current threshold (2.38) is modified by a term

$$\frac{qN_\delta T_\delta T_{\text{setback}}}{\varepsilon_r}$$

where T_{setback} is the thickness of the InGaAs setback layer between the base and the grade. The Kirk current threshold becomes:

$$J_c = \left[qN_c v_c + \frac{(\Phi_{bi} + V_{\text{applied}}) 2\varepsilon_r v_c}{T_c^2} - 2\frac{N_\delta T_\delta (T_{\text{setback}} + T_{\text{grade}})}{T_c^2} \right] (1/\Gamma) \qquad (2.47)$$

The correction from (2.46) is incorporated as Γ.

2.5.7 The setback layer

The conduction-band potential drop across the setback layer is

$$\Delta\phi_{\text{setback}} \cong (V_{\text{applied}} + \phi_{bi})\frac{T_{\text{setback}}}{T_c} + \frac{qN_\delta T_\delta T_{\text{setback}}}{\varepsilon_r} + \frac{(qN_c - J(x)/v_c)\,T_c T_{\text{setback}}}{2\varepsilon_r} \tag{2.48}$$

In the following numerical examples $T_c = 215$ nm, $T_{\text{setback}} = 20$ nm, $T_\delta = 3$ nm and $N_\delta = 3 \cdot 10^{18}\text{cm}^{-3}$ [69]. The second term, of magnitude 0.25 V, is present due to the change in the conduction band quasi-field at the interface between the grade and the setback layers. The setback-layer potential drop provides electrons incident on the grade with kinetic energy , reducing the likelihood of current blocking. Only for $V_{cb} > 5.0$ volts does $q\Delta\phi_{\text{setback}}$ exceed $E_{g,InGaAs}m_h^*/(m_h^* - m_e^*)$, the energy required for impact ionization in the InGaAs setback layer.

Since InP-collector DHBT's generally exhibit breakdown $V_{br,CEO} \cong (25\,\text{to}\,\text{V}/\mu\text{m}\cdot T_c$ [68], impact ionization in the setback layer should have minimal effect on $V_{br,ceo}$. Similarly, breakdown by impact ionization will occur in the InP collector at a smaller V_{bc} than is required for breakdown in the grade. One important observation from (2.48) is the $1/T_c$ dependence in the first term. A thicker collector needs a thicker setback layer to maintain the same potential drop over the setback layer (figure 2.25), and to provide the incoming electrons with the same velocity. The δ-doping can also be tuned to provide the desired potential drop over the setback layer, but if the δ-doping is increased too much incomplete depletion in the collector will occur. The maximum collector doping is set by equation 2.40. This dependence could be an explanation with two DHBT with the same base, setback layer and grade (the new 24 nm thick one) but different collector thickness had very different f_τ. The device with thinner collector (200 nm) had an f_τ of 207 GHz, while the device with thicker collector (300 nm) had an f_τ of only 116 GHz [71,98].

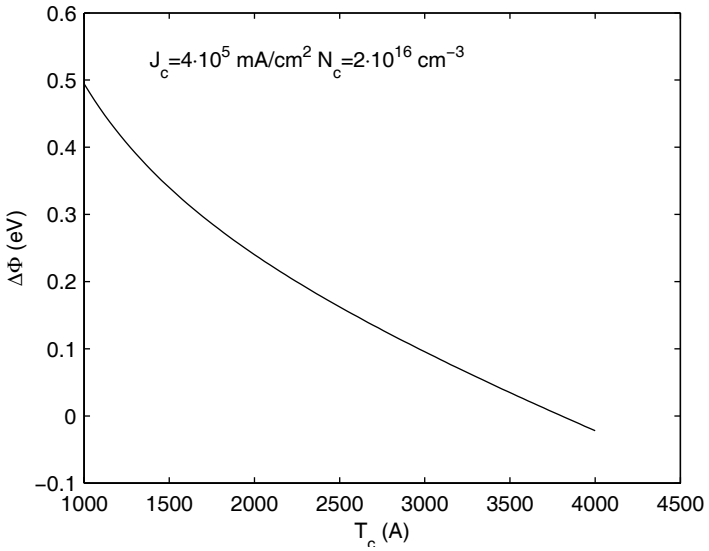

Figure 2.25. Potential drop over a 20 nm setback layer as a function of collector thickness, with $V_{\text{applied}} = 1.5$ V.

Chapter 3

Design of InP transistors

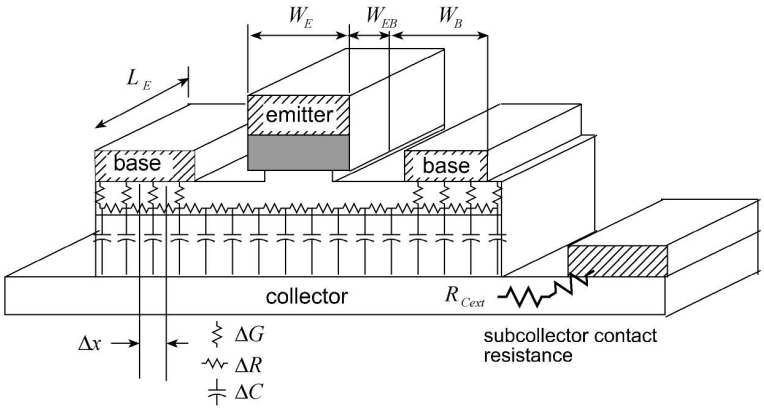

Figure 3.1. The distributed network model

3.1 Simulation of distributed network model of HBT

Simulations were performed on an early stage to judge the feasibility of achieving high f_τ and f_{max} with a Mesa HBT compared to a Transfered Substrate HBT. The model used a distributed network for the intrinsic and extrinsic base–collector regions (figure 3.1) and used the same input parameters as transfered substrate design models. The regions beneath the base contacts are treated as RC–networks, and the region directly beneath the emitter is composed of 128 Spice models of HBT's in parallel. The model was used to extract S–parameters and f_τ and f_{max} (figures 3.2,3.3 and 3.4). The key findings where that the base contact resistance ρ_c and current density J_c need to be as low and as high as possible , respectively (figure 3.2). The simulations encouraged us to develop a mesa HBT with very good base ohmics and a narrow base contact. Undercutting the base–collector mesa is a method to decrease the capacitance [74], especially it will reduce C_{cbx}, the external base–collector capacitance, which is of importance for digital circuit design. We simulated the effect of this using the mesh–model and found that *if* the base contact resistance is good the influence of undercutting is not very large (figure 3.4). Undercutting poses a danger of creating process failures due the base contacts breaking off, and on the basis on the simulations we decided to keep the undercut small, on the order of $0.1 - 0.2$ μm. The design focused on achieving a narrow mesa structure enabled by good lithography (section 4.2) and improved

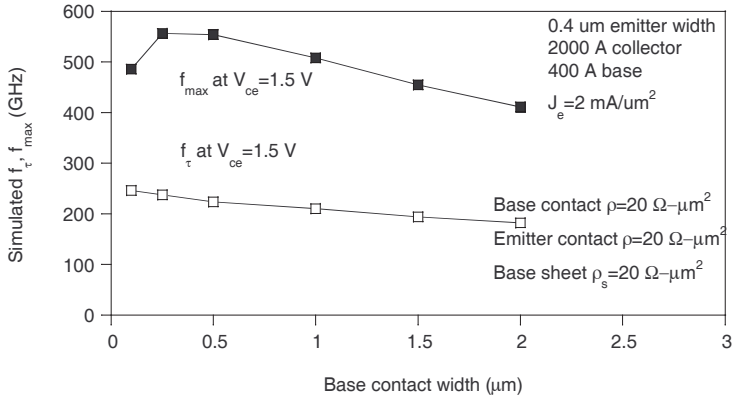

Figure 3.2. Effect of base contact width simulated with distributed network model

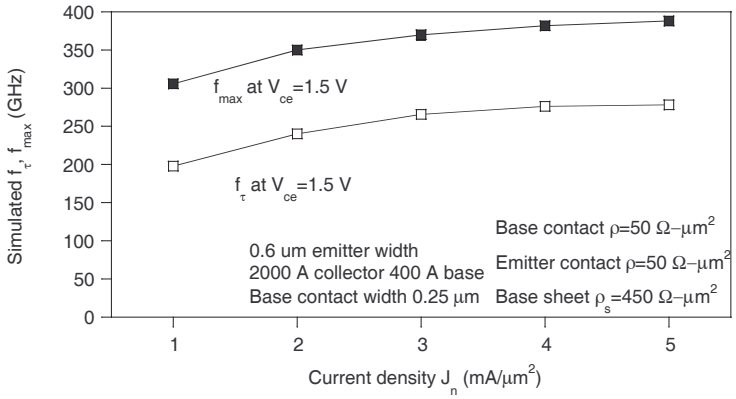

Figure 3.3. Effect of current density simulated with distributed network model

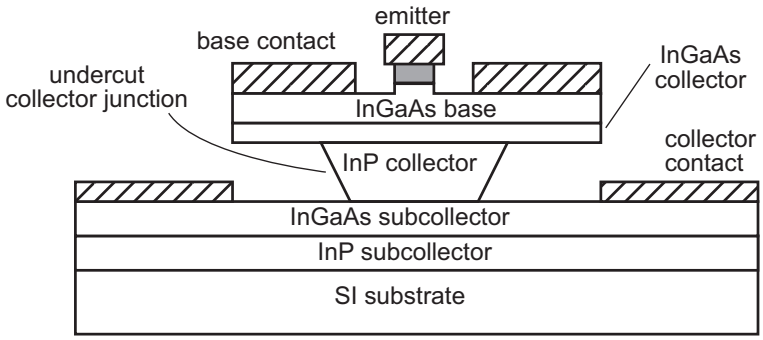

Figure 3.4. Schematic of undercut HBT

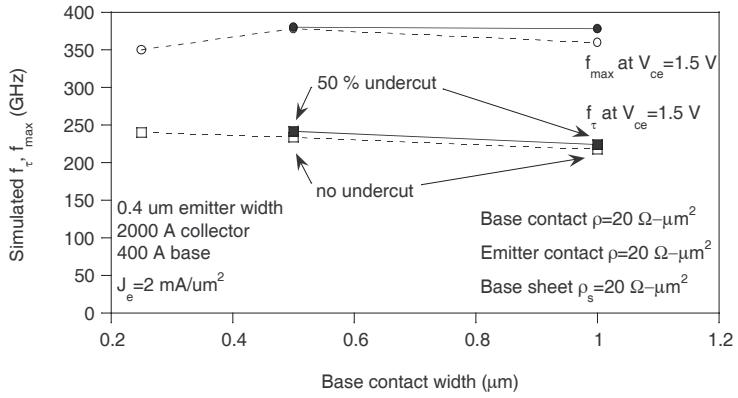

Figure 3.5. Simulation results for undercut HBT . The undercut results correspond to the situation where 50 % of the base contact is undercut

base ohmics (section 2.4).

3.2 Emitter design

The emitter resistance can be expressed as:

$$\begin{aligned} R_{ex} &= \frac{\rho_{c,e}}{L_e W_{e,contact}} + \frac{\rho_{cap} T_{cap}}{L_e W_{e,contact}} \\ &+ \frac{\rho_{e2} T_{e2}}{L_e W_{e,junct}} + \frac{\rho_{e1} T_{e1}}{L_e W_e} \end{aligned} \tag{3.1}$$

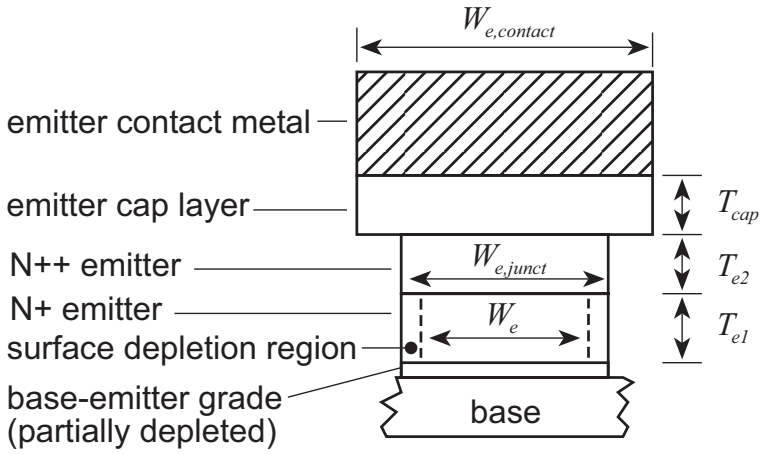

Figure 3.6. Schematic of the emitter in a HBT.

When the work at UCSB started the emitter resistances were on the order of $50\Omega - \mu\text{m}^2$ [107], compared to $27\Omega - \mu\text{m}^2$ [105] and $15\Omega - \mu\text{m}^2$, [27] for other HBT's reported in the literature. To address this problem the emitter cap doping was increased from $1 \cdot 10^{19}$ cm^{-3} to $3 \cdot 10^{19}$ cm^{-3}, the InAlAs/InGaAs emitter–cap to emitter grade was removed at the same time as the N^+ InP doping was increased to $3 \cdot 10^{19}$ cm^{-3} (see tables A.2 and A.3, A.4). The grade was judged unnecessary on the grounds of the doping level, both materials are degenerate and as can be seen in figure 2.8 the electron concentration in the interface hardly changes at high doping levels such as $1 \cdot 10^{19}$ cm^{-3}. Were the junction doping level lower, $1 \cdot 10^{17}$ cm^{-3}, a grade would have been necessary (figure 2.8). Another suspected part of the high emitter resistance at UCSB was surface depletion along the edges of the emitter. Surface (edge) depletion regions of width $(2\epsilon\phi/qN_{e1})^{1/2}$, where N_{e1} is the N layer doping and ϕ is the surface Fermi energy relative to that of the bulk semiconductor will be formed. This occurs in all emitter layers but is only significant in the low–doped parts. The thickness of this layer was chosen to 30 nm, which is still enough to ensure depletion in the emitter–base region. The old $n-$ emitter region was composed of 70 nm InAlAs doped at $8 \cdot 10^{17}$ cm^{-3}, and the drawback here is that every 10 nm of undepleted InAlAs offers a resistivity near 1.8 $\Omega/\mu m^2$, or for a typical junction area of 5 μm^2, $\approx 1.5\ \Omega$.

It was therefore decided to keep the n- emitter thickness as thin as 30 nm, which still is larger than the estimated and measured depletion region width.

The emitter can be made by InP or $In_{48}Al_{52}As$. InAlAs offers certain advantages: reduced surface leakage and good passivation with SiN [6]. Also, because the InAlAs bandgap is higher than InP (1.48 eV compared to 1.35 eV), and the emitter–base junction needs to be graded, a larger valence band discontinuity is possible. The determining factor however, was that the carbon doped base neces-

sitated wet etching of the HBT, and the InAlAs/InGaAs wet etch always created a large undercut, unsuitable for high performance devices, and so we decided to use InP. InP offers better resistivity and heat transfer, but passivation is more difficult [88]. A transfered substrate HBT have a major problem in the fact the heat transfer takes place through the emitter junction – a very narrow region. We employ polyimid passivation which does not contribute any significant amount to the thermal transfer. Other passivation schemes that employ harder dielectrics such as SiC [38] offers enhanced thermal heat–sinking. Mesa DHBT on the other hand have a much larger area for heat–sinking in the base–collector junction, and a DHBT with InP collector and a minimum amount of InGaAs in the subcollector have much enhanced thermal conductivity. A 0.7 by 8 μm emitter area HBT have the following thermal resistance:

- Transfered substrate InP emitter: 9 kK/W
- Transfered substrate InAlAs emitter: 13 kK/W
- Mesa InP collector: 2 kK/W

The thermal budget is clearly a problem for transfered substrate HBT's. But even in a mesa HBT about 30 % of the heat might be removed through the emitter region, according to simulations on the geometry for our HBT's, [84, 93].

3.3 Base design

The first base designs were using the UCSB base line device: 40 nm bandgap graded InGaAs base. The indium composition went from lattice matched 53 % at the base–collector interface and 44 % at the base–emitter interface. This creates a quasi field in the base as discussed in section 2.19. The base was doped with beryllium at first, and then with carbon. In order to reduce the base sheet and contact resistance carbon doping was actively pursued but the results were not as good as expected as the gain was lower than expected.

SIMS analysis of base layers grown at UCSB revealed a slope in the carbon concentration in the base – the carbon concentration decreased towards the emitter, despite the Ga concentration increasing, which should increase the carbon incorporation [20]. The carbon concentration went from $\approx 8{\cdot}10^{19}$ cm^{-3} to $\approx 1.8{\cdot}10^{20}$ cm^{-3}. The reason probably had to do with a change in growth temperature over the base. This however, led to the first calculations of base transit time effects caused by doping variations and showed that an electric field was created by the doping gradient. The size of the field was not negligible compared to the bandgap induced quasi field of $\approx$50 meV. In table 3.1 the calculated base transit time is calculated for the following three structures. A structure with constant doping and bandgap, a structure with constant doping and a bandgap grade – the intended case – and in the last column a structure with graded bandgap but with a carbon retrograde, as inferred from SIMS data. From $\beta = \tau_n/\tau_b$ we get a gain difference for the

Table 3.1. Influence of carbon retro–grade: transit times

Base thickness (nm)	Constant base τ_b (ps)	Bandgap grade τ_b (ps)	Retro grade τ_b(ps)
30	0.22	0.12	0.22
40	0.35	0.19	0.34

Table 3.2. Influence of carbon retro–grade: relative gain

Base thickness (nm)	Constant base Gain	Bandgap grade Gain	Retro grade Gain
30	1	1.83	1.00
40	1	1.84	1.03

retro graded case compared to the bandgap graded case which was our base line structure, of 54 % and 56 % for 30 and 40 nm thick bases (table 3.2).

While the problem with carbon incorporation at UCSB was solved, carbon doped wafers also became available as research samples from IQE Inc [76] and were used for the later structures [1]. Initial carbon doped HBT's showed lower gain than corresponding Be doped HBT's [76]. To improve gain it was decided to switch from a graded base–emitter junction to an abrupt base–emitter junction, and to use either a bandgap graded base (table A.4) or a doping graded base (table A.3). An abrupt base–emitter junction will increase gain at the cost of increased drive voltage, something that is not good for integrated circuits, but the current gain would otherwise have become forbiddingly low.

If the current gain is limited by Auger recombination in the base, and not by co–called ballistic electrons traversing the base, the gain is a function of base doping, thickness, Auger recombination coefficient and electron minority mobility. Ballistic electrons can be included in the expression (equation 3.2) by increasing the electron minority mobility.

$$\beta = \frac{2k_bTA\mu_{e,base}}{N_a^2T_b^2} \tag{3.2}$$

Where $A = 4{\cdot}10^{-29}$ cm^6/s is the Auger recombination coefficient [47,77]. Figure 3.8 show the Auger recombination limited gain in a HBT as a function of doping and base width. In the derivation the electron mobility in the base has been tuned to our measured gain for 30 and 25 nm thick bases. The electron minority mobility seems to be smaller in carbon–doped InGaAs than in beryllium doped InGaAs, and even smaller still in carbon doped GaAsSb, [66]

3.3.1 Base transit time calculations

The calculation tool described in section 2.19 was used to evaluate possible base designs. The first task was to investigate the current bandgap graded device, and see

[1] I am very grateful the excellent epi IQE provided

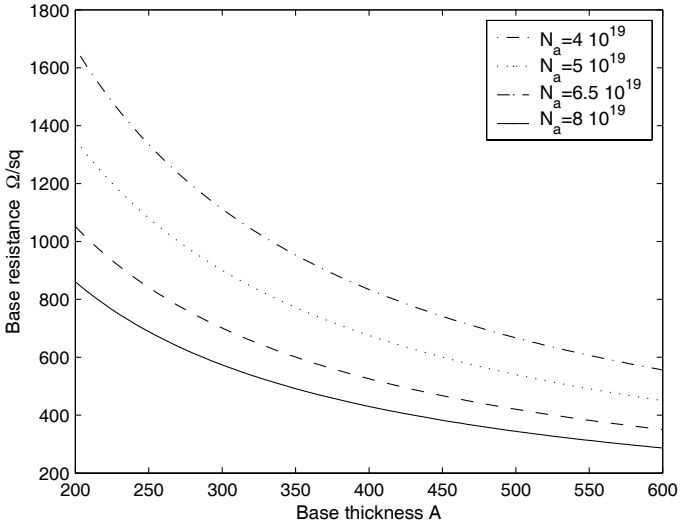

Figure 3.7. Calculated base resistance for different average doping levels

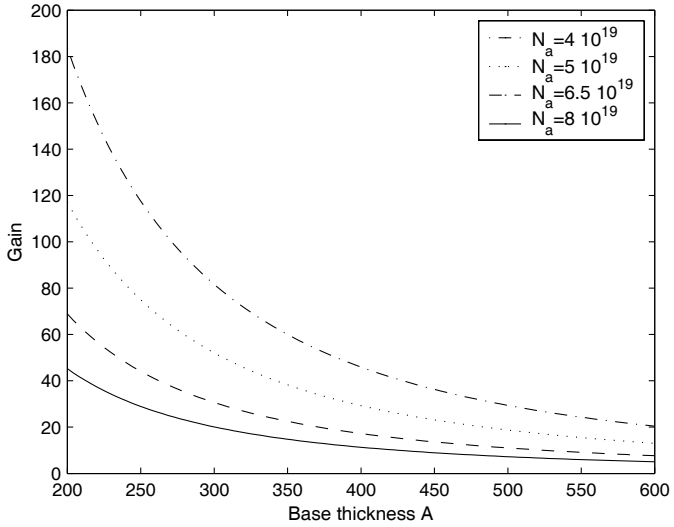

Figure 3.8. Calculated Auger recombination limited current gain

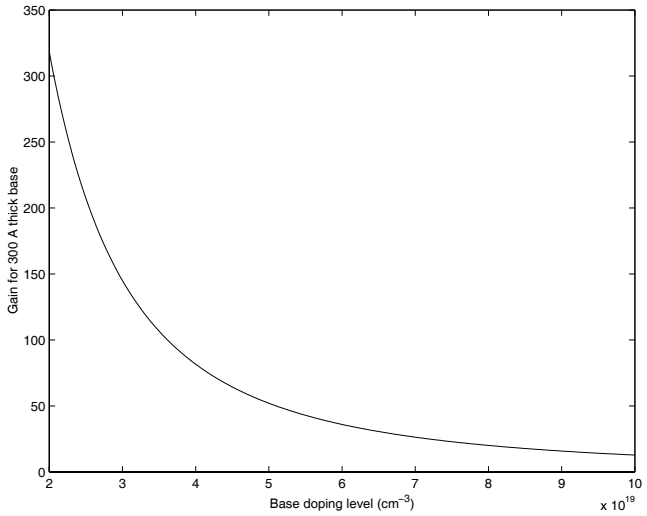

Figure 3.9. Calculated Auger recombination limited current gain as a function of doping

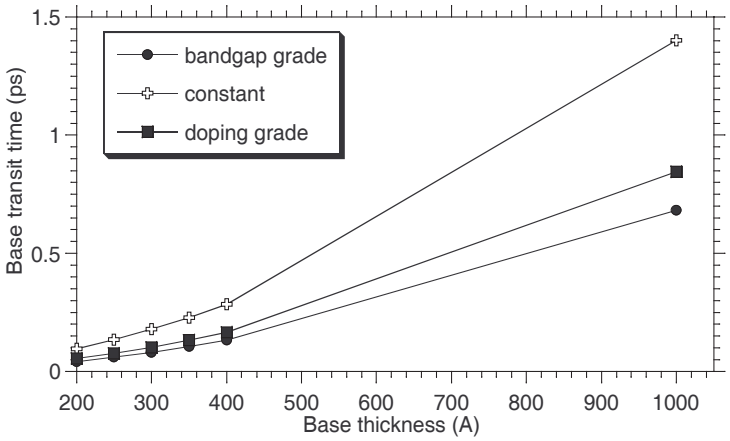

Figure 3.10. Calculated base transit time for different base configurations. The base grades are from DHBT-17 and DHBT-18

Table 3.3. Different base transit time and base sheet resistance for different base configurations. T_b=30nm

Base type	Calc. τ_b (ps)	Calc. ρ_s ($\Omega/\square$)	Meas.ρ_s ($\Omega/\square$)
$N_d = 8 \cdot 10^{19}$, In: 44 $\rightarrow$ 53	0.12	580	604-624
$N_d = 5 \cdot 10^{19}$, In: 44 $\rightarrow$ 53	0.09	905	
$N_d = 8 - 5 \cdot 10^{19}$, In: 53	0.16	703	515-616
$N_d = 8 - 4 \cdot 10^{19}$, In: 53	0.14	760	
$N_d = 8 - 3 \cdot 10^{19}$, In: 53	0.12	826	
$N_d = 8 - 1 \cdot 10^{19}$, In: 53	0.09	1001	
$N_d = 10 - 7 \cdot 10^{19}$, In: 53	0.17	542	
$N_d = 5 - 2 \cdot 10^{19}$, In: 53	0.12	1269	
$N_d = 8 - 5 \cdot 10^{19}$, In: 44 $\rightarrow$ 53	0.07	705	

what transit time that resulted in, and how strong the base electric field was. Given that, doping graded designs were examined to see what values of doping should be selected. Figure 2.18 show the effective base electric field from a bandgap graded and doping graded base. A doping gradient was chosen to provide a quasi field matching the field induced by the bandgap graded devices.

Table 3.3 shows the calculated base transit time for a base with constant doping and bandgap, a base with graded bandgap and constant doping (In content going from 44.5 % at the emitter base junction to 53 % at the base –collector junction) and a base with lattice matched InGaAs and a doping variation linearly graded from $8 \cdot 10^{19}$ cm^{-3} to $5 \cdot 10^{19}$ cm^{-3}. The base transit time is broken down into the internal transit time and the exit time. The calculated base sheet resistance is also shown. Tables 3.4 and 3.5 show the calculated base transit times for two types of InGaAs base, DHBT-17 (doping graded) and DHBT-18 (bandgap graded).

So called ballistic injection, where some electrons after tunneling through the emitter-base conduction band barrier have a velocity given by $sqrt2\Delta E_c/m^*$ (see section 2.3.3) is not included in the calculation. For thick bases the electrons with excessive velocity will quickly slow down and the overall influence of the abrupt emitter will be small.

Figures 3.11,3.12 and 3.13 show the calculated base transit times for the following two different types of base:

- Bandgap graded $N_a = 8 \cdot 10^{19}$ cm^{-3}. $In_{44}Ga_{56}As \rightarrow In_{53}Ga_{47}As$
- Doping graded $N_a = 8 \cdot 10^{19}$ to $5 \cdot 10^{19}$ cm^{-3} $\approx In_{53}Ga_{47}As$

These two base configurations are the basis for the fast HBT's presented in this thesis. Included in table 3.3 is the measured base sheet resistance from two HBT wafers, DHBT-17 and DHBT-18. From the ratio of base transit time one would expect the gain of the bandgap graded HBT to be larger by a factor $0.16/0.12 = 1.33$, but large area HBT measurements show the gain is 18 for the bandgap graded

Table 3.4. Base transit time and base sheet resistance for bandgap graded base.

Base thickness (nm)	$\tau_{b,internal}$ (ps)	τ_{exit} (ps)	$\tau_{b,total}$ (ps)	Calculated ρ_s ($\Omega/\square$)
20	0.039	0.046	0.086	1055
25	0.061	0.058	0.119	845
30	0.089	0.069	0.158	603
40	0.158	0.092	0.25	528

Table 3.5. Base transit time and base sheet resistance for doping graded base.

Base thickness (nm)	$\tau_{b,internal}$ (ps)	τ_{exit} (ps)	$\tau_{b,total}$ (ps)	Calculated ρ_s ($\Omega/\square$)
20	0.029	0.033	0.062	1048
25	0.045	0.041	0.086	840
30	0.065	0.050	0.115	600
40	0.115	0.066	0.181	524

HBT and 25 for the doping graded HBT, a ratio of $18/25 = 0.72$, while the base sheet resistance ρ_s is smaller for the dopant graded sample. The doping graded HBT show better gain and lower resistivity than expected. Gain is effected by the electron lifetime, which is growth dependent [78]. The InGaAs layers are grown lattice–matched with compensation for carbon doping, as verified by X–ray.

