UNIVERSITY OF CALIFORNIA

Santa Barbara

# Submicron InP-based Heterojunction Bipolar Transistors

A Dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Electrical and Computer Engineering

by

Miguel Urteaga

Committee in charge:

Professor Mark Rodwell, Chair

Professor Evelyn Hu

Professor Stephen Long

Professor Umesh Mishra

December 2003

The dissertation of Miguel Urteaga is approved.

_____

Evelyn Hu

_____

Stephen Long

_____

Umesh Mishra

_____

Mark Rodwell, Committee Chair

November 2003

Submicron InP-based Heterojunction Bipolar Transistors

ACKNOWLEDGEMENTS

I was fortunate to arrive at UCSB and join a research group that had a tremendous base of knowledge and resources. The development of the InP HBT technology has depended on the efforts and dedication of five generations of Ph.D. students. The work described in this thesis benefited from the contributions of past group members and collaborations with my current lab colleagues. I'd like to acknowledge and thank past and current labmates Dennis Scott, Yun Wei, Zach Griffith, Mattias Dahlström, Yingda Dong, Vamsi Paidi, Navin Parthasarathy, Youngmin Kim, PK Sundararajan, Thomas Mathew, Shri Jaganathan, Sangmin Lee and James Guthrie.

The success of the research group has been in large measure due to the leadership and guidance of Professor Mark Rodwell. His enthusiasm for research has been contagious, and his technical contributions to the work in this thesis cannot be understated. Additionally, he has done a tremendous job in acquiring and maintaining the funding necessary for a university group to perform meaningful research in an expensive field.

The access to high quality device epitaxy was essential to the results obtained in this work. I'd like to thank Dennis Scott, Yingda Dong, Youngmin Kim and Amy Liu and her colleagues at IQE for providing epitaxial material for experiments and device processing. I'd also like to thank the staff at UCSB who provided support for the experimental work performed there. Jack Whaley, Brian Thibeault, Bob Hill, Neil Baker and Don Freeborn were always available for help with cleanroom equipment and processes. Bill Mitchell and Ernie Caine helped with the electron-

beam lithography work used in the transferred-substrate HBT process. Mark Cornish and Jan Lofvander provided support for the Materials department SEM and FIB tools that were extremely valuable for diagnosing process problems.

The scaled-mesa HBTs described in this work were fabricated both at UCSB and at Rockwell Scientific Company. I'd like to thank Bobby Brar, Richard Pierson, Petra Rowell, Scott Newell, and the cleanroom staff at RSC for their support of these efforts.

I'd also like to thank my Ph.D. committee members Professor Evelyn Hu, Professor Stephen Long and Professor Umesh Mishra. I appreciate the comments and suggestions I have received from them regarding this work.

Finally, I would like to thank my friends and family members for their support, particularly my wife Ruby for having the patience to stick with me during my four year California adventure.

VITA OF MIGUEL URTEAGA
November 2003

EDUCATION

Bachelor of Applied Science in Engineering Physics, Simon Fraser University, Burnaby, British Columbia, Canada, May 1999

Master of Science in Electrical and Computer Engineering, University of California, Santa Barbara, June 2001

Doctor of Philosophy in Electrical and Computer Engineering, University of California, Santa Barbara, January 2004(expected)

PUBLICATIONS

M. Urteaga, M.J.W. Rodwell, "Power gain singularities in transferred-substrate InAlAs/InGaAs HBT's," *IEEE Transactions on Electron Devices*, vol. 50, no. 7, July 2003.

M. Urteaga, D. Scott, S. Krishnan, Y. Wei, M. Dahlström, Z. Griffith, N. Parthasarathy and M.J.W. Rodwell, "G-band (140-220 GHz) InP-based HBT amplifiers," *IEEE Journal of Solid State Circuits*, vol. 38, no. 9, September 2003.

M.Urteaga, S.Krishnan, D.Scott, Y.Wei, M.Dahlström, S.Lee, M.J.W.Rodwell, "Submicron InP -based HBTs for ultra-high frequency amplifiers," *International Journal of High Speed Electronics and Systems*

M. Urteaga, D. Scott, S. Krishnan, Y. Wei, M. Dahlström, Z. Griffith, N. Parthasarathy, M.J.W. Rodwell, "Multi-stage G-band (140-220 GHz) InP HBT Amplifiers,"*GaAsIC Symposium Conference Digest,* Monterey, CA, Oct. 2002.

M. Urteaga, D. Scott, T. Mathew, S. Krishnan, Y. Wei, M. Dahlström, M. Rodwell, "Characteristics of Submicron HBTs in the 140-220 GHz Band," *Proceedings 2001 Device Research Conference*, University of Notre Dame, IN, June 2002.

M. Urteaga, D. Scott, T. Mathew, S. Krishnan, Y. Wei, M. Rodwell "Single-stage G-band HBT Amplifier with 6.3 dB Gain at 175 GHz, " *GaAsIC Symposium Conference Digest,* Baltimore, MA, Oct. 2001.

M. Urteaga, D. Scott, T. Mathew, S. Krishnan, Y. Wei and M.J.W. Rodwell, “185 GHz Monolithic Amplifier in InGaAs/InA1As Transferred-Substrate HBT Technology,” *IEEE MTT-S 2001 International Microwave Symposium Digest*, Vol. 3, Phoenix Civic Center, Phoenix, AZ, May. .

M. Dahlström, X.-M. Fang, D. Lubyshev, M. Urteaga, S. Krishnan, N. Parthasarathy, Y.M. Kim, Y. Wu, J.M. Fastenau, W.K. Liu, M. Rodwell, “Wideband DHBTs using a graded carbon-doped InGaAs base,” *IEEE Electron Device Letters*, vo. 24, no. 7, July 2003.

M. Dahlström, Z. Griffith, M. Urteaga, M.J.W. Rodwell, X-M. Feng, D. Lubyshev, Y. Wu, J.M. Fastenau, W.K. Liu, “InGaAs/InP DHBT’s with > 370 GHz ft and fmax using a Graded Carbon-Doped Base” *Proceedings 2003 Device Research Conference*, Salt Lake City, UT, June 2003.

Y.M. Kim, M. Urteaga, M.J.W. Rodwell, A.C. Gossard, “High Speed, Low Leakage Current InP/$In_{0.53}Ga_{0.47}As$/InP metamorphic double heterojunction bipolar transistors,” *Electronics Letters*, Vol. 38, No. 21, October 2002.

S. Krishnan, Z. Griffith, M. Urteaga, Y. Wei, D. Scott, M. Dahlström, N. Parthasarathy, M. J. W. Rodwell, “87 GHz static frequency dividers in an InP-based mesa DHBT technology,” *GaAsIC Symposium Conference Digest*, Monterey, CA, Oct. 2002.

M.J.W. Rodwell, D. Scott, M. Urteaga, M. Dahlström, S. Krishnan, Z. Griffith, Y. Wei, N. Parthasarathy, Y.M. Kim, "Submicron InP bipolar transistors: scaling laws, technology roadmaps, advanced fabrication processes," *Proceedings 2002 International Conference on Solid State Devices and Materials*, Nagoya, Japan, September 2002.

M. J. W. Rodwell, S. Krishnan, M. Urteaga, Z. Griffith, M. Dahlström, Y. Wei, D. Scott, N. Parthasarathy, Y-M Kim, S. Lee, “Interconnects in 50-100 Ghz integrated circuits,” *Proceedings of the General Assembly of International Union of Radio Sciences*, Maastricht, Netherlands, August 2002.

Y. Wei, M. Urteaga, Z. Griffith, D. Scott, S. Xie, V. Paidi, N. Parthasarathy, M. Rodwell, “75 GHz, 80 mW InP DHBT power amplifier,” *IEEE Radio Frequency Integrated Circuits Symposium Technical Digest*, Philadelphia, PA, June 8-13, 2002.

Y. Wei, S. Lee, P.K. Sundararajan, M. Dahlström, M. Urteaga, M. Rodwell, “W-band InP/InGaAs/InP DHBT MMIC power amplifiers,” *IEEE MTT-S International Microwave Symposium Digest*, Seattle, WA, June 2002.

D. Scott, H. Xing, S. Krishnan, M. Urteaga, N. Parthasarathy and M. Rodwell, "InAlAs/InGaAs/InP DHBTs with polycrystalline InAs extrinsic emitter regrowth," *Proceedings 2002 IEEE Device Research Conference*, Santa Barbara, CA, June 2002.

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

Y. Wei, S. Lee, P.K. Sundararajan, M. Dahlström, M. Urteaga, M. Rodwell, "High current (100mA) InP/InGaAs/InP DHBTs with 330 GHz fmax," *Conference Digest Indium Phosphide and Related Materials Conference*, Stockholm, Sweden, May 2002.

M. Dahlström, M. Urteaga, S. Krishnan, N. Parthasarathy, M.J.W. Rodwell, X. M. Fang, D. Lubyshev, Y. Wu, J. M. Fastenau, W.K. Liu, "Ultra-Wideband DHBTs using a Graded Carbon-Doped InGaAs Base," *Conference Digest Indium Phosphide and Related Materials Conference*, Stockholm, Sweden, May 2002.

M. J.W. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S. C. Martin, R.P. Smith, S. Jaganathan, S. Krishnan, S. I. Long, R. Pullela, B. Agarwal, U. Bhattacharya, L. Samoska, and M. Dahlström. "Submicron Scaling of HBTs," *IEEE Transactions on Electron Devices*, Special Issue on the History of the Bipolar Junction Transistor, vol. 48, no. 11, Nov. 2001.

S. Jaganathan, S. Krishnan, D. Mensa, T. Mathew, Y.Betser, Y. Wei, D. Scott, M. Urteaga, M. Rodwell. "An 18 GHz continuous time sigma-delta analog-digital converter implemented in InP transferred substrate HBT technology," *IEEE Journal of Solid State Circuits*, Vol. 36, No. 9, September 2001.

S. Lee, H. J. Kim, M. Urteaga, S. Krishnan, Y. Wei, M. Dahlström and M. Rodwell, "Transferred-Substrate InP/InGaAs/InP Double Heterojunction Bipolar Transistors with fmax=425 GHz," *IEE Electronics Letters*, Vol. 37, No. 17, Aug. 2001.

T. Mathew, H.J. Kim, S. Jaganathan, D. Scott, S. Krishnan, Y. Wei, M. Urteaga, M.J.W. Rodwell and S. Long, "75 GHz ECL static frequency divider using InAIAs/InGaAs HBTs," *IEE Electronic Letters*, Vol. 37, No. 11, May 2001.

M.J.W. Rodwell, M. Urteaga, Y. Betser, D. Scott, M. Dahlström, S. Lee, S. Krishnan, T. Mathew. S. Jaganathan. Y. Wei, D. Mensa, J. Guthrie, R. Pullela, Q. Lee, B. Agarwal, U. Bhattacharya, S. Long "Scaling of InGaAs/InAIAs HBTs for High Speed Mixed-Signal and mm-Wave ICs," *International Journal of High Speed Electronics and Systems*, Vol. 11, No. 1, 2001.

ABSTRACT

Submicron InP-based Heterojunction Bipolar Transistors

by

Miguel Urteaga

This work examines the design and performance of submicron heterojunction bipolar transistors (HBTs) in the InP-based material system. Device results from two unique transistor geometries are considered. A transferred-substrate process has been used to aggressively scale the extrinsic collector-base capacitance of the transistor by lithographically patterning both sides of the device epitaxy. Deep submicron transferred-substrate HBTs have demonstrated peaking and in some cases singularities in the measured unilateral power gain ($U$). Associated with these measurements are negative resistance trends in the device output and feedback conductances, trends that cannot be modeled by standard HBT models. A potential explanation of the observed characteristics is electron velocity modulation in the collector-base junction. A theoretical model for capacitance cancellation by electron velocity modulation is developed, and its correlation with experimental data examined.

Because of their unique characteristics, the power gain cutoff frequency $f_{max}$ of transferred-substrate devices cannot be confidently extrapolated from low frequency device measurements. However, high levels of transistor power gain have been measured in the 140-220 GHz frequency band. Small-signal amplifiers have been fabricated in this frequency band, and a gain of 6.3 dB at 175 GHz has been demonstrated from a single-transistor design.

The second device topology considered in this work is an aggressively scaled mesa-HBT. The process flow and device epitaxy have been tailored for application towards digital logic design. Important characteristics for digital logic transistors are high current density operation, low extrinsic collector-base capacitance, low extrinsic contact resistances, and high device yield. The mesa-HBT process flow uses dielectric sidewall spacers and a tungsten-based base Ohmic contact to form the self-aligned base-emitter junction. A trench isolation process has also been developed to reduce the extrinsic collector-base capacitance of the transistor. Scaled mesa-HBTs have been realized operating at current densities > 5 mA/μm$^2$ with a simultaneous $f_t$ and $f_{max}$ of close to 300 GHz.

TABLE OF CONTENTS

LIST OF FIGURES

# Chapter 1 Introduction

The indium phosphide-based material system offers many desirable characteristics for the fabrication of wide bandwidth heterojunction bipolar transistors (HBTs). Compared to the silicon material system, the InP system offers higher electron mobilities, higher usable acceptor doping levels and a greater flexibility in engineering both the emitter-base and collector-base heterojunctions. Despite these considerable material properties advantages, the bandwidths of commercial InP-based HBTs, and particularly, the bandwidths of InP-based digital circuits are comparable to those of their Si/SiGe counterparts.

Device scaling- the reduction of lithographic features and epitaxial layer thicknesses- is necessary to extend transistor bandwidths. Si/SiGe bipolar transistors are scaled laterally to deep submicron dimensions through advanced processing and fabrication techniques. By contrast, InP-based HBTs are typically fabricated with junction dimensions >1μm, and a standard III-V HBT process flow makes further device scaling difficult.

In this work, the development and characterization of submicron InP-based HBTs is presented. Work has been performed using two unique device topologies. A transferred-substrate approach is used to lithographically define submicron emitter-base and collector-base junctions on both sides of the device epitaxy. This aggressive processing approach enables a dramatic reduction in collector-base capacitance, with a correlated increase in the transistor's maximum frequency of oscillation $f_{max}$. High

values of transistor power gain have been measured to frequencies up to 220 GHz.

The transferred-substrate technology demonstrates the high speed potential for a submicron InP HBT, however, the underlying fabrication steps in the process make it difficult, if not impossible, to scale for larger levels of integration and production. The levels of yield and integration achievable in the transferred substrate process has been limited to circuits with ~100 transistors fabricated on quarters of 2" wafers in a university research lab environment.

The second device topology considered in this work is a scaled mesa-HBT technology. For this device, fabrication steps have been developed aimed at improving the performance and manufacturability of a submicron device. The process relies on well-established dielectric sidewall spacer and dry etch processes to form a self-aligned base-emitter junction. A trench etch and dielectric refill has been incorporated to substantially reduce the extrinsic collector-base capacitance of the transistor. Additionally, the device epitaxy has been tailored to support a submicron process, with device parameters that are optimized for high digital logic speeds and not necessarily traditional transistor figures-of-merit ($f_t$, $f_{max}$).

The thesis is divided into two parts. The first section deals with the design, fabrication and characterization of transferred-substrate HBTs and circuits. In Chapter 2, the transferred-substrate technology is described and general scaling laws for HBTs are reviewed. The measurement and characterization of submicron HBTs presents numerous challenges. Chapter 3 describes the measurement and calibration methods used to make on-wafer network analyzer measurements to frequencies up to

220 GHz. In Chapter 4, the measured characteristics of submicron transferred-substrate HBTs are presented. In some highly scaled devices, negative resistance trends are observed in the reverse transmission and output conductance of the device, leading to a measured singularity in the transistor unilateral power gain. A theory to explain this phenomenon due to electron velocity modulation in the collector space charge region is presented. Chapter 5 describes the design and performance of millimeter-wave amplifiers designed in the 140-220 GHz (G-band) frequency range.

In the second section of the thesis, the development towards a scalable/manufacturable HBT process is described. In Chapter 6, HBT parameters of critical importance for digital circuit design are highlighted, and the scaled mesa-HBT process flow is presented. In Chapter 7, results from the scaled HBT technology are presented.

# Chapter 2 Transferred-Substrate HBTs

The transferred-substrate technology described in this section has been in development at UCSB since 1994. Device and circuit development in the technology has been the subject of 10 Ph.D. theses, and numerous graduate and post-doctoral students have made contributions to the technology. The work described here most closely resembles the research performed by Q. Lee [1], in that it has focused on extending the maximum frequency of oscillation ($f_{max}$) of the transistor through lateral device scaling.

In this chapter, the transferred-substrate technology is described in detail. General scaling laws for HBTs are first described in terms of the transistor figures-of-merit $f_t$ and $f_{max}$. An extended transistor Tee-model is presented to describe the distributed nature of the collector-base junction characteristics and more accurately describe the transistor's $f_{max}$. The transferred-substrate process flow is then presented, and key process and epitaxial features for the transistors used in this work are highlighted.

## *2.1 Device Scaling*

The transistor bandwidth is considered in terms two RF figures-of-merit, the current gain cutoff frequency $f_t$ and the maximum frequency of oscillation $f_{max}$. These figures-of-merit have been traditionally used as metrics to evaluate a technology's potential for RF and high-speed circuit applications. However, this analysis must be performed with care, as the relative contributions of the transit time and RC charging delay elements that determine a circuit bandwidth may differ considerably from those

that determine $f_t$ and $f_{max}$. This is particularly true for digital logic speed. A problem that will be considered in further detail in Chapter 6. Despite these limitations, an analysis of the delay elements in the expressions for $f_t$ and $f_{max}$ is beneficial in that it does show the interaction between delay terms as the transistor geometry is laterally and vertically scaled.

### 2.1.1 Scaling for $f_t$

For an HBT well described by a hybrid-p small-signal equivalent circuit model, the current gain cutoff frequency is given by

$$\frac{1}{2pf_t} = t_c + t_b + \frac{kT}{qI_c}\left(C_{je} + C_{cb}\right) + \left(R_{ex} + R_c\right)C_{cb} \qquad \text{Eqn. 2.1}$$

where $t_c$ and $t_b$ are the collector and base transit times, respectively, $C_{je}$ is the emitter-base junction capacitance, $C_{cb}$ is the total collector base junction capacitance, and $R_{ex}$ and $R_c$ are the extrinsic emitter and collector resistances, respectively.

Consider first the base transit time $t_b$. The introduction of a quasi-electric field in the base region, through compositional or doping grading, can significantly reduce the base transit time. For the HBT results presented in this work, a compositional grading of the InGaAs base from 45.5% Ga content at the base-emitter junction to 53.2% at the collector-base junction is used to introduce ~52 meV bandgap grading. If a linear grading of the base bandgap energy with position is used then [2]

$$t_b = \frac{T_b^2}{D_n}\left(\frac{kT}{\Delta E}\right) - \frac{T_b^2}{D_n}\left(\frac{kT}{\Delta E}\right)^2\left(1 - e^{-\Delta E/kT}\right) + \left(\frac{T_b}{v_{exit}}\right)\left(\frac{kT}{\Delta E}\right)\left(1 - e^{-\Delta E/kT}\right) \qquad \text{Eqn. 2.2}$$

where $?E$ is the grading of the base bandgap energy, $T_b$ is the base thickness, $D_n$ is the base minority carrier diffusivity, and $v_{exit}$ is the base exit velocity (of the order of $\left(kT/m^*\right)^{1/2}$, where $m^*$ is the electron effective mass in the base [2]). Note that the 52 meV bandgap grading described above is enough to reduce $t_b$ by ~2:1 for a 400 Å InGaAs base ($D_n$ = 40cm$^2$/sec, $v_{exit}$ = 3 x 10$^7$ cm$^2$/sec). Eqn. 2.2 indicates $t_b \propto T_b^2$ for a thick base layer or large $v_{exit}$, although the $v_{exit}$ term may add a significant correction for thin base devices.

