UNIVERSITY of CALIFORNIA

Santa Barbara

Highly Scaled InP/InGaAs DHBTs Beyond 1 THz Bandwidth

A Dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy

in

Electrical and Computer Engineering

by

Johann Christian Rode

Committee in Charge:

> Professor Mark J. W. Rodwell, Chair
> Professor Umesh K. Mishra
> Professor Christopher Palmstrøm
> Professor Robert York
> Doctor Miguel Urteaga

March 2015

The Dissertation of Johann Christian Rode is approved.

_____

Professor Umesh K. Mishra

_____

Professor Christopher Palmstrøm

_____

Professor Robert York

_____

Doctor Miguel Urteaga

_____

Professor Mark J. W. Rodwell, Committee Chair

March 2015

Curriculum Vitæ
Johann Christian Rode

Personal

Aug. 1982   Born in Paderborn, Germany.

Education

2014   Ph.D. in Electrical and Computer Engineering, University of California, Santa Barbara.

2009   Diplom in Electrical Engineering (*with distinction*), Universität Karlsruhe, Germany.

Publications

[1] **J. C. Rode**, H.-W. Chiang, P. Choudhary, V. Jain, B.J. Thibeault; W.J. Mitchell; M.J.W. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J.M. Fastenau, A.W.K. Liu, "*Indium Phosphide Heterobipolar Transistor Technology Beyond 1 THz Bandwidth,*" in preparation.

[2] **J. C. Rode**, H.-W. Chiang, P. Choudhary, V. Jain, B.J. Thibeault; W.J. Mitchell; M.J.W. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J.M. Fastenau, A.W.K. Liu, "*An InGaAs/InP DHBT With Simultaneous $f_\tau/f_{\max}$ 404/901 GHz and 4.3 V Breakdown Voltage,*" Electron Devices Society, IEEE Journal of the , vol.3, no.1, pp.54,57, Jan. 2015.

[3] H. W. Chiang, **J. C. Rode**, P. Choudhary, and M. J. W. Rodwell, "*Interfacial transport and current gain in InGaAs/InP DHBTs for THz applications,*" 72nd IEEE Device Research Conference, 2014.

[4] H. W. Chiang, **J. C. Rode**, P. Choudhary, and M. J. W. Rodwell, "*Lateral Carrier Diffusion and Current Gain in THz InGaAs/InP DHBTs,*" J. Appl. Phys., vol. 115, no. 3, pp. 034513, 2014.

[5] S. Daneshgar, H.-C. Park, **J. C. Rode**, Z. Griffith, M. Urteaga, B.-S. Kim, M.J.W. Rodwell, "*High efficiency W-band power amplifiers using ring-shaped sub-quarter-wavelength power combining technique,*" Microwave Symposium (IMS), 2014 IEEE MTT-S International, vol., no., pp.1,4, 1-6 June 2014.

[6] H.-C. Park, S. Daneshgar, **J. C. Rode**, Z. Griffith, M. Urteaga, B.-S. Kim, M. Rodwell, "*An 81 GHz, 470 mW, 1.1 mm^2 InP HBT power amplifier with 41 series power combining using sub-quarter-wavelength baluns,*" Microwave Symposium (IMS), 2014 IEEE MTT-S International , vol., no., pp.1,4, 1-6 June 2014.

[7] H.-C. Park, S. Daneshgar, **J. C. Rode**, Z. Griffith, M. Urteaga, B.-S. Kim, M. Rodwell, "*30% PAE W-Band InP Power Amplifiers Using Sub-Quarter-Wavelength Baluns for Series-Connected Power-Combining,*" Compound Semiconductor Integrated Circuit Symposium (CSICS), 2013 IEEE , vol., no., pp.1,4, 13-16 Oct. 2013.

[8] M. J. W. Rodwell, M. Seo, J. Hacker, A. Young, Z. Griffith, R. Pierson, M. Urteaga, **J. C. Rode**, H. W. Chiang, V. Jain, E. Lobisser, T. Reed, A. Baraskar, J. J. M. Law, A. D. Carter, S. Lee, D. C. Elias, B. J. Thibeault, W. J. Mitchell, S. Stemmer, A. C. Gossard, S. Mehotra, M. Povolotskyi, G. Klimeck, "*Transistor and IC design for Sub-mm-wave and THz ICs,*" 2012 European Microwave IC Conference, 2012.

[9] M. J. W. Rodwell, **J. C. Rode**, H. W. Chiang, P. Choudhary, T. Reed, E. Bloch, S. Danesgar, H-C Park, A. C. Gossard, B. J. Thibeault, W. J. Mitchell, M. Urteaga, Z. Griffith, J. Hacker, M. Seo, B. Brar, "*THz Indium Phosphide Bipolar Transistor Technology,*" 2012 IEEE Compound Semiconductor Integrated Circuit Symposium, 2012.

[10] E. Lobisser, **J. C. Rode**, V. Jain, H. W. Chiang, A. Baraskar, W. J. Mitchell, B. J. Thibeault, Mark J. W. Rodwell, "*InGaAs/InP DHBTs with Emitter and Base Defined through Electron-Beam Lithography for Reduced Ccb and Increased RF Cut-off Frequency,*" 39th International Symposium on Compound Semiconductors, 2012.

[11] V. Jain, **J. C. Rode**, H. W. Chiang, A. Baraskar, E. Lobisser, B. J. Thibeault, M. J. W. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. M. Fastenau, W. K. Liu, "*1.0 THz f_{max} InP DHBTs in a refractory emitter and self-aligned base process for reduced base access resistance,*" 69th IEEE Device Research Conference, 2011.

[12] M.H. Beringer, S. Illy, J. Jin, S. Kern, **J. C. Rode**, M. Thumm, "*Further design steps towards a 4 MW 170 GHz coaxial-cavity gyrotron,*" Infrared, Millimeter, and Terahertz Waves, 2009. IRMMW-THz 2009. 34th International Conference on , vol., no., pp.1,2, 21-25 Sept. 2009.

[13] **J. C. Rode**, M.H. Beringer; S. Illy, S. Kern, B. Piosczyk, M. Thumm, "*Design study of magnetron injection guns for a 4 MW 170 GHz coaxial gyrotron,*" Vacuum Electronics Conference, 2009. IVEC '09. IEEE International , vol., no., pp.96,97, 28-30 April 2009.

[14] R. Bonk, R. Brenot, C. Meuer, T. Vallaitis, A. Tussupov, **J. C. Rode**, et al., "*1.3/1.5 m QD-SOAs for WDM/TDM GPON with extended reach and large upstream/downstream dynamic range. In Optical Fiber Communication Conference (p. OWQ1)*", Optical Society of America, 2009.

Abstract

Highly Scaled InP/InGaAs DHBTs

Beyond 1 THz Bandwidth

by

Johann Christian Rode

This work examines the efforts pursued to extend the bandwidth of InP-based DHBTs above 1 THz. Aggressive lithographic and epitaxial scaling of key device dimensions and simultaneous reduction of contact resistivities have enabled increased RF bandwidths by reduction of device RC and transit delays. A fabrication process for forming base electrodes and base/collector mesas of highly scaled transistors has been developed that exploits superior resolution (10 nm) and alignment (sub-30 nm) of electron beam lithography. Ultra-low resistance, thermally stable base contacts are critical for extended f_{max} bandwidth: a novel dual-deposition base metalization technique is presented that removes contaminating lithographic processes from the formation of the base contact, thereby enabling low resistivity contacts ($4\,\Omega\,\mathrm{\mu m}^2$) to ultra-thin base layers (20 nm). The composite base metal stack exploits an ultra-thin layer of platinum that controllably reacts with base, yielding low contact resistivity, as well as a thick refractory diffusion barrier which permits stable operation at high current densities and elevated temperatures. Reduction in emitter-base surface leakage and subsequent increase of current gain was achieved by passivating emitter-base

semiconductor surfaces with conformally grown ALD Al_2O_3. RF bandwidth limiting parasitics associated to the perimeter of highly scaled transistors have been identified and significantly reduced, among which are high sheet resistance of base electrodes, excess undercut of emitter stripes and improperly scaled base posts. At 100 nm collector thickness, the breakdown voltage of the transistor BV_{CEO} has been increased to more than 4.1 V by passivating base/collector surfaces.

With the technology improvements discussed, transistors with f_τ of 480 GHz and $f_{\max}$ in excess of 1 THz have been demonstrated at 200 nm emitter width and 80 nm single-sided base contact width. Transistors at the same emitter width, but 30 nm base contact width exhibit f_τ of 550 GHz and $f_{\max}$ of 850 GHz. Estimations from a finite element model predict higher bandwidth on smaller footprint transistors. However, inadequacies of RF calibration structures prevent $f_{\max}$ extraction on these devices.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The demand for submm-wave radio systems [1,2], high data rate communication systems [3] and high performance signal processing systems [4,5] drives the development of high bandwidth transistors [6]. Despite the cost advantage of highly integrated RF CMOS circuits in the matured Si technology, heterojunction bipolar transistors in Si/SiGe and InGaP/GaAs material systems remain the prime choice for commercial RF designs that require both high linearity, power-added efficiency and breakdown simultaneously while delivering high output power.

Heterojunction bipolar transistors in the InGaAs/InP material system exhibit highest RF bandwidth at a given lithographic feature size: the low effective carrier mass in InGaAs enables fast diffusive transport through the base (InGaAs: $0.045\,m_0$, GaAs: $0.067\,m_0$, SiGe: $0.12\,m_0$). Electrons are swept through the InP collector in excess of their Fermi velocity. The advantageous band alignment between the InP

emitter and InGaAs base allows for high base doping $>1 \times 10^{20}/\text{cm}^3$ for low base contact resistance while retaining high emitter injection efficiency. The dielectric strength of the InP collector enables high breakdown voltage.

In this work, InP HBTs with triple-mesa structure are discussed: fabrication begins with epitaxial growth of collector, base and emitter semiconductor by a commercial vendor. Tall emitter contacts are deposited and electrically isolated in a SiN_x sidewall process. The emitter semiconductor mesa is isolated with selective wet etches. Base electrodes are deposited around the emitters in a self-aligned process, minimizing the gap between electrodes and active device region. The base/collector semiconductor mesa is patterned with selective wet etches, and the collector terminal is fabricated by metalizing the subcollector. Finally, posts are deposited, the devices are isolated and planarized in a low-ε dielectric. Coplanar-like wiring structures are fabricated that enable RF characterization of HBT devices.

The RF performance of mesa HBTs is increased by means of scaling [7]: transit delays are lowered by thinning epitaxial base and collector layers, $(2\pi f_\tau)^{-1} \approx \tau_c + \tau_b + RC$. Concurrently, RC charging delays are reduced by lithographically narrowing emitter and base/collector widths while maintaining constant parasitic resistances R_{ex}, R_{bb}, device current I_c and transconductance g_m. Successful scaling requires lithographic resolution for both emitter and base to be less than a quarter of the smallest emitter width, while base-to-emitter alignment must be better than a third of an emitter width in order to simultaneously obtain low base access resistance

R_{bb} and base-collector capacitance C_{cb} for optimum RF performance. Major scaling challenges arise from fabricating ohmic contacts with lowest contact resistitivies to both emitter and base. The emitter and base metalization must also sustain high device current densities $J_e > 30\,\mathrm{mA/\mu m^2}$ at elevated junction temperatures $\sim 400\,\mathrm{K}$ without degradation by either electromigration or thermal decomposition processes.

Chapter 2 briefly introduces basic design considerations, figures of merit and scaling laws for triple-mesa RF HBTs. With shrinking base contact resistivity, high sheet resistivity of base electrodes $R_{sh,\mathrm{base}}$ causes voltage drops along the length of the emitter, degrading f_max bandwidth with increasing emitter length L_e: a simple calculation is presented that estimates this effect as a function of $R_{sh,\mathrm{base}}$ and L_e. A simple finite-element circuit is introduced that has been used to verify this effect as well as to quantify the deterioration of RF bandwidths due to various fabrication issues. In chapter 3, the fabrication process and improvements that have enhanced device performance and yield are discussed. The process for forming base electrodes and base/collector mesas has been transitioned from i-line projection to electron beam lithography, yielding better than $30\,\mathrm{nm}$ base-to-emitter alignment at $10\,\mathrm{nm}$ resolution. In an effort to improve base contact resistivity, a dual-deposition base metalization process is presented that has lithographic processes removed from the formation of base contacts in order to retain pristine semiconductor surfaces. The metal composition of the base electrodes has been modified to simultaneously yield thermally stable low resistivity contacts while reducing the sheet resistivity of the

base electrodes for increased f_{max} bandwidth. As part of the dual-deposition process, base/emitter semiconductor surfaces are passivated with Al_2O_3 for increased current gain and reduced base access resistance. In identifying further limitations to the fabrication process of highly-scaled devices, additional key improvements have been made to increase device performance, among which are scaled base posts, reduced emitter end undercut and base/collector passivation with SiN_x. A process for fabricating scaled TLM structures is introduced as means to quickly evaluate the effects of thermal anneals, process contaminants and modifications to base metal composition on base contact resistivity. Unlike non-scaled TLM structures, the scaled TLM fabrication resembles HBTs in both fabrication processes and dimensions. Chapter 4 discusses results of HBT fabrication campaigns and scaled TLMs. Simultaneous f_τ and f_{max} of 0.48 THz and 1.07 THz have been achieved on a sample that has been fabricated with discussed process improvements: smaller footprint devices with higher RF bandwidth f_τ 0.51 THz have been fabricated, but inadequacies in the calibration methods and structures prevent accurate extraction of f_{max} bandwidth.

References

[1] J. Hacker, M. Urteaga, M. Seo, A. Skalare, and R. Lin, "InP HBT amplifier MMICs operating to 0.67 THz," in *Microwave Symposium Digest (IMS), 2013 IEEE MTT-S International*, pp. 1–3, IEEE, 2013.

[2] M. Seo, M. Urteaga, J. Hacker, A. Young, A. Skalare, R. Lin, and M. Rodwell, "A 600 GHz InP HBT amplifier using cross-coupled feedback stabilization and dual-differential power combining," in *Microwave Symposium Digest (IMS), 2013 IEEE MTT-S International*, pp. 1–3, June 2013.

[3] E. Bloch, H.-c. Park, Z. Griffith, M. Urteaga, D. Ritter, and M. J. Rodwell, "A 107 GHz 55 dB-Ohm InP broadband transimpedance amplifier ic for high-speed optical communication links," in *Compound Semiconductor Integrated Circuit Symposium (CSICS), 2013 IEEE*, pp. 1–4, IEEE, 2013.

[4] B. Oyama, D. Ching, K. Thai, A. Gutierrez-Aitken, and V. Patel, "InP HBT/Si CMOS-based 13-b 1.33-Gsps digital-to-analog converter with >70-dB SFDR," *Solid-State Circuits, IEEE Journal of*, vol. 48, pp. 2265–2272, Oct 2013.

[5] Z. Xu, J. McArdle-Moore, T. Oh, S. Kim, S. Chen, Y. Royter, M. Lau, I. Valles, D. Hitko, and J. Li, "A 0.8/2.4 GHz tunable active band pass filter in InP/Si BiCMOS technology," *Microwave and Wireless Components Letters, IEEE*, vol. 24, pp. 47–49, Jan 2014.

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

[7] M. J. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S. C. Martin, R. Smith, *et al.*, "Submicron scaling of HBTs," *Electron Devices, IEEE Transactions on*, vol. 48, no. 11, pp. 2606–2624, 2001.

Chapter 2

InP Bipolar Transistor Design

In this chapter, the principle of operation of triple-mesa RF HBTs, essential device parameters and trade-offs between different design goals are presented.

2.1 Principle of Operation

An npn heterobipolar transistor structure is fabricated by epitaxially growing a wide bandgap n- collector, a narrow bandgap p+++ base and a wide bandgap n+ emitter in sequence. Ohmic contacts to emitter, base, and collector are formed. A band diagram of such structure under bias is shown in figure 2.1.

In forward-active operation mode, the base-collector diode is reverse-biased while the base-emitter diode is forward-biased. Electrons are swept vertically from the emitter into the base which is thinner than their diffusion length. Most electrons diffuse through the base with only a small fraction recombining with holes. The
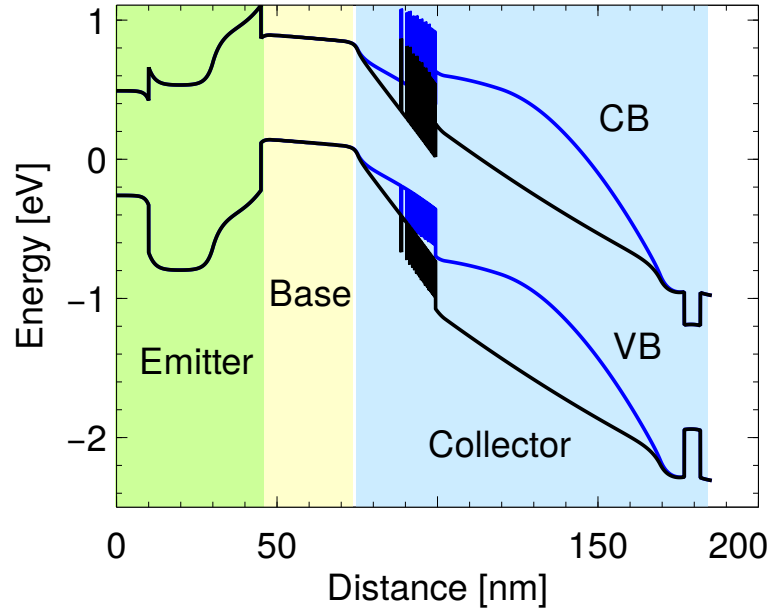
Figure 2.1: Band diagram of heterobipolar transistor under bias with valence band (VB) and conduction band (CB), showing bands with no (black) and Kirk (blue) current density in the collector.

electrons are then swept across the collector by the electric field of the reverse-biased junction. The electron concentration at the metallurgical base-emitter interface and subsequently the collector current can be modified by changing the (input) potential across the base-emitter diode: the collector current is approximately independent to changes of the base-collector potential (output), achieving transistor behavior.

Holes are confined to the base by the heterointerfaces: potential barriers in the valence band and differences in effective mass restrict holes from flowing to either emitter or collector, thereby suppressing parasitic hole currents.
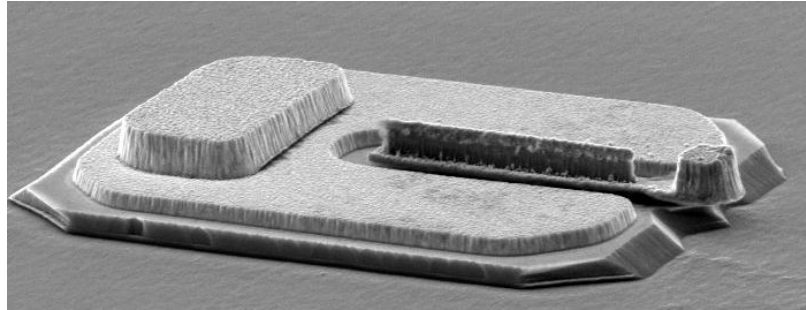
Figure 2.2: Angled scanning electron micrograph (SEM) of a fabricated heterojunction bipolar transistor prior to BCB planarization.

2.2 Device Topology

At UCSB, triple-mesa HBTs in the InP/InGaAs material system are researched. Emitter contacts are formed on a highly doped, low bandgap emitter cap InGaAs layer. The emitter is isolated, and base electrodes are deposited around the emitter in a self-aligned process: close spacing of base electrodes to active regions of the device $\approx 15\,\mathrm{nm}$ reduces gap resistance terms, thereby minimizing critical base access resistance. Base/collector mesas are formed in selective wet etches. The collector is contacted with a highly doped, thick and thus conductive subcollector. Although the subcollector has similar gap resistance terms associated to it, the conductivity of this layer is high enough to be mostly irrelevant for RF performance. A large area, non-self-aligned horseshoe-shaped contact to the subcollector is therefore formed. Devices are isolated in the third mesa etch. Figure 2.2 shows a scanning electron micrograph of a fabricated triple mesa HBT prior to planarization in a low-ε dielectric benzocyclobutene (BCB).

While current flowing across emitter contacts is swept vertically into the device,
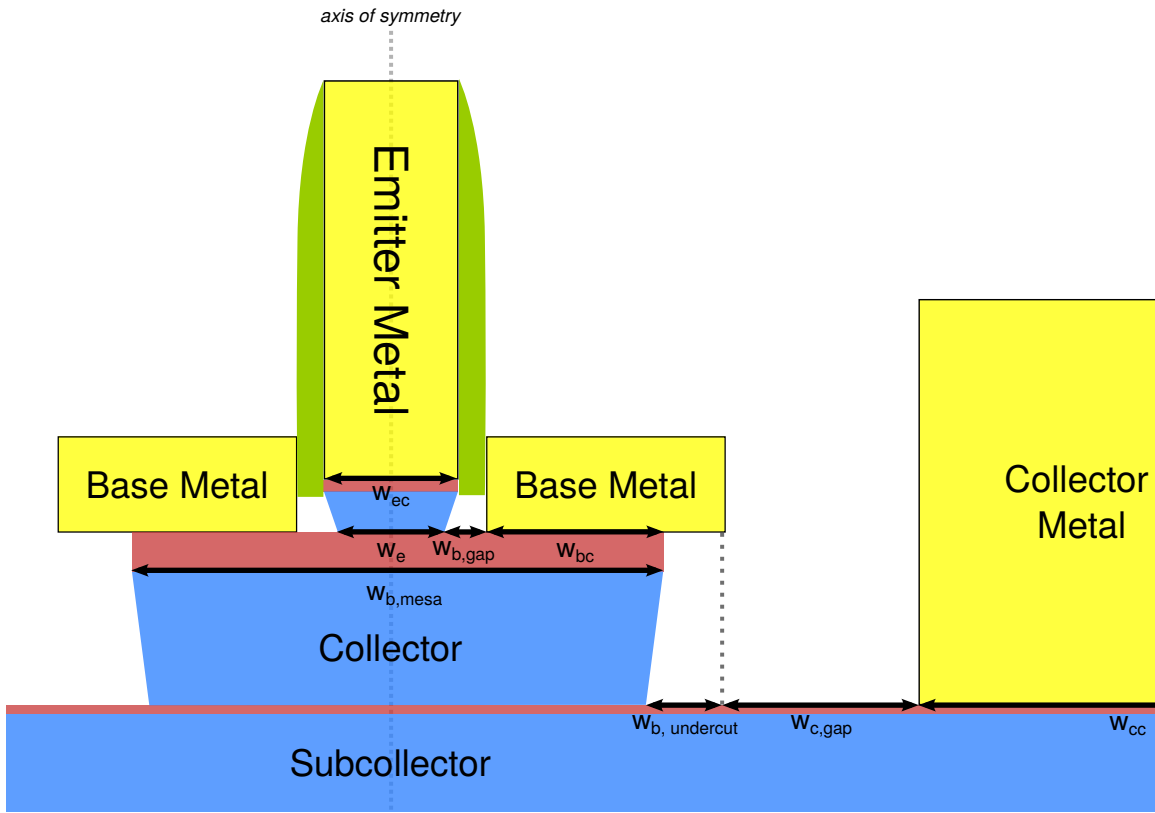
Figure 2.3: Cross-sectional schematic of a triple-mesa HBT with metal (yellow), low bandgap semiconductor (red) and wide bandgap semiconductor (blue). Key dimensions labeled.

current through planar base and subcollector contacts changes direction from vertical flow at the metal-semiconductor interface to horizontal flow into the device, imposing limitations on minimum access resistance that can be attained by enlarging metal contacts [1]. Key device dimensions are illustrated in figure 2.3: the emitter contact width w_{ec}, the emitter junction width w_e, the overlap of the base metal with the base semiconductor w_{bc}, the gap between base metal and active device $w_{b,\mathrm{gap}}$, the total width of the base/collector mesa $w_{b,\mathrm{mesa}}$, and the single-sided undercut of the base

mesa $w_{b,\mathrm{undercut}}$, the gap between base-collector mesa and collector contact $w_{b,\mathrm{undercut}}+$

$w_{c,\mathrm{gap}}$, and the extent of the collector metal w_{cc}.

2.2.1   Emitter Design

Crucial for high f_τ bandwidth is low emitter access resistivity. The emitter is

therefore capped with a highly doped layer of low bandgap material to enable low

resistivity ohmic contacts [2]. The InP emitter layers below contain a thin n+ region

which clamps the extent of the space charge region, followed the n- space charge

region itself.

Among processing considerations, the extent of the depletion zone t_e should be

chosen to find the optimum between low emitter-base junction capacitance $C_{je} \propto 1/t_e$

as well as low space charge zone resistance that adds to emitter access resistance:

$$\rho_{sc} = \frac{1}{q} \int\limits_{t_e} \frac{1}{\mu_n(z)\, n(z)}\, dz \tag{2.1}$$

with the elementary charge q, the electron mobility μ_n and local charge density $n(z)$.

The doping of of the space charge region n_{de} should be high enough to support

operation at and above Kirk current density $J_{e,Kirk} \approx 2 J_{c,Kirk}$ at which optimum

transport is achieved in the collector. If the doping of the depletion zone has been

chosen too low, injected carriers can screen the electric field in the space charge region

so the injection point for the emitter-to-base electron current is moved away from the

heterointerface: for any additional change in ∂V_{be}, the change in base current ∂I_b
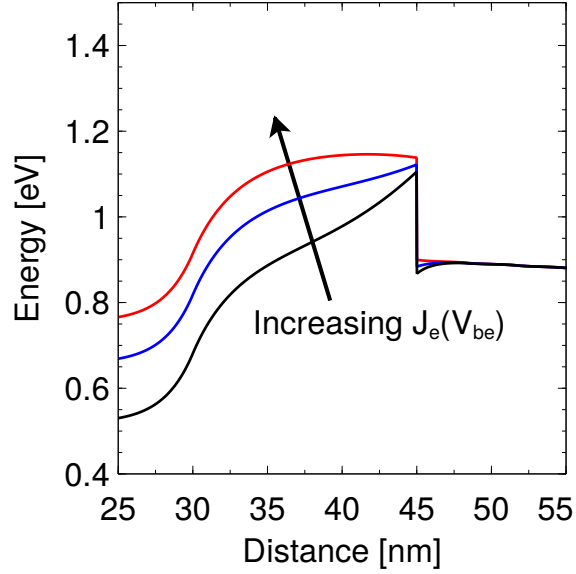
Figure 2.4: Conduction band profile at the abrupt emitter base junction for increasing $J_e(V_{be})$. As J_e is increased, a barrier for the electrons is formed in emitter that does not coincide with the metallurgical heterointerface anymore.

and subsequently collector current ∂I_c is no longer determined by energy difference $E_F - E_c$ at the heterointerface, but by maximum height of the barrier (see Figure 2.4). Transconductance $g_m = \partial I_c / \partial V_{be}$ and subsequently RF bandwidth is compromised in this operating regime.

At high current densities, Boltzmann carrier statistics remain no longer valid: degenerate Fermi-Dirac carrier statistics must be used to calculate current transport. Assuming specular conduction across the heterointerface (no reflection, unidirectional flow), it can be shown that the current density is [3]

$$J_{e,\text{Fermi}-\text{Dirac}} = \frac{q\,m^*}{2\pi^2\hbar^3}(k_B T)^2 \int\limits_0^\infty \frac{x}{1 + \exp(x - \eta_f)}\,dx, \qquad (2.2)$$

11

with the effective electron mass m^*, the reduced Planck constant $\hbar = h/2\pi$, the Boltzmann constant k_B, the junction temperature T, and normalized Fermi energy $\eta_f = E_f/k_B T$.

In the regime where Boltzmann statistics are valid, the emitter current density is

$$J_{e,\text{Boltzmann}} = \frac{q\, m^*}{2\pi^2 \hbar^3} (k_B T)^2 \exp(\eta_f) \tag{2.3}$$

.

A closed-form approximation for the transconductance normalized to emitter area as a function of relative Fermi level can be found [4]:

$$g_m = \frac{\partial I_c / A_e}{\partial V_{be}} \approx \frac{\partial J_e}{\partial V_{be}} = \frac{q\, m^*}{2\pi^2 \hbar^3} (k_B T) \ln \left[1 + \exp\left(\frac{E_f - E_c}{k_B T} \right) \right] \tag{2.4}$$

with the collector current I_c.

Figure 2.5 shows a plot of the transconductance as a function of emitter current density calculated from Boltzmann and Fermi-Dirac statistics at different junction temperatures. Transconductance can be improved at a given current density by using a material with higher effective density of states / higher effective mass m^*.

The transconductance of abrupt heterointerfaces is deteriorated by tunneling: a significant portion of electrons in the space charge region can tunnel through the triangular potential barrier and contribute to a parasitic collector current. This is reflected by high collector ideality $\eta_c > 1$ in Gummel characteristics and reduced
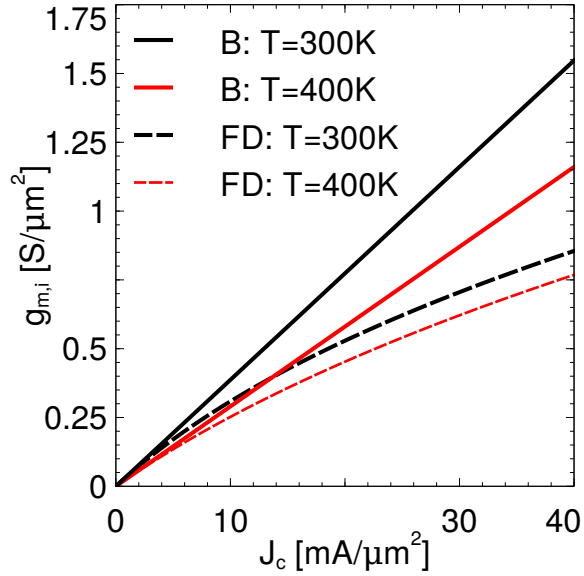
Figure 2.5: Transconductance calculated with Boltzmann (B) and Fermi-Dirac (FD) statistics at temperatures T=300 K and 400 K.

charging time $(C_{je} + C_{cb})/g_m$.

Quantum reflection of carriers at the heterojunction, barrier modulation effects and quasi-Fermi level drops diminish transconductance further at high current densities [5].

Elevated operating temperatures and high current densities necessitate thermally stable metal contacts that are impervious to electromigration. In the UCSB fabrication technology, an emitter metalization process for a composite refractory Mo/W/TiW metal stack has been established that is stable to current densities up to $60\,\mathrm{mA/\mu m^2}$ and can deliver a total access resistance to the emitter of less than $3\,\Omega\,\mu m^2$.

2.2.2 Base Design

The power gain cutoff frequency f_{max} is very sensitive to the resistivity of the ohmic contact between base electrode and semiconductor. Previous experiments on test samples have indicated that high base doping n_a is crucial for producing a low ohmic contact [6,7]. High doping, however, decreases current gain $\beta = \tau_n/\tau_b$ (electron/hole carrier lifetime τ_n/τ_b) mainly due to Auger recombination $\tau_n \propto n_{a,\text{effective}}^{-2}$. Simultaneous reduction of base thickness t_b is therefore required to maintain current gain when base doping is increased.

The base transit time and, by extension, current gain bandwidth f_τ can be enhanced by a quasi-electric field: the slope of the semiconductor energy bands is adjusted to improve electron transport by either grading the doping concentration or by varying the composition of the base semiconductor alloy throughout the base. While compositional grading decouples the quasi-electric field from the doping concentration, it introduces additional challenge of lattice-matching base semiconductor layers and has been therefore not used in this work.

The enhanced base transit time can be written as [8]

$$\tau_b = \frac{t_b^2}{D_n}\frac{k_B T}{\Delta E_C}[1 - \frac{k_B T}{\Delta E_C}(1 - \exp(-\frac{\Delta E_C}{kT}))] + \frac{t_b}{v_{\text{exit}}}\frac{k_B T}{\Delta E_C}(1 - \exp(\frac{\Delta E_C}{kT})), \quad (2.5)$$

with the conduction band slope ΔE_C, electron diffusivity D_n and exit velocity of minority carriers into the collector v_{exit}.

High doping concentrations in the base cause contraction of the bandgap [9] and subsequently modifications to the the conduction band slope. For accurate prediction of the base transit time, it must be therefore taken into account. High doping concentration also changes the lattice constant slightly: the In:Ga ratio is therefore adjusted during growth to ensure lattice match to InP.

In addition to improving base transit time, the quasi-electric field also drives electrons away from the base surface, thereby reducing base-emitter surface leakage and increasing current gain.

2.2.3   Collector Design

Low collector transit time is essential for attaining high f_τ: careful considerations must be therefore taken for designing the collector.

In epitaxial design of wafers presented in this work, the collector is comprised of a setback region, a superlattice grade, a pulse doping and a drift collector region. The superlattice chirped between InGaAs and InAlAs provides a smooth grading to the bands from the InGaAs base to the InP drift collector. The setback layer provides carriers with sufficient energy to traverse the grade. The pulse doping layer forms a dipole to restore fields across the graded region.

Optimum transport is attained at Kirk threshold current density: the charges comprising the collector current screen out the collector doping such that the electric

field at the base side of the collector is zero. This current can be written as

$$J_{c,\text{Kirk}} = \frac{2\varepsilon\varepsilon_0 v_{\text{eff}}}{t_c^2}(\varphi_{bi} + V_{cb}) + qn_c v_{\text{eff}}, \tag{2.6}$$

with the effective carrier velocity in the collector v_{eff} , the collector thickness t_c, the built-in potential φ_{bi} and the collector doping concentration n_c.

The collector should be fully depleted when no base-collector voltage is applied, $V_{cb} = 0$: this limits the maximum doping concentration to

$$n_{c,\text{max}} = \frac{2\varepsilon\varepsilon_0 \varphi_{bi}}{q\, t_c^2} \tag{2.7}$$

. For $n_c = n_{c,\text{max}}$, equation 2.6 can be rewritten as

$$J_{c,\text{Kirk}} = \frac{4\varepsilon\varepsilon_0 v_{\text{eff}}}{t_c^2}(\varphi_{bi} + V_{cb}) \tag{2.8}$$

.

Further limitation on the maximum doping concentration and subsequently the Kirk current density arises from the pulse doping layer [10].

In the Kirk regime, electrons sweep through the the first part of the collector in near flatband conditions without scattering. The effective carrier velocity as defined by the charge control model

$$\tau_c \equiv \frac{t_c}{2\, v_{eff}} \tag{2.9}$$

can exceed 3×10^7 cm/s in 100 nm thick InGaAs/InP collectors.

However, transport is severely degraded if only a small portion of electrons accumulates enough energy to scatter from low effective mass Γ into high effective mass L valleys (≈ 0.6 eV Γ-L separation for InP) [11].

The subcollector has a thin layer of n++ InGaAs to yield low resistivity ohmic contacts to the collector electrodes. A certain thickness of this layer is desirable to reduce sensitivity to contaminants that have accumulated on the sample surface from prior processing and overetching of the base/collector mesa. However, the heat conductivity of InGaAs is an order of magnitude worse than InP: a layer too thick would therefore thermally isolate the ambient substrate from the collector in which most of the heat is generated during device operation, causing degradation of carrier transport and early device failure.

2.3 TLM Structures

Transfer length method (TLM) structures enable the extraction of the contact resistivity between metal electrodes and semiconductor [1]. A set of metal pads with varying spacing { $w_{\mathrm{gap},1}$, $w_{\mathrm{gap},2}$, ... } is deposited onto the semiconductor surface. After fabrication, the resistance between each pair of adjacent pads $R(w_{\mathrm{gap}})$ is measured using four-terminal sensing at current densities which HBTs are operated at. From the set of measured resistances, the contact resistivity of the metal-semiconductor interface and the sheet resistance of the semiconductor in the gap can be extracted.

The resistance R of metal-semiconductor-metal structures as a function of gap spacing w_{gap} is:

$$R(w_{\text{gap}}) = 2R_c + R_{sh}\frac{w_{\text{gap}}}{L_{\text{pad}}}, \tag{2.10}$$

with the contact resistance R_c, the dimension of the pad L_{pad} transversal to the current flow and the sheet resistance of the semiconductor between the pads R_{sh}.