Upon inspection of equation 2.19 it is apparent that the diffusivity has a strong temperature dependence, $D_{n,b} = k_B T/q\mu_{n,b}$. Since the other factors in equation 2.19 only have a weak temperature dependence, the effect of a temperature gradient through the base was examined. Is there such a temperature gradient? According to [21] the temperature difference for DHBT-17 at UCSB over the base was around 17 °C, with the region beneath the base the hottest. The results of the simulations are shown in figure 3.15 and show an influence of the temperature. The

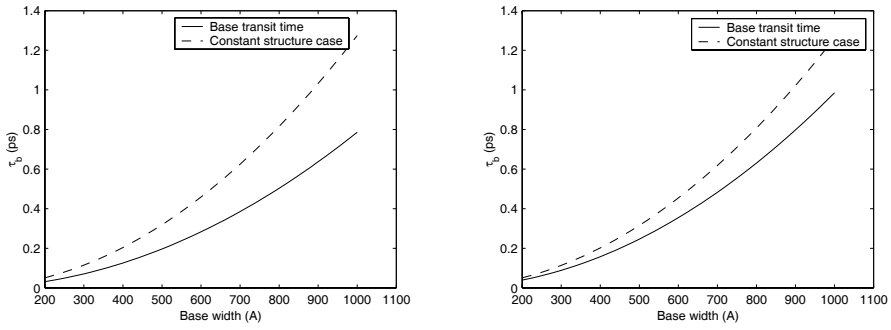

Figure 3.11. Calculated internal base transit time from equation 2.22 for bandgap graded base (left) and doping graded base (right)

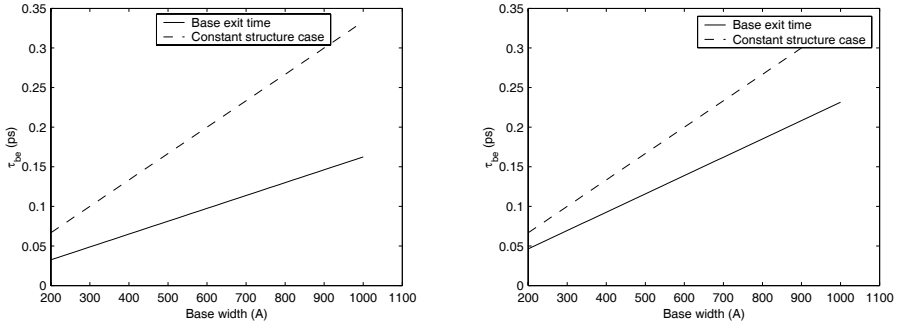

Figure 3.12. Calculated base exit time from equation 2.23 for bandgap graded base (left) and doping graded base (right)

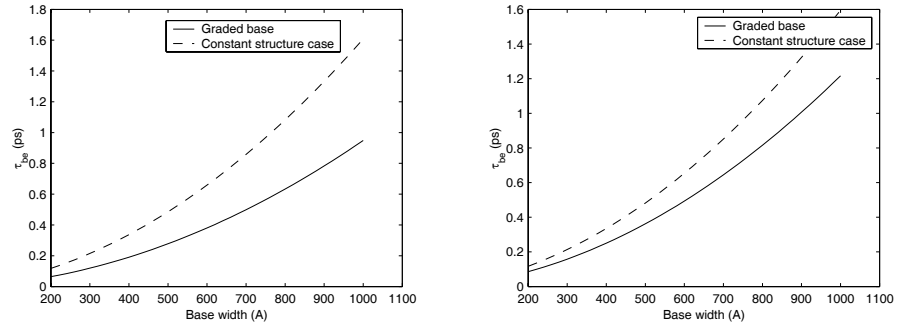

Figure 3.13. Calculated total base transit time from equations 2.22 and 2.23 for bandgap graded base (left) and doping graded base (right)

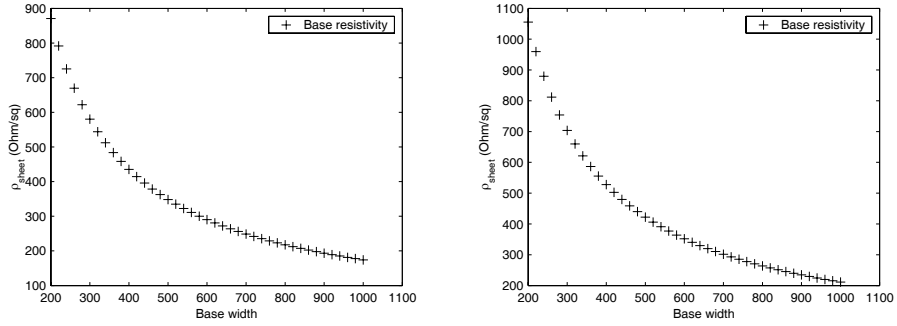

Figure 3.14. Calculated base resistance for bandgap graded base (left) and doping graded base (right)

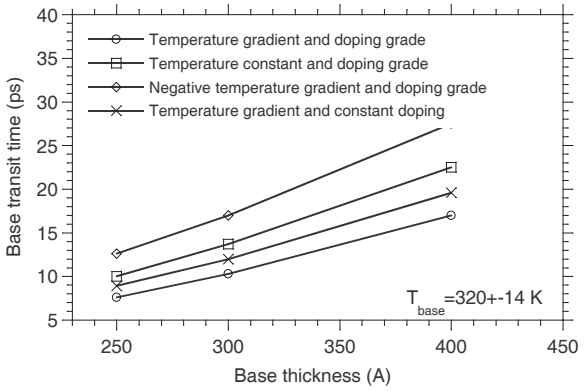

Figure 3.15. Calculated base transit time with the influence of temperature

situation of the temperature gradient would be different in a transfered substrate HBT, where the major heat flow takes place through the emitter. The emitter is a narrow region with large thermal resistance, and the effect of the temperature gradient would increase the base transit time, and thus reduce the gain of the device. While gain is not critically important for devices optimized for RF–performance, circuit design when the gain is low becomes increasingly difficult. We have chosen 20 to be a lower limit for the acceptable gain. As can be seen in figure 3.9 increased doping quickly reduces the gain, and base grading schemes are necessary for improving the gain. The gain increase when changing from a 30 nm base to a 25 nm base was only 27/25, with variation over the wafers.

3.4 Grade and collector designs

The grade in use late 2000 was a 48 nm thick InAlAs/InGaAs super lattice, with a lattice period of 1.5 nm (figure 3.16). These became wafers DHBT 1 and DHBT 2, designed and processed by K. Sundarajan [91]. Our notion was that the grade was unnecessarily thick, with an expected low electron velocity. The grade did not have a setback region, but started directly after the base. The first InAlAs layers were 0.2, 0.4, 0.6 ... Å thick, surrounded by InGaAs such that the period thickness was 15 Å. However, since the nearest neighbor distance in InAlAs/InGaAs lattice matched to InP is 2.4 Å, this meant that in reality no InAlAs was grown over most of the wafer area for the first part of the grade. Only when the InAlAs layers became close to 2.4 Å could we expect coverage. This effectively meant that the first 10-15 nm of the grade region was an InGaAs setback region!

This was not immediately realized and the next grade designs focused on making the grade thinner (figure 3.18), and a 10 nm and a 20 nm design was grown, with the not quite appropriate δ-doping level according to equation 2.28 where instead of

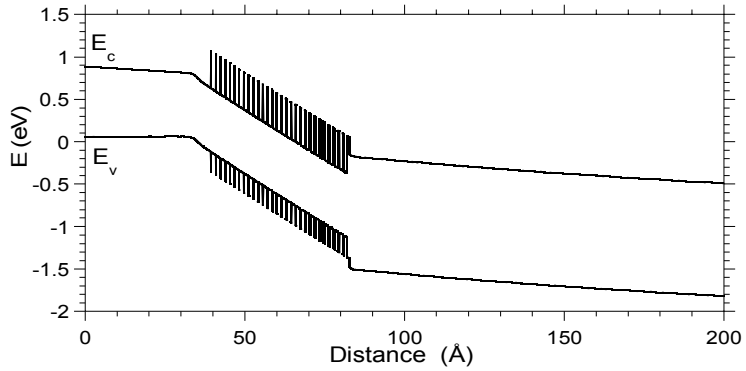

Figure 3.16. The first grade, 48 nm thick with no setback

ΔE_c the total bandgap difference ΔE_g was used. This made the δ-doping roughly twice the proper amount.

$$N_\delta T_\delta = \frac{\varepsilon_r \Delta E_c}{q^2 T_{grade}} \tag{3.3}$$

The thickness T_δ was chosen to 3 nm after discussion with our MBE–growers. These became wafers DHBT-3 and DHBT-4. The new grades were designed according to an algorithm that minimized the total deviation from linear grade between InGaAs and InP, by mixing InGaAs and InAlAs layers of different thickness. An additional constraint was that no layer was allowed to be thinner than 1/2 lattice period [2]. Figure 3.17 shows how the 20 nm thick grade was constructed. The first designs showed disappointing RF and DC performance (page 88), and one suspicion was Be–out diffusion into the grade layers as the debacle with the δ-doping had not yet been discovered.

To compensate for that thin setback layers were incorporated in new HBT layer structures . One design that was tried was a composite–collector design in which $\sim 1/3$ was InGaAs, then a grade and then InP for the rest (DHBT-6 and DHBT-9). The idea was the region next to the base should have high–velocity InGaAs, and the rest InP for breakdown considerations (equation 2.33).

The characteristics was still disappointing and electron trapping in the super lattice was suspected (equation 2.29).

Based on [67] the new super lattice period was fixed to 1.5 nm with a thinnest layer thickness of 0.15 nm. The thickness was increased to 24 nm to reduce the

[2] You can select any growth period, but the growers insist on maintaining the interatomic distance.

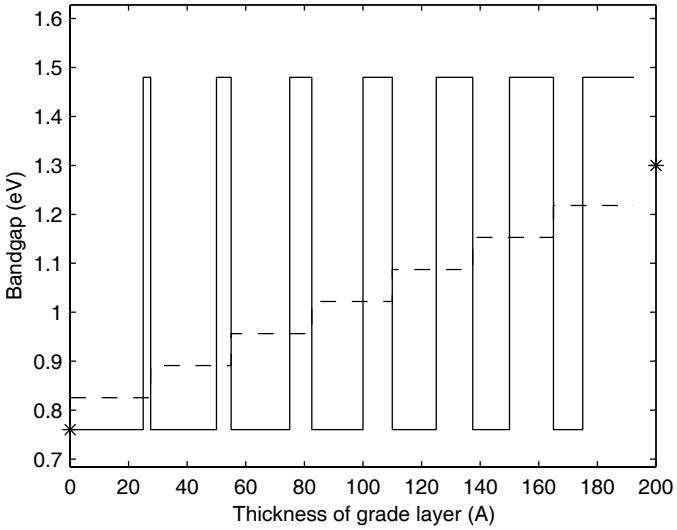

Figure 3.17. The 20 nm thick grade

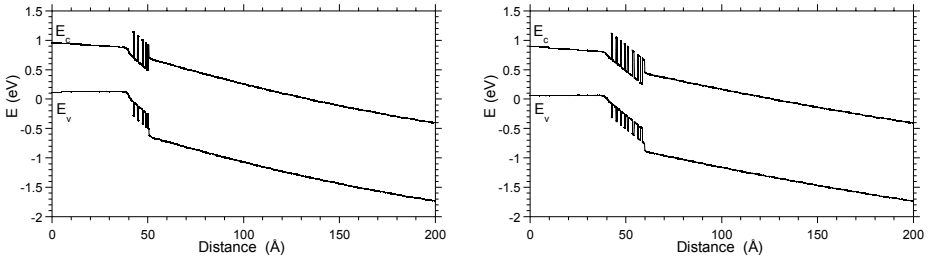

Figure 3.18. New grade designs 10 nm thick used in DHBT-3,5,6 (left) and 20 nm thick used in DHBT-4 (right)

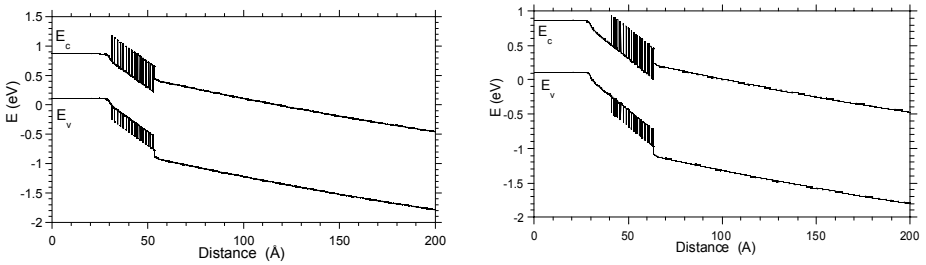

Figure 3.19. New grade designs 24 nm thick with no setback region (left) and 10 nm setback region, used in MHBT-1 (right)

δ-doping necessary, and the proper doping was selected. The first design used a 10 nm thick setback layer. Using the 1–D Poisson simulation tool *BandProf*, [92] the conduction band was designed so the potential drop before the grade was facilitate electron tunneling through a suspected barrier, figure 3.19. The performance was very promising with a 207 GHz f_τ for a metamorphic DHBT made by Dr. Kim, [98].

Simulations were done to take the Kirk effect into account. *BandProf* was used and the collector current density was simulated by replacing the collector doping N_c with an effective collector doping $N_{c,eff} = N_c - J_c/qv_{sat}$. To improve the high current tolerance the setback layer thickness was increased to 20 nm. The final grade is shown in figure 3.20.

Thinning of the collector was done *a priori* to reduce the collector transit time, with the knowledge that the maximum allowed current density would increase as $\sim 1/Tc^2$. As discussed earlier the maximum allowable collector doping is given by:

$$N_c \leq \frac{2\varepsilon_r \left[\Phi_{bi} + V_{bc}\right]}{qT_c^2} - \frac{2N_\delta T_\delta T_{setback}}{T_c^2} \tag{3.4}$$

Figure 3.21 is based on equation 3.4 and shows the maximum permissable collector doping level as a function of collector thickness and applied bias V_{bc}.

- $T_c = 300$ nm $\rightarrow N_c \leq 1.1 \cdot 10^{16}$ cm^{-3}
- $T_c = 200$ nm $\rightarrow N_c \leq 2.5 \cdot 10^{16}$ cm^{-3}
- $T_c = 150$ nm $\rightarrow N_c \leq 9.9 \cdot 10^{16}$ cm^{-3}

For our 217 nm thick collectors we chose the collector doping to be $N_c = 2 \cdot 10^{16}$ cm^{-3} and for our 150 nm thick devices we chose $N_c = 3 \cdot 10^{16}$ cm^{-3}. It is wise to leave a margin or error for the doping levels.

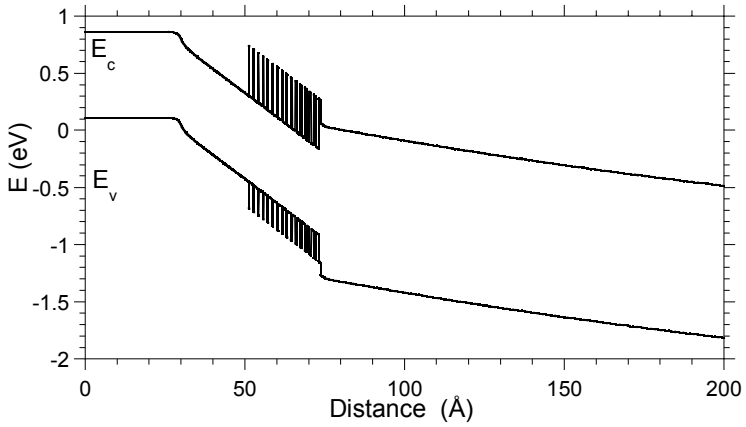

Figure 3.20. The final grade design, used in DHBT-17 onwards

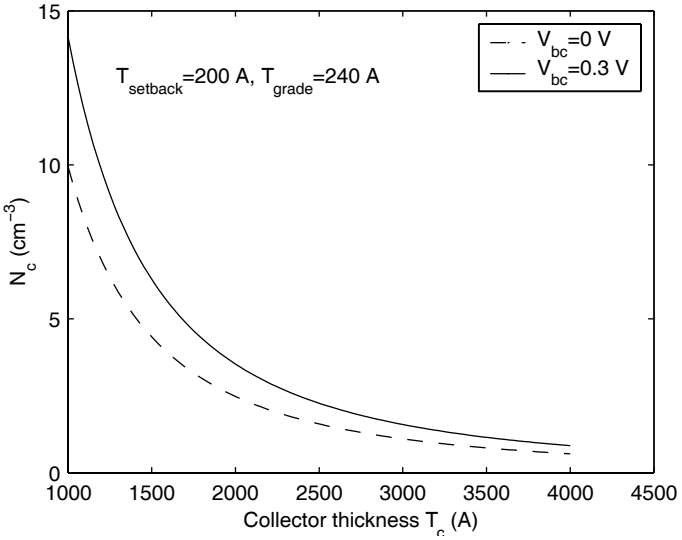

Figure 3.21. Maximum allowed collector doping level. $\Phi_{bi} = 0.96$ eV.

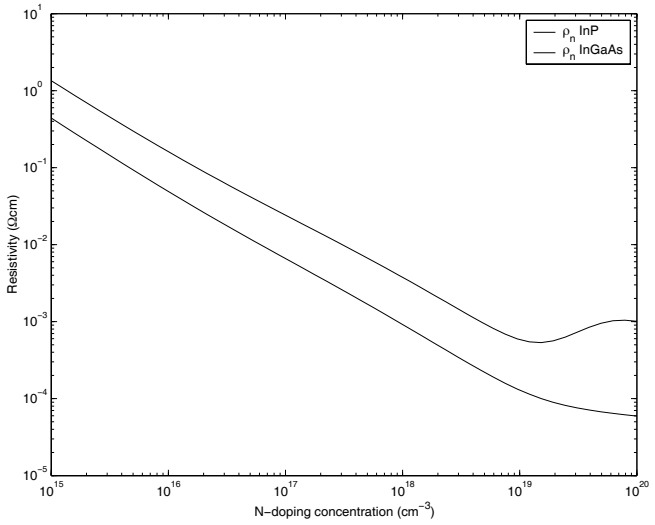

Figure 3.22. Resistivity of n–InP and n–InGaAs. The fit for n–InP above $1 \cdot 10^{19}$ cm^{-3} is extrapolated.

3.5 Subcollector design

In order to run the mesa HBT's at a current density high enough for them to compete with transfered substrate HBT's in circuit speed all obstacles have to be removed. Through proper base–collector design the Kirk threshold can be increased manyfold, and through using an InP collector a major thermal resistance is removed. The thermal conductivity of InP is (figure 2.7) $\sim$ 68 W/mK, while InGaAs has a thermal conductivity $\sim$ 5 W/mK. It is clear that changing the collector is not enough, the subcollector need to be optimized as well [93]. A subcollector that employs 50 nm InGaAs and 200 nm InP has a thermal resistance that is 26 % of the same thickness in an all-InGaAs subcollector. If the InGaAs thickness is reduced further, as was done in designs DHBT-19 and onwards, the thermal resistance is 19 % for a subcollector with 25 nm InGaAs and 300 nm InP compared to an all-InGaAs subcollector. It is possible to reduce the InGaAs portion even more, but at some point the contact resistance will begin to suffer and there is a risk of losing wet etch selectivity. To make sure the electrical resistivity did not suffer from the change to a composite subcollector a simple calculation was carried out (figure 3.23), where the resistivity data are from [15]. The InGaAs-InP junction is assumed to have no resistance due to the high doping level. The simulation result was used to fix the width of the subcollector contact in the mask-set to 2.6 μm.

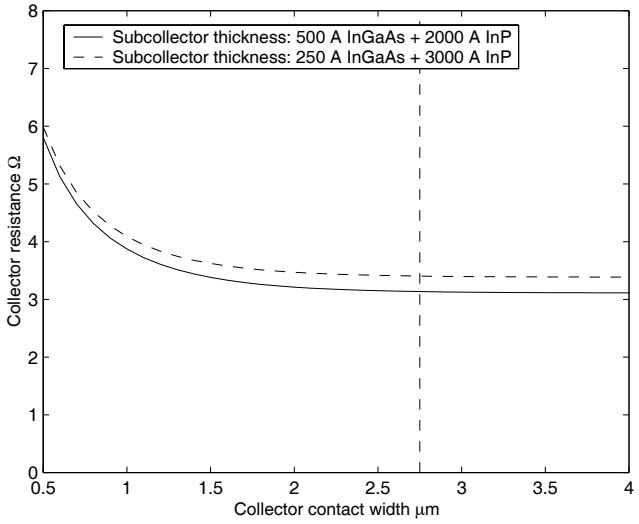

Figure 3.23. Collector resistance for composite InGaAs/InP subcollector. The line indicates the chosen collector width

The resistance in the subcollector can be calculated in the same way as the base resistance (equation 2.25):

$$\begin{aligned} R_{cc} &= R_{c,cont} + R_{c,gap} + R_{c,spread} \\ R_{c,cont} &= \sqrt{\rho_{c,s}\rho_{c,c}}/2L_c \\ R_{c,gap} &= \rho_{c,s}W_{bc-gap}/2L_{bc} \\ R_{c,spread} &= \rho_{c,s}W_{bc}/12L_{bc}. \end{aligned} \tag{3.5}$$

The calculated resistance is on the order of 3-4 Ω. L_c is the collector contact length, W_{bc-gap} is the spacing between the base–collector mesa and the collector contact, and L_{bc} and W_{bc} are the length and the width of the base–collector mesa, respectively. As our contacts are horse–shoe shaped the actual resistance is slightly smaller, with an extra contact access length contribution from the interconnecting section on the order of W_{bc}/L_{bc}.

3.6 Design of RF waveguides

Transfered substrate HBT's have a very clean RF environment [1]. The waveguides are on top of 5 μm BCB with an unbroken metal ground plane beneath. The same solution is not readily available for a mesa HBT. Of the available solutions we chose

Table 3.6. CPW calibration structures properties

Frequency GHz	Length μm	Delay ps
5-45	1204	9.93
70-140	345	2.85
140-220	171	1.41

narrow coplanar waveguides (CPW) directly on the semi–insulating InP substrate. It would have been possible to place the waveguides on polyimid but the varying thickness would lead to different propagation properties in different parts of the wafer. Polyimid is spun on and we expected a thickness variation over the wafer.

Using *Agilent's Linecalc* utility the necessary properties for 50 Ω lines were calculated. One design consideration was the need to keep the inner conductor as narrow as possible due to the risk of crosstalk between input and output due to substrate mode coupling. At the actual transistor very wide inner conductors would lead to capacitive coupling between the terminals. A constraint for how narrow the waveguide can be is set by the current density design rule. Given a maximum current density for sustained operation in interconnects of 5 mA/μm^2, a 7 μm wide conductor can operate at 35 mA. The latest HBT's reported in this thesis operate at current levels exceeding this. Keeping the dimensions compact also reduces the size of the CPW modes, in order to reduce substrate coupling and resistive losses. The real part of the mode propagation constant — the effective dielectric constant — is calculated to 6.12.

The mask set is equipped with on-wafer transmission-line extended reference planes and line-reflect-line calibration structures for 5-45, 70-110 and 140-220 GHz measurements [6,60]. Every measurement band requires transmission through lines at the geometric center frequency for a proper calibration of CPW through lines:

The mask set was designed with a large number of different transistor dimensions, encompassing emitter widths from 0.4 μm to 2 μm, and a majority of these were prepared for RF–measurements (figure 3.24). They are intended for GSG–type microwave probes, with a probe spacing of 75 or 150 μm. To reduce probe to probe coupling the contact pads are 230 μm away from the transistor. The DC resistance of the signal line is about 1 Ω. In the latest revision of the mask-set the probe pitch was reduced to 75 μm in order to decrease the likelihood of probe to probe coupling.

During measurements on the first high–speed devices, measurement problem were encountered probably arising from:

- Substrate coupling to fundamental mode
- Possible coupling to lateral micro–strip modes

The first problem was dealt with by thinning the substrate down to 120 μm, and performing the measurements on a piece of ferrite filled epoxy. The reduction of the

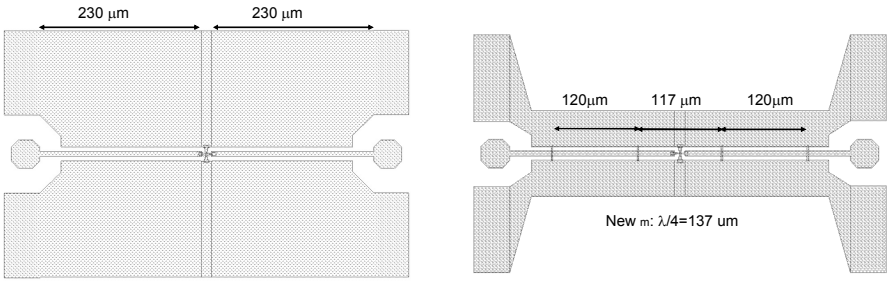

Figure 3.24. Wide CPW of the first iteration (left). Narrow CPW of the second iteration (right)

substrate thickness pushes the frequency of the fundamental mode above 220 GHz, the maximum measurement frequency. The frequency of the fundamental, or slab modes, are given by [94]:

$$f_n = \frac{n}{2T_{substrate}\sqrt{\varepsilon_d\mu_d - \varepsilon_0\mu_0}}, \quad n = 1, 2, 3, \ldots \tag{3.6}$$

The cut-off frequency of the first surface wave modes becomes, for a 600 μm thick InP substrate:

- TE_0,$TM_0 \rightarrow f_{c0} = 0$ Hz — the parasitic micro–strip mode
- TE_1,$TM_1 \rightarrow f_{c0} = 96.2$ GHz
- TE_2,$TM_2 \rightarrow f_{c0} = 192.2$ GHz
- TE_3,$TM_3 \rightarrow f_{c0} = 288.5$ GHz
- TE_4,$TM_4 \rightarrow f_{c0} = 384.6$ GHz

If the substrate is reduced to 150 μm, f_{c1} is increased to 433 GHz, sufficiently high above the maximum measurement frequency. The ferrite provides a lossy interface for the fundamental substrate mode in order to reduce the coupling efficiency [97].

The second problem is the coupling to laterally propagating micro–strip modes at the CPW ground plane edges. The condition for possible mode excitation is:

$$D_{ground} \leq \frac{\lambda}{2} = \frac{c_0}{2\sqrt{\varepsilon_r} f_{stop}} \tag{3.7}$$

where λ is the mode wavelength the highest frequency of interest, f_{stop}. For 220 GHz D_{ground} becomes 189 μm. The CPW ground planes were thus redesigned to have a width significantly below this. The ground planes now extend 50 μm on each side of the inner conductor.

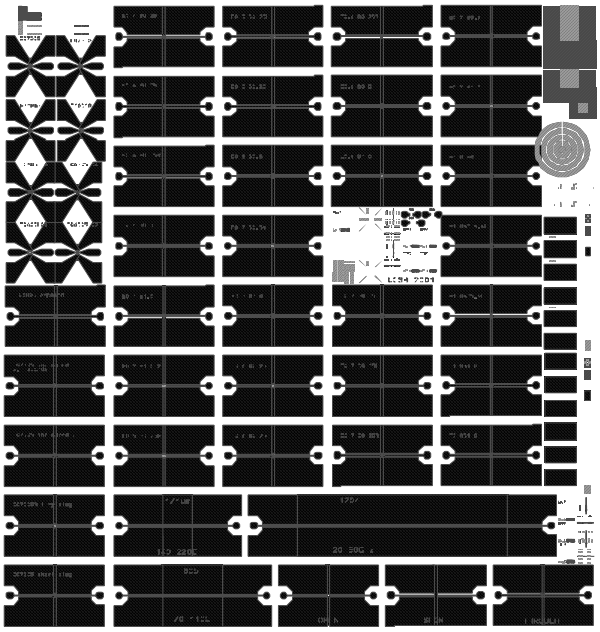

Figure 3.25. First iteration of mask set, with 35 sub-μm HBT's of different size

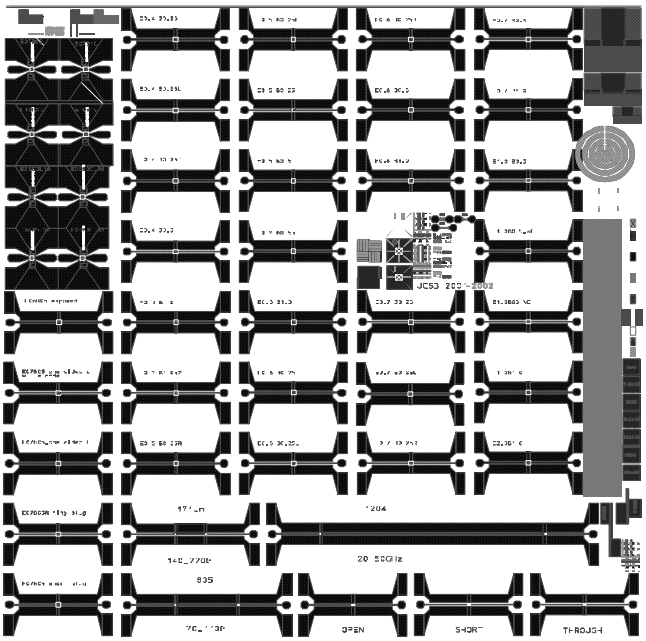

Figure 3.26. Second iteration of mask set with narrow waveguides to reduce parasitic coupling

The new mask set design also incorporated air bridges between the ground planes to suppress the slotline modes between them. The distance between air bridges was chosen to be less than half a wavelength at the maximum measurement frequency:

$$\Delta L_{air} \leq \frac{\lambda_{eff}}{2} = \frac{c_0}{2nf_{stop}} \tag{3.8}$$

The mode propagation constant n is roughly 7. At 220 GHz ΔL_{air} becomes 258 μm. In the design the air bridges were place closer to each other, 120 and 117 μm.

3.7 Mask set designs

The mask set contains 35 different HBT's, with the majority being 8 μm long and 0.4, 0.5, 0.6, 0.7 and 1.0 μm wide. The length of 8 μm was chosen to compensate for a rather high base contact resistance of 50 $\Omega\mu\text{m}^2$, and still provide a low base resistance, as $R_{bb} \sim 1/L_e$ (equation 2.25). The potential drop from the center of the emitter contact the ends is negligible due to the thick emitter metal employed. The effect was simulated with a distributed mesh model, and calculated the relative potential drop in the *metal* compared to the potential drop in the metal–semiconductor *contact resistance*. Assuming an unrealistically high and pessimistic metal resistance of 5 $\Omega - cm$, and a contact resistance $\rho_c = 25\ \Omega\mu\text{m}^2$, the relative potential drop in the metal is only 1.5 % for a 8 μm long emitter contact, seen from the center to each end. The emitter metal resistance is judged to have no influence and un-even turn-on is not expected. Base contacts extend 0.25, 0.5 or 1.0 μm on each side of the emitter. The range of base contact widths was chosen to span $L_d = \sqrt{\rho_c/\rho_s}$, the base transfer length, to empirically decide which width provided the maximum performance. The base is contacted through a base plug region that extends through polyimid and is level with the emitter contact (figure 4.8). By connecting the base via a plug in this way the extrinsic base–collector capacitance is reduced, since the contact to the base is separated by the polyimid. The mask dimensions of the plug are 1×2 μm, while in reality the area shrinks somewhat due to lithography. The area is $\sim 1.5\ \mu\text{m}^2$, and with the design rule $J \leq 5$ mA/mum^2 and $\beta \approx 25$, an I_b of around 1 mA provides plenty of margin current density–wise as the maximum allowed current would be 7.5 mA. Another place where the current density in the metal must be checked is the beginning of the base contact, right near the base plug. This is where the base current density is maximum. In the first iteration of the base mask the width of this region was the same as the width of the base, and in the second iteration it was always 0.5 μm wide. For the 0.5 μm wide case the maximum current is again on the order of 1 mA, and the cross–section area of the base metal is $2 \times 0.5 \times 0.1\ \mu\text{m}^2$. The resulting current density is 10-20 mA/μm^2, which is an area of concern. When the mask set was designed, collector currents seldom exceeded 15 mA, and with the higher gain of those lower doped HBT's, this resulted in base currents a factor of two or more lower than what we see today.

All but one HBT have a double–sided, horse–shoe shaped collector contact. The separation base–collector contact is only 0.5 μm but with automatic alignment that is usually not a problem. The width of the collector contact is determined in section 3.5, and results are shown in figure 3.23. The width of the collector contact is constrained by the parasitic capacitance to the emitter interconnect.