Electrons entering from the base and traveling through the collector space charge region introduce a displacement current at the collector terminal. The mean delay of this displacement current defines the collector transit time. To first order in frequency, the collector transit time is given by [3, 4]

$$t_c = \int_0^{T_c} \frac{\left(1 - x/T_c\right)dx}{v(x)} \equiv \frac{T_c}{2v_{eff}} \qquad \text{Eqn. 2.3}$$

where $T_c$ is the collector thickness, $v(x)$ is the position dependent electron velocity, and $v_{eff}$ is used to describe an effective electron velocity ($v_{eff} = v_x$, for constant velocity profile).

Electrons entering the collector from the base undergo ballistic transport [5], where they may travel a significant portion of the collector region before obtaining sufficient kinetic energy to scatter to a higher effective mass, and hence lower effective velocity, satellite conduction valley (G-L valley separation: 0.55 eV for InGaAs [6], 0.6 eV for InP [7]). The velocity profile in the collector is often modeled

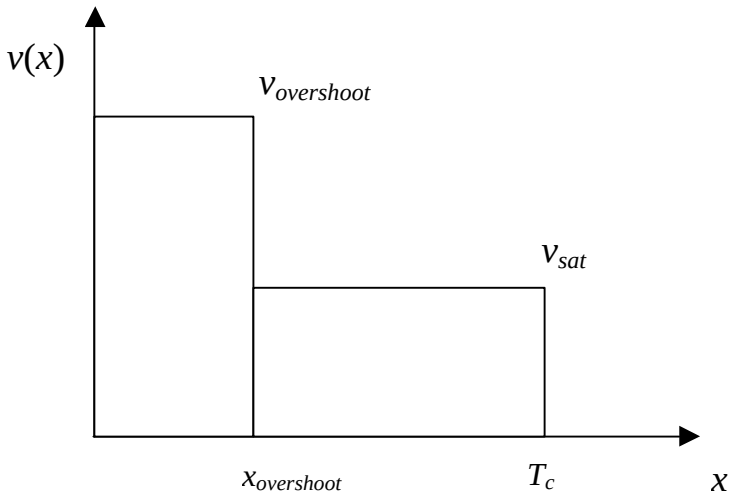



Figure 2.1: Two-step model of electron velocity profile in collector region of InP-based HBT.

by a two-step velocity profile (Figure 2.1) consisting of a high velocity ($v_{overshoot}$) region entering the collector followed after some appropriate distance by a slower velocity ($v_{sat}$) region that covers the remainder of the collector [3]. This velocity profile model is supported by Monte Carlo simulations of collector transport [8].

The described velocity profile in the collector is beneficial, since, given the form of Eqn. 2.3, the collector transit time depends most strongly on the velocity of electrons close to the collector-base junction boundary. In fact, extracted values of $v_{eff}$ may be higher than saturated drift velocity of the collector material. For the transferred-substrate devices presented in this work a $v_{eff}$ of 3.1 x $10^7$ cm/sec is extracted from device measurements compared to a reported $v_{sat}$ of 1 x $10^7$ cm/sec for InGaAs at comparable doping levels to the collector of an InP HBT [9].

Regardless of the exact velocity profile, the collector transit time is generally assumed proportional to the collector thickness.

The collector thickness also plays a critical role in determining the collector capacitance charging time $[kT/qI_c]C_{cb}$. Clearly, increasing the transistor's current density will reduce the delay term; however, the maximum current density is limited by the onset of the Kirk effect, or base pushout [10]. At high current densities, the electron space charge in the collector screens the bound donor charge resulting in a reduction in electric field near the base-collector junction. When the current increases so that the electric field at the base-collector junction boundary is reduced to zero, the Kirk threshold is reached.

In a single-heterojunction device, holes are no longer confined to the base and will enter the collector region resulting in an increase in the base transit time and the collector-base junction capacitance. It has been speculated that the high frequency performance of a single-heterojunction device may be enhanced if operated slightly in the Kirk regime. This improvement would result from a decrease in collector transit time due to reduction of the electric field (enhanced ballistic transport) near the base-collector junction [11]. In a double heterojunction device, the valence band barrier prevents holes from entering the collector region. With no flow of positive charge entering the collector, the conduction band in the collector will continue to bend upwards and eventually present a barrier to current flow entering the collector. A collapse in current gain, and increase in collector transit time is observed when the

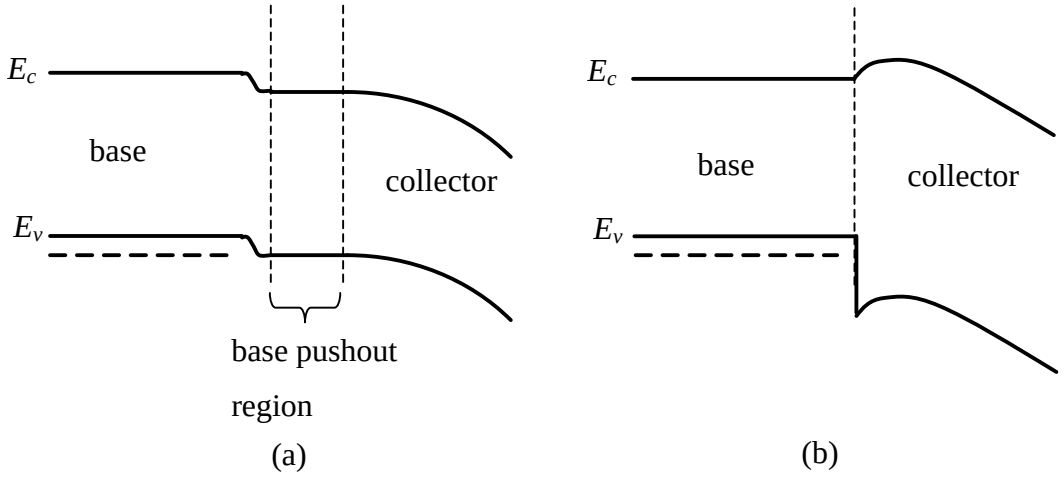



Figure 2.2: Simplified band diagrams for transistors operating in Kirk regime illustrating (a) base pushout for single heterojunction device and (b) current blocking in double heterojunction device.

Kirk threshold is reached in double-heterojunction devices. Figure 2.2 shows simplified band diagrams illustrating Kirk regime operation for single and double heterojunction junction devices.

The criteria for the Kirk threshold is determined by solving Poisson's equation with the presence of bound donor charge and mobile electron charge in the collector space charge region. Under the assumption that the collector doping is selected to result in a fully depleted collector at zero applied collector-base voltage (a desirable characteristic for digital logic applications), the condition for the onset of base pushout is

$$I_{c,\max} = A_e \left(V_{cb} + V_{bi}\right) 4ev_{eff} \left/ T_c^2 \right. \qquad \text{Eqn. 2.4}$$

where $A_e$ is the emitter area, $V_{cb}$ is the applied collector-base junction voltage, and $V_{bi}$ is the built in junction potential. This derivation assumes zero current spreading in the collector region, a condition that will increase the maximum current density for a practical device [12].

The collector-base capacitance is given by the dielectric capacitance $C_{cb} = eA_c/T_c$, where $A_c$ is the collector area. Combining this with Eqn. 2.4 gives, $[kT/qI_c]C_{cb} \propto T_c(A_c/A_e)$. Decreasing the collector thickness can thus reduce the delay term, however, the operating current density must increase to the square of the reduction. The increase in current density presents challenges for thermal management of high performance HBTs.

While the $[kT/qI_c]C_{cb}$ time constant can be reduced by decreasing the collector thickness and increasing the operating current density, the $R_{ex}C_{cb}$ delay term is not as easily managed. The extrinsic emitter resistance has contributions from the Ohmic contact resistance and the bulk resistance of the underlying emitter semiconductor layers. The shape and structure of the emitter layers determines the exact relationship of the terms. For example, the undercut of the emitter semiconductor during a wet etch process generally makes the area of the underlying semiconductor layers smaller than that of the defined contact. Rather than considering all terms in detail, the emitter resistance is often defined as

$$R_{ex} \equiv r_e/A_e \qquad \text{Eqn. 2.5}$$

where $?_e$ is a fitted parameter.

If the emitter contacting and semiconductor areas are kept equal (i.e. no undercut), we expect $?_e$ to remain constant regardless of emitter width. For devices fabricated at UCSB, it has been observed that $?_e$ tends to increase with decreasing emitter dimensions particularly for submicron devices. Possible reasons for this trend are an attack of the emitter metal contact during wet chemical etching, or a lateral depletion of the $n^-$ emitter semiconductor layer due to Fermi-level pinning at the surface. Typical values of $?_e$ obtained for InP-based HBTs are on the order of 10-50 O-$\mu m^2$.

With a known emitter resistivity, the collector capacitance charging delay term is then given by $C_{cb}R_{ex} = eA_c r_e / T_c A_e$ . One sees that even if $?_e$ is assumed to remain constant with decreasing emitter width, the delay term presents challenges to scaling. In order to reduce the delay term as the collector thickness is vertically scaled, improvements in the emitter resisistivity ($?_e$) or the collector-to-emitter area ratio ($A_c/A_e$) must be realized. An idealized HBT structure will have an $A_c/A_e$ ratio approaching 1. For the highly scalable transferred-substrate devices presented in this work, the ratio is typically ~2. The process flow used to fabricate a standard III-V mesa-HBT presents severe challenges to reducing this ratio, and approaches to reducing $A_c/A_e$ for mesa-devices will be discussed later.

Reducing the emitter resistivity is also difficult. Improvements in epitaxial material (higher achievable doping levels) or Ohmic contact metallurgies are

necessary but not easily achieved. Si/SiGe HBTs benefit from a emitter regrowth process that forms a low resistance polycrystalline contact above the contact. The contact area may be several times larger than the active emitter junction, and thus reduces the Ohmic contact resistance. Similar regrowth approaches are currently being pursued in InP-based technologies [13].

The collector resistance ($R_c$) of an InP HBT is typically less than half that of the emitter, and the contribution of the $R_c C_{cb}$ delay term to Eqn. 2.1 should be small. The unique geometry of a transferred-substrate device allows for a zero series resistance Schottky collector contact to be formed directly under the emitter-base junction. In a mesa-HBT the collector resistance has contributions from the Ohmic contact resistance the spreading resistance underneath the base mesa, and the gap resistance between the base mesa and the collector contact. For a mesa-HBT with collector contacts that run down both sides of the base mesa, the collector resistance is given by

$$R_c = \frac{\sqrt{r_c R_{sheet}}}{2L} + \frac{W_{mesa} R_{sheet}}{12L} + \frac{W_{gap} R_{sheet}}{2L} \qquad \text{Eqn. 2.6}$$

where $R_{sheet}$ is the sheet resistance of the subcollector, $?_c$ is the contact resistivity, $W_{mesa}$ is the width of the base mesa, $W_{gap}$ is the spacing between the collector contact and the base mesa, and $L$ is the length of the base mesa. Reduction of $R_c$ can be achieved by increasing the thickness of the sub-collector layer to reduce $R_{sheet}$.

However, in a typical mesa-HBT process, this approach comes with a loss of device planarity since the sub-collector must be removed to isolate devices.

In addition to its influence on RF performance, the collector resistance also plays a role in determining the maximum operating current density of the transistor. The $I_cR_c$ voltage drop across the resistance decreases the potential applied across the base-collector junction that determines the Kirk threshold. In fact, the term $V_{cb}$ in Eqn. 2.4 should be more accurately replaced by ($V_{cb}$ - $I_{c,max}R_c$). Clearly, controlling $R_c$ becomes more import as device current densities are increased.

Of all of the delay terms in Eqn. 2.1, the relationship between the emitter charging time constant $[kT/qI_c]C_{je}$ and the physical HBT parameters is the least straightforward. If one were to assume $C_{je}$ were simply a depletion capacitance, it would be reasonable to expect that this charging time could be minimized simply by making the emitter-base depletion region very thick through use of very low emitter doping and a large bandgap grading region. However, a thick depletion region will exhibit increased charge storage and an increase in ideality factor due to larger gradients in the quasi-Fermi level.

Charge storage effects in the emitter depletion region can be accounted for using an approach similar to that used to derive the collector transit time. Integral relations for the emitter charging time constant have been developed in [14] and [15]; however, these expressions have not been considered in detail in this work. If the emitter depletion region is thin, and doped heavily enough to support high current

density operation, the emitter depletion capacitance is determined by the depletion thickness and $[kT/qI_c]C_{je} = [kT/qI_c]eA_e/T_{depletion}$. Increasing the operating current density of the transistor thus reduces this time constant.

### 2.1.2 Scaling for $f_{max}$

The maximum frequency of oscillation of an HBT depends on the $f_t$ of the transistor and the distributed base-collector junction parasitics. An HBT $f_{max}$ is often approximated as $f_{\max} \cong \sqrt{f_t/8pR_{bb}C_{cb}}$, where $R_{bb}C_{cb}$ is the product of the base resistance and the full collector-base junction capacitance. This approximation is typically quite poor, as extrinsic collector-base capacitance that is not charged through the base resistance will not affect the transistor $f_{max}$. A more accurate determination of $f_{max}$ requires a consideration of the distributed collector-base network.

First consider the HBT base resistance $R_{bb}$. The base resistance consists of contributions from the contact resistance $R_{cont}$, the gap resistance between the emitter junction and base contact $R_{gap}$, and the spreading resistance underneath the emitter $R_{spread}$. Given a base sheet resistance $?_s$ and specific contact resistance $?_c$, the total base resistance is given by

$$R_b = R_{cont} + R_{gap} + R_{spread} \qquad \text{Eqn. 2.7}$$

$$R_{cont} = \frac{\sqrt{r_s r_c}}{2L_e}$$

$$R_{gap} = \frac{W_{be} r_s}{2L_e}$$

$$R_{spread} = \frac{W_e r_s}{12L_e}$$

where $W_e$ and $L_e$ are the emitter junction width and length respectively, and $W_{be}$ is the gap spacing between the emitter junction and the base contact.

Analyzing the terms in Eqn. 2.7 one sees that while the spreading resistance can be reduced by decreasing the emitter width, the remaining resistive terms will not be affected. Thus, reduction in the $R_{bb}C_{cb}$ product is limited unless the base-collector junction width is reduced or improvements are made in the base sheet and contact resistances. The challenge of scaling the collector-base junction width has generally frustrated the scaling of III-V mesa-HBTs for simultaneous high $f_t$ and $f_{max}$. The transferred-substrate process provided a means to simultaneously scaled both the emitter-base and collector-base junction to submicron dimenstions.

Figure 2.3 shows a schematic diagram of a transferred-substrate HBT with a distributed R-C network representing the collector-base junction parasitics. The schematic demonstrates the contributions of both the lateral spreading and vertical contact resistances. An accurate approach to modeling an HBT's $f_{max}$ is to enter the distributed model with an appropriately small grid spacing into a microwave circuit simulator. Alternatively, an analytic expression for $f_{max}$ can be developed from hand analysis of the distributed network. An analytical model for determining $f_{max}$ of a

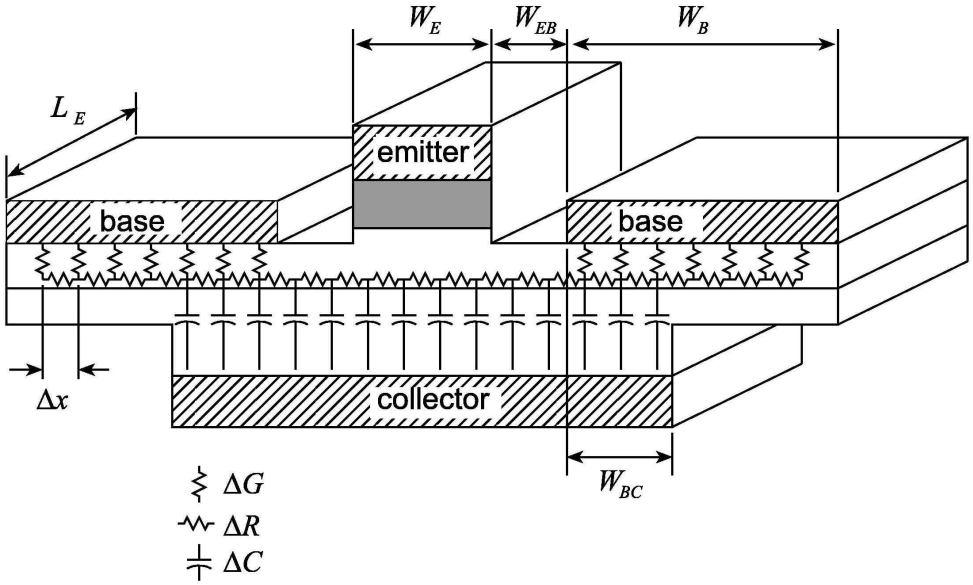



Figure 2.3: Cross-section of transferred-substrate HBT with distributed circuit elements modeling base-collector junction parasitics. The mesh-model can be entered in circuit simulator to accurately predict transistor $f_{max}$.

mesa-HBT was developed by Vaidyanathan and Pulfrey [16] This model can be extended to apply to a transferred-substrate HBT [1,15].

Figure 2.4 shows an equivalent small-signal Tee-circuit model for a transferred-substrate HBT. The model contains components common to a standard Tee-model [17]; however, the base-collector capacitance and base resistance have been partitioned to a number of elements to better model the distributed network of Figure 2.3. Three capacitances are defined with respect to the base-collector capacitance. $C_{cb,e} = eL_eW_e/T_c$ is the capacitance of the base-collector junction lying under the emitter. $C_{cb,gap} = 2eL_eW_{eb}/T_c$ is the capacitance of the junction lying

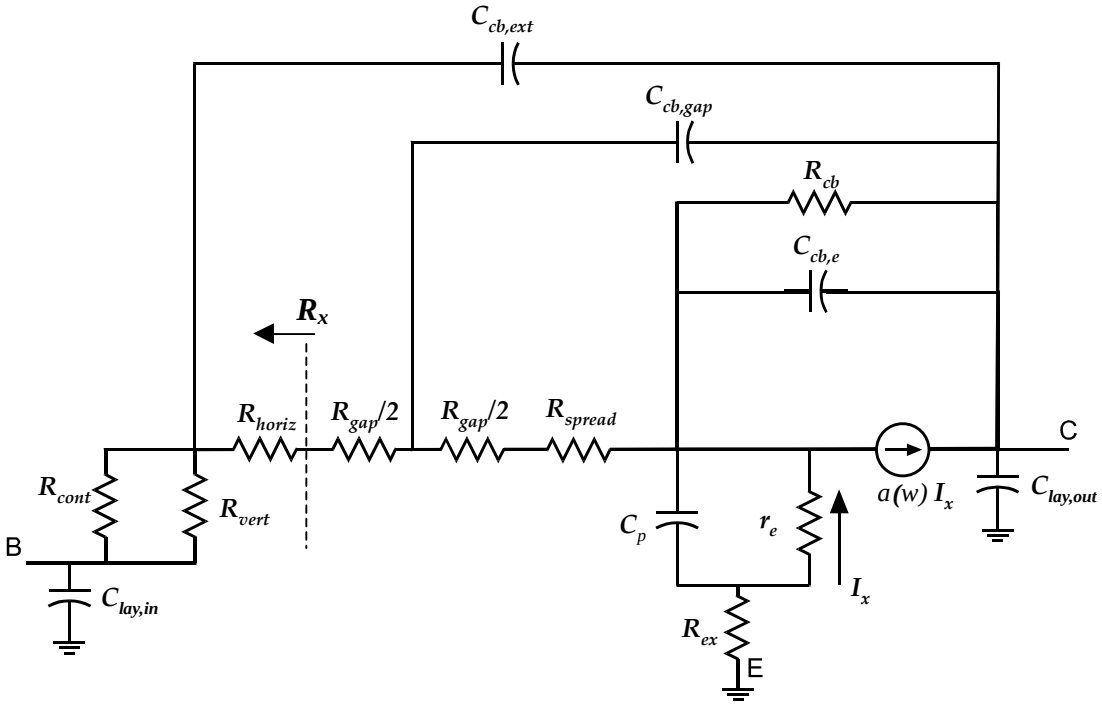



Figure 2.4: Small-signal Tee-model of transferred-substrate HBT with additional elements to describe base-collector junction parasitics.

underneath the gap between the emitter junction and base metal.