A 2D model has been developed that describes the potential distribution underneath the metal contacts and subsequently the total contact resistance R_c [12]. The similarity of the partial differential equations to those of transmission lines gave rise to the term *transmission line model*, also abbreviated as TLM. The contact resistance can be written as

$$R_c = \frac{\sqrt{R_{sh}\,\rho_c}}{2\,L_{\text{pad}}}\coth\left(\frac{w_{\text{pad}}}{L_t}\right) \tag{2.11}$$

with the sheet resistance of the semiconductor underneath the contact R_{sh}, the specific contact resistance ρ_c, the dimension of the pad transverse to the gap L_{pad} and the contact width w_{pad}. Within a transfer length

$$L_t = \sqrt{\rho_c/R_{sh}}, \tag{2.12}$$

the voltage has dropped to $1/e$ ($\approx 36.7\,\%$), and most of the current is passed through this section.

Three special cases are of interest:

- $w_{\text{pad}} > 3\,L_t$. The contact width is much larger than the transfer length. Equation 2.11 reduces to $(\lim_{x \to +\infty} \coth(x) = 1)$

$$R_c = \frac{\sqrt{R_{sh}\,\rho_c}}{L_{\text{pad}}} = \frac{\rho_c}{L_t\,L_{\text{pad}}} \tag{2.13}$$

The contact resistance is independent of contact pad width w_{gap}: a increase of contact pad width w_{gap} will therefore not reduce the contact resistance.

- $w_{bc} \ll L_t$. The contact width is much smaller than the transfer length. The contact resistance becomes area-limited, similar to metal-semiconductor contacts with current flow perpendicular to the interface. Equation 2.11 reduces to $(\coth(x) \approx 1/x \text{ for } |x| < 0.5)$

$$R_c = \frac{\rho_c}{L_{\text{pad}}\,w_{\text{pad}}} \tag{2.14}$$

- $w_{bc} \approx L_t$. The contact resistance can be approximated with the first two terms of the Laurent series: $\coth(x) \approx 1/x + x/3$ for $|x| \approx 1$. The values of two equivalent resistors $R_{c1} + R_{c2} = R_c$ are thus

$$R_{c1} = \frac{\rho_c}{L_{\text{pad}}\,w_{\text{pad}}}, \quad R_{c2} = \frac{R_{sh}\,w_{\text{pad}}}{3\,L_{\text{pad}}} \tag{2.15}$$

In the HBT process, two types of TLM structures are fabricated alongside transistors to monitor base contact resistivity: non-pinched (Figure 2.6) and pinched
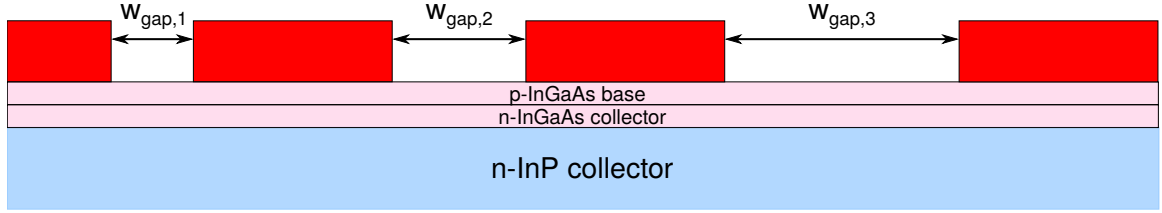
Figure 2.6: Non-pinched TLM structure with exposed base semiconductor surface.
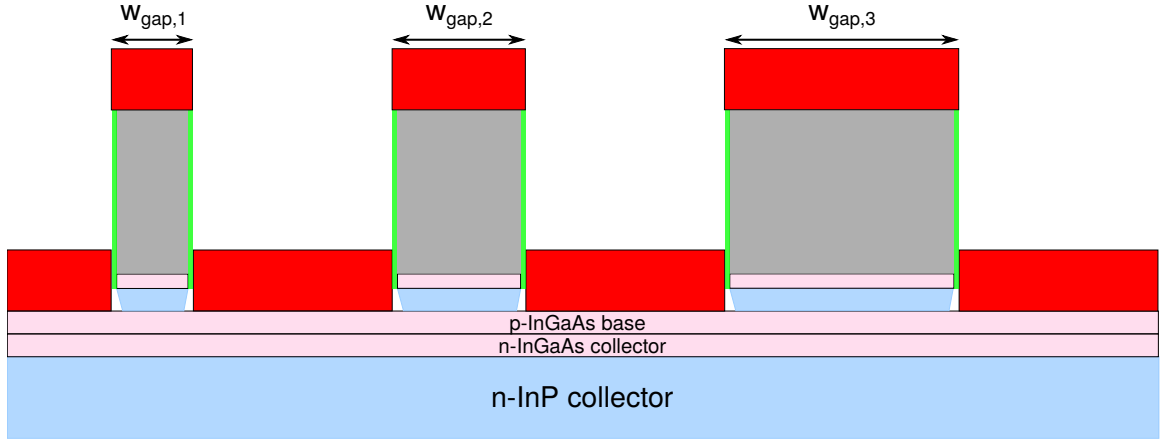
Figure 2.7: Pinched TLM structure with gaps defined by emitter metal.

(Figure 2.7). Non-pinched TLM structures have gaps defined in resist with the semiconductor surface in the gap exposed: the surface is therefore depleted and can suffer from process damage (e.g. oxidation in thermal processes) that can create surface states. The sheet resistance $R_{sh,\text{non}-\text{pinched}}$ extracted from measurements of non-pinched TLM structures is therefore higher than the sheet resistance underneath the metal pads:

$$R_{\text{non}-\text{pinched}}(w_{\text{gap}}) = 2R_c + R_{sh,\text{non}-\text{pinched}}\frac{w_{\text{gap}}}{L_{\text{pad}}}, \qquad (2.16)$$

The gap of pinched TLMs is defined by emitter stripes with the TLM pads de-

posited around the emitter gap in a self-aligned process: the base semiconductor remains encapsulated and has a sheet resistance $R_{sh,\text{pinched}}$ similar to the sheet resistance underneath the metal contacts. The measured resistance is

$$R_{\text{pinched}}(w_{\text{gap}}) = 2(R_c + R_{b,\text{undercut}}) + R_{sh,\text{pinched}} \frac{w_{\text{gap}}}{L_{\text{pad}}}. \tag{2.17}$$

The additional resistance $R_{b,\text{undercut}} = R_{sh,\text{non}-\text{pinched}}\, w_{\text{undercut}}/L_{\text{pad}}$ arises due to the gap between TLM electrodes and emitter semiconductor that results from the emitter semiconductor undercut and sidewall thickness.

With the contact width w_{gap} much larger than the transfer length, the contact resistance becomes (cf. equation 2.13)

$$R_c = \frac{\rho_c}{L_t L_{\text{pad}}} = \frac{\sqrt{R_{sh,\text{pinched}}\, \rho_{c,base}}}{L_{\text{pad}}} \tag{2.18}$$

The contact resistivity ρ_c can be therefore extracted with slope of equation 2.17 and the intersect of equation 2.16:

$$\rho_c = \frac{R_c^2 L_{\text{pad}}^2}{R_{sh,\text{pinched}}} \tag{2.19}$$

The measurements assume uniform contact resistivity across the area onto which the TLM pads are deposited. Further errors arise from uncertainty in the determination of gap spacing by SEM, non-uniform pad edges as a result of a lift-off process and

parasitic currents at pad corners that have been inadvertently rounded in lithographic

processes.

2.4   Equivalent Circuit Model

An equivalent hybrid-π model of a bipolar transistor is shown in figure 2.8. This

model is a first order approximation of the equivalent Tee circuit and has been sim-

plified by lumping RC networks corresponding to distributed capacitances and resis-

tances spread across the device, but it accurately represents the physical device in

the small signal regime and can be used to quantify essential device parameters. At

its core, a voltage controlled current source models the current gain of the transistor.

The remaining parameters account for various physical effects and will be described

in detail.

2.4.1   Emitter Access Resistance R_{ex}

The emitter access resistance represents the resistance that an electron encounters

while traversing from the top of the emitter metalization up to the metallurgical

emitter-base junction. It includes the emitter metal resistance

$$R_{em,\text{metal}} \approx \rho_{s,em,\text{metal}} \cdot \frac{T_{em,\text{metal}}}{A_{ec}}, \qquad (2.20)$$
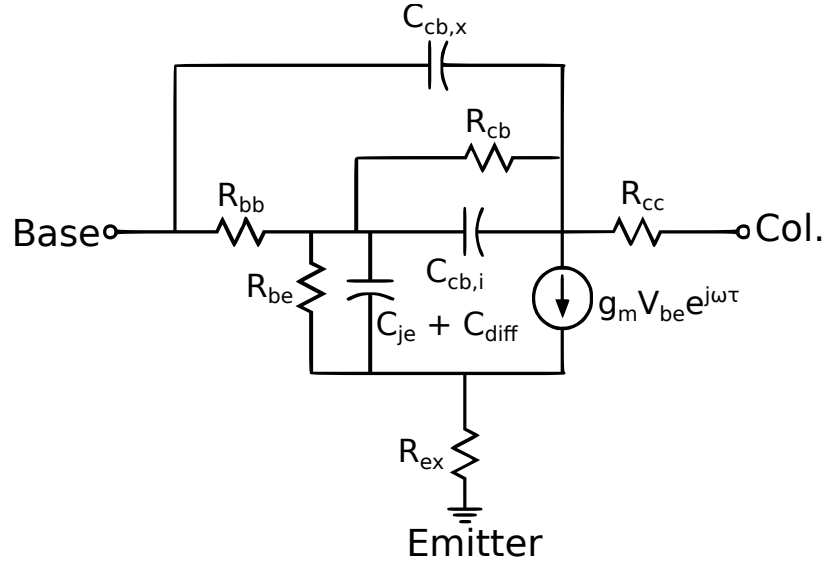
Figure 2.8: Hybrid-π model of a bipolar transistor.

with the specific resistivity of the emitter metal $\rho_{s,em,\mathrm{metal}}$, the emitter contact area $A_{ec} = L_e w_{ec}$, the emitter length L_e, and the emitter metal thickness $T_{em,\mathrm{metal}}$. An additional constituent is the emitter contact resistance

$$R_{con,em} = \frac{\rho_{em}}{A_{ec}} \tag{2.21}$$

with ρ_{em} the specific contact resistance between emitter metalization and semicon-ductor. The resistance of emitter space charge region of thickness t_{dep} adds a term [13]

$$R_{sc,em} = \frac{1}{q} \cdot \frac{\partial \Delta E_{fn}}{\partial I_e}, \tag{2.22}$$

with the emitter current I_e and the drop of the electron quasi-Fermi level

$$\Delta E_{fn} = \int\limits_{t_{dep}} \frac{J_e}{\mu_n(z)\, n(z)} dz, \qquad (2.23)$$

with the emitter current density $J_e = I_e/A_e$, the emitter area $A_e = L_e\, w_e$. The total emitter access resistance is the sum

$$R_{ex} = R_{em,\mathrm{metal}} + R_{con,em} + R_{sc,em}. \qquad (2.24)$$

Normalizing the emitter access resistance to the device area gives a device-independent figure that allows quantitative comparison of emitter access technologies:

$$\rho_{ex,xs} = R_{ex} \cdot A_{ec} \qquad (2.25)$$

A typical value for the upper limit of the sheet resistance of the composite emitter metal stack in the UCSB process is $0.8\,\Omega/\square$. For a transistor with $T_{em,\mathrm{metal}} = 500\,\mathrm{nm}$, the contribution from the normalized finite resistance of the emitter metal ranges below $0.3\,\Omega\,\mu\mathrm{m}^2$. For a well designed emitter, the normalized resistance of the emitter space charge region is below $0.1\,\Omega\,\mu\mathrm{m}^2$ [13]. The bulk of the emitter access resistance can therefore be attributed to the contact resistance which is $\approx 4\,\Omega\,\mu\mathrm{m}^2$ or less for the transistors discussed in this work.
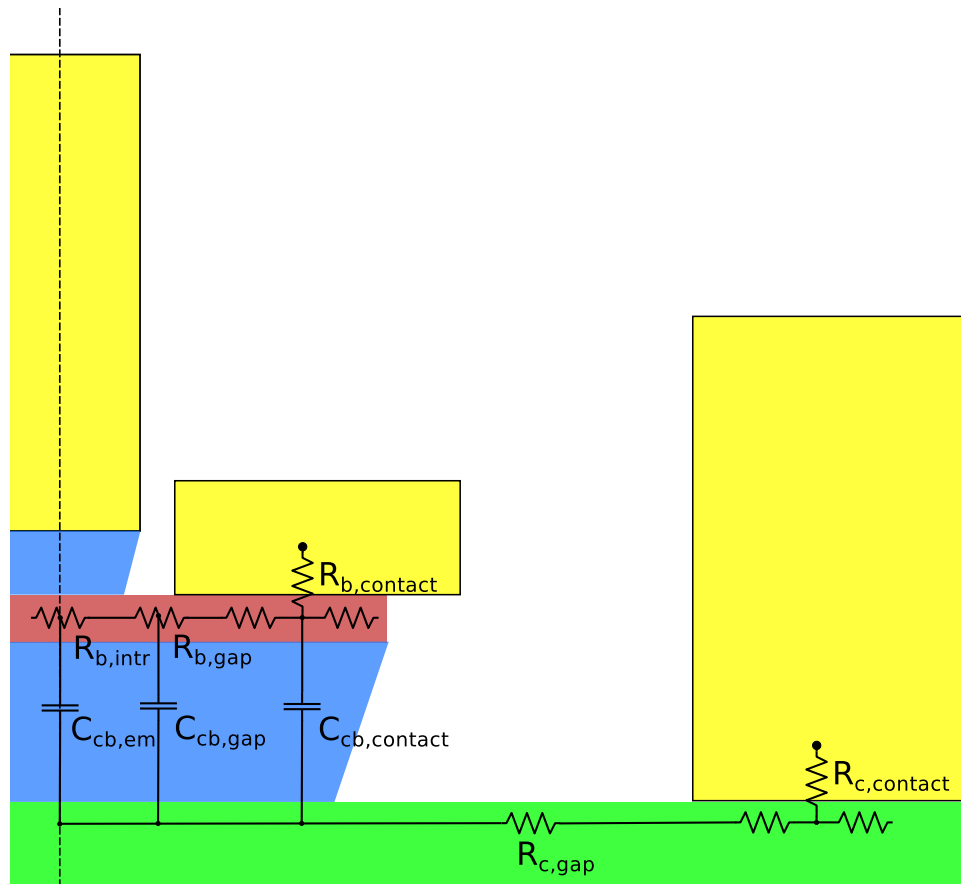
Figure 2.9: Equivalent circuit overlaid to a cross-sectional illustration of the active transistor across the emitter.

2.4.2   Base Access Resistance R_{bb}

Current from the base metal that surrounds both sides of the emitter enters the semiconductor vertically and then traverses horizontally to the active part of the device, encountering the sheet resistance of the metal, the specific contact resistance between metal and semiconductor and the sheet resistance of the semiconductor. The base access resistance consists thus out of multiple contributions (see Figure 2.9):
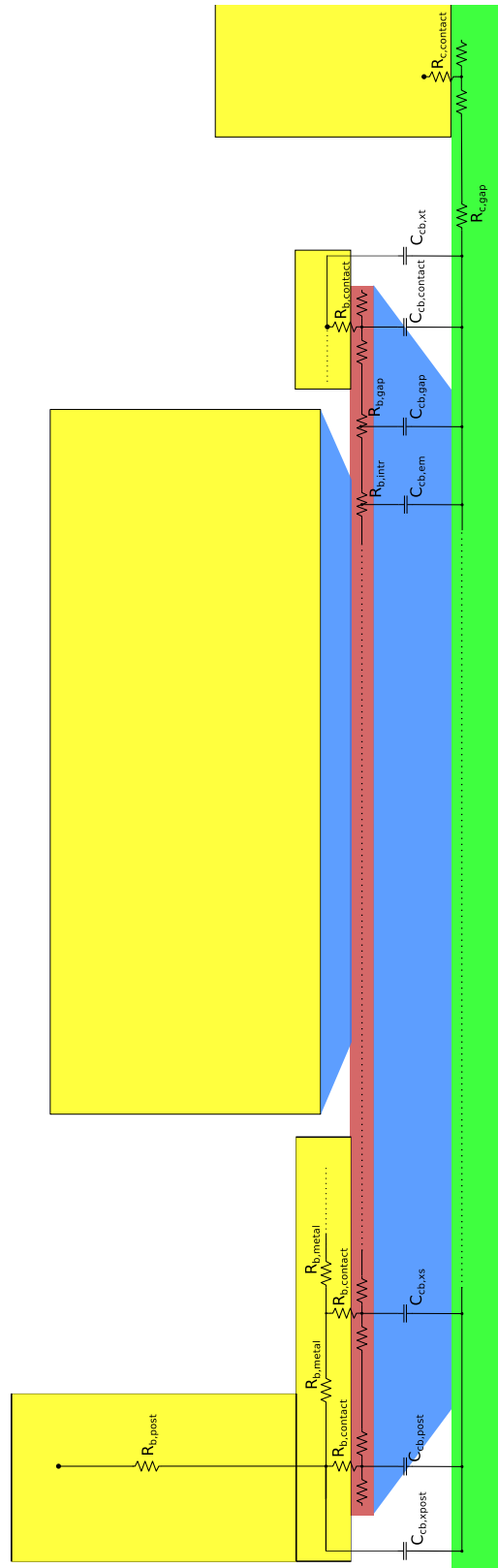
Figure 2.10: Equivalent circuit overlaid to a cross-sectional illustration of the device along the emitter.

Base Contact Resistance. The base contact resistance can be expressed as

$$R_{b,\text{contact}} = \frac{R_{sh,\text{base}}\,\rho_{b,\text{contact}}}{2L_e} \coth\left(\frac{w_{bc}}{L_t}\right), \tag{2.26}$$

with the single-sided overlap of base electrode and semiconductor w_{bc}, the contact resistivity $\rho_{b,\text{contact}}$, the sheet resistance underneath the contact $R_{sh,\text{base}}$, the transfer length L_t and emitter length L_e.

For the devices presented in this work, a typical value for the sheet resistance of the base semiconductor layer is $750\,\Omega/\square$, while the contact resistance is $4\,\Omega\,\mu\text{m}^2$. The transfer length is hence $70\,\text{nm}$.

Base Gap Resistance. The gap resistance reflects the sheet resistance of the exposed semiconductor between base metalization and the active device area underneath the emitter:

$$R_{b,\text{gap}} = R_{sh,\text{gap}}\frac{w_{b,\text{gap}}}{2\,L_e}, \tag{2.27}$$

with the sheet resistance of the exposed base semiconductor $R_{sh,\text{gap}} \approx R_{sh,\text{unpinched}}$ (equation 2.16).

Intrinsic Base Resistance. The resistance of the base semiconductor underneath the emitter can be expressed as [14]

$$R_{b,\text{intr}} = R_{sh}\frac{w_e}{12\,L_e}. \tag{2.28}$$

Base Metal Resistance. The resistance of the base metal contributes as

$$R_{b,\text{met}} = R_{sh,\text{bmet}} \frac{w_{bc}}{6\, L_e}, \tag{2.29}$$

with the sheet resistance of base metal $R_{sh,\text{bmet}}$.

The total base access resistance includes all contributions:

$$R_{bb} = R_{b,\text{contact}} + R_{b,\text{gap}} + R_{b,\text{intr}} + R_{b,\text{met}} \tag{2.30}$$

The sheet resistance of the base semiconductor can be obtained from TLM measurements (section 2.3) and verified numerically:

$$R_{sh} = \left(q \int_{0}^{t_{\text{base}}} \mu_p(p)\, p(x) dx \right)^{-1} \tag{2.31}$$

Fitting parameters for the doping dependent hole mobility $\mu_p(p)$ of p-InGaAs with carbon doping concentration n_a have been obtained from hall measurements of previous growths:

$$\mu_p(p) = 5448\, \frac{\text{cm}^2}{\text{V s}} - 107.3\, \frac{\text{cm}^2}{\text{V s}} \cdot \ln\left(n_a \cdot \text{cm}^3 \right) \tag{2.32}$$

The depletion depth of the exposed p-InGaAs base semiconductor can be calculated under the assumption that the Fermi level is pinned $\Delta E_c \approx 0.2\,\text{eV}$ below the
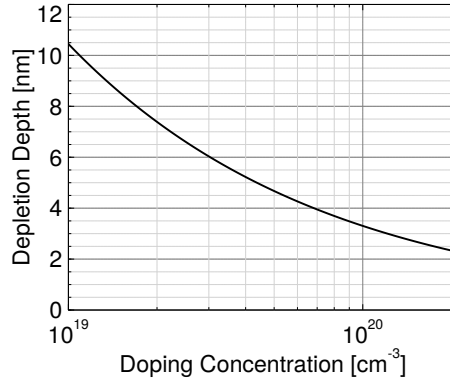
Figure 2.11: Depletion depth of p-InGaAs as a function of doping concentration assuming Fermi level pinning $0.2\,\text{eV}$ below the conduction band.

conduction band edge [15]. The depletion potential is

$$\varphi_{\text{dep}} = \frac{1}{q}\left[\,(E_v - E_F) + E_g - \Delta E_c\,\right],\tag{2.33}$$

with Fermi level offset to the valence band $E_v - E_F$, and the semiconductor band gap E_g.

Assuming Schottky boundary conditions, the depletion depth is

$$t_{\text{dep}} = \sqrt{\frac{2\,\varepsilon_0 \varepsilon_{\text{InGaAs}}\, \varphi_{\text{dep}}}{p}},\tag{2.34}$$

with the doping concentration p, the vacuum permittivity ε_0, and the relative dielectric permittivity of InGaAs $\varepsilon_{\text{InGaAs}}$.

For highly doped p-InGaAs $p \approx 9 \times 10^{19}\,\text{cm}^{-3}$, the Fermi level is $\approx 0.14\,\text{eV}$ below the valence band: $E_v - E_F \approx 0.14\,\text{eV}$. The depletion depth as a function of the initial doping is plotted in figure 2.11.
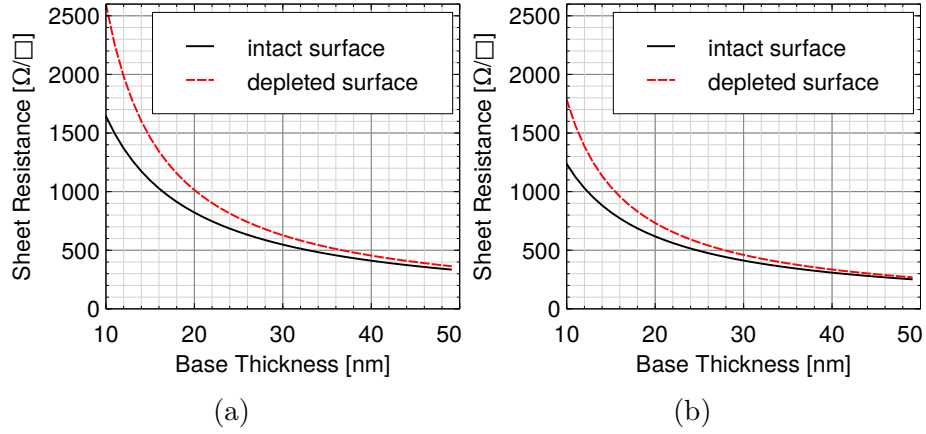
Figure 2.12: Sheet resistance of p-InGaAs semiconductor doped from (a) $9 \times 10^{19}/\text{cm}^3$ to $4 \times 10^{19}/\text{cm}^3$ and (b) $12 \times 10^{19}/\text{cm}^3$ to $8 \times 10^{19}/\text{cm}^3$ as a function of layer thickness. Solid: no surface depletion. Dotted: surface depleted.

With the depletion depth, a lower boundary for the sheet resistance of exposed base semiconductor can be calculated:

$$R_{sh,\text{exposed}} = \left(q \int_{t_{dep}(p)}^{T_{\text{base}}} \mu_p(p(x))\, p(x) dx \right)^{-1} \tag{2.35}$$

The resistivity calculated from above equation does not include surface damage due to subsequent processing: the sheet resistance extracted from unpinched TLM measurements is therefore higher.

2.4.3 Collector Access Resistance R_{cc}

Similar to the base access resistance, the collector access resistance is comprised of multiple contributions:

Collector Contact Resistance.   The contact resistance between the collector met-

alization and the subcollector semiconductor can be expressed as:

$$R_{c,\text{contact}} = \frac{R_{sh,c}\,\rho_{c,\text{contact}}}{2L_e} \coth\left(\frac{w_{cc}}{L_t}\right), \qquad (2.36)$$

with the specific contact resistance $\rho_{c,\text{contact}}$, single-sided collector contact width w_{cc},

sheet resistance of the subcollector semiconductor $R_{sh,c}$, and the collector transfer

length $L_t = \sqrt{\rho_{c,\text{contact}}/R_{sh,c}}$. In the UCSB process, the contact is much wider than

the transfer length $W_{cc} > L_t$ ($R_{sh,c} \approx 16.5\,\Omega/\square$, $\rho_{c,\text{contact}} \approx 10\,\Omega\,\mu m^2$, $W_{cc} \approx 2\,\mu m$) so

the expression can be simplified:

$$R_{c,\text{contact}} = \frac{\rho_{c,\text{contact}}}{2\,L_t\,L_e} \qquad (2.37)$$

Surface damage to the subcollector can be neglected for the calculation of the

sheet resistance because it is sufficiently thick (300 nm).

Subcollector Resistance.   The resistance of the subcollector semiconductor layer

between collector metal contacts and base/collector mesa is

$$R_{c,\text{gap}} = R_{sh,c}\frac{w_{c,\text{gap}}}{2\,L_e}, \qquad (2.38)$$

with the gap between collector metalization and device mesa $w_{c,\text{gap}}$.
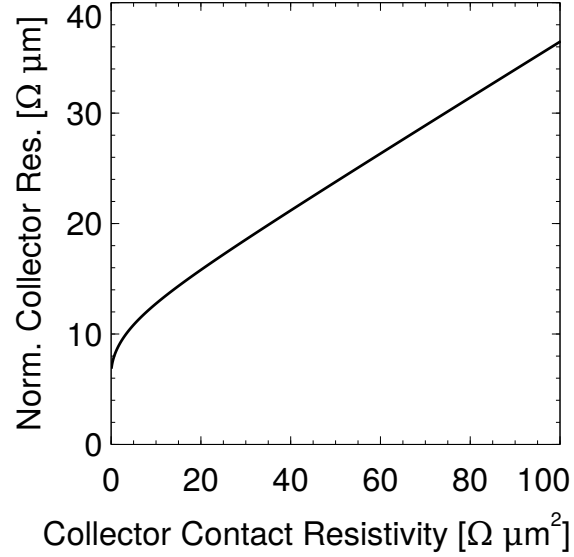
Figure 2.13: Normalized collector resistance $R'_{cc} = R_{cc} L_e$ as a function of specific contact resistance $\rho_{c,\text{contact}}$ assuming $R_{sh} = 16\,\Omega/\square$, $w_{cc} = 2\,\mu\text{m}$, $w_{c,gap} = 750\,\text{nm}$, $w_{mesa} = 260\,\text{nm}$.

Intrinsic Resistance. The resistance of the subcollector semiconductor underneath the base/collector mesa is

$$R_{c,\text{intr}} = R_{sh,c}\frac{w_{\text{mesa}}}{12\,L_e}. \tag{2.39}$$

The collector access resistance is the sum:

$$R_{cc} = R_{c,\text{contact}} + R_{c,\text{gap}} + R_{c,\text{intr}} \tag{2.40}$$

Figure 2.13 plots the total collector access resistance normalized to emitter length as a function of contact resistance.

2.4.4 Base-Collector Capacitance C_{cb}

Multiple components contribute to the capacitance of the base-collector region C_{cb} (cf. Figure 2.9): the contribution from in the active device region is the sum of the capacitance underneath the base metalization $C_{cb,\text{cont}}$, the capacitance in the gap between base metal and the active device part $C_{cb,\text{gap}}$, the capacitance underneath the emitter $C_{cb,em}$, and finally the fringing capacitance below and around the undercut base contact $C_{cb,xt}$:

$$C_{cb,\text{cont}} = 2\varepsilon_0\,\varepsilon_{SC}\,\frac{w_{bc}\,L_e}{t_c}, \qquad\qquad C_{cb,\text{gap}} = 2\varepsilon_0\,\varepsilon_{SC}\,\frac{w_{b,\text{gap}}\,L_e}{t_c}, \qquad (2.41)$$

$$C_{cb,em} = \varepsilon_0\,\varepsilon_{SC}\,\frac{w_{b,\text{gap}}\,L_e}{t_c}, \qquad\qquad C_{cb,xt} = 2\gamma_0\,\varepsilon_0\,\varepsilon_{BCB}\,\frac{w_{b,\text{undercut}}\,L_e}{t_c}, \qquad (2.42)$$

with $1 < \gamma_0 < 1.5$ a factor accounting for the fringing fields, the dielectric permittivity ε_0, the effective permittivity of the base/collector semiconductor ε_{SC} and the collector thickness t_c.

Additionally, the regions underneath the base post, between base post and emitter and at the emitter end (see Figure 2.10) add to C_{cb}:

$$C_{cb,\text{post}} = \varepsilon_0\,\varepsilon_{SC}\,\frac{A_{BP}}{t_c}, \qquad\qquad C_{cb,xs} = \varepsilon_0\,\varepsilon_{SC}\,\frac{w_{xs}\,L_{xs}}{t_c}, \qquad (2.43)$$

$$C_{cb,\text{xpost}} = \gamma_0\,\varepsilon_0\,\varepsilon_{BCB}\,\frac{(L_{BP} - 2\,w_{b,\text{undercut}})\,w_{b,\text{undercutT}} + w_{\text{undercutT}}L_{BP}}{t_c}, \qquad (2.44)$$

with $1 < \gamma_1 < 1.5$ a factor accounting for the fringing fields, the area of the base post A_{BP}, the width w_{xs} and length L_{xs} of the region between base post and the active device, the base post diameter d_{BP} and the base post undercut $w_{\mathrm{undercutT}}$.

The total base-collector capacitance is therefore comprised out of a contribution that is independent of the emitter length L_e and a part that is proportional to the emitter length.

For the hybrid-π model, the base-collector capacitance is partitioned into an intrinsic part $C_{cb,i}$ and an extrinsic part $C_{cb,x}$: the intrinsic part is derived from the charging time constant τ_{cb} that links f_τ to f_{max}.

2.4.5   Base-Emitter Resistance R_{be}

The base-emitter resistance represents the differential resistance at the bias point:

$$R_{be} = \left. \frac{\partial V_{be}}{\partial I_e} \right|_{V_{be}}, \tag{2.45}$$

with the emitter-base voltage V_{be} and the emitter current I_e.

Under the assumption that the current gain $\beta = I_c/I_b$ is independent of V_{be} ($\partial \beta / \partial V_{be} = 0$), and β is sufficiently large ($I_c \approx I_e$), the resistance can be written as

$$R_{be} = \frac{\beta}{g_m}, \tag{2.46}$$

with the transconductance $g_m = \partial I_c / \partial V_{be}$.

2.4.6   Base-Collector Resistance R_{cb}

R_{cb} has unclear physical correspondence, but is used to accurately fit measured Y_{21} data.

2.4.7   Base-Emitter Capacitance C_{be}

The base-emitter capacitance is comprised of two contributions: the fictitious diffusion capacitance is due to carrier charge storage in forward-operation mode

$$C_{\text{diff}} = \frac{\partial I_e}{\partial V_{be}}(\tau_b + \tau_c) \approx g_m(\tau_b + \tau_c). \tag{2.47}$$

. The capacitance associated with charges separated by the emitter space charge zone and emitter sidewalls is C_{je}. The total base-emitter capacitance is therefore

$$C_{be} = C_{\text{diff}} + C_{je} \tag{2.48}$$

.

2.5   Figures of Merit

With the collector short-circuited, it is found that the frequency-dependent current gain has single pole form

$$h(f) = \frac{h_0}{1 + jf/f_{\text{3dB}}}, \tag{2.49}$$
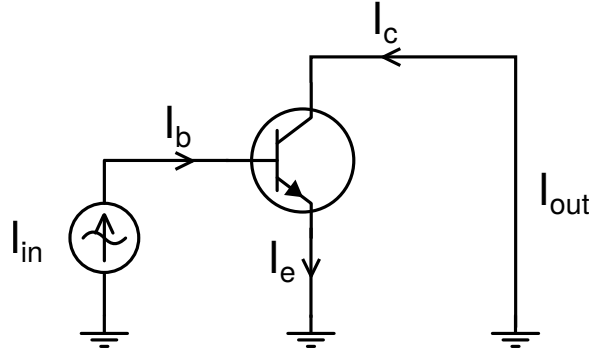
Figure 2.14: Bipolar transistor with short-circuited output biased with base current source.

with $f_{3\text{dB}}$ the frequency at which $|h(f_{3\text{dB}})| = h_0/\sqrt{2}$.

The current gain cutoff frequency f_τ is the frequency at which current gain reaches unity, $|h(f_\tau)| = 1$:

$$f_\tau = \sqrt{(h_0^2 - 1)}f_{3\text{dB}} \approx h_0\, f_{3\text{dB}} \tag{2.50}$$

From nodal analysis of the hybrid π-circuit in figure 2.8, an expression for the current gain cutoff frequency can be found:

$$\frac{1}{2\pi f_\tau} = \tau_{ec} = \tau_b + \tau_c + \left(\frac{\eta k_B T}{q I_c}\right)C_{je} + \left(\frac{\eta k_B T}{q I_c} + R_{ex} + R_c\right)C_{cb} \tag{2.51}$$

The current gain cut-off frequency can be easily obtained from RF measurements and allows extraction of intrinsic device parameters, e.g. transit times and transconductance.

At frequencies above current gain cutoff, voltage and therefore power gain can still be achieved in an amplifier design. This gives rise to the definition of the maximum

frequency of oscillation f_max: at this frequency, all power gains (maximum stable / maximum available / unilateral) of the device become unity. In practice, f_max can be extracted from unilateral gain U since it has a functional form identical to $|h(f)|$ (cf. equation 2.49), exhibiting $-20\,\text{dB}$ per decade roll-off beyond f_3dB.

The power gain cutoff frequency f_max is linked to the current gain cutoff frequency f_τ with a time constant τ_{cb} associated to the charging time of a distributed RC network in the base-collector region [16]:

$$f_{max} = \sqrt{\frac{f_\tau}{8\pi\tau_{cb}}} = \sqrt{\frac{f_\tau}{8\pi R_{bb}C_{cb,eff}}}. \tag{2.52}$$

In figure 2.9, the schematic of an equivalent circuit is overlaid over a cross-section of a transistor. With the method of time constants, the charging time τ_{cb} of this cross-sectional model can be determined:

$$\begin{aligned}
\tau_{cb} =&\, C_{cb,\text{contact}}\frac{\rho_c}{w_{bc}L_e}+ \\
& C_{cb,\text{gap}}\left(R_{b,\text{contact}}+\frac{1}{2}R_{b,\text{gap}}\right)+ \\
& C_{cb,\text{em}}\left(R_{b,\text{contact}}+R_{b,\text{gap}}+\frac{1}{2}R_{b,\text{intr}}\right)
\end{aligned} \tag{2.53}$$

In the calculation of the cross-sectional model, perfect conductance of the base electrode along the length of the emitter is assumed. However, in a real transistor, the finite conductance will degrade f_max bandwidth particularly on long emitter length
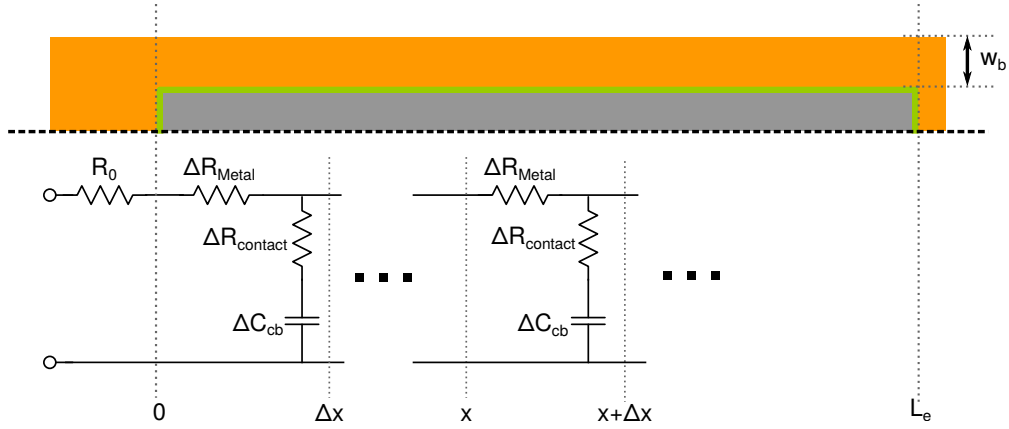
Figure 2.15: Circuit schematics for estimating charging time constant τ_{cb} due to voltage drops along the emitter stripe.

devices (Figure 2.10) . A simplified calculation will be used to quantify this effect.