The emitter is connected to ground through an interconnect metal bridge (figure 4.9). In the first iteration of the mask set the bridge was 2 μm wide, and the width was constrained by fear of parasitic capacitive coupling to base and collector. Calculations showed, however, that the capacitance is small if the polyimid thickness is large enough, and by reducing the collector metal thickness, and increasing the emitter metal thickness the separation could be increased to 0.8-1 μm. Simulations performed by Dr. Harrison at the University of Nottingham [93] showed that about 30 % of the heat generated in the HBT was heat–sinked through the emitter contact. By making the bridge wider the amount of heat transfered through the emitter could be increased and the device current density increased. The bridge covers 5.5 μm out of an 8.0 μm emitter contact.

Chapter 4

Processing

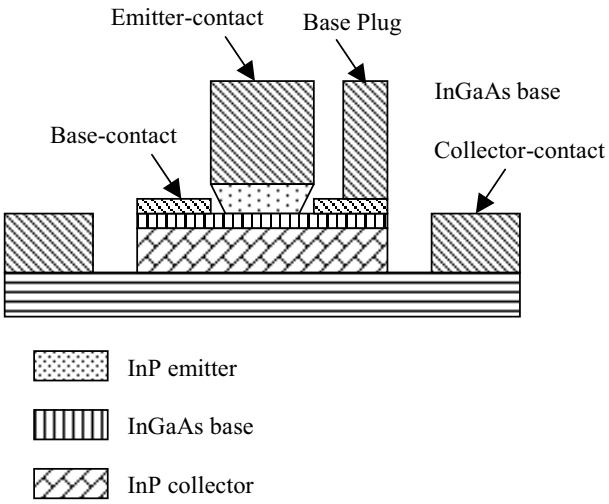

Figure 4.1. Schematic of a mesa HBT

4.1 Overview of the process

4.1.1 Choice of process

The process is based on previous work by members of the Rodwell group (particularly S. Krishnan) and the process for the mesa DHBT is the result of the marriage between that and the author's experience. The process is based upon a number of limitations set by materials, available machines, and the need to be conservative as the process was upgraded. The choice of an all–wet etched process was done due to the fact that the RIE capable of etching InP used hydrogen. Hydrogen containing plasmas are not readily compatible with carbon doped InGaAs bases (Appendix B.1, [88]). This necessitated wet etching of the InP. The choice of etching solution was given to a phosphoric acid: hydrochloric acid mixture that provides very good selectivity to InGaAs, and a controllable emitter undercut [89]. The choice of ashing procedure is described in section 4.2.1 and the choice of base metal in section 2.4. Photoresist SPR 955-09 was chosen instead of SPR 950-08 due to the deeper depth of focus for the former, and that was in part the reason to choose nLOF instead of AZ 5214.

4.1.2 The process

After wafer cleaving and marking with relevant information an initial solvent clean is performed. Following a dehydration bake – part of all photo resist steps and will not implicitly mentioned here after, for more in depth coverage see appendix C –

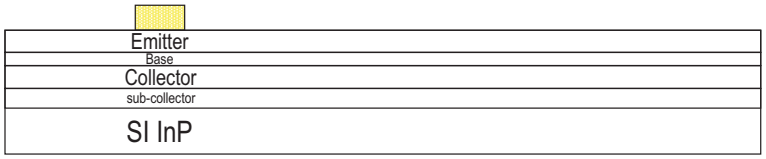

Figure 4.2. Emitter contact.

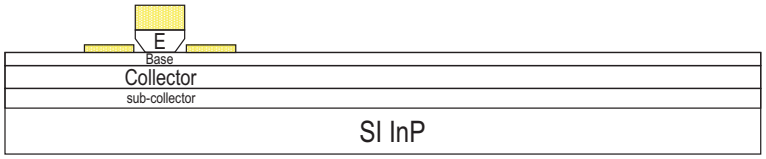

Figure 4.3. Etch down to base and lift–off base contact.

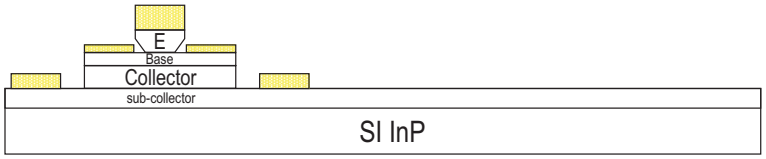

Figure 4.4. Etch down to collector and lift–off collector contact.

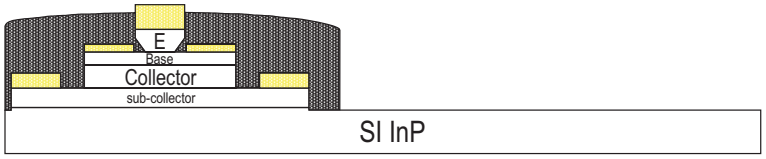

Figure 4.5. Planarize with polyimid and etch–back so emitters and base–plugs are exposed.

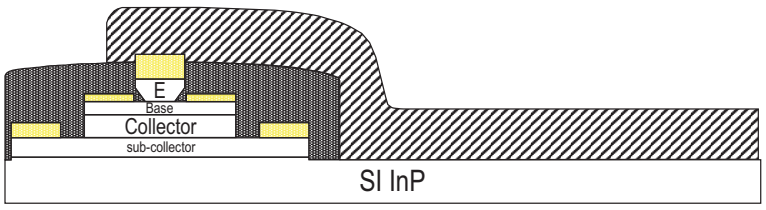

Figure 4.6. Open via–hole for collector, and contact terminals with interconnect metal

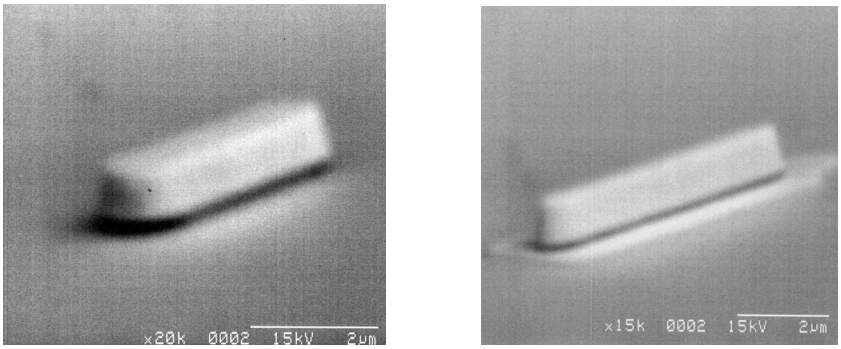

Figure 4.7. After the emitter–base etch (left). After the base contact deposition (right)

emitter lithography is performed (figure 4.2). The step is done with SPR 955-09 with the new CEM top hardening procedure. Following development UV–ozone ashing is performed and pumpdown for E–beam evaporation is done. Care has to be taken during mounting of the sample in the evaporator chamber to ensure that the wafer is perpendicular to the source, to provide better lift–off. Evaporation starts in the low 10^{-7} Torr regime, and the shutter is opened only after the melting of source provides a pressure drop in the chamber, with the exception for gold where this pressure drop does not occur. Lift–off is done in 1165, preferably overnight in a covered beaker. Take has to be taken during lift-off – if the wafer is taken up into air while metal is still adhering to the surface that metal will be virtually impossible to remove. After solvent clean the wafer is inspected and SEM's are possibly taken. Following an ashing step the emitter cap and emitter wet etching steps are performed. The ashing step is done using a plasma etcher – the idea is that the more aggressive ashing will help remove small thin strands sometimes hanging from the emitter metal.

The first etch will remove InGaAs and stop on InP. InGaAs has a greenish color and InP has a reddish color – if the layers are thick enough. A trick to estimate the exact etching time is to dip a sacrificial corner of the wafer for 10 seconds, and then etch the entire wafer. For a selective etch the final color of the wafer will occur first in the corner, and etching is stopped when the color of the wafer is the same all over. If undercut is desired the etching is performed a bit longer. The etching will stop on the InGaAs base, or if the base-emitter junction is graded, in the grade (figure 4.3).

Base lithography is performed using nLOF and local alignment. This is a critical step: the alignment needs to be very good and rework is undesirable. The base is very thin (25–30 nm) and every iteration removes a thin layer of base through oxidation. Following inspection of the lithography – the undercut should be visible as an extra dark line around the photo resist edge – UV ozone ashing is performed. Following an NH_4OH oxide etch E–beam evaporation takes place. It is important

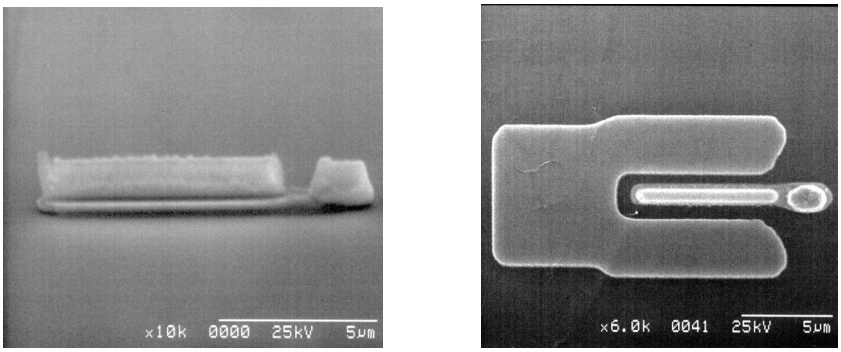

Figure 4.8. After the base–collector etch, with base plug in place(left). After collector contact deposition (right)

that the E–beam evaporator thickness monitor is working properly – if too much first layer Pd is deposited shorting of the base–collector junction can occur upon annealing as the Pd reacts with the semiconductor. Also, if too much metal as a whole is deposited the base metal will come into contact with the emitter cap or emitter metal and a base–emitter short will occur. Lift–off is performed in 1165 and following a solvent clean the sample is annealed using Rapid Thermal Annealing (RTA) for 1 minute at 300 °C. In this annealing step the Pd will react with the InGaAs as described in section 2.4.2. The color of the base metal will change from bright gold to darker gold. The anneal chamber is baked out with nitrogen at 120 °C for 3 minutes with the sample inside before ramping the temperature up to 300 °C to drive out oxygen and water vapor

The next step is to place the base plug on the end of the base. The thickness of the plug is such that it becomes level with the emitter metal and the metal thickness is therefore measured. Base plug lithography is done using the same procedure as the emitter lithography, but with a slightly higher dose and development time, to ensure proper exposure of the $1 \times 2\,\mu$m plug. Metal deposition and lift–off is done as for the emitter. The next step is to etch the base–collector junction down to the subcollector. In an InP/InGaAs DHBT this step is made simple due to the excellent etching selectivity present between InP and InGaAs. First, it is necessary to cover the emitter-base region with a photo resist mask. This lithography is done with SPR 518, a 1.9 μm thick positive photo resist with good wet etch resistance. Following development and an ashing step the two step wet etching sequence takes place. The first etch etches the base, the InGaAs setback and the InGaAs/InAlAs super lattice grade (check table A.3). The second step etches InP down to the subcollector. This step is over etched to provide 0.1-0.2 μm undercut in the base–collector junction, to reduce the collector capacitance [90]. The change of color when etching and the measured depth is monitored. The photo resist is rinsed away with 1165, and following a cleaning step collector lithography takes place. The procedure is similar to the base lithography and can be performed using either

nLOF 2020 or Shipley AZ 5214. A word of caution is needed here: compared to AZ 5214 nLOF provides metal lines 0.1 μm wider. This needs to be taken into account when doing mask design. Measurement of TLM pattern for the base can now take place. Measurement of the true TLM separation needs to be done. For the very low contact resistances obtained the TLM metal pad separation plays a large roll for the contact resistance value and needs to be determined. Collector metal deposition of Ti(20 nm)Pd(20 nm)Au(300 nm) is performed (figure 4.4). The thickness of the metal is limited by the extrinsic base–collector and emitter–collector capacitance. The emitter is contacted by a metal line running over the polyimid, separated by a layer of polyimid, and the if the collector metal is too thick there is even a risk of shorting. This happened is the first mask revision when the base plug and the collector metal were the same step, and the thickness of the collector metal was ≈ 1 μm.

The next step is mesa isolation: the HBT mesa is covered with photo resist like in the base–collector etch. This lithography is done with SPR 518. Following development and an ashing step the two step wet etching sequence takes place. The first etch etches the InGaAs subcollector. (table A.3). The second step etches InP down through the InP n^+ subcollector $\approx$ 100 nm down into the semi–insulating substrate. The depth is monitored and the resistance between two collector metal pads is measured to ensure the subcollector has been etched completely. The photo resist is rinsed away with 1165 and a solvent clean. Following a dehydration bake an adhesion promoter is spun unto the wafer, followed by polyimid (HD Microsystems Pyralin PI 2556), (figure 4.9). Care has to be observed to avoid bubbles forming as they will lead to craters of polyimid. The backside needs to be protected from polyimid as the wafer will otherwise be glued to the holder in the polyimid bake. The polyimid is baked in an oven under N_2 atmosphere for 6 hours in temperatures up to 270 °C. The bake hardens the polyimid into a plastic dielectric material. After the bake a thick layer of photoresist AZ 4330 is spun on the wafer. The purpose of the resist is to planarize the surface for the ensuing polyimid blanket back–etch. The etch is done in an RIE with oxygen plasma at 10 mTorr pressure. The etching is monitored by a HeNe laser and a photodetector. As the photo resist and the polyimid are being etched the reflected laser signal undergoes sinusoidal oscillations and after $\approx$17 cycles the top of the emitter and base plugs are exposed and the etch is halted. The sample is patterned with SPR 518 so the HBT is masked and surrounding areas. The polyimid in the surrounding areas are etched in a plasma etcher in oxygen at 300 mTorr with the photo resist as mask. The final step is interconnect metalization, (figure 4.6). Lithography is done with nLOF 2020 and 1 μm (Ti(20 nm)Pd(20 nm)Au(1000 nm)) are deposited using E–beam. Following lift–off the samples are ready for measurement.

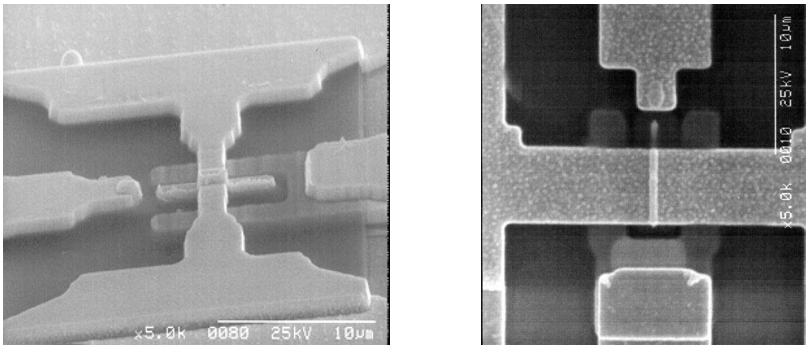

Figure 4.9. Interconnect metal contacts the device (left). Wider metal for improved thermal heat sinking (right)

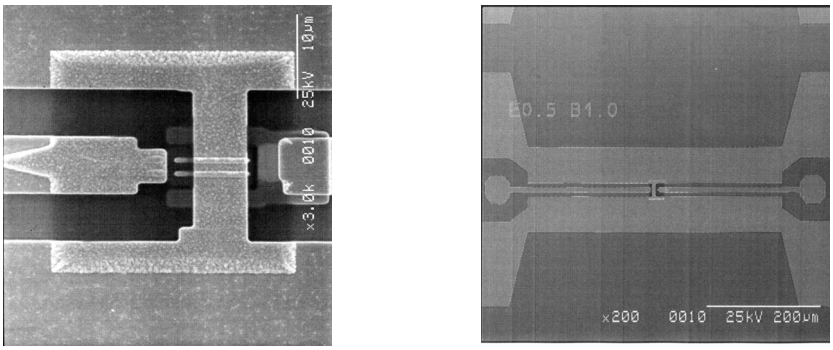

Figure 4.10. Interconnect metal to double emitter HBT (left). Overview of HBT with probing pads and waveguides (right)

4.2 Process improvements

The purpose of this chapter is to illustrate some of the problems addressed in the course of making these HBT.

4.2.1 Ozone

Initially, all HBT work at UCSB used an oxygen plasma etcher for photoresist residue removal (ashing or descumming). In a plasma etcher the semiconductor surface is exposed to energetic ions and damage of the surface might occur [79]. In advanced LSI labs ashing is often made using systems that do not expose the surface to energetic ions, such as downstream descummers. It was felt that the plasma etcher might not be optimal for sensible contacts, and for the base contact step ashing was not performed at all [80], instead the development was carried out twice in an attempt to remove photo resist residue. Based on research showing ultraviolet ozone plasma treatment (UV–ozone) being a gentle yet effective way of cleaning the surface, and the fact that the UCSB clean room actually had a UV–ozone photo reactor, all critical ashing steps were transfered to this system [86]. UV–ozone is critical for achieving the extremely low base contact resistance we have observed [81]. Base contacts cleaned by UV–ozone have a contact resistance $1 \cdot 10^{-8}$ to $5 \cdot 10^{-7}$ $\Omega - \mathrm{cm}^2$ while base contacts cleaned by oxygen plasma etcher have a contact resistance $1 \cdot 10^{-6}$ to $1 \cdot 10^{-5}$ $\Omega - \mathrm{cm}^2$. The resist etching speed is roughly 1 nm/min. Reports indicate the oxide created by the treatment is denser than usual and that it can also be used for surface passivation [86].

4.2.2 Resists

Lift–off with positive photoresist

All the metalization steps for the HBTs presented in this thesis were done by E–beam evaporation and lift–off. Lift–off is a technique in which the photoresist profile gets a negative slope. During E–beam deposition, the deposited metal follows line–of–sight and a small region will be masked by the overhanging photoresist. When the sample is placed in a solvent the photoresist dissolves, and a defined metal structure remains. The emitter metalization step is particularly demanding: the line width is as narrow as 0.5 μm and the emitter metal is 900 nm thick. As shown in figure 4.11 the metal edges are ragged and can lead to problems later in the process [87]. Experiments with top surface hardening took place together with Z. Griffith, based on [82]. Dipping the photoresist (SPR 955-09) in developer (MF-701) or spinning on contrast enhancement layer (CEM) can both create cross linking of the top part of the photoresist, and create the desired undercut profile upon developing. A timed surface hardening step with CEM turned out to give very good resist profiles, enabling 0.5 μm thin lines 1 μm thick with smooth edges (figure 4.11).

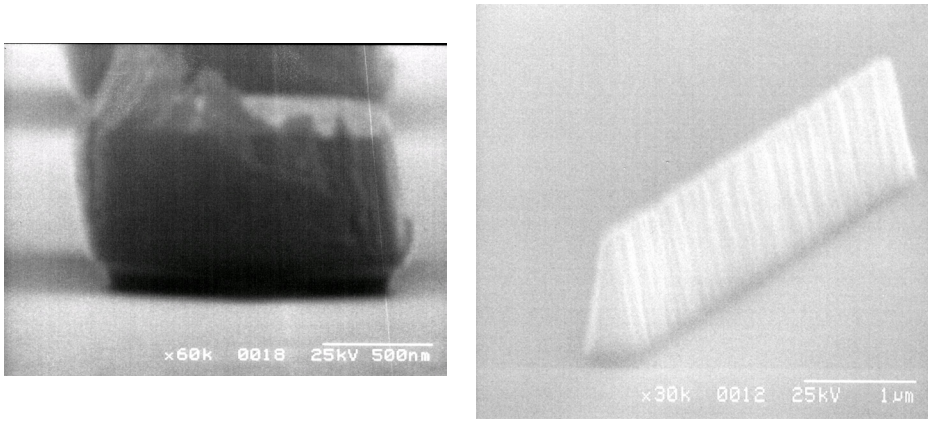

Figure 4.11. Poor lift–off profile (left) and improved lift–off profile. The emitter is 0.5 μm wide and 1 μm high (right). Photo courtesy of Z. Griffth.

Lift–off with negative photoresist

Negative photo resists work by a double chemical reaction step: the regions that are to be patterned are actually masked in the lithography step, and the surrounding areas exposed. Following a bake step that hardens the illuminated area the entire wafer is exposed and developed. The non–illuminated regions have not their photo resist removed, and if done correctly, the photo resist sidewall profile is negative. Evaporation and lift–off is done as described in section 4.2.2. This was performed with Shipley AZ 5214 but the reproducibility was poor, the necessary exposure, bake, and development times had to be adjusted on a weekly basis. An alternative was sought for and was found in Shipley nLOF 2020, a new type of negative photo resist that does not need a flood exposure step [83]. The new process is capable of defining 0.6 μm wide 100 nm thick lines, and for the interconnect metalization step it is capable of 1 μm thick lines with excellent side wall profile and step coverage, figure refnLOFimage. The only disadvantage with nLOF is the difficulty to remove it. Double baths of warm 1165 are not always enough and stronger strippers are necessary. Certain, such as PRX-127, attack Ti metalizations and etch hard baked polyimid.

4.2.3 Resist removal

Initially photo resist removal was done with acetone, followed by warm methanol and warm propanol. Disregarding the fire danger, the resist removal was often incomplete, leaving hazy shades of photo resist over the sample. An alternative was found in *Photoresist Remover 1165*. Especially when warmed to 80°C 1165 provides a marked improvement in apparent cleanliness. It is now used instead of acetone for photoresist removal.

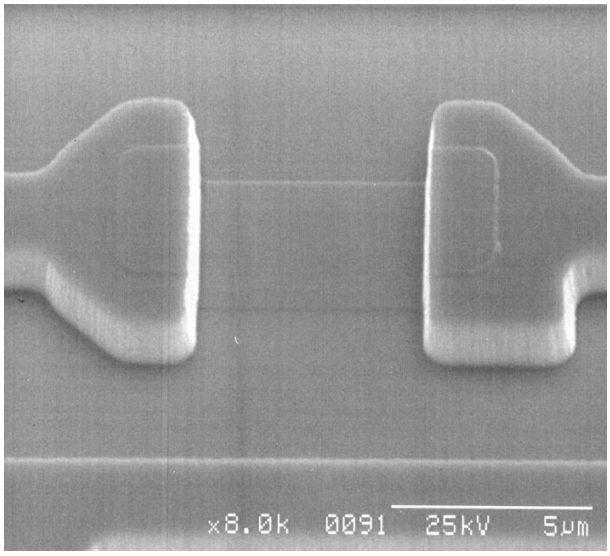

Figure 4.12. Metalizations done with improved negative photoresist nLOF.

4.2.4 Metal purity

The UCSB clean room is by it's nature a very busy environment and minor mishaps happen. Contact resistance to the base is critical to our device performance (equation 2.25, section 2.4) and a way to supervise the metal purity in the E–beam evaporators was needed. Co–depositing on laboratory type glass slides was done, and the metal resistance was then characterized using 4-point probing. The results are presented in section 5.3.2.

4.2.5 Stepper optimization

The stepper at UCSB was at the onset of this work used in manual mode – the substrate alignment was manipulated with a joystick – and rework was often necessary to achieve better than 0.5 μm alignment. To facility fabrication of mesa HBT's either the tolerances for base alignment had to be relaxed or a stepper improvement was needed. It turned out that the stepper is equipped with a automatic alignment system, that reads alignment targets on the mask and on the wafer and automatically optimizes the alignment. This is a standard feature on steppers in semiconductor industry. After some dedication the automatic alignment procedure (*local alignment*) was calibrated and helps the HBT work at UCSB achieve better than 0.2 μm alignment almost all the time. 0.1 μm alignment is commonly observed.

Chapter 5

Results

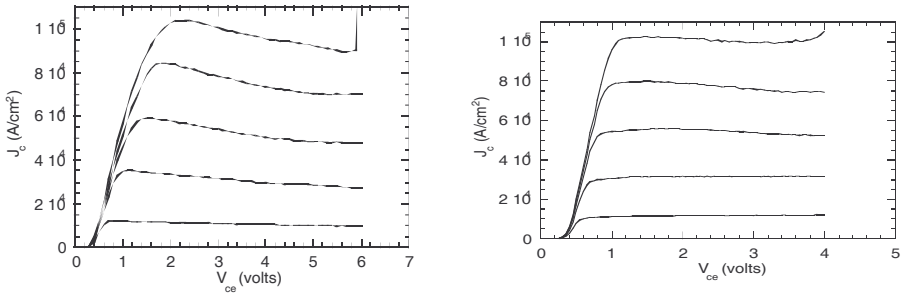

Figure 5.1. DC characteristics for DHBT-1 (left) and DHBT-2 (right) with 48 nm grade

5.1 Early designs : Grade problems

The starting designs had a 48 nm thick base–collector grade, without an explicit setback layer (table A.2). For a summary of all structures reported here, see table A.1. The devices were processed by K. Sundarajan using the transfered substrate technique. One sample had a 300 nm thick collector (DHBT-1), the other had a 200 nm thick collector (DHBT-2). The bases were 40 nm thick, Be doped to $4 \cdot 10^{19}$ cm^{-3}. The first HBT (DHBT-1) had a f_τ=165 GHz and f_{max} = 300 GHz, at a current density of 1 mA/μm^2. The second HBT (DHBT-2) had a f_τ=215 GHz and f_{max} = 210 GHz, at a current density of 1 mA/μm^2. These were the first successful DHBT made at UCSB, [95].

The following designs had a much thinner base–collector grade, 10 and 20 nm, without an explicit setback layer. The purpose of thinning the grade was to increase the collector velocity. The expectation was that the 48 nm thick super–lattice layer presented an substantial electron obstacle and was therefore redesigned and thinned down. One sample had a 170 nm thick collector (DHBT-3) with a 20 nm grade, and one sample had a 300 nm thick collector (DHBT-5) with a 10 nm grade. The bases were 40 nm thick , Be doped to $4 \cdot 10^{19}$ cm^{-3}. The first HBT (DHBT-3) had a f_τ=120 GHz and f_{max} = 200 GHz, at a current density of $\approx$ 1 mA/μm^2. The second HBT (DHBT-5) had a f_τ = 128 GHz and f_{max} = 120 GHz , at a current density of 1.3mA/μm^2. The RF performance of these devices was very disappointing, and the current density for highest f_τ and f_{max} was very low, indicative of poor electron transport in the collector.

As an attempt to increase the electron velocity the composite collector was introduced, in which 1/2 to 1/3 of the collector was InGaAs, followed by a grade and an InP layer. The purpose of the InGaAs layer was to provide superior electron transport, and the role of the InP layer was to increase the breakdown voltage of the HBT.

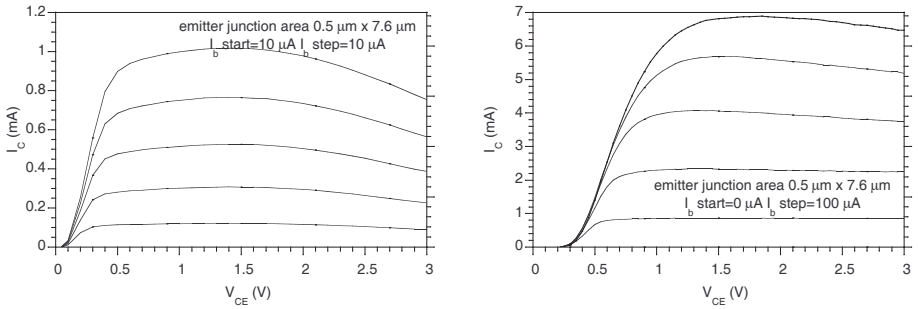

Figure 5.2. DC characteristics for DHBT-3 (left) and DHBT-5 (right) with 10 and 20 nm grade

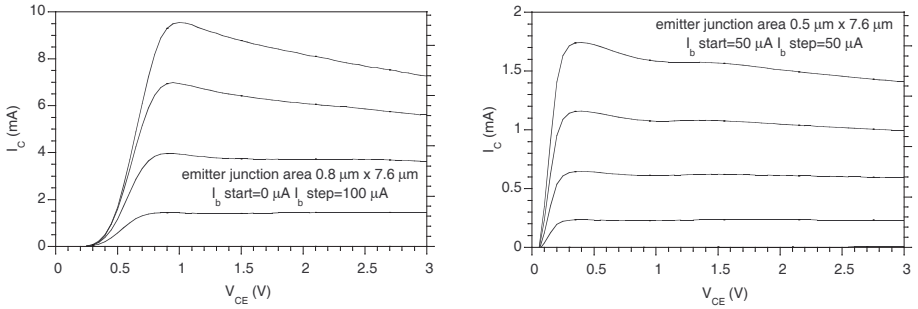

Figure 5.3. DC characteristics for DHBT-6 (left) and DHBT-9 (right) with 10 nm grade and InGaAs/InP collector

DHBT-6 had a combined collector InGaAs/InP, where the first 60 nm were InGaAs, followed by a 10 nm grade and 70 nm InP.

DHBT-9 had a combined collector InGaAs/InP, where the first 70 nm were InGaAs, followed by a 10 nm grade and 70 nm InP. The bases were 40 nm thick, Be doped to $4 \cdot 10^{19}$ cm^{-3}. The first HBT (DHBT-3) had a f_τ=126 GHz and f_{max} = 177 GHz, at a current density as low as 0.8 mA/μm^2. The second HBT (DHBT-5) had an f_τ=114 GHz and f_{max} = 144 GHz, at a current density of 1.6 mA/μm^2. The results were still unsatisfactory. Upon subsequent analysis of DHBT-3 to -9 certain factors stand out:

- Super lattice period up to 2.4 nm
- No setback layer, greatly decreasing the Kirk threshold
- Too strong δ-doping
- Very thin grade

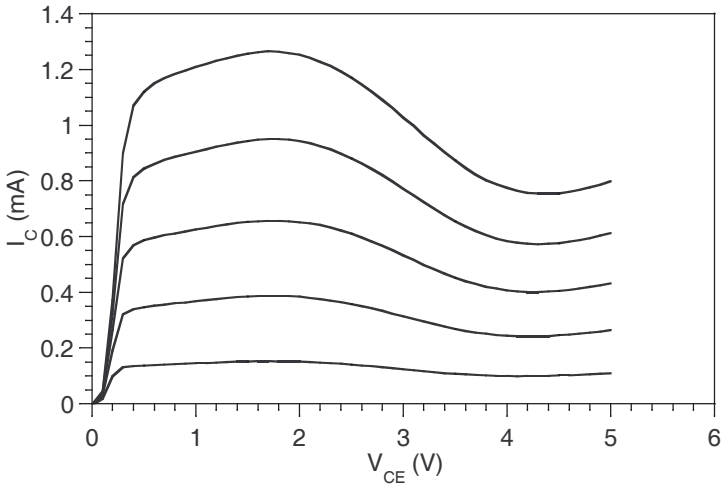

Figure 5.4. DHBT-3 showing evidence of current blocking with increased voltage

It is possible trapping could occur for grade layers 2.4 nm thick, especially if the grade is placed adjacent to the base with no setback layer. With a setback layer the electrons have more energy, reducing the likelihood for trapping (figure 5.4). A HBT without a setback layer is also sensitive for dopant out–diffusion from the base. The δ-doping was mistakenly designed for the total bandgap difference between InGaAs and InP, not just the conduction band difference. This made the doping roughly twice the required amount, and resulted in the conduction band dropping rapidly in the region between the base and the δ-doping.

$$N_\delta T_\delta = \frac{\varepsilon_r \Delta E_c}{q^2 T_{grade}} \tag{5.1}$$

One possible explanation for the low gain is Zener tunneling in the base–collector grade region since the electric field in the setback and grade region is higher than normal because of the higher doping. Distortions such as band to band recombination or $\Gamma - L$ scattering can subtract from the collector current and increase the collector transit time.