$C_{cb,ext} = 2eL_eW_{bc}/T_c$ is the capacitance of the junction lying under the base Ohmic contacts. Components of the base resistance are described in Eqn. 2.7, with the exception of two additional resistances. $R_{vert} = r_v/2W_{bc}L_e$ represents the vertical contact resistance over the path $W_{bc}$, and $R_{horiz} = r_sW_{bc}/2L_e$ represents the lateral sheet resistance over the same path.

The $R_{vert}/R_{horiz}$ network approximates the distributed network charging $C_{cb,ext}$ in the mesh model of Figure 2.3. This approximation is valid under the condition that $W_{bc}$ is less than the base Ohmic contact transfer length ($L_{contact}$). If this condition

is not met, then $R_{vert}$ and $R_{horiz}$ must be replaced with a finite element ladder network with a larger number of discrete elements. Additionally, the model assumes that $W_b >> L_{contact}$. Both of the aforementioned assumptions are generally satisfied for the transferred-substrate devices described in this work.

In the model of Figure 2.4, the charging resistances seen by $C_{cb,e}$ and $C_{cb,gap}$ contains the component $R_x = R_{cont} \| R_{vert} + R_{horiz} = R_{cont}$. While the simplified lumped element model approximates Figure 2.3 only if the $W_{bc} < L_{contact}$, the relationship $R_x = R_{cont}$ is generally true provided $W_b >> L_{contact}$. The same holds true for the expressions for $f_{max}$ presented below.

For a zero collector series resistance transferred-substrate HBT, the Vaidyanathan and Pulfrey model reduces to [15, 16]

$$f_{max} = \sqrt{\frac{f_t'}{8\pi t_{cb}}} \qquad \text{Eqn. 2.8}$$

$$\frac{1}{2\pi f_t'} = t_b + t_c + \frac{kT}{qI_c}\left(C_{je} + C_{cb}\right)$$

$$t_{cb} = C_{cb,e}\left(R_{cont} + R_{gap} + R_{spread}\right) + C_{cb,gap}\left(R_{cont} + R_{gap}\right) + C_{cb,ext}\left(R_{cont} \| R_{vert}\right)$$

The model of Figure 2.4 more accurately represents the distributed nature of the collector-base junction parasitics than a standard hybrid-p circuit model. However, one can approximate this network with a hybrid-p model (Figure 2.5) if the internal collector base capacitance $C_{cbi}$ is selected such that the correct transistor $f_{max}$ is obtained. The collector-base capacitance elements in Figure 2.5 are then given by

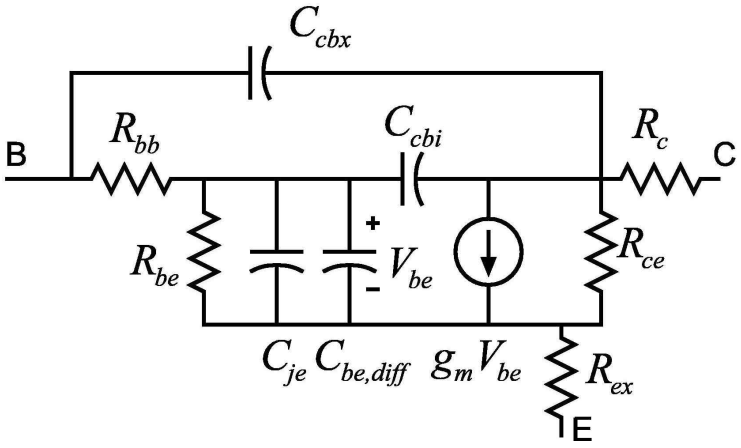



Figure 2.5: HBT hybrid-p equivalent circuit model. Circuit elements $C_{cbi}$ and $C_{cbx}$ are determined from time constant of Eqn. 2.8.

$C_{cbi} = t_{cb}/R_{bb}$ and $C_{cbx} = C_{cb} - C_{cbi}$, where $C_{cb} = C_{cb,e} + C_{cb,gap} + C_{cb,ext}$.

The hybrid-p model is useful when a compact representation of the transistor is desired. However, in the transformation from the Tee-model to the hybrid-p model some first order frequency approximations are made. In the hybrid-p model, the collector delay contributes to the base-emitter admittance through the diffusion capacitance $C_{be,diff} = g_m(t_c + t_b)$. Representing this contribution as a lumped capacitance assumes a first-order expansion of the collector transport factor $a_c(w) = \frac{\sin(wt_c)}{wt_c} e^{-jwt_c}$. Further lumped circuit elements could in principal be added to the hybrid-p model to more accurately model the base-emitter admittance at higher frequencies. However, given the added complexity, the use of the Tee-model may be preferred. In analyzing the device results presented in Chapter 4, the Tee-model representation is employed.

## 2.2 Transferred-substrate HBT Process Flow

The transferred-substrate process enables an aggressive scaling of the collector-base junction capacitance that cannot be realized with a conventional mesa-HBT geometry. The large values of transistor $f_{max}$ achieved with the devices have demonstrated the potential of a low parasitic HBT technology for millimeter-wave amplification. The primary disadvantage of the technology is the complexity of the process flow and the limited yield and levels of integration that have been achieved. In this section, details of the transferred-substrate technology are described. The epitaxial layer structure for the devices used in this work will be presented first. The device process flow will then be described and specific process steps that present yield limitations for the technology will be pointed out.

### 2.2.1 Epitaxial Layer Design

The device layer structure for the transistors fabricated in this work is shown in Table 2.1. The material used in this work was grown by molecular beam epitaxy (MBE) on Fe-doped semi-insulating InP substrates. Material was grown in the MBE laboratory at UCSB and purchased from the commercial epitaxy vendor IQE. The layer structure is described from the top down. A 1000 Å InGaAs emitter provides a good contacting layer for the emitter Ohmic contact. The emitter cap is followed by a chirped-super lattice (CSL) InGaAs/InAlAs grade that removes the conduction band discontinuity between the emitter cap and the wideband gap InAlAs emitter. The total emitter thickness is kept relatively large (>2000 Å) to facilitate the formation of self-aligned base Ohmic contacts.

| Layer | Composition | Dopant | Thickness |
|---|---|---|---|
| Emitter Cap | $In_{0.53}Ga_{0.47}$ As | Si: 1x $10^{19}$ $cm^{-3}$ | 1000 Å |
| Emitter Cap Grade | $In_{0.53}Ga_{0.47}$ As / $In_{0.52}Al_{0.48}As$ | Si: 1x $10^{19}$ $cm^{-3}$ | 200 Å |
| $N^{+}$ Emitter | $In_{0.52}Al_{0.48}As$ | Si: 1x $10^{19}$ $cm^{-3}$ | 700Å |
| $N^{-}$ Emitter | $In_{0.52}Al_{0.48}As$ | Si: 8x $10^{17}$ $cm^{-3}$ | 500Å |
| Base-Emitter Grade | $In_{0.45}Ga_{0.55}$ As / $In_{0.52}Al_{0.48}As$ | Si: 8x $10^{17}$ $cm^{-3}$ | 233 Å |
| Base-Emitter Grade | $In_{0.45}Ga_{0.55}$ As / $In_{0.52}Al_{0.48}As$ | Be: 2x $10^{18}$ $cm^{-3}$ | 67 Å |
| Base | $In_{0.45}Ga_{0.55}As$ ? $In_{0.53}Ga_{0.47}$ As | Be: 4x $10^{19}$ $cm^{-3}$ | 400 Å |
| Collector Setback | $In_{0.53}Ga_{0.47}$ As | Si: 1x $10^{16}$ $cm^{-3}$ | 400 Å |
| Collector Pulse Doping | $In_{0.53}Ga_{0.47}$ As | Si: 1x $10^{18}$ $cm^{-3}$ | 50 Å |
| Collector | $In_{0.53}Ga_{0.47}$ As | Si: 1x $10^{16}$ $cm^{-3}$ | 2550 Å |
| Growth Buffer | $In_{0.52}Al_{0.48}As$ | U.I.D | 2500 Å |
| Substrate | InP | Semi-Insulating | |

Table 2.1: Transferred-substrate HBT epitaxial layer structure.

An InGaAs/InAlAs CSL grade with a 33 Å period is inserted between the base and emitter junctions. The addition of the grade reduces the turn-on voltage of the junction by the amount of the conduction band discontinuity (~0.4V). The last few periods of the grade are beryllium doped to suppress out-diffusion of beryllium from the base to the emitter [18]. The base is 400 Å thick and linearly graded in 50 Å intervals from $In_{0.45}$ $Ga_{0.55}As$ to $In_{0.53}$ $Ga_{0.47}$ As (lattice matched composition). The grading introduces a 50 meV quasi-electric field in the base.

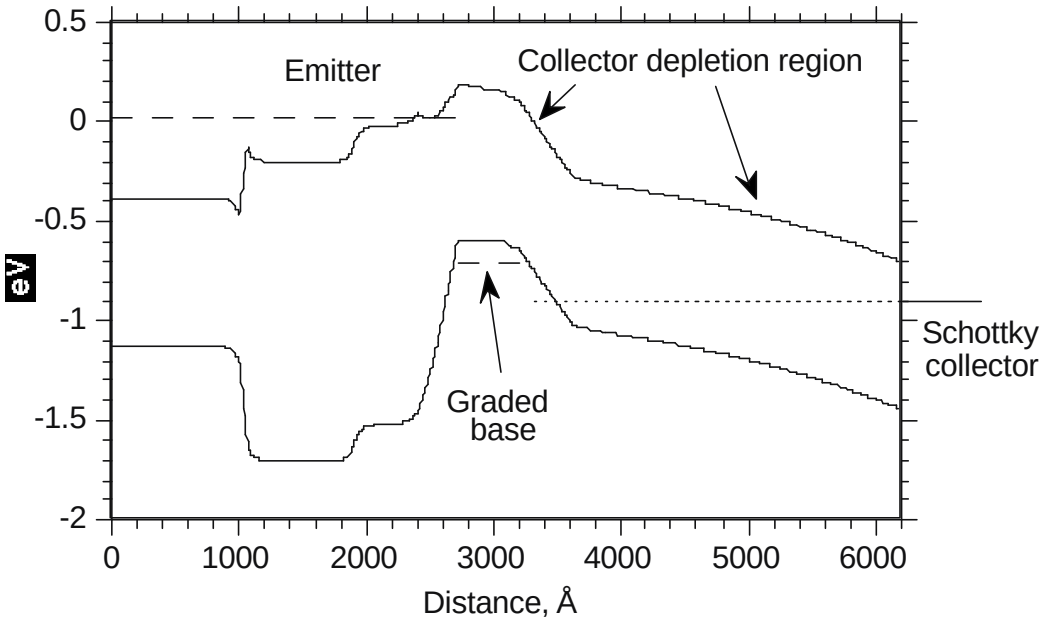



Figure 2.6: Band diagram of single-heterojunction transferred-substrate device with device epitaxy described by Table 2.1. Bias conditions: $V_{be}$ = 0.7 V, $V_{ce}$ = 0.9 V.

The transferred-substrate HBTs described in this work are single-heterojunction devices with a total collector thickness of 3000 Å. A 50 Å collector pulse doping layer located 400 Å from the base-collector junction is inserted to delay the onset of base pushout [18]. The pulse doping layer is nominally doped at 1x $10^{18}$ $cm^{-3}$. This doping was used in the epitaxy for the amplifier results presented in Chapter 5. However, for the device results presented in Chapter 4 a doping of 1x $10^{17}$ $cm^{-3}$ was mistakenly used. The device structure is grown on an InAlAs buffer layerthat is unintentionally doped. During the substrate transfer step this buffer layer i removed to expose the collector epitaxy. A band diagram of the layer structure in forward active mode of operation is shown in Figure 2.6.

### 2.2.2 Process Flow

#### Emitter Contact Deposition and Mesa Etch

The process flow begins with the definition of the emitter Ohmic contacts. For the transistors described in this work, electron-beam (e-beam) lithography was used to define submicron emitter stripes. A bi-layer PMMA (polymethyl methacrylate) positive e-beam resist is used to form an undercut resist profile and facilitate metal liftoff. Lithography was performed using the UCSB JEOL JBX-5D11 e-beam writer. The emitter metal was deposited in an electron-beam evaporation system. The emitter metal stack is Ti/Pt/Au/Si (200Å/400Å/9000Å/500Å). The top Si layer protects the emitter metal during the emitter mesa dry etch.

A self-aligned emitter mesa etch is performed using the emitter Ohmic contact as the etch mask. The etch is a combination dry/wet etch that is designed to minimize the undercut of the emitter semiconductor. A methane($CH_4$)/hydrogen/argon (MHA) reactive ion etch (RIE) is performed to etch through the emitter cap and grade and stop in the $N^-$ InAlAs emitter region.

The MHA etch is known for producing a large number of polymer byproducts, particularly when etching InAlAs [19]. In the HBT emitter etch, these byproducts may stay on the sidewalls of the emitter semiconductor and remain there during the wet chemical etch impeding the undercut of the emitter semiconductor. This can lead to the base-emitter short circuits when the self-aligned base Ohmic contact is evaporated. In the past, polymer removal was performed using an oxygen ($O_2$) RIE after the emitter etch cycle [1]. Near the end of the experimental work

performed in this thesis, this etch process began to fail due to some unknown changes in the RIE system. A new process was developed where short duration $O_2$ RIE polymer removal cycles were added after MHA etch cycles in five minute intervals. Using this process, polymer etch products were removed from the emitter sidewalls and an emitter undercut was achieved during the wet etch process. Figure 2.7 shows the emitter etch profile using the cycled RIE etch process.

In the HBT emitter etch process, laser endpoint detection is used to stop the dry etch in the $N^-$ InAlAs emitter region. The remaining emitter semiconductor is etched using a wet chemical etch process. The InAlAs layers are etched using a hydrobromic acid/acetic acid/hydrochloric acid/DI water mixture. The etch stops selectively in the base-emitter grade when a sufficient InGaAs composition is reached. A non-selective citric acid based etch with a slow etch rate (~10 Å/sec) is used to etch the remainder of the base-emitter grade and etch ~100 Å into the base semiconductor. At completion of the emitter mesa etch, the total undercut of the emitter semiconductor layers is ~0.1 μm on each side of the emitter contact (see Figure 2.7). While the undercut may seem like a relatively small amount, the total represents a limit for deep submicron scaling of the emitter-base junction.

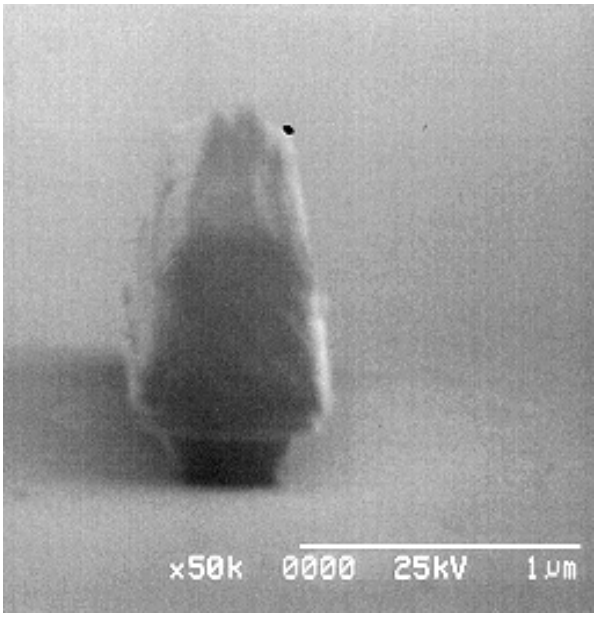



Figure 2.7: Emitter etch profile after cycled MHA/$O_2$ RIE process followed by wet chemical etch.

The disadvantage of the MHA dry etch is its incompatibility with carbon-doped InGaAs layers. While beryllium-doped material was used for the transferred-substrate devices reported in this work, carbon is a preferred donor for InGaAs layers. Compared to beryllium, carbon has higher achievable activated doping levels and a lower diffusion coefficient [18]. Hydrogen passivation of carbon-doped and a lower diffusion coefficient [18]. Hydrogen passivation of carbon-doped InGaAs is a well-known phenomenon. In metalorganic vapor deposition (MOCVD) growth of InGaAs epitaxial layers, hydrogen passivation limits the activated dopant concentration that can be achieved in carbon-doped layers [20, 21]. Hydrogen passivation of carbon-doped InGaAs base layers is also observed when performing a MHA RIE of the emitter semiconductor.

In MOCVD grown InGaAs, a high temperature post-growth anneal may be used to drive out the incorporated hydrogen and activate the carbon dopant [22,23]. Experiments were performed to determine whether such an anneal could be used to for MHA passivated InGaAs layers. Samples were prepared with an InGaAs/InAlAs emitter layer on top of a 1000 Å carbon-doped InGaAs layer. A five minute MHA RIE was used to etch into the InAlAs layer. The remaining InAlAs layer was then etched using a selective wet chemical etch that stopped on the carbon-doped InGaAs layer. High temperature anneals of the samples were performed at increasing temperatures and Hall measurements were then made to determine the bulk resistivity of the InGaAs layer. The Hall measurements were compared to a sample that did not have MHA RIE exposure.

Table 2.2 shows the bulk resistivity measurements of the InGaAs samples. For anneal temperatures less than 400 C, the InGaAs samples produced N-type Hall measurements with very high bulk resistivity values. It was found that an anneal temperature of at least 400° C was required to begin reactivation of the carbon donors, and a 30 minute anneal at 450° C was sufficient to reestablish the baseline doping level of the material. The temperature and length of time of the anneal required to reactivate the carbon donors are too high to be compatible with a non-refractory emitter Ohmic contact metallization. In some MOCVD grown InP HBT processes, W- or WSi refractory emitter Ohmic contacts are used and a carbon activation anneal is performed after emitter mesa formation [24].

| Anneal Temperature | Anneal Time | Resistivity |
|---|---|---|
| NO MHA Exposure | | 38.1 O-μm |
| 300° C | 10 min | N-type |
| 350° C | 10 min | N-type |
| 400° C | 10 min | 254.0 O-μm |
| 450° C | 10 min | 59.6 O-μm |
| 450° C | 30 min | 34.7 O-μm |

Table 2.2: Bulk resistivity of C-doped InGaAs layers after exposure to MHA RIE and subsequent high temperature anneal.

### Base Contact Deposition and Mesa Etch

The undercut of the emitter semiconductor during the emitter mesa etch allows the emitter contact to be used as a shadow mask for the evaporation of the base Ohmic metal. The base contact pattern is defined using an image reversal photolithography step, with the pattern defined as to surround the emitter contact. The base metal is then deposited using electron-beam evaporation. The line-of-site deposition profile of the evaporator prevents base metal from coming in contact with the undercut emitter semiconductor. In order to avoid base-emitter short circuits, the base metal thickness must be considerably less than the emitter semiconductor thickness. Even when this condition is met, metal strands that are commonly seen in metal liftoff processes may cause base-emitter short circuits. These short circuits have been one of the primary yield limitations observed in the transferred-substrate process. The probability of the failure has been observed to vary from process run to process run

[25], with variation in emitter liftoff profile and/or emitter mesa etch profile likely influencing the yield.

The quality of the base Ohmic contact is influenced by the choice of base metallization and the treatment of the base surface before metal evaporation. For the majority of the transferred-substrate work reported in this thesis, a non-alloyed Ti/Pt/Au (200Å /400 Å/800 Å) base metallization was used, and the wafer surface was treated with a $O_2$ plasma etch and dilute ammonium hydroxide etch just prior to entering the evaporator. Using this process, Ohmic contact resistivities ranging from 50-150 O-$\mu m^2$ were generally obtained, with typical results tending towards the higher end of the range.

Work performed at UCSB by M. Dahlström resulted in significant improvements in base Ohmic contact resistivity and run-to-run uniformity [26]. Process changes were made to both the metallization scheme and the sample surface preparation. In the revised process flow, a UV Ozone treatment is used in place of the $O_2$ plasma descum to remove residual photoresist from the wafer surface. It is believed that the $O_2$ plasma treatment that had been used may have caused damage to the InGaAs base layer. By contrast, UV Ozone treatment of InP compounds has been found to produce a stoichiometric oxide on the semiconductor surface that reduces surface defects [27]. After UV Ozone treatment, a dilute ammonium hydroxide etch is used prior to loading the sample in the evaporator.