Under the assumption that the gap resistance $R_{b,\text{gap}}$ and the intrinsic semiconductor resistance $R_{b,\text{intr}}$ are negligible, and that the base contact width is less than the transfer length (cf. equation 2.14), equation 2.53 can be rewritten as:

$$\tau_{cb,\text{approx}} := C_{cb,\text{approx}} R_{\text{contact}} = C_{cb,\text{approx}} \frac{\rho_c}{w_{bc} L_e} \tag{2.54}$$

The approximate collector-base capacitance can now be linked to physical dimensions:

$$C_{cb,\text{approx}} = \frac{\varepsilon_0 \varepsilon_{CB} \left(w_{bc} + w_{\text{gap}} + w_{em}/2 \right)}{t_c} \tag{2.55}$$

with the effective dielectric constant of the base/collector semiconductor ε_{CB} and dimensions as shown in Figure 2.3.

The total charging time constant for a simplified equivalent circuit reflecting the

base sheet resistance $R_{\mathrm{metal}}(L) = R_{sh,\mathrm{metal}} L/w_b$ is

$$\tau_{cb,\mathrm{3D}} = \int\limits_{0}^{L_e} \left(R_0 + R_{\mathrm{metal}}(L) + dR_{\mathrm{contact}}\right) dC_{cb,\mathrm{approx}} \tag{2.56}$$

RC delays associated to the region between emitter and base post, emitter end and fringing fields are omitted in this estimation.

The integration yields three terms $\tau_{cb,\mathrm{3D}} =: \tau_{cb,0} + \tau_{cb,1} + \tau_{cb,2}$. The first term $\tau_{cb,0}$ is identical to equation 2.54:

$$\tau_{cb,0} = \frac{\varepsilon_0 \varepsilon_{BC} \rho_c}{t_c} \frac{w_{bc} + w_{gap} + w_e/2}{w_{bc}} = \tau_{cb,\mathrm{approx}} \propto L_e^0 \tag{2.57}$$

The second term is proportional to the emitter length L_e, describing a charging delay due to the full device capacitance $C_{cb,\mathrm{approx}}$ charged through base post resistance R_0:

$$\tau_{cb,1} = \frac{\varepsilon_0 \varepsilon_{BC} R_0 L_e}{t_c} \frac{w_{bc} + w_{gap} + w_e/2}{w_{bc}} = R_0 C_{cb,\mathrm{approx}} \propto L_e^1 \tag{2.58}$$

The third term has a quadratic dependency on the emitter length L_e:

$$\tau_{cb,2} = \frac{\varepsilon_0 \varepsilon_{BC} R_{sh,\mathrm{metal}} L_e^2}{2t_c} \frac{w_{bc} + w_{gap} + w_e/2}{w_b} \propto L_e^2 \tag{2.59}$$

Comparing equations 2.59 to 2.57, a fictitious contact resistance due to finite base
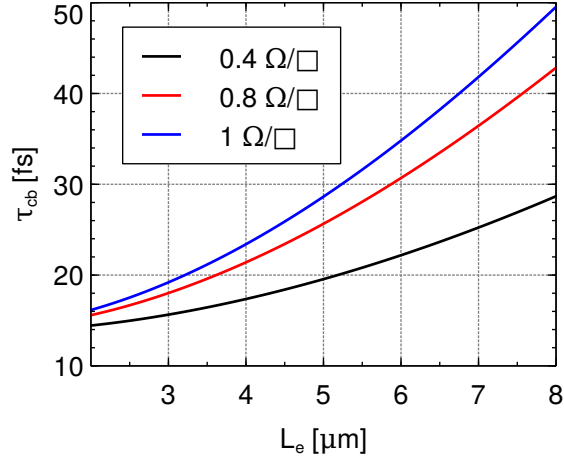
Figure 2.16: Numerical calculations of charging delay τ_{cb} as a function of emitter length L_e for different base metal sheet resistance R_{sh}. A finite element model has been used that has been matched to a fabricated transistor.

metal sheet resistance can be defined:

$$\rho_{b,\text{metal}} := \frac{w_{bc}}{w_b} \frac{R_{sh,\text{metal}} L_e^2}{2} \tag{2.60}$$

When the base sheet resistance is too high or there is insufficient base/collector mesa undercut, $w_b \approx w_{bc}$, this additional charging delay can severely limit RF bandwidth. Figure 2.16 shows numerically calculated τ_{cb} as a function of emitter length for a set of different base metal sheet resistances R_{sh}. The model described in section 2.6 has been used for these calculations. The quadratic increase of τ_{cb} with L_e confirms the trends predicted by above calculations.

With the results of the distributed circuit model, a geometry-dependent fitting factor $c \approx 1.5$ that depends on the different relative composition of the base-collector parasitics can be found to match to the fictitious contact resistivity when assumptions

made for above derivations have been violated:

$$\rho_{b,\text{metal}} = \frac{w_{bc}}{w_b} \frac{R_{sh,\text{metal}} L_e^2}{c} \tag{2.61}$$

2.6   Distributed Circuit Model

In order to evaluate effects of processing issues on RF bandwidths and to accurately extract key device parameters from fabricated devices, a finite-element circuit model has been developed that reflects the distributed nature of parasitics. The values of all circuit elements used in this model are derived from three sets of parameters: the first set is comprised of parameters specific to the epitaxial wafer design, e.g. collector thickness t_c, normalized transconductance g_m/A_{ej}, et cetera. The second set encompasses parameters that have been realized in the fabrication of the sample, e.g. contact resistivities, the base-collector undercut, the end undercut of emitter stripes, et cetera. The third set specifies the dimensions of a transistor.

Figure 2.17 shows a simplified circuit model along the length of the transistor (compare to Figure 2.10). The left section of the circuit represents the parasitics associated with the base post regions. The active device is partitioned into N slices of L_e/N length of the 2D transistor model shown in 2.18. The base terminals of the slices are connected through resistors that are associated to the base electrode resistance. Emitter end undercut is modeled by placing the 2D transistor model of
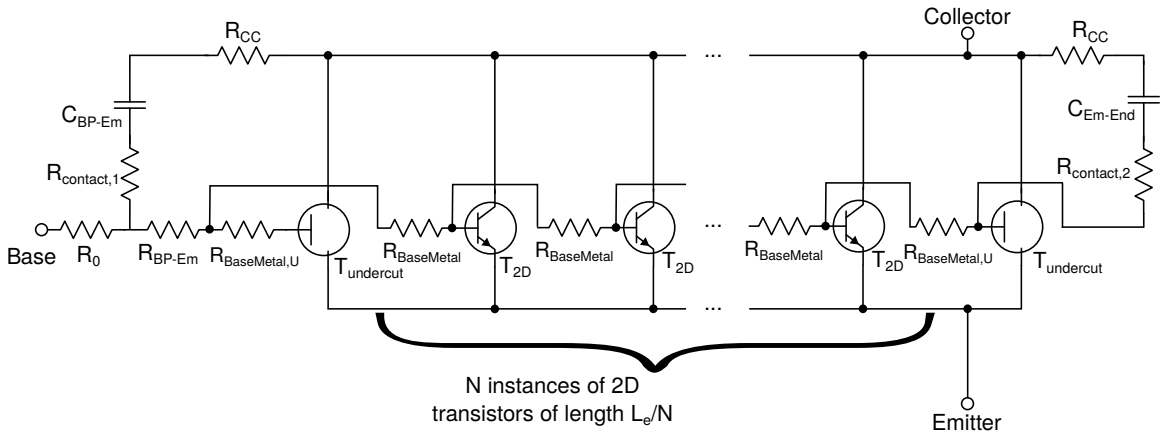
Figure 2.17: Finite element circuit model along the length of the transistor.
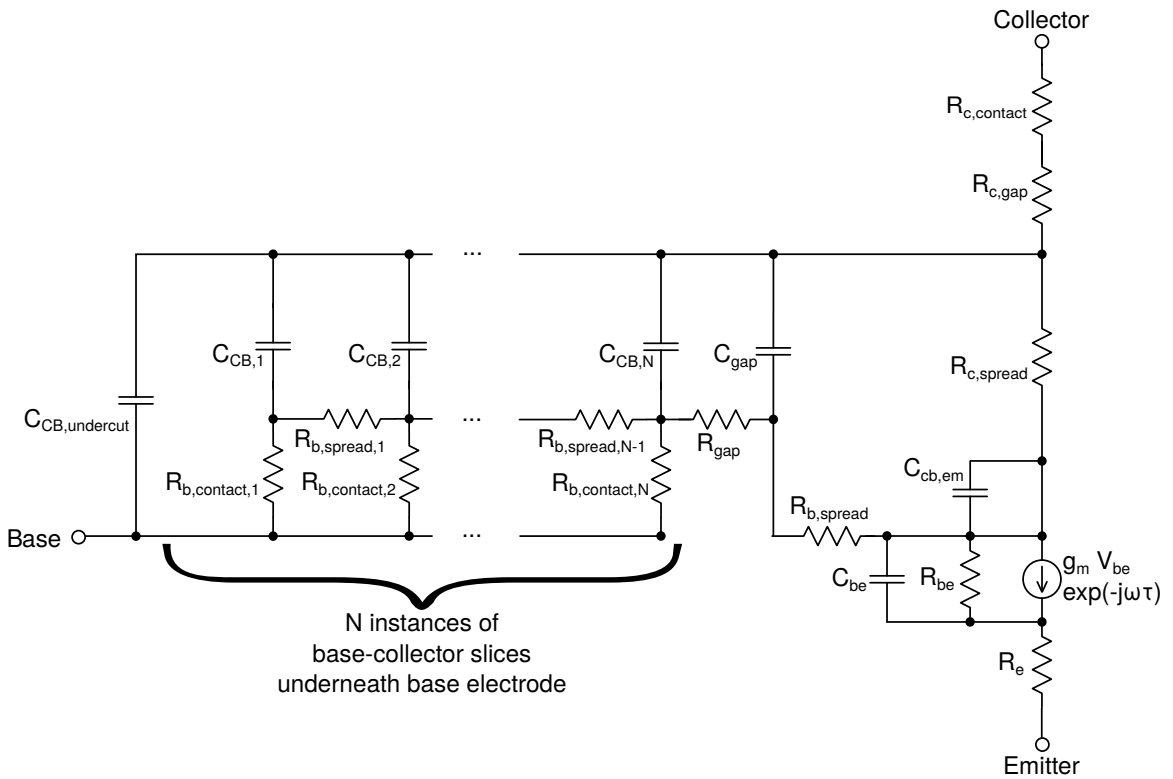
Figure 2.18: Finite element circuit model along the width of the transistor T_{2D}.

length L_{undercut} at both ends of the emitter stripe with the active part removed. The parasitics of the emitter end opposite to the base post is reflected by $R_{\text{contact,2}}$ and $C_{\text{Em-End}}$.

The 2D transistor model shown in Figure 2.18 is mostly identical to the model derived in [10]. Only the meshing of the base electrode regions has been made denser, and the transistor model has been substituted with a voltage-controlled current source parallel to a resistor R_{be}.

2.7   Scaling Laws

RF bandwidth is closely tied to RC charging delays and transit times as discussed above. For improving bandwidth by a scaling factor of γ, all transit and charging delays must be reduced by γ while maintaining current density I_e/L_e and resistances [17]. This is achieved by epitaxially thinning the base t_b by γ^c ($0.5 < c < 1$, [18,19]) and the collector t_c by γ, assuming constant effective carrier velocities across scaling generations. This doubles capacitances and hence RC delays, necessitating a reduction of emitter and base/collector area by γ^2 for desired reduction of C by γ. This is done by shrinking emitter w_e and base/collector mesa widths for reasons of heat dissipation [20]. For maintaining device access resistance R, contact resistivities must also be scaled by $1/\gamma^2$.

High base doping n_a is key for low contact resistivity $\rho_{c,\text{base}}$. However, high doping also degrades current gain due to Auger recombination $\propto n_a^{-2}$. Moreover, the

Design Parameter	Scaling Law
Collector Depletion Layer Thickness t_c	$\gamma^{-1} : 1$
Base Thickness t_b	$\gamma^{-\frac{1}{2}} : 1$
Emitter-base junction width t_e	$\gamma^{-2} : 1$
Base-collector junction width w_{bj}	$\gamma^{-2} : 1$
Emitter access resistivity $\rho_{c,\mathrm{emitter}}$	$\gamma^{-2} : 1$
Base contact resistivity $\rho_{c,\mathrm{base}}$	$\gamma^{-2} : 1$
Emitter current density J_e	$\gamma^{2} : 1$
Emitter length L_e	$\gamma : 0$

Table 2.1: Approximate HBT scaling laws .

perimeter-to-area ratio increases with γ, exacerbating emitter-base surface leakage currents [21]. Maintaining current gain requires therefore scaling of the base thickness beyond $\gamma^{0.5}$. A technology that decouples the base doping of the intrinsic base from the doping of the base semiconductor to which ohmic contacts are formed could potentially alleviate this constraint. At the same time, the conductivity of the base electrode must be improved by γ^4 as a consequence of simultaneous reduction of base width by γ^2 and base contact resistivity by γ^2.

Kirk current density in the collector is increased by γ^c, $1.5 < c < 2$ (equation 2.8): while the collector thickness is reduced by γ and $J_{\mathrm{Kirk}} \propto t_c^{-2}$, the bias voltage of the collector-base diode V_{cb} is also reduced so current densities scale with less than γ^2. Emitter contacts must sustain current densities in excess of $J_{e,\mathrm{Kirk}} = k J_{c,\mathrm{Kirk}}$ with the current spreading factor $k \approx 2$. In first order approximation, transconductance g_m stays constant for constant I_c. However, degeneracy effects will degrade g_m at

increased operating current densities. A thinner base and steeper bandgap grade as a result of higher quasi-electric field reduces current spreading in the base and collector, alleviating emitter current density scaling, i.e. k can become smaller than 2.

To first order, the electric field in the collector required to initiate avalanche breakdown will decrease by γ^3 with collector doping increased by γ^2 (equation 2.7) and collector thickness reduced by γ. Despite reduced bias voltage, scaling will thus reduce the safe range of operation.

References

[1] H. Berger, "Models for contacts to planar devices," *Solid-State Electronics*, vol. 15, no. 2, pp. 145 – 158, 1972.

[2] A. K. Baraskar, M. A. Wistey, V. Jain, U. Singisetti, G. Burek, B. J. Thibeault, Y.-J. Lee, A. Gossard, and M. J. Rodwell, "Ultralow resistance, nonalloyed ohmic contacts to n-InGaAs," *Journal of Vacuum Science Technology B: Microelectronics and Nanometer Structures*, vol. 27, pp. 2036–2039, Jul 2009.

[3] R. Tsu and L. Esaki, "Tunneling in a finite superlattice," *Applied Physics Letters*, vol. 22, pp. 562–564, Jun 1973.

[4] N. Machida, Y. Miyamoto, and K. Furuya, "Minimum emitter charging time for heterojunction bipolar transistors," in *Indium Phosphide and Related Materials Conference Proceedings, 2006 International Conference on*, pp. 325–328, 2006.

[5] V. Jain and M. J. Rodwell, "Transconductance degradation in near-THz InP double-heterojunction bipolar transistors," *Electron Device Letters, IEEE*, vol. 32, pp. 1068–1070, Aug 2011.

[6] A. Baraskar, V. Jain, M. A. Wistey, E. Lobisser, B. J. Thibeault, A. C. Gossard, and M. J. Rodwell, "In-situ and ex-situ ohmic contacts to heavily doped p-InGaAs," in *16th Int. Conf. Molecular Beam Epitaxy*, 2010.

[7] J. Lin, S. Yu, and S. Mohney, "Characterization of low-resistance ohmic contacts to n-and p-type InGaAs," *Journal of Applied Physics*, vol. 114, no. 4, p. 044504, 2013.

[8] H. Krömer, "Zur Theorie des Diffusions- und des Drifttransistors," *Arch. Elektr. Uebertrag.(AEUe)*, vol. 8, pp. 223–228, 1954.

[9] J. Lopez-Gonzalez and L. Prat, "The importance of bandgap narrowing distribution between the conduction and valence bands in abrupt HBTs," *Electron Devices, IEEE Transactions on*, vol. 44, pp. 1046–1051, Jul 1997.

[10] Z. M. Griffith, *Ultra high speed InGaAs/InP DHBT devices and circuits*. University of California, Santa Barbara, 2005.

[11] T. Ishibashi, "Influence of electron velocity overshoot on collector transit times of HBTs," *Electron Devices, IEEE Transactions on*, vol. 37, pp. 2103–2105, Sep 1990.

[12] H. Berger, "Models for contacts to planar devices," *Solid-State Electronics*, vol. 15, no. 2, pp. 145–158, 1972.

REFERENCES

[13] V. Jain, *InP DHBTs in a Refractory Emitter Process for THz Electronics*. University of California, Santa Barbara, 2011.

[14] W. Liu, *Handbook of III-V heterojunction bipolar transistors*. Wiley, 1998.

[15] H. Hasegawa, "Fermi level pinning and schottky barrier height control at metal-semiconductor interfaces of InP and related materials," *Japanese Journal of Applied Physics*, vol. 38, no. 2S, p. 1098, 1999.

[16] M. Vaidyanathan and D. Pulfrey, "Extrapolated fmax of heterojunction bipolar transistors," *Electron Devices, IEEE Transactions on*, vol. 46, pp. 301–309, Feb 1999.

[17] M. J. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S. C. Martin, R. Smith, *et al.*, "Submicron scaling of HBTs," *Electron Devices, IEEE Transactions on*, vol. 48, no. 11, pp. 2606–2624, 2001.

[18] P. Dodd and M. Lundstrom, "Minority electron transport in InP/InGaAs heterojunction bipolar transistors," *Applied Physics Letters*, vol. 61, pp. 465–467, Jul 1992.

[19] H. Ito, S. Yamahata, and K. Kurishima, "Evaluation of base transit time in ultra-thin carbon-doped base InP/InGaAs heterojunction bipolar transistors," *Electronics Letters*, vol. 32, pp. 1413–1415, Jul 1996.

[20] M. J. Rodwell, M. Le, and B. Brar, "InP bipolar ICs: scaling roadmaps, frequency limits, manufacturable technologies," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 271–286, 2008.

[21] H.-W. Chiang, J. C. Rode, P. Choudhary, and M. J. Rodwell, "Lateral carrier diffusion and current gain in terahertz InGaAs/InP double-heterojunction bipolar transistors," *Journal of Applied Physics*, vol. 115, no. 3, p. 034513, 2014.

Chapter 3

InP HBT Fabrication Technology

In the previous chapter, the design of heterobipolar transistors has been discussed. With epitaxial and lithographic scaling as means to attain higher RF bandwidths, main challenges lie in the fabrication technology for highly scaled devices. At the time this work began (March 2010), the dominant limitation of RF performance was the use of optical lithography for forming base electrodes and base/collector mesas: alignment tolerances ($\approx 100\,\mathrm{nm}$) and resolution ($> 250\,\mathrm{nm}$) of the i-line projection lithography stepper were inadequate for the sub-200 nm scaling generation. While superior resolution (sub-20 nm) of electron beam lithography (EBL) has already been exploited for forming emitters, attaining repeatable base-to-emitter alignment better than 30 nm has required optimization of writing strategies and design fracturing for EBL. Moreover, the 500 nm tall emitters necessitate thick EBL resist for the formation of base electrodes and mesas, adding further complexity to the process. The high

sensitivity of EBL resist to damage from stray electrons and radiation in electron beam evaporators has posed another challenge to overcome for successful integration of a complete EBL base formation process [1, 2].

It has been observed that prior processing and lithographic chemicals have introduced contaminants on the base semiconductor surface, limiting attainable contact resistivities to the base and subsequently f_{max} bandwidth. This has motivated the development of a process for manufacturing scaled TLMs that closely resemble HBTs in dimensions and fabrication. The fast turnaround of the scaled TLM process has enabled quick iterative debugging, leading to the development of a novel dual-deposition process that removes lithographic processing from the formation of metal-to-base semiconductor contacts. The process yields low resistance base contact that are thermally stable and impervious to electromigration at elevated operating temperatures.

Conventionally fabricated devices suffered from damage to exposed semiconductor surfaces between base electrodes and emitters, resulting in reduced current gain and reliability. A composite Al_2O_3/SiN_x passivation process has been developed that encapsulates these regions right immediately formation of base contacts.

Further limitations of the fabrication process that limit bandwidth of highly-scaled devices have been mitigated:

- Emitter End Undercut: Increased emitter thickness to emitter width ratio causes rapid undercut of the emitter stripe ends, deteriorating f_τ bandwidth

by reducing the active device area while maintaining capacitances.

- Accidental Deposition of Base Metal onto Emitter Sidewalls: Base metal accidentally deposited onto emitter sidewalls due to misalignment in the evaporator or suboptimal emitter metal shape result in increased C_{be}, limiting RF performance.

- Base Post Scaling: Non-scaled base posts constitute a significant fraction of C_{cb} particularly on small footprint devices, limiting RF bandwidth.

- Mechanical Stress in Thermal Processing: Different coefficients of thermal expansion between benzocyclobutene (BCB) and other materials on the sample surface cause mechanical stress particularly during quick thermal ramp-ups and -downs, resulting in poor yield and reliability.

- Low surface-assisted breakdown: Transistors passivated with BCB have exhibited low breakdown BV_{CEO}.

In this chapter, the UCSB fabrication process of HBTs is briefly outlined. Substantial key enhancements to the technology are presented that have enabled fabrication of THz bandwidth transistors.
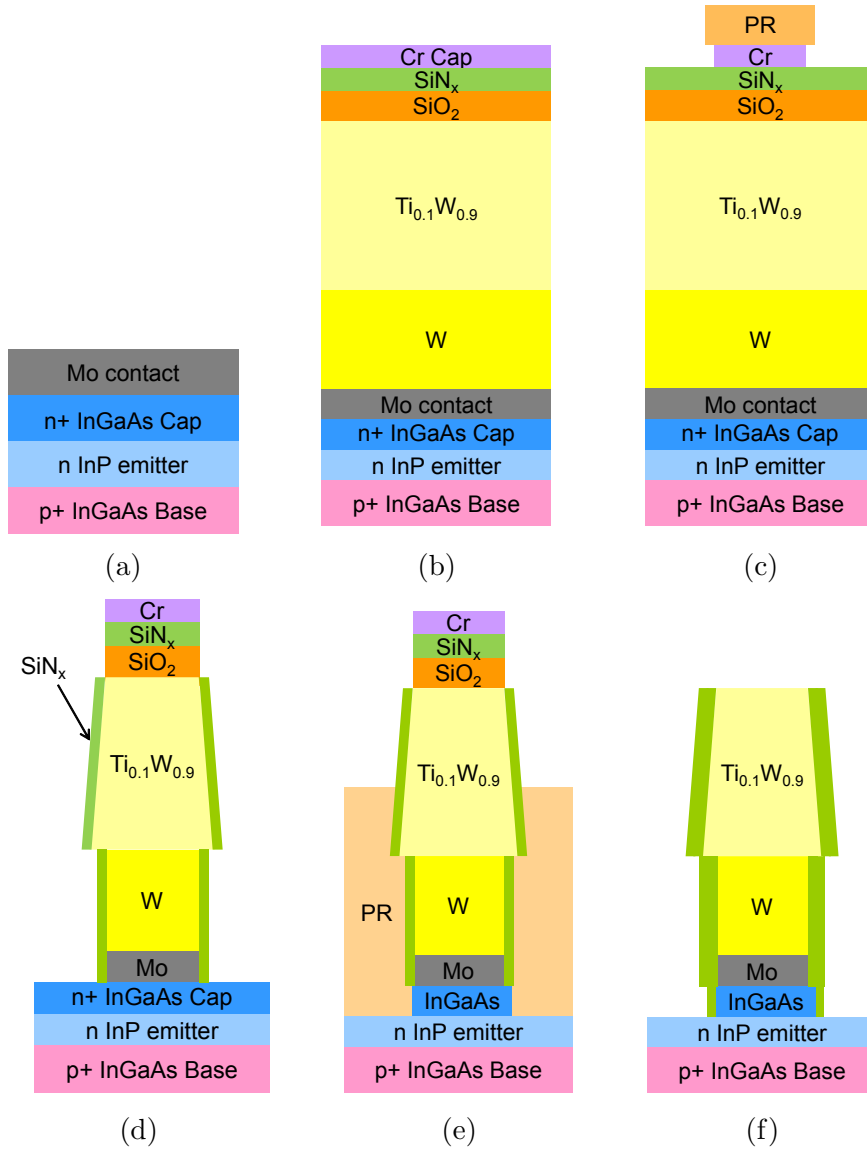
Figure 3.1: HBT Process: Emitter metal and sidewall formation.

3.1   Process Overview

Samples are cleaved from a 4" InP wafer epitaxially grown by a commercial vendor.
The top surface of the sample is oxidized in UV O_3, and the oxide is removed in diluted
hydrochloric acid (HCl). Immediately after the etch, the sample is transported to an
electron beam evaporator. Below a pressure of 5×10^{-7} Torr, a 20 nm thick layer of
the refractory metal molybdenum (Mo) is evaporated (Figure 3.1a).

A composite layer of tungsten (W) / titanium-tungsten (TiW, 10 % titanium by
weight) is sputtered onto the sample. The sputtering process has been calibrated to
deposit a stress-compensated film with an absolute stress modulus below 200 MPa.
80 nm thick SiO_2 and 40 nm thick SiN_x films are deposited in a PECVD process. A
40 nm thick chrome layer is evaporated (Figure 3.1b).

Photoresist is spun onto the chrome and it is patterned in an EBL process. The
pattern is transferred into the chrome in an anisotropic Cl_2/O_2 etch (Figure 3.1c).

The photoresist is stripped in N-Methyl-2-pyrrolidone (NMP). Residues are re-
moved in an oxygen plasma. The emitter metal is dry-etched in a vertical ICP
process with chrome as a hard mask [3]. A PECVD SiN_x sidewall is formed (≈ 30 nm
as deposited, Figure 3.1d).

The emitter InGaAs cap is removed in a short etch in a dilution of peroxide and
phosphoric acid H_2O_2:H_3PO_4:H_2O 1:1:25. Photoresist is spun onto the sample and
burned back in an oxygen plasma until the top part of the emitter stripes are exposed.
The sample is submerged in buffered hydrofluoric acid (BHF): the SiO_2 layer above
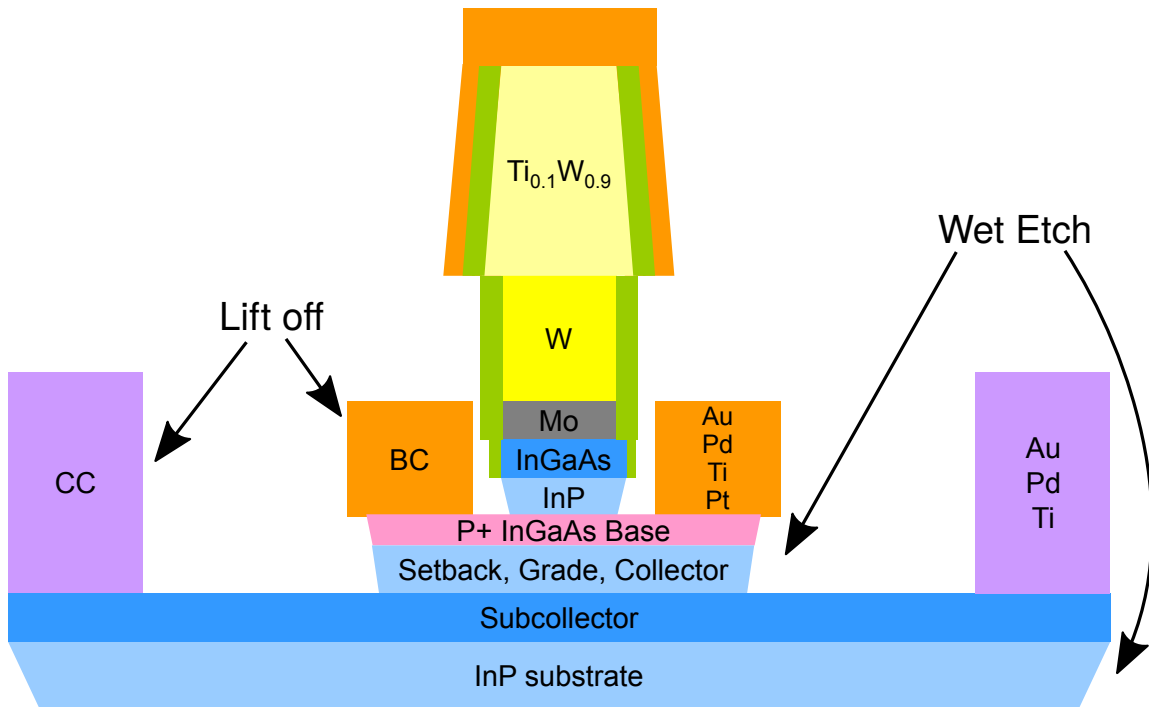
Figure 3.2: HBT Process: Base and collector electrode formation.

the TiW metal is removed so the chrome hard masks that have been damaged in the dry etch are flushed off the emitters (Figure 3.1e).

Because the BHF etch has partially removed the first sidewall, a second sidewall is formed (Figure 3.1f). The emitter semiconductor is removed in a wet etch H_3PO_4:HCl 4:1. Base electrodes and posts are formed in lift-off processes. The base/collector mesa is formed in selective wet etches. Another set of lift-offs forms collector contacts and posts. The devices are isolated in wet etches (3.2).

At this point, the front-end process is completed. Back-end fabrication starts with the application of benzocyclobutene (BCB), a low-ε dielectric. The BCB is cured by slowly heating the sample to $250\,°C$ and maintaining this temperature for $1\,h$. The

BCB is then ashed back in CF_4/O_2 to expose emitters and posts. A SiN_x layer is deposited to enhance metal 1 adhesion. Openings to emitters and posts are masked with photoresist and transferred into SiN_x with a CF_4/O_2 dry etch. Finally, 1 μm thick Ti/Au/Ti contact pads are formed in a lift-off process.

3.2   Emitter Process Improvements

3.2.1   Emitter Shape

In the presence of exposed refractory metals on the sample surface, processing chemicals can cause rapid corrosion of non-noble electrode metals. While the emitter sidewalls should fully encapsulate the emitter metal by design (Figure 3.1f), metal can be still exposed at weak points: When the emitter stripe is drawn as a rectangle on the lithographic mask, the high resolution of electron beam lithography process transfers this design pattern with almost perfectly orthogonal edges (Figure 3.3a). The ICP etch process for fabricating sidewalls removes the sidewall material preferentially at these edges, leaving the emitter metal exposed. In the mask layout, the edges of all emitters have been therefore rounded to establish complete sidewall coverage (Figure 3.3b).

(a)

(b)
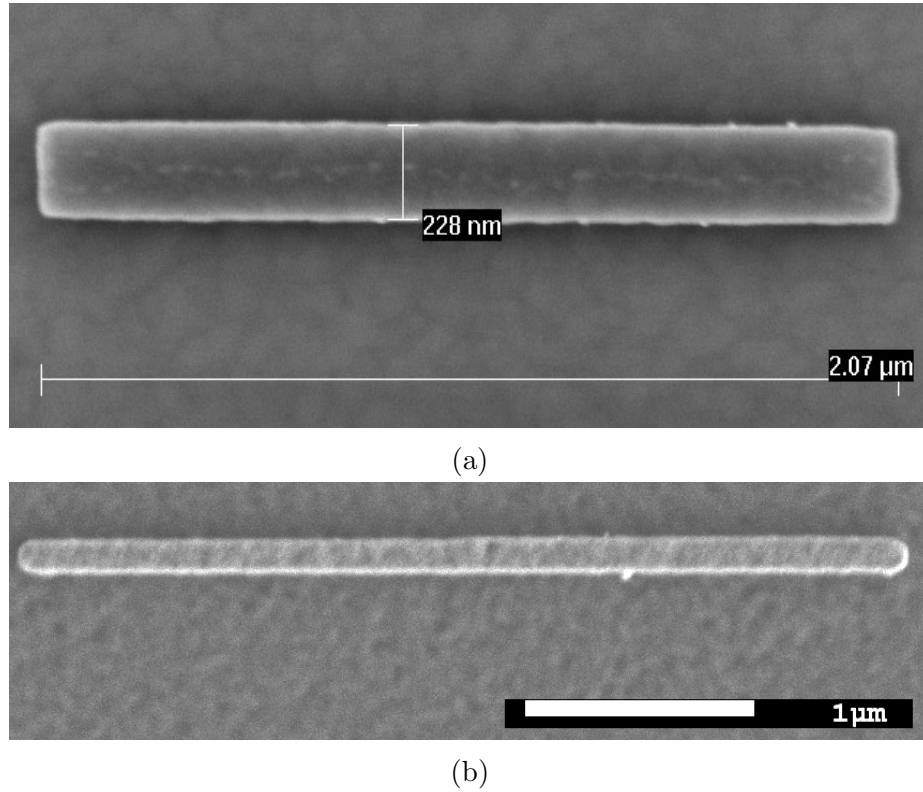
Figure 3.3: Transferred emitter pattern into chrome hard mask without (a) and with (b) rounding the edges of emitter stripes.

(a)   (b)
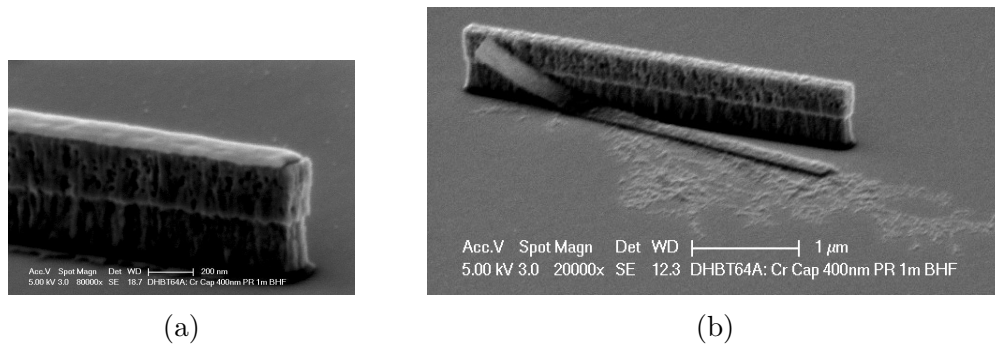
Figure 3.4: Failure modes of the chrome hard mask removal process: (a) chrome hard mask has collapsed onto emitter, (b) chrome hard mask has fallen onto the semiconductor, and damaged photoresist has contaminated the semiconductor surface.
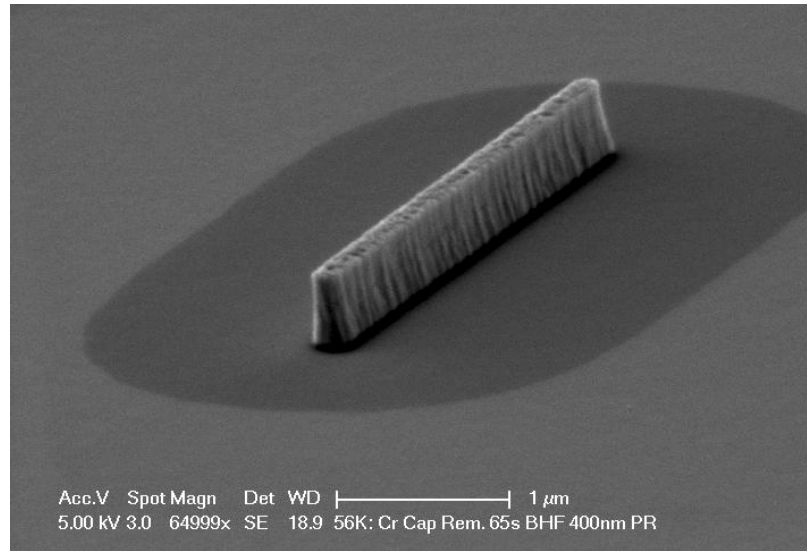
Figure 3.5: Discoloration of surface around emitters after stripping of the planarization resist in the chrome hard mask removal process.