Evidence of current blocking is seen in figure 5.4. The shape of the I-V characteristic is independent of current level, indicating the non-ideal shape is not a thermal effect. The current level is still very low, less than 1 mA for a junction size of roughly 4 mA/μm^2.

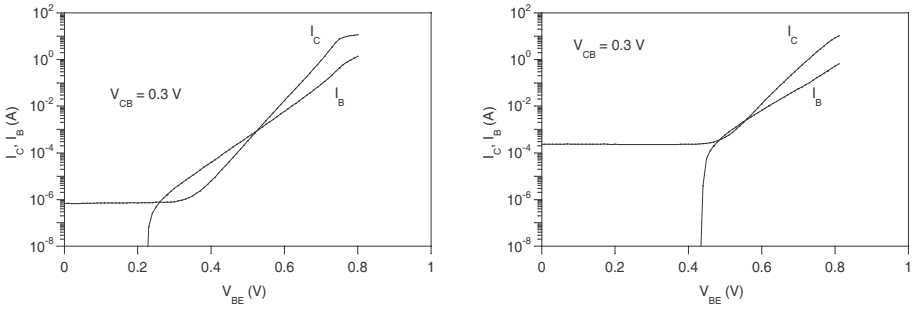

Figure 5.5. Gummel plots DHBT-5 (left) and DHBT-6 (right)

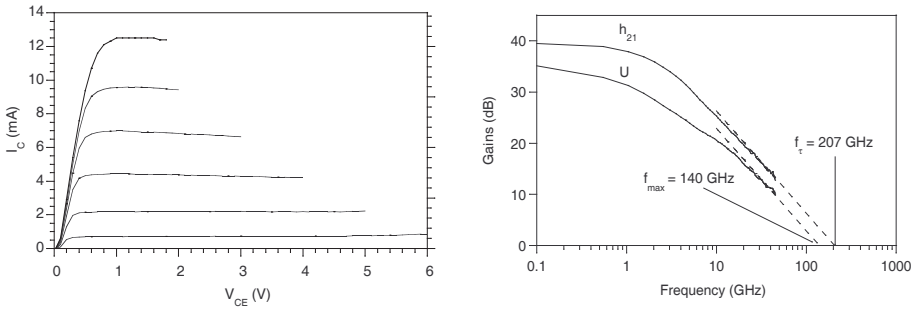

Figure 5.6. DC characteristics of the first device with the new grade (left), RF gain and cut–off frequencies (right)

5.2 Late designs: good grade

The first designs to thin the grade and make faster transistors were not successful. A new design was introduced with a 24 nm grade with the correct δ-doping level according to equation 5.1. A setback layer of 10 nm and later 20 nm was introduced. The first successfully manufactured device with the new grade was processed by Dr. Kim at UCSB [98]. These devices were grown on a GaAs substrate with a metamorphic buffer layer to accommodate the lattice mismatch. The base were 40 nm thick , Be doped to $4 \cdot 10^{19}$ cm^{-3}. The collector was 200 nm thick, of which the setback layer was 10 nm and the grade 24 nm . The collector doping level was $2 \cdot 10^{16}$ cm^{-3}. The HBT (DHBT-M1) had a f_τ=207 GHz and f_{max} = 140 GHz, at a current density of 4 mA/μm^2. The current density is suddenly much higher, an indication of a successful grade design. The current gain cutoff frequency f_τ of 210 GHz represent a respectable value, especially for a metamorphic wafer. These values were substantially higher than for previous designs.

Two new wafers were designed using the new grade and a 20 nm setback layer. Based on calculations of base transit time (section 3.3) the base was thinned to

30 nm, and doped with carbon in order to increase the doping level and compensate for the reduced thickness in terms of base sheet resistance. One wafer used a doping graded base (DHBT-17) and one used a bandgap graded base (DHBT-18). The wafer structure is shown in tables A.3 and A.4. Device data for the doping graded and the bandgap graded showed better device performance for the doping graded device, and subsequent device structures used doping graded bases. DHBT-18 (figure 5.8) showed poor passivation, low, current-dependant gain and base-emitter leakage.

DC gain for DHBT-17 was 22-25, while DHBT-18 showed a lower gain of around 18 at high current densities. DHBT-17 showed very impressive performance with a record-performance f_{max} of 400 GHz and one of the best f_τ of 282 GHz reported at that time [99]. The current density was 4 mA/μm^2 and $V_{ce} = 1.7$ V. DHBT-18 showed an f_τ of 250 GHz and an f_{max} of 356 GHz, but had problems with poor passivation, a problem that would plague subsequent devices. The current density was 4.9 mA/μm^2 and $V_{ce} = 1.7$ V.

To increase the device speed and maximum current density two new designs were made with a 150 nm thick collector. DHBT-19 had a 25 nm thick carbon doped base with an abrupt base–emitter junction (table A.5). DHBT-20 was otherwise identical except for a grade between the emitter and base (table A.6). DHBT-21 was designed identical to DHBT-19 but with a thinner collector of 100 nm, but failed to yield devices.

DHBT-19 and DHBT-20 exhibited a gain of 10-30, with a maximum f_τ=280 GHz for DHBT-19 and f_τ=300 GHz for DHBT-20. Gain was 30 % higher for the abrupt device DHBT-19 than for the graded DHBT-20, which suggests ballistic injection. The current density in these devices is very high. A HBT with an 0.4×7.6 μm emitter junction had an f_τ=237 GHz and an f_{max}=255 GHz at a current density of 11.5 mA/μm^2 with $V_{ce} = 1.3$ V. This one of the absolutely highest values reported. Ida et. al. also report HBT operation over 200 GHz at a current density higher than 10 mA/μm^2 [3]. HBTs from wafers DHBT-19 and DHBT-20 experience repeatable device failure at 35 mA, or 70 mA for double emitter devices. This indicates the failure is not in the coplanar waveguides connecting the devices, but instead in some region inherent to each device, as the base plug or the emitter contact.

Processing was complicated by a number of factors such as stepper lithography machine problems, and poor passivation with polyimid, resulting in base–emitter leakage [101]. The result for DHBT-19 was for 2 μm wide emitters, which were the only ones to survive processing due to problems with passivation and emitter removal.

5.3 DC–measurements

Devices were evaluated for their I-V characteristics. Non-flat I-V curves requires an explanation: certain devices – especially transfered substrate devices – are sensitive to device heating, which shows up as a negative slope. A device with a poor grade

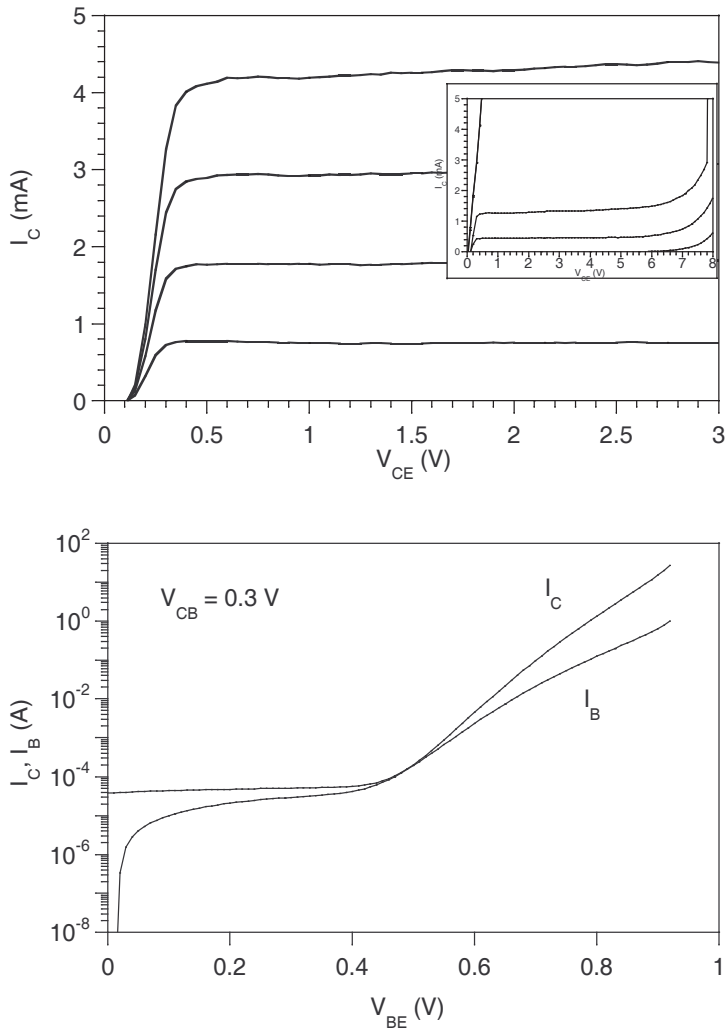

Figure 5.7. DC characteristics of the first device with doping graded carbon base

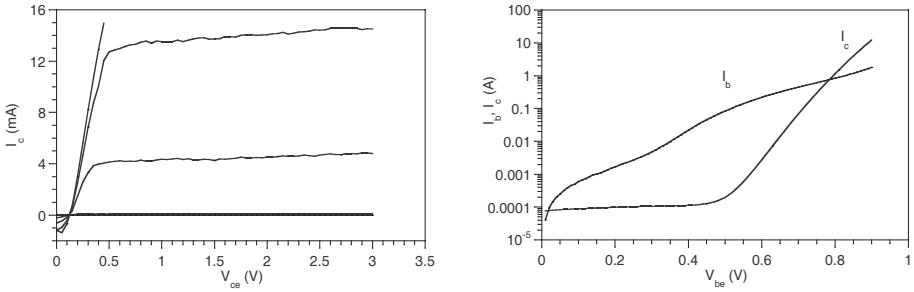

Figure 5.8. DC characteristics of DHBT-18 with bandgap graded carbon doped base, showing poor passivation

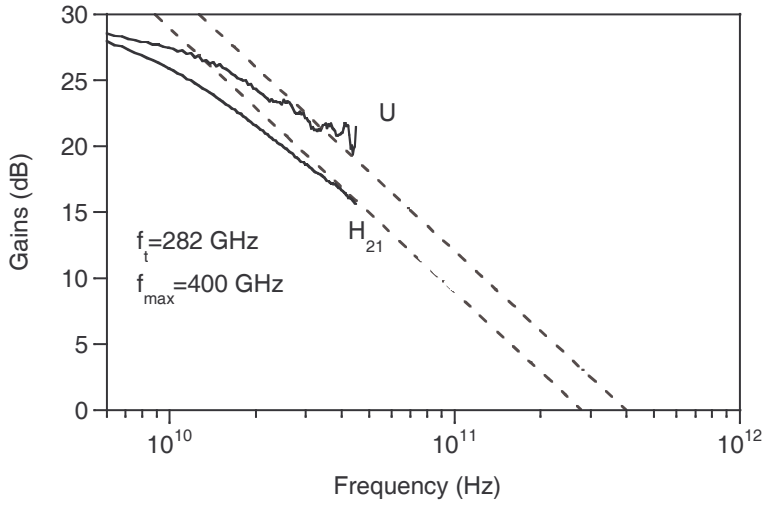

Figure 5.9. RF characteristics of the first device with doping graded carbon base, DHBT-17

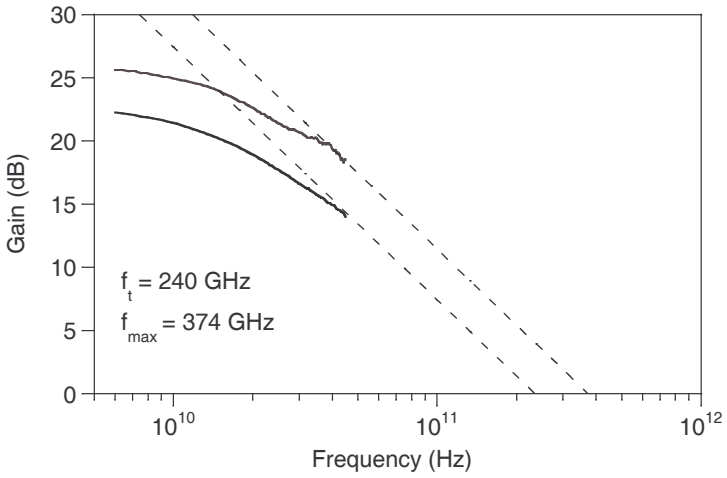

Figure 5.10. RF characteristics of DHBT-18 with bandgap graded carbon doped base

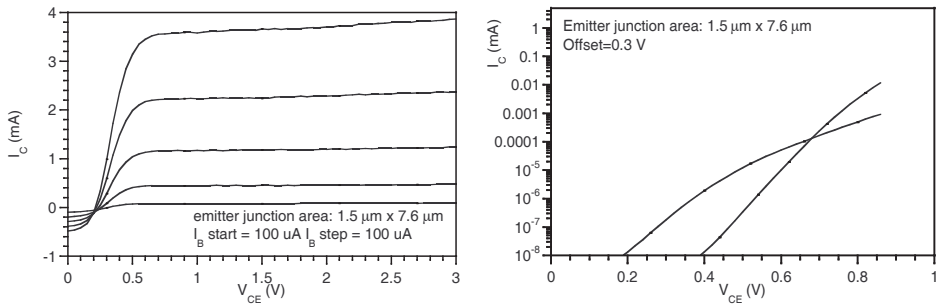

Figure 5.11. DC characteristics of DHBT-19 device with 150 nm collector

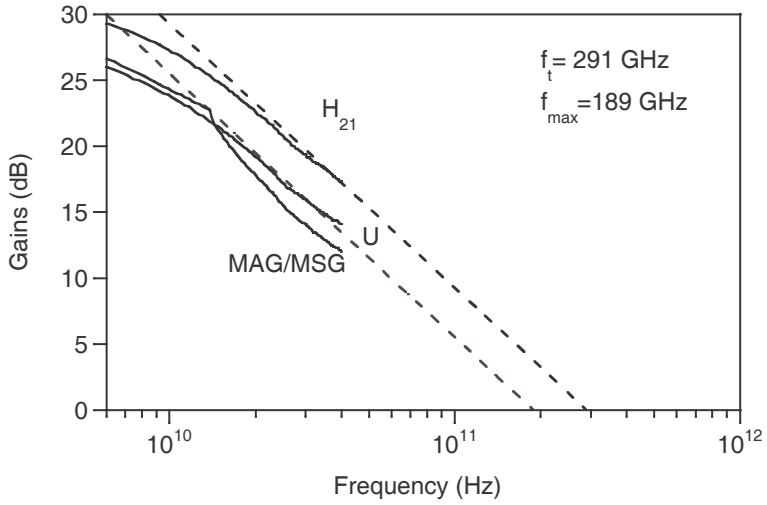

Figure 5.12. RF characteristics of DHBT-19 device with 150 nm collector

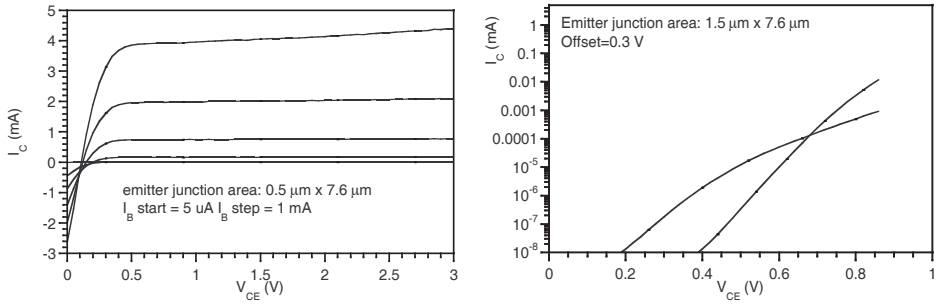

Figure 5.13. DC characteristics of DHBT-20 device with 150 nm collector

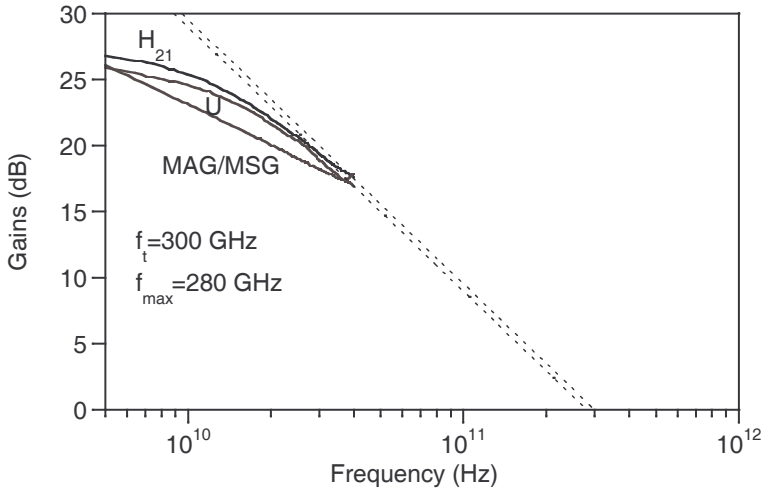

Figure 5.14. RF characteristics of DHBT-20 device with 150 nm collector

design might show undulations in the I-V curve as in DHBT-9 (figure 5.3). A low Kirk threshold shows as a drop in gain at higher V_{ce}, stronger for increased current density, as for DHBT-3 figure 5.2. In the gummel measurements information about the type of current injection is available as the slope for I_b and I_c. A third type of measurement performed are emitter resistance measurements using the fly–back method [12]. In this method a base current is driven through the emitter–base diode as a function of V_{ce}. The inverse slope of the curve corresponds to the emitter resistance.

Table 5.1. Emitter and collector TLM results

Type	ρ_s $\Omega/\square$	Contact ρ_c $\Omega-\mu m^2$
Emitter DHBT-18	20	11
Emitter DHBT-19	21	8
Subcollector DHBT-18	7	13
Subcollector DHBT-19	6	7

5.3.1 TLM–measurements

During the process transmission–line–measurements are performed to evaluate the process. The most critical contact is the base contact, and evaluation of the base resistance is done using two types of patterns: with the emitter left between the pads, and with the emitter removed and the base exposed. The measured resistance depends on the width of the structure L_{TLM} and the separation between the pads.

$$\begin{aligned} R_{TLM} &= \frac{\rho_s x}{L_{TLM}} + 2R_c + 2R_{probe} \\ \rho_s &= \frac{dR_{TLM}}{dx} L_{TLM} + 2R_{probe} \\ \rho_c &= \frac{1}{\rho_s} \left(\frac{R_{TLM}(0)}{2} L_{TLM} \right)^2 \end{aligned} \tag{5.2}$$

The width of the pattern L_{TLM} is 25 μm in our pattern, and the inter pad distance x is $1, 2, 3, 4, 5$ μm. The probe resistance is zero when using a 4–point probe measurement set–up. The sheet resistance correlates to the derivate of the measured resistance. $R_{TLM}(0)$ is the constant in the linear fit $y = ax + b$. For very small values of ρ_c the separation between contact x pads needs to be measured, as a small deviation gives a large error in ρ_c. The sheet resistance is less sensitive to variations in pad separations – the difference between designed pad spacing and measured is typically due to photo resist or metal evaporation influence, and offsets all spacings roughly the same amount.

A special circular pattern exists for measuring the emitter resistance without having to perform an isolation etch. Normally isolation of the structure is required so that current can only flow directly from contact pad to contact pad. If the measurement patterns are circular inside each other this can be achieved without isolation, as the direction of the current flow is well–determined [12]. Base TLM results are shown in table 5.5. The difference between extrinsic and intrinsic base sheet resistance is significant for devices with thin bases and abrupt base–emitter junction. The intrinsic emitter resistance is always lower than the extrinsic and it seems as if base material is missing in the extrinsic region. Assuming the intrinsic

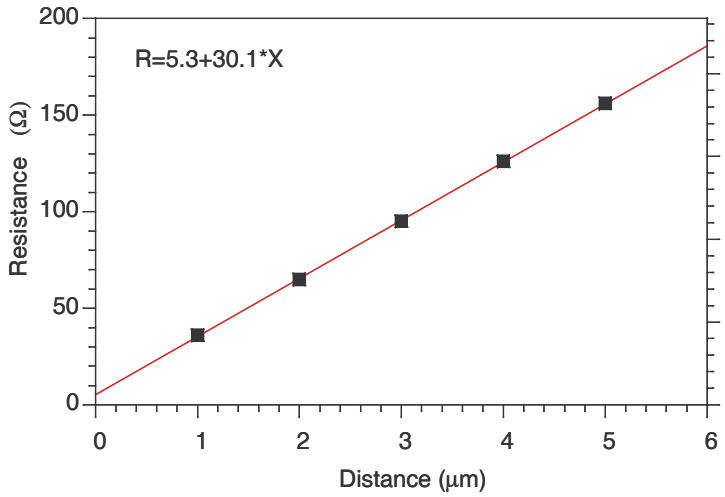

Figure 5.15. TLM data from DHBT-17. $\rho_s = 722\ \Omega/\square$ and $\rho_c \approx 8\ \Omega\mu\text{m}^2$

base thickness is T_b the extrinsic base thickness becomes

$$T_{b,extrinsic} = T_b \frac{R_{intrinsic}}{R_{extrinsic}} \tag{5.3}$$

Summarizing the results from 14 different layer structures and more than 25 different samples $\Delta T_b = T_b - T_{b,extrinsic} \approx 5.1$ nm for devices with exposed base–emitter junction. For devices with graded base–emitter junction partly or wholly remaining $\Delta T_b = 1.7$ nm . This can be explained by a combination of surface etching and surface depletion (equation 2.2). Solving Poisson's equation for the base, assuming a surface pinning potential of -0.2 eV the depletion width becomes 3.5-4 nm . The rest of the base might have etched or oxidized away, as selectivity for wet etches is only a relative term. Oxidation, development and oxide stripping in subsequent process steps can all remove base material. For GaAsSb bases the situation is complicated by etching of the base by common photoresist developers.

Thin bases are very sensitive to this effect and the devices with 25 nm base had 800-1000 $\Omega/\square$ sheet resistance in the extrinsic region, while the intrinsic region had 600-750 $\Omega/\square$.

Devices with a grade or a ledge protecting the base are less sensitive to this. DHBT-20 also had a 25 nm base but had the base–emitter grade left in place, with the base–contact annealed through the grade region, and had $\approx 730\ \Omega/\square$ extrinsic base sheet resistance with an intrinsic sheet resistance $\approx 725\ \Omega/\square$. The small increase in base sheet resistance in the extrinsic region can be explained by surface depletion extending through the low–doped grade.

Table 5.2. Metal resistivity measured after E–beam evaporation

Metal	Theory ρ_m $\mu\Omega$–cm	Measured ρ_m $\mu\Omega$–cm
Pd	10.8	13-21
Ti	42	81
Au	2.2-2.4	2.5-3.3

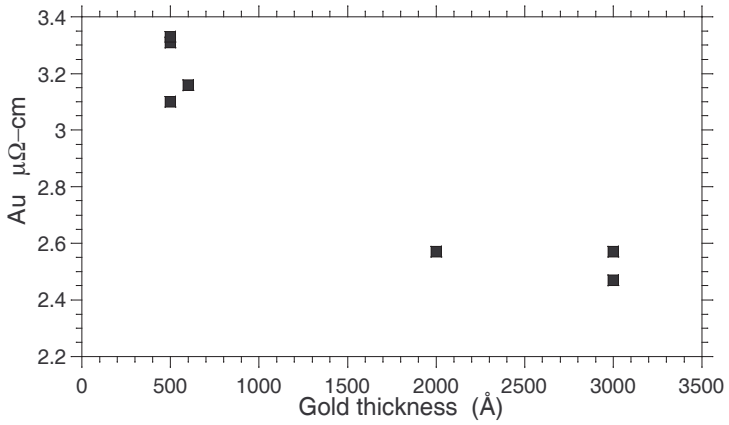

Figure 5.16. Extracted resistivity for gold thin films

5.3.2 Metal resistance

As described in section 4.2.4 the metal resistance of the contacts was characterized. The resistance is of concern for thin, small contacts as we discovered the resistance increase more rapidly than $1/T_{Au}$ would predict for gold thickness below 1000 Å.

The resistivity increases for very thin layers , as can be seen in figure 5.16. The data was gathered using 4–point probe measurements. It is not clear if it is the metal grain size that is responsible or the actual limitation by surface scattering in thin layers. This is of concern for the base contact. The base contact is realized as Pd(3 nm)/Ti(15 nm)/Pd(15 nm)/Au(60 nm) and the layers are thin enough for the resistance to increase. A typical measured base contact metal resistance is 0.5 $\Omega/\square$.If the resistance did not increase with decreasing metal thickness the base contact resistance would 0.35 $\Omega/\square$, but measured data show 0.45-0.55 $\Omega/\square$. This is not negligible for a very narrow base contact. The resistive drop from the interconnect to the other end of the base contact (see figure 1.2) can be estimated as $R_{bm} = \rho_m W_e/W_b$, with the following results: As can be seen in table 5.3 the base metal resistance is rather large compared to the base resistance R_{bb} (6-20 Ω). This is a fundamental limitation in a narrow mesa HBT with self–aligned base. Increasing the width of the base metal makes the base–collector capacitance larger,

Table 5.3. Resistance in base metal

Base metal width (μm)	R_{bm} Ω
0.25	7.2-8.8
0.5	3.6-4.4
1.0	1.8-2.2

while the metal thickness is limited due to the risk of shorting to the emitter.

5.4 S–parameter measurements

5.4.1 The measurement method

Device measurements are performed with a network analyzer in the 5 to 45 GHz measurement range, in the 70 to 110 GHz range and in the 140 to 220 GHz range using ground-signal-ground microwave probes. The design of the waveguides is described in section 3.6.

For analysis of device characteristics the 5 to 45 GHz measurement is most important, as it permits simplifying the extraction procedure by canceling higher order terms in frequency.

Calibration is done using the Line-Reflect-Line (LRL) method in which on-wafer microwave open, short and different through-lines are used to move the microwave plane adjacent to the transistor, where the base and collector are contacted. No de-embedding of probe capacitance is necessary in this measurement method. Several re-calibrations are necessary, and after every hour of measurement or so. DC bias conditions are recorded during the measurement.

5.4.2 The extraction method

Upon actual measurement of the device the small-signal gain H_{21} and Mason's unilateral gain U are directly plotted versus frequency to see how the bias conditions effect the device. From a -20 dB/decade extrapolation f_τ and f_{max} are determined, as in figure 5.9.

For determination of more parameters such as emitter resistance, base-collector capacitance, extraction is done to fit an equivalent circuit model (figure 5.17), [100].

The first step is to record, from S-parameters, Y_{21} at low frequencies. At low frequencies Y_{21} is can be expressed in an equivalent circuit model as:

$$Y_{21} = \left(R_{ex} R_{bb}/\beta + \frac{n k_b T}{q I_e} \right)^{-1} \tag{5.4}$$

By plotting $1/Y_{21}$ against the inverse current as in figure 5.18, R_{ex} can be approximately determined. The constant term is not strictly R_{ex}, as there is a R_{bb}/β term.

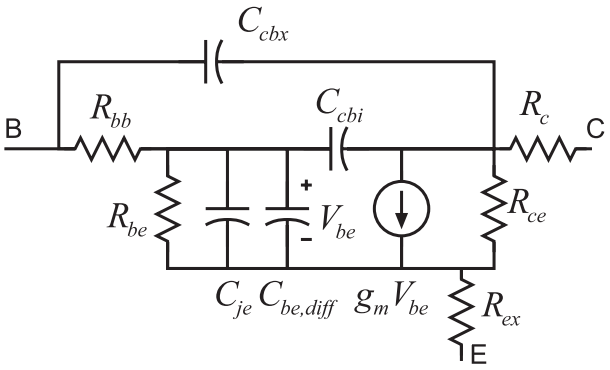

Figure 5.17. Equivalent circuit model

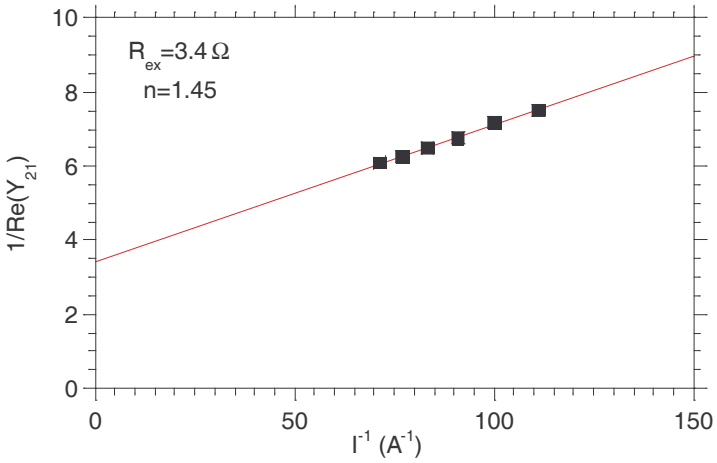

Figure 5.18. Extraction of R_{ex} and n from Y_{21}

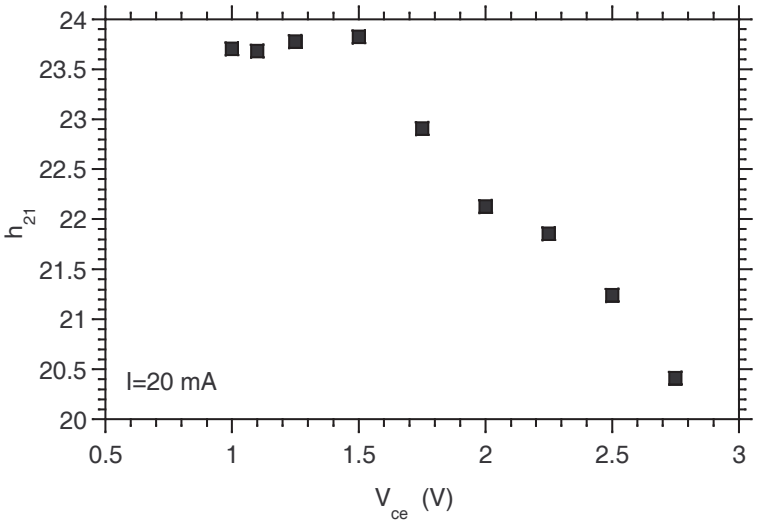

Figure 5.19. Extraction of H_{21} at 6 GHz for DHBT-17

These terms can also be determined and the R_{ex} is updated accordingly.