The improved base contact scheme uses a Pd/Ti/Pd/Au (30 Å/200Å /400 Å/800 Å) metal stack. It has been found that the insertion of thin palladium (Pd) or

platinum (Pt) layers in contact with the semiconductor surface improves contacts to p-type InGaAs [28]. In addition to having high work functions that are favorable to forming p-type Ohmic contacts, these metals tend to react with the semiconductor surface upon deposition facilitating the removal of the native oxide [29]. Using the improved base contact process, contact resistivities in the range of 10 $\Omega$-$\mu m^2$ are typically measured using the transmission line method (TLM). In this range of values, determining a precise value of the contact resistance is difficult due to uncertainty in the exact TLM gap dimensions.

After depositing the base contact metal, the base mesa can be defined. The base mesa etch mask is aligned to the base contact pattern and protects the base-emitter junction during the mesa etch. For the single-heterojunction devices, described in this work the base mesa etch is performed using a chlorine-based RIE. The etch goes through the InGaAs collector layer and is stopped somewhere in the InAlAs buffer. The etch depth is controlled by monitoring the optical signal from a laser interferometer system. Figure 2.8 shows a schematic cross-section of a HBT after the base-mesa step. A SEM image of a transferred-substrate after the base mesa isolation is shown in Figure 2.9.

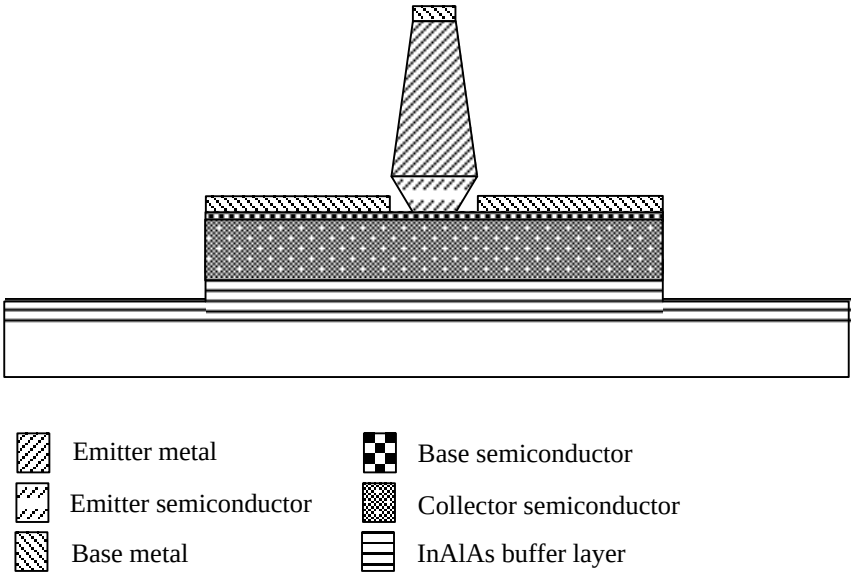



Figure 2.8: Transferred-substrate HBT cross-section after base mesa isolation etch.

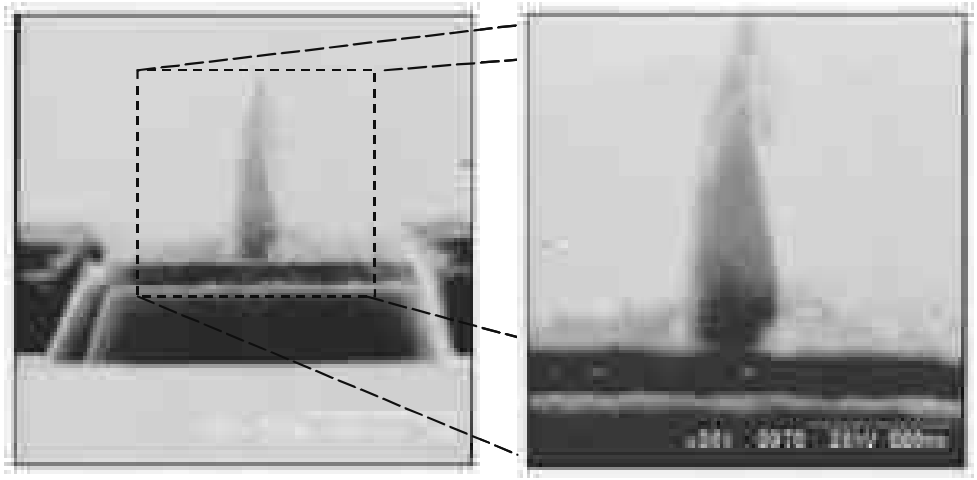


Figure 2.9: SEM image of submicron transferred-substrate HBT after base mesa isolation. Image taken by D. Scott.

### Device Passivation and First Level Interconnect

The transistor is passivated using the spin-on-polymer polyimide, a product that has been found to provide low surface leakage passivation of InP-based semiconductors [30]. The polymer is blanket deposited on the wafer at a thickness of ~1.8 μm, and oven-cured in a $N_2$ ambient. A planarization and etchback process is used to level the polyimide surface and expose the emitter post.

A thick photoresist (~2.3 μm) is spun onto the wafer to help planarize the underlying topology. The photoresist/polymide stack is then etched in a $O_2$ RIE system with the etch parameters adjusted to provide equal etch rates of the two materials. A laser interferometer system is used to monitor the etchback. By counting the number of interference cycles, the etchback depth is controlled so that the emitter post is exposed out of the polyimide. The polymide is the patterned using a photoresist mask. Mesas of polymide are left around individual devices and a via in the polymide is formed to contact the base metal.

After patterning the polyimide the first level of interconnect metal is defined by photolithography and e-beam evaporation. The interconnect metallization is a Ti/Au (200 Å/10,000 Å) stack. Interconnect metal is brought over the polyimide mesa to contact the emitter post and base contact via. Step coverage over the mesa is aided by angling the sample on a rotating stage during the metal evaporation. Wiring for the collector contact is also brought into proximity of the device, and this interconnect is contacted after substrate transfer. Figure 2.10 shows a schematic

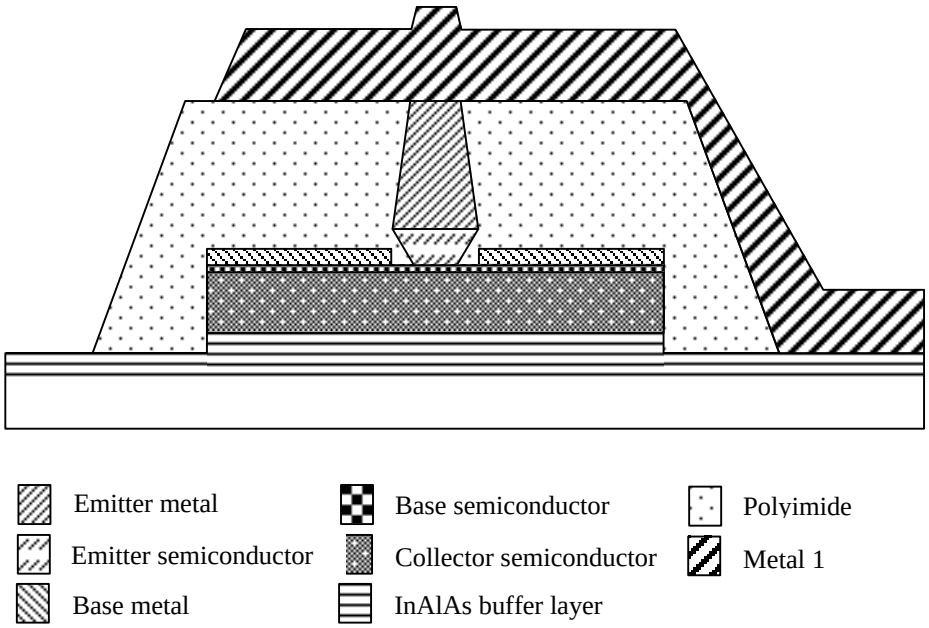



Figure 2.10: Transferred-substrate HBT cross-section after deposition of Metal 1 interconnect.

cross-section of the transistor after polyimide passivation and deposition of the interconnect metal.

### Circuit Element Definition

When fabricating transistors solely for device measurements, a single interconnect level is sufficient. In an IC process, such as that used to fabricate the amplifiers described in this work, an additional interconnect layer as well as thin-film resistors and metal-insulator-metal (MIM) capacitors are required.

Thin-film NiCr resistors are fabricated using e-beam evaporation, and the targeted sheet resistance of the resistors is 50 O/square. The resistor metallization is

deposited prior to the first level interconnect which is subsequently used to contact the resistors.

A 4000 Å $Si_xN_y$ layer serves as the insulator for MIM capacitors and is deposited by plasma enhanced chemical vapor deposition (PECVD). The $Si_xN_y$ is blanket deposited over the entire wafer and then pattern etched to remain only where necessary (capacitor structures, metal crossovers). The bottom plate of the capacitor is formed by the first level interconnect and the top plate by the second level interconnect. The MIM capacitors have a nominal capacitance value of 0.16 fF/ $\mu m^2$.

In addition to serving as the bottom capacitor plate, the second level interconnect may also be used as an additional level of wiring in an IC design. The importance of this interconnect level is more pronounced in densely integrated digital IC designs, rather than the microwave designs described in this work. The second level interconnect serves an important role in the thermal shunt via that is present under the devices, a process step that will be described shortly.

### BCB Dielectric and Ground Plane Definition

Until this point, the transferred-substrate process has resembled a standard III-V mesa-HBT process flow, with the obvious exception being the absence of the collector contact definition. Preparation for substrate transfer begins with the deposition of a thick spin-on-polymer that will serve as the dielectric medium for on-wafer transmission lines. Bezocyclobutene (BCB) is spun onto the wafer at a thickness of ~6 μm and oven-cured in an $N_2$ ambient. In addition to having excellent planarization properties, BCB also has a low dielectric constant ($e_r$ ~ 2.7) and

exhibits low dielectric loss at microwave frequencies [31], two desirable properties for forming on-wafer transmission lines.

A $CF_4/O_2$ RIE process is used to etchback the BCB to a final thickness of 5 μm while forming via holes in the dielectric. The via holes are necessary to make contact to the metal ground plane that will be deposited on top of the BCB. For the device measurements and amplifier results presented in this work, devices are configured in a common-emitter configuration. In this configuration, the BCB via is placed directly over the HBT footprint and contacts the first interconnect metal.

Obviously, other device configurations are required for different circuit applications, and the emitter of the device cannot always be grounded. However, even in these instances, the poor thermal conductivity of BCB requires that a via be placed over the device to serve as a thermal shunt. The thermal via is realized by placing a MIM capacitor above the device and patterning the via above the second level interconnect. This configuration results in a large parasitic layout capacitance to ground that may reduce circuit bandwidth in many applications. Additionally, short circuits through the $Si_xN_y$ above the device have been found to severely limit transistor yield [25].

After patterning the BCB, a Ti/Au seed layer is blanket deposited on the wafer using sputter deposition. The conformal deposition of the sputter process ensures continuous film coverage in the via holes. An Au ground plane is then electroplated over the entire wafer to a final thickness of ~5 μm. A cross-sectional schematic of the HBT after the ground plane deposition is shown in Figure 2.11.

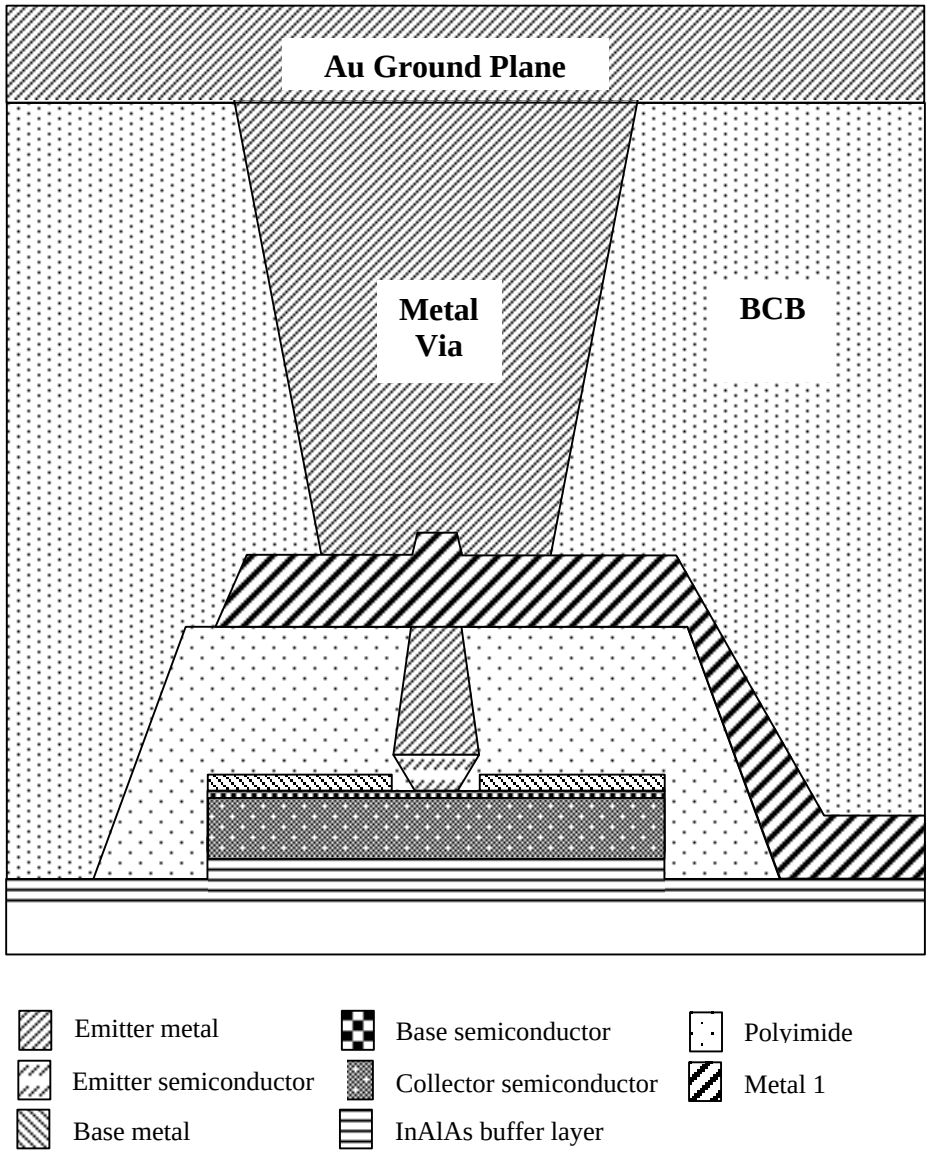



Figure 2.11: Transferred-substrate HBT cross-section after ground plane deposition.

### Substrate Transfer

Substrate transfer begins with the bonding of the InP substrate to a carrier wafer. In this work, both GaAs and later AlN substrates have been used as the carrier wafer.

Bonding is performed using a flip-chip bonding system. The InP wafer is brought up to temperature on one of the bonding machines chucks, and an indium/lead (In/Pb 60%/40%) based solder is spread evenly over the surface of the gold plated ground plane. This step in the process requires a certain degree of manual dexterity, as the solder is spread over the wafer using the wooden end of a cotton swab. A thin and uniform application of the solder is essential to ensure the sample is free of voids and air bubbles after the substrate transfer.

After applying the solder over the InP wafer, the carrier wafer, which has also had a Ti/Au seed layer deposited on it, is placed metal side down on top of the InP wafer. The bonder is then used to clamp the two wafers together with an appropriate prescribed force. The resulting InP/carrier wafer sandwich is then allowed to cool while under pressure completing the bonding step.

The substrate transfer is performed by selectively removing the InP substrate, leaving the device epitaxy and BCB dielectric bonded to the carrier wafer. The InP substrate is etched in a 4:1 HCl:DI water solution with complete removal of a 500 μm substrate typically taking ~ 1 hour. For single-heterojunction devices, the substrate-transfer etch stops selectively on the InGaAs collector. Double heterojunction devices fabricated in the transferred-substrate process required the addition of multiple-etch stops above the InP collector layer. Even with the addition of these etch stops, the substrate transfer etch often attacked the collector semiconductor presenting a severe yield limitation for double heterojunction devices in the technology [32].

As the description of the process flow suggests, the transferred-substrate process was not easily scaled to large scales of integration. Attempts were made to develop a copper plating process that would eliminate the mechanical bonding step and allow fabrication of larger samples [33]. Despite some limited success, refining this process proved too onerous given the limitations of a university research environment. Given these limitations, process runs were routinely performed on quarters of 2” or 3” wafers.

### Collector Contact Definition

After substrate transfer, the collector contact is formed on the exposed collector semiconductor. One of the challenges of the transferred-substrate process was the alignment of the collector contact to the underlying emitter-base junction. The mismatch of thermal expansion coefficients between InP and the carrier substrate would lead to a lateral expansion or shrinkage of patterns after substrate transfer. A shrinkage of approximately 1 part in 3000 as observed when bonding to GaAs wafers, and an expansion of approximately 1 part in 10000 was observed when bonding to AlN carrier wafers. Such dimensional changes are significant when submicron alignments are required. When using optical stepper lithography, the dimensional change would have to be accounted for in the collector mask design and the stepper array dimensions.

In the work described here, electron beam lithography was used to define collector contacts. In this process, the e-beam writer could perform local alignments for each individual transistor. Local alignment marks formed in the emitter contact

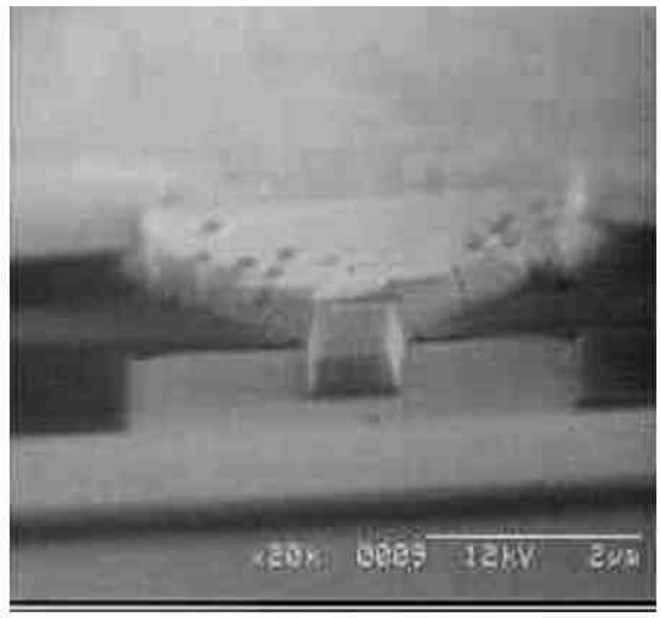


Figure 2.12: SEM image of transferred-substrate HBT collector contact written by electron beam lithography.

layer were placed within 100 μm of each device, and an alignment was performed before each collector contact was written. Excellent alignment tolerance (<0.1 μm) was typically achieved despite the dimensional variation in the samples.

The collector contacts were formed using the same lithography process previously described for the emitter. A Ti/Pt/Au (200Å /400 Å/4000 Å) metal stack is deposited by e-beam evaporation. An SEM photograph of a collector contact after liftoff is shown in Figure 2.12. After lift-off of the collector contacts, a collector recess etch is typically performed to remove some of the excess collector semiconductor and reduce fringing fields. The recess etch is performed self-aligned to the collector contact using a citric acid based etch. Typically, ~1000 Å of collector semiconductor

is removed during the etch, and this is found to lead to a dramatic improvement in high frequency device performance. The recess etch completes the transferred-substrate process. A schematic cross-section of the final device geometry is shown in Figure 2.13

## *2.3 Conclusions*

In this chapter, general scaling laws for HBTs were reviewed and a transferred-substrate HBT process was described. The substrate transfer process enables the base-collector junction to be scaled to submicron dimensions, thus addressing one of key impediments to increasing a transistor's $f_{max}$. The technology has demonstrated state-of-the art transistor power gain at mm-wave frequencies, and high-gain tuned amplifiers have been designed in the 140-220 GHz frequency band. These results will be described in the following two chapters.