3.2.2   Chrome Hard Mask Removal

A common failure mode that has diminished yield was insufficient removal of the chrome hard mask. On a large portion of devices, the chrome hard mask would not float off the emitters when the SiO_2 underlayer was being removed in BHF, but instead collapse back onto the emitter (Figure 3.4a) or onto close-by semiconductor (Figure 3.4b), resulting in either open emitter terminals or high base contact resistance. Furthermore, photoresist skin that has been been damaged in the ashing process would not be fully removed, contaminating critical semiconductor surfaces around the emitter and subsequently deteriorating the base contact when sinking down in the wet etch of the InP emitter.

In addition, surface discolorations around emitters have been observed after stripping the photoresist (Figure 3.5): this indicates that BHF acid has crept through

the photoresist and attacked the surface around emitters, invalidating the intended purpose of the planarization process.

The following process changes have been implemented to enhance yield:

- A layer of SiN_x has been inserted between SiO_2 and the chrome hard mask. The etch rate of SiN_x in BHF is ≈ 8 times slower than SiO_2 so a thin layer of SiN_x remains underneath the chrome hard mask when the SiO_2 film is being etched off. Electrostatic effects between the chrome hard mask and emitter metal are diminished.

- Small amounts of the nonionic surfactant Tergitol have been added to BHF to enhance the removal process.

- The photoresist planarization process has been abandoned: the chrome caps are removed in a BHF etch immediately after the emitter metal dry etch.

3.2.3   Surface Preparation for Emitter Wet Etch

Prior to wet etching the InP emitter, the semiconductor surface is prepared by removing oxides in a dilution solution of ammonium hydroxide (NH_4OH) : water 1:10. This step enhances repeatability of the InP wet etch. However, investigations into potential contamination introduced by emitter processing have revealed the presence of unknown compounds in the field and on the sensitive surfaces around the emitters after the InP emitter wet etch (Figure 3.6a). A consecutive solvent clean of the sample
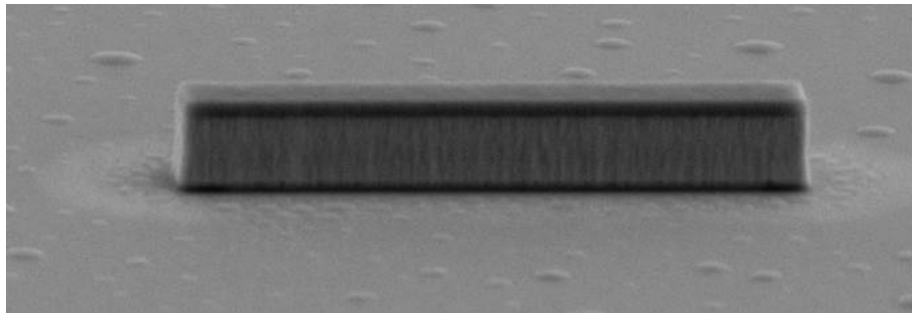
(a)
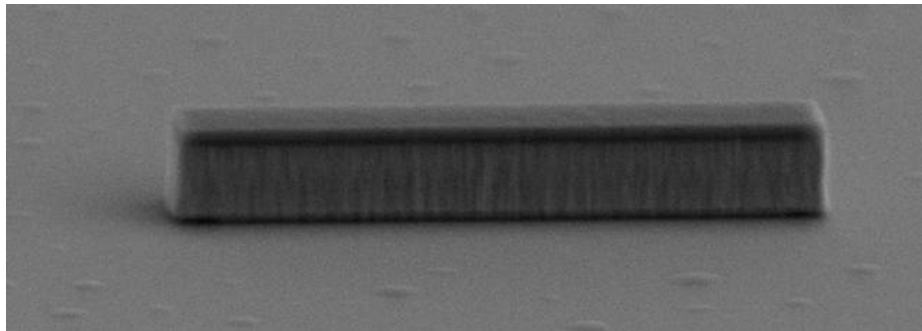
(b)

Figure 3.6: Contamination of the semiconductor surface around emitter features as a result of diluted NH$_4$OH surface treatment (a) immediately after the wet etch and (b) after a solvent clean (rinse in acetone, isopropyl alcohol, deionized water).
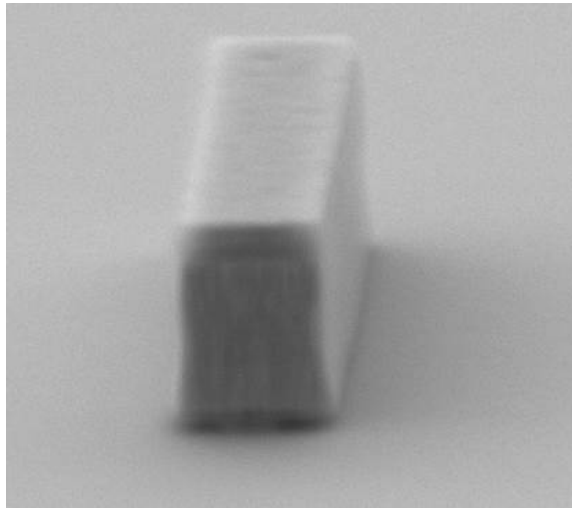
Figure 3.7: Emitter after oxidation/oxide removal cycles and InP wet etch showing uniform base surface without any signs of contamination.

has only partially removed contaminants (Figure 3.6b), severely limiting minimum attainable base contact resistivity.

The following changes have been made to the process to mitigate these effects:

- Surface preparation in ammonia has been substituted with multiple oxidation and oxide removal cycles: the sample surface is oxidized in a UV O_3 reactor, and the oxide is removed in diluted hydrochloric acid (1:10 $HCl:H_2O$).

- Bench lights are shut off during any wet etches and rinses to reduce photoelectric effects [4].

Figure 3.7 shows an emitter on a fabrication campaign where fixes described above have been implemented: the uniform base surface shows no signs of contamination.

slow etch facet

fast
etch
facets

n InP emitter

fast
etch
facets
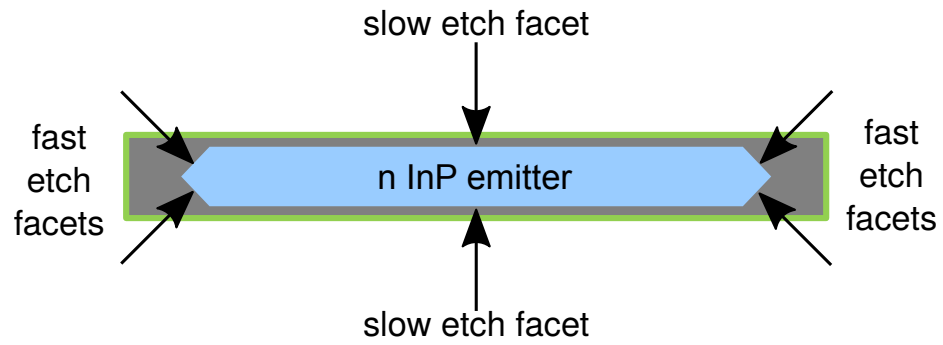
slow etch facet

Figure 3.8: Bottom-up view of the emitter with slow and fast etch facets.

1100 nm   220 nm

(a)

830 nm   50 nm
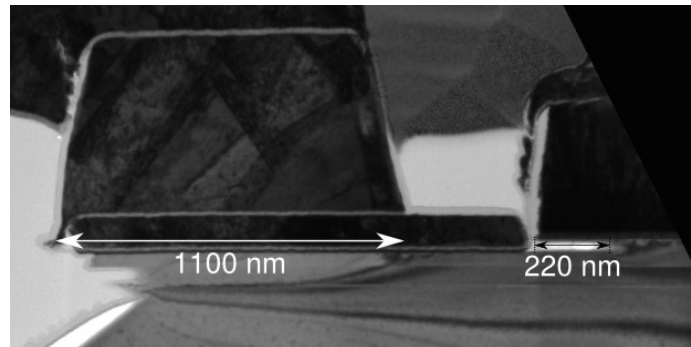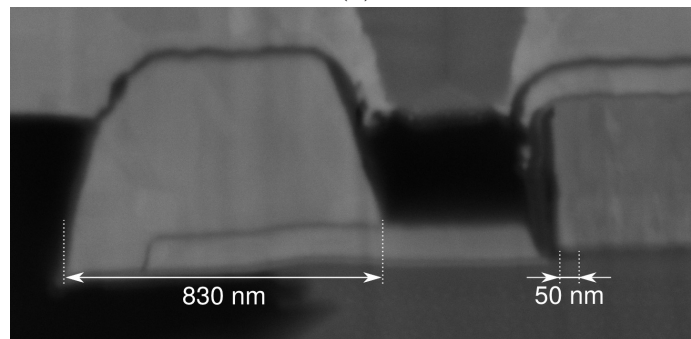
(b)

Figure 3.9: Emitter end and base post undercut/scaling between (a) regular (a) and (b) improved process.

3.2.4   Emitter End Undercut

With shrinking emitter width w_e, the difference in etch rates between crystal facets exacerbates undercut of emitter stripes from their ends (Figure 3.8). The undercut reduces the active device area while keeping capacitances C_{be} and C_{cb} approximately constant, degrading f_τ and $f_{\max}$ bandwidth. In conjunction with the surface preparation described above that removes a fraction of the InP emitter, the wet etch time has been reduced by 40 % to 5 s to control the undercut.

Figure 3.9 indicate significant reduction of the single-sided emitter end undercut from 220 nm to 50 nm as a result of above changes.

3.3   Base Process Improvements

With aggressive scaling, high resolution lithographic processes are required for fabricating transistors. While emitters have been formed exploiting the advanced resolution of electron beam lithography at the time this work started (March 2010), base fabrication still relied on i-line projection lithography with limited resolution (300 nm, wavelength λ=365.4 nm) and emitter-to-base alignment (150 nm). Base access resistance R_{bb} depends very closely on the base electrodes formed symmetrically around the emitter: Any misalignment will add to C_{cb} without significantly reducing R_{bb}, deteriorating $f_{\max}$ bandwidth. At sub-200 nm emitter widths, both resolution and emitter-to-base alignment tolerance of i-line lithography are inadequate. This

has motivated the transition of base formation processes to electron beam lithography with sub-20 nm resolution and alignment better than 30 nm.

3.3.1   Electron Beam Lithography Strategies

With resolution requirements of the HBT process far below the capabilities of the EBL tool (JEOL 6300), the writing strategy has been optimized for minimizing the exposure time: the tool has been operated in 4th lens mode with the largest deflector field size (500 µm) to minimize the number of stage movements during an exposure. Also, the beam current has been set to 2 nA, being the optimum between short pixel dwelling time and small step size (resolution).

Marks for global (200 µm·4 µm, 1 per die) and local alignment (10 µm·2 µm, 64 per die) have been added to the mask for emitter lithography. The writing time between alignment mark detection and subsequently the misalignment caused by system drift is reduced by splitting the e-beam reticle into quarters. Moreover, the lithography pattern is partitioned into a primary and secondary set: the primary set containing only alignment-critical, small area transistor features is written first, allowing for frequent re-alignment by mark detection. The secondary set is comprised of large control structures such as Vernier marks and profilometer pads insensitive to misalignment and takes the bulk of the pattern exposure time.

Scattering processes of beam electrons with resist and substrate cause double exposure of nearby resist, making it difficult to manually assign correct exposure

doses to areas with densely packed features on the reticle. This so-called proximity effect can be corrected numerically, greatly reducing the complexity of transferring the design pattern into resist. This numerical correction has been therefore used for base electrode lithography processes.

3.3.2   Base Metal Formation by Electron Beam Lithography

While regular EBL processes use resist that is very thin ($< 100\,\mathrm{nm}$) to obtain highest resolution ($< 10\,\mathrm{nm}$), the height of the emitter metal and semiconductor ($\approx 500\,\mathrm{nm}$) requires resist of similar thickness to reduce local thickness variation effects while constraints on the resolution are more relaxed ($< 50\,\mathrm{nm}$). Also, the resist sidewall should have a negative slope to enable lift-off processes. Due to the sensitivity of f_max bandwidth on the base contact, it is important that the resist develops fully out without leaving residues on the base semiconductor surface that will degrade base contact resistivity.

UV6 Process

An EBL lift-off process using thick resist has been developed by Felix Recht and Dan Denninghoff for use in high aspect ratio T-gates [5]: the process uses MicroChem UV6, a positive tone resist that has been initially designed for UV exposure, but can also be exposed by electron beam lithography. The process has been adjusted to meet the requirements of the HBT process: the resist is spun onto samples at 3000 rpm to a
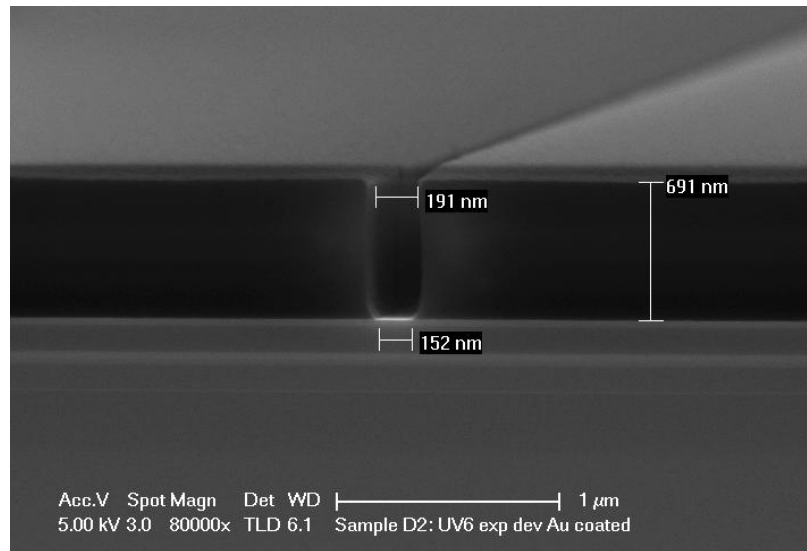
Figure 3.10: Cross-sectional SEM of exposed and developed UV6 resist coated with a thin layer of Au for enhanced contrast.

thickness of 700 nm (Figure 3.10). The resist has a chemical amplification mechanism requiring a post bake immediately after exposure. The development chemistry (2.38 % TMAH) is identical to optical lithography processes.

While a lift-off process with this resist can produce base electrodes with desired features, it is prone to fail when the sample is irradiated in x-rays and electrons during e-beam evaporation of the base electrodes: the resist sidewalls will either only partially dissolve in photoresist stripper, leaving residues behind, or completely fail to dissolve, destroying the sample (Figure 3.11). An alternative EBL process has been therefore developed.

(a)	(b)
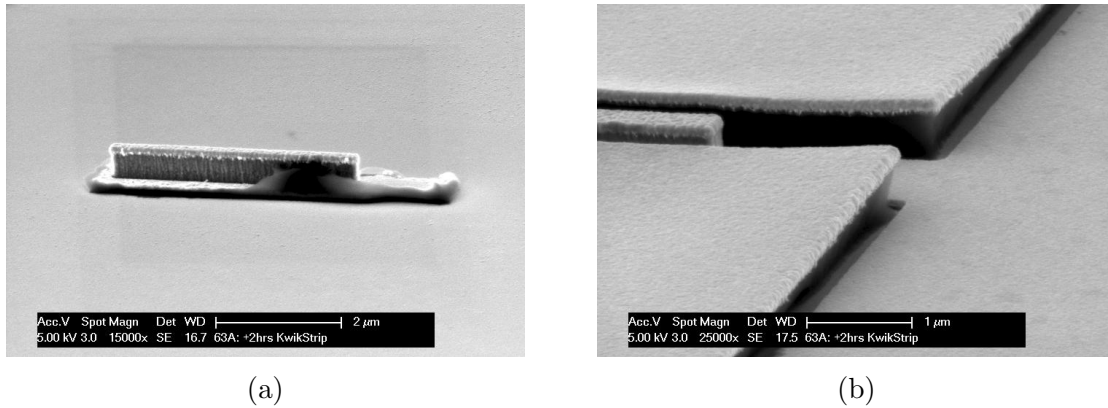
Figure 3.11: UV6 resist damaged by electron and x-ray radiation during electron beam evaporation. (a) Residues around lifted-off base electrode, (b) lift-off failure due to damaged resist sidewalls that failed to be removed in photoresist stripper.
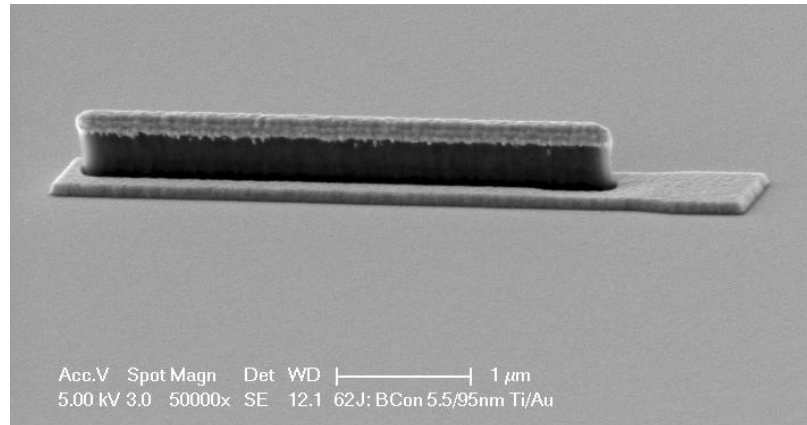
Figure 3.12: Emitter with base electrode lifted off using PMGI/ZEP.
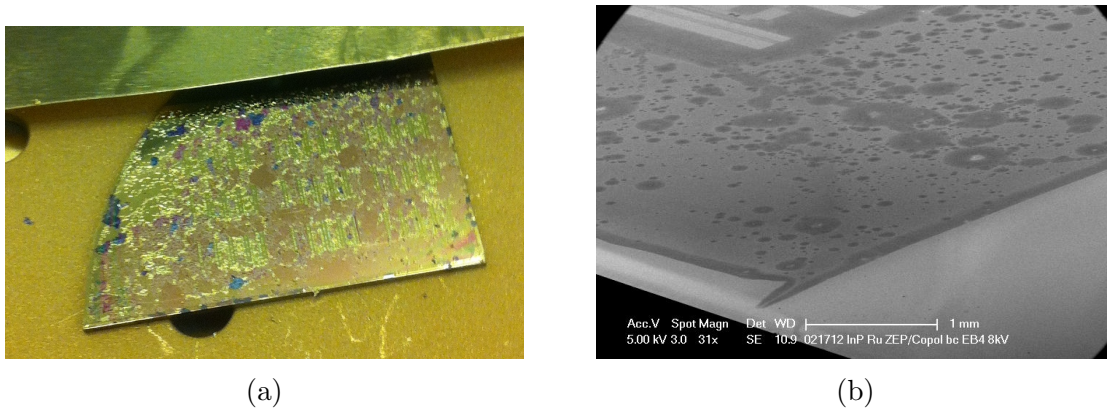
Figure 3.13: ZEP resist damaged by electron and x-ray radiation during electron beam evaporation. (a) photograph of a sample immediately after evaporation showing resist blistering, (b) low magnification SEM of a sample after stripping blistered resist.

PMGI/ZEP520 Process

A standard positive tone dual-layer resist process has been adjusted: PMGI is spun onto the sample at 3000 rpm to a thickness of 400 nm. The PMGI is baked at 180 °C for 3 min. ZEP520:Anisole 1:1 or, alternatively, CSAR:Anisole is spun onto the sample to a thickness of 200 nm. After another prebake at 180 °C for 3 min, the sample is exposed. ZEP/CSAR is developed using amyl acetate: it has been found that development with methyl isobutyl ketone (MIBK) : isopropyl alcohol (IPA) leaves residues on the underlayer PMGI that will collapse onto the sensitive base surface. The underlayer is developed out in 2.38 % TMAH. The underlayer has low sensitivity to electron beam exposure so it is isotropically washed out during development, creating a desired overhang of ZEP/CSAR. This greatly improves lift-off (Figure 3.12). However, the resist is also very sensitive to damage during e-beam evaporation: the resist blisters when irradiated with x-rays and electrons, ruining the lift-off. It can

also form compounds on the sample surface that cannot be removed anymore with either strippers or oxygen plasma treatment.

Enhancing Electron Beam Deposition with E-Beam Resist

Both ZEP/CSAR and UV6 are very sensitive to damage from x-rays and electrons during metal deposition in electron beam evaporators. Cheng has found that carbon contamination of gold sources can greatly increase emission of electrons onto the sample [1, 2]. The presence of a nickel source in the hearth adds further to electron irradiation: Nickel is magnetic at room temperatures and can deflect the electrons from beam onto the sample.

As a remedy, the gold source been placed into a crucible made out of tungsten. Small amounts of tantalum have been added to the gold source that acts as getter for carbon contaminants. A graphite spacer was placed between the gold crucible and the hearth to reduce the heat conductance between source and hearth, decreasing the power required for evaporation. Furthermore, nickel was removed from the system for every deposition. This has reduced radiation to an acceptable level so this catastrophic lift-off failure mechanism has been eliminated.

3.3.3   Base Mesa Formation by Electron Beam Lithography

A thicker dilution of the negative mAN-2400 resist that has already been in use for the emitter lithography has been adopted for protecting the base/emitter regions
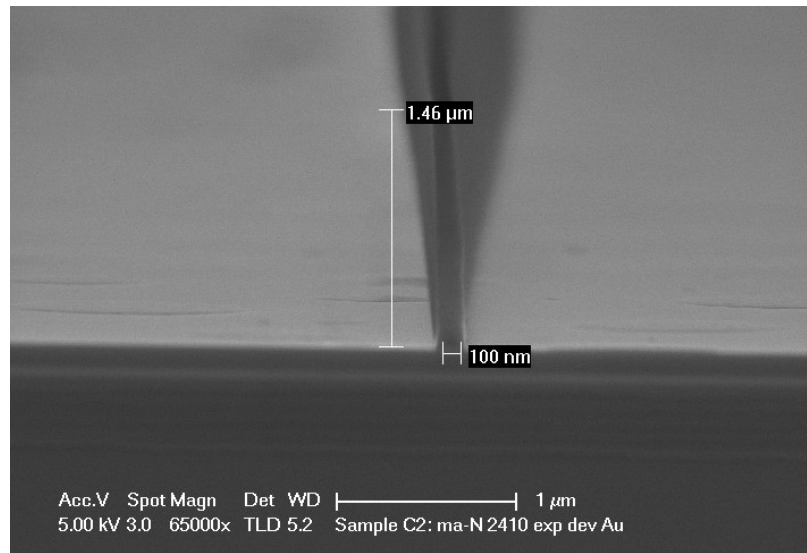
Figure 3.14: Cross-sectional SEM of a 100 nm wide line in maN2410.

during the wet etches that form the base/collector mesa. The resist spins on at 1.4 μm thickness and yields aspect ratio in excess of 10:1 (see Figure 3.14).

3.3.4   Dual-Deposition Base Metal Process

Base contact resistance is very critical for $f_{\max}$ bandwidth. Lithographic processes introduce contamination to the base semiconductor, limiting the attainable contact resistance. This has motivated the development of a two-step base metalization process that decouples deposition of base contact metals from the formation of base electrodes. In the first step, a blanket base contact layer is deposited, removing any lithographic processing. In the second step, base pads are lifted-off. The blanket metal deposited in the field is removed in a dry etch.

Prior to wet etching the InP emitter, the surface is cleaned by removing the
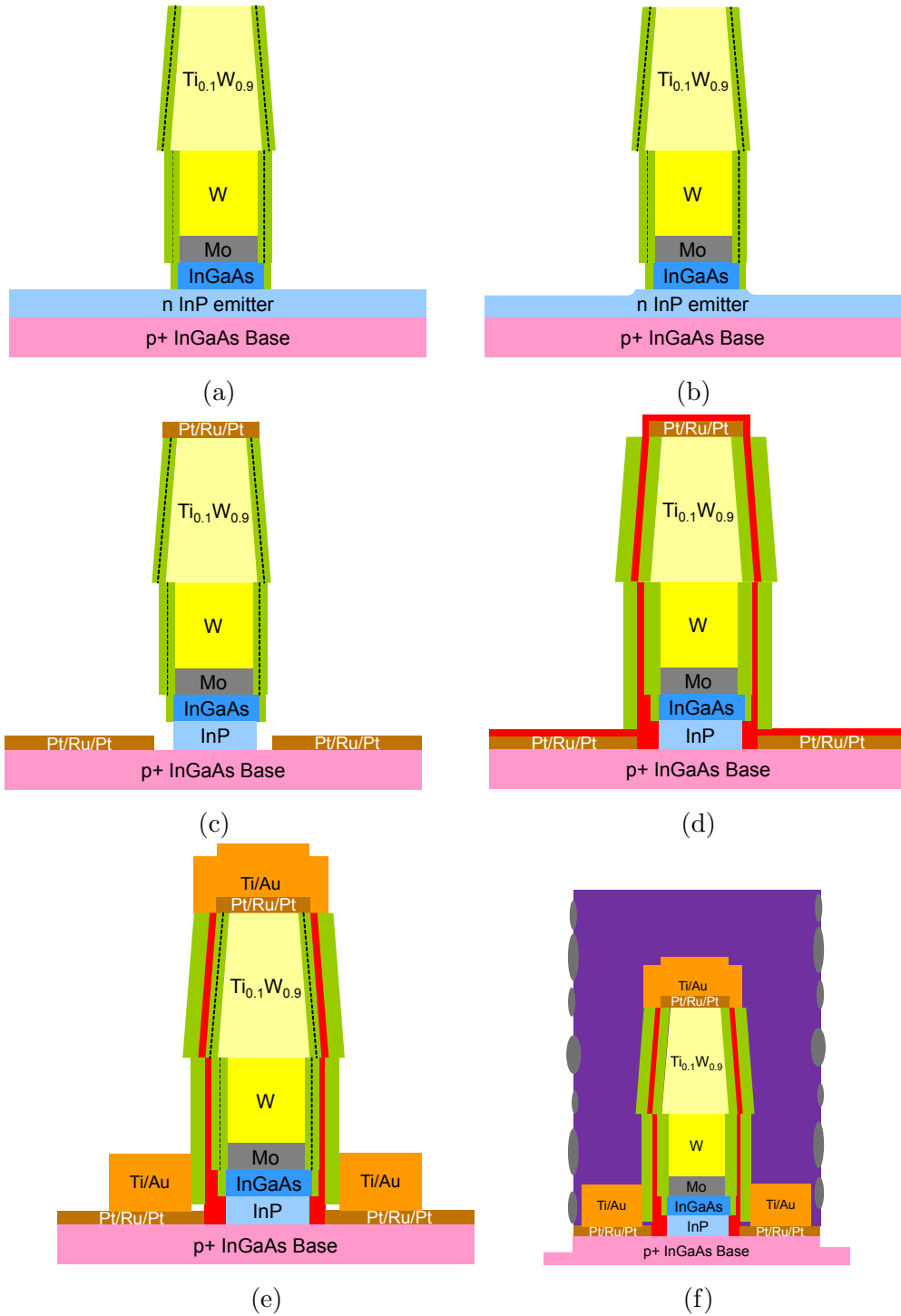
Figure 3.15: Dual deposited base metalization process flow.

topmost layers with multiple cycles of oxidation in UV ozone and subsequent oxide removal in diluted HCl (Figure 3.15b, also compare to section 3.2.3). The 30 nm thick InP emitter is removed in a 5 s wet etch in 4:1 H_3PO_4:HCl. The etch time has been shortened by 40 % to reduce lateral emitter undercut along fast etch planes from 220 nm per side to $\approx$ 50 nm (Figure 3.9a).

After a solvent clean for removing etch residues and short deoxidizing dip in 1:10 HCl:DI, the sample is immediately loaded to an electron beam evaporator. When a pressure below 6×10^{-7} Torr has been reached, a blanket metal stack of Pt/Ru/Pt is evaporated onto the sample (Figure 3.15c). The initial thin layer of platinum is deposited at very low rates (0.1 Å/s) to improve surface coverage: this layer will controllably sink into the base semiconductor, moving the ohmic metal-semiconductor contact away from the surface. The 15 nm thick refractory ruthenium acts as a thermally stable diffusion barrier for upper metal layers. The 2 nm thick topmost noble metal platinum layer encapsulates Ru, protecting metals exposed on the sample surface from galvanic corrosion in processing chemicals. This metal stack has a sheet resistance of 25 $\Omega/\square$, 20:1 less than the sheet resistance of the base semiconductor.

After initial metal deposition, a 10 nm thin layer of Al_2O_3 is conformally deposited in a thermal ALD process. The alumina protects the exposed base regions between base metalization and emitter semiconductor from damage in subsequent processing and passivates the emitter-base surface. A 20 nm thick PECVD SiN_x sidewall is formed (Figure 3.15d), enabling the removal of Al_2O_3 in the field in a wet etch either

in 2.38% TMAH (etch rate 3 nm/min) or 1:50 BHF:DI (etch rate $\approx$30 nm/min). The

sidewall also increases the spacing between accidental base metal deposits on the

emitter sidewalls and the emitter metal, reducing the base-emitter capacitance C_{be}

and preventing short-outs.

Base metal pads of Ti/Au 5/95 nm are then lifted-off in a standard, bi-layer

electron beam lithography process (Figure 3.15e). Prior to metal deposition, resist

residues are removed in an oxygen plasma (20 s, 100 W, 300 mTorr): the presence of

the base contact metal diminishes the risk of damaging the base semiconductor. The

sheet resistance of the full composite base metal electrode is $\approx 0.4\,\Omega/\square$, approximately

half the sheet resistance of conventional lifted-off base metal stacks. After deposit-

ing base posts, emitter/base regions are protected utilizing electron beam lithogra-

phy with 1 μm thick resist (microposit maN-2410). A Cl_2/O_2 dry etch (20/5 sccm,

0.67 Pa, 400 W RF, 100 W ICP, 40 s) is used to remove the topmost blanket base layers

Pt/Ru [6]. Without breaking vacuum, a short sputtering etch in Ar/Cl_2 (45/5 sccm,

1 Pa, 600 W RF, 150 W ICP, 20 s) is performed to remove non-volatile etching prod-

ucts from the field. The addition of Cl_2 enables removal of etch redeposits on the

resist sidewalls in subsequent wet etches [7]. However, the sputtering etch is non-

selective to InGaAs, removing $\approx$ 25 nm of the base/collector. The wet etching times

of the base/collector mesa are adjusted accordingly (Figure 3.15f). Plasma-damaged

resist residues collapse onto emitter and base post during stripping and are removed

prior to deposition of interconnect metals in a short Ar sputter (20 s, 300 W ICP,
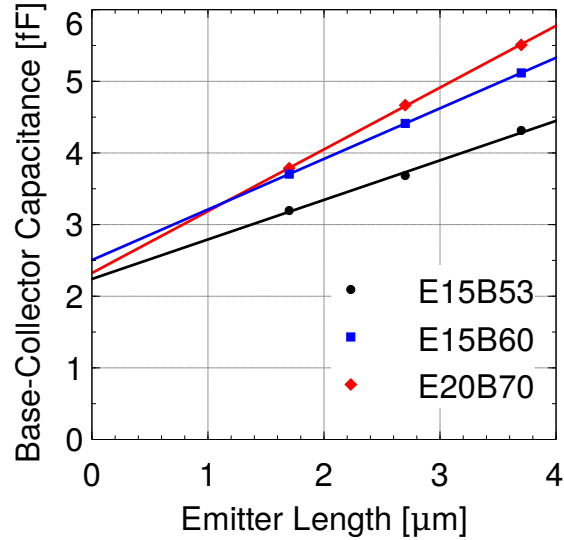
Figure 3.16: Base-collector capacitance at Kirk current density of a set of devices with identical widths as a function of emitter length. The intercept gives an estimate for the capacitance of the base post $C_{cb,\text{post}} \approx 2.2\,\text{fF}$.

50 W RF, 20 sccm, 1 Pa).

3.3.5   Base Post Scaling

Prior base post fabrication processes have established the use of a lift-off underlayer LOL1000 and 1.2 µm thick lift-off resist NLOF5510 for the fabrication of base posts with 15 /15 /550 nm Ti/Pd/Au metal stack and a diameter of 1.1 µm [8]. While the addition of a lift-off layer to the process had alleviated some yield issues, it remained unreliable: a common failure mode was ripping out almost all base posts in certain areas of the sample during the lift-off process. Also, the base post had not been scaled properly for fears of decreasing process reliability even further.

At 1.1 µm diameter, controllably undercutting the base post without damaging

the active part of the device remained challenging (Figure3.9a). A large fraction of the base-collector capacitance was therefore due to the base post (Figure 3.16), thereby limiting RF bandwidth.

To enhance base post adhesion, the metal stack has been changed to Ti/Au: Pd has been omitted. The diameter of the base post has also been decreased to $\approx$ 800 nm. The 40 % area reduction has enabled almost complete undercut of the base post (Figure 3.9b). With these changes, failure at base post lift-off has not occurred.

3.4   Collector Process Improvements

Among emitter, base and subcollector contacts, the subcollector InGaAs layer accumulates most contaminants from prior processing. However, the horseshoe-shaped collector metal has also the largest contact area, making the collector less sensitive to higher contact resistivities (see Figure 2.13). In order to reliably yield contact resistivies below $20 \, \Omega \, \mu m^2$, the process has been modified to include a surface clean prior to collector contact deposition: the sample is oxidized in UV O_3, and the oxide is removed in diluted hydrochloric acid. The sample is rinsed in DI water and $20 / 20 / 250 \, nm$ Ti/Pd/Au contacts are lifted-off.

3.5   Backend Improvements

After finishing front-end fabrication, devices are planarized in BCB and contacted with interconnect structures on metal 1. Very low yield due to open emitter-base junctions has been observed which is attributed to thermo-mechanical fatigue. Several process changes have been made to restore yield.

3.5.1   Device Passivation

In previous process campaigns, the sample has been deoxidized in diluted NH_4OH prior to the application of BCB. The surface preparation was changed to hydrochloric acid (compare to section 3.2.3), and $30\,nm$ thick PECVD SiN_x has been grown prior to application of BCB in order to encapsulate all structures, providing additional mechanical support during thermal cycling. The SiN_x layer has also enhanced breakdown from $3.7\,V$ [9] to $4.3\,V$ [10].

3.5.2   Low Temperature Nitride

A layer of SiN_x is used between BCB and metal 1 to enhance metal adhesion. In previous process campaigns, this layer was grown by PECVD at $250\,°C$. The PECVD process required fast thermal ramps, inducing high mechanical stress on the sample. The PECVD process has been substituted with a sputtering process at either $100\,°C$ or room temperature.

3.5.3   Tapered Base Feed Line

In old designs of the interconnect metalization, the feed for the base post had

significant overlap capacitance with the collector semiconductor. The feed has been

tapered to reduce this capacitance.