By plotting the small signal gain H_{21} versus frequency we can get β from the value at 5-6 GHz. Figure 5.19 also show how the gain is decreasing at higher bias. The current is 20 mA or 4.9 mA/μm^2, with $V_{ce} = 1.7$ V.

The base-collector capacitance can be determined from:

$$Y_{12} = \left(1/R_{cb} + \omega^2 C_{cbi}^2 R_{bb}\right) + j\omega \left(C_{cbx} + C_{cbi}\right) \tag{5.5}$$

In the limit of low frequency (ω^2 terms small) we get:

$$\Re(Y_{12}) \approx (1/R_{cb}) \tag{5.6}$$

$$\Im(Y_{12}) \approx \omega \left(C_{cbx} + C_{cbi}\right) \tag{5.7}$$

It will be shown later how the capacitance changes with bias, and also how the change in capacitance varies with device size.

The final part is to determine the transit time and the charging capacitive terms associated with it. From the expression for f_τ we get:

$$\frac{1}{2\pi f_\tau} = \tau_{bc} + R_{ex}\left(C_{cbx} + C_{cbi}\right) + \left(\frac{nk_bT}{qI_e}(C_{cb} + C_{je} + C_{poly})\right) \tag{5.8}$$

Plotting $1/2\pi f_\tau$ versus frequency the two terms can be determined. The measurement can be complicated by two things, current dependant gain will reduce $f\tau$ at low currents, and current dependant base–collector transit time.

Table 5.4. Breakdown of delay terms: DHBT-20. I_e=30 mA, and the ideality factor $n = 1.39$

Delay term	Delay (fs)	Relative delay %
Meas. f_τ	294	-
Meas. τ_{ec}	541	-
$R_{ex}C_{cb}$	61.5	11
$R_{ex}C_{poly}$	19.3	4
τ_{bc}	361	67
$(k_bT/qI_e)\,C_{je}$	59.8	11
$(k_bT/qI_e)\,C_{cb}$	26.5	5
$(k_bT/qI_e)\,C_{poly}$	8.3	2
Sum:	537	100
Ekv. f_τ	297	

5.5 Device results from DHBT-1 to 21

5.5.1 Extraction of delay terms

The different terms in (5.8) are broken down and analyzed. C_{cb} is available from the Y_{12} measurement, and by using a default value of $C_{poly} = 5$ fF the base emitter junction capacitance C_{je} can be calculated. As an example, the different delay terms for a HBT from DHBT-20 is shown.

An analysis as shown in table 5.4 is invaluable to provide information about the dominant delay terms. What is very important for improving device performance is to increase the current in order to reduce the different charging capacitances. One observation regarding the ideality factor needs to be done. Large area devices for DHBT-20 showed an ideality factor ~ 1.12, which is as expected for a graded base-emitter junction. But the small devices has $n \approx 1.39$. The small area HBT's showed poor passivation (figure 5.13) and the higher ideality factor is most likely related to surface leakage currents [12, 101].

5.5.2 Collector current spreading

Figure 5.20 show how an extraction of collector velocity can take place. The dotted line correspond to a collector saturation velocity of $\approx 4.2 \cdot 10^5$ m/s, and the solid line correspond to $\approx 3.0 \cdot 10^5$ m/s. The data is from DHBT-17 with a 215 nm thick collector. The extraction is based on the Kirk current equation (2.38), and assumes *no* current spreading. The collector velocity is obviously higher than bulk value of $\approx 3 \cdot 10^5$ m/s.

But, in the derivation current spreading was neglected, which is an oversimplification. According to equation 2.46 in section 2.5 current spreading does occur,

Table 5.5. Summary of device performance: the base

Device	T_b (nm)	N_a cm^{-3}	Gain	intrinsic ρ_s $\Omega/\square$	extrinsic ρ_s $\Omega/\square$	Contact ρ_c $\Omega-\mu m^2$
DHBT-1	40	$4 \cdot 10^{19}$				
DHBT-2	40	$4 \cdot 10^{19}$				
DHBT-3	40	$4 \cdot 10^{19}$	19			
DHBT-5	40	$5 \cdot 10^{19}$	7			
DHBT-6	40	$4 \cdot 10^{19}$	12	485		20
SHBT-1	30	$1.2 \cdot 10^{20}$	6	399	485	72
DHBT-9	40	$4 \cdot 10^{19}$	7	520	541	86
DHBT-10	30	$8 \cdot 10^{19}$		647	705	71
MHBT-10	40	$4 \cdot 10^{19}$	76	1020	1020	
DHBT-17	30	$8 \rightarrow 5 \cdot 10^{19}$	25	555	677	1
DHBT-18	30	$8 \cdot 10^{19}$	18	604	752	6
DHBT-19	25	$8 \rightarrow 5 \cdot 10^{19}$	27	745	798	30
DHBT-20	25	$8 \rightarrow 5 \cdot 10^{19}$	15	725	730	9
DHBT-21	25	$8 \rightarrow 5 \cdot 10^{19}$		693	842	33

Table 5.6. Summary of device performance: RF

Device	T_c nm	f_τ (GHz)	f_{max} GHz	at J_c $mA/\mu m^2$	V_{BCEO} V	Type
DHBT-1	300 nm	165	300	1	9	T.S
DHBT-2	200 nm	215	210	1	6	T.S
DHBT-3	176 nm	120	200		8.2	T.S
DHBT-5	300 nm	128	120	1.3		T.S
DHBT-6	146 nm	187	224	0.8		T.S
SHBT-1	200 nm					T.S
DHBT-9	150 nm	114	144	1.6		Mesa
DHBT-10	197 nm					Mesa
MHBT-1	200 nm	207	140	4.1	5.5	Mesa
DHBT-17	217 nm	282	400	4	7.5	Mesa
DHBT-18	217 nm	250	356	4.9		Mesa
DHBT-19	150 nm	298	190	4.4		Mesa
DHBT-20	150 nm	300	280	6.4		Mesa

Table 5.7. Summary of device performance: extracted from DC and RF measurements

Device	T_c nm	$N_{a,base}$ cm^{-3}	$\mu_{h,base}$ cm^2/Vs	$\mu_{e,base}$ cm^2/Vs	τ_{bc} (fs)	Type
DHBT-5	300 nm	$4 \cdot 10^{19}$			517	T.S/bandgap
DHBT-6	146 nm	$4 \cdot 10^{19}$	80.4	240	436	T.S/bandgap
SHBT-1	200 nm	$1.2 \cdot 10^{20}$	43.5	608		
DHBT-9	150 nm	$4 \cdot 10^{19}$	65	140		Mesa/bandgap
DHBT-10	197 nm	$8 \cdot 10^{19}$	40.2			bandgap
MHBT-1	200 nm	$4 \cdot 10^{19}$	38.2	1520	5.5	Mesa/bandgap
DHBT-17	217 nm	$8 \rightarrow 5 \cdot 10^{19}$	57.7	743	204	Mesa/doping
DHBT-18	217 nm	$8 \cdot 10^{19}$	43	810	331	Mesa/bandgap
DHBT-19	150 nm	$8 \rightarrow 5 \cdot 10^{19}$	49.9	557	274	Mesa/doping
DHBT-20	150 nm	$8 \rightarrow 5 \cdot 10^{19}$	56.5	309	230	Mesa/doping
DHBT-21	100 nm	$8 \rightarrow 5 \cdot 10^{19}$	65.0			Mesa/doping

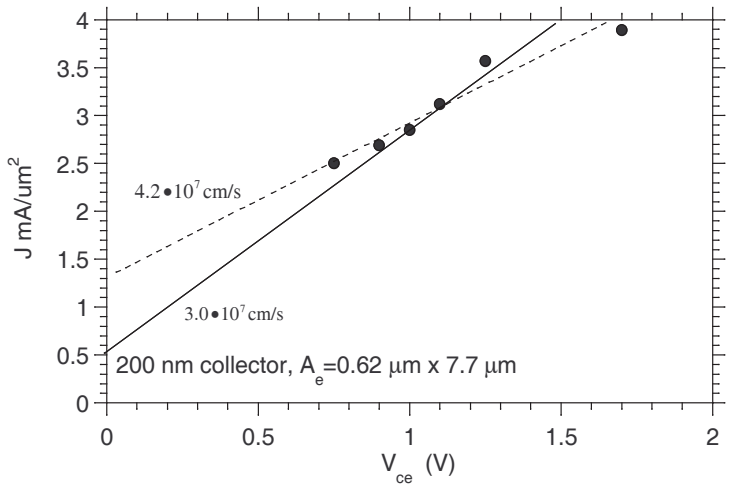

Figure 5.20. Kirk threshold for from 5 % drop-off in f_τ for DHBT-17

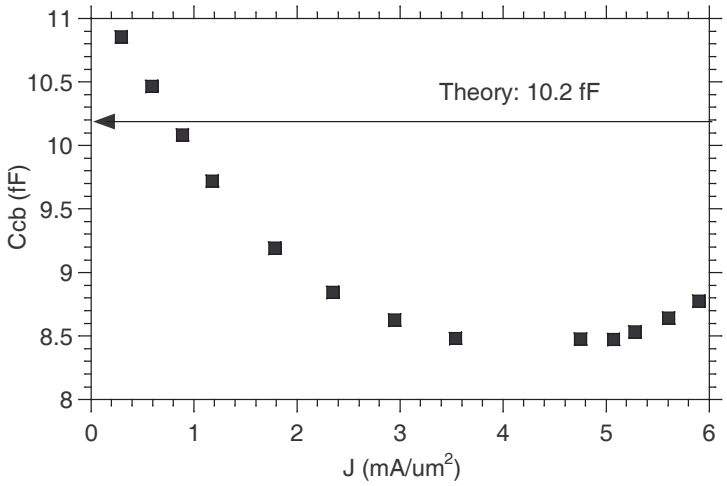

Figure 5.21. C_{cb} extracted from from DHBT-18. The line indicated the theoretically calculated C_{cb}

with the data in figure 2.23 as evidence. In equation 2.46 a current spreading factor Γ was defined, and the effective collector current density was expressed as:

$$J_{eff} = \frac{J_e}{\Gamma} \tag{5.9}$$

For emitter with a width near the lateral diffusion width L_d the correction is significant (figure 2.22). The importance for the velocity extraction is in the $J_e \sim \Gamma v_{sat}$ term. Current spreading makes Γ larger than one, and the extracted velocity needs to divided by Γ for obtaining the true collector saturation velocity.

5.5.3 Capacitance cancelation

At high collector current densities it has been observed that the collector capacitance C_{bc} deceases with current to a value much lower than predicted by geometry (figure 5.21) and doping ($C_{cb} = \epsilon A_{bc}/T_c$), and we observe it on every HBT, in different amounts. This was theoretically predicted by Camnitz and Moll [72], and first experimentally observed by Betser and Ritter [96]. In III-V semiconductors the electron velocity peaks at certain electric field (see e.g. [16]) and decrease rapidly thereafter. This is accredited to $\Gamma - L$ scattering (section 2.1). The effect of decreasing electron velocity with increased electric field is to increase the collector space charge and thereby screen the base from variations in applied collector voltage, and in consequence to reduce the capacitance. The capacitance modulation is predicted to lead a -40 dB/decade roll-off in H_{21} above $\sim 1/2\pi\tau_c$, the frequency corresponding to the collector transit time. For a 217 nm thick collector this corresponds to

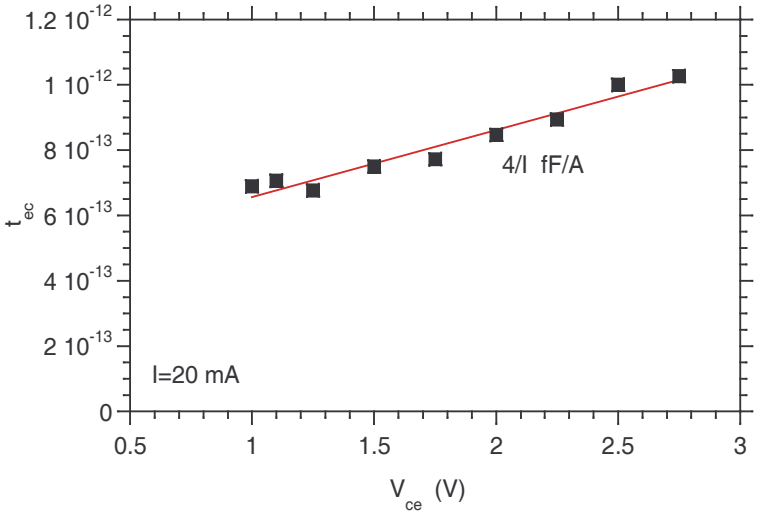

Figure 5.22. τ_{ec} extracted from from DHBT-17

440 GHz using equation 2.31 with $v_{sat} = 3 \cdot 10^5$ m/s. Betser and Ritter give the following expression for the capacitance reduction:

$$C_{cb} = \frac{\varepsilon_r A_e}{T_c} - \tau_c \left| \frac{dI_c}{dV_{bc}} \right|_{T_c} - I_c \left| \frac{d\tau_c}{dV_{bc}} \right|_{T_c} \tag{5.10}$$

The first part of this expression is the classical capacitor expression, as defined by the junction width , or if fully depleted T_c. The second term corresponds to the output resistance of the HBT, which, judging from our I-V curves, is a very small term. The extracted output resistance is on the order of some thousand ohms. The third term is the electron velocity modulation term, corresponding to the increase in collector transit time as the bias is increased, (figure 5.22)

To accommodate for current spreading the emitter junction area A_e should be replaced with the effective emitter area, A_e/Γ. Current spreads laterally during transport through the collector, flowing through a region of width $\sim (W_e + 2L_d)$ as described in section 2.5.6. The capacitance cancelation works in this region, strongly reducing the capacitance. Accordingly, the strongest modulation in capacitance would be for HBT's with comparably small extrinsic collector regions — the current should flow in as much of the collector as possible. Compelling evidence for this is shown in figure 5.28 where devices with narrow emitters and wide base-collector region show less variation in C_{cb} than devices with comparably wider emitters. Figure 5.23 shows the decrease in C_{cb} as the current density is increased.

A cancelation in the base-collector capacitance could also occur from modulation of the collector depletion region thickness. The injected electrons are equivalent in

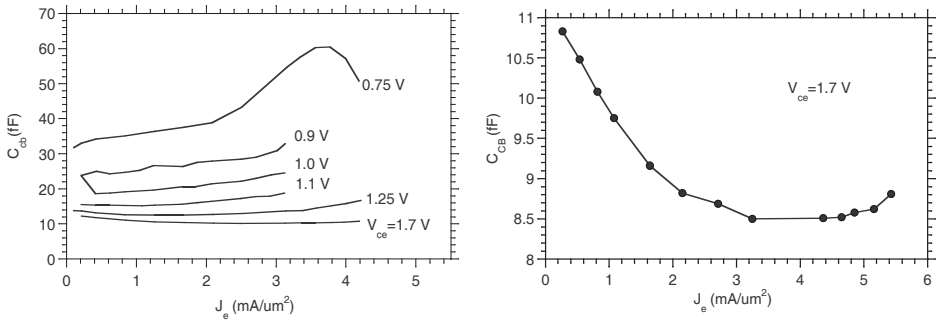

Figure 5.23. C_{cb} as a function of current density for DHBT-17 (left) and DHBT-18 (right)

effect as a p-doping and effectively reduces the collector doping level, and thereby increases the width of the depletion region.

$$x_d = \sqrt{\frac{2\varepsilon_r \left(V_{applied} + V_{bi}\right)}{qN_{c,eff}}} \tag{5.11}$$

But the collector is not uniformly doped, the subcollector doping level is three orders of magnitude higher than the collector doping level as well as the injected electron charge density, and once the depletion region reaches the subcollector realistic current densities will not change the depletion region width. What we would expect to observe is then strong capacitance modulation when the collector is not fully depleted and weak — none — when the applied bias in increased as to fully deplete the collector.

Figures 5.26 and 5.27 show a different situation. When the collector voltage bias is increased the capacitance modulation is larger, not smaller. Once full depletion has been obtained the capacitance reduction is very close to constant. The capacitance modulation will probably show a weak dependence on V_{cb} even in full depletion — the magnitude of modulation is determined by the $d\tau_c/dV_{cb}$ term, which corresponds to the exact art of electron velocity modulation by applied bias.

Can we determine this term? Yes, we can extract the dependence of base–collector transit time on applied bias from S-parameter measurements. Most properly we should use the extracted collector delay time τ_c. This has been done in figure 5.25, using τ_{bc} from two different bias points. The drawback of this method is that only two points were available.

An approximate method for which more data is available is to plot $\tau_{ec} = 1/2\pi f_\tau$ over applied bias. This assumes $d\tau_c/dV_{cb} \approx d\tau_{ce}/dV_{ce}$, which is true in a non–leaky HBT where the base transit time does not depend on V_{bc}. Figure 5.24 shows how f_τ changes with applied bias. Observe how for V_{ce} =1.7 V the measured f_τ is lower at the very lowest current levels. This indicates that a stronger applied bias leads to a larger base–collector transit time which is an indication of $\Gamma - L$ scattering.

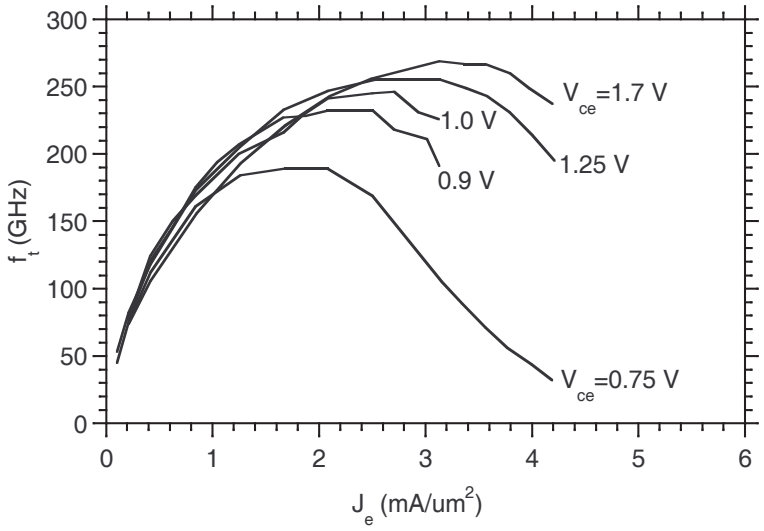

Figure 5.24. Variation of f_τ as a function of bias for DHBT-17

As the current density is increased the collector velocity appears to increase. Figure 5.22 shows how the emitter-to-collector delay time τ_{ec} increases with increasing V_{bc}. This is in full agreement with $\Gamma - L$ scattering. A change in shape of the conduction band near the base is expected at high current injection (figure 2.21) which should lead to less $\Gamma - L$ scattering in the region next to the base, which is the region most important to transit time.

From the data show in figure 5.22, DHBT-20 shows 4 fF/A of capacitance cancelation for *that* device geometry, with its amount of current spreading. Figure 5.26 shows the decrease in C_{cb} as the current density is increased for two different devices. One device has a 0.7 μm wide emitter contact (the top curves). Two curves are shown; one for V_{ce}=1.5 V and one for V_{ce}=1.7 V. The maximum sustainable current density is larger for V_{ce}=1.7 V than V_{ce}=1.5 V. The lower curve is for a device with 0.5 μm wide emitter contact. The overall capacitance reduction is lower, *and* the maximum sustainable current density is larger than for the 0.7 μm HBT, in agreement with the theory of current spreading.

Figure 5.27 shows the capacitance decrease for DHBT-20 as a function of bias and current density. Compared to DHBT-17 and DHBT-18 the maximum sustainable current density is large: the collector is thinner and the Kirk threshold is higher.

Figure 5.28 shows a compilation of the ratio of capacitance reduction as a function of emitter-to-base mesa ration, W_e/W_{bc}. The C_{cb} reduction is stronger at higher V_{cb}, which can be seen in figure 5.28 when data points are stacked vertically. The trend shows that for HBTs where the emitter covers much of the collector the total capacitance reduction is larger than for devices with large extrinsic base–

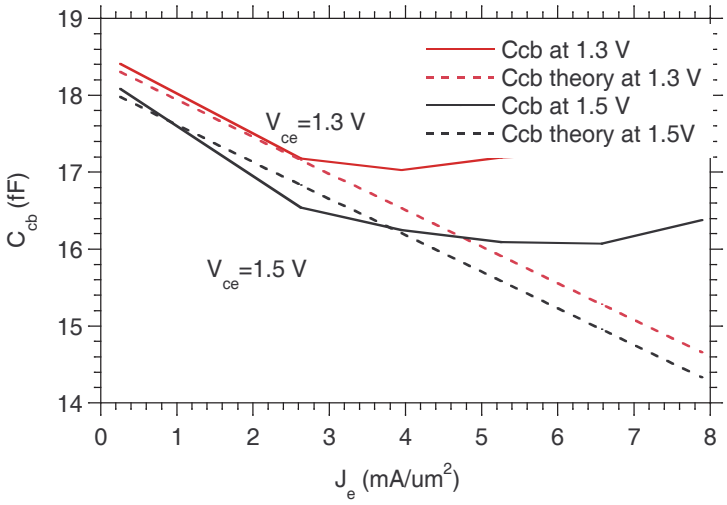

Figure 5.25. C_{cb} extracted and predicted for DHBT-20

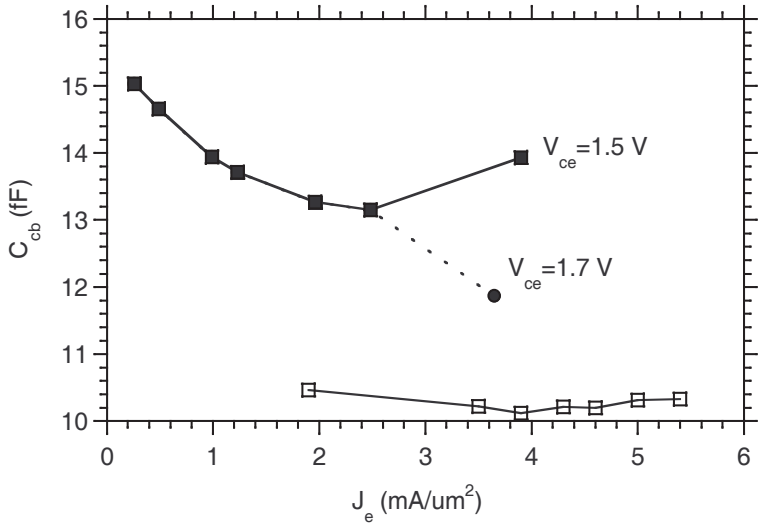

Figure 5.26. C_{cb} data from DHBT-17. The upper curves are for a 0.54 μm wide emitter, and the lower for a 0.34 μm wide emitter, with a larger extrinsic base–collector capacitance

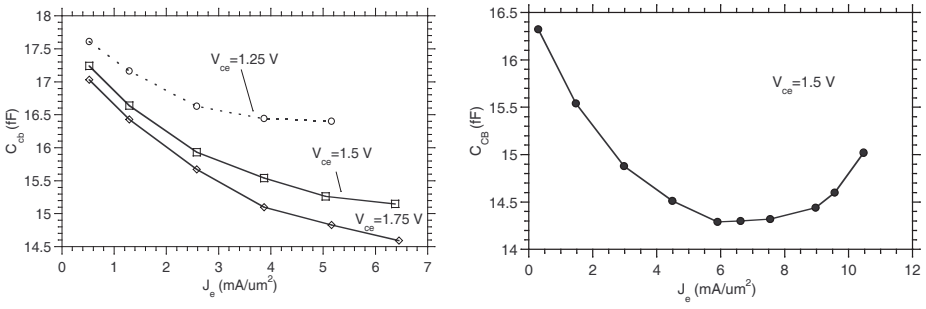

Figure 5.27. C_{cb} as a function of bias (left) and current density (right) for DHBT-20

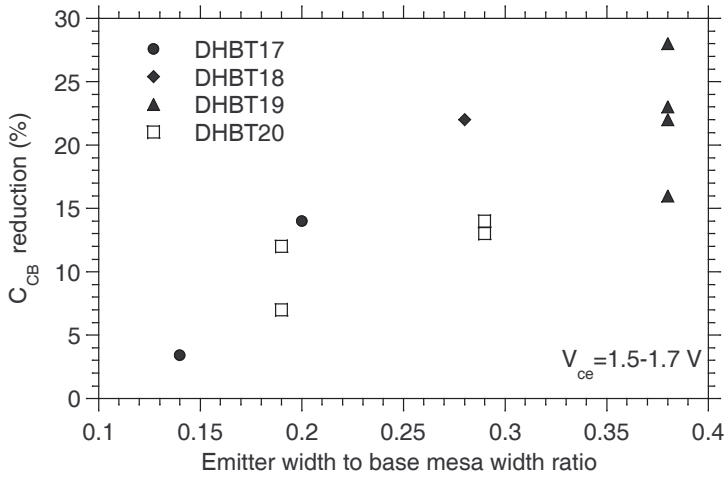

Figure 5.28. Ratio of capacitance reduction from several devices

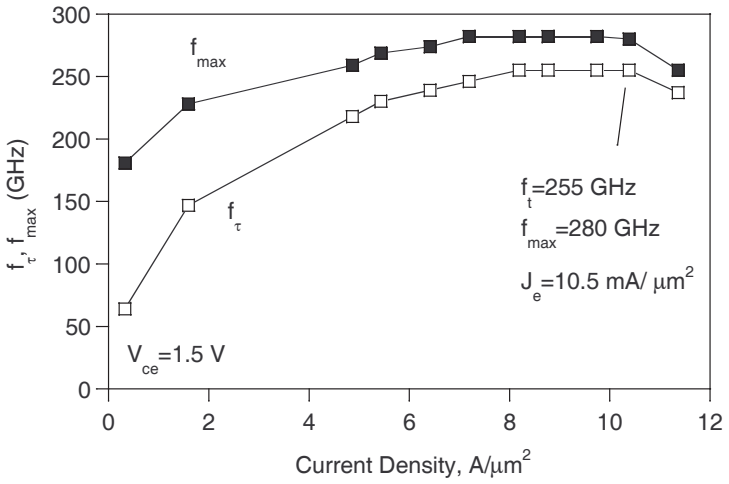

Figure 5.29. Variation of f_τ as a function of bias for DHBT-20

collector regions. If we extrapolate the trend in C_{cb} reduction with geometry we would reduce C_{cb} by ~ 65 % when the emitter and collector overlap fully. The total reduction can increase even more if the current could be increased. The data in figure 5.28 are from devices with different collector thickness and doping, and the devices thus have different J_e where C_{cb} starts to increase.

5.5.4 Maximum current density

Table 5.6 show at which current density the best f_τ was reached. As can be seen in figure 6.2 the trend is positive, and confirms that switching to mesa HBT from transfered substrate and working on removing thermal barriers has been successful. It also illustrates the good transport properties of our collector grade, which permits good device performance at these high current densities.

The maximum current density scales up with decreasing emitter width. There are two possible reasons: current spreading and better thermal heat sinking. A pulsed measurement would be necessary to disconcert which is the major effect. An indication that the thermal effects are not limiting is the fact that f_τ keeps increasing for higher power levels as seen to left in figure 5.31.