Key disadvantages of the transferred-substrate technology are the complexity of the process flow and the limited yield and levels of integration that could be achieved. As described previously, the process of performing the substrate transfer was labor intensive and prone to periodic failures. Efforts to improve the substrate transfer process were limited by the resources available in a university research cleanroom environment. In addition to the difficulties associated with substrate transfer, the technology also suffered from yield limitations that are inherent to many mesa-HBT technologies. Chief among these was the self-aligned base-emitter contact process where control of the emitter undercut was extremely challenging for submicron devices. Periodic base-emitter short circuits would often be observed due

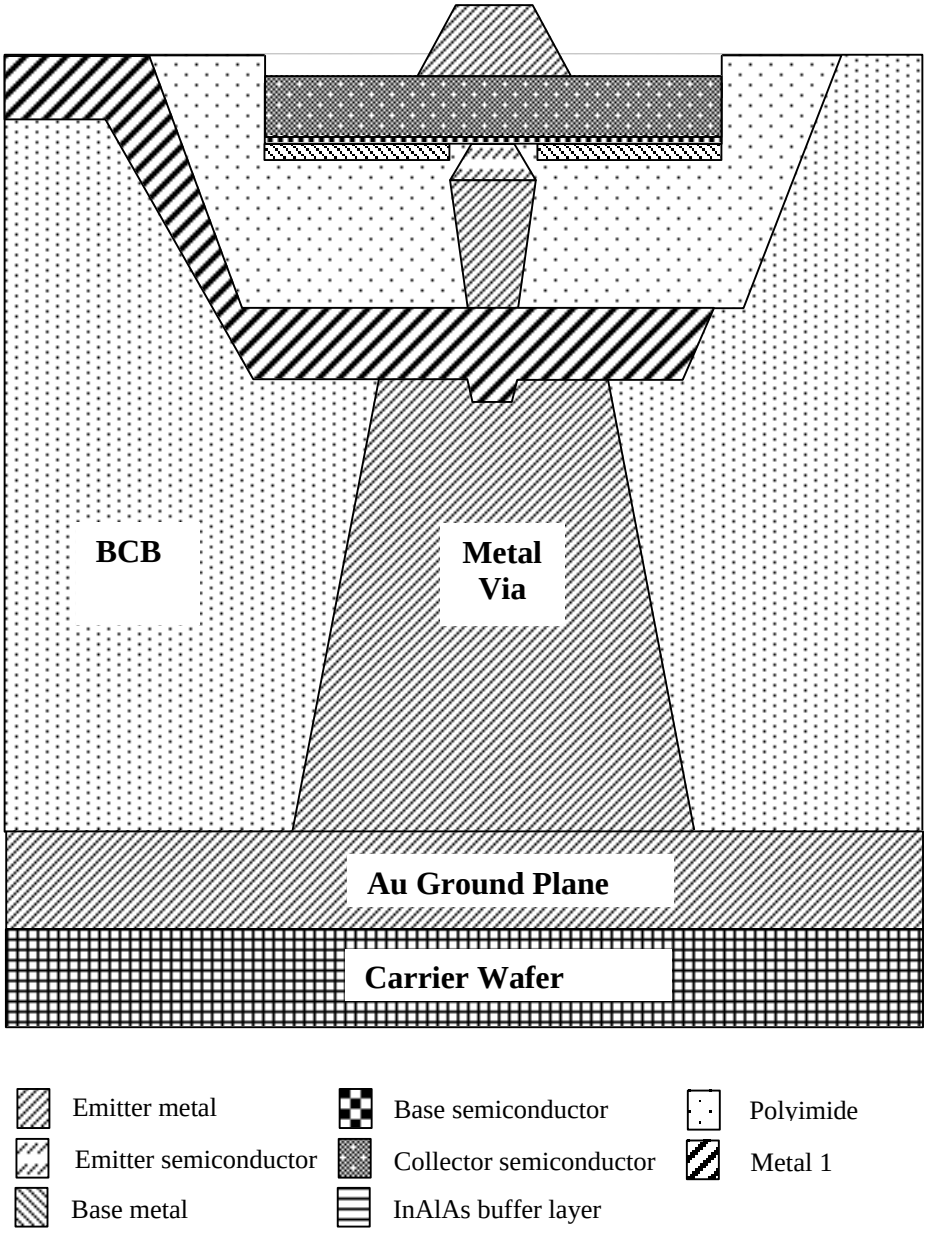



Figure 2.13: Transferred-substrate HBT cross-section after collector contact definition

to incomplete mesa etching or the presence of metal strands from liftoff processes.

From a transistor performance perspective, it should be noted that the benefits of the technology are gained from a dramatic reduction of the $R_{bb}C_{cb}$ time constant. While this reduction benefited $f_{max}$-limited tuned circuit applications, the advantages for mixed-signal and digital circuits are not as great. As described earlier in this chapter, scaling of all transistor parasitics (transit times and *RC* time constants) requires both vertical scaling of transistor dimensions and lateral scaling of the device epitaxy. Simultaneously, the operating current density of the transistor must increase. The thermal environment of the transferred-substrate process limits the operating current density of the transistor and requires the addition of a thermal shunt via underneath each device. The added parasitic capacitance to ground associated with the via does not scale easily with device dimensions. When these factors are taken into consideration, the advantages gained from substrate-transfer become much less. In Chapter 6 of this work, a mesa-HBT technology will be described that seeks to address the scaling and integration issues associated with mixed-signal and digital circuit applications.

## REFERENCES


1. Q. Lee, *Ultra-high Bandwidth Heterojunction Bipolar Transistors and Millimeter-wave Digital Integrated Circuits*, Ph.D. Dissertation, University of California Santa Barbara, June 1999.

2. H. Kroemer, "Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region," *Solid State Electronics*, vol. 28, 1985, pp. 1101-1103.

3. S. Laux, W. Lee, " Collector signal delay in the presence of velocity overshoot," *IEEE Electron Device Leters*, vol. 11, no. 4, Nov. 1990, pp. 174-176.

4. T. Ishisbashi, "Influence of electron velocity overshoot on collector transit times of HBTs," *IEEE Transactions on Electron Devices*, vol. 37, no. 9, Sept. 1990, pp. 2103-2105.

5. T. Ishibashi, "Nonequilibrium electron transport in HBTs," *IEEE Transactions on Electron Devices*, vol. 48, no. 11, Nov. 2001, pp. 2595-2605.

6. M. Littlejohn, K.W. Kim, H. Tian, "High-field transport in InGaAs and related heterostructures," in *Properties of Lattice Matched and Strained Indium Gallium Arsenide*, P. Bhattacharya, editor, INSPEC, 1991, London.

7. E.P. O'Reilly, "Band structure of InP: Overview," in *Properties of Indium Phosphide*, INSPEC, 1991, London.

8. H. Nakajima, T. Ishibashi, "Monte carlo analysis of nonequilbrium electron transport in InAlGaAs/InGaAs HBT's," *IEEE Transactions on Electron Devices*, vol. 40, no. 11, Nov. 1993, pp. 1950-1956.

9. Data found on website of Ioffe Physico-Technical Institute, *www.ioffe.rssi.ru/SVA/NSM/* St. Petersburg, Russian Federation.

10. C.T. Kirk, "A theory of transistor cutoff frequency ($f_t$) fall-off at high current density," *IEEE Transactions on Electron Devices*, ED-9, 1962, pp. 164.

11. R. Katoh, “Charge-control analysis of collector transist time for (AlGa)As/GaAs HBT’s under a high injection condition,” *IEEE Transactions on Electron Devices*, vol. 17, no. 10, Oct. 1990, pp. 2176-2182.

12. P.J. Zampardi, D.-S. Pan, “Delay of Kirk effect doe to collector current spreading in heterojunction bipolar transistors,” *IEEE Electron Device Letters*, vol. 17, no. 10, Oct. 1996, pp. 470-472.

13. D. Scott, H. Xing, S. Krishnan, M. Urteaga, N. Parthasarathy and M. Rodwell, “InAlAs/InGaAs/InP DHBTs with Polycrystalline InAs Extrinsic Emitter Regrowth,” *Conference Digest 2002 Device Research Conference*, June 2002, Santa Barbara, CA,

14. M.J.W. Rodwell, M. Urteaga, Y. Betser, D. Scott, M. Dahlstrom, S. Lee, S. Krishnan, T. Mathew. S. Jaganathan. Y. Wei, D. Mensa, J. Guthrie, R. Pullela, Q. Lee, B. Agarwal, U. Bhattacharya, S. Long “Scaling of InGaAs/InAIAs HBTs for High Speed Mixed-Signal and mm-Wave ICs”, *International Journal of High Speed Electronics and Systems*, vol. 11, no. 1, pp. 159-215.

15. M.J.W. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S.C. Martin, R.P. Smith, S. Jaganathan, S. Krishnan, R. Pullela, B. Agarawal, U. Bhattacharya, L. Samoska, M. Dahlström, “Submicron scaling of HBTs,” *IEEE Transactions on Electron Devices*, vol. 48, no. 11, Nov. 2001. pp. 2602-2624.

16. M. Vaidynathan, D. L. Pulfrey, “Extrapolated $f_{max}$ of heterojunction bipolar transistors,” *IEEE Transaction on Electron Devices*, vol. 46, no. 2, Feb. 1999.

17. A.P. Laser, D. L. Pulfrey, “Reconciliation of methods for estimating $f_{max}$ for microwave heterojunction transistors,” *IEEE Transactions on Electron Devices*, vol. 38, no. 8, August, 1991, pp. 1685-72.

18. D. Mensa, *Improved Current-Gain Cutoff Frequency and High Gain-Bandwidth Amplifiers in Transferred-Substrate HBT Technology*, Ph.D. Dissertation, University of California Santa Barbara, Sept. 1999.

19. J.E. Schramm, *Reactive Ion Etching of Indium-based Compounds Using Methane/Hydrogen/Argon*, Ph.D. Dissertation, University of California Santa Barbara, June 1995.

20. S. A. Stockman, A.W. Hanson, C.M. Colomb, M.T. Fresina, J.E. Baker, G.E. Stillman, "A Comparison of TMGa and TEGa for Low-Temperature Metalorganic Chemical Vapor Deposition Growth of $CCl_4$-Doped InGaAs," *Journal of Electronic Materials*, vol. 23, no. 8, Aug. 1994, pp. 791-800.

21. H. Ito, S. Yamahata, N. Shigekawa, K. Kurishima, "Heavily carbon doped base InP/InGaAs heterojunction bipolar transistors grown by two-step metalorganic vapor deposition," *Japanese Journal of Applied Physics*, vol. 35, no. 12A, December 1996, pp. 6139-6144.

22. Y.-J. Chen, J.M. Kuo, and B.H. Kear, "De-hydrogenation studies of carbon-doped $In_{0.53}Ga_{0.47}As$ grown by gas-source MBE and their applications to InP/ $In_{0.53}Ga_{0.47}As$ HBTs," *Proceedings Indium Phosphide and Related Materials Conference*, May 2000, pp. 37-40.

23. N. Watanabe, S. Yamahata, T. Kobayashi, "Hydrogen removal by annealing from C-doped InGaAs grown on InP by metalorganic chemical vapor deposition," *Journal of Crystal Growth*, no. 200, 1999, pp. 599-602.

24. K. Kurishima, S. Yamahata, H. Nakajima, H. Ito and Y. Ishii, "Peformance and stability of MOVPE-grown carbon-doped InP/InGaAs HBT's dehydrogenated by an anneal after emitter mesa formation," *Japanese Journal of Applied Physics*, vol. 37, no. 3B, March 1998, pp. 1353-1358.

25. T. Mathew, *High Speed Digital ICs in Transferred-Substrate HBT Technology*, Ph.D. Dissertation, University of California Santa Barbara, August 2001.

26. M. Dahlström, *Ultra High Speed InP Heterojunction Bipolar Transistors*, Ph.D. Dissertation, Royal Institute of Technology, Stockholm, Sweeden, 2003.

27. R. Driad, Z.-H, Lu, S. Laframboise, D. Scansen, W.R. McKinnon, S.P. McAlister, "Reduction of surface recombination of InGaAs/InP heterostructures using UV-irradiation and ozone," *Japanese Journal of Applied Physics*, 38, 1999, pp. 1124-1127.

28. J.S. Yu, S.H. Kim, T.I. Kim, “PtTiPtAu and PdTiPtAu ohmic contacts to p-InGaAs,”*Proceedings of IEEE Twenty-Fourth International Symposium on Compound Semiconductors*, San Diego, CA, Sept. 1997, pp. 175-178.

29. D.G. Ivey, “Platinum metals in Ohmic contacts to III-V semicontuctors,” *Platinum Metals Review*, no. 43, 1999, pp. 2-12.

30. D. Caffin, L. Bricard, J.L. Courant, L.S. How Kee Chun, B. Lescaut, A.M. Duchenois, M. Meghelli, J.L. Benchimol, P. Launay, “Passivation of InP-based HBTs for high bit rate circuit applications,” *Conference Digest1997 Indium Phosphide and Related Materials Conference*, May, Cape Cod, MA, pp. 637-640.

31. Data found on Dow Chemicals website, *http://www.dow.com/cyclotene/over.htm*.

32. Y. Wei, *Widebandwidth Power Heterojunction Bipolar Transistors and Amplifiers*, Ph.D. Dissertation, University of California Santa Barbara, March 2003.

33. J. Guthrie, *Metal Substrate Process and W-band Power Amplifier Integrated Circuits in Transferred-Substrate HBT Technology*, Ph.D. Dissertation, University of California Santa Barbara, March 2000.

# Chapter 3 High Frequency Device Measurements

Highly-scaled HBTs have extremely small reverse transmission characteristics and low shunt output conductances making device measurements and model extraction challenging. Accurate and repeatable on-wafer measurements require a well-characterized measurement environment. State-of-the-art transistor bandwidths far exceed the DC-50 GHz bandwidths covered by typical commercial vector network analyzers (VNAs). Presently, VNA test set extensions are available covering frequencies up to 325 GHz and extending on-wafer measurement methods to these frequencies presents numerous challenges. In this chapter, VNA calibration methods for on-wafer transistor measurements are discussed. A Through-Reflect-Line (TRL) calibration method using extended references planes is presented and applied to the measurements of submicron HBTs. The benefits and limitations of the TRL calibration method versus other on-wafer calibrations are considered.

## *3.1 On-wafer Vector Network Analyzer Calibration*

The goal of a two-port VNA calibration is to place the measurement reference planes precisely at the input and output of the device-under-test (DUT). In an on-wafer measurement environment, this requires removing from the measurements the effects of delays and losses associated with microwave cabling, microwave wafer probes and the on-wafer transmission line network that the DUT is embedded in. A network analyzer calibration is performed by measuring a set of defined calibration standards. From measurements of these standards, a set of error correction coefficients is

generated and applied to calibrate subsequent measurements. An analysis of the derivation and formulation of the error correction terms is beyond the scope of this work, and the reader is referred to [1, 2, 3, 4] for the mathematical details of various VNA calibrations. Here, VNA calibration issues specific to on-wafer transistor measurements are considered.

### *3.2 On-Wafer Probe-to-Probe Coupling*

A standard two-port VNA calibration involves the determination of 12 error correction terms. These terms include a port-to-port isolation term that accounts for cross-talk between the two measurements ports that is not related to the DUT. An accurate determination of the isolation error correction term requires that the unwanted cross-talk between ports does not vary when measuring each of the calibration standards or the device-under-test. In an on-wafer measurement environment, port-to-port crosstalk can be large due to radiative coupling between on-wafer probes. The radiation, and hence coupling, will depend on the impedance presented to the probes, and the conditions for an accurate isolation calibration are difficult, if not impossible, to achieve on-wafer.

Highly scaled transistors have extremely small reverse transmission characteristics, and probe-to-probe coupling that is not accounted for in a measurement calibration can easily corrupt device measurements. Probe-to-probe leakage can be accounted for using more complicated 15- or 16- term VNA error corrections that involve the measurement of further calibration standards than required for a 12-term calibration [3,4]. However, these corrections require precise

characterization of calibration standards, and such characterizations are difficult to achieve for on-wafer elements, particularly at mm-wave frequencies.

An alternative approach to deal with probe-to-probe coupling is to ensure that the magnitude of the coupling is much less than the reverse transmission characteristics of the device that is being measured. This can be accomplished by embedding the transistor on-wafer between long lengths of transmission line, thus physically separating the on-wafer probes. Probe-to-probe isolation that is at least 20 dB lower than that of the measured transistor should be sufficient for accurate device characterization [5]. In this work, a probe-to-probe separation of ~500 μm has been used for most of the presented device measurements. A chip photograph of a transistor test structure is shown in Figure 3.1. The figure shows a device fabricated in the transferred-substrate HBT technology embedded in lengths of on-wafer microstrip transmission line.

### *3.3 On-wafer versus Probe-tip Calibration*

Two approaches are commonly used for on-wafer device measurement calibration. In one approach, calibration standards are realized on a separate calibration substrate. These calibration substrates are available commercially covering frequencies up to 110 GHz and are typically fabricated using thin-film processes on alumina ($Al_2O_3$) substrates. In the second on-wafer calibration approach, custom calibration standards are realized on the active device substrate.

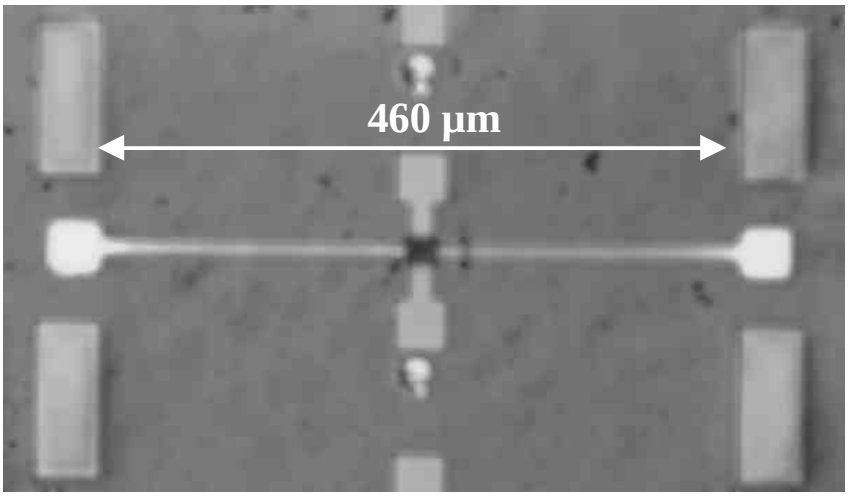



Figure 3.1: Transferred-substrate HBT embedded in on-wafer transmission line network to minimize probe-to-probe coupling during device measurements.

When using the calibration substrate approach, the calibration is designed to place the measurement references planes at the wafer-probe tips, and therefore this approach is often referred to as a "probe-tip" calibration. This calibration approach is commonly used for on-wafer device measurements and offers the advantage of having well-characterized precision calibration standards. However, there are numerous disadvantages to performing the calibration on a different substrate than the device-under-test.

As previously discussed, transistors are generally embedded on-wafer in a transmission line network, and minimizing probe-to-probe coupling requires that the physical length of the structure is quite large. If a probe-tip calibration is performed, the effect of the embedding structure must be eliminated from the transistor measurements. An approximate approach often used to account for the embedding network is to measure the capacitance of an open circuit network and subtract this

capacitance from device measurements. This approach can lead to considerable error as the value of the capacitance may be of the same order of the input capacitance of a submicron device. This approach also ignores the series resistance and series inductance of the embedding network, both of which may significantly effect device measurements.

A more precise determination of the electrical characteristics may be determined by measuring open circuit and short circuit test structures and subtracting the Y-parameters and Z-parameters, respectively, of the measured networks from the device measurements [6]. However, this approach must assume that the physical length of the embedding network is small relative to the propagation wavelength at the measurement frequency. Other approaches for accounting for the embedding network include: modeling of the network with electromagnetic simulations, or fitting measurements of the network to an equivalent lumped element circuit model.