$C_{cb,extra}$

Device Isolation Mask

Base Feed

Collector Feed

$C_{cb,extra}$

Emitter
Feed

(a)

Base Feed

Collector Feed

Emitter
Feed

(b)
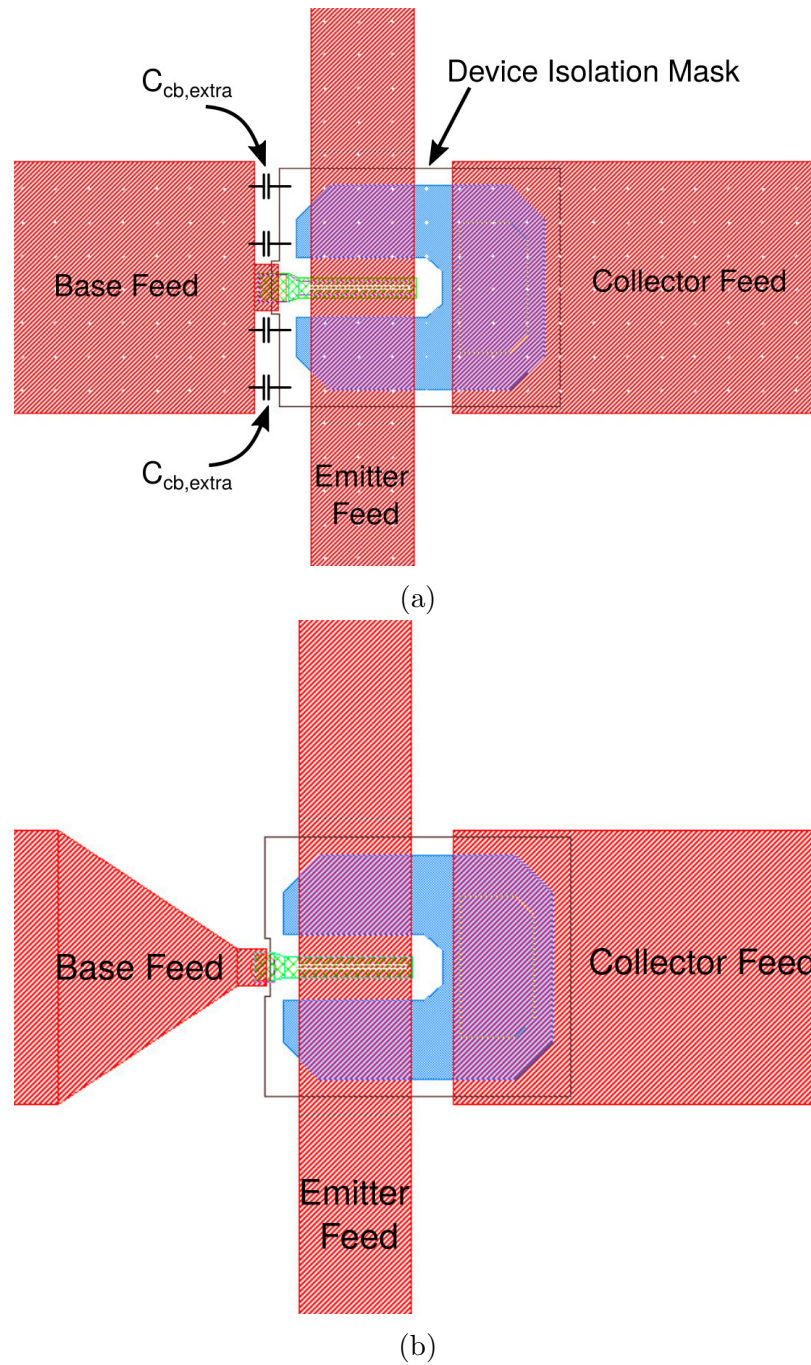
Figure 3.17: Metal 1 mask layout adjustments to reduce overlap capacitance between base feed line and subcollector: (a) before adjustment, (b) after.

3.6   Scaled TLM Process

A fast-turnaround process for fabricating scaled TLMs (section 2.3) has been developed to quickly iterate on process changes for improving base contact resistivity. In contrast to conventional TLM fabrication [11], lithographic processes and key dimensions of scaled TLMs have been designed to have high similarity to the HBTs. The epitaxial design of wafers onto which scaled TLMs have been fabricated is also either very similar or identical to epitaxial HBT wafers.

Fabrication for unpinched TLMs that have resist-defined gaps starts with lift-off of contact pads in an electron beam lithography process (Figure 3.18a, also see section 3.3.2). Large area pads of 20/500 nm Ti/Au that overlap the contact metal are then lifted off with i-line lithography to enable needle probing (Figure 3.18b) . The TLMs are isolated using the established EBL process for masking the TLM gap regions (3.18c, section 3.3.3).

The current force pads have been designed to provide symmetrical current feed to the TLM gap, while voltage sense pads are connected close to the gap to reduce the influence of parasitic voltage drops in the metal. Gap and pad width dimensions are obtained from SEMs, and the resistance between pads is measured with four terminal sensing.

The process for pinched TLMs is very similar to HBT fabrication: emitter metal is deposited, emitter stripes are formed and etched and a sidewall is deposited (see section 3.1). The contact metal is deposited around the emitter in a self-aligned
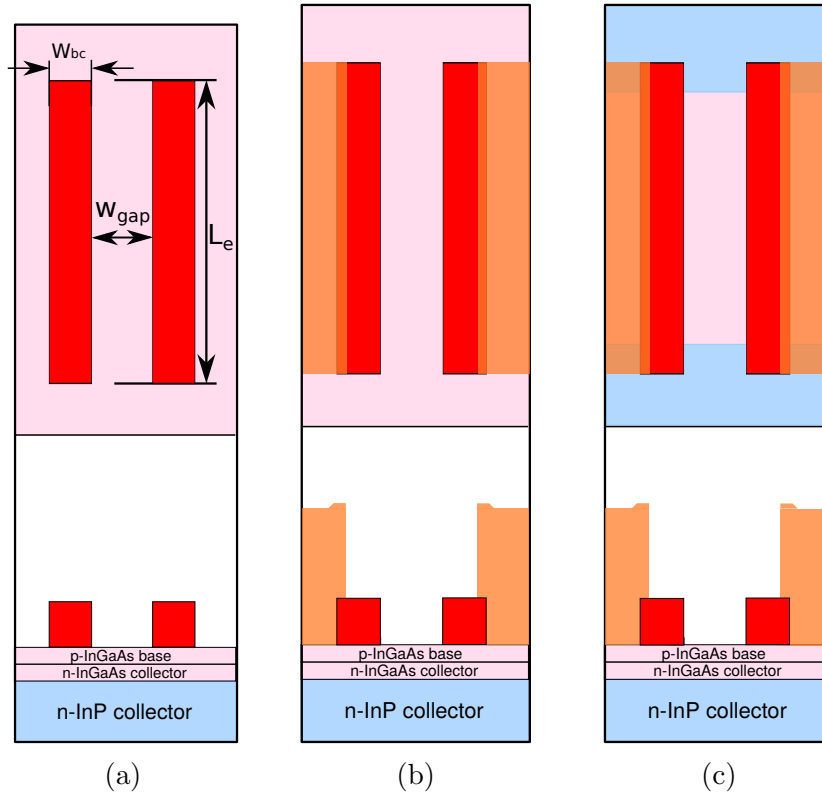
Figure 3.18: Top-down and cross-sectional view of scaled TLM process for unpinched TLMs. Gap spacing $w_{gap} \approx w_e$, contact pad width $\approx w_{bc}$, pad dimensions perpendicular to gap L_e.
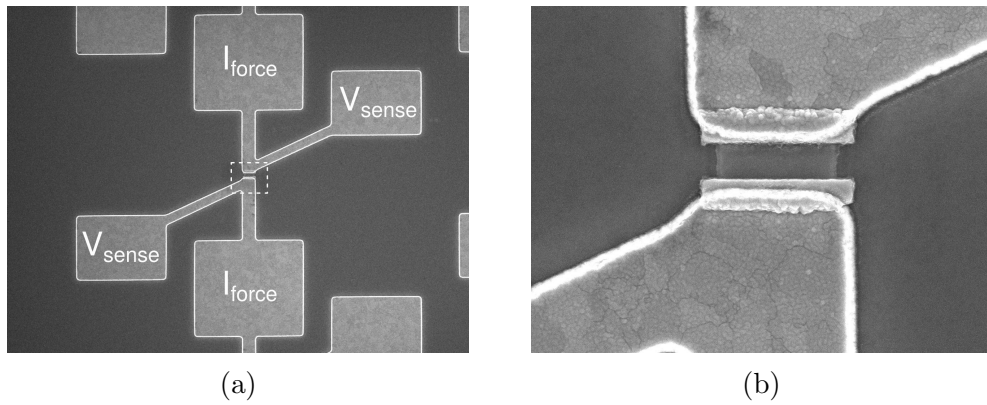
Figure 3.19: Top-down SEM of fabricated unpinched TLMs: (a) low magnification image showing probe pads and dotted outline of (b) higher magnification image showing TLM gap region.
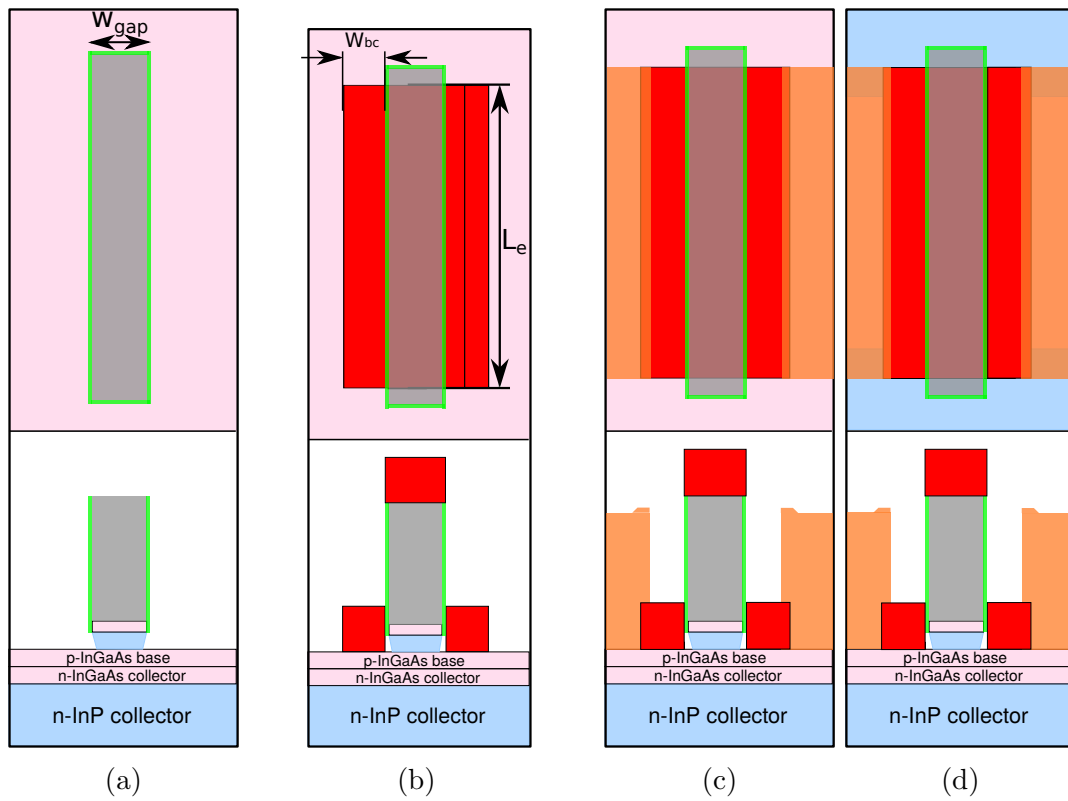
Figure 3.20: Scaled TLM process with TLM gaps defined by emitters.

process (Figure 3.20a). The remainder of the process is identical to the non-scaled

TLM process.

References

[1] K. Cheng, M. Le, D. Mitchell, and L. Hanes, "Effects of electron radiation generated during e-beam evaporation on photoresist liftoff process," in *CS MANTECH Conference*, 2010.

[2] Y. Pao, K. Tran, C. Shih, and N. Hardy, "Solution to the e-beam gate resist blistering problem of 0.15 micron PHEMTs," in *CS MANTECH Conference*, 1999.

[3] V. Jain, E. Lobisser, A. Baraskar, B. J. Thibeault, M. J. W. Rodwell, Z. Griffith, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. M. Fastenau, and W. K. Liu, "InGaAs/InP DHBTs in a dry-etched refractory metal emitter process demonstrating simultaneous ft/fmax 430/800 GHz," *Electron Device Letters, IEEE*, vol. 32, pp. 24–26, Jan 2011.

[4] T. Krämer, *High-Speed InP Heterojunction Bipolar Transistors and Integrated Circuits in Transferred Substrate Technology*. Cuvillier, 2010.

[5] D. J. Denninghoff, S. Dasgupta, J. Lu, S. Keller, and U. K. Mishra, "Design of high-aspect-ratio T-gates on n-polar GaN/AlGaN MIS-HEMTs for high fmax," *Electron Device Letters, IEEE*, vol. 33, no. 6, pp. 785–787, 2012.

[6] C. Hsu, J. Coburn, and D. Graves, "Etching of ruthenium coatings in O2-and Cl2-containing plasmas," *Journal of Vacuum Science & Technology A*, vol. 24, no. 1, pp. 1–8, 2006.

[7] T. Shibano, K. Nakamura, T. Takenaga, and K. Ono, "Platinum etching in Ar/Cl2 plasmas with a photoresist mask," *Journal of Vacuum Science Technology A: Vacuum, Surfaces, and Films*, vol. 17, pp. 799–804, May 1999.

[8] V. Jain, *InP DHBTs in a Refractory Emitter Process for THz Electronics*. University of California, Santa Barbara, 2011.

[9] V. Jain, J. Rode, H.-W. Chiang, A. Baraskar, E. Lobisser, B. J. Thibeault, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and W. Liu, "1.0 THz fmax InP DHBTs in a refractory emitter and self-aligned base process for reduced base access resistance," in *Device Research Conference (DRC), 2011 69th Annual*, pp. 271–272, June 2011.

[10] J. Rode, H.-W. Chiang, P. Choudhary, V. Jain, B. Thibeault, W. Mitchell, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and A. Liu, "An InGaAs/InP DHBT with simultaneous ft/fmax 404/901GHz and 4.3V breakdown voltage," *Electron Devices Society, IEEE Journal of the*, vol. 3, pp. 54–57, Jan 2015.

[11] A. Baraskar, V. Jain, M. A. Wistey, E. Lobisser, B. J. Thibeault, A. C. Gossard, and M. J. Rodwell, "In-situ and ex-situ ohmic contacts to heavily doped p-InGaAs," in *16th Int. Conf. Molecular Beam Epitaxy*, 2010.

Chapter 4

Experimental Results

In this chapter, measurement methodology is briefly outline. Device results of samples that have been fabricated with the improved fabrication processes described previously are presented. Sample HBT56J features a 30 nm base and a 100 nm collector. Base electrodes and base/collector mesas have been formed using using electron beam lithography and base/collector passivation with SiN_x [1]. Several key issues that have limited RF bandwidth to f_τ / f_{max} of 400/900 GHz have been identified, among which is high base contact resistance. The quick-turnaround scaled TLM process has been exploited to investigate factors limiting base resistivity. Important results of the scaled TLM process are presented, culminating in the development of an advanced dual-deposition base metalization process (see section 3.3.4). Two samples with identical epitaxial design are shown that have been fabricated using this process and other process enhancements: HBT64C and HBT64J feature a 20 nm base with increased

doping and a 100 nm collector. The highest extractable f_{max} bandwidth 1070 GHz

at f_τ of 480 GHz has been achieved on HBT64J. We suspect higher bandwidth of

transistors with smaller junctions, but extraction of f_{max} bandwidth is not possible

on these devices due to inadequacies of the RF calibration structures.

4.1   Device Measurement

4.1.1   DC Characteristics

The resistance of pinched and unpinched base TLMs and collector TLMs is mea-

sured with a DC semiconductor parameter analyzer. TLM pads are contacted with

needle probes. Current is forced through the TLMs while the voltage drop is simulta-

neously measured with high-impedance differential voltage measurement units (four

terminal sensing). The current is swept between negative and positive maximum

operating current density. The resistance is extracted from a least-squares fit to a

linear function. The measurement error estimated from the deviation of the fit to

the measured data is less than 1 % of the total resistance. Gap spacings have been

obtained from SEMs that have been taken prior to BCB planarization.

DC measurements of a transistor (Gummel, common-emitter I/V, breakdown,

etc.) are taken with microwave probes that make a mostly reflection-free contact

to the co-planar waveguide structures into which the transistors are embedded. The

measurement instrument is connected through DC ports of bias tees that are mounted

on the microwave probes. The RF ports of the bias tees are terminated.

4.1.2   Small-signal Microwave Characteristics

Measurements of device microwave characteristics are crucial for determining RF bandwidths and intrinsic device parameters. Highly-scaled HBTs have very small reverse transmission characteristics $\propto 1/A_e$ [2], making accurate small-signal measurements of such devices challenging. The bandwidth of transistors exceeds the capabilities of state-of-the-art vector network analyzers (VNAs). However, measurements from $0.5\,\mathrm{GHz}$ to $67\,\mathrm{GHz}$ give sufficient indication of bandwidths since the $3\,\mathrm{dB}$ frequency at which current h_{21} and unilateral gains U start to roll off at $20\,\mathrm{dB}$ per decade is below $30\,\mathrm{GHz}$, depending on the device geometry.

A two step process is used to correct for the parasitic effects of cabling, probes and on-wafer interconnect structures: in the first step, reference planes are carried to the probe tips by LRRM calibration on a commercial $\mathrm{Al_2O_3}$ substrate: known through, short, open and match structures on this substrate are measured, and error terms are calculated from these measurements with which the cable/probe delays and losses can be stripped from measured S parameters [3]. Unlike other calibration methods (e.g. SOLT) that rely on well-defined short and open structures over the entire frequency range, LRRM calibration does not require carefully realized short and open, making it less ill-conditioned for frequencies above $20\,\mathrm{GHz}$. Also, reproducibility and repeatability of measurements taken after LRRM calibration usually surpasses those

obtained after SOLT calibration [4]. The calibration is validated by measuring a transmission line structure on the Al_2O_3 substrate: when the insertion loss is less than $-40\,dB$ over the entire frequency range at both ports and the phase of S_{12} and S_{21} is linear, the calibration is considered valid.

In the second step, on-wafer open and short structures are measured [5]: an open interconnect structure has the device removed. For the short structure, the device is also not present, but its footprint is filled with interconnect metal. The transistor parameters Y_{DUT} can be obtained by subtracting open Y_{open} and short $Z_{short} = (Y_{open} - Y_{short})^{-1}$ from the measured parameters $Y_{measured}$:

$$Y_{DUT} = \left((Y_{measured} - Y_{open})^{-1} - Z_{short}\right)^{-1} \tag{4.1}$$

In previous interconnect designs [6], the ground plane was shared between neighboring CPW structures, introducing undesired resonance effects in the measured unilateral gain U at frequencies above $30\,GHz$ that add ambiguity to the extraction of f_{max}. By separating ground planes of adjacent coplanar waveguide structures, these resonance effects have been suppressed.

In the calibration of the VNA, it is assumed that the RF signal propagates in a single well-established mode between probe and RF structures that is identical between calibration substrate and wafer. This assumption is violated, introducing an error to the measurement. In addition, the co-planar waveguide structures used

for the transistors discussed in this chapter are not well-isolated from the substrate, potentially exciting parasitic modes at resonant frequencies that further deteriorate RF measurements.

Insufficient isolation between probes adds another source of error to the measurement: close probe spacing $\approx 170\,\mu$m required to minimize signal line losses and ensure single mode operation causes probe-to-probe coupling, decreasing the accuracy of device characterization.

These limitations have motivated the development of a process for fabricating microstrip lines: the back-end process is extended to produce an additional layer of BCB and metal [7]. Metal 1 is used as ground plane, shielding signal lines from the substrate. Probe spacing is increased to $280\,\mu$m. Most importantly, a complete set of calibration structures is fabricated along with transistor wiring that enables *on-wafer* multi-line TRL (through, reflect, line) calibration, eliminating the ambiguity of the two-step calibration process. Further accuracy improvements can be attained by reducing the pitch of RF probes. While this process has been established on test samples and incorporated into mask layouts with which several samples have been fabricated, it has not yet been executed on HBT samples.

T (nm)	Material	Doping (cm^{-3})	Description
10	In$_{0.53}$Ga$_{0.47}$As	8×10^{19} : Si	Emitter Cap
20	InP	5×10^{19} : Si	Emitter
15	InP	2×10^{18} : Si	Emitter
30	In$_{\approx 0.5}$Ga$_{\approx 0.5}$As	9–5×10^{19} : C	Base
13.5	In$_{0.53}$Ga$_{0.47}$As	5×10^{16} : Si	Setback
16.5	InGaAs/InAlAs	5×10^{16} : Si	B-C Grade
3	InP	3.6×10^{18} : Si	Pulse Doping
67	InP	5×10^{16} : Si	Drift Collector
7.5	InP	2×10^{19} : Si	Sub-Collector
5	In$_{0.53}$Ga$_{0.47}$As	4×10^{19} : Si	Sub-Collector
300	InP	1×10^{19} : Si	Sub-Collector
5	In$_{0.53}$Ga$_{0.47}$As	NID	Etch Stop
≈ 625k	SI InP		Substrate

Table 4.1: Epitaxial Structure Design of HBT56.

4.2 HBT 56

4.2.1 Epitaxial Design

The wafer HBT56 has been grown by solid source molecular beam epitaxy on a 4" InP substrate by IQE. The n-In$_{0.53}$Ga$_{0.47}$As emitter cap is highly doped for low emitter resistance R_{ex}. The 30 nm thick base is doping-graded from 9–5×10^{19}/cm^3, resulting in 55 meV conduction band slope. Numerical calculations indicate a base transit time $\tau_b \approx 71$ fs. The 100 nm thick collector is comprised of a 13.5 nm setback, a 16.5 nm chirped superlattice InGaAs/InAlAs grade and a 67 nm drift collector region (Table 4.1). The collector doping 5×10^{16}/cm^3 is slightly above the maximum
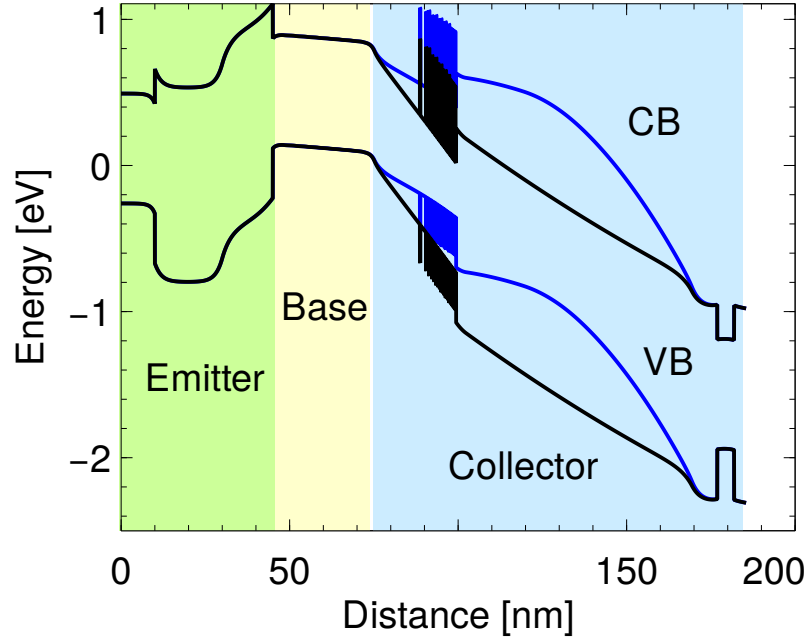
Figure 4.1: Numerically calculated band structure of HBT56 for J_e=0 mA/µm^2 (black) and J_e=24 mA/µm^2 (blue), V_{cb}=1 V, V_{be}=1 V, V_{cb} =0.7 V. A current spreading factor of $J_e/J_c \approx 2$ was assumed.

doping concentration 4.2×10^{16}/cm^3 at which the collector would be fully depleted

at zero bias.

Figure 4.1 shows the band diagram for a transistor under bias at zero and at Kirk

collector current density.

4.2.2   Fabrication

The sample has a composite emitter stack of 20/250/250 nm Mo/TiW/W. The

thickness of first and second emitter sidewall has been both 30 nm as deposited.

For the base electrode, 3.5 nm Pt has been lifted off in the UCSB cleanroom. The

remainder of the stack 12/17/70 nm Ti/Pd/Au has been evaporated at Teledyne
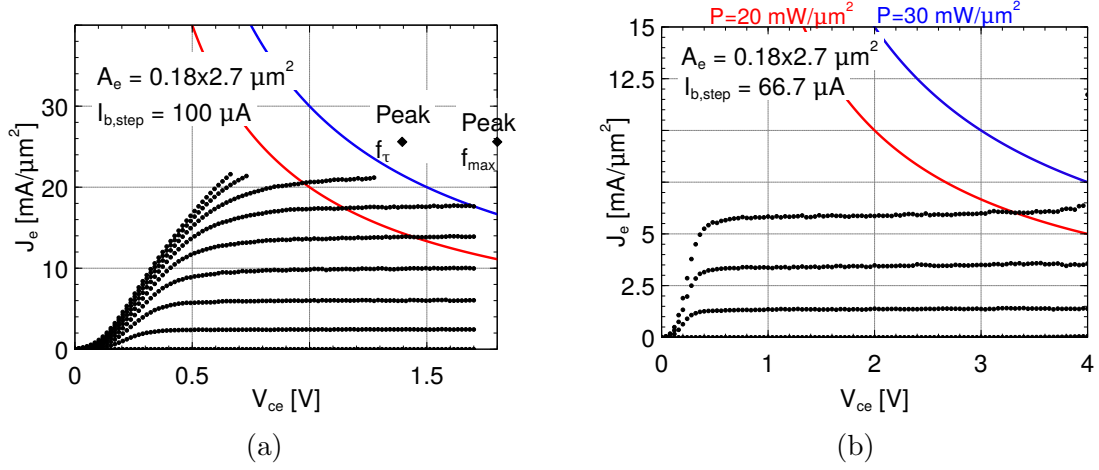
Figure 4.2: Common emitter I–V characteristics for a transistor with emitter junction area 180 nm x 2.7 µm for (a) low voltage and (b) high voltage operation.

Scientific due to issues with the evaporators in the UCSB cleanroom at the time of fabrication (see section 3.3.2). Prior to deposition of collector contacts, the surface was prepared with a dip in diluted hydrochloric acid after development of photoresist: oxidation in UV O_3 has been omitted.

The sample has been passivated with a dip in diluted hydrochloric acid followed by a PECVD SiN_x deposition prior to BCB planarization.

4.2.3   DC Characteristics

Extractions from base TLM measurements indicate a base contact resistivity $\rho_c \approx 9\,\Omega\,\mu m^2$. The sheet resistance of unpinched and pinched base TLM structures is estimated at 1200 and 810 $\Omega/\square$, indicating process damage to the extrinsic base regions. Collector TLMs show unusually high contact resistivity $\rho_{coll} \approx 55\,\Omega\,\mu m^2$. Emitter access resistivity extracted from RF measurements is $\rho_{em,xs} \approx 3.5\,\Omega\,\mu m^2$.
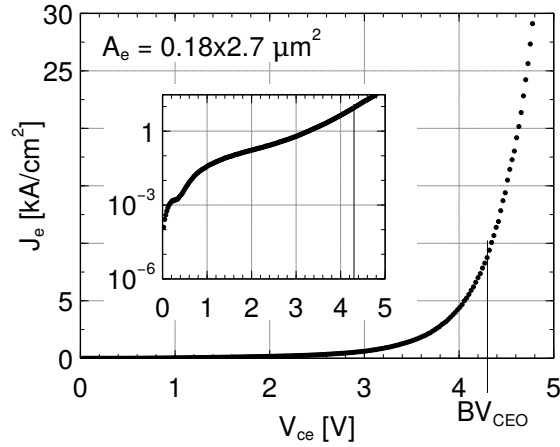
Figure 4.3: Common-emitter breakdown measurement with floating base of a transistor with junction area 180 nm x 2.7 µm. BV_{ceo} 4.3 V has been extracted when the emitter current density is $10\,kA/cm^2$.
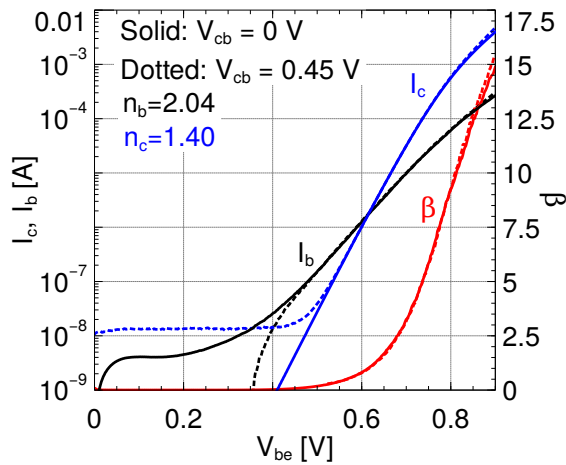
Figure 4.4: Gummel characteristics for an HBT with 180 nm x 2.7 µm emitter junction area.
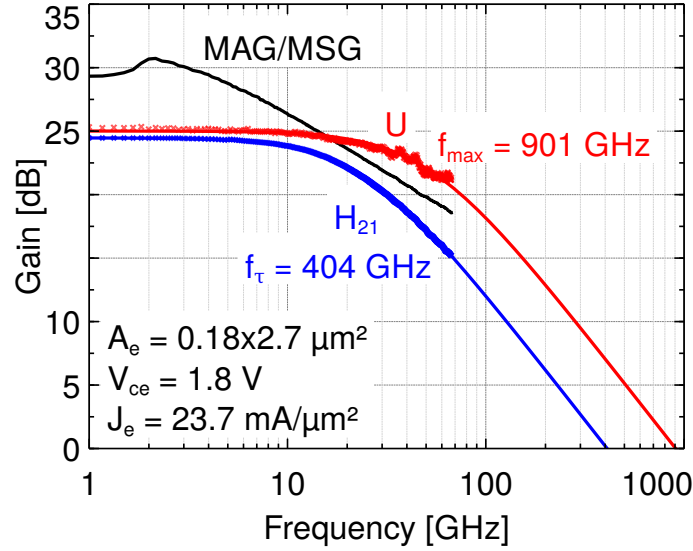
Figure 4.5: Microwave gains for an HBT with $180\,\text{nm}$ x $2.7\,\mu\text{m}$ emitter junction area. Single pole fit yields f_τ $401\,\text{GHz}$, $f_{\max}$ $901\,\text{GHz}$.

Common I/V characteristics for low and high voltage operation are shown in Figure 4.2. Figure 4.3 shows breakdown measurement with common-emitter breakdown voltage $BV_{ceo} = 4.3\,\text{V}$ at $J_e = 10\,\text{kA/cm}^2$. From devices with emitter-base junction geometries identical to the device with best RF performance, a peak current gain β of 16 can be observed (Figure 4.4).

4.2.4   Microwave Characteristics

Figure 4.5 shows microwave current gain H_{21}, unilateral gain U and maximum available / maximum stable gain MAG/MSG for a device with $180\,\text{nm}$ wide emitter junction. Peak RF performance was obtained at $J_e = 23.7\,\text{mA/}\mu\text{m}^2$, $V_{ce} = 1.8\,\text{V}$, $I_c = 10\,\text{mA}$, $V_{cb} = 0.88\,\text{V}$, $P/A_{ej} = 42.7\,\text{mW/}\mu\text{m}^2$, $C_{cb}/I_c = 270\,\text{ps/V}$. Single pole fit yields f_τ $401\,\text{GHz}$, $f_{\max}$ $901\,\text{GHz}$. Kirk effect is observed when f_τ falls to 95 % of its
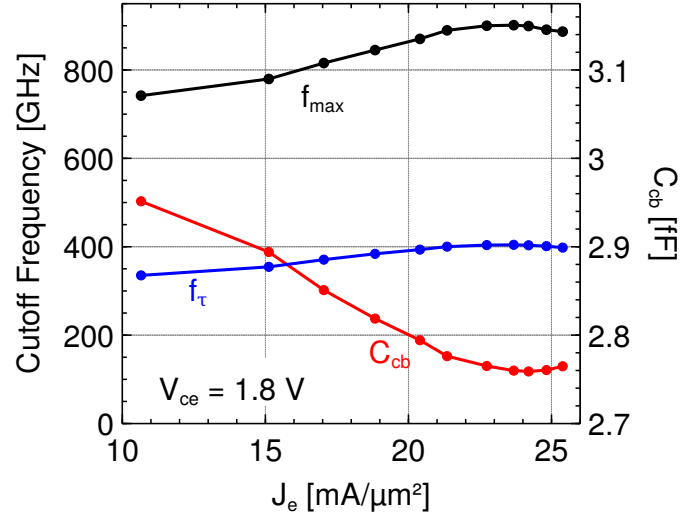
Figure 4.6: Variation of f_τ, f_{max} and C_{cb} with J_e at $V_{ce} = 1.8\,\mathrm{V}$ for an HBT with 180 nm x 2.7 µm emitter junction area and 310 nm base-collector mesa width.

peak value (Figure 4.6).

A small signal equivalent hybrid-π circuit has been developed from RF measurements exhibiting good agreement between measured and simulated S parameters (Figure 4.7).

4.2.5   TEM Analysis

Cross-sectional transmission electron micrographs have been obtained to determine exact device dimensions and information about the fabrication process (Figure 4.8). The emitter junction width has been 20 nm wider than drawn due to variations in the dry etch. The Pt base contact has sunk $\approx 4\,\mathrm{nm}$ into the base. Accidental deposition of base metal onto emitter sidewalls and damage to the semiconductor between base electrodes and emitter is observed. The base/collector mesa has been

$C_{cb,x}$ = 3.03 fF

R_{cb} = 15.5 kΩ

R_{bb} = 27.3 Ω $C_{cb,i}$ = 0.81 fF R_{cc} = 1.8 Ω

Base   Col.

R_{be} = 56 Ω

C_{cg} = 3.7 fF

0.26S V_{be} $e^{-j\omega\,0.210ps}$

C_{je} + C_{diff} = 2.4 fF + 53.5 fF τ_b = 71 fS

R_{ex} = 5.0 Ω

Emitter

(a)

$S_{21}/10$ S_{12}x5

measured
simulated

S_{11} S_{22}
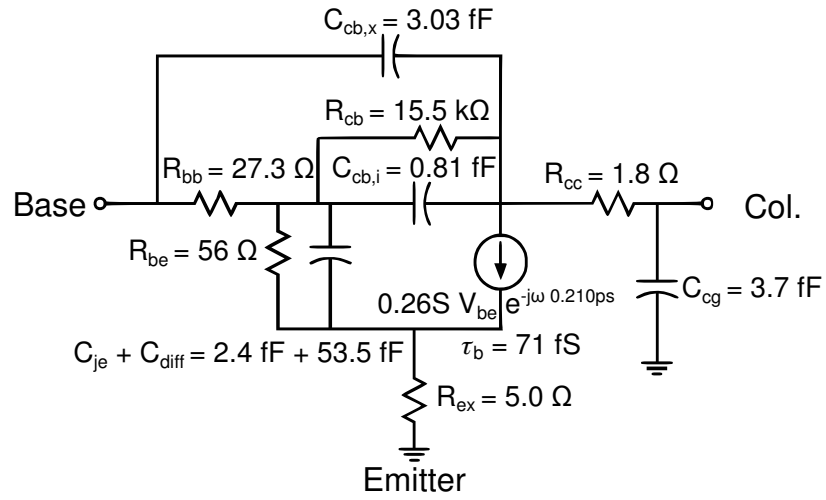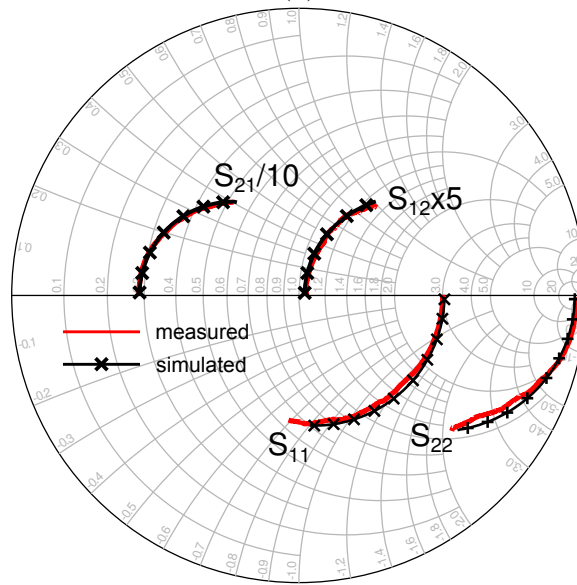
(b)

Figure 4.7: (a) A hybrid-π equivalent circuit for the HBT at peak f_{max} performance. (b) Comparison of (solid line) measured S-parameters of Figure 4.5 and (x) simulated S-parameters from 0.5 to 67 GHz .

94
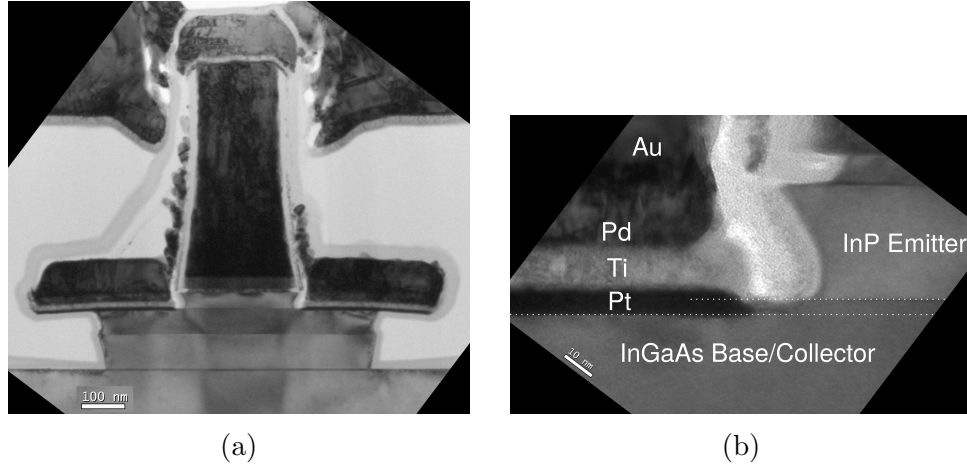
Figure 4.8: Cross-sectional TEMs of (a) the entire device, (b) magnified at the emitter-base region. Emitter junction width $w_e = 240\,\text{nm}$, single-sided base metal width $w_{bm} = 220\,\text{nm}$, single-sided base mesa undercut $w_{bmu} = 125\,\text{nm}$, emitter-base contact spacing $w_{Gap} \approx 12\,\text{nm}$.

undercut $w_{bmu} = 125\,\text{nm}$. Emitter-to-base misalignment is less than $20\,\text{nm}$.