5.5.5 Extraction of material parameters

Using DC data hole and electron base mobility was extracted, and from extracted S–parameter data an estimate of the collector transit time τ_c could be made, figure 5.32. For the hole mobility in the base equation 5.12 was used, and for the

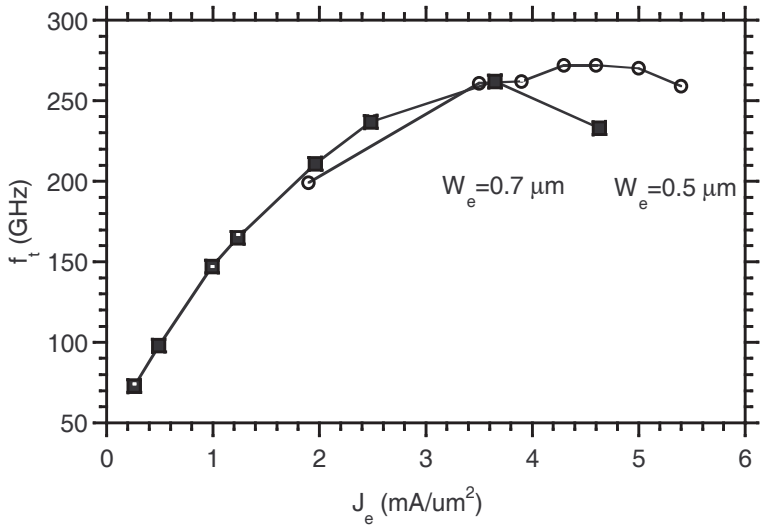

Figure 5.30. Variation of f_τ as a function of emitter width for DHBT-17

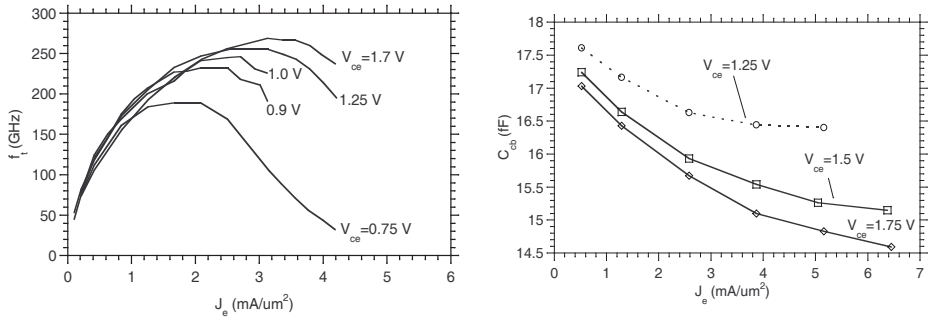

Figure 5.31. Trend in f_τ as a function of V_{ce} for DHBT-17(left) and current density for C_{cb} increase for DHBT-20 (right)

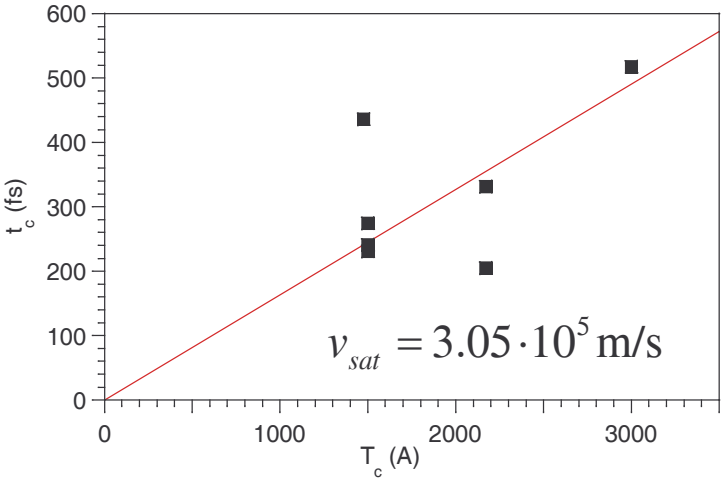

Figure 5.32. Collector velocity extracted from τ_c vs. collector thickness

electron minority mobility in the base equation 5.13 was used. For the extraction the expression for a constant collector velocity was used. Fitting the data with a line through origo a collector velocity of $v_c = 3.05 \cdot 10^5$ cm/s is extracted. A free fit gives a collector velocity of $v_c = 4.46 \cdot 10^5$ cm/s, both of which are within the realm of reported velocities [1, 5]

Another way to extract the collector velocity is to use the Kirk current threshold measurements. In the expression for the Kirk threshold, equation 2.38, there is a J_c/v_c dependence. Using this dependence a value of the collector velocity can be obtained by plotting the Kirk current threshold versus collector voltage bias. This extraction, done on data from DHBT-17, provides a value of $v_c = 4.2 \cdot 10^5$ cm/s, (figure 5.33). This appearantly high value of collector velocity corresponds to the velocity without taking current spreading into effect. The extracted velocity corresponds to $v_{c,\text{eff}} = v_c/\Gamma$, and the amount of current spreading Γ needs to be corrected for. From 2.45 we get $1/\Gamma = 1.15$ for this device geometry. With a basic saturation velocity of $v_c = 3.0 \cdot 10^5$ cm/s the apparent velocity $v_{c,\text{eff}}$ with current spreading becomes $v_{c,\text{eff}} = v_c/\Gamma = 3.0 \cdot 10^5 \times 1.15 = 3.62 \cdot 10^5$ cm/s. The two different methods for collector velocity extraction provides us with $v_{c1} = 3.05$ to $v_{c2} = 4.2 \cdot 10^5$ cm/s. If the collector velocity would be constant the value of the current spreading factor Γ would then the ratio of the two velocities, $v_{c2}/v_{c1} = 1.37$, which is higher than the calculated value of current spreading (1.15). It is fully expected that the collector velocity is higher than $3.0 \cdot 10^5$ cm/s for thin collectors, and the resulting velocity seems to be due to both current spreading and velocity increase; the velocity seems to have increased on the order of 16 %.

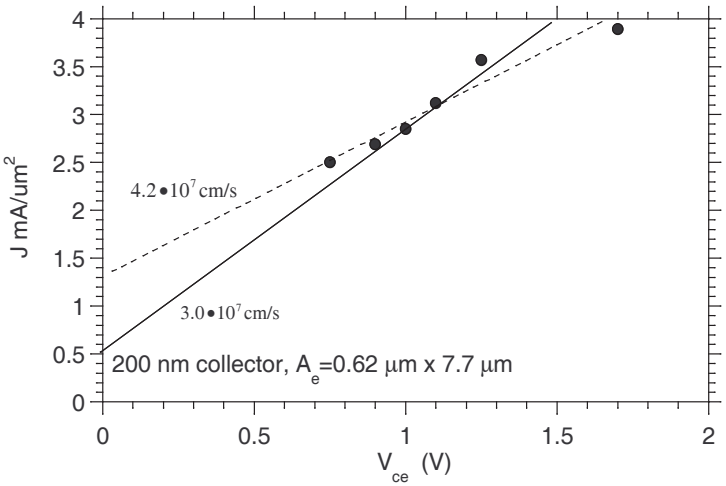

Figure 5.33. Collector velocity extracted the Kirk current condition

From the device data presented in table 5.5 the base hole mobility was calculated, (figure 5.34).

$$\mu_h = \frac{1}{qN_a\rho_s T_b} \tag{5.12}$$

The values for base thickness or doping level were not confirmed experimentally. The data is also shown in figure 5.35 as a function of doping.

If the HBT gain is dominated by recombination the gain is expressed by the relation:

$$\beta = \frac{2k_b T A \mu_{e,base}}{N_a^2 T_b^2} \tag{5.13}$$

Where $A = 4 \cdot 10^{-29}$ cm^6/s is the Auger recombination coefficient [47]. From this the electron minority mobility can be extracted, (figure 5.34), but this is only a rough estimate. The gain is effected by several issues such as base grading, base-collector current blocking and an abrupt or graded emitter. The hole majority and electron minority mobilities used to calculate base transit times is also shown.

It is interesting to compare the calculated base and collector transit times as they are of the same order of magnitude but always extracted as a whole, the base–collector transit time. To separate the base transit time from the collector would be beneficial but the author is unaware of a method for this. The base transit time in indirectly available from the current gain but requires at least two separate structures with different base thickness for determination. A 30 nm graded base has a base transit time $\approx 110 - 150$ fs, while the collector transit time is ≈ 280 fs for a 200 nm collector. For time being the base transit time is calculated and then subtracted away from the total base–collector delay to provide the collector delay.

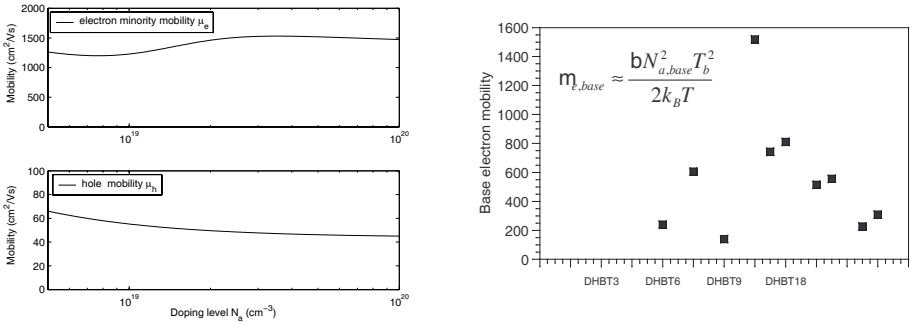

Figure 5.34. Calculated base hole and electron mobility (left) (from [15]) compared to extracted electron minority mobility (right)

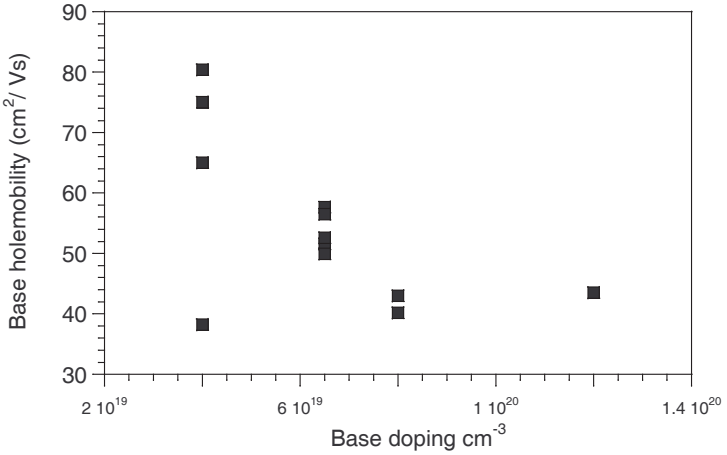

Figure 5.35. Extracted base hole mobility as a function of doping

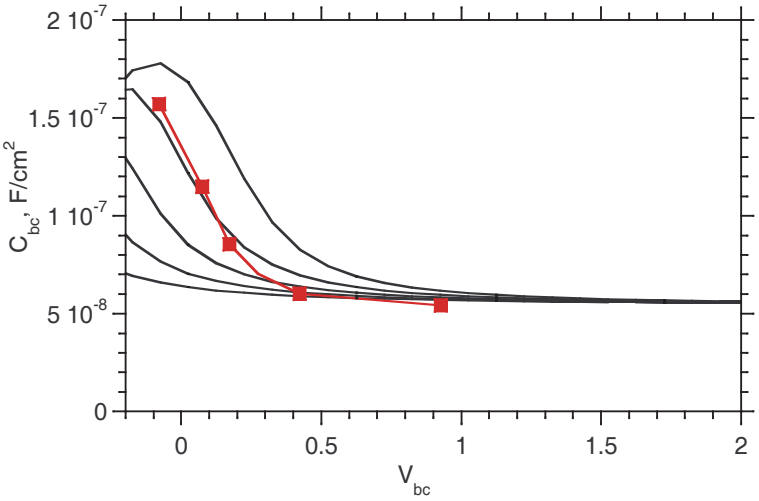

Figure 5.36. The measured base–collector capacitance compared to simulation results. The different curves correspond to $1 \ldots 5 \cdot 10^{18}$ cm^{-3}

.

5.5.6 Discussion on DHBT-17 and DHBT-18

When comparing DHBT-17 and DHBT-18 one observes that DHBT-17 has lower base sheet resistivity than DHBT-18, but also higher gain. This is unexpected since gain $\sim 1/N_a^2$, and for a device with the same base thickness and the same base quasi field one would expect the gain to be lower for the less resistive device. X–ray of the HBT indicate the InGaAs peak is lattice matched to InP for both wafers, and the In content is thus very close to 53 %. DHBT-17 show side lobes characteristic of good interfaces while DHBT-18 has a wider peak, which is natural since the base composition is graded. However, there are several InGaAs layer in the layer structure and the base layer might not be responsible for the side peaks. Thus, the unexpected gain and resistivity does not come from an unusual lattice composition, and while a thicker base could explain the resistivity it is not in agreement with the observed gain. The base properties DHBT-17 seems to be un usually good, and the growth would be worthwhile to repeat.

5.5.7 The δ-doping in DHBT-17

One cause for concern with DHBT-17 was the lack of complete collector depletion at $V_{cb} = 0$. The measured capacitance is shown in figure 5.36 compared to *BandProf* simulations with different amount of δ-doping. The match suggest the δ-doping was closer to $4 \cdot 10^{18}$ cm^{-3} than $3 \cdot 10^{18}$ cm^{-3}. The MBE machine used for growth of this wafer was overdue for calibration, and analysis showed the growth prior had received too much n-dopant.

Chapter 6

Conclusions

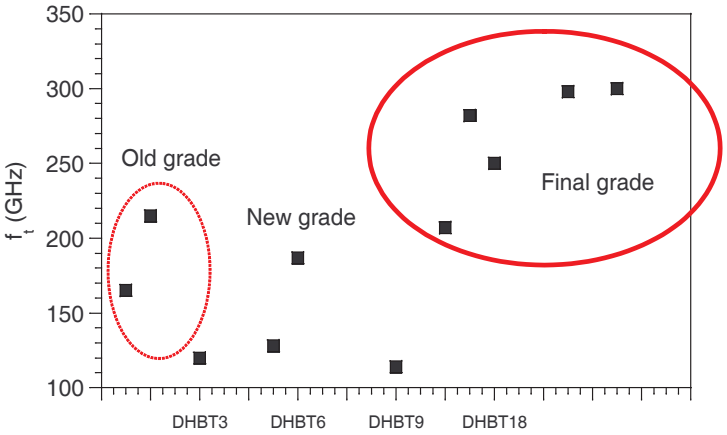

Figure 6.1. Evolution of f_τ of different DHBT designed by the author

6.1 Observations on the manufactured HBT

When looking back at the different HBT's designed and tortuously manufactured it is clear that MHBT-1 represent a substantial improvement. f_τ and $J_c(peak)$ increased suddenly to a new level from then on, (figure 6.1 and figure 6.2). The key reason for that is the new base–collector grade design with a 20 nm set–back layer that provides a higher electron velocity in the collector with no observable current blocking. It also allowed a drastically higher current density – which also improves f_τ by reduction of charging capacitance delay times. The increase in current density greatly benefits device performance, and it is due to a number of reasons apart from the new base–collector grade: small device dimensions improve heat flow, wide emitter contact bridge for heat removal in the emitter direction, and the use of InP instead of ternaries whenever possible in the collector and subcollector region.

The trend in f_{max} is not as clear (figure 6.3), but DHBT-17 showed as unsurpassed f_{max} for mesa HBTs. The base contact resistance to DHBT-17 and DHBT-18 was phenomenal, with some of best reported values of p-contact resistance, and the collector was thicker than for the latest HBT's, which all contributed to the record f_{max}. The latest iteration showed higher f_τ - as expected, but their thinner collectors made for a lower f_{max}. The contact resistance has also been difficult to repeat to just that level that was achieved for DHBT-17 and DHBT-18. Since the base contact resistance is very important for mesa HBT's, close attention has to be paid to all steps in the processing for the base contact. Factors like UV-ozone lamp intensity, RTA-annealer calibration and metal purity controls are some necessary actions. Poor passivation plagued the latest process runs, and it is clear it degraded the DC as well as the RF performance. Instead of an $f_\tau \leq 300$ GHz, and

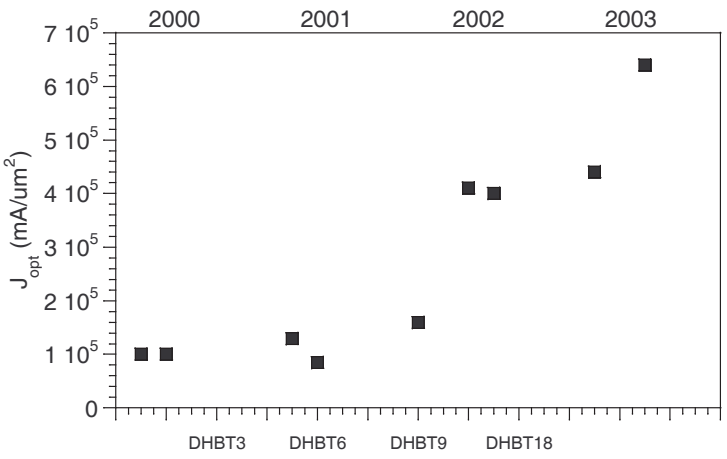

Figure 6.2. Evolution of J_c for the highest f_τ of different DHBT designed by the author

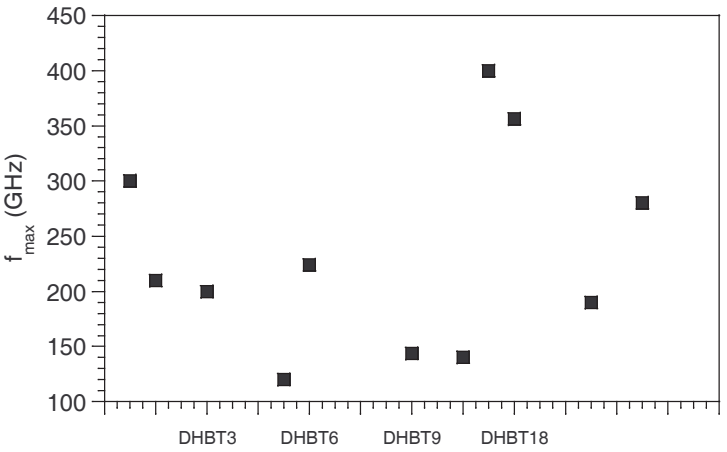

Figure 6.3. Evolution of f_{max} of different DHBT designed by the author

an $f_{max} \leq 280$ GHz we would have seen $f_\tau \leq 340$ GHz, and an $f_{max} \leq 310$ GHz, based on simulation results with a leakage resistor inserted between the base and the emitter.

Of the different obstacles for higher speed the emitter resistance is more and more becoming our main target. We observe $\rho_{ex} \approx 20 - 25\ \Omega - \mu\text{m}^2$. This constitutes an significant delay in f_τ through it's coupling with C_{je} and C_{bc}. Not only is it a problem as it is — it also hampers further emitter scaling. Direct replacement of the InGaAs cap with InAs is not possible due to the wet etching solution used for InP would also etch InAs.

One of the most positive characteristics with my devices is strong improvement in current density while maintaining good cut-off frequencies. The collector thickness plays an important role in keeping the Kirk threshold high — while HBT's from wafer DHBT-17 with a 217 nm collector occasionally could operate at current densities up to 10 mA/μm^2 the device performance at that point was terrible (f_τ=77 GHz). The thinner collector for DHBT-19 and DHBT-20 allows efficient high power operations. A HBT had an f_τ=237 GHz and an f_{max}=255 GHz at a current density of 11.5 mA/μm^2.

In fact, the current is now so high that the reliability of contact regions is a concern: the current densities reached are higher than expected or the mask set designed for. At the time of designing the mask set the best device performance was observed at a current density of $1-2\text{mA}\mu\text{m}^2$. The current levels are now much higher, especially for large or multi-finger devices. An example: a double-emitter HBT (two 0.7 μm by 8 μm wide emitters) operate at 70 mA with a 6 μm wide waveguide for the collector current, which translates to 12 mAμm^2 but our design guide-line is 5 mA/μm^2.

However, the goal of 300 GHz HBT's operating at 10 mA/μm^2 does not seem too distant.

6.2 Conclusions

A brief overview:

- Influence of the degenerate doping levels are visible in doping grades and transistor turn–on
- Keep the base-collector grade super lattice period to 1.5 nm, not more
- Maintain a setback layer on the order of 20 nm between the grade and the base
- Use Pd-based base ohmics
- Dope the top portion of the base very highly
- Use UV-ozone treatment for critical contacts

- Do not make the base too thin as surface depletion and etching reduces it, unless it is covered by a ledge or a grade.
- Minimize the amount of ternaries especially in the subcollector but generally in all regions where heat flows
- Design for high current densities and pay attention to the collector doping level
- Work on reducing the emitter resistance
- Try to shrink the emitter to $\sim 0.1 - 0.3\ \mu m$ to aquire higher allowable current density for "free"
- If by some technique the size of the base–collector junction can be chosen freely, allow it to be one diffusion width on each side for the sake of current diffusion
- In order to maximize the capacitance reduction at high currents try to reduce the size of the base–collector junction

The influence of the degenerate doping levels are visible in the transistor turn–on bias. The highly doped base increases the built–in voltage in the emitter–base junction. The device turns on for a highly doped device is noticeably higher (section 2.2.2). The effect of the degenerate base doping is indirectly visible in the good gain and RF–characteristics achieved in those devices.

After a number of iterations and some simple quantum mechanical calculations then conclusion has been reached to keep the base-collector grade super lattice period to 1.5 nm, not more. The super lattice period cannot be much less since the interatomic distance is 0.25 nm, and only a limited number of atomic layers can be packed into a grade layer.

It has been found that to maintain a setback layer on the order of 20 nm between the grade and the base is important for the collector electron transport. The setback layer keeps the electrons energized as they enter the lattice, decreasing the risk of possible trapping. Further, as the current density increases and the band near the grade starts to bend upwards, the setback layer helps keep the band beneath the base conduction band level. For very thick collectors the thickness of the setback layer should be increased to maintain the same potential drop before the grade. An alternative method is to increase the δ-doping level, but that can lead to incomplete depletion of the collector at zero bias.

The use of Pd-based base ohmics has been instrumental in keeping the base resistance down. The measured contact resistances are orders of magnitude better than most other.

The Pd-ohmics certainly are most important, but it is also critical to dope the top portion of the base very heavily. An advantage of base grading schemes is that they permit an increase in base doping level without hurting the gain of the HBT too much.

UV–ozone treatment has been found to be critical for the contact resistance. It seems as if the UV–ozone treatment is very gentle in terms of introducing defects compared to the plasma etchers otherwise available in the UCSB clean room.

The latest HBT's had a 25 nm thick base and it was found that the extrinsic base sheet resistance was noticeably higher than the intrinsic base sheet resistance. An exposed base surface should not be too thin to avoid this, and 30 nm for now seems like a minimum thickness. If the lateral extent of the exposed base surface is very small the magnitude of this problem decreases given that the base contact resistance is very good.

The thermal resistance of ternaries and quarternaries is much higher than the thermal resistance of InP. The amount of ternaries should be minimized, especially in the subcollector but generally in all regions where heat flows. The thermal resistance of the passivation layers will become more important with shrinking device size. The current passivation layers — polyimid, BCB and SiN — all have high thermal resistance.

The design has to be made for high current densities and high currents. With a the current density of 10 $\mathrm{mA}/\mu\mathrm{m}^2$ attention has to paid not only to the collector doping level but also to the current density in interconnects and contacts.

The evidence for current spreading, and the increase in maximum current density it brings along should be exploited. An narrow emitter with a width comparable to the extent of the current spreading will benefit the most. The emitter resistance will scale up if the width is decreased, and this is a most important problem.

This also means that in order to provide room for the current spreading the size of the base–collector junction should be on the order of one diffusion width on each side.

In order to maximize the capacitance reduction at high currents try to reduce the size of the base–collector junction, but a reduction below the limit discussed above will mean a reduction in the maximum possible current density.

6.3 Current problems

A summary of current problems:

- The emitter resistance must be reduced
- Poor passivation of the emitter–base junction degraded several recent devices
- The stepper had for a long time random errors due an upgrade which reduced yield
- The likelihood of getting good or decent results from a wafer is too low due to processing problems

The emitter resistance provides a major obstacle to reduction of the transistor size, and alone provides a substantial delay term for transistor speed. Currently

the delay term associated with the emitter resistance ($R_{ex}C_{cb}$ and $R_{ex}C_{layout}$) constitute 6-15 % of the total delay terms in f_τ, while the delay terms associated with junction discharging times ($\sim C_{je}/I_e$, C_{layout}/I_e and C_{cb}/I_e) constitute 5-20 % of the total delay terms, less so for HBT's operating at a higher current density. With shrinking emitter size and increasing current density the relative influence of the emitter resistance delay terms will increase. The collector also provide a parasitic resistance with an associated delay term but do not increase with scaling as R_{ex} does, in fact, with a narrower emitter, the collector resistance will decrease (equation 3.5). The sheet resistance of the subcollector can be decreased by making the highly doped InP subcollector layer thicker but that also makes the transistor more non–planar with added processing difficulty.

The poor passivation of the emitter–base junction that degraded several recent devices is a major nuisance and degraded the device performance, including the RF performance. It is possible that the simple maneuver of replacing the polyimid bottle will improve the passivation, as well as abstaining from the use of dangerous photo resist strippers such as PRX-127, but it is also possible that thinning the low–doped emitter part provides a lower leakage resistance path on the emitter surface.

The stepper problems seem to have been fixed with an upgrade of the upgrade, that hopefully should solve the problems of bad alignment. Results so far (March 2003) indicate that is so, and that will contribute to the device yield problem.

Poor device yield will probably always plague a university clean room facility working with aggressively scaled and tuned device processing. I guess it is in the nature of the beast. It would however have been very useful to tune the device design as there are a number of open issues.

6.4 Outlook

6.4.1 The physics of the base

In the base there a number of questions remaining: What is the exact influence of the base grade, especially as the base thickness decreases? Experiments would also be useful to extract the relative influence in base transit time, and gain, from abrupt emitter–base junctions and from base grades. Does the grade matter for thin bases with abrupt base–emitters? Finding a linear dependence of gain on base thickness is not enough because the base exit term τ_{exit} also have such a linear dependence. In growth experiments our gain was increased after we introduced doping graded bases, but the growth conditions were optimized on the same time as well so we are lacking those data points!

Why is the gain lower for carbon doped devices? Do we have carbon double bonds with a huge electron scattering cross section?

What is the influence of strain on the line-up for InP-InGaAs? Is it more useful to increase the In content to compensate for the small size of the carbon atoms or is preferable to increase the Ga content since carbon doping in InGaAs actually is carbon doping of dilute GaAs?

6.4.2 The physics of the collector

Many interesting questions open up in the collector related to the electron transport and field dynamics: The energy of an electron affects how fast it will travel, but if it has too much energy it will slow down. All electrons in the collector effect the electric field in the collector but the slow ones will reside longer and effect it more. Up to a certain limit the mobility of the electrons seem to increase due to the way the change the electric field. Electrons will also diffuse laterally and the amount of lateral spread depends seems to be dynamic. When the electron right beneath the emitter start to experience current blocking they will move more laterally , reducing the current blocking. How does all this work out? Can $\Gamma - L$ scattering be removed in very thin collectors if the applied bias is small enough so that electrons can never achieve the necessary energy? How much of the energy does an electron retain after traversing the base?

Apart of the above mentioned I would like to investigate the best design of the emitter and base. What is the appropriate emitter doping, high for increased gain or low for decreased surface leakage and indirectly higher gain?

6.4.3 The coming devices

In the short term I hope and expect to see individual devices as well as small scale circuits such as ADC and modulator drivers using the mesa technology. The device performance is such that it is only to put them together into circuits and achieve record-breaking results. That is unfortunately much more difficult than it sounds. The mesa DHBT's are also very suitable for power devices with their strong current handling capability at high frequencies and I have high hopes for the on-going design. We will now apply the experience gained in this work to design new HBT's, and by the time of the defense one results might be in. Specifically I want to address the emitter resistance problem by increasing the indium content in the emitter cap layer. Unfortunately we cannot switch completely to InAs in our current process because the InP wet etch would remove the emitters, unless a sidewall protection layer is used ... We will investigate InGaAs, with a small proportion of Ga to provide etch selectivity, as an emitter cap material. I also hope to have the opportunity to grow a wafer with a grade composed of both bandgap grading and doping grading. Simulation results indicate very good performance (table 3.3, last entry) and a trade-off with base thickness seems possible.

On a longer term I am a firm believer in regrown emitters. In order to achieve a narrow base–emitter junction without suffering completely in emitter resistance the top part must be regrown and flare out into a wider region where contacting

can take place. I also belive that the future of InP in the commercial marketplace is dependant on mimicking the advances in process technology and scaling that the Si industry has shown, especially the other HBTs, SiGe. I believe that very narrow emitters and base mesa offers certain advantages because of current spreading and capacitance calculations. The thermal budget is also more forgiving for a small device. Good passivation will be increasingly important as the device dimensions are scaled down and we need to have better control over the surface, especially in the emitter-region.

Santa Barbara April 8, 2003

Mattias Dahlström

Appendix A

Summary of device structures

Table A.1. Summary of HBT structures

Device	In %	T_b nm	N_a cm^{-3}	T_c nm	T_{setback} nm	T_{grade} nm
DHBT-1	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	300	0	48
DHBT-2	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	200	0	48
DHBT-3	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	176	0	10
DHBT-4	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	176	0	20
DHBT-5	$44 \rightarrow 53$	40	$5 \cdot 10^{19}$	300	0	10
DHBT-6	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	148	60	10
SHBT-1	53	30	$1.2 \cdot 10^{20}$	200	-	-
DHBT-9	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	150	70	10
DHBT-10	$44 \rightarrow 53$	30	$5 \cdot 10^{19}$	200	10	20
MHBT-10	$44 \rightarrow 53$	40	$4 \cdot 10^{19}$	200	10	24
DHBT-17	53	30	$8 \rightarrow 5 \cdot 10^{19}$	217	20	24
DHBT-18	$44 \rightarrow 53$	30	$8 \cdot 10^{19}$	217	20	24
DHBT-19	53	25	$8 \rightarrow 5 \cdot 10^{19}$	150	20	24
DHBT-20	53	25	$8 \rightarrow 5 \cdot 10^{19}$	150	20	24
DHBT-21	53	25	$8 \rightarrow 5 \cdot 10^{19}$	100	20	24

Table A.2. Previous DHBT layer structure with 300 nm collector (DHBT 2)

Material	Doping(cm^{-3})	Thickness(nm)
n-InGaAs	$1 \cdot 10^{19}$	100
n-InGaAlAs	$1 \cdot 10^{19}$	20
n-InP	$1 \cdot 10^{19}$	30
n-InP	$8 \cdot 10^{17}$	70
n-InGaAlAs	$8 \cdot 10^{17}$	23.3
p-InGaAlAs	$2 \cdot 10^{18}$	6.6
p-InGaAs	$4 \cdot 10^{19}$	40
n-InAlGaAs	$1 \cdot 10^{16}$	48
n-InP	$2 \cdot 10^{18}$	2
n-InP	$1 \cdot 10^{16}$	250
n-InGaAs	$1 \cdot 10^{19}$	75
InGaAs	UID	150
InAlAs	UID	250

Table A.3. Graded doping layer structure with 215 nm collector (DHBT-17)

Material	Doping (cm^{-3})	Thickness (nm)
n-InGaAs	$3 \cdot 10^{19}$	40
n-InP	$3 \cdot 10^{19}$	80
n-InP	$8 \cdot 10^{17}$	10
n-InP	$5 \cdot 10^{17}$	30
p-InGaAs	$8 \rightarrow 5 \cdot 10^{19}$	30
n-InGaAs	$2 \cdot 10^{16}$	20
n-InAlGaAs	$2 \cdot 10^{16}$	24
n-InP	$3 \cdot 10^{18}$	3
n-InP	$2 \cdot 10^{16}$	170
n-InP	$1.5 \cdot 10^{19}$	50
n-InGaAs	$3 \cdot 10^{19}$	50
n-InP	$3 \cdot 10^{19}$	200
SI-InP	UID	

Table A.4. Graded bandgap layer structure with 215 nm collector (DHBT-18)

Material	Doping (cm^{-3})	Thickness (nm)
n-InGaAs	$3 \cdot 10^{19}$	40
n-InP	$3 \cdot 10^{19}$	80
n-InP	$8 \cdot 10^{17}$	10
n-InP	$5 \cdot 10^{17}$	30
p-InGaAs (graded)	$6 \cdot 10^{19}$	30
n-InGaAs	$2 \cdot 10^{16}$	20
n-InAlGaAs	$2 \cdot 10^{16}$	24
n-InP	$3 \cdot 10^{18}$	3
n-InP	$2 \cdot 10^{16}$	170
n-InP	$1.5 \cdot 10^{19}$	50
n-InGaAs	$3 \cdot 10^{19}$	50
n-InP	$3 \cdot 10^{19}$	200
SI-InP	UID	

Table A.5. Graded doping layer structure with 150 nm collector (DHBT-19)

Material	Doping (cm^{-3})	Thickness (nm)
n-InGaAs	$3 \cdot 10^{19}$	40
n-InP	$3 \cdot 10^{19}$	80
n-InP	$8 \cdot 10^{17}$	10
n-InP	$3 \cdot 10^{17}$	30
p-InGaAs	$8 - 5 \cdot 10^{19}$	25
n-InGaAs	$3 \cdot 10^{16}$	20
n-InAlGaAs	$3 \cdot 10^{16}$	24
n-InP	$3 \cdot 10^{18}$	3
n-InP	$3 \cdot 10^{16}$	100
n-InP	$1.5 \cdot 10^{19}$	50
n-InGaAs	$2 \cdot 10^{19}$	25
n-InP	$2 \cdot 10^{19}$	300
SI-InP	UID	

Table A.6. Graded doping and graded emitter–base layer structure with 150 nm collector(DHBT-20)

Material	Doping (cm^{-3})	Thickness (nm)
n-InGaAs	$3 \cdot 10^{19}$	40
n-InP	$3 \cdot 10^{19}$	80
n-InP	$8 \cdot 10^{17}$	10
n-InP	$5 \cdot 10^{17}$	20
n-InAlGaAs	$5 \cdot 10^{17}$	16
p-InGaAs	$8 - 5 \cdot 10^{19}$	25
n-InGaAs	$3 \cdot 10^{16}$	20
n-InAlGaAs	$3 \cdot 10^{16}$	24
n-InP	$3 \cdot 10^{18}$	3
n-InP	$3 \cdot 10^{16}$	100
n-InP	$1.5 \cdot 10^{19}$	5
n-InGaAs	$2 \cdot 10^{19}$	25
n-InP	$2 \cdot 10^{19}$	300
SI-InP	UID	

Appendix B

Theory of carbon doping of InGaAs

B.1 Theory of carbon doping of InGaAs

Carbon as a dopant has very low diffusivity compared to Zn and Be, and is also non-toxic. Carbon doping of GaAs is now an established technique, and efforts are now being done to dope InGaAs (lattice matched to InP) with carbon [108,111,112,117]. This short paper describes how doping with carbon works, as of current knowledge.