A further drawback of the probe-tip calibration approach is that calibration substrates generally have a different transmission line environment than that used for the device-under-test. A standard VNA calibration assumes that only a single electromagnetic propagation mode exists at the calibration reference planes for both measurement and calibration [7]. The discontinuity at the probe/wafer interface does not meet these conditions and the field distribution at the discontinuity will depend on the transmission line environment that is being coupled into. As such, the probe-tip calibration realized on the calibration substrate need not apply to the substrate of

the device-under-test. The discrepancies are expected to increase at higher frequencies as the wavelength approaches the size of the probe tips [7].

The alternative to a probe-tip calibration is to realize custom calibration standards on the device substrate. Using this approach, calibration standards can be realized in the same embedding network used for the active devices and the measurement reference planes can be placed at the device terminals. The drawback to this approach is that the realization of known calibration standards on a device substrate may be challenging and involve modifications to the standard process flow. The choice of the appropriate calibration method can ease these challenges.

### *3.4 Calibration Methods*

There are several different VNA calibration routines that can be employed. The calibrations differ in the calibration standards that are measured and in the assumptions made regarding the standards. Calibration methods are generally named according to the standards that are measured, and common VNA calibrations include Short-Open-Load-Through (SOLT), Through-Reflect-Line (TRL), Line-Reflect-Match (LRM) and Line-Reflect-Reflect-Match (LRRM).

Of the calibration methods, the SOLT calibration is most commonly used for full two-port VNA calibrations, particularly in a coaxial measurement environment. The SOLT calibration requires an accurate model for all of the standards used in the calibration. The models generally include the capacitance of the Open standard due to fringing fields, the inductance of the Short standard due to a finite length ground connection, the inductance of the resistive Load standard, and the propagation delay

of the Through standard. Coaxial calibration kits can be purchased that provide accurate measurements of the model parameters. Similarly, commercially available calibration substrates provide SOLT model parameters for the purposes of probe-tip calibrations.

The effort required to obtain accurate models of calibration standards in a custom on-wafer environment may be justified for large volume manufacturing. An advantage of the SOLT calibration is that the elements are physically small, and the calibration standards can be designed to keep a fixed probe-to-probe spacing. For automated wafer probing, this feature is desirable. Techniques to model SOLT standards using electromagnetic simulations and measurements using TRL calibrations have been presented in [8,9], and these approaches have been used for device measurements to frequencies up to 110 GHz.

The Line-Reflect-Match (LRM) and Line-Reflect-Reflect-Match (LRRM) calibration methods require less knowledge of the calibration standards than required for an SOLT calibration. The additional Reflect standard measured in a LRRM eases the restriction required for a LRM calibration that two known and electrically identical Match standards be realized [10]. Additionally, the LRRM calibration extracts the inductance of the load standard, so that only the DC resistance of the standard need be known. The LRRM calibration is well suited for on-wafer device measurements, particularly in a circuit technology where thin-film resistors can be realized without the addition of a process mask step. Additionally, like the SOLT calibration the LRRM standards can be realized with a fixed probe-to-probe spacing

The LRRM calibration is generally not available in standard VNA hardware (Agilent 8510, for example), although it can be implemented in commercial VNA control software [11].

Like the LRRM calibration, the Through-Reflect-Line (TRL) calibration does not require an accurate characterization of all of the calibration standards. The calibration uses two transmission line standards one of which is designated "Through", and the other of which is designated "Line". The Line standard differs from the Through line by some electrical length $?L$. The Reflect standard may be an open or short circuit termination.

The solution for the error terms in a TRL calibration is overdetermined, and the reflection coefficient of the Reflect standard and the propagation constant of the Line standard can be determined from calibration measurements. The only parameter that must be known is the characteristic impedance ($Z_0$) of the Line standard. This characteristic impedance becomes the reference impedance for calibrated measurements, and it is important to realize that this impedance has frequency dependent real and imaginary parts. Methods for accounting for the frequency dependence of the characteristic impedance are discussed in the following section.

An often cited disadvantage of the TRL calibration is that one Line standard can only cover a 1:8 frequency span, with the ideal $?L$ being a quarter-wavelength at the center of the span. As such, multiple Line standards are required to cover larger frequency ranges, and low frequency standards can take up a large amount of valuable wafer area. Multiple Line standards may also be used to provide

measurement redundancy in a band. A multi-line TRL calibration technique can be used to reduce errors due to probe-placement repeatability [12].

For the measurement results presented in this work, standard TRL (single Line standard) calibrations were used. In the following section, specific issues related to the calibration are addressed and measurements designed to verify the calibration accuracy are presented.

## *3.5 TRL Calibration Methods and Verification*

### 3.5.1 Complex Characteristic Impedance Corrections

In the presence of conductor and substrate losses, the characteristic impedance of a TEM transmission line can be expressed in terms of its equivalent circuit parameters such that

$$Z_0 = \sqrt{\frac{R + jwL}{G + jwC}} \qquad \text{Eqn. 3.1}$$

where $R$ is the resistance per unit length, $L$ is the inductance per unit length, $G$ is the conductance per unit length, and $C$ is the capacitance per unit length of the transmission line.

As discussed in the previous section, a TRL calibration is referenced to the characteristic impedance of the Line standard, and as represented in Eqn. 3.1, the impedance will be complex in the presence of resistive or substrate losses (finite $R$ or $G$). In an on-wafer environment, line dimensions are small and resistive losses are high. Due to the skin effect, resistive losses tend to increase proportionally to the

square root of the frequency, $R \propto \sqrt{w}$ [13]. If substrate losses are ignored ($G$=0), the imaginary part of Eqn. 3.1 will tend to decrease with increasing frequency. As such, accounting for the complex characteristic impedance of the Line standard is particularly important at low frequencies.

The two most commonly used types of on-wafer transmission lines are microstrip and coplanar waveguide (CPW). Both transmission lines cannot support pure TEM modes of propagation due to the presence of inhomogeneous dielectrics. However, both transmission lines do support so-called "quasi-TEM" modes. A quasi-TEM mode has finite longitudinal components of the electric and magnetic fields, but these components are small relative to the transverse components and the mode closely resembles a TEM mode [14]. A quasi-TEM mode will exhibit a frequency dependent characteristic impedance and propagation constant because the longitudinal components of the fields tend to increase with increasing frequency. The change in $Z_0$ due to this effect in the frequency bands of interest for device measurements is not as large as the change due to resistive losses. However, the change in $Z_0$ due to mode variation can still be modeled and used to improve the accuracy of a TRL calibration.

Methods to accurately model the characteristic impedance of on-wafer transmission lines have been extensively investigated by Williams and Marks [15, 16]. As a byproduct of the TRL calibration, the propagation constant (*?*) of the Line standard is obtained. The propagation constant of a TEM transmission line is expressed in terms of its equivalent circuit parameters as

$$g = \sqrt{(R + jwL)(G + jwC)} \qquad \text{Eqn. 3.2}$$

Eqn. 3.1 and Eqn. 3.2 can be combined to express the equivalent circuit parameters in terms of $Z_0$ and ?.

$$Z_0 g = R + jwL \qquad \text{Eqn. 3.3}$$

$$Z_0 / g = G + jwC \qquad \text{Eqn. 3.4}$$

In a III-V on-wafer wiring environment, we expect the substrate losses to be small at the frequencies of interest for device modeling. If *j? C>>G*, Eqn. 3.3 and Eqn. 3.4 indicate that a measure of the phase of $Z_0$ can be obtained from a measure of the phase of ?. Additionally, if the capacitance per unit length (*C*) is known the magnitude of $Z_0$ can also be determined [15].

Marks and Williams note that the dependence of *C* on frequency and conductivity is typically weak, and *C* may be approximated by the DC capacitance ($C_0$). They further propose a transmission line capacitance measurement based on a low frequency (5MHz to 1 GHz) extraction from a TRL calibration [16]. This approach is limited in an IC design environment because at such low frequencies, the Line lengths required for a TRL calibration are excessively long.

The approach used in this work was to model the characteristic impedance of the transmission line standards using electromagnetic simulation software. To verify the simulations, the complex propagation constant determined from the simulation was compared to the extracted propagation constant from TRL calibrations. A planar method-of-moments electromagnetic (EM) simulator was used to model on-wafer

transmission lines [17]. For the measurements presented in this work, mircrostrip and CPW transmission lines were used for transferred-substrate and mesa-HBT measurements, respectively. EM simulations of both types of transmission lines have been performed.

For EM simulations, transmission line geometries and substrate layer thicknesses were estimated from mask dimensions and process conditions. S-parameter simulations were performed from 6-220 GHz, and from the simulations, the complex characteristic impedance and propagation constant of the transmission lines could be extracted. Figure 3.2 shows the real and imaginary parts of the characteristic impedance simulated for a microstrip transmission line in the transferred-substrate HBT wiring environment. As predicted from Eqn. 3.1, the magnitude of the imaginary part of $Z_0$ is seen to decrease with increasing frequency. The real part of $Z_0$ shows a significant variation at low frequencies and begins to plateau towards a value of 56 O. The transmission line was designed with an intended characteristic impedance of 50 O. However, processing variations for the microstrip line simulated in Figure 3.2 resulted in the BCB substrate thickness being slightly thicker than intended (5.3 μm versus 5.0 μm).

Transistor measurements were made using an Agilent 8510 network analyzer. S-parameter measurements were made with a TRL calibration referenced to the characteristic impedance of the Line standard. A MATLAB software program was written to take the measured S-parameters and convert them to a reference impedance of 50 O. This reference impedance was determined from

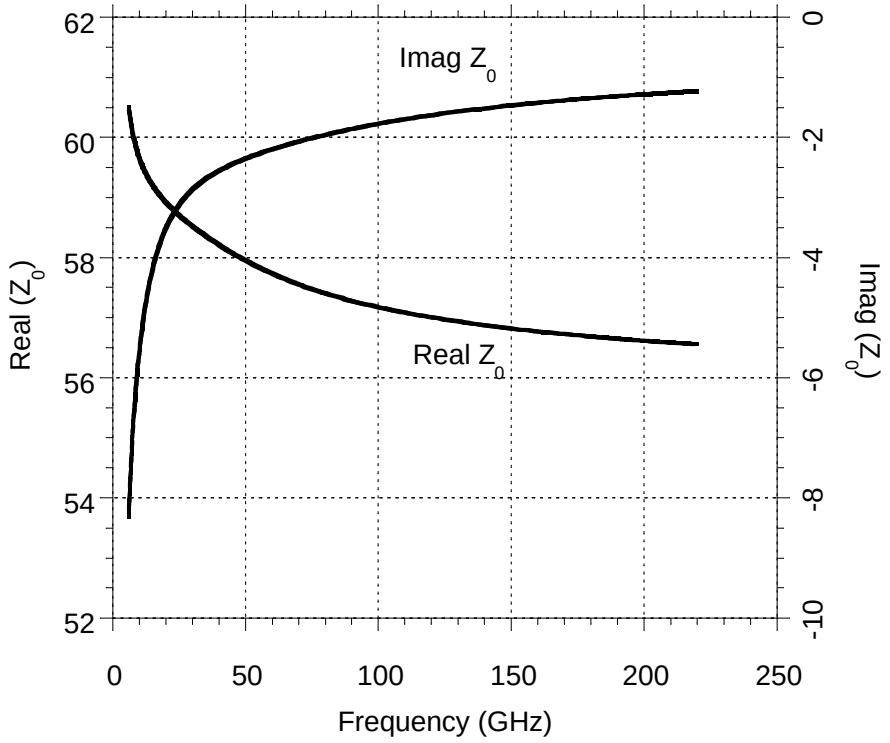



Figure 3.2: EM simulation of complex characteristic impedance ($Z_0$) for microstrip transmission line in transferred-substrate wiring environment.

EM simulations of the Line standard. The source code for the software is included in Appendix A.

As demonstrated in Figure 3.2, the variation in $Z_0$ decreases with increasing frequency. Therefore, the effects of adding the complex impedance correction to a TRL calibration are most evident at low frequencies. This is illustrated in Figure 3.3 which shows the measured $S_{11}$ and $S_{22}$ of a transferred-substrate HBT with and without the application of the complex $Z_0$ correction in the 6-40 GHz, and 75-110 GHz frequency bands. The correction at low frequencies is particularly important

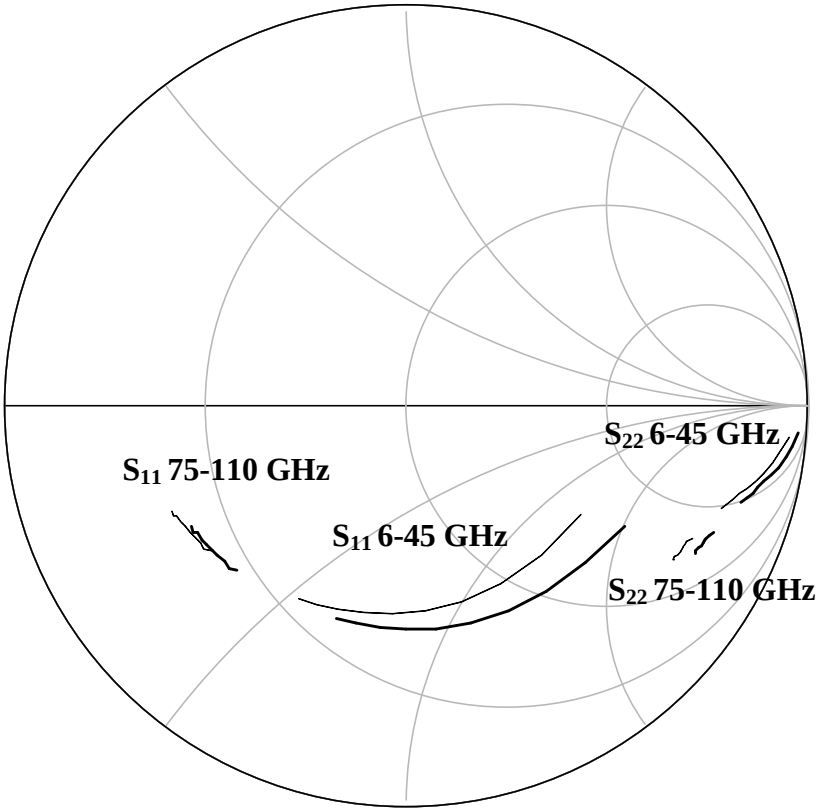



Figure 3.3: Measured S-parameters of transferred-substrate HBT with (solid line) and without (dashed line) complex impedance correction applied to TRL calibration.

since the extraction of certain transistor equivalent circuit model parameters involves the observation of low frequency asymptotic behavior.

As described above, the TRL calibration cannot be used to determine the characteristic impedance of the Line standard. However, the calibration does provide a measure of the complex propagation constant of the Line standard, and as described by Eqn. 3.3 and Eqn. 3.4, the propagation constant is related to $Z_0$ through the electrical parameters of the transmission line. Therefore, comparing the propagation constant extracted from the TRL calibration with that determined from

EM simulations of the Line standard gives an indication of the accuracy of the applied complex impedance correction. Figure 3.4 shows the real part (*a*) and imaginary part normalized to frequency (*ß*/frequency) of the complex propagation constant simulated for a microstrip transmission line in the transferred-substrate process. Also, shown on the graphs are the measured parameters extracted from TRL calibrations performed in the 6-40 GHz, 75-110 GHz, and 140-220 GHz frequency bands.

Good agreement between the measured and simulated phase term (*ß*/frequency) is observed across all three measured frequency bands. Good agreement is also seen between the measured and simulated *a* in the 6-40 GHz and 75-110 GHz bands. In the 140-220 GHz band, a large discrepancy between measurement and simulations is observed. The difference suggests an error in the amplitude measurement of the Line standard after performing the calibration. The measurement of *a* is very sensitive to small variations in the magnitude measurement. A comparison of $S_{21}$ for the measured and simulated Line standard shows <0.3 dB magnitude variation across the 140-220 GHz frequency band. Possible sources for the magnitude variation are excessive probe-to-probe coupling, or coupling to nearby circuit elements during claibration. Later in this chapter, measurements of Open and Short standards in the 140-220 GHz band will be presented that show similar magnitude variations. In the context of the TRL characteristic impedance correction, it is again noted that the importance of the correction diminishes with increasing frequency. The difference between

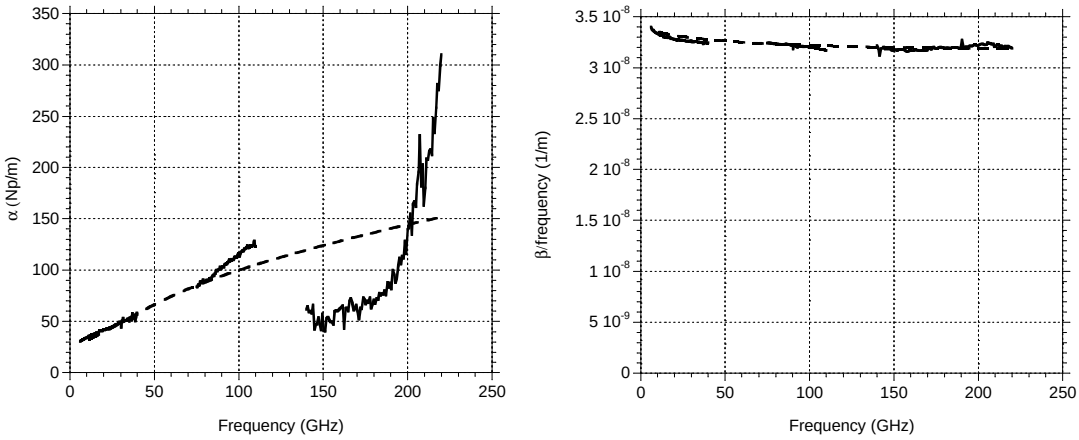



Figure 3.4: Measured (solid) and simulated (dashed) real (*a*) and imaginary parts (*ß*/frequency) of propagation constant of microstrip transmission line fabricated in transferred-substrate HBT technology.

measurements and simulations inthe 140-220 GHz band likely speaks more to the accuracy of the calibration rather than the validity of the complex impedance correction.

### 3.5.2 Higher Order Modes

The transferred-substrate process provides a thin low-loss microstrip wiring environment. The properties of the wiring environment ensure a single-mode propagation environment up to 220 GHz. Mesa-HBTs fabricated at UCSB have used coplanar waveguide (CPW) wiring, and this environment is more difficult to characterize. In addition to the desired coplanar mode, conductor backed CPW can also support slotline and microstrip modes of propagation. A schematic showing the excitation of the different CPW modes is shown in Figure 3.5.

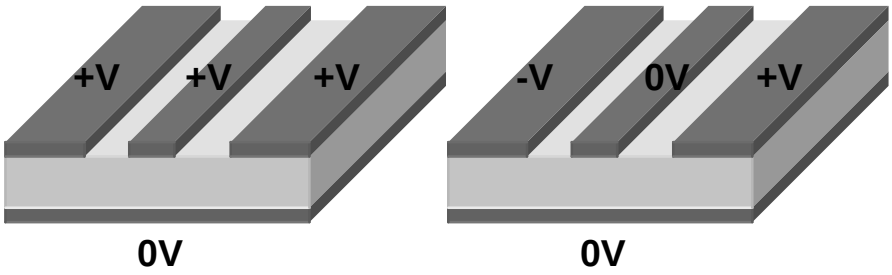



Figure 3.5: Excitation of parasitic microstrip and slotline modes in CPW transmission line.

In addition to the aforementioned modes, a dielectric slab substrate mode can also be supported. The substrate mode does not have a low frequency cutoff, and low frequency leakage into the substrate mode can be modeled as substrate loss. More seriously at higher frequencies, a mode matching condition may be reached ($?_{CPW} = ?_{substrate}$, $v_{CPW} = v_{substrate}$) between the CPW and substrate modes. Under this condition, excessive energy exchange between the two modes will occur. Note that the substrate modes are also present in a microstrip environment; however, the synchronous coupling condition is a function of the substrate height and the frequency at which it occurs should be outside of the measurement frequency range for transferred-substrate devices.