4.2.6   Discussion

A sample from the same epitaxial wafer has been fabricated exhibiting f_τ of $480\,\text{GHz}$ and $f_{\max}$ of $1\,\text{THz}$ at large base-collector junction width due to accidental misalignment [8]. Although sample HBT56J shows excellent alignment at reduced base-collector width, the observed $f_\tau / f_{\max}$ bandwidth has been lower.

High base contact resistivity has been identified as the main detractor of $f_{\max}$ bandwidth: the interruption of base metal deposition due to fabrication circumstances has caused increased contact resistivity $\rho_c > 9\,\Omega\,\mu\text{m}^2$. Although a few devices have exhibited $f_{\max}$ bandwidth of $900\,\text{GHz}$, the bulk of the measured devices showed $700\,\text{GHz}$ $f_{\max}$, indicating inconsistencies with base ohmics across the sample. Approximately
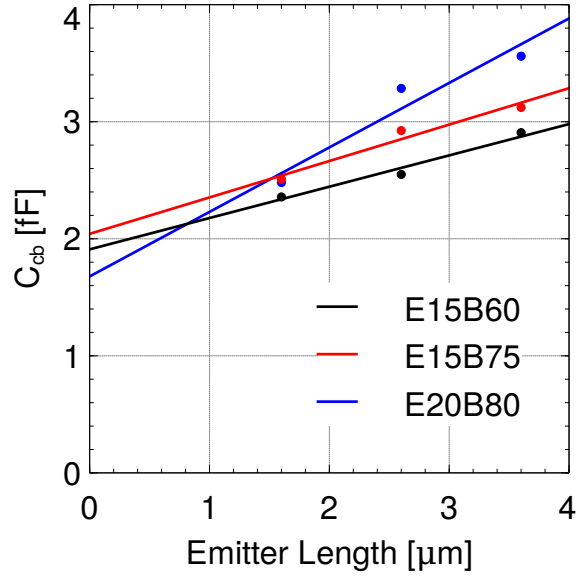
Figure 4.9: Variation of base-collector capacitance C_{cb} versus emitter length L_e for different device geometries. Intersect $C_{cb,\text{excess}} \approx 2\,\text{fF}$.

$3\,\Omega\,\mu\text{m}^2$ of the total extracted contact resisitivity on $L_e = 2.7\,\mu\text{m}$ transistors is due to the finite sheet resistance of the base metal $\approx 0.8\,\Omega/\square$. The $f_{\max}$ bandwidth has been further limited by excess capacitance caused by insufficiently undercut base posts (Figure 4.9). The RF bandwidth of devices with smallest junction area is reduced due to disproportionally high C_{cb}.

Accidental deposits of base metal onto the emitter sidewall have increased C_{be}, while emitter end undercut $200\,\text{nm}$ per side has decreased the active devices area, further reducing RF bandwidths. In addition, emitter end undercut $\approx 150\,\text{nm}$ per side has reduced C_{cb}/I_c.

High collector contact resistivity is due to surface contamination. A more thorough surface cleaning procedure (section 3.4) has therefore been integrated into the process.

The breakdown voltage has been increased from $3.7\,\mathrm{V}$ to $4.3\,\mathrm{V}$ as a result of the enhanced surface passivation process (see section 3.5.1). The passivation process has also improved yield.

4.3   Scaled TLM Results

With f_{max} bandwidth of HBT56J below expectations, a series of experiments exploiting the fast-turnaround scaled TLM process (section 3.6) has been conducted to find root causes of high base contact resistivity. At $100\,\mathrm{nm}$ collector thickness HBT generation, HBT base contact fabrication needs to yield contact resistivity below $4\,\Omega\,\mu\mathrm{m}^2$ [10]. While the contact resistivity of refractory metals to highly doped layers of thick p-InGaAs is low (Figure 4.10), integration into the HBT process has been unsuccessful: an HBT sample fabricated with a Ru contact to a $18\,\mathrm{nm}$ thick p-InGaAs base p-doped from $14\text{--}9 \times 10^{19}/\mathrm{cm}^3$ has exhibited contact resistivity of $19.2\,\Omega\,\mu\mathrm{m}^2$ (Figure 4.11), indicating contamination at the interface from prior processing. Previous TLM experiments at UCSB and published reports [11] indicate better contact resistivity to p-InGaAs with Pd or Pt than with refractories. The scaled TLM experiments have been therefore executed with Pt as base contact metal that has a shallower interdiffusion depth than Pd.

The effect of processing prior to base metal deposition has been quantified by fabricating scaled TLMs without emitter fabrication. In order to assess the thermal stability of the contacts, samples have been measured before and after a thermal
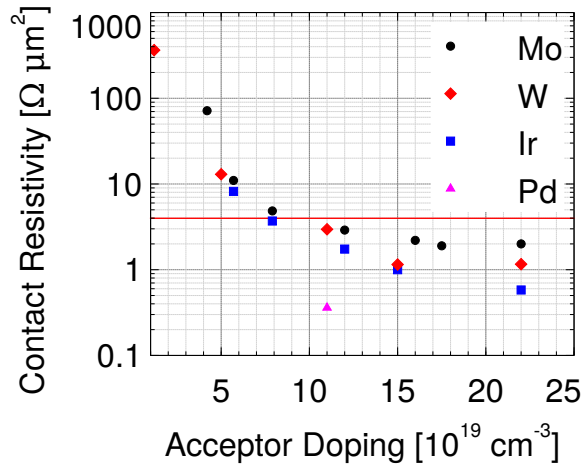
Figure 4.10: Variation of contact resistivity ρ_c to p-InGaAs at different doping levels for metals Mo, W, Ir and Pd [9]. The red line indicates the upper limit for 100 nm collector thickness HBT node $\rho_c = 4\,\Omega\,\mu m^2$.
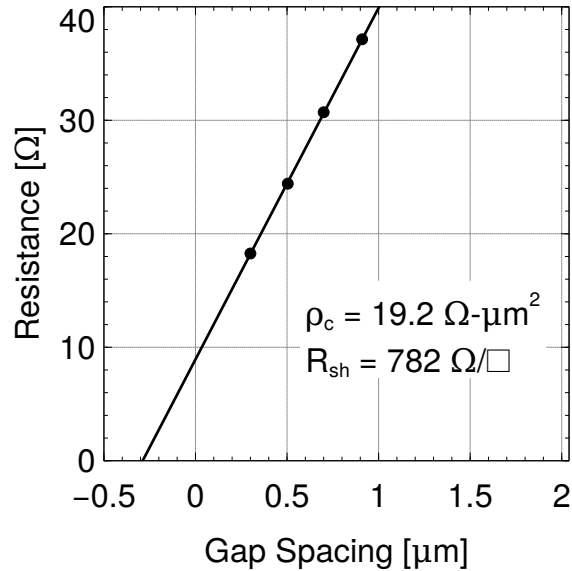
Figure 4.11: Measured base TLM resistance as a function of gap spacing on sample HBT65D: 15 nm Ru contact to 18 nm p-InGaAs base with a doping grade from 14– $9 \times 10^{19}/cm^3$.

(a)   (b)
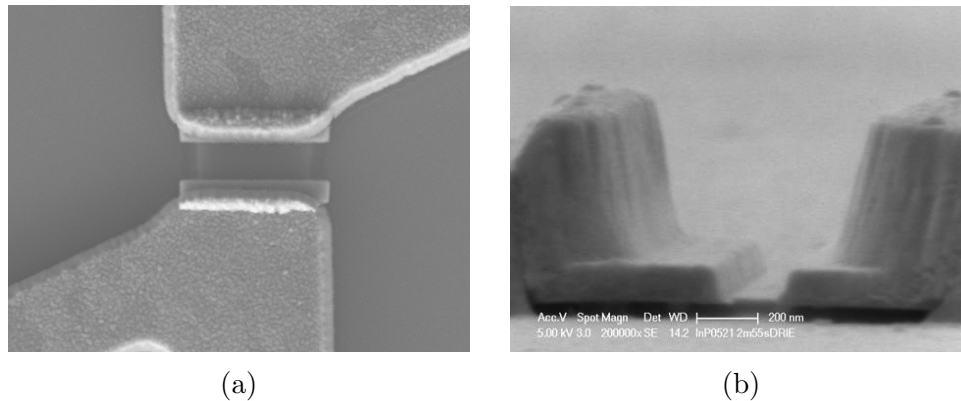
Figure 4.12: (a) Top-down and (b) 85° SEMs of a scaled TLM on sample 121217E fabricated without emitter processing.

anneal that is identical to the cure required for BCB planarization. Changes to the composition of the base metal stack have been made so as to attempt improving base contact resistivity.

4.3.1   Sample 121217E

The fabrication process of scaled TLM and extraction of contact resistivity has been verified on a sample with $25\,\text{nm}$ thick p-InGaAs doped at $\approx 1.4 \times 10^{20}/\text{cm}^3$ on an InP buffer. The sample has been grown in the UCSB MBE on a 2" InP substrate. The sample has been much higher doped than HBT base layers and lacks the doping grade, but it provides a baseline for subsequent scaled TLM samples and a point of comparison to TLMs fabricated with i-line lithography [9]. The TLM pads have been deposited using a PMGI/ZEP EBL lift-off process with a dip in diluted hydrochloric acid as surface preparation just prior to evaporation. Figure 4.12 shows SEMs of TLM structures after fabrication.
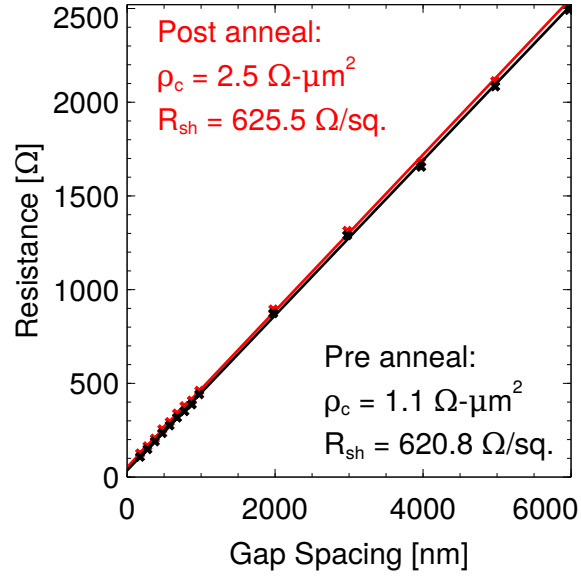
Figure 4.13: Measured TLM resistance as a function of pad spacing for the sample 121217E.

The extracted contact resistivity prior to annealing was $1.1\,\Omega\,\mu m^2$. After the $1\,h$ $250\,°C$ anneal, the contact resistivity has increased to $2.5\,\Omega\,\mu m^2$. The achieved contact resistivity is sufficiently low for the HBT scaling node discussed in this work: however, with p-InGaAs doping of the sample much higher than the doping level of an HBT base, the experimental result is not directly transferable to the HBT process. Also, with the extracted contact resistivity significantly above estimations derived from prior results (Figure 4.10), surface contamination from lithographic processing is suspected.
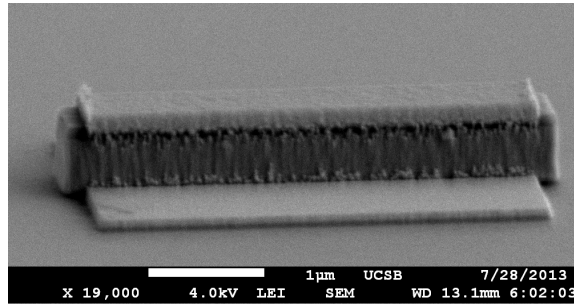
Figure 4.14: SEM of pinched TLM structure on TLMv3D after deposition of TLM electrodes.

4.3.2 Scaled TLM Samples with Emitter Processing

Scaled TLMs with emitter-defined gaps have been fabricated: the sample has been cleaved from a commercially grown wafer with HBT layer structure. The base on this sample is 25 nm thick p-InGaAs with a doping grade from 9×10^{19} to $5 \times 10^{19}/\text{cm}^3$. Emitter metal has been deposited and emitter stripes have been defined in a dry etch process (section 3.1). A single 30 nm thick sidewall has been deposited after etching the InGaAs emitter cap. The hardmask was not removed. After sidewall formation, the sample containing multiple dice has been cleaved into several pieces on which several different scaled TLM fabrication campaigns have been executed.

TLMv3D

On this sample, the emitter surface has been prepared with an oxide removal dip 10 s in $NH_4OH:H_2O$ 1:10 followed by 60 s H_2O rinse. The emitter was removed in a wet etch 9 s $H_3PO_4:HCl$ 4:1. After emitter wet etch, residues have been observed with SEM (Figure 3.6). A solvent clean has partially removed the residues from the
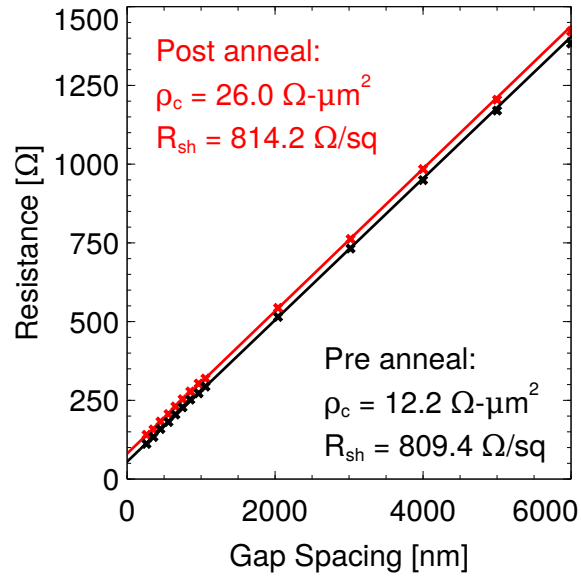
Figure 4.15: Measured TLM resistance as a function of pad spacing for the sample TLMv3D.

critical base surface around the emitter. For the base electrode metal, a stack of 15/12/16/65 nm Pt/Ti/Pd/Au has been lifted off (Figure 4.14) after a deoxidizing dip in HCl:H_2O 1:10.

Before the thermal anneal, the measured contact resistivity was $12\,\Omega\,\mu m^2$. After the anneal, the contact resistivity has increased to $26\,\Omega\,\mu m^2$ (Figure 4.15).

TLMv3C

Prior to the wet etch, the emitter surface has been oxidized in a UV ozone reactor. The oxide was removed with a 10 s dip in diluted HCl:H_2O 1:10 followed by a 60 s H_2O rinse, substituting the basic oxide removal solution with an acidic one. Wet bench lights have been shut off during the oxide removal and wet etches. SEMs of the sample show no indication of surface contamination (Figure 3.7). The remainder
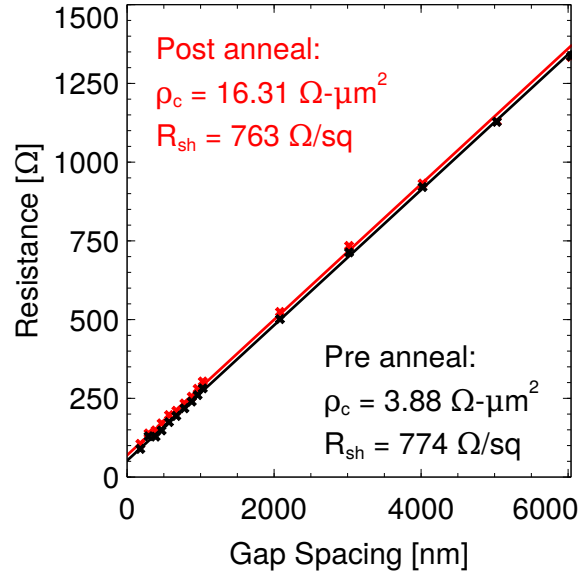
Figure 4.16: Measured TLM resistance as a function of pad spacing for the sample TLMv3C.

of the fabrication process is identical to TLMv3D.

Before the thermal anneal, the measured contact resistivity was $4\,\Omega\,\mu m^2$, indicating an improvement due to changed surface preparation. After the anneal, the contact resistivity has increased to $16\,\Omega\,\mu m^2$ (Figure 4.16), suggesting thermal contact degradation due to deep interdiffusion of $15\,nm$ Pt with the InGaAs base.

TLMv3O

On TLMv3O, the improved surface preparation has been retained. The thickness of the initial Pt contact layer of the base electrode has been reduced to $3\,nm$. Instead of Ti, a layer of Al of unknown thickness has been accidentally deposited.

Before the thermal anneal, the measured contact resistivity was $5\,\Omega\,\mu m^2$. After the anneal, TLM structures have exhibited low resistance with almost no dependency
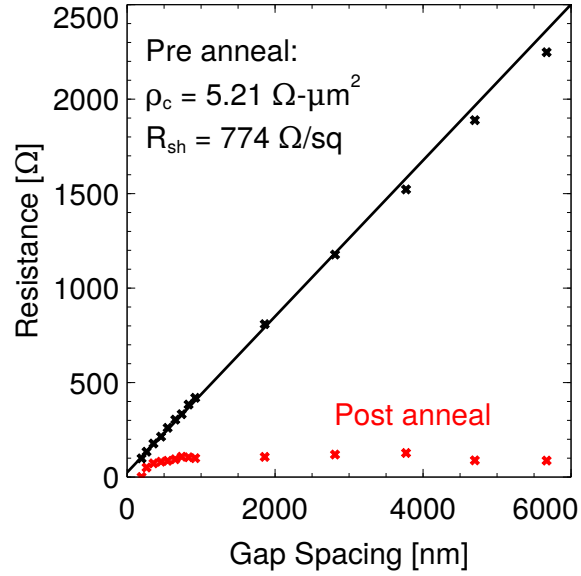
Figure 4.17: Measured TLM resistance as a function of pad spacing for the sample TLMv3O.

on gap spacing, indicating that base electrode metal has diffused through the entire

25 nm thick base (Figure 4.17).

TLMv3R

On this sample, the correct metal stack has been deposited. Measured contact

resistivity was $3\,\Omega\,\mu m^2$ before annealing and $10\,\Omega\,\mu m^2$ after the thermal bake.

4.3.3   Discussion

The results of TLMv3R and TLMv3O indicate that an initial base contact layer

of 3 nm Pt fully reacts with the base during thermal anneals, losing its property as a

diffusion barrier between upper metal layers and InGaAs. With the initial layer of Pt

chosen too thick (TLMv3C), ohmic contacts on doping-graded bases are deteriorated

as a result of deep interdiffusion [12] that can even potentially short out the base-collector junctions.

In the literature [13], formation of Ti-As compounds at elevated temperatures is reported that degrade the ohmic interface. We suspect that with standard lifted-off 3/12/16/65 nm Pt/Ti/Pd/Au base electrodes, Pt-Ti-As compounds are formed during thermal processes that deteriorate the contact. A novel base metal process has been therefore developed that maintains a reactive Pt contact layer to overcome surface contaminants while exploiting the thermally stable refractory metal ruthenium (Ru) as a diffusion barrier (section 3.3.4). The initial base contact metal is deposited without any lithographic processing to minimize contamination. After forming base contacts, base and emitter semiconductor surfaces are passivated with Al_2O_3 that is thermally grown in an atomic layer deposition process, providing complete surface coverage due to conformal growth.

4.3.4   Scaled TLM Sample with Emitter and Dual-Deposited Base Processing

On sample TLMv3B, the dual-deposition base metalization process has been executed (Figure 4.18). Before the thermal anneal, the measured contact resistivity was $1.5\,\Omega\,\mu m^2$, increasing to $3\,\Omega\,\mu m^2$ after the anneal (Figure 4.19). The contact resistivity yielded on this sample is sufficiently so this process has been used on all subsequent HBT fabrication campaigns.
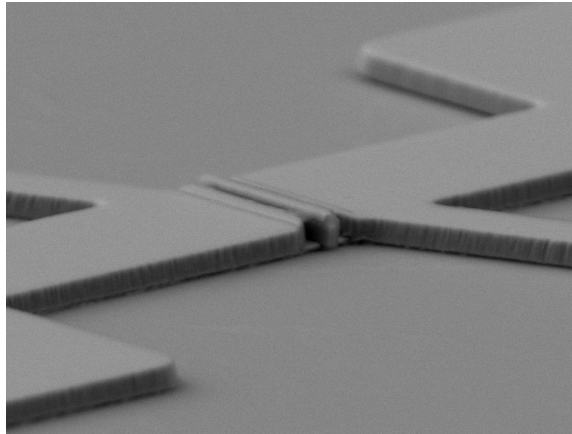
Figure 4.18: SEM of pinched TLM structure on TLMv3B at 80° after fabrication.
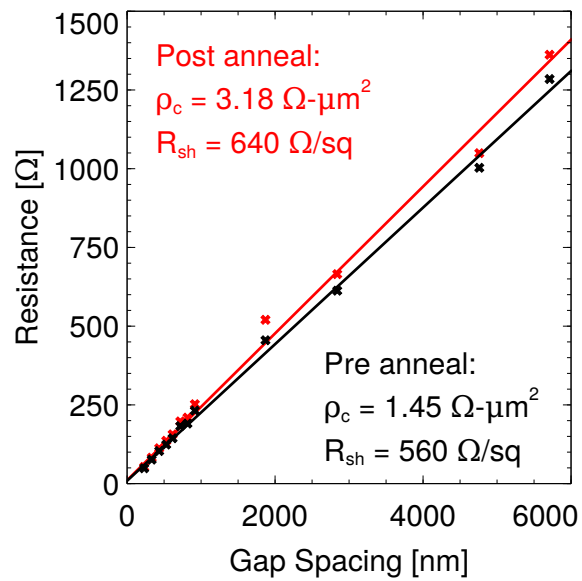
Figure 4.19: Measured TLM resistance as a function of pad spacing for the sample TLMv3B.

4.4   HBT 64

4.4.1   Epitaxial Design

The wafer HBT64 has been grown by solid source molecular beam epitaxy on a 4" InP substrate by IQE. In comparison to HBT56, the doping of the space charge region in the emitter has been increased from $2 \times 10^{18}/\mathrm{cm}^3$ to $5 \times 10^{18}/\mathrm{cm}^3$ to reduce source starvation effects. Base thickness was reduced from $30\,\mathrm{nm}$ to $20\,\mathrm{nm}$ while doping has been simultaneously increased from $9\text{--}5 \times 10^{19}/\mathrm{cm}^3$ to $11\text{--}7 \times 10^{19}/\mathrm{cm}^3$, resulting in $90\,\mathrm{meV}$ conduction band slope. Numerical calculations indicate a base transit time $\tau_b \approx 46\,\mathrm{fs}$. The collector and subcollector design has not been changed between HBT56 and HBT64. Table 4.2 lists the epitaxial layer structure. Figure 4.20 shows the band diagram under bias at zero and at Kirk collector current density.

4.4.2   Fabrication

Two samples have been fabricated from the wafer: HBT64C and HBT64J. The emitter hardmask removal on 64C has been executed with photoresist planarization. On 64J, the photoresist planarization has been skipped. Two emitter sidewalls ($30\,\mathrm{nm}$ as deposited) have been formed on both samples. The emitter etch has been prepared on both samples with oxidation in a UV O_3 reactor and oxide removal in HCl:DI. Emitter etching time has been reduced from $8\,\mathrm{s}$ to $5\,\mathrm{s}$ to successfully reduce emitter end undercut (refer to section 3.2.4). The dual base metal process has been executed

T (nm)	Material	Doping (cm^{-3})	Description
10	In$_{0.53}$Ga$_{0.47}$As	8×10^{19} : Si	Emitter Cap
15	InP	5×10^{19} : Si	Emitter
15	InP	5×10^{18} : Si	Emitter
20	In$_{\approx 0.5}$Ga$_{\approx 0.5}$As	11-7 $\times 10^{19}$: C	Base
13.5	In$_{0.53}$Ga$_{0.47}$As	5×10^{16} : Si	Setback
16.5	InGaAs/InAlAs	5×10^{16} : Si	B-C Grade
3	InP	3.6×10^{18} : Si	Pulse Doping
67	InP	5×10^{16} : Si	Drift Collector
7.5	InP	2×10^{19} : Si	Sub-Collector
7.5	In$_{0.53}$Ga$_{0.47}$As	4×10^{19} : Si	Sub-Collector
300	InP	1×10^{19} : Si	Sub-Collector
3.5	In$_{0.53}$Ga$_{0.47}$As	NID	Etch Stop
$\approx$ 625k	SI InP		Substrate

Table 4.2: Epitaxial Structure Design of HBT64.
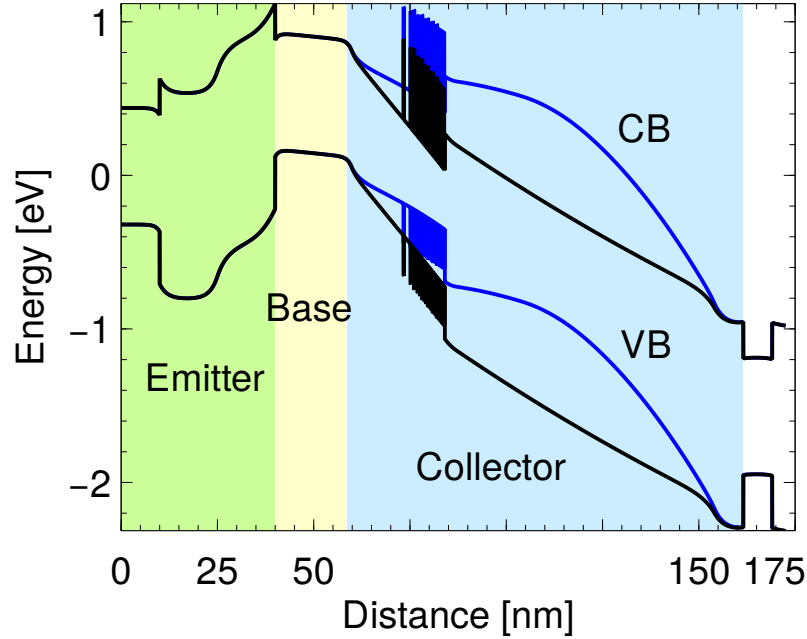
Figure 4.20: Numerically calculated band structure of HBT64 for J_e=0 mA/µm^2 (black) and J_e=18 mA/µm^2 (blue), V_{cb}=1 V, V_{be}=1 V, V_{cb} =0.7 V. A current spreading factor of $J_e/J_c \approx 1.5$ was assumed.
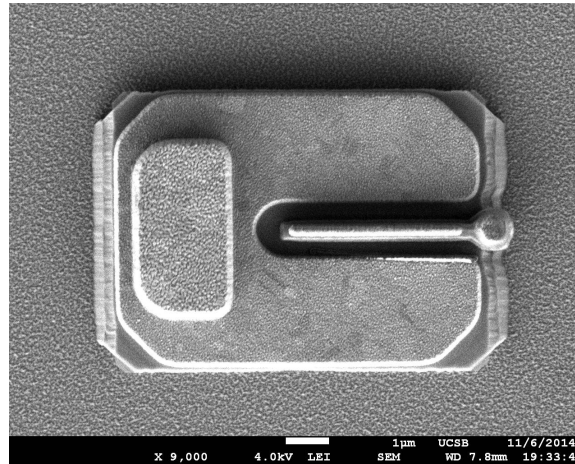
Figure 4.21: Top-down SEM of a fabricated transistor on sample HBT64J before planarization indicating rough surface in the field.

(a)   (b)
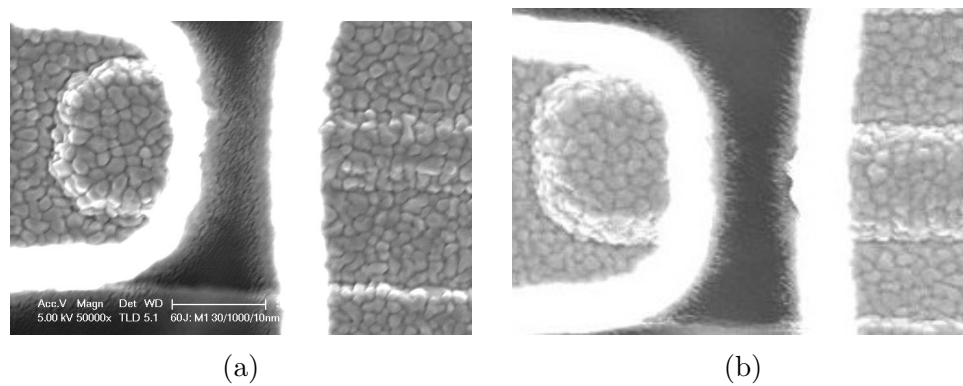
Figure 4.22: Accidental emitter-to-base shorts on metal 1 interconnect layer (a) before and (b) after dry etch.

on both samples, only varying the thickness of the initial Pt 1 nm on 64C, 2 nm

on 64J (section 3.3.4). A passivation layer of thermal Al_2O_3 grown at $200\,°C$ for

≈ 90 ALD cycles has been formed, resulting in 10 nm thick films. A single base

sidewall (30 nm as deposited) was formed in a standard process. The Al_2O_3 has

been removed in a wet etch of 1:50 $BHF:H_2O$ for $\approx 30\,s$. Base electrodes of 5/95 nm

have been lifted off in a PMGI/CSAR electron beam lithography process: prior to

metal deposition, photoresist residues have been removed with an oxygen plasma.

Reduced diameter base posts (section 3.3.5) have been deposited. The base contact

metal in the field has been removed in a composite dry etch of 40 s Cl_2/O_2 and 25 s

Cl_2/Ar. Base/collector mesas have been isolated, collector electrodes were formed

after surface clean (section 3.4), collector posts have been lifted off. The transistors

were isolated in a selective wet etch: on 64J, the wet etch has left residues in the

field on parts of the sample (Figure 4.21): the Ar sputtering step of the base dry etch

has been too short, leaving contaminants on the surface that sank down during wet

etches. A nitride layer was grown prior to BCB planarization (section 3.5.1). Prior

to metal 1 lithography, photoresist residues have been sputtered off base posts and

emitter stripes. Accidental shorts of metal 1 between base and emitter interconnect

electrodes due to poor photoresist adhesion in that region required a 20 s sputtering

etch (20 sccm Ar at 1 Pa, 300 W RF, 50 W bias) to clear out excess metal (Figure 4.22).
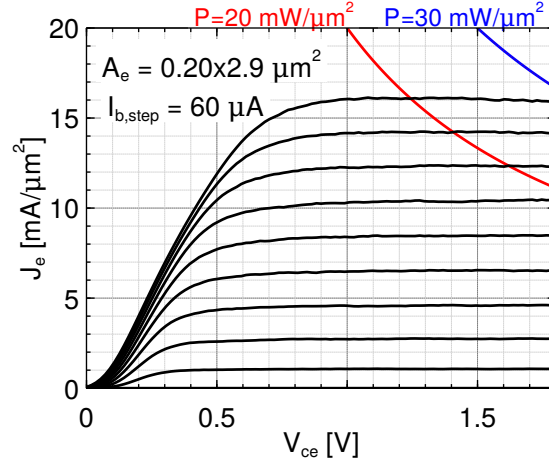
Figure 4.23: Common-emitter characteristics of a transistor with junction area $200\,\text{nm}$ x $2.9\,\mu\text{m}$.

4.4.3   DC Characteristics

Base TLM structures have been inadvertently damaged during fabrication: measured resistances do not correlate with gap width. Collector TLMs indicate contact resistivity $\rho_{\text{coll}} \approx 16\,\Omega\,\mu\text{m}^2$. Emitter access resistivity extracted from RF measurements is $\rho_{\text{em,xs}} \approx 3\,\Omega\,\mu\text{m}^2$.

Gummel and common-emitter DC characteristics are shown in Figures 4.23 and 4.24. Peak current gain of the device with highest measurable f_{max} bandwidth is ≈ 16. Base and collector ideality are similar to sample HBT56.

Floating base breakdown measurements indicate $BV_{ceo} = 4.1\,\text{V}$ on sample 64J (Figure 4.25).
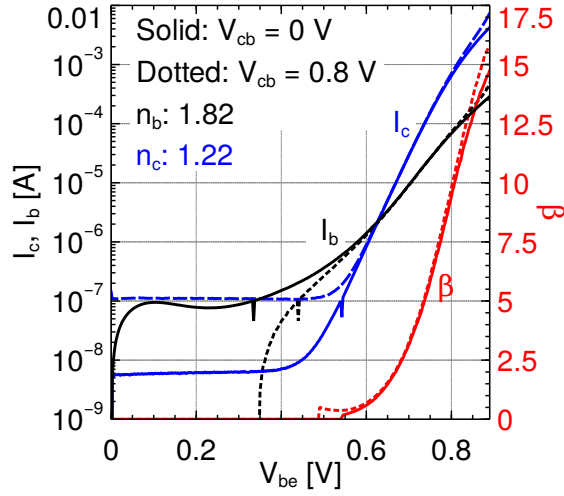
Figure 4.24: Gummel characteristics for an HBT with $200\,\text{nm}$ x $2.9\,\mu\text{m}$ emitter junction area.
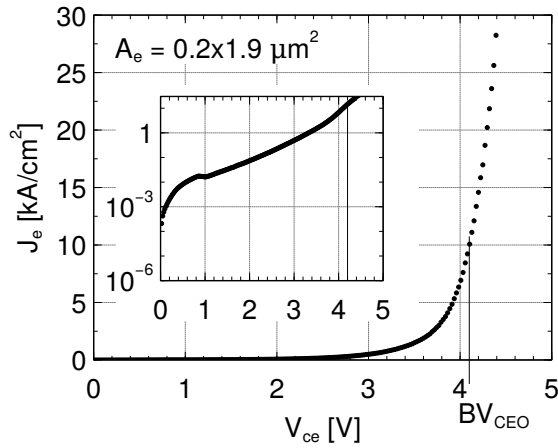
Figure 4.25: Common-emitter breakdown measurement with floating base of a transistor with junction area $200\,\text{nm}$ x $1.9\,\mu\text{m}$. $BV_{ceo} = 4.1\,\text{V}$ has been extracted when the emitter current density is $10\,\text{kA/cm}^2$.
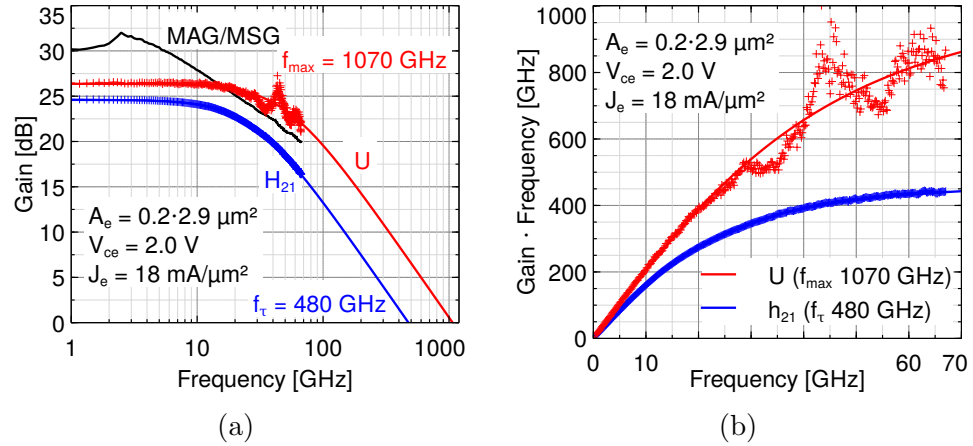
Figure 4.26: Microwave gains for an HBT with 200 nm x 2.9 µm emitter junction area. (a) Double logarithmic plot, (b) linear plot of gain·frequency vs frequency. Single pole fit yields f_τ 470 GHz, $f_{\max}$ 1070 GHz.
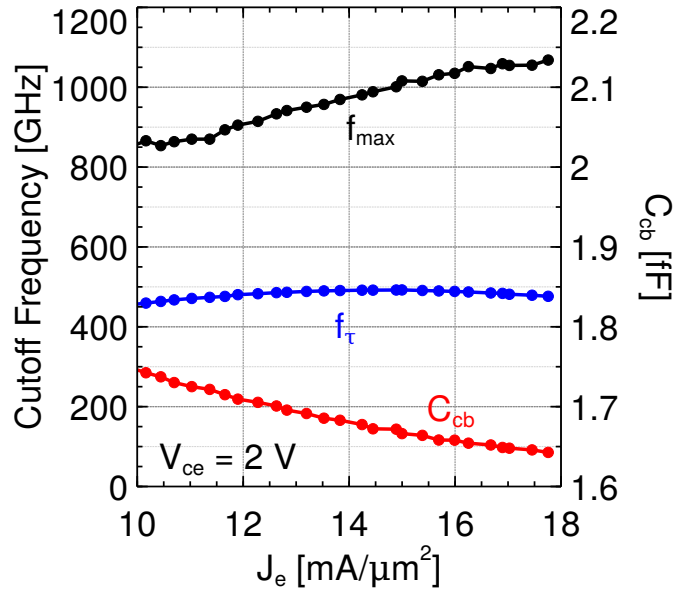
Figure 4.27: Variation of f_τ, f_{max} and C_{cb} with J_e at $V_{ce} = 2.0$ V for an HBT with 200 nm x 2.9 µm emitter junction area and 390 nm base-collector mesa width.