B.1.1 Fundamentals

Carbon is a group IV element and can thus be at either a group III position (In or Ga) or at a group V position (As). If carbon is at a group III position it will act as a donor (n-dopant) and if it is at a group V position it will act as an acceptor (p-dopant) [109]. Only carbon at a group V position will act as a p-dopant, which is the desired use here, and it is important to minimize the amount not in V-position. In GaAs carbon will prefer to be at the group V position but in InAs at the group III position. The reason for this is the difference in bonding energy. The Ga-C bond is stronger than the As-C bond, which is stronger than the In-C bond. The crystal lattice efforts to minimize the free energy will result in the following preferred structure: Ga-C-Ga-As-Ga-C-Ga etc. in GaAs and In-As-C-As-In-As-C-As in InAs and with a similar result in InP (figure B.1). In fact, carbon doping of InGaAs can be regarded as carbon doping of GaAs in a InAs lattice [115]. This can be observed in experiments with different Ga/In fraction - the more Ga the higher p-dope level. This makes it probable that the high doping levels in GaAs of more than 10^{20} cm^{-3} will not quite be repeated in InGaAs. Shirikashi *et al.* [114] reported a hole concentration of $1.5 \cdot 10^{21}$ cm^{-3} in GaAs HBT's using MOMBE grown at 390-490 °C with elemental metal sources and TBP/TMG. However, not all the carbon successfully incorporated at a group V site acts as an acceptor, and the reasons for that is hydrogen passivation and, to a lesser extent, interstitial carbon [116–118], figure B.2.

B.1.2 Hydrogen passivation

Hydrogen passivation of dopants is a well-known phenomenon in several materials and in C-doped InGaAs hydrogen can passivate more than 50 % of the carbon. The hydrogen is incorporated during the epitaxy from H2, AsH3 or metal-organics. In GaAs the metal-organics have been used as the carbon source, and in that case hydrogen can come directly from the carbohydrate. For InGaAs CC4 or CBr4 are used instead so that is not the source. The hydrogen-carbon bond is cracked at temperature above 500 °C and the carbon can be activated by a post-growth anneal in an inert atmosphere [108,110,111,119]. One complication for heterostructures is that an InP layer on top of the p-doped InGaAs [108,116] hinders the hydrogen out-diffusion. The reason for this seems to be that the hydrogen diffuses as H+ and the built-in electrical field between the InP and the InGaAs retard the hydrogen. One way of circumventing this is to increase the temperature to more than 600 °C, which

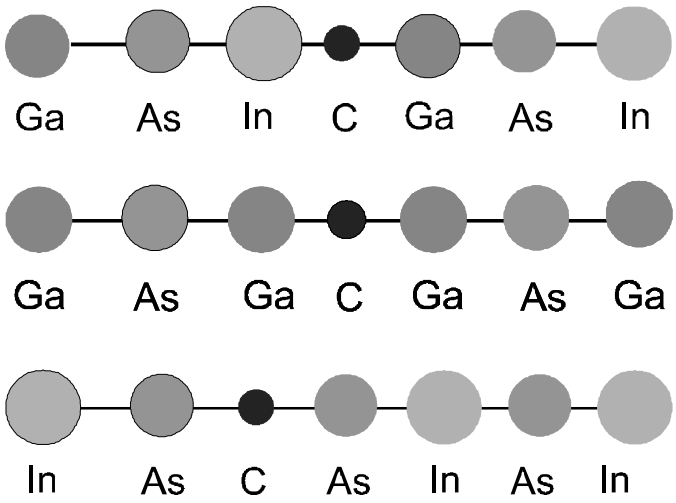

Figure B.1. The position of carbon in GaAs and InAs.

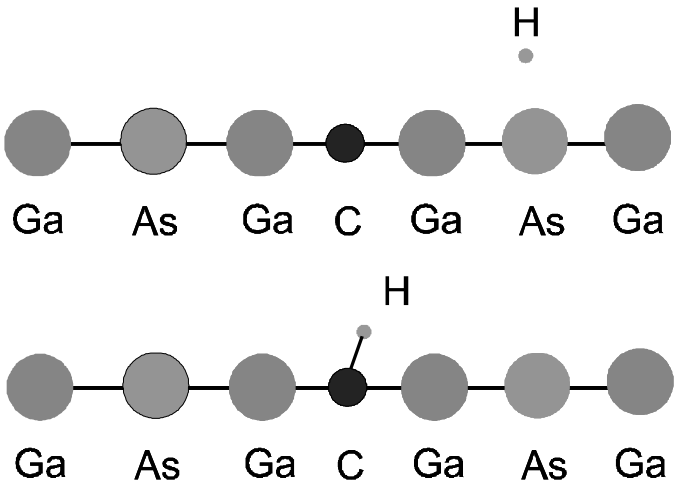

Figure B.2. Hydrogen passivates the carbon

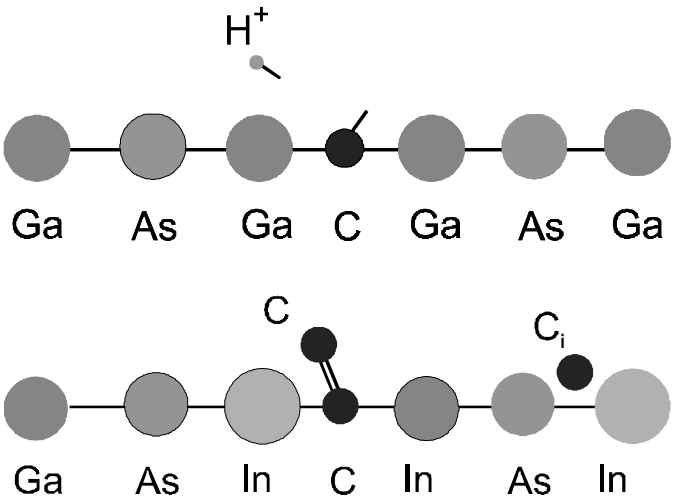

Figure B.3. The carbon is reactivated by annealing (top). Double carbon bonds lowers the mobility (bottom).

gives the protons enough kinetic energy to go through the InP layer. Another way is to anneal before the InP layer is grown, but some hydrogen will diffuse back into the InGaAs during the InP growth and in the case of HBT's, this anneal means a growth interruption at the critical emitter-base interface [113]. Very hypothetically it may be possible to tune the built-in electric field during the anneal by using an applied voltage or photonically induced carriers. A method successful in GaN is to use an e-beam to crack the hydrogen-dopant bond - this was tried here but did not work.

B.1.3 Interstitial carbon

While annealing at temperatures higher than 600 °C result in close to 100 % activation of carbon sitting at the correct group V position it has been observed that an increasing number of carbon atoms are in an interstitial position, resulting in lower than expected doping levels and also reduced mobility. Growing an InP layer on top of the p-InGaAs enhances this process

Figure B.3 Carbon interstitials or carbon-carbon bonds formed at high C-concentration and temperature. The carbon is believed to occupy either an interstitial position or to form a double bond with another carbon atom in the (100) direction [109] as shown in figure B.3. This process is dependent of accessible group V material to kick-out the carbon, and that is the reason it is enhanced by an InP cap layer. The underlying reason for all this is the lattice contraction that occurs when a smaller carbon atom occupies an As-site, as experimentally confirmed by Watanabe *et al.* [110]. At carbon concentrations of more than 10^{20} cm^{-3} the stress

in the lattice built up by this is so large that group V material - if available -will kick out the carbon if the temperature is high enough, relieving some of the stress.

The effective doping level available in GaAs or InGaAs will not be much higher than $1.2 \cdot 10^{21}$ cm^{-3} unless some way of solving the stress built-up is found. If not, C-doped InGaAs with doping levels close to that will be sensitive to temperatures of more than 600 °C when covered by InP. This is of great concern for processing. Unless the material is grown in very hydrogen poor conditions -which MOCVD is not - a post growth anneal in an inert atmosphere is necessary. The temperature for this is 450-500 °C if covered by InGaAs and 550-650 °C if covered by InP. It might be necessary to anneal after dry etching or PECVD because of possible in-diffusion of hydrogen.

Appendix C

Process Flow

Align 3" IQE material with emitters parallell to the major flat of the wafer
Align 2" UCSB material with emitters perpendicular to the major flat of the wafer

A.1 **Emitter Contacts (Mask Layer 1)**

A Solvent Cleaning

(a) Check the resistivity of the D.I. water. It should be > 17 MΩ.
(b) Solvent clean with acetone, methanol and propanol.
(c) DI rinse
(d) Blow dry with N_2.
(e) Dehydration bake, 120°C, 30 min in petri dish without cover.

B Photoresist Application and Exposure

(a) Cool down in HMDS chamber 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply SPR 955-0.9 with syringe and filter to cover wafer.
(d) Spin at 2.5 krpm for 30 sec.
(e) Soft Bake, 90°C, 1 min. on hot plate.
(f) Apply CEM with syringe and filter to cover wafer.
(g) Wait 1 min.
(h) Spin at 4 krpm for 30 sec.
(i) Expose for 2.15 seconds, focus around 2.

C Development

(a) Post Bake, 111°C, 1 min.
(b) Rinse in running D.I., 30 sec.

(c) Develop in MF-701 for 2 min, 10 sec.

(d) Rinse in running D.I., 1 min.

D UV Ozone Descum of Photoresist

(a) Run UV-ozone chamber for at least 20 minutes to warm up the lamp.

(b) Run with wafers in petri-dish for 10 minutes.

(c) Resist adhesion bake hotplate 111°C 1 min.

E Evaporation

(a) Mix a dilute solution of HCl : H_2O :: 1 : 10.

(b) Dip in dilute HCl for 15 sec.

(c) Rinse in DI for 3 min.

(d) Blow dry with N_2.

(e) Load ebeam 4. Inspect Ti, Pd and Au crucibles.

(f) Pump down to below 1 x 10-6 torr.

(g) Deposit material:

Material	Thickness(Å)	Dep. Rate (Å/sec)
Ti	200	1-2
Pt	400	1-2
Au	8000	2-7

F Liftoff –>> DO NOT LET ACE OR 1165 DRY ON WAFER! <<–

(a) Soak wafer in 1165 beaker until metal comes loose.

(b) If the liftoff is stubborn, leave the wafer soaking overnight. Seal the top of the beaker with foil. Leave name and number.

(c) 1165 squirt with pipette. Remove the metal.

(d) Keep in new 1165 kept at 80°C for 5 min.

(e) Rinse with ACE, METH, ISO (squirt bottle).

(f) Rinse in running DI water for 1 min.

(g) Blow dry with N_2.

(h) Check under microscope, then Dektak thickness of metal. Check width of metal lines. Check if 0.4 μm lines came through.

A.3 Base Contact etch (no mask required)

A Oxygen Plasma Descum

(a) 300mT of O_2.

(b) power = 100W at low frequency.

(c) run for 30 seconds.

B Surface Prep

(a) Mix a dilute solution of NH_4OH : H_2O :: 1 : 10
(b) Dip in dilute NH_4OH for 10 sec.
(c) Rinse in DI for 1 min.
(d) Blow dry with N_2.

C InGaAs Emitter-cap Wet Etch

(a) Mix a solution of H_2O_2 : H_3PO_4 : H_2O :: 1 : 1 : 25
(b) Etch for 23-24 seconds. Colorchange complete after 10-15 secs.
(c) Dektak to make sure that the etch depth is about 400 Å.

D InP Emitter Wet Etch

(a) Mix a solution of HCl : H_2O :: 1 : 4
(b) Etch for 22-24 seconds
(c) Dektak to make sure that the etch depth has increased by about 1200 Å

E Nonselective etch: Only for samples with emitter-base grade

(a) Mix etchant as follows: 55ml of 1M citric acid in 220ml DI. Mix well. Add 5ml H_2O_2. Mix well. Add 1ml Phosphoric acid. Mix well.
(b) Stirring at 200 rpm etch for 35 seconds by suspending the wafer in a basket.
(c) Rinse in DI for 3 min.
(d) Blow dry with N_2.
(e) You should have etched $\sim$ 300 Åof semiconductor or $\sim$ 100 Åinto the base.

A.4 Base Contact (Mask Layer 2)

A Dehydration bake, 120°C, 30 min. in petri dish without cover

B Photoresist Application and Exposure

(a) Cool down in HMDS chamber 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply nLOF2020 with syringe and filter to cover wafer.
(d) Spin at 4 krpm for 30 sec.
(e) Hot Plate Bake, 11°C, 1 min.
(f) Expose for 0.43 sec., focus offset of 2.
(g) Hot Plate Bake, 112°C, 1min.

C Development

(a) Develop in MF-701 for 1 minute 45 sec.
(b) Rinse in running DI water for 1 min.

(c) Blow dry with N_2.

(d) Observe under microscope. If there is a suspicion that some photoresist is remaining may go back into developer for 5 - 10 secs longer.

D UV Ozone Descum of Photoresist

(a) Run UV-ozone chamber for at least 20 minutes to warm up the lamp.

(b) Run with wafers in petri-dish for 10 minutes.

(c) load sample in the reactor

(d) run for 10 minutes

(e) Resist adhesion bake hotplate 111°C 1 min.

E Evaporation Ebeam 4

(a) Mix a dilute solution of NH_4OH : H_2O :: 1 : 10

(b) Dip in dilute NH_4OH for 15 sec.

(c) **No DI Rinse**

(d) Blow dry with N_2.

(e) Load wafer. Include a glass slide for metal resistance measurement.

(f) Pump down to below 1 x 10-6 torr.

(g) Deposit material:

Material	Thickness(Å)	Dep. Rate (Å/sec)
Pd	25-30	1-2
Ti	150	1-2
Pd	150	1-2
Au	700	3-5

F Liftoff –>> DO NOT LET ACE OR 1165 DRY ON WAFER! <<–

(a) Soak wafer in beaker of 1165 until metal comes loose.

(b) If the liftoff is stubborn, leave the wafer soaking overnight. Seal the top of the beaker with foil. Leave name and number.

(c) 1165 squirt with pipette. Remove the metal.

(d) Keep in new 1165 kept at 80°C for 5 min.

(e) Rinse with ACE, METH, ISO (squirt bottle).

(f) Rinse in running DI water for 1 min.

(g) Blow dry with N_2.

(h) Check under microscope, then Dektak thickness of metal.

G RTA for 60 seconds at 300°C; program 300/1'p. Run a test program first.

A.5 Base Plug (Mask Layer 3)

A Solvent Cleaning

(a) Check the resistivity of the D.I. water. It should be > 17 MΩ.
(b) Cold ACE 3 min.
(c) Cold METH 3 min.
(d) Cold ISO 3 min.
(e) Running DI 3 min.
(f) Blow dry with N_2.
(g) Dehydration bake, 120°C, 30 min in petri dish without cover.

B Photoresist Application and Exposure

(a) Cool down in HMDS chamber 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply SPR 955-0.9 with syringe and filter to cover wafer.
(d) Spin at 2.5 krpm for 30 sec.
(e) Soft Bake, 90°C, 1 min. on hot plate.
(f) Apply CEM with syringe and filter to cover wafer.
(g) Wait 1 min.
(h) Spin at 4 krpm for 30 sec.
(i) Expose for 2.2 seconds, focus around 2.

C Development

(a) Develop in MF-701 for 2 min.
(b) Rinse in running DI water for 1 min.
(c) Blow dry with N_2.
(d) Observe under microscope. If there is a suspicion that some photoresist is remaining may go back into developer for 5 - 10 secs longer.

D UV Ozone Descum of Photoresist

(a) Run UV-ozone chamber for at least 20 minutes to warm up the lamp.
(b) Run with wafers in petri-dish for 10 minutes.
(c) Resist adhesion bake hotplate 111°C 1 min.

E Evaporation in Ebeam 1 or 4

(a) Mix a dilute solution of HCl : H_2O :: 1 : 10
(b) Dip in dilute HCl for 15 sec.
(c) Rinse in running DI wafer for 3 min.
(d) Blow dry with N_2.
(e) Place wafer on rotating stage. Level it.
(f) A new crystal should always be used in Ebeam 1.
(g) Pump down to below 1 x 10-6 torr.
(h) Deposit material:

Material	Thickness(Å)	Dep. Rate (Å/sec)
Ti	200	1-2
Pd	400	1-2
Au	9800	3-7

F Liftoff –>> DO NOT LET ACE OR 1165 DRY ON WAFER! <<–

(a) Soak wafer in beaker of 1165 until metal comes loose.
(b) If the liftoff is stubborn, leave the wafer soaking overnight. Seal the top of the beaker with foil. Leave name and number.
(c) 1165 squirt with pipette. Remove the metal.
(d) Keep in new 1165 kept at 80°C for 5 min.
(e) Rinse with ACE, METH, ISO (squirt bottle).
(f) Rinse in running DI water for 1 min.
(g) Blow dry with N_2.
(h) Check under microscope, then Dektak thickness of metal.

A.6 Base-Mesa Isolation (Mask Layer 4)

A Solvent Cleaning

(a) Check the resistivity of the D.I. water. It should be > 17 MΩ.
(b) Cold ACE 3 min.
(c) Cold METH 3 min.
(d) Cold ISO 3 min.
(e) Running DI 3 min.
(f) Blow dry with N_2
(g) Dehydration bake, 120°C, 30 min. in petri dish without cover.

B Photoresist Application and Exposure

(a) Cool down in HMDS chamber 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply SPR 518-A with syringe and filter to cover wafer.
(d) Spin at 4.0 krpm for 30 sec.
(e) Hot Plate Bake, 90°C, 1 minute.
(f) Expose for 1.2 sec, focus offset of 20.
(g) Hot Plate Bake, 110°C, 1 minute.

C Development

(a) Develop in full beaker of MF-701 for 1 min, 30 seconds.
(b) Rinse in running DI water for 1 min.

(c) Blow dry with N_2.

(d) Observe under microscope. If there is a suspicion that some photoresist is remaining may go back into developer for 5 - 10 secs longer.

D Oxygen Plasma Descum of Photoresist and Hardbake

(a) 300mT of O_2.

(b) power = 100W at low frequency.

(c) run for 30 seconds.

(d) Resist adhesion bake hotplate 111°C 1 min.

E Base-Mesa Isolation Etch

(a) Mix a solution of H_2O_2 : H_3PO_4 : H_2O :: 1 : 1 : 25

(b) Etch (the InGaAs base and InAlGaAs grade) for 35 seconds

(c) Dektak to make sure that the etch depth is about 500 Å.

(d) Mix a solution of HCl : H_2O :: 1 : 4

(e) Etch (the InP collector) for 25 seconds

InP thickness (Å)	Etch time (sec)	Time until final color change (sec)
600	30	18
1000	35	20
1500	45	24

(f) Dektak to make sure that the etch depth has increased by the expected amount.

F Resist strip

(a) Remove photoresist with warm 1165.

(b) Rinse by spraying with ACE, METH and ISO.

(c) Rinse in running DI water for 3 min.

(d) Blow dry with N_2.

A.7 Collector Contacts (Mask Layer 5)

A Solvent Cleaning

(a) Check the resistivity of the D.I. water. It should be > 17 MΩ.

(b) Cold ACE 3 min.

(c) Cold METH 3 min.

(d) Cold ISO 3 min.

(e) Running DI 3 min.

(f) Blow dry with N_2.

(g) Dehydration bake, 120°C, 30 min in petri dish without cover.

B Photoresist Application and Exposure

(a) Cool down in HMDS chamber 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply nLOF2020 with syringe and filter to cover wafer.
(d) Spin at 4 krpm for 30 sec.
(e) Hot Plate Bake, 111°C, 1 min.
(f) Expose for 0.45 sec., focus offset of 2.
(g) Hot Plate Bake, 112°C, 1min.

C Development

(a) Develop in MF-701 for 1 min 45 sec.
(b) Rinse in running DI water for 1 min.
(c) Blow dry with N_2.
(d) Observe under microscope. If there is a suspicion that some photoresist is remaining may go back into developer for 5 - 10 secs longer.

D UV Ozone Descum of Photoresist

(a) Run UV-ozone chamber for at least 20 minutes to warm up the lamp.
(b) Run with wafers in petri-dish for 10 minutes.
(c) load sample in the reactor
(d) run for 10 minutes
(e) Resist adhesion bake hotplate 111°C 1 min.

E Evaporation

(a) Mix a dilute solution of HCl : H_2O :: 1 : 10
(b) Dip in dilute HCl for 15 sec.
(c) Rinse in running DI wafer for 3 min.
(d) Blow dry with N_2.
(e) Place wafer in Ebeam 4.
(f) Deposit material:

Material	Thickness(Å)	Dep. Rate (Å/sec)
Ti	200	1-2
Pd	400	1-2
Au	3000	5-7

F Liftoff –>> DO NOT LET ACE OR 1165 DRY ON WAFER! <<–

(a) Soak wafer in beaker of 1165 until metal comes loose.
(b) If the liftoff is stubborn, leave the wafer soaking overnight. Seal the top of the beaker with foil. Leave name and number.
(c) 1165 squirt with pipette. Remove the metal.

(d) Keep in new 1165 kept at 80°C for 5 min.
(e) Rinse with ACE, METH, ISO (squirt bottle).
(f) Rinse in running DI water for 1 min.
(g) Blow dry with N_2.
(h) Check under microscope, then Dektak thickness of metal.

A.8 Device Isolation (Mask Layer 6)

A Solvent Cleaning

(a) Check the resistivity of the D.I. water. It should be > 17 MΩ.
(b) Cold ACE 3 min.
(c) Cold METH 3 min.
(d) Cold ISO 3 min.
(e) Running DI 3 min.
(f) Blow dry with N_2
(g) Dehydration bake, 120°C, 30 min. in petri dish without cover.

B Photoresist Application and Exposure

(a) Cool down in HMDS chamber 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N2.
(c) Apply SPR 518-A with syringe and filter to cover wafer.
(d) Spin at 3.0 krpm for 30 sec.
(e) Hot Plate Bake, 90°C, 1 minute.
(f) Expose for 1.3 sec, focus offset of 2.
(g) Hot Plate Bake, 110°C, 1 minute.

C Development

(a) Develop in full beaker of MF-701 for 1 min, 30 seconds.
(b) Rinse in running DI water for 3 min.
(c) Blow dry with N_2.
(d) Observe under microscope. If there is a suspicion that some photoresist is remaining may go back into developer for 5 - 10 secs longer.

D Oxygen Plasma Descum of Photoresist and Hardbake

(a) 300mT of O_2.
(b) power = 100W at low frequency.
(c) run for 60 seconds.

E Device Isolation Etch

(a) Mix a solution of $H_2O_2 : H_3PO_4 : H_2O :: 1 : 1 : 25$
(b) Etch (the InGaAs sub-collector) for 20 seconds

(c) Dektak to make sure that the etch depth is about 250 Å.
(d) Mix a solution of HCl : H_2O :: 1 : 4
(e) Etch (the InP sub-collector and the InP buffer) for 45 seconds
(f) Dektak to make sure the SI-substrate has been reached.
(g) Check Isolation by measuring leakage between two collector pads. The resistance should be higher than 10 MΩ

F Resist strip

(a) Remove photoresist with acetone in liftoff beaker for 3 minutes.
(b) Rinse by spraying with METH and ISO.
(c) Rinse in running DI water for 3 min.
(d) Blow dry with N_2.

G Oxygen Plasma Photoresist Removal (optional, at your discretion)

(a) 300mT of O_2.
(b) power = 100W at low frequency.
(c) run for 30 seconds.
(d) examine under microscope to ensure photoresist removal.

H Characterization of process so far.

(a) Dektak etch depth, measure TLMs.

A.9 Poly Planarize (Mask Layer 7)

A Solvent Cleaning

(a) **Get Poly out of refrigerator, warm up for at least 2 hours!!**
(b) Check the resistivity of the D.I. water. It should be > 17MΩ.
(c) Cold ACE 3 min.
(d) Cold METH 3 min.
(e) Cold ISO 3 min.
(f) Running DI 3 min.
(g) Blow dry with N_2.
(h) Dehydration bake, 120°C, 30 min. in petri dish without cover.

B Poly Spin & Cure

(a) Mix adhesion promoter in designated beaker (∼ 0.3 mL of VM-651 from bottle using plastic dropper in 300 mL of D.I.). Stir then get a new dropper.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply adhesion promoter to cover wafer.
(d) Let it sit on the wafer for 20 seconds.

(e) Spin at 3.0 krpm for 60 sec.
(f) Hot plate bake, 125°C, 3 minutes.
(g) Apply DuPont Ployimide (PI 2556) to cover wafer with pipette. Take care to avoid bubbles.
(h) Let sit on wafer for 20 seconds.
(i) Spin at 2.5 krpm for 30 sec (to give 1.8 μm film).
(j) Hard bake polyimide using program P3 in the BLUE-M OVEN. Always place the wafers on foil and check the underside for polyimid.
   i. hold at 90°C for 60 min.
   ii. ramp at 4°C per min. to 150°C.
   iii. hold at 150°C for 60 min.
   iv. ramp at 4°C per min. to 230°C.
   v. hold at 230°C for 60 min.
   vi. ramp at 4°C per min. to 170°C.

C Photoresist Application and Exposure

(a) Remove from oven when below 50°C.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Apply AZ 4330 with pipette. Take care to avoid bubbles.
(d) Spin at 5.0 krpm for 30 sec.
(e) Soft Bake, 90°C, 30 min. in petri dish without cover.
(f) Hard Bake, 120°C, 30 min. in petri dish without cover.

D Polyimide Etchback

(a) Load RIE#1 according to instructions.
(b) Pump down to low E-6.
(c) Set up laser monitor.
   i. Look for diffraction pattern to identify beam. Laser signal should be about 500mV. Use outermost spot of laser beam.
   ii. Set up chart recorder for 1 hour and ~700mV.
(d) Etch conditions:
   i. O_2 flow rate 7.0 sccm.
   ii. chamber pressure = 10 mTorr.
   iii. P=60W.
(e) Etch for 17.5 cycles.

F Photoresist Application and Exposure

(a) Wafer on spinner chuck with vacuum, blow with N_2.
(b) Apply SPR 518-A with syringe and filter to cover wafer.
(c) Spin at 2.5 krpm for 30 sec.

(d) Hot Plate Bake, 90°C, 1 minute.
(e) Expose for 1.3 sec, focus offset of +2. Must use global alignment.
(f) Hot Plate Bake, 110°C, 1 minute.

G Development

(a) Develop in MF-701 for 1 min, 30 seconds.
(b) Rinse in running D.I. for 1 minutes.
(c) Hard Bake, 120°C, hotplate 5 min.

I Etch Poly Poly etch for 30 minutes, PEII-A, 300 mT O_2, 100 W.

J Resist strip

(a) Flood Expose at 7.5 mW/cm^2 for 2 minutes.
(b) Remove resist in MF-701, 2 minutes.
(c) Rinse in running D.I. for 3 minutes.
(d) Blow dry with N_2.
(e) Inspect under microscope.
(f) Load wafer in SEM and look at 1um emitter fingers to see if they are clear.
(g) If not, etch two minutes in PEII-A at 100W, 300 mT then SEM again.

K Reflow Bake Hot plate bake, 250°C, 10 minutes to reflow polyimide and complete the cure. Note: even after reflow bake is polyimid attacked by strong strippers such as PRX-127.

A.10 Metal-1 (Mask Layer 9)

A Solvent Cleaning

(a) Check the resistivity of the D.I. water. It should be $> 17M\Omega$.
(b) Cold ACE 3 min.
(c) Cold METH 3 min.
(d) Cold ISO 3 min.
(e) Running DI 3 min.
(f) Blow dry with N_2.
(g) Dehydration bake, 120°C, 30 min. in petri dish without cover.

B Photoresist Application and Exposure

(a) Cool down after dehydration, 3 min.
(b) Wafer on spinner chuck with vacuum, blow with N_2.
(c) Spin on HMDS at 4 krpm.
(d) Apply nLOF2020 with syringe and filter to cover wafer.
(e) Spin at 3 krpm for 30 sec.

(f) Hot Plate Bake, 111°C, 1 min.

(g) Expose for 0.45 sec, focus offset of 18.

(h) Hot Plate Bake, 112°C, 1 min 15 seconds.

C Development

(a) Develop in MF-701 for 2 minutes.

(b) Rinse in running DI water for 1 min.

(c) Blow dry with N_2.

(d) Observe under microscope. If there is a suspicion that some photoresist is remaining may go back into developer for 5 - 10 secs longer.