A VNA calibration assumes that only a single-propagation mode exists at the measurement reference planes. An accurate TRL calibration in a CPW wiring environment requires the suppression of the higher order modes inherent to the

structure. The slotline mode can be suppressed with the addition of periodic grounding straps connecting the two ground planes. These straps are generally realized using an airbridge process. Propagation of the microstrip mode requires a conductor backing, and this mode can be suppressed by mounting the wafer on a dielectric (microwave absorber, or Teflon) rather than on a metal probe station chuck. Substrate modes can also attenuated with the addition of microwave absorber underneath the substrate. Additionally, since the mode matching condition is related to the substrate height, the wafer can be thinned to increase the synchronous coupling frequency. A conservative guideline to avoid coupling to substrate modes is to ensure $h < 0.12\ ?_d$, where $?_d$ is the wavelength in the dielectric [18]. For a standard 500 μm thick InP substrate, the guideline suggests that coupling to substrate modes may occur for frequencies as low as 20 GHz.

Mesa-HBT device measurements made in this work utilized 500 μm InP substrates with the wafers placed on microwave absorber during measurements The measurement structures did not use ground plane strapping. Figure 3.6 shows the real and imaginary parts of the CPW line propagation constant extracted from a TRL calibration. Also shown in the figure are the results from EM simulations of the transmission line. Measurements in the 6-40 GHz band appear well behaved and match simulations closely. Large discrepancies are observed over the 75-110 GHz band an indication that higher order mode coupling may be present. Device measurements in the 75-110 GHz also show artifacts that are consistent with a poor calibration.

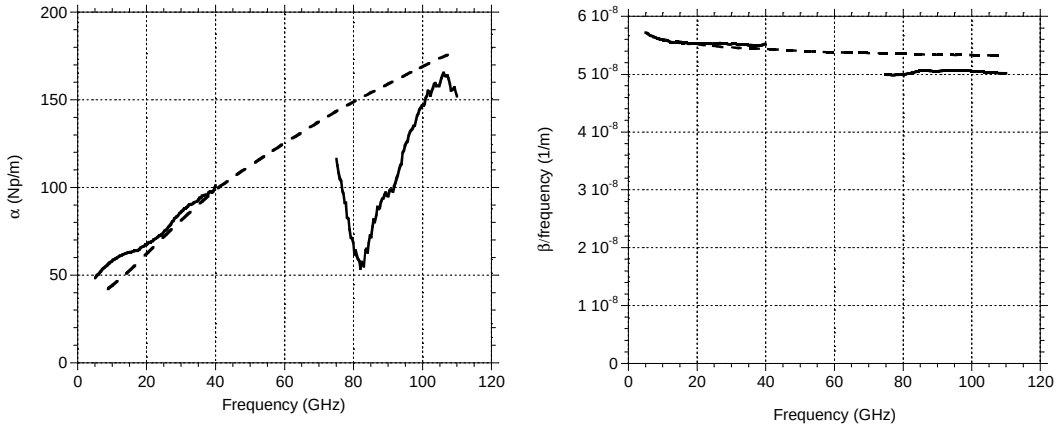



Figure 3.6: Measured (solid) and simulated (dashed) real (*a*) and imaginary parts (*ß*/frequency) of propagation constant of coplanar waveguide transmission line fabricated in mesa-HBT technology.

### 3.5.3 Calibration Verification

Quantitatively assessing the accuracy of a VNA calibration is difficult, particularly in an on-wafer environment where the realization of standards with known electrical characteristics is challenging. The measurement of the TRL calibration Line standard and the comparison with EM simulations as described in the previous section provides qualitative evidence that a good calibration has been achieved. As shown in the previous sections, this measurement has been used to indicate problems with the calibration in the 140-220 GHz band for a transferred-substrate microstrip calibration and in the 75-110 GHz band for a mesa-HBT CPW calibration. Additionally, the open and short standards can be measured after a calibration. Since a TRL calibration does not specify the reflection coefficient of the

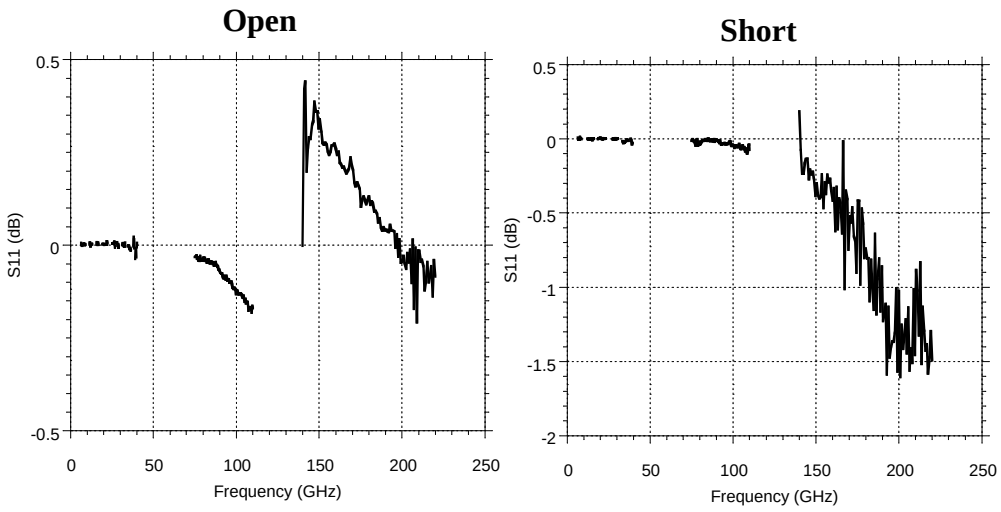



Figure 3.7: Measured reflection coefficient of open and short calibration standards after TRL calibration in transferred-substrate microstrip environment.

Reflect standard, measurements of these structures provides evidence as to whether a good calibration has been achieved. With the low parasitic wiring environment provided in the transferred-substrate process, it is expected that the characteristics of the open and short standards closely resemble those of ideal circuit elements. Figure 3.7 shows the magnitude of the reflection coefficient for the short and open standards measured in the transferred-substrate wiring environment. The measurements appear well-behaved in the 6-40 GHz and 75-110 GHz frequency bands but show a large variation in the 140-220 GHz band. This is consistent with the measurement of the propagation constant presented earlier.

In Chapter 5, amplifier measurements are presented in the 140-220 GHz frequency band. For these measurements, small variations in amplitude will not have a large effect on the amplifier characteristics, and the calibration appears well suited for these applications. For submicron device measurements, a greater level of

calibration accuracy is required, particularly in the measurement of the reverse transmission coefficient $S_{12}$. In the following chapter, device measurements are presented that show a large deviation in $S_{12}$ from trends established in the lower frequency bands. For these reasons, device measurements in the 140-220 GHz band must be treated with caution. They are presented for completeness but will not be used to make specific conclusions regarding transistor behavior.

## REFERENCES


1. G.F. Engen and C.A. Hoer, “Thru-reflect-line: An improved technique for calibrating the dual automatic network analyzer,” *IEEE Transaction on Microwave Theory and Techniques*, vol. 27, Dec. 1979, pp. 987-993.

2. F. Purroy, L, Pradell, “New theoretical analysis of LRRM calibration technique for vector network analyzers,” *IEEE Transaction on Instrumentation and Measurement*, vol. 50, no. 5, Oct. 2001, pp. 1307-1314.

3. H. Heuermann, B. Schiek, “15-term self-calibration methods for error-correction of on-wafer measurements,” *IEEE Transaction on Instrumentation and Measurement*, vol. 46, no. 5, Oct. 1997, pp. 1105-1110.

4. J.V. Butler, D.K. Rytting, M.F. Iskander, R.D. Pollard, M. Vanden Bossche, “16-term error model and calibration procedure for on-wafer network analysis measurements,” *IEEE Transaction on Microwave Theory and Techniques*, vol. 39, no. 12, Dec. 1991, pp. 2211-2217.

5. J. Pla, W. Struble, F. Colomb, “On-wafer calibration techniques for measurement of microwave circuit and devices on thin substrates,” *1995 IEEE MTT Symposium Digest*, Orlando, FL, May, pp. 1045-1048.

6. M.C. Koolen, J.A. Geelen, M.P. Versleijen, “An improved de-embedding technique for on-wafer high-frequency characterization,” *1991 IEEE Bipolar Circuits and Technology Meeting*, pp. 188-191.

7. D.F. Williams, R.B. Marks, “Calibrating on-wafer probes to the probe tips,” *40th ARFTG Conference Digest*, Dec. 1992, pp. 136-143.

8. M. DuFault, A.K. Sharma, “A novel calibration verification procedure for millimeter-wave measurements,” *1996 IEEE MTT Symposium Digest*, pp. 1391-1394.

9. W.M. Okamura, M. DuFault, A.K. Sharma, “A comprehensive millimeter-wave calibration development and verification approach,” *2000 IEEE MTT Symposium Digest*, pp. 1477- 1480.

10. F. Purroy, L. Pradell, “New theoretical analysis of the LRRM calibration technique for vector network analyzers,” *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 5, Oct. 2001, pp. 1307-1314.

11. WinCal TM, Calibration Software for Vector Network Analyzers, Cascade MicroTech, Beaverton, Oregon.

12. R.B. Marks, “A multiline method of network analyzer calibration,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, no. 7, July 1991, pp. 1205-1215.

13. D. Pozar, *Microwave Engineering*, Addison-Wesley Publishing Company, 1993, pp. 26.

14. R.K. Hoffman, *Handbook of Microwave Integrated Circuits*, Artech House , Norwood, MA, 1987, Chapter 2.

15. R.B. Marks, D.F. Williams, “Characteristic impedance determination using propagation constant measurement,” *IEEE Microwave and Guided Wave Letters*, vol. 1, no. 6, June 1991, pp. 141-143.

16. D.F. Williams, R.B. Marks, “Transmission line capacitance measurement,” *IEEE Microwave and Guided Wave Letters*, vol. 1, no 9, Sept. 1991, pp. 243-245.

17. *Agilent Momentum Software*, Palo Alto, CA: Agilent Technologies.

18. M. Riaziat, R. Majidi-Ahy, I-J Feng, “Propagation modes and dispersion characteristics of coplanar waveguide,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 38, no. 3, March 1990, pp. 245-251.

# Chapter 4 Power Gain Singularities in Transferred-Substrate HBTs

The transferred-substrate process permits submicron scaling of the emitter-base and collector-base junctions resulting in a reduction in the base resistance and collector-base capacitance and thereby increasing the transistor power gain cutoff frequency $f_{max}$. Transferred-substrate HBTs fabricated at UCSB have been reported with progressively decreasing junction dimension, as a result of which progressively increasing high frequency power gains have been observed [1, 2, 3].

The observed trend of increasing device gain with scaling is in fact more rapid than the simple geometric scaling theory presented in Chapter 2 predicts. In [3], a deep submicron transferred-substrate HBT was reported with a measured 20 dB unilateral power gain at 100 GHz. A -20 dB/decade extrapolation of the measured transistor power gain predicts a $f_{max}$ of ~1 THz, although it was noted in [3] that such a large extrapolation must be treated with caution. Based on the device geometry and the measured base sheet and contact resistivities, the scaling theory in Chapter 2 predicts a transistor power gain cutoff frequency of 420 GHz.

The observed high-frequency gain of these devices is in part due to a substantial (and experimentally observed) rapid decrease in the collector-base capacitance with increasing collector bias current. A reduction in collector-base capacitance with increased current due to electron velocity modulation in the collector depletion region was predicted by Moll and Camnitz [4], and further

investigated by Betser and Ritter [5] and is one possible cause of the observed capacitance variation.

In this chapter, further measurements of submicron transferred-substrate HBTs are presented. For these devices, peaking, and in some cases singularities, are observed in the high-frequency unilateral power gain ($U$). In some devices, $U$ is unbounded over a full 20-110 GHz bandwidth. Associated with these high observed power gains are a rapid decrease in collector-base capacitance with increasing bias current, and trends towards negative conductance in the device output admittance (common-emitter $Y_{22}$) and a positive conductance in the reverse transmission (common-emitter $Y_{12}$). Trends that cannot be predicted with standard HBT circuit models. A consequence of the observation of unbounded $U$ is that $f_{max}$ of the devices cannot be extrapolated from measurements.

An extended version of the Moll/Camnitz velocity modulation model may explain the observed high frequency characteristics. In this model, the high frequency dynamics of the electron redistribution in the collector region are modeled, and it will be shown that electron velocity modulation in the collector produces both a frequency-dependent reduction in the collector base capacitance and a high-frequency negative conductance between collector and base. The developed model is added to small-signal HBT circuit model, and circuit simulations are compared to the measured device results.

## *4.1 Capacitance Cancellation by Electron Velocity Modulation*

Device measurements presented in the following section will show evidence of negative resistance effects in the output and feedback conductance of a transferred-substrate HBT. At the bias conditions where these trends are observed, a significant reduction in the collector-base capacitance from its DC value is also observed. The reduction in collector-base capacitance observed in InP-based HBTs with increasing collector current has been attributed to the modulation of electron velocity in the collector space charge region [4,5]. The derivations of [4,5] were performed assuming that the collector space charge could respond instantaneously to changes in the collector-base voltage. In this section, a model to describe the redistribution of the electrons in the collector space charge region is presented. In analyzing the dynamics of the problem, it is found that that in addition to the predicted reduction in collector-base capacitance, negative resistance terms are also observed in the base-collector admittance.

### 4.1.1 Capacitance Cancellation Theory

In advanced III-V HBTs, electrons entering the collector space charge region experience ballistic transport and may travel a significant fraction of the collector at a higher velocity than the saturated electron velocity of the bulk semiconductor [6]. The electric field profile in the collector influences the electron velocity profile since the kinetic energy of the electrons determines the scattering probability to lower effective velocity satellite conduction band valleys. At higher applied collector-base voltages, InP HBTs will exhibit larger collector transit times due to a lower effective

velocity. Modulation of the collector velocity profile changes the collector space charge profile due to the redistribution of mobile electrons. Capacitance cancellation arises because modulation of mobile charge in the collector screens the base and collector terminals from changes in the electric field.

The dynamics of capacitance cancellation are considered under the assumption of a collector electron velocity that is a function of $V_{CB}$ but does not vary with position within the collector. Having in the preceding discussion considered the importance of ballistic transport effects in the collector, it is recognized that the resulting analysis is therefore only approximate. However, the goal of this work is to present a theory as to the origin of negative resistance effects in InGaAs-collector HBTs through velocity modulation and not to develop an exact model.

The dynamics of capacitance cancellation will be analyzed in the time domain. Consider an HBT operating with a DC collector-base voltage $V_{CB,0}$ and a DC collector current $I_{C,0}$. For the given bias conditions, electrons in the space-charge region travel with a velocity $v_o$. To simplify further calculations, an inverse velocity $s_o = 1/v_o$ is also defined.

The collector region has a thickness $W_C$ and is uniformly doped at a concentration of $N_D$. It is assumed that the collector region is fully depleted at the applied DC bias conditions, such that $W_C$ is not modulated by small changes in the applied collector base voltage. To further simplify the analysis, it is also assumed that the base and subcollector regions are heavily doped and undepleted. Under these

assumptions, the entire collector-base voltage is dropped across the collector space charge region. The total charge density in the region is given by

$$n(x) = qN_D - J_{CB}s_o \qquad \text{Eqn. 4.1}$$

where $J_{CB} = I_{C,0} / A_E$ is the current density entering the collector region from the base.

The redistribution of electron charge in the collector region is considered in response to a collector voltage step $?V_{CB}$ applied at $t = 0$. The change in electron velocity caused by the applied $?V_{CB}$ is assumed to occur instantaneously and uniformly across the collector space charge region. The inverse velocity after application of the voltage step is given by

$$s(x) = s_0 + \Delta s \qquad \text{Eqn. 4.2}$$

where $\Delta s = \Delta V_{CB}\left(\partial s / \partial V_{CB}\right)$. Note that an increase in the collector-base voltage resulting in a decrease in the electron velocity corresponds to an increase in the inverse velocity.

Although the electron velocity is assumed to change instantaneously the electron charge in the collector cannot be instantaneously redistributed, and at $t = 0^+$ the charge distribution is still described by the unperturbed electron velocity (Eqn. 4.1). The redistribution of the electron space charge is assumed to take place as a moving charge front as illustrated in Figure 4.1. The charge front enters at the

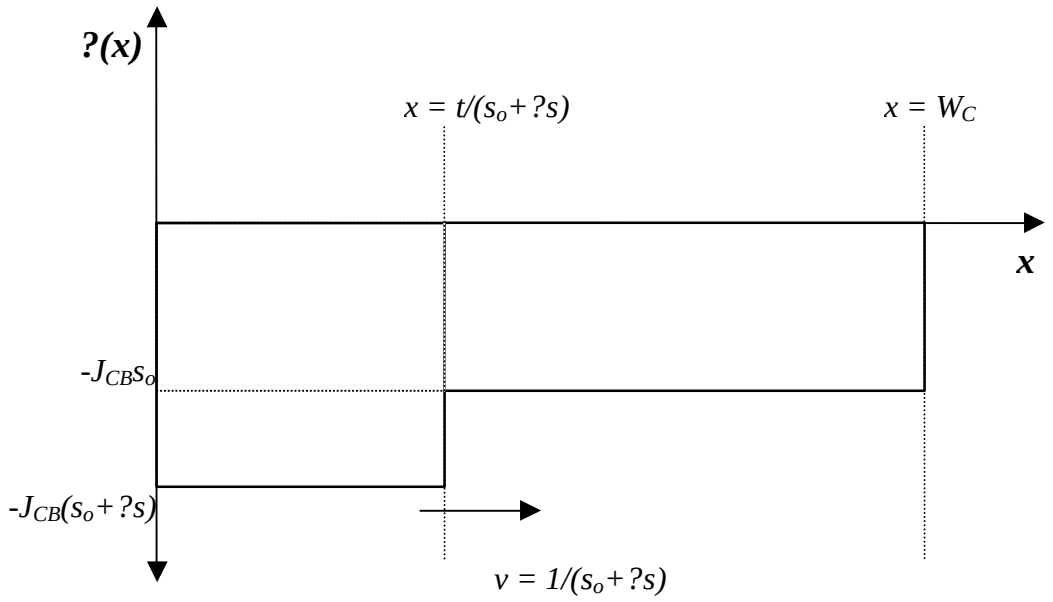



Figure 4.1: Electron space charge density in collector region at time *t* after application of collector-base voltage step.

collector-base junction and travels through the collector at the perturbed electron velocity, such that the charge distribution at time *t* is given by

$$n(x) = \begin{cases} qN_D - J_{CB}(s_o + \Delta s) & \text{for } 0 \le x < t/(s_o + \Delta s) \\ qN_D - J_{CB}s_o & \text{for } t/(s_o + \Delta s) \le x \le W_C \end{cases} \qquad \text{Eqn. 4.3}$$

and at $t = W_C(s_0 + \Delta s)$, the collector charge density has reached its uniform steady-state $n(x) = qN_D - J_{CB}(s_o + \Delta s)$.

Modeling the electron charge redistribution as a uniform moving charge front assumes that the density of electrons does not in itself perturb the electron velocity. This assumption may not hold at high current densities, where the density of electrons is comparable to the collector donor charge. Extending calculations to

include the effects of current modulation on the electron velocity may be necessary to more accurately model the collector-base admittance at high current densities. However, such a calculation is beyond the scope of this work.

The time evolution of the collector space charge distribution can be used to determine the time dependence of the collector current. The mobile electron charge in the collector region can be viewed as a collection of traveling sheets of charge. The displacement current generated at the collector terminal from a sheet of charge with density $?_s$ traveling at velocity $v$ through the depleted collector is given by

$$J_d = \frac{-r_s v}{W_C} \qquad \text{Eqn. 4.4}$$

where $J_d$ is defined to be a positive current if flowing into the collector terminal.