4.4.4   Microwave Characteristics

Microwave gains have been obtained from S-parameter measurements from 500 MHz to 67 GHz. On sample 64J, devices with heavily undercut base/collector mesas exhibit highest measurable RF performance (Figure 4.26): a least-squares fit to a single-pole roll-off shows f_τ 480 GHz, f_{max} 1070 GHz for a device with emitter junction area 200 nm ·2.9 µm at V_{ce} =2 V, J_e =18 mA/µm^2, P/A_{ej} =36 mW/µm^2, V_{cb} = 1.13 V, C_{cb}/I_c = 173 fs/V. The goodness of fit has been validated on a linear scale plot. Figure 4.27 shows variation of f_τ, f_{max} and C_{cb}on J_e. This device has not been electrically stressed: Kirk effect has been observed on other devices at $J_e = 19$ mA/µm^2 where f_τ drops to 95 % of its peak value.

Smaller footprint devices on 64J exhibit higher f_τ, but calibration artifacts in measured Mason's unilateral gain prevent reliable extraction of f_{max} bandwidth (Figure 4.28).

A small signal equivalent hybrid-π circuit has been developed from RF measurements of the device shown in Figure 4.26 exhibiting good agreement between measured and simulated S parameters (Figure 4.29).

On sample 64C, f_τ 480 GHz and f_{max} 910 GHz bandwidths have been observed on a transistor with emitter junction area 220 nm ·2.9 µm at V_{ce} =1.85 V, $J_e = 18$ mA/µm^2, $P/A_{ej} = 33.3$ mW/µm^2, $V_{cb} = 0.98$ V, $C_{cb}/I_c = 305$ fs/V.
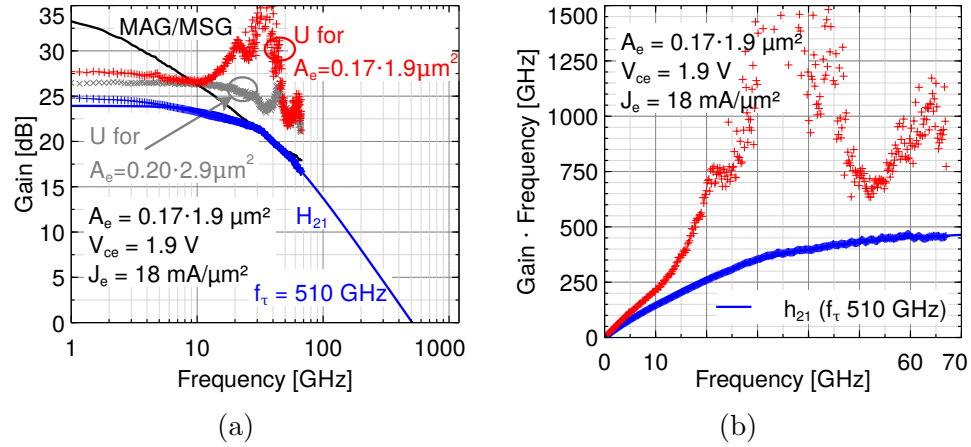
Figure 4.28: Microwave gains for an HBT with 170 nm x 1.9 µm emitter junction area. (a) Double logarithmic plot with unilateral gain of device in figure 4.26 drawn for comparison, (b) linear plot of gain·frequency vs frequency. Single pole fit yields f_τ 510 GHz.

4.4.5   TEM Analysis

Transmission electron micrographs have been obtained from sample HBT64A, a sample fabricated by Han-Wei Chiang using the conventional lifted-off base process exhibiting a maximum f_{max} 667 GHz and f_τ 476 GHz, as well as HBT64C and HBT64J. A composite image created from four separate TEMs comparing HBT64A with lifted-off base contacts to HBT64J with dual deposited base contacts is shown in Figure 4.31. The entire cross-section and a cutout of the emitter-base region is shown in Figure 4.33 for HBT64C and in Figure 4.32 for HBT64J. The interdiffusion depth of the base metal is ≈ 3.1 nm for 1 nm thick Pt base contact layer on HBT64C and ≈ 6 nm for 2 nm thick Pt base contact layer on HBT64J.
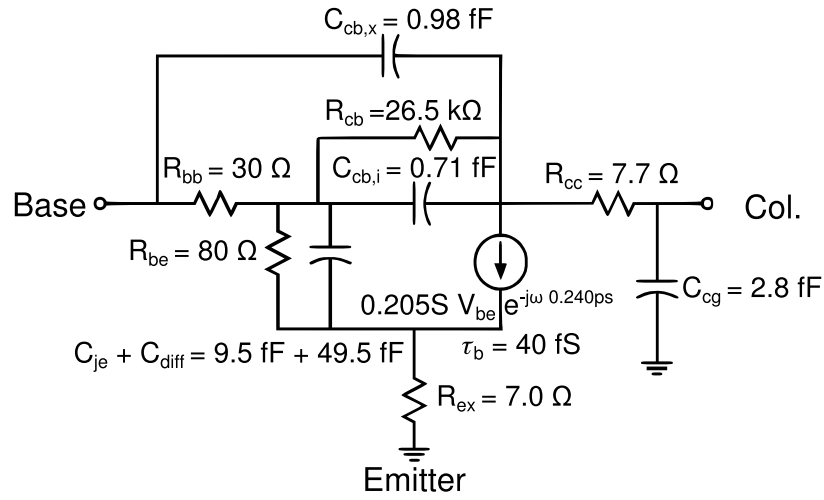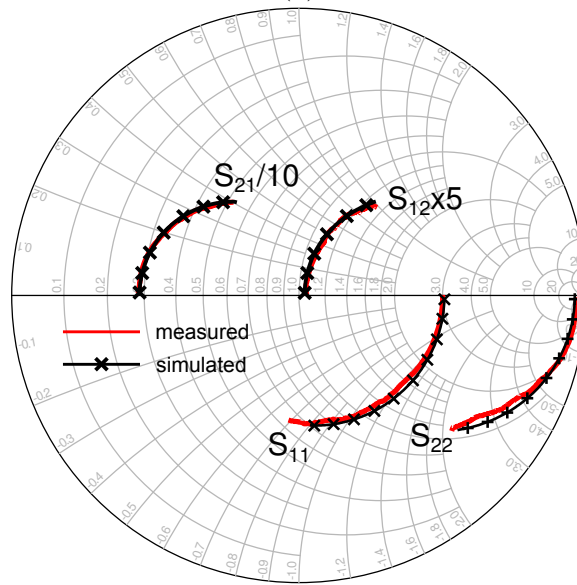
Figure 4.29: (a) A hybrid-π equivalent circuit for the HBT at peak f_{max} performance. (b) Comparison of (solid line) measured S-parameters of Figure 4.26 and (x) simulated S-parameters from 0.5 to 67 GHz .

(a)   (b)
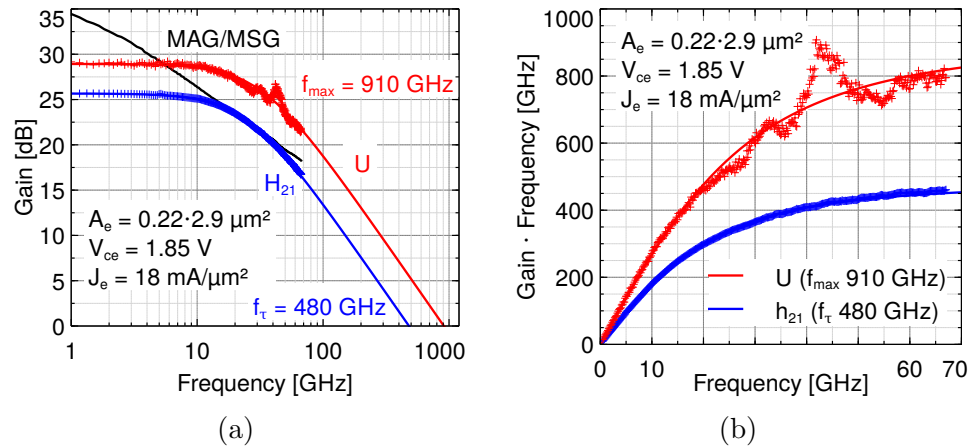
Figure 4.30: Microwave gains for an HBT with 220 nm x 2.9 µm emitter junction area. (a) Double logarithmic plot, (b) linear plot of gain·frequency vs frequency. Single pole fit yields f_τ 480 GHz, f_{max} 910 GHz.
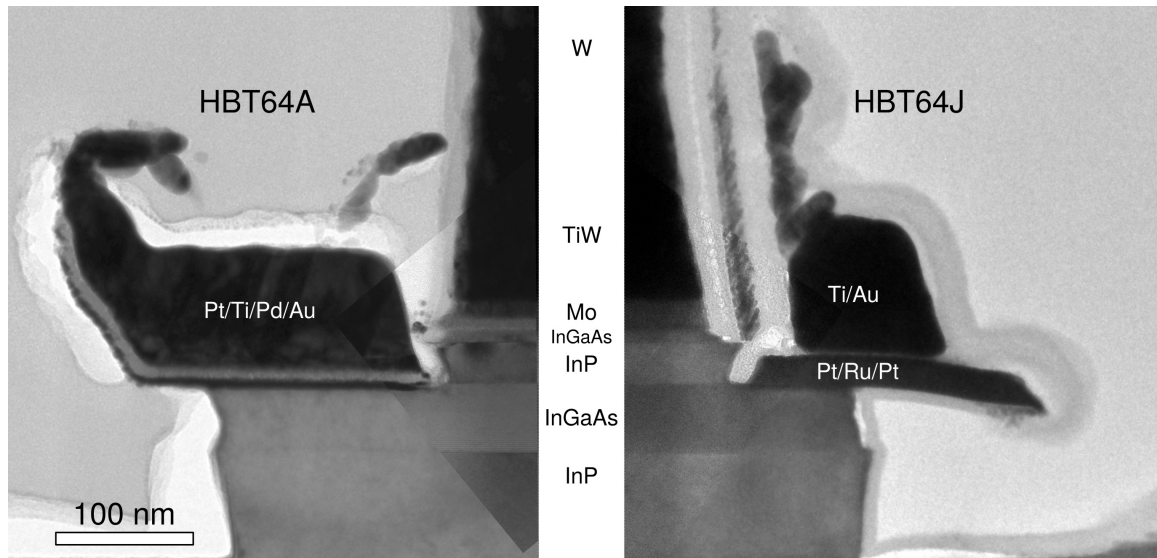
Figure 4.31: Composite TEMs of sample with base fabricated using lifted-off base contact technology (left, HBT64A) and dual-deposited base contact technology (right, HBT64J).

117

(a)   (b)
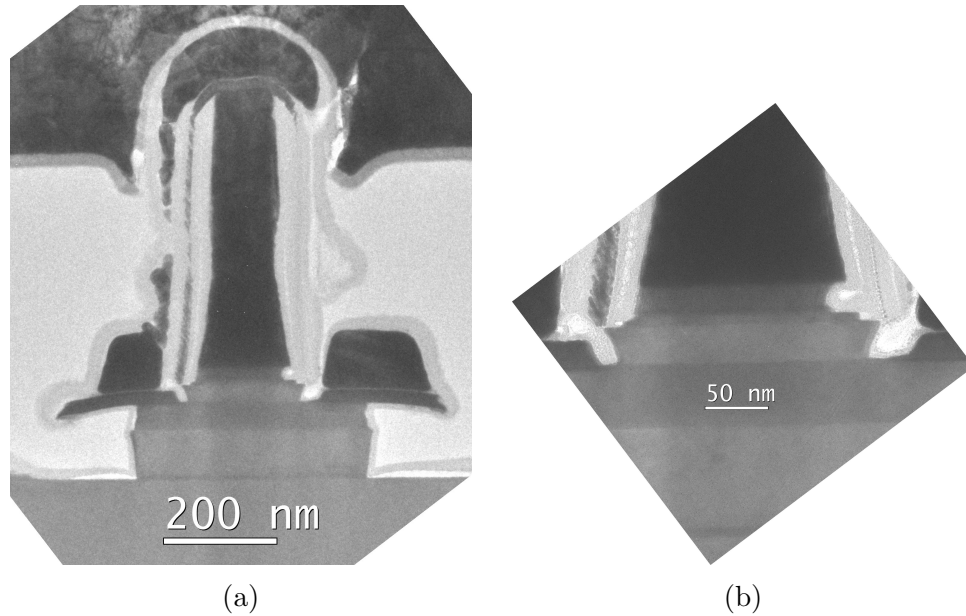
Figure 4.32: Cross-sectional TEM of (a) the complete HBT with $200\,\text{nm}$ x $2.9\,\mu\text{m}$ emitter junction area and (b) the emitter-base region of the same device on sample 64J.

4.4.6   Discussion

Very high $f_{\max}$ bandwidth has been observed on HBT64J that was enabled by improved base ohmics due to the dual-deposition base metal process: RF measurements indicate a total base contact resistivity of $\approx 4\,\Omega\,\mu\text{m}^2$, better than any previous sample with lifted-off base contact metalization has yielded. The extracted base contact resistivity on 64C is $\approx 4\,\Omega\,\mu\text{m}^2$. Incomplete surface coverage of $1\,\text{nm}$ Pt contact layer has increased base contact resistivity on 64C despite shallower interdiffusion that results in higher doping at the metal-semiconductor interface.

The sheet resistivity of the base electrodes has been reduced from $1\,\Omega/\square$ on 56J to $0.4\,\Omega/\square$ on 64C and 64J as a result of dual-deposition base metal process. With

118
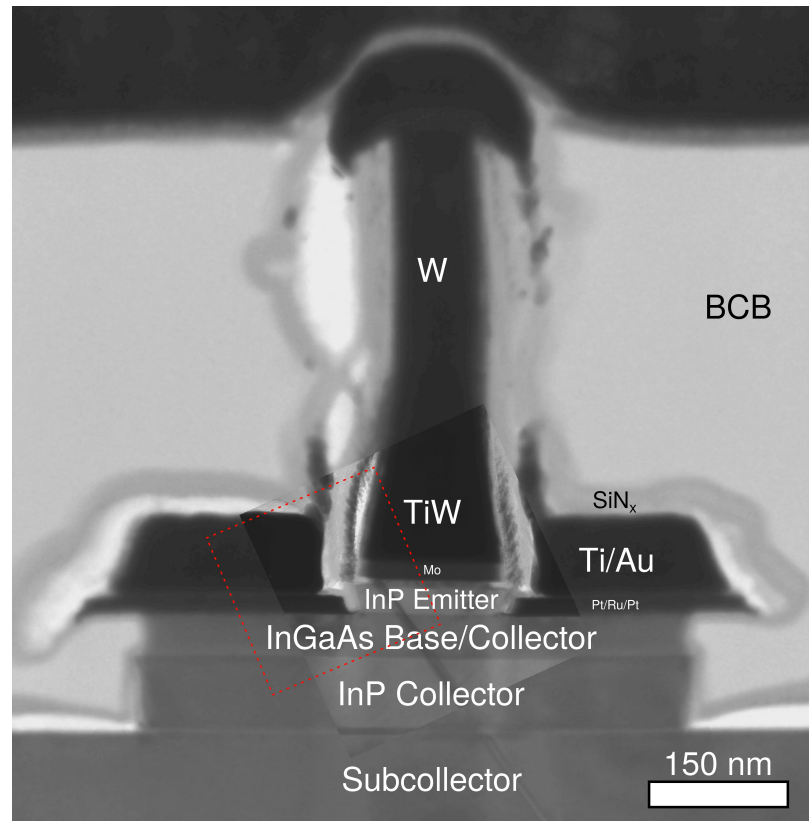
(a)
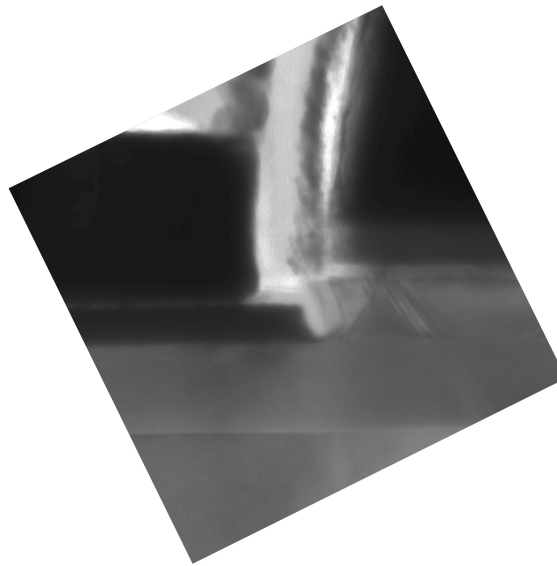
(b)

Figure 4.33: Cross-sectional TEM of (a) the complete HBT with 150 nm x 2.9 µm emitter junction area and (b) higher-magnification cutaway of the emitter-base region marked with red dotted rectangle in (a) on sample 64C.

the heavily undercut base/collector mesa on 64J, the fictitious contact resistivity due to the finite resistance of the base electrodes has been reduced from $\approx 3\,\Omega\,\mu m^2$ on 56J to $\approx 0.7\,\Omega\,\mu m^2$ on 64J and $\approx 1.4\,\Omega\,\mu m^2$ on 64C on highest RF bandwidth HBTs ($L_e = 3\,\mu m$).

Reduction in parasitic capacitances achieved by shrinking the base post, undercutting the base post, reducing the undercut of emitter ends and a thinned base has further improved highest f_τ bandwidth from $440\,GHz$ on 56J to $550\,GHz$ on both 64J and 64C.

Current gain has been maintained on 64J and 64C despite increased base doping and subsequently reduced bulk current gain: estimations for the current gain due to bulk recombinations predict a $40\,\%$ decrease from 56 to 64 ($\beta_{\mathrm{bulk}} = \tau_n/\tau_b \propto t_b^{-0.5} n_a^{-2}$, minority carrier lifetime $\tau_n \propto n_a^{-2}$ due to Auger from n_a base doping concentration, base transit time $\tau_b \propto t_b^{0.5}$ with base thickness t_b). The Al_2O_3 layer deposited as part of the dual-deposition base metalization process has completely covered all sensitive base/emitter surfaces, thereby passivating and protecting it from damage in subsequent processing. Steeper base doping grade has further reduced emitter-base surface currents, resulting in maintained overall current gain.

Wet etching issues during base/collector mesa and device isolation have decreased yield on 64J. We suspect either incomplete removal of etch residues during the sputtering base metal dry etch or contamination of etchants/glassware.

Reduced InP emitter wet etching times and poor dimension control of the emit-

ter metal dry etch have increased junction widths, yielding only devices with $w_e >$

160 nm. Narrower junction width HBTs are expected to exhibit higher bandwidths.

Insufficient calibration structures have limited $f_{\max}$ extraction on small footprint

devices. The mismatch of $\Re(Y_{12})^{-1}$ to the $50\,\Omega$ measurement system has limited

measurement accuracy.

References

[1] J. Rode, H.-W. Chiang, P. Choudhary, V. Jain, B. Thibeault, W. Mitchell, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and A. Liu, "An InGaAs/InP DHBT with simultaneous ft/fmax 404/901GHz and 4.3V breakdown voltage," *Electron Devices Society, IEEE Journal of the*, vol. 3, pp. 54–57, Jan 2015.

[2] M. Urteaga, *Submicron InP-based Heterojunction Bipolar Transistors*. 2003.

[3] F. Purroy and L. Pradell, "New theoretical analysis of the lrrm calibration technique for vector network analyzers," *Instrumentation and Measurement, IEEE Transactions on*, vol. 50, pp. 1307–1314, Oct 2001.

[4] A. M. Safwat and L. Hayden, "Sensitivity analysis of calibration standards for solt and lrrm," in *58th ARFTG Conf. Dig*, vol. 40, pp. 1–10, 2001.

[5] M. Koolen, J. Geelen, and M. Versleijen, "An improved de-embedding technique for on-wafer high-frequency characterization," in *Bipolar Circuits and Technology Meeting, 1991., Proceedings of the 1991*, pp. 188–191, IEEE, 1991.

[6] V. Jain, *InP DHBTs in a Refractory Emitter Process for THz Electronics*. University of California, Santa Barbara, 2011.

[7] H.-W. Chiang, *Design and Fabrication of Sub-100 nm Base-Emitter Junctions of THz InP DHBTs*. 2015.

[8] V. Jain, J. Rode, H.-W. Chiang, A. Baraskar, E. Lobisser, B. J. Thibeault, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and W. Liu, "1.0 THz fmax InP DHBTs in a refractory emitter and self-aligned base process for reduced base access resistance," in *Device Research Conference (DRC), 2011 69th Annual*, pp. 271–272, June 2011.

[9] A. Baraskar, *Development of Ultra-Low Resistance Ohmic Contacts for indium gallium arsenide/indium phosphide HBTs*. 2011.

[10] M. J. Rodwell, M. Le, and B. Brar, "InP bipolar ICs: scaling roadmaps, frequency limits, manufacturable technologies," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 271–286, 2008.

[11] J. Lin, S. Yu, and S. Mohney, "Characterization of low-resistance ohmic contacts to n-and p-type InGaAs," *Journal of Applied Physics*, vol. 114, no. 4, p. 044504, 2013.

[12] E.-F. Chor, D. Zhang, H. Gong, W. K. Chong, and S. Y. Ong, "Electrical characterization, metallurgical investigation, and thermal stability studies of (pd, ti, au)-based ohmic contacts," *Journal of Applied Physics*, vol. 87, pp. 2437–2444, Mar 2000.

[13] D. Ivey, P. Jian, R. Bruce, and G. Knight, "Microstructural analysis of Au/Pt/Ti contacts to p-type InGaAs," *Journal of Materials Science: Materials in Electronics*, vol. 6, no. 4, pp. 219–227, 1995.

Chapter 5

Conclusions

5.1 Accomplishments

On the path to highly-scaled heterobipolar transistors with highest RF bandwidths, key features have been added to the manufacturing process that have enabled significant narrowing of base/collector widths while yielding low contact contact resistivities to both emitter and base. At the same time, device parasitics have been significantly reduced.

Electron beam lithography has been exploited for the formation of well-aligned, narrow base electrodes and base/collector mesas. High resolution (10 nm) and excellent emitter-to-base alignment (sub-30 nm) has been achieved by optimization of pattern writing strategies and proximity effect correction. Radiation damage of EBL resist has been identified as root cause of failure in base lift-off processes. With ra-

diation doses reduced to acceptable levels, two samples have been fabricated yielding substantially increased $f_{\max}$ bandwidth [1, 2].

Extraneously high base contact resistivity has been identified as the main limitation on $f_{\max}$ bandwidth. This has motivated the development of a scaled TLM process that is similar to both HBT fabrication and dimensions, but has very fast turn-around times. The scaled TLM process has enabled a series of experiments that have shed light on multiple issues limiting base contact resistivity: we infer from the results that contaminants to the base semiconductor have been inadvertently introduced in prior processing. Furthermore, it has been found that the Pt contact layer to the base is an insufficient diffusion barrier between the InGaAs base and metals of the base electrodes.

The results of scaled TLM experiments have motivated the development of a novel dual-deposition base metalization process that shows superior contact resistivity and thermal stability in comparison to conventional {Pd,Pt}/Ti/Pd/Au contacts: immediately after removing the InP emitter, base contact metalization is deposited without any lithographic patterning so as to maintain a pristine semiconductor surface for the ohmic interface. A composite metal stack of platinum and ruthenium is exploited: a controllably shallow reaction between an ultrathin layer of platinum and the base semiconductor moves the ohmic interface away from the surface. With the upper metal layer ruthenium as a diffusion barrier, low contact resistivity below $4\,\Omega\,\mu m^2$ is maintained even after high temperature processing.

It has been shown that finite sheet resistance R_{sh} of the base electrodes imposes further limitations on $f_{\max}$ bandwidth: an expression relating the sheet resistance to a fictitious base contact resistivity $\rho_{\text{BaseMetal}} \propto R_{sh}L_e^2$ has been derived and numerically verified with a finite-element circuit modeling the 3D transistor. For the HBT with highest RF bandwidth on sample 56J, approximately $3\,\Omega\,\mu\text{m}^2$ of the total measured base contact resistivity is only due to the finite resistance of the base electrodes. Fabrication of thicker, more conductive base electrodes has been enabled by the dual-deposited base metalization process: the addition of a third sidewall mitigates risk of base-emitter shorts caused by accidental base metal deposits onto thin emitter sidewalls while also decreasing C_{be}. With aggressive undercut of the base/collector mesa, the effect of base electrode resistance on $f_{\max}$ bandwidth has been reduced further: the fictitious contact resistivity on samples 64C and 64J has been decreased to $0.4\,\Omega\,\mu\text{m}^2$.

Further deficiencies of the fabrication process that limit RF bandwidth have been identified and resolved: on highly scaled devices, the ends of the emitter stripe have been severely undercut along fast etch facets. This has reduced the active device area, increasing C_{cb}/I_c and subsequently degrading RF bandwidth particularly on small footprint transistors. Non-scaled base posts have amounted for $> 60\,\%$ of C_{cb} capacitance, greatly reducing transit frequencies. The fabrication process has been modified to mitigate both issues.

Passivation of base/collector junctions with a dip in diluted hydrochloric acid

and subsequent encapsulation with PECVD SiN_x has increased breakdown voltage BV_{CEO} from 3.7 V [1] to 4.3 V [2].

The design and process improvements have been demonstrated on sample 64J on which the base has been thinned to 20 nm while the doping grade was increased to 11–$7 \times 10^{19}/cm^3$. Transistors exhibited f_τ 480 GHz and $f_{\max}$ well above 1 THz at $w_e = 200$ nm emitter width. With simultaneous emitter access resistivity $3\,\Omega\,\mu m^2$ and base contact resistivity $4\,\Omega\,\mu m^2$, key roadmap milestones for the 100 nm thick collector HBT generation have been met [3].

5.2   Future Work

As of 2015, no immediate physical constraints are observable that would prevent the realization of further scaled heterobipolar transistors with RF bandwidths well beyond of what has been demonstrated in this work: scaling laws remain valid. However, the fabrication of such devices will continue to be challenging.

Precise control of key device dimensions is critical for scaling. In this work, the dry-etch process for forming the high aspect ratio $\approx$ 5:1 refractory metal emitter electrodes has shown poor repeatability: despite various efforts to calibrate and stabilize the process, the width of the emitter electrode remains very sensitive to ill-controlled external parameters affecting the dry etch. In addition, the emitter junction has been inadvertently widened further as a result of shortened wet etch times for reducing the undercut of emitter ends. Both effects have resulted in increased emitter width by

$\approx 100\,\mathrm{nm}$. A modified process is therefore required for forming emitter metal stripes and etching emitter semiconductor layers that will restore precise emitter width control. Efforts are underway at UCSB to develop a new emitter metal process: emitter stripes are formed by etching high aspect ratio trenches etched into amorphous Si. These trenches are then filled with conformally grown metal. The sample is planarized and silicon in the field is removed. Unlike the composite TiW/W dry etch used in this work, high aspect ratio dry etching processes for amorphous silicon have a high degree of control over sidewall slopes and subsequently emitter width. Repeatable formation of narrow and deep trenches have been demonstrated.

The electron beam lithography process that has been developed for base formation will remain usable for several scaling generations: with adjustments to the pattern writing conditions at the expense of writing time and lithographic process parameters (resist thickness etc.), base electrodes can be formed with base-to-emitters alignment errors below $10\,\mathrm{nm}$ and simultaneously enhanced resolution.

Decreased base contact resistivity requires simultaneously reduced base electrode sheet resistance: adjustments to the composition and thickness of the third sidewall around the emitter metal can enable lift-off deposition of thicker, more conductive base electrodes that will become necessary with advanced scaling nodes.

Reducing ohmic contact resistivity for emitter and base electrodes remains crucial for successful scaling. Better emitter contact resistivity is obtainable by substituting the InGaAs emitter cap either with a thin layer of lattice-mismatched lower bandgap

InAs [4], or with InGaAs that has higher In content [5].

Base contact resistivity can be improved by increasing base doping [4] at the expense of current gain. With additional current gain degradation due to increased perimeter to area ratio $\propto 1/w_e$ between successive scaling generations, maintaining or improving current gain requires reduced doping concentration of the intrinsic base (underneath the emitter) and very high doping of the extrinsic base (underneath the contacts). Decoupling base doping concentrations is possible by regrowing either highly doped base semiconductor, or the moderately doped emitter semiconductor.

The dual-deposition base metalization process has demonstrated low contact resistivity. The surface morphology of the base semiconductor prior to formation of ohmic contacts is not well understood though: further investigations are therefore required that might give insight into chemical reaction mechanisms of prior processing steps, resulting in more effective surface preparation. Also, the dual deposition process requires a sputtering dry etch that quickly removes InGaAs alongside Pt: substituting the sputtering step with a more chemical dry etch will allow selective removal of Pt without potentially damaging the semiconductor on which subcollector contacts are deposited.

With increased perimeter to area ratio, surface effects will become more dominant on scaled devices, necessitating improved surface passivation techniques to suppress such effects.

Inadequate calibration methods and waveguide structures on samples 64C and

64J have caused artifacts in the measurements of Mason's unilateral gain on small footprint devices, invalidating attempts to reliably extract $f_{\max}$ bandwidth. Back-end processing has been therefore extended to enable fabrication of microstrip waveguide structures by addition of a second metal layer [6], enabling full on-wafer multi-LRRM calibration and de-embedding. Probe spacing is greatly increased with microstrip waveguide structures, further enhancing measurement accuracy.

References

[1] V. Jain, J. Rode, H.-W. Chiang, A. Baraskar, E. Lobisser, B. J. Thibeault, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and W. Liu, "1.0 THz fmax InP DHBTs in a refractory emitter and self-aligned base process for reduced base access resistance," in *Device Research Conference (DRC), 2011 69th Annual*, pp. 271–272, June 2011.

[2] J. Rode, H.-W. Chiang, P. Choudhary, V. Jain, B. Thibeault, W. Mitchell, M. Rodwell, M. Urteaga, D. Loubychev, A. Snyder, Y. Wu, J. Fastenau, and A. Liu, "An InGaAs/InP DHBT with simultaneous ft/fmax 404/901GHz and 4.3V breakdown voltage," *Electron Devices Society, IEEE Journal of the*, vol. 3, pp. 54–57, Jan 2015.

[3] M. J. Rodwell, M. Le, and B. Brar, "InP bipolar ICs: scaling roadmaps, frequency limits, manufacturable technologies," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 271–286, 2008.

[4] A. Baraskar, *Development of Ultra-Low Resistance Ohmic Contacts for indium gallium arsenide/indium phosphide HBTs.* 2011.

[5] Z. M. Griffith, *Ultra high speed InGaAs/InP DHBT devices and circuits.* University of California, Santa Barbara, 2005.

[6] H.-W. Chiang, *Design and Fabrication of Sub-100 nm Base-Emitter Junctions of THz InP DHBTs.* 2015.

Appendix A

Terahertz HBT Process Flow

This appendix describes the state of the DHBT mesa process flow with lifted off base contacts, as of March 2015.

A.1   Sample Preparation

- Cleave 4" wafer into piece parts: emitters must be parallel to the major flat. With a recticle size of $4\,\text{mm} \times 5.5\,\text{mm}$ and a 5x5 array of dice, piece parts should be $\approx 26\,\text{mm} \times 30\,\text{mm}$.

- Make light marks with a scribe on the back of the sample indicating the direction of the major flat. Also mark the sample designation.

A.2   Emitter Formation

If possible, the complete emitter metalization should be performed in immediate succession. Up until emitter contacts have been deposited, you should wear a surgical mask.

A.2.1   Emitter Surface Preparation

- E-beam #1 evaporator: change crystal monitor, align wafer chuck to source, load molybdenum (Mo) source, pump to below $1 \times 10^{-6}\,\text{Torr}$. Degas Mo source, i.e. evaporate $10\,\text{nm}$ while manually sweeping the beam over a large surface area. The source must not spit at evaporation rates $1.5\,\text{Å/s}$: if it does, melt it in further. Wait at least $30\,\text{min}$ prior to venting.

- Perform standard solvent clean on sample: $3\,\text{min}$ acetone, $3\,\text{min}$ isopropyl alcohol (IPA), and $3\,\text{min}$ lightly flowing DI H_2O. Resistance of DI water must be better than $16\,\text{M}\Omega$.

- Place sample in $110\,°C$ oven for $10\,min$ to dehydrate the surface.

- UV O_3 oven: clean reaction chamber and wafer holder with IPA and non-shedding wipe. Let the reactor run empty for $15\,min$. Load sample and run in UV O_3 for $10\,min$.

- At the acid bench, clean DI and HCl beakers and wafer basket with DI. Prepare a beaker with 1:10 HCl:H_2O. Use wafer basket to stir solution.

- Vent e-beam #1. Verify that the source is free of contaminants and has a shiny surface.

- Dip sample in solution for $10\,s$, gently agitating. Immediately rinse for $60\,s$ with H_2O.

- *Immediately* load sample in e-beam #1 evaporator.

A.2.2 Emitter Mo Evaporation

- Pump e-beam #1 to pressure below $5 \times 10^{-7}\,Torr$.

- Evaporate $20\,nm$ Mo onto the sample at $0.5–1\,Å/s$. Do not sweep the electron beam and constantly monitor evaporation rate and the source during the evaporation: if spitting occurs, close the shutter immediately and re-melt the source.

A.2.3 Emitter W/TiW Calibration and Deposition

Prior to depositing the emitter metal on the actual sample, one or more calibration runs may be necessary to find the right pressures for the W and TiW depositions to reduce the stress of the film below $\sim 200\,MPa$ — a value empirically found to be sufficiently low to promote good emitter yield. There is a discrepancy between Si and InP wafers: deposition conditions that result in $\approx +1000\,MPa$ stress on Si wafers are stress-free on InP.

Note: TiW source is 10% Ti by *weight*, not atomic composition.

- Vent load lock chamber of sputter #4 and inspect carrier chuck. Make sure that all screw holes are covered. Load carrier chuck and pump load lock.

- Condition the chamber with $30\,min$ blanket W deposition.

- Measure initial wafer curvature in Tencor Stress Measurement tool, then load dummy 2" Si wafer and deposit composite W/TiW film.

- Vent, and measure the stress after deposition: if it is between $800\,\mathrm{MPa}$ and $1200\,\mathrm{MPa}$, proceed with deposition on a 2" InP piece. Also measure sheet resistivity of the full wafer: a sheet resistivity $\approx 0.55\,\Omega/\square$ on the Si wafer indicates a good film. On 2" InP, the sheet resistivity should be $\approx 0.8\,\Omega/\square$.

- If the stress is greater than $800\,\mathrm{MPa}$, take cross-sectional SEMs of the sample to measure the film thickness, then recalibrate the tool. Previous data that links stress to argon pressure has been recorded.

- Once a stable W/TiW recipe has been established, load the actual samples and deposit the composite metal film on the surface.