D UV Ozone Descum of Photoresist

(a) Run UV-ozone chamber for at least 20 minutes to warm up the lamp.

(b) Run with wafers in petri-dish for 10 minutes.

(c) Resist adhesion bake hotplate 111°C 1 min.

E Evaporation in Ebeam 4

(a) Mix a dilute solution of HCl : H_2O :: 1 : 10.

(b) Dip in dilute HCl for 15 sec.

(c) Rinse in DI for 3 min.

(d) Blow dry with N_2.

(e) Load ebeam 4.

(f) Pump down to below 1 x 10-6 torr.

(g) Deposit material:

Material	Thickness(Å)	Dep. Rate (Å/sec)
Ti	200	2-3
Pd	200	2-3
Au	9000	2-7

F Liftoff –>> DO NOT LET ACE OR 1165 DRY ON WAFER! <<–

(a) Soak wafer in beaker of 1165 until metal comes loose.

(b) If the liftoff is stubborn, leave the wafer soaking overnight. Seal the top of the beaker with foil. Leave name and number.

(c) 1165 squirt with pipette. Remove the metal.

(d) Keep in new 1165 kept at 80°C for 5 min.

(e) Rinse with ACE, METH, ISO (squirt bottle).

(f) Rinse in running DI water for 1 min.

(g) Blow dry with N_2.

(h) Check under microscope, then Dektak thickness of metal.

A.11 **Airbridges (Mask Layer 10-11)**

A Proceed with optional airbridge process after RF measurement

Bibliography

[1] M. J. W. Rodwell, M. Urteaga, Y. Betser, T. Mathew, S. Krishnan, D. Scott, S. Jaganathan, D. Mensa, J. Guthrie, R. Pullela, Q. Lee, B. Agarwal, U. Bhattacharya, S. Long, S. C. Martin and R. P. Smith, "Scaling of InGaAs/InAlAs HBTs for High Speed Mixed-Signal and mm-Wave ICs", *International Journal of High Speed Electronics and Systems*, vol. 11, (no. 1), pp. 159-216, 2001

[2] T. Enoki, E. Sano, T. Ishibashi, "Prospects of InP-based IC technologies for 100-Gbit/s -class lightwave communications systems", *International Journal of High Speed Electronics and Systems*, Vol. 11, No. 1, pp. 137-158, 2001

[3] M. Ida, K. Kurishima and N. Watanabe, "Over 300 GHz f_τ and f_{max} InP/InGaAs Double Heterojunction Bipolar Transistors with a thin Pseudomorphic Base", *IEEE Electron Device Letters*, vol 23, (no. 12), pp. 694-696, Dec. 2002

[4] M. W. Dvorak, C. R. Bolognesi, O. J. Pitts and S. P. Watkins, "300 GHz InP/GaAsSb/InP double HBTs with high current capability and $BV_{CEO} > 6V$", *IEEE Electron Device Letters*, vol 22, (no. 8), pp. 361-3, Aug. 2001

[5] M. Ida, K. Kurishima, N. Watanabe, T. Enoki, "InP/InGaAs DHBTs with 341-GHz f_τ at high current density of over 800 kA/cm^2", *International Electron Devices Meeting 2001*, pp. 35.4.

[6] M.J.W. Rodwell, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S.C. Martin, R. P. Smith, S. Jaganathan, T. Mathew, P. Krishnan, S. Long, R. Pullela, B. Agarwal, U. Bhattacharya, L. Samoska"Submicron Scaling of Heterojunction Bipolar Transistors for THz Device Bandwidths", *IEEE TED Special Issue on the History of the Bipolar Junction Transistor*

[7] B. Jaganathan, M. Khater, F. Pagette, J. -S. Rieh, D. Angell, H. Chen, J. Florkey, F. Golan, D. R. Greenberg, R. Groves, S. J. Jeng, J. Johnson, E. Mengistu, K. T. Schonenberg, C. M. Schnabel, P. Smith, A. Stricker, D. Ahlgren, G. Freeman, K. Stein and S. Subbanna, "Self-Aligned SiGe NPN

Transistors with 285 GHz f_{max} and 207 GHz f_τ in a Manufacturable Technology", *IEEE Electron Device Letters*, vol. 23, no. 5, pp. 258-260, May 2002

[8] S. J. Jeng, B. Jaganathan, J. S. Rieh, J. Johnson, K. T. Schonenberg, D. Greenberg, A. Stricker, H. Chen, M. Khater, D. Ahlgren, G. Freeman, K. Stein, S. Subbanna, "A 210-GHz f_τ SiGe HBT with a non-self-aligned structure", *IEE Electron Device Letters*, vol 22, (no. 11), pp. 542-4, Nov. 2001

[9] H. Kroemer, "Heterostructure Bipolar Transistors and Integrated Circuits", IEEE Proceedings, Vol. 70, No. 1, January 1982, pp. 13-25

[10] P. Asbeck, F. Chang, K.-C. Wang, G. Sullivan, and D. Cheung, "GaAs-based Heterojunction Bipolar Transistors for Very High Performance Electronic Circuits", IEEE Proceedings, vol. 81 (12), pp. 1709-1726, December 1993.

[11] M. Sokolich, D. P. Docter, Y.K. Brown, A.R. Kamer, J.F. Jensen, W.E. Stanchina, S. Thomas III, C. H. Fields, D. A. Ahmari, M. Lui, R. Martinez, J. Duvall, "A low power 52.9 GHz static frequency divider in a manufacturable 180 GHz AlInAs/InGaAs HBT IC technology", in *IEEE GaAs IC Symposium Technical Digest*, Nov. 1-4, 1998, Atlanta, Ga. , pp. 117-120.

[12] W. Liu, "III-V Heterojunction Bipolar Transistors", *Wiley–Interscince* 1998.

[13] M.J.W. Rodwell, S.T. Allen, R.Y. Yu, M.G. Case, M. Reddy, E. Carman, J. Pusl, M. Kamegawa, Y. Konishi, and R. Pullela, "Active and Nonlinear Wave Propagation Devices in Ultrafast Electronics and Optoelectronics", IEEE Proceedings, Vol. 82, No. 7, pp. 1037-1058, July 1994.

[14] Y. Matsuoka, S. Yamahata, K. Kurishima and H. Ito, "Ultrahigh-speed InP/InGaAs Double-Heterostructure Bipolar Transistors and Analysis of Their Operation", *Japanese Journal of Applied Physics*, vol. 35, pp.5646-5654, 1996.

[15] V. Palankovski, "Simulation of Heterojunction Bipolar Transistors", *Ph.D. Thesis*, Technischen Universität Wien, Fakultät für Elektrotechnik und Informationstechnik, Dec 2000

[16] "NSM Archive - Physical Properties of Semiconductors" *available online: http://www.ioffe.ru/SVA/NSM/Semicond/*

[17] I. Vurgaftmanm J.R. Meyer, L.R. Ram-Mohan "Band parameters for III-V compound semiconductors and their alloys", *Journal of Applied Physics*, Vol. 89, No. 11, 1 June 2001.

[18] C. Kittel, *Introduction to Solid State Physics*, 7th Ed., Wiley, (1996).

[19] T. Ishibashi, "Influence of electron velocity overshoot on collector transit times of HBTs", IEEE Transactions on Electron Devices, vol. 37, no. 9, pp. 2103-2105, September 1990.

[20] D.M. Lubushev J. Neal, W.Z. Cai, M. Micovic, T.S. Mayer, D. L. Miller "MBE Growth of Near–Infrared InGaAs Photodetectors with Carbon Tetrabromide as a p–type Dopant", *The 24'th International Symposium on Compound Semiconductors*, 8-9 September 1997.

[21] Ian Harrison, U. of Nottingham, *private communication*, UCSB 2002.

[22] AC.D. Latham, R. Jones, S. Öberg, P.R. Briddon, "Density–functional calulations of carbon doping in III-V compund semiconductors", *Physical Review B.*, Vol. 63 2001.

[23] Y-M. Kim,*private communication*, UCSB 2001.

[24] D. Scott, *private communication*, UCSB 2001.

[25] M. Bächthhold, J.G. Korvink, J. Funk, H. Baltes, "Simulation of p-n junctions in Highly Doped Semiconductors at Equilibrium Conditions", *Physical Electronics Laboratory Report, ETH Höngerberg*, May 7 1997.

[26] S. Yamahata, K. Kurishima, H. Ito and Y. Matsuoka, "Over-220-GHz-f_τ-and-f_{max} InP/InGaAs double-heterojunction bipolar transistors with a new hexagonal-shaped emitter", *GaAs IC Symp. Tech. Dig.*, 1995, pp. 163-166.

[27] M. Sokolich, D.P. Docter, Y.K. Brown, A.R. Kramer, J.F. Jensen, W.E. Stanchina, S. Thomas III, C.H. Fields, D.A. Ahmari, M. Lui, R. Martinez, J.Duvall, "A Low Power 52.9 GHz Static Divider Implemented in a Manufacturable 180 GHz AlInAs/InGaAs HBT IC Technology ", *IEEE GaAs IC Symp. Tech. Dig.*, pp. 117-120, 1998.

[28] J. Fastenau, Growth calibratrions performed by IQE

[29] B.G. Streetman, *Solid State Electronic Devices*, third edition, Prentice-Hall, 1990.

[30] S. Laux, W. Lee, "Collector signal delay in the presence of velocity overshoot", IEEE Electron Device Letters, vol. 11, No. 4, pp. 174-176, 1990

[31] H. Kroemer, "Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region", *Solid State Electronics* Vol. 28, No.11 pp.1101-1103 1985.

[32] T. Oka K. Ouchi, K. Mochizuki, "Characterization of InGaP/GaAs Heterojunction Bipolar Transistors with a Heavily doped base", *Jpn. J. of Applied Phys.*, Vol. 40, 2001

[33] H. Kroemer, "ECE 221 course notes", UCSB.

[34] S. Yamahata, Y. Matsuoka, T. Ishibashi, "Ultrahigh-speed AlGaAs/GaAs ballistic collection transistors using carbon as a p-type dopant", Electronics Letters, Vol. 29, No. 22, 28 October 1993, pp. 1996-1997

[35] M. Ito, S. Yamahata, K. Kurishima," Evaluation of base transit time in ultrathin carbon doped base InP/InGaAs heterojunction bipolar transistors", *Electronic Letters*, Vol. 32, No. 15, 1996.

[36] J.L. Benchimol, J. Mba, B. Sermage, M. Riet, S. Blayac, P. Berduager, A.M. Duchenois, P. Andre, J. Thuret, C. Gonzales, A. Konczykowska, "Investigation of carbon-doped base materials grown by CBE for Al-free InP HBTs", *Journal of Crystal Growth*, 209, 2000.

[37] Y. Betser, D. Ritter, "High Emitter Efficiency in InP/InGaAs HBTs with ultra high base doping levels", *IEEE Electron Device Letters* Vol. 16, No.3 1995

[38] A paper about passivation with Si, SiO, SiN, Sic

[39] U. Bhattacharya "Transferred-substrate heterojunction bipolar transistors", *Ph. D. Dissertation*, University of California, Santa Barbara, 1996.

[40] V. Patri, J. Kumar, "Profile Design Considerations for Minizing Base Transit Time in SiGe HBTs" *IEEE Transaction on Electron Devices*, Vol. 45, no. 8, Aug. 1998.

[41] D. Streit, M. Hafizi, D. Umemoto, J.R. Velebeir, L. Tran, A. Oki, M. Kim, S. Wang, C. Kim, L. Sadwick, R.J. Hwu, "Effect of exponentially graded base doping on the performance of GaAs/AlGaAs heterojunction bipolar transistors", *IEEE Electron Device Letters* Vol. 12, No. 5, May 1991

[42] K.H. Kwok, "Analytical expressions of base transit time for SiGe HBTs with retrograde base profiles", *Solid State Electronics*, Vol. 43, 1999.

[43] D. Streit, A. Oki, L. Tran, D. Umemoto, K. Kobayashi "High performance HBT's with built-in base fields: exponentially graded doping vs. graded composition", *Conference proceedings*

[44] V. Palankovski, G. Kaiblinger-Grujin, S. Selberherr, "Study of dopant dependant band gap narrowing in compound semiconductor devices ", *Materials Science and Engineering B* , B66, 1999.

[45] J. Lopez-Gonzalez, Lluis Prat, "The importance of Bandgap Narrowing Distribution Between the Conduction and Valence Band in Abrupt HBTs", *IEEE Transaction on Electron Devices*, Vol. 44, no. 7, July 1997.

[46] H. Kroemer , "Heterojunction Bipolar Transistors: What should we built? ", *J. Vac. Sci. Tech. B*, Vol.1, No.2, April-June 1983

[47] M. Okhubo, J. Osabe, T. Nimomiya,“ Compositionally graded C-doped InGaAs Base in InP/InGaAs D-HBTs grown by MOCVD with Low Base Sheet Resistance and High Current Gains”, *Proceedings of conference*, ????

[48] H.Wang, G.I. Ng, “ Avalananche Multiplication in InP/InGaAs double heterojunction Bipolar Transistors with Composite Collector”, *IEEE Trans. on Electron Devices* V ol. 47 No. 6, June 2000

[49] R. Teissier, J.L Pelouard, F. Mollot, “ Direct measurement of ballistic electron distribution and relaxation length in InP-based heterojunction bipolar transistors using photoluminescence spectroscopy”, *Appl. Phys. Lett.*, vol 72, pp 2730-2732, 1998.

[50] E.F. Chor, D. Zhang, H. Gong, W.K. Chong, S.Y. Ong “ Electrical characterization, metallurgical investigation, and thermal stability studies of (Pd, Ti, Au)-based ohmic contacts ”, . *Journal of Applied Physics*, vol.87, (no.5), AIP, 1 March 2000. p.2437-44

[51] W.K. Chong , E.F. Chor, C.H. Cheng, S.J. Chua, “Electrical characterization, metallurgical investigation, and thermal stability studies of (Pd, Ti, Au)-based ohmic contacts ”, *IEEE*, 1998

[52] J.S. Yu, S.H. Kim, T.I. Kim, “ PtTiPtAu and PdTiPtAu ohmic contacts to p-InGaAs”, *Proceedings of the IEEE Twenty-Fourth International Symposium on Compound Semiconductors* , San Diego, CA, USA, 8-11 Sept. 1997

[53] K.S Sandhu, A.E. Staton-Bevan, M.A. Crouch, “ The role of the Pd diffusion barrier in Au/Pd/Ti Ohmic contact structures to p-GaAs for HBT applications.”, *Proceedings of the Institute of Physics Electron Microscopy and Analysis Group Conference* , Bristol, UK, 12-15 Sept. 1995

[54] A. Kobayashi, T.Sakurai, T. Hashizume, “ An atomistic study of the GaAs-Pd interface”, *Journal of Applied Physics* , 59, 15 May 1996.

[55] E. Nebauer, M. Mai, J. Würfl and W. Österle, “Au/Pt/Ti/Pt base contacts to n-InGaP/p^+–GaAs for self–passivating HBT ledge structures, *Semicond. Sci. Technol.* Vol.15, pp.818-822, 2000.

[56] , T. Usagawa, M. Kobayashi, P.D. Rabinzohn, A. Ihara, M. Kawata, T. Yamada, E. Tokumitsu, M. Konagai, K. Takahashi, “Extremely low non-alloyed specific contact resistance $\mathrm{rho_c}(10^{-8}\Omega\mathrm{cm}^2)$ to metalorganic molecular beam epitaxy grown super heavily C-doped $10^{21}cm^{-3}$ p^{++} GaAs”, *J. Appl. Physics* ,69 (12) 15 June 1991

[57] R. Caron-Popowich, J. Washburn, T. Sands, A.S. Kaplan, “Phase formation in the Pd-InP system”, textit J. Appl. Physics , Vol. 64, 15 November 1988.

[58] D. Rosenfeld, S. Alterovitz, "Carrier Transit Time Through a Base with Dopant Dependent Mobility", *IEEE Transaction if Electron Devices*, Vol. 41, No 5, May 1994

[59] K.H. Kwok, " Analytical expressions of base transit time for SiGe HBTs with retrograde base profile", *Solid State Electronics* , 43 (1999) pp.275-283

[60] D. Mensa, Q. Lee, R. Pullela, B. Agarwal, J. Guthrie, S. Jaganathan, M. Rodwell, "Baseband Amplifiers in Transferred-Substrate HBT Technology", *GaAs IC Symp. Tech. Dig.*, submitted for presentation, 1998.

[61] Q. Lee, S. C. Martin, D. Mensa, R. P. Smith, J. Guthrie and M. J. W. Rodwell, "Submicron transferred-substrate heterojunction bipolar transistors", *IEEE Electron Device Letters*, vol. 20, no. 8, pp. 396-8, Aug. 99

[62] Q. Lee, D. Mensa, J. Guthrie, S. Jaganathan, T. Mathew, Y. Betser, S. Krishnan, S. Ceran, M. J. W. Rodwell, "66 GHz static frequency divider in transferred-substrate HBT technology", *1999 IEEE Radio Frequency Integrated Circuits Symposium*, Anaheim, CA, USA, pp. 87-90, June 1999

[63] P. J. Zampardi and D. Pan, "Delay of Kirk Effect Due to Collector Current Spreading in Heterojunction Bipolar Transistors" , *IEEE Electron Device Letters*, Vol.17 No.10 October 1996.

[64] J. Mba, D. Caffin, A.M. Duchenois, M. Riet, J.L. Benchimol, P. Launay, J. Godin, A. Scavennec " Low-Power operation in InP-based DHBT's for high bit rate circuit applications: reduction of saturation voltage (V_{sat}) ", *10th International Conference on Indium Phosphide and Related Materials* , 11-15 May 1998 Tsukuba, Japan.

[65] K. Yang, G. Munns, G. Haddad, " High f_{max} InP Double Heterojunction Bipolar Transistros with Chirped InGaAs/InP Superlattice Base-Collector Junction grown by CBE", *IEEE Electron Device Letters* , Vol 18, No 11 November 1997

[66] M. Dvorak, O.J. Pitts, S.P. Watkins, C.R. Bolognesi, "Abrubt Junction InP/GaAsSb/InP Double Heterojunction Bipolar Transistors with f_t as High as 250 GHz and $BV_{CEO} > 6$ V", *Proceedings of IEDM 2000* ,2000.

[67] C. Nguyen, T. Liu, M. Chen, R. Virk, M. Chen, "Bandgap engineered InP-based power double heterojunction bipolar transistors", *1997 International Conference on Indium Phosphide and Related Materials*, Cape Cod, MA, USA, pp. 15-19, 1997

[68] K. Kurishima, M. Ida, M. Watanabe, "InP Double Heterojunction Bipolar Transistors with a Carbon-Doped Graded Base", *Conference Proceedings of Solid State Devices and Materials*, Nagoya 2002

[69] M. Dahlstrm, X.-M. Fang, D. Lubyshev, M. Urteaga, S. Krishnan, N. Parthasarathy, Y.M. Kim Y. Wu, J.M. Fastenau, W.K. Liu, and M.J.W. Rodwell, "Wideband DHBTs using a Graded Carbon-Doped InGaAs Base", *submitted to Electron Device Letters*

[70] W. Liu and D.S. Pan, "A proposed Collector Design of Double Heterojunction Bipolar Transistors for Power Applications", *IEEE Electron Device Letters*, Vol. 16, No.7, July 1995

[71] S. Lee, M. Urteaga, Y. Wei, Y. Kim, M. Dahlström, S. Krishnan, and M. Rodwell, "Ultra High fmax InP/InGaAs/InP Transferred Substrate DHBTs", *2002 IEEE Device Research Conference*, June 24-26, Santa Barbara, 2002.

[72] L. H. Camnitz and N. Moll, "An Analysis of the Cutoff-Frequency Behavior of Microwave Heterojunction Bipolar Transistors ", In *Compound Semiconductor Transistors* , edited by S. Tiwari, pp. 21-45, IEEE Press, Piscataway, 1992.

[73] C. T. Kirk, "A theory of transistor cutoff frequency(f_T) fall-off at high current density", *IEEE Transactions on Electron Devices*, ED-9, p. 164, 1962

[74] W. Liu, D. Hill, H. F. Chau, J. Sweder, T. Nagle and J. Delany, "Laterally etched undercut (LEU) technique to reduce base-collector capacitance in heterojunction bipolar transistors", *IEEE GaAs IC Symp. Tech. Dig.*, pp. 167-170, 1995.

[75] *To the reviewer:*this paper, on undercut-collector HBTs, is to be published in the 1999 International Electron Device Meeting. It is by the TRW group. A full reference will be available before the review cycle is completed on this manuscript.

[76] IQE Inc, www.iqeepi.com.

[77] K. Kurishiam, H. Nakajima, S. Yamahata, T. Kobayashi, Y. Matsouka, "Effects of a Compositionally-Graded $In_xGa_{1-x}As$ Base in Abrupt–Emitter InP/InGaAs Heterojunction Bipolar Transistors", *Jpn. Journal of Applied Physics*, Vol. 34, Part 1, No. 2B, pp. 1221-1227, 1995

[78] J.L. Benchimol, J. Mba, A.M. Duchenois, B. Sermage, P. Launay, D. Caffin, M. Meghelli, M. Juhel, " CBE growth of carbon doped InGaAs/InP HBT's for 25 GBit/s circuits", *Journal of Crystal Growth*, Vol. 188, 1998

[79] E. Skogen, *Private communction*,UCSB 2002.

[80] S. Krishnan, "T.S DHBT Processflow sheet",*Rodwell Group*, UCSB 2000.

[81] M. Urteaga, *Private communction regarding contact problems*,UCSB 2002

[82] R. Redd, D. Maurer, L.S. Klingbeil, "Revitalization of Single Layer Lift-off for Finer Resolution and Challenging Topography", *Proceedings of 2001 GaAs MANTECH* , 2001

[83] M. Toukhy, S. Mullen, P.H. Lu, G. Espin, J. Griffith and D. Maurer, "Simple process for 0.5μm Lift-off Applications in the GaAs and MEMS Industries", *Proceedings of 2002 GaAs MANTECH* , 2002

[84] A.R.Reid, T.C. Kleckner, M.K. Jackson and P.J. Zampardi, "Weak Scaling of Thermal Resistance in AlGaAs/GaAs Heterojunction Bipolar Transistors", *Proceedings of 2002 GaAs MANTECH* , 2002

[85] M. Rohner, I. Schnyder, D. Huber, H. Jäckel and C. Bergamaschi , "Gain limititions of scaled InP/InGaAs heterojunction bipolar transistors", *J. Appl. Physics*, No. 11. June 2001

[86] N. Matine ,G. Soerensen, C.R. Bolognesi, D. DiSanto, X. Xu, S.P. Watkins " Electrical Stress damage reversal in non-passivated fully self-aligned InP HBTs by ozone surface treatment ", *Eletronic Letters* , 1999 Vol. 35 No. 25

[87] T. Mathew, "High-Speed Digital ICs in Transferred-Substrate HBT Technology", *Ph. D. thesis*, UC Santa Barbara, Sep. 2001

[88] F. Ren, R.A. Hamm, J.R. Lothian, "Passivation of carbon doping in InGaAs during ECR-CVD of SiN_x", *Solid State Electronics*, Vol. 39, No. 5 1996

[89] M. Dahlström "Etching study for regrown InP MQW lasers", *unpublished* 1997

[90] V. Myiamoto, J.M.M Rios, A.G. Dentai, S. Chandrasekar " Reduction of Base–Collector Capacitance by Undercutting the Collector and Subcollector in GaInAs/InP DHBT's" *IEEE Electron Device Letters*, Vol. 17 No. 3 March 1996.

[91] S. Krishnan "A Continuous-Time Sigma-Delta A-D Converter in an InP-based HBT Technology", *Ph. D. Dissertation*, University of California, Santa Barbara, 2002.

[92] Device simulation tool from Prof. W. Frensley

[93] I. Harrison, M. Dahlström, S. Krishnan, Z. Griffith, Y.M Kim, M.J.W. Rodwell, "Thermal limitations of InP HBTs in 80 and 160Gbit integrated circuits" *to be submitted*

[94] K.M. Noujeim, "Private communication", 2002

[95] S. Krishnan, M. Dahlström, T. Mathew, Y. Wei, D. Scott, M. Urteaga and M.J.W. Rodwell "Transfered substrate DHBT with a fmax over 300 GHz", *IPRM 2001*, Nagano, Japan, May 2001

[96] Yoram Betser and Dan Ritter, "Reduction of the base collector capacitance in InP/GaInAs heterojunction bipolar transistors due to electron velocity modulation", *IEEE Trans. Electron. Dev.*, vol. 46 , no. 4, April 1999

[97] K. Bellenhoff and W. Heinrich, "Exitcation of parasitic parallell–plate line mode at coplanar discontinuities", *IEEE MTT-S Digest*, pp. 1789-1792, 1997

[98] Y.M. Kim, M. Dahlström, S. Lee, M.J.W. Rodwell, A.C. Gossard "High-Performance InP/$In_{0.53}Ga_{0.47}As$/InP Double Heterojunction Bipolar Transistors on GaAs Substrates", *Electron Device Letters*, Vol 23., No.6, June 2002

[99] M. Dahlström, M. Urteaga, S. Krishnan, N. Parthasarathy, M. J. W. Rodwell, X. M. Fang, D. Lubyshev, Y. Wu, J. M. Fastenau and W. K. Liu, "Ultra-Wideband DHBTs using a Graded Carbon-Doped InGaAs base", *post-deadline news, IPRM 2002*, Stockholm, Sweden, May 2002

[100] M.J. Rodwell "ECE 202 course notes", *ECE Dept. UCSB*

[101] Y.F. Yang, C.C Hsu, E.S. Yang, "Surface Recombination Current in InGaP/GaAs Heterostructure-Emitter Bipolar Transistors", *J. Appl. Phys.*, vol. 35, pp.5646-5654, 1996.

[102] M. Vaidyanathan and D. L. Pulfrey, "Extrapolated f_{max} of heterojunction bipolar transistors", IEEE Transactions on Electron Devices, Vol. 46, No.2, February 1999.

[103] M. Wurzer; T.F. Meister; I. Schafer; H. Knapp; J. Bock; R. Stengl; K. Aufinger; M. Franosch; M. Rest; M. Moller; H. -M. Rein; A. Felder," 42 GHz static frequency divider in a Si/SiGe bipolar technology", *IEEE International Solid-State Circuits Conference dig.*, pp.122-3, 1997.

[104] Y. Matsuoka, S. Yamahata, K. kurishima and H. Ito, "Ultrahigh-speed InP/InGaAs Double-Heterostructure Bipolar Transistors and Analysis of Their Operation", *J. Appl. Phys.*, vol. 35, pp.5646-5654, 1996.s

[105] S. Yamahata, K. Kurishima, H. Nakajima, T. Kobayashi and Y. Matsuoka, "Ultra-high f_{max} and f_{τ} InP/InGaAs double-heterojunction bipolar transistors with step-graded InGaAsP collector", *IEEE GaAs IC Symp. Tech. Dig.*, pp. 345-348, 1994.

[106] T. Ishibashi, H. Nakajima, H. Ito, S. Yamahata and Y. Matsuoka, "Suppressed base-widening in AlGaAs/GaAs ballistic collection transistors", *Device Research Conf. Tech. Dig.*, 1990, pp. VIIB-3.

[107] D. Mensa, "Improved Current-Gain Cutoff Frequency and High Gain-Bandwidth Amplifiers in Transferred-Substrate HBT Technology", *Ph. D. Thesis*, UC Santa Barbara, Sep. 1999

[108] H. Ito, S. Yamahata, N. Shigekawa, K. Kurishima; *Jpn J. Appl. Phys* Vol 35 (1996) pp. 6139-6144

[109] B.H. Cheong, K.J. Chang, "Compensation and diffusion mechanisms of carbon dopants in GaAs" *Phys. Rev. B.*, Vol. 49, n. 24 pp. 17436-17469, 1994

[110] K. Watanabe, H. Yamazaki; *J. Appl. Phys.* 74 (9) 5587-5595

[111] R. A. Hamm, R. Malik, D. Humphrey, R. Ryan, S. Chandrasekar, L. Lunardi, M. Geva, "Carbon doping of $Ga_{47}In_{53}As$ using carbontetrabromide by metalorganic molecular beam epitaxy for InP-based heterostructure bipolar transistor devices" *Appl. Phys. Letters* Vol. 67 p. 2226, 1995

[112] R. Driad, F. Alexandre, J.L. Benchimol, B. Jusserand, B. Sermage, M. Juhel, P. Launay,"Improved stability of C-doped GaAs grown by chemical beam epitaxy for heterojunction bipolar transistor applications" *Journal of Crystal Growth* Vol. 158, pp. 210-216, 1996

[113] D. Keiper, R. Westphalen, G. Landgren; *HMA Report 1997*

[114] J. Shirahashi, T. Azuma, F. Fukuchi, M. Konagai, K. Takahashi, "InGaP/GaAs Heterojunction Bipolar Transistors with an Ultra-High Carbon-Doped Base " *Jpn. J. Appl. Phys.* Vol. 34 (1995) pp. 1204-1207

[115] H. Ito and K. Kurishima,"Influence of gallium sources on carbon incorporation efficiency into InGaAs grown by metalorganic chemical vapor deposition ", *Journal of Crystal Growth*, Vol. 15, pp.215-221, 1996

[116] Q.J. Hartmann, H. Hwangbo, A. Yung, D.A. Ahmari, M.T. Fresina, J.E. Baker, G.E. Stillmam, "Removal of hydrogen from the base of carbon-doped In0.49Ga0.51P/GaAs heterojunction bipolar transistors by ex situ annealing and the effects on device characteristics" *Applied Physics Letters* Vol. 68 (7) 12 , pp. 982-984,Feb 1996

[117] T.B. Joyce, S.P. Westwater, P.J. Goodhew, R.E. Pritchard, "Growth of carbon-doped GaAs, AlGaAs and InGaAs by chemical beam epitaxy and the application of in-situ monitoring ", *Journal of Crystal Growth* Vol. 164, Issues 1-4, Pages 371-376, July 1996

[118] C. Caneau, R. Bhat, S. Goswami, M.A. Koza, "OMVPE Grown GaInAs:C for HBTs" *Journal of Elec. Materials*, Vol 25, No 3, pp.491-495, 1996

[119] T.P. Chin, P.D. Kirchner, J.M. Woodall and C.W. Tu, "Highly carbon-doped p-type $Ga_{0.5}In_{0.5}As$ and $Ga_{0.5}In_{0.5}P$ by carbon tetrachloride in gas-source molecular beam epitaxy" *Appl. Phys. Lett.* 59 (22) 1991.