A.2.4   Sacrificial SiO$_2$/SiN$_\mathrm{x}$ Deposition

- Vent PlasmaTherm PECVD chamber. Wipe inside walls with IPA-soaked shedding wipe. Pump chamber and run standard $60\,\mathrm{min}$ SiO$_2$ clean.

- Vent chamber and load samples in center of chamber (maximum 2 samples at a time). Pump down.

- Run standard $80\,\mathrm{nm}$ SiO$_2$ deposition. Without venting, run standard $40\,\mathrm{nm}$ SiN$_\mathrm{x}$ deposition.

- Unload samples and let them slowly cool down for a few minutes.

A.2.5   Chrome Hardmask Deposition

- Vent e-beam evaporator #1 and load chrome. Load samples and carefully blow off dust with nitrogen gun.

- Pump down below $3 \times 10^{-6}\,\mathrm{Torr}$.

- Deposit $40\,\mathrm{nm}$ of Cr at 1–$2\,\text{Å}/\mathrm{s}$.

A.2.6   Emitter Lithography

- Clean sample with standard solvent clean and dehydration bake.

- Oxidize chrome surface for better resist adhesion with $30\,\mathrm{s}$ O$_2$ plasma in the PE-II etcher ($300\,\mathrm{mTorr}$ pressure at $100\,\mathrm{W}$ power).

- Check o-ring and surface of spinner chuck. Place sample on spinner chuck and turn vacuum on. Nudge sample with tweezers at edge to verify vacuum is holding it in place. Test-spin the sample without photoresist to make sure that the sample is held tightly. If the sample feels loose or the trial spin gives a

vacuum error, remove the sample. Cleaning the back of the sample and cleaning the spinner chuck may be necessary.

- Cut end off a clean pipette and attach it to the end of a syringe. Withdraw enough ma-N 2403 resist (or ma-N 2405:thinner 1:1) to approximately fill half of the pipette stem, and then pull the syringe back to bring all the photoresist into the syringe. Remove the pipette end from the syringe, and attach a $0.2\,\mu m$ filter.

- Push the photoresist through the filter, covering the sample uniformly in droplets. The first drops should form a puddle at the center of the sample that will be extended to cover the entire sample.

- Spin for $30\,s$ at $4000\,rpm$ (recipe #7). If any large spots or solid particulates in the resist appear near the center of the sample, strip the photoresist in 1165 (NMP) heated to $80\,°C$. First heat the stripper, then submerse the sample for $15\,min$, followed by a solvent clean. Redo the photoresist application.

- Softbake at $90\,°C$ for $90\,s$.

- Load the sample in the #2/2E cassette for the electron beam writer. Turn the cassette upside down on the table, so the handle is on the left side, and the 2E holder is in the upper left corner. Load the sample so that the major flat is aligned horizontally, i.e. parallel to the left-right axis. Nudge the sample until it is centered vertically in the 2E window. You can use the alignment stage to verify that the sample is loaded at exactly $0°$ angle.

- Load cassette into the autoloader, pump for $\approx 20\,min$, and perform necessary calibrations and exposure. Double-check the fractured pattern for executing exposure.

- Develop in beaker of AZ300MIF ($2.38\,\%$ aqueous TMAH by weight) for 35–$40\,s$ with gentle agitation every $10\,s$, then rinse in DI H_2O for $3\,min$.

- Verify lithography came out as expected under optical microscope: the marks used for the local stepper stepper alignment should show no signs of bulging.

A.2.7   Emitter Hardmask Etch

- Take a reasonably shiny looking 6" Si carrier wafer from the "Cl_2/O_2" section of the Rodwell ICP carrier wafer box. Clean it with acetone from a spray bottle, then wipe dry with a non-shedding wipe. Repeat with isopropyl alcohol from a spray bottle, and dry the wafer with the N_2 gun.

- Load the wafer in one of the ICP's cassettes, and run the standard O_2 clean (recipe #121) for 10 min.

- Clean a second wafer from the "Cl_2/O_2" portion of the box identically to the first.

- Using the dropper in the bottle, place a dewdrop sized droplet of Santovac oil in the center of the wafer. Place the sample on top of this, and use two wooden swabs to press on opposite corners of the sample to bring it flush with the sample surface.

- Load the wafer in the other ICP cassette and run a Cl_2/O_2 etch with 26/4 sccm of gas flow, 1 Pa pressure, and 400/15 W of source and bias power for 125 s.

- When the sample returns from the etcher, the surface should look uniform, indicating all the Cr in the field has been etched away to expose the SiO_2 below. Immediately transfer the sample to a beaker of DI H_2O, and take it to the developer benches to rinse for 2 min.

- Transfer the sample to a beaker of 1165, pre-heated in the 80° bath. Leave the sample and beaker of 1165 in the bath for 60 min, then transfer the sample to IPA and then DI H_2O rinse for 3 min each.

- Descum the sample for 30 s in the O_2 only PE-II asher at 300 mTorr and 100 W.

- Inspect the Cr etch mask with top-down SEM – verify the field and features are clear of photoresist scum or Cr particles, the stitching offsets within features are ~10 nm or less, and record the actual widths of the emitters after the etch. If the etch is unsatisfactory in any way, redo the Cl_2/O_2 etch on the entire sample, redeposit a new Cr layer, and redo the lithography and etch.

A.2.8   Emitter Contact Dry Etch

- If necessary, change the gas from CHF_3 to Ar.

- Prepare two 6" carrier wafers for SF_6/Ar etches as described above.

- Load the wafer in one of the ICP's cassettes, and run the standard O_2 clean (recipe #121) for 10 min. If the plasma does not look dim and white during the O_2 clean, the chamber has not been fully cleaned, and a longer O_2 clean is necessary.

- Check the parameters of the high power etch #161: 20/5 sccm SF_6/Ar flow rate, 1 Pa pressure, 600 W source power, 200 W bias power, etch time $\approx$ 3 min:5 s.

- Check the parameters of the low power etch #197: 5/5 sccm SF_6/Ar flow rate, 0.5 Pa pressure, 600 W source power, 15 W bias power, etch time $\approx$ 1 min:20 s.

- Load the wafer in the other cassette, and run the composite high power / low power SF_6/Ar etch. The plasma should look light blue/violet and must not flicker.

- When the first wafer is returned from the system, mount the sample on the carrier wafer.

- Run the composite etch with the sample when the quartz temperature has reached 55 °C or below.

- Immediately rinse the sample, while still attached to the carrier wafer, with H_2O from a spray bottle. Use a wooden swab to push the sample off the carrier wafer.

- Immediately transfer the sample to a beaker of DI H_2O, and rinse for 2 min at the developer bench. Follow with a standard solvent clean, and dry with N_2 at <20 psi.

- Inspect the device in angled SEM. Check device yield for full range of emitter widths, verify emitter profile is sufficiently vertical, and inspect field for etch completion.

If you want to run multiple samples, always run an oxygen clean and chamber conditioning etch before running the next sample. The etch is accelerated when being executed back-to-back.

The given SF_6/Ar recipes above and CF_4/O_2 recipes below are frequently altered by other users. Make sure to verify *all* settings — source and bias powers, gas flow rates, pressure, and duration of etch. Frequently, low-pressure recipes such as #134 will have a higher pressure during the ignition stage. This is normal, but double-check the pressure drops to the proper value during the etch step.

A.2.9 Etch Hardmask Removal

- Perform a standard solvent clean and dehydration bake on the sample.

- Mix buffered hydrofluoric acid with two drops of Tergitol placed with a pipette. Stir the solution with the etch baskets.

- Submerge the HBT sample for 65 s in BHF, mildly agitating the solution every 10 s at the beginning of the etch, and constantly at last 20 s of the etch.

- Rinse the HBT sample with DI H_2O for 3 min.

- Inspect the device in angled SEM. Make sure that the hardmask has not fallen back onto emitters or onto surface close to emitters.

A.2.10 First Sidewall Formation

- Vent PlasmaTherm PECVD chamber. Clean chamber as described above.

- Pump and run 60 min SiN_x clean.

- Vent chamber and place single 2" Si wafer in center of chamber.

- Run standard 100 nm SiN_x deposition on Si wafer.

- Prepare solution of 1:10 $HCl:H_2O$. Just prior to depositing the nitride film, dip the sample in the hydrochloric acid solution and rinse with DI H_2O for 60 s.

- Vent, remove Si wafer, place sample in the center of chamber and run standard 30 nm SiN_x deposition.

- Cleave Si to get a piece with area similar to the HBT sample. Determine SiN_x thickness with ellipsometer at three different points across the surface.

- Prepare two 6" carrier wafers for CF_4/O_2 etches as described above.

- Clean carrier wafers with solvents.

- Load a wafer in one of the ICP's cassettes, and run the standard O_2 clean (recipe #121) for 10 min.

- Check the parameters of the etch #135: 20/5 sccm CF_4/O_2 flow rate, 1 Pa pressure, 500 W source power, 100 W bias power, etch time 5 min.

- Check the parameters of the etch #133: 20/2 sccm CF_4/O_2 flow rate, 0.3 Pa pressure, 25 W source power, 18 W bias power, etch time $\approx$ 3 min:30 s depending on etch rate (see below).

- Load the wafer in the other cassette, and run chamber preconditioning etch #135. The chemistry also etches silicon: if the carrier wafer surface becomes too dull after the etch, replace the wafer.

- Mount the Si piece onto a carrier wafer and run etch #133 for 4 min after the quartz temperature has dropped to 55 °C or below.

- When the dummy sample is returned, rinse with DI H_2O and remove from the carrier wafer with wooden swabs. Return to the ellipsometer and re-measure the film thickness at the same three points. Calculate an average etch rate based on the before- and after-etch thicknesses.

- Mount the HBT sample on a carrier wafer.

- Based on the calculated etch rate, run the low-powered CF_4/O_2 etch for long enough to etch 120% of deposited SiN_x thickness. Since the etch is mostly chemical and operates at fairly low bias power, overetching does little damage to the sample surface.

- After etch completion, spray the sample with DI H_2O and remove it. Solvent clean the sample.

- Inspect the sample in angled SEM. The surface around emitters should be clean all across the sample.

A.2.11 Emitter InGaAs Wet Etch

- At acid bench, prepare beakers with $NH_4OH:H_2O$ 1:10 and 1:1:25 $H_3PO_4:H_2O_2:H_2O$. Stir the dilutions of hydrochloric acid and ammonia manually with the designated wafer baskets in order to mix them up. Place the magnetic stirrer in the beaker and put the beaker on the hotplate, enabling stirring of the InGaAs etchant at 200 rpm for 2 min.

- Submerge the HBT sample in the $NH_4OH:H_2O$ for 10 s, and immediately rinse under DI H_2O for another 60 s. Remove InGaAs etchant from stirring plate. Dry with N_2 at <20 psi.

- Dip sample in InGaAs etchant for 4 s. Rinse under DI H_2O for 3 min. Dry with N_2 at <20 psi.

- Inspect under optical microscope to make sure field looks uniform.

Note: The profilometer has insufficient resolution. It is therefore not possible to accurately measure ≈ 15 nm height difference that has resulted from the wet etch.

A.2.12 Second Sidewall Formation

Deposit 30 nm sidewall as described in section A.2.10.

A.3 Base Formation

The base contact formation should be completed within a single day, starting from the InP wet etch up until the deposition of the Al_2O_3 sidewall. Up until base contacts have been deposited, you should wear a surgical mask.

A.3.1 Emitter InP Etch / Base Contact Formation

- Prepare e-beam evaporator #1: verify alignment between source and sample holder. Change crystal monitor. Place ruthenium (Ru) and platinum (Pt) sources into hearth slots. Program thickness monitor for Ru: density 12.362, z-ratio 0.182, tooling 142. Pump to below 1×10^{-6} Torr.

- Degas Ru source: evaporate ≈ 5 nm while manually sweeping electron beam over source. Stable, spit-free evaporation should be possible at 1 Å/s while maintaining a pressure 1×10^{-6} Torr or below. If not, form a bigger melt of source material. Take a note of the current required to initiate evaporation. Wait at least 30 min before venting the tool to allow source to cool down.

- Take top-down SEMs: measure the width of all emitter sizes and determine the discrepancy between the emitter width as etched and as designed. This width will be used later to adjust the width of base electrodes.

- Solvent clean and dehydrate the HBT sample.

- Clean the reaction chamber and sample stage of the UV O_3 reactor with IPA and a non-shedding wipe. Run empty for 30 min.

- Prepare beaker with 1:10 HCl:H_2O, stirring the solution with the wafer basket.

- Run UV O_3 reactor with sample for 10 min. Remove sample, dip in HCl:H_2O for 10 s, rinse in DI H_2O for 60 second, blow dry with N_2 gun. Make sure that the sample backside is dry.

- Place sample in UV O_3 reactor: rotate sample by 180° in comparison to previous oxidation run. Re-run for 10 min.

- At the acid bench, prepare a beaker of 4:1 H3PO$_4$:HCl. Add magnetic stirrer to the 4:1 H3PO$_4$:HCl beaker, and place on a *room temperature hot plate* set to stir at 200 rpm.

- Vent e-beam #1. Prepare sample holder.

- After completing the second oxidation process, take sample and de-oxidize it in diluted HCl as described above. *Optional:* repeat oxidation and oxide removal cycle.

- Shut off bench lights. Submerge the sample for 4–5 s into the InP etchant. Rinse for 3 min in DI H_2O.

- *Optional:* perform solvent clean on the sample.

- De-oxidize the sample with diluted 1:10 HCl:H_2O 10 s, rinse with DI H_2O for 60 s. Blow-dry the sample thoroughly with N_2 gun.

- Immediately mount sample onto holder for e-beam #1, place holder in evaporation chamber and pump down to at least 5×10^{-7} Torr.

- Evaporate 2 nm Pt at 0.1 Å/s: increase beam current slowly until pressure rises slightly and the evaporation starts.

- Evaporate 15 nm Ru at 0.5 Å/s, using your previous note about the beam current required for evaporation as a starting point.

- Evaporate 2 nm Pt at 0.5 Å/s: the beam current required for this layer is significantly less than for the previous Pt layer.

- Vent the tool, retrieve sample.

The base contact resisitivity is very sensitive to external contaminants. It is essential that the processing time is minimized between removal of the InP emitter semiconductor and deposition of the base contacts in order to reduce contamination of the exposed base surface.

A.3.2 Base Al_2O_3 Passivation

- Initiate cool-down of Oxford atomic layer deposition (ALD) tool to 200 °C.

- Measure native oxide thickness of 2" Si wafer with ellipsometer at multiple points.

- Grow 120 cycles of water Al_2O_3 on Si wafer. After completion, remeasure sample with ellipsometer to determine growth rate ($\approx$ 1.2 Å/cycle). Keep the Si wafer.

- Grow 10 nm Al_2O_3 on HBT sample.

- After completion, heat the ALD to 300 °C.

Note: Fabrication can be stopped at this point and resumed later.

A.3.3 Base Sidewall Formation

Fabricate a 30 nm sidewall as discussed in section A.2.10.

A.3.4   Base Electrode Lithography

- Adjust mask layout for emitter width variation (section A.3.1).

- Solvent clean the sample.

- Prepare a dilution of 1:50 BHF:DI. Etch silicon wafer with Al_2O_3 (section A.3.2) for 30 s. Remeasure thickness, determine etch rate.

- Etch HBT sample, calculating the etch time assuming 120 % deposited film thickness.

- Solvent clean the sample, dehydrate.

- Spin on PMGI SF8:thinner 3:1 at 3000 rpm, 300 rpm/s, 60 s. Apply the resist using syringes with 0.2 μm filters.

- Prebake resist on hotplate 180 °C 3 min.

- Spin on CSAR:anisole 2:1 at 3000 rpm, 300 rpm/s, 60 s. Apply the resist using syringes with 0.2 μm filters.

- Prebake resist on hotplate 180 °C 3 min.

- Load sample to the #2/2E cassette for the e-beam writer as precisely in the same place as it was for the emitter write as possible. Move the cassette, still facedown, to the glass jig and perform the necessary manual rotational corrections. Measure distances from sample edge to alignment mark on R3C1.

- Load cassette in the e-beam writer, and perform necessary calibrations and alignments. Make sure that the base width adjustments are correct. Verify global (SETWFR) and local (CHIPAL) alignment. Expose the sample.

- After exposure, vent the autoloader and remove sample from cassette.

- Develop CSAR with 50 s amyl acetate, then rinse with IPA 60 s.

- Check features under microscope with green filter in light path.

- Develop PMGI with 25 s AZ300MIF.

- Check undercut of features microscope with green filter in light path.

A.3.5 Base Electrode Liftoff

- Prepare aluminum foil masks: cut $230\,\text{mm} \times 310\,\text{mm}$ window into pieces of aluminum foil to cover sample edges during evaporation. The foil will mask the sample during evaporation to facilitate the liftoff.

- Prepare e-beam #4: change crystal monitor, remove nickel source, replace public titanium (Ti) source with Rodwell titanium source.

- Remove photoresist residue from sample with $10\,\text{s}$ O_2 plasma in PE-II: $100\,\text{W}$ power, $300\,\text{mTorr}$ pressure.

- Fix sample to holder, taping down the prepared Al foil mask with Kapton tape. Pump to $2 \times 10^{-6}\,\text{Torr}$ or below.

- Place beaker with 1165 covered with Al foil into $80\,°\text{C}$ bath.

- Evaporate $50\,\text{Å}$ titanium at $0.5\,\text{Å/s}$, then a $950\,\text{Å}$ or thicker layer of gold (Au) at $2\,\text{Å/s}$.

- Vent the tool, retrieve the sample, submerge in heated 1165 beaker for $1\,\text{h}$.

- Perform lift-off, clean sample with IPA, DI.

- Inspect with top-down SEM to verify base contact dimension and alignment, and emitter yield. Inspect with angled SEM to verify clean lifted off base metal.

A.3.6 Base Post Lithography

- Measure the heights of the emitter and base contact DEKTAK pads with the DEKTAK Profilometer on several different dice across the sample.

- Perform the standard solvent clean and dehydration bake on the sample.

- Spin on LOL1000 at $4000\,\text{rpm}$ $30\,\text{s}$ or, alternatively, PMGI:thinner 1:2 at $3000\,\text{rpm}$ $30\,\text{s}$.

- Place on hotplate at $180\,°\text{C}$ for $3\,\text{min}$.

- Program spinner recipe #0 for $1800\,\text{rpm}$ $40\,\text{s}$ $350\,\text{rpm/s}$.

- Remove sample and let cool for $1\,\text{min}$. Return sample to spinner chuck and verify the vacuum is holding it in place.

- Use pipette to cover sample uniformly in nLOF 5510. Spin with recipe #0.

- Soft-bake $90\,°\text{C}$ on hotplate for $60\,\text{s}$.

- Load sample in stepper using 76.2 mm, 635 μm chuck.

- Expose using Base Post mask in the GCA AutoStepper for 0.24 s or less. Dice R3C2 and R3C4 are used for alignment. You can reduce the spacing between base posts and emitter by giving a negative pass shift in X, e.g. $X = -0.000\,075$ mm.

- Post-bake for 60 s on 110 °C hotplate.

- Develop for 1 min:45 s in beaker of AZ300 MIF, with no agitation. Rinse in DI H_2O for 3 min.

- Carefully inspect sample under optical microscope. The offset between base post verniers and emitter verniers should be <100 nm. Adjust focus slightly to confirm LOL has undercut underneath the opening in the nLOF, which should look like a lighter, fuzzy ring around the well defined opening in the top of the resist. If resist is unsatisfactory or misaligned by more than 100 nm, write down the offset, and strip the photoresist for 30 min in 1165 pre-heated at 80 °C, followed by 3 min IPA, DI rinse. Then redo lithography after incorporating a pass shift into the exposure job file.

A.3.7 Base Post Liftoff

- Prepare e-beam #4 and Al foil mask as described in section A.3.5.

- Remove photoresist residue from sample with 20 s O_2 plasma in PE-II: 100 W power, 300 mTorr pressure.

- Pump to 3×10^{-6} Torr or below.

- Place beaker with 1165 covered with Al foil into 80 °C bath.

- Evaporate 15 nm Ti at 1 Å/s.

- Evaporate sufficiently thick layer of Au (≈ 600 nm) so that the base post is ≈ 100 nm above the top of the emitter stripe.

- Vent the tool, retrieve the sample, submerge in heated 1165 beaker for 1 h.

- Perform lift-off, clean sample with IPA, DI.

- Inspect sample with SEM. Verify post height, alignment and yield.

A.3.8 Base Mesa Lithography

- Perform the standard solvent clean and dehydration bake on the sample.

- Apply man2410 at 3000 rpm, 450 rpm/s, 60 s. Apply the resist using syringes with 0.2 µm filters.

- Softbake at 90 °C for 2 min:30 s.

- Load sample to the #2/2E cassette for the e-beam writer as precisely in the same place as it was for the emitter write as possible. Move the cassette, still facedown, to the glass jig and perform the necessary manual rotational corrections. Measure distances from sample edge to alignment mark on R2C1.

- Load cassette in the e-beam writer, and perform necessary calibrations and alignments. Make sure that the base width adjustments are correct. Verify global (SETWFR) and local (CHIPAL) alignment. Expose the sample.

- Develop in beaker of AZ300MIF for 2 min:15 s with gentle agitation every 30 s, then rinse in DI H_2O for 3 min.

- Inspect sample under optical microscope for alignment, focus checkers.

A.3.9 Base Mesa Etch

- Obtain height of emitter and base electrode pads with profilometer.

- Prepare two 6" carrier wafers for Cl_2/O_2 etches as described above.

- Load the wafer in one of the ICP's cassettes, and run the standard O_2 clean (recipe #121) for 10 min. If the plasma does not look dim and white during the O_2 clean, the chamber has not been fully cleaned, and a longer O_2 clean is necessary.

- Check the parameters of the Ru etch #166: 20/5 sccm Cl_2/O_2 flow rate, 0.67 Pa pressure, 400 W source power, 100 W bias power, etch time $\approx$ 40 s.

- Check the parameters of the Pt etch #136: 5/45 sccm Cl_2/Ar flow rate, 1.3 Pa pressure, 600 W source power, 150 W bias power, etch time 35 s.

- Load carrier wafer into cassette, pre-condition chamber by running combined etches #166 and #136.

- Mount sample onto carrier wafer, run combined etch when the quartz temperature is below 55 °C.

- After etch, spray water over sample, remove sample from carrier, rinse sample for 3 min with DI H_2O. Blow-dry sample, then soak part of a few shedding wipes with acetone and clean backside of sample from santovac oil by carefully moving sample over soaked wipes using wooden sticks. Re-rinse samples again and confirm that the backside is reasonably clean.

- Remeasure height of emitter and base electrode pads with profilometer: the dry etches remove parts of the chrome hard mask on the emitter and gold from the base electrode as well as InGaAs base semiconductor. For the conditions described above, $\approx 25\,nm$ of InGaAs are removed: wet etch times have to be adjusted accordingly.

- At the acid bench, prepare beakers of 1:10 $HCl{:}H_2O$, 1:1:25 $H_3PO_4{:}H_2O_2{:}H_2O$, 4:1 $H_3PO_4{:}HCl$ and stir them.

- Prepare sample surface with a 10 s dip in diluted HCl, then 60 s rinse in DI H_2O.

- Etch sample for $\approx 15\,s$ (etch rate $\approx 2\,nm/s$) in InGaAs etchant. Overetching is advantageous to reduce base metal resistance by undercutting base/collector mesa, but has to be well-controlled to avoid excessive undercut. Rinse in DI H_2O for 3 min. The surface should look uniform.

- Remeasure height of emitter and base electrode pads with profilometer.

- Etch drift collector with InP etchant: 100 nm collector designs require $\approx 33\,s$. Bubbles should uniformly form on the surface during the etch.

- Remeasure height of emitter and base electrode pads with profilometer, confirming etch depth.

- Remove the resist by submerging the sample in 1165 pre-heated to 80 °C for 1 h.

- Clean sample with IPA, DI H_2O.

- Inspect with angled SEM to observe mesa undercut and etch completion, as well as to verify photoresist has been removed. The field should be clean.

A.4 Collector Formation

A.4.1 Collector Electrode Lithography

- Perform the standard solvent clean and dehydration bake on the sample.

- Spin nLOF 2020 for 30 s at 3500 rpm (recipe #6).

- Soft-bake on 110 °C hotplate for 60 s.

- Expose collector electrode mask with 0.16 s dose as described in section A.3.6.

- Develop for 120 s in beaker of AZ300MIF, with gentle agitation every 30 s. Rinse in DI H$_2$O for 2 min.

- Inspect under optical microscope. Offset between collector electrode verniers and emitter verniers should be <300 nm. If there is larger misalignment, re-work the resist as described above.

A.4.2 Collector Electrode Liftoff

- Prepare e-beam #4 and Al foil mask as described in section A.3.5.

- Prepare subcollector surface with 10 min UV O$_3$ oxidation and oxide removal in 1:10 HCl:DI as discussed above.

- Immediately load sample, then pump to 3×10^{-6} Torr or below.

- Place beaker with 1165 covered with Al foil into 80 °C bath.

- Evaporate 20 nm Ti at 1 Å/s.

- Evaporate 20 nm Pd at 1 Å/s.

- Evaporate 250 nm Au at 4 Å/s.

- Vent the tool, retrieve the sample, submerge in heated 1165 beaker for 1 h.

- Perform lift-off, clean sample with IPA, DI.

- Inspect sample with SEM. Verify clean lift-off.

A.4.3 Device Isolation Lithography

- Perform the standard solvent clean and dehydration bake on the sample.

- Spin SPR955-0.9 at 4000 rpm for 30 s (recipe #7).

- Soft-bake the sample for 90 s on the 90 °C hotplate.

- Expose device isolation mask with 0.27 s dose as described in section A.3.6.

- Post-bake the sample for 60 s on the 110 °C hotplate.

- Develop for 60 s in beaker of AZ300MIF, with gentle agitation every 30 s. Rinse in DI H$_2$O for 3 min.

- Inspect under optical microscope. Offset between device isolation verniers and emitter verniers should be <150 nm. If there is larger misalignment, re-work the resist as described above.

A.4.4 Device Isolation Etch

- Oxidize the sample surface and remove contaminants with 10 s O_2 plasma in PE-II: 100 W power, 300 mTorr pressure.

- Determine height of emitter pad, resist pad using profilometer.

- At the acid bench, prepare beakers with 1:10 NH_4OH:DI, 1:1:25 H_3PO_4:H_2O_2:H_2O, 4:1 H_3PO_4:HCl. Stir diluted HCl with wafer basket, InGaAs and InP etchants with magnetic stirrer and stirring plates.

- Deoxidize the surface by submerging sample in 1:10 NH_4OH:DI for 10 s, followed by a 60 s DI H_2O rinse.

- Remove the InGaAs subcollector with 20 s etch in H_3PO_4:H_2O_2:H_2O solution. Rinse with DI H_2O for 3 min. The surface should look uniform after drying the sample with N_2 gun.

- Remove the InP subcollector with 43 s etch in H_3PO_4:HCl solution. Rinse with DI H_2O for 3 min. Phosphine bubbles should uniformly appear on the sample surface during the etch. The sample surface should look uniform after drying the sample with N_2 gun.

- Determine height of emitter pad, resist pad using profilometer, confirming the etch depth (subcollector thickness).

- Remove the InGaAs etch stop layer with 10 s etch in H_3PO_4:H_2O_2:H_2O solution. Rinse with DI H_2O for 3 min. The surface should look uniform after drying the sample with N_2 gun

- Etch into the InP substrate with 15 s submersion in H_3PO_4:HCl solution. Rinse with DI H_2O for 3 min.

- Determine height of emitter pad, resist pad using profilometer, confirming the etch depth (≈ 200 nm).

- Strip the resist by placing it into pre-heated 1165 for 60 min.

- Clean sample with IPA, DI H_2O.

- Inspect with angled SEM to observe mesa undercut and etch completion, as well as to verify all photoresist has been removed.

A.4.5   Collector Post Formation

Collector post lithography is identical to the lithography for collector electrode formation (section A.4.1) with the exception of exposing a different mask. The deposition of the post is also identical to the deposition of the base post: the total thickness of the collector post should be such that it is 150 nm above the emitter metal, e.g. 20 nm Ti/630 nm Au for a 100 nm collector design.

A.5   Back-End Fabrication

A.5.1   SiN$_x$ Passivation and BCB Planarization

- Vent PlasmaTherm PECVD chamber. Clean chamber as described above.

- Pump and run 60 min SiN$_x$ clean.

- Take top-down SEMs to determine TLM gaps, surface quality around devices.

- Take plenty of angled SEMs of devices, TLMs, test structures.

- Deposit standard 100 nm SiN$_x$ without wafer loaded to condition the chamber.

- Perform the standard solvent clean and dehydration bake on the sample.

- Oxidize the sample surface with 10 min UV O$_3$.

- Clean sample holder used for the BCB annealing oven with IPA.

- Verify that recipe #5 of the annealing oven does the following four-step recipe: 5 min ramp to 50 °C, 5 min soak. 15 min ramp to 100 °C, 15 min soak. 15 min ramp to 150 °C, 15 min soak. 60 min rise to 250 °C, 60 min soak. Non-forced cooldown to ambient temperature.

- Set N$_2$ flow to 100 on the annealing oven.

- Remove oxide by dipping sample in 1:10 HCl:DI H$_2$O, followed by 60 s DI H$_2$O rinse.

- Immediately load sample into PECVD and deposit 30 nm SiN$_x$ film.

- Retrieve sample from PECVD after deposition, confirm adhesion of SiN$_x$ film with microscope.

- Prepare two syringes of BCB 3022-46 with 0.2 μm filters.

- Program recipe 30 s spin at 1500 rpm, 150 rpm/sec into position #0.

- Place sample onto spinner chuck, test vacuum, apply BCB: a lot of force is required due to the high viscosity of BCB. If the filter breaks, continue with the other syringe. Wait 30 s after the entire surface has been covered with BCB, then spin.

- Take sample to blue oven and start recipe #5. Reduce N_2 flow to 60.

- Allow 6-8 hrs for cycle to complete and return to room temperature. *Do not* remove sample until oven temperature is below 100°.

A.5.2 BCB Ashing

- Use Nanometrics Reflectometer with recipe #10 and a dielectric constant of 1.6 to verify the photoresist thickness. It should be $\approx 44\,000\,\text{Å}$.

- Verify the temperature of the Panasonic ICP ashing chamber is 50 °C.

- Prepare two carrier wafers from the CF_4/O_2 section of the Rodwell box.

- Load the wafer in one of the ICP's cassettes, and run a 10 min CF_4/O_2 ash with 50/200 sccm gas flow, 40 Pa pressure, and 1000 W power (recipe #308).

- Mount sample on wafer, load the wafer in an ICP cassette and run a 4 min CF_4/O_2 ash with the same parameter as the chamber condition.

- When the wafer is returned, immediately rinse the sample, while still attached to the carrier wafer, with H_2O from a spray bottle. Gently remove the sample from the wafer with a wooden swab, and transfer the sample to a beaker of DI H_2O, and rinse for 2 m at the developer bench. Follow with a standard solvent clean, and dry with N_2 at <20 psi.

- Measure the thickness of the BCB using the Nanometrics. Calculate an etch rate from the measurement before etching, although this should only be used as a rough estimate. Total BCB height will be ≈ 200 nm below the emitter pad step height obtained from the last profilometer measurement during device isolation.

- With the ashing rate, repeat the last three steps, recalculating the ashing rate after every iteration. The cumulative ash time should be between 6 min:30 s and 8 min.

- When the BCB height is within 400 nm of the expected final thickness, inspect the device in the SEM using the 45° holder, at 2 kV accelerating voltage to minimize charging. Check if emitters and posts are through the BCB — they will appear bright and come sharply into focus, while the field is darker and blurry. Check both edge and center dice.

Note: It takes at least 4–5 ashing cycles to obtain the desired BCB thickness. The last iterations have to have rather short than long ashing time in order not to overetch.

A.5.3   Contact Via Deposition

- Perform the standard solvent clean and dehydration bake on the sample.

- Vent load lock chamber of sputter #4 and inspect carrier chuck. Make sure that all screw holes are covered. Load carrier chuck and pump load lock.

- Condition the chamber with 20 min blanket SiO_2 deposition at 20 °C.

- Run 1800 s SiN_x deposition at 20 °C on 2" Si wafer

- Perform the standard solvent clean and dehydration bake on the sample.

- Unload Si wafer, load sample along with another Si witness wafer to sputter #4.

- Measure SiN_x film thickness with ellipsometer. Adjust deposition time of sputter #4 recipe to deposit 60 nm thick film.

- Deposit film on HBT sample.

- Unload sample after deposition.

A.5.4   Contact Via Lithography

Contact via lithography is identical to the lithography for device isolation (section A.4.3) with the exception of exposing a different mask.

A.5.5   Contact Via Etch

- Cleave Si witness wafer to get a piece with area similar to the HBT sample. Determine SiN_x thickness with ellipsometer at three different points across the surface.

- Prepare two 6" carrier wafers for CF_4/O_2 etches as described above.

- Clean carrier wafers with solvents.

- Load a wafer in one of the ICP's cassettes, and run the standard O_2 clean (recipe #121) for 10 min.

- Check the parameters of the etch #135: 20/5 sccm CF_4/O_2 flow rate, 1 Pa pressure, 500 W source power, 100 W bias power, etch time 5 min.

- Check the parameters of the etch #133: $20/2$ sccm CF_4/O_2 flow rate, 0.3 Pa pressure, 25 W source power, 18 W bias power, etch time ≈ 3 min:30 s depending on etch rate (see below).

- Load the wafer in the other cassette, and run chamber preconditioning etch #135. The chemistry also etches silicon: if the carrier wafer surface becomes too dull after the etch, replace the wafer.

- Mount the Si piece onto a carrier wafer and run etch #133 for 4 min after the quartz temperature has dropped to 55 °C or below.

- When the dummy sample is returned, rinse with DI H_2O and remove from the carrier wafer with wooden swabs. Return to the ellipsometer and re-measure the film thickness at the same three points. Calculate an average etch rate based on the before- and after-etch thicknesses. The etch rate of sputtered SiN_x films is less than that of PECVD SiN_x films (≈ 8 nm/min).

- Mount the HBT sample on a carrier wafer.

- Based on the calculated etch rate, run the low-powered CF_4/O_2 etch for long enough to etch 120 % of deposited SiN_x thickness. Since the etch is mostly chemical and operates at fairly low bias power, overetching does little damage to the sample surface.

- After etch completion, spray the sample with DI H_2O and remove it. The surface will be hydrophobic. Place the sample in acetone for 3 min, then IPA for 3 min, then in pre-heated 1165 for 60 min.

- Clean sample with IPA, DI H_2O.

- Remove resist residues with 30 s O_2 plasma in the PE-II etcher (300 mTorr pressure at 100 W power.

- Inspect the sample in 45° angled SEM. The surface around posts should be clean all across the sample.

A.5.6   Post Cleanup Sputter

- Perform the standard solvent clean and dehydration bake on the sample.

- If necessary, change gases from CHF_3 to Ar.

- Prepare two 6" carrier wafers for SF_6/Ar etches as described above.

- Clean carrier wafers with solvents.

- Load a wafer in one of the ICP's cassettes, and run the standard O_2 clean (recipe #121) for 10 min.

- Check the parameters of the etch #161: 20 sccm Ar flow rate, 1 Pa pressure, 300 W source power, 50 W bias power, etch time 20 s.

- Run etch #161 for 5 min with empty carrier wafer to condition the chamber.

- Mount sample onto carrier wafer.

- Run etch #161 for 20 s to remove resist residues from the base dry etch.

- After etch completion, spray the sample with DI H_2O and remove it.

- Perform the standard solvent clean on the sample.

- Inspect the sample in 45° angled SEM at 2 kV. Base and emitter terminals should be free of contaminants.

A.5.7 Metal 1 Formation

Metal 1 lithography is identical to the lithography for collector electrode formation (section A.4.1) with the exception of exposing a different mask. A film of 30 nm Ti, 1000 nm Au and 10 nm Ti is deposited. Should difficulties arise during the lift-off process, the photoresist stripper AZ300T can be used instead of 1165 to more aggressively dissolve the photoresist. After lift-off, the sample should be inspected with top-down SEMs: the region between base post and emitter must be metal free.

A.5.8 Metal 1 Cleanup Sputter

In the past, photoresist did not adhere well in the region between base post and emitter: during evaporation, emitter-base shorts have been formed by gold creeping underneath the resist. The shorts can be eliminated by sputtering off a thin gold film in a process identical to the one described in section A.5.6.