At $t = 0^+$, the electron charge in the collector is uniformly distributed, and the region can be divided into sheets of charge of thickness ? x and charge density $r_s = -J_{CB} s_o \Delta x$. It is again noted that at this time the electrons are assumed to be traveling at the perturbed electron velocity while the unperturbed electron velocity is used to describe the charge distribution. The total collector current at $t = 0^+$ can be determined by summing the displacement current contributions from all of the sheets of charge in the collector region while taking $\Delta x \to 0$. The change in collector current is then given by

$$\Delta J_C\big|_{t=0^+} = \int_0^{W_C} \frac{J_{CB} s_o}{W_C (s_o + \Delta s)} dx - J_{CB} = \frac{J_{CB} s_o}{(s_o + \Delta s)} - J_{CB} \approx \frac{-J_{CB} \Delta s}{s_o} \qquad \text{Eqn. 4.5}$$

Eqn. 4.5 shows that immediately after the application of the collector-base voltage step the collector current has been reduced from its initial value $J_{CB}$, indicating a trend toward negative conductance. It is important to note that $J_{CB}$, the electron current entering the collector region, will not vary with the applied $?V_{CB}$. The base region in III-V HBTs is highly doped and modulation of the collector-base voltage will not cause a significant change in the base width, and hence, the electron current entering the collector will stay constant. This conclusion is supported by lack of Early effects in typical III-V HBTs.

Given the time-dependent charge density described by Eqn. 4.3, the change in current at the collector terminal can be determined using the same formalism used to derive Eqn. 4.5

$$\Delta J_C(t) = \begin{cases} \dfrac{-J_{CB}\Delta s}{s_o} + \dfrac{J_{CB}\Delta s}{W_C s_o^2} t & \text{for } 0 < t \le W_C(s_o + \Delta s) \\ 0 & \text{for } W_C(s_o + \Delta s) < t \end{cases} \qquad \text{Eqn. 4.6}$$

Eqn. 4.6 shows that the collector current initially drops and then linearly increases back to its DC value $J_{CB}$, which it reaches at $t = W_C(s_0 + \Delta s)$ when the electron charge front has filled the collector region. The time dependent waveforms of the collector-base voltage and collector current are shown in Figure 4.2.

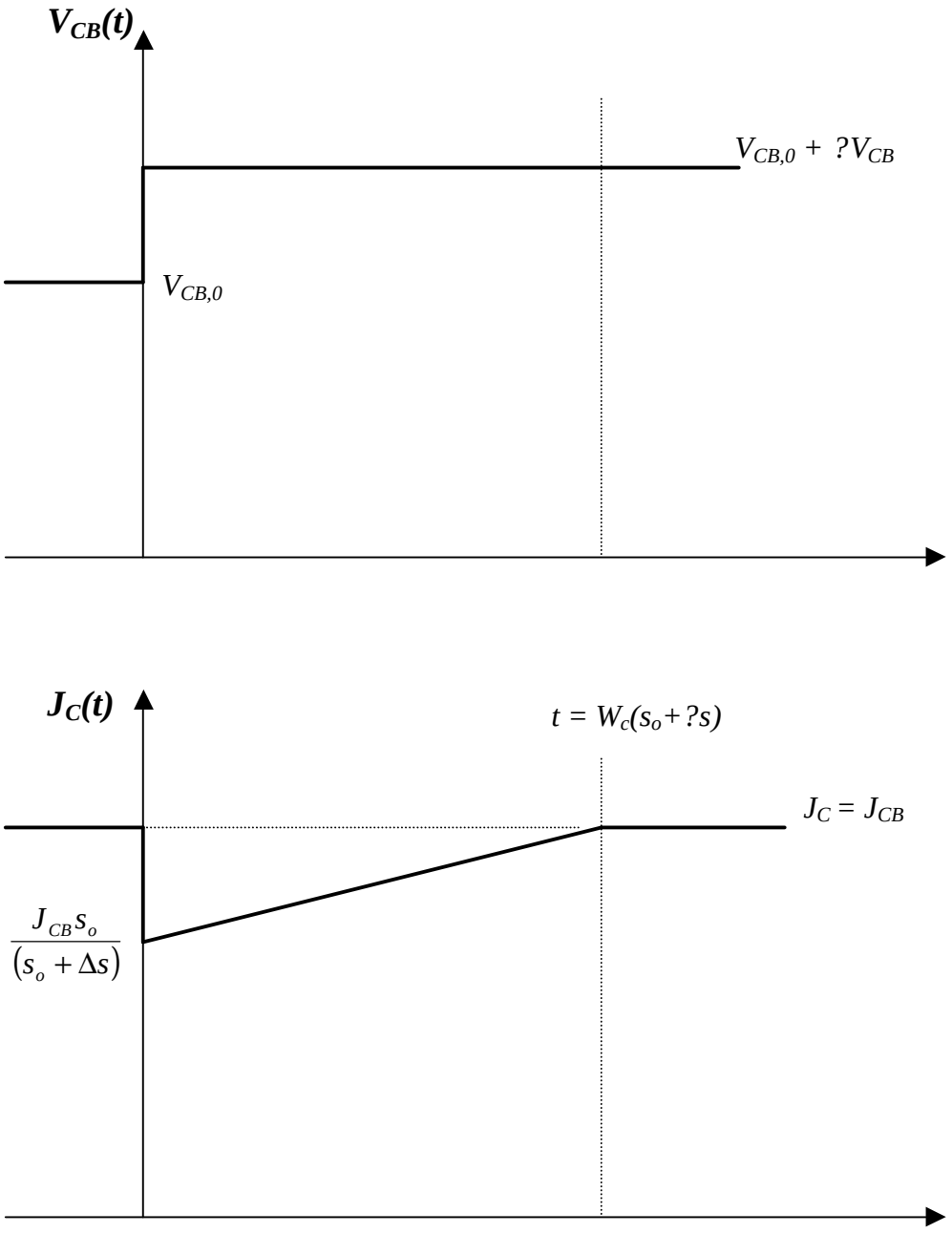



Figure 4.2: Time dependence of collector-base voltage and collector current for small step signal applied to collector-base voltage.

The time dependence of the collector current has been calculated for a step change in the applied collector-base voltage. Using Fourier techniques, the step response can be used to calculate the frequency response of the collector current to an applied collector-base voltage, and an effective collector-base admittance $Y_{cb} = \partial I_c / \partial V_{CB}\big|_{I_E}$ can be determined. Consider a voltage impulse applied at time *t=0* described by $\Delta V_{CB}(t) = (v_{CB}\Delta t)d(t)$, where $v_{CB}$ is the magnitude of the applied impulse and *?t* is the duration of the impulse.

The impulse response of the change in collector current can be determined from the derivative of the step response given by Eqn. 4.6. The impulse response of the collector current is then given by

$$\Delta J_C(t) = \frac{-J_{CB}v_{CB}\Delta t}{s_0}\frac{\partial s}{\partial V_{CB}}d(t) + \frac{J_{CB}v_{CB}\Delta t}{s_o^2 W_C}\frac{\partial s}{\partial V_{CB}}rect\left(\frac{t}{Ws_o} - 1/2\right) \qquad \text{Eqn. 4.7}$$

This expression can be rewritten to include the collector transit time. Under the assumption of an electron velocity that does not vary with position in the collector, the collector transit time is given by $t_c = W_C s/2$ and $\frac{\partial t_c}{\partial V_{CB}} = \frac{\partial}{\partial V_{CB}}(W_C s/2) = \frac{W_C}{2}\frac{\partial s}{\partial V_{CB}}$. Substituting these expressions into Eqn. 4.7 gives

$$\Delta J_C(t) = \frac{-J_{CB}v_{CB}\Delta t}{t_c}\frac{\partial t_c}{\partial V_{CB}}d(t) + \frac{J_{CB}v_{CB}\Delta t}{2t_c^{\,2}}\frac{\partial t_c}{\partial V_{CB}}rect\left(\frac{t}{2t_c} - 1/2\right) \qquad \text{Eqn. 4.8}$$

The small-signal collector base admittance is found by taking the Fourier transform of the impulse response

$$Y_{CB}(w) = \frac{\Delta I_C(w)}{\Delta V_{CB}(w)} = \frac{\Im[A_E \Delta J_C(t)]}{\Im[\Delta V_{CB}(t)]} = \frac{-I_{C,0}}{t_c} \frac{\partial t_c}{\partial V_{CB}} \left( 1 - e^{-jwt_c} \frac{\sin wt_c}{wt_c} \right) \qquad \text{Eqn. 4.9}$$

Note that this derivation has been performed under the condition of a constant current $J_{CB}$ entering the collector depletion regions, as is necessary for deriving the small signal base-collector admittance. Expanding Eqn. 4.9 to second order in frequency gives

$$Y_{cb} = -jwI_{E,0} \frac{\partial t_c}{\partial V_{cb}} \left( 1 - jw \frac{2t_c}{3} \right) + \mathrm{O}(w^3) + \ldots \approx \frac{-jwC_{cb,canc}}{1 + jw \frac{2t_c}{3}} \qquad \text{Eqn. 4.10}$$

Eqn. 4.10 represents transfer function of an equivalent circuit network that can be used to describe $Y_{cb}$ at low frequencies. This network is described by a negative capacitance of magnitude $C_{cb,canc} = I_{C,0} \partial t_c / \partial V_{CB}$ in series with a negative resistance of magnitude $2t_c / (3C_{cb,canc})$. The network appears in parallel with the dielectric capacitance of the intrinsic collector-base junction. Eqn. 4.10 is useful for modeling the effects of velocity modulation at low frequencies, and provides physical insight in describing these effects. However, for the simulations presented in the remainder of this work, the full frequency dependent expression for $Y_{cb}$ (Eqn. 4.9) is used.

To better illustrate the frequency dependence of the base-collector admittance, Figure 4.3 plots the equivalent collector-base capacitance ($C_{cb,eff}$ = *Im(*$Y_{cb}$*)/?* ) and real part of $Y_{cb}$ calculated from Eqn. 4.9. The data is plotted to 300 GHz and assumes a low frequency capacitance cancellation of 2 fF and a collector

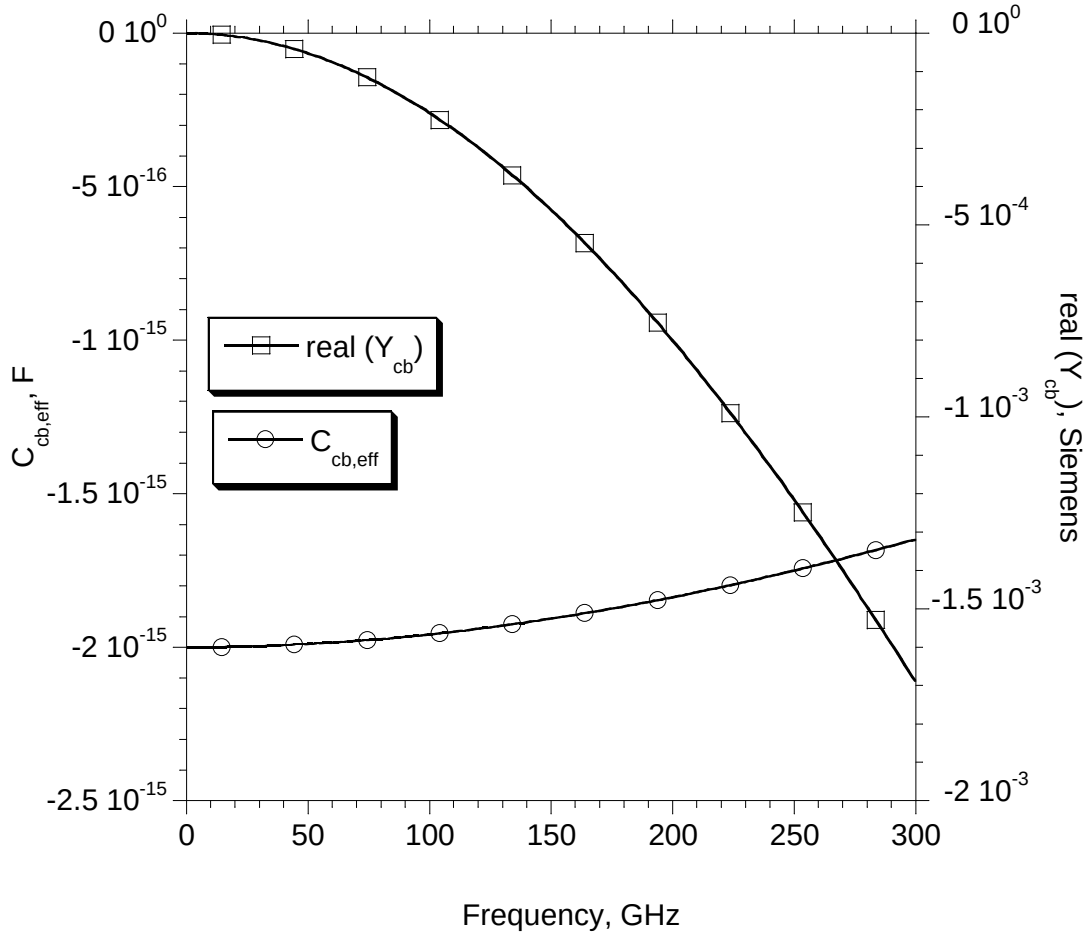



Figure 4.3: Effective collector-base capacitance ($C_{cb,eff}$) and real part of $Y_{cb}$ determined from Eqn. 4.9 assuming $t_c$ = 0.5 psec and $C_{cb,canc}$ = 2 fF.

transit time of 0.5 psec. It is seen that with increasing frequency the effective negative collector-base capacitance decreases slightly, while the real part of $Y_{cb}$ is observed to become increasingly negative. Later it will be shown that the trend towards negative conductance in $Y_{cb}$ will be consistent with device measurements of transferred-substrate HBTs.

A static derivation of capacitance cancellation through electron velocity modulation was proposed by Moll and Camnitz [4] and more fully developed by Betser and Ritter [5]. In the static derivation, the intrinsic collector-base capacitance of the HBT is given by

$$C_{cb} = \frac{eA_E}{W_C} - I_C \left. \frac{\partial t_C}{\partial V_{CB}} \right|_{W_C} \qquad \text{Eqn. 4.11}$$

In this expression, a term related to the base-collector output conductance has been ignored, which is shown by Betser and Ritter to be insignificant for practical HBTs [5]. Eqn. 4.10 and Eqn. 4.11 show that as expected the results of the dynamic analysis and static analysis are consistent to first order in frequency.

The static derivation of capacitance cancellation is performed using the charge control approximation that assumes the electron distribution in the collector can change instantaneously in response to changes in the base-collector voltage. Clearly, as has been described here, this is not the case. Electrons must enter the collector from the base and travel with a finite velocity. Therefore, one expects the derivation of Eqn. 4.11 to begin to fail at frequencies approaching the inverse of the collector transit time. The dynamic derivation that has been performed shows that the degree of capacitance cancellation does indeed decrease with increasing frequency, and more importantly, that negative resistance effects due to electron velocity modulation may be significant at relatively low frequencies.

## 4.2 Device Measurements

The measurements of one specific transistor will be presented to demonstrate the power gain singularities and negative resistance effects that have been observed in some transferred-substrate devices. The device was fabricated with emitter junction dimensions of 0.3 x 18 $\mu m^2$, and collector stripe dimensions of 0.7 x 18.4 $\mu m^2$. The transistor layer structure is the same as that shown in Table 2.1 with one exception. The exception being that the delta doping pulse in the collector was doped at 1 x $10^{17}$ $cm^{-3}$ instead of the desired 1 x $10^{18}$ $cm^{-3}$.

S-parameter measurements were made in 6-45 GHz, 75-110 GHz, and 140-220 GHz frequency bands. The measurement and calibration methods described in Chapter 3 were used to measure the devices. The transistors were embedded in a microstrip transmission line network that provided a on-wafer probe-to-probe separation of 460 μm, and a TRL calibration was used to move the measurement reference planes to the device terminals. The characteristic impedance correction described in Chapter 3 was applied to all measurements.

Figure 4.4 shows the unilateral power gain ($U$), the short circuit current gain ($h_{21}$), and the maximum stable gain (the device is not unconditionally stable over any portion of the measured frequency range) of the device measured at a bias condition of $V_{CE}$ = 1.1V and $I_C$ = 5mA. The unilateral power gain is observed to increase becoming negative at ~20 GHz, and remaining negative across the entire 75-110 GHz band.

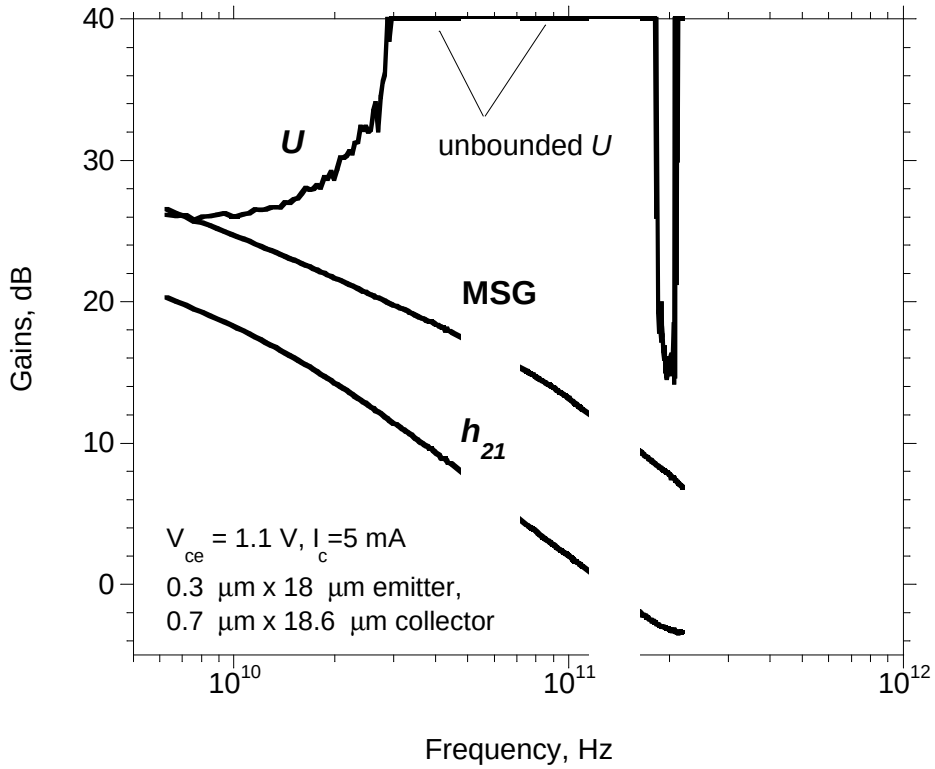



Figure 4.4: Measured unilateral power gain (*U*), short circuit current gain ($h_{21}$) and maximum available gain (*MAG*) of submicron transferred-substrate HBT.

The cause and consequences of a negative unilateral power gain can be better understood by considering the expression for Mason's unilateral power gain in terms of the transistors Y-parameters [7]

$$U = \frac{\left|Y_{21} - Y_{12}\right|^2}{4\left(G_{11}G_{22} - G_{21}G_{12}\right)} \qquad \text{Eqn. 4.12}$$

where $G_{11}$, $G_{12}$, $G_{21}$, and $G_{22}$ are the real parts of the networks Y-parameters.

Eqn. 4.12 shows that a negative unilateral power gain may be obtained in the presence of a negative output conductance ($G_{22}$) or a positive feedback component ($G_{12}$). Both of these trends have been observed in transistor measurements over

portions of the frequency band where negative unilateral power gain is observed. Later in this section, bias dependent measurements of $G_{12}$ and $G_{22}$ are presented.

From Eqn. 4.12 one also sees that if the unilateral power gain of a network is negative, an appropriate positive shunt conductance may be added to the input or output port of the network such that the denominator of Eqn. 4.12 goes to zero. Therefore, negative $U$ is equivalent to unbounded (infinite) power gain. The unilateral power gain of a network represents the gain available if the network is unilateralized (reverse transmission $S_{12} = 0$) using lossless shunt and series feedback. For a network with a negative unilateral power gain, one expects the unilateralized network to exhibit a negative input or output impedance thus being able to support one-port oscillations. In practice, a unilateralized transistor cannot be realized in an on-wafer environment at microwave frequencies. In addition to the losses inherent in on-wafer reactive elements, any attempt to unilaterize the device at a particular frequency will likely make the transistor highly unstable at other frequencies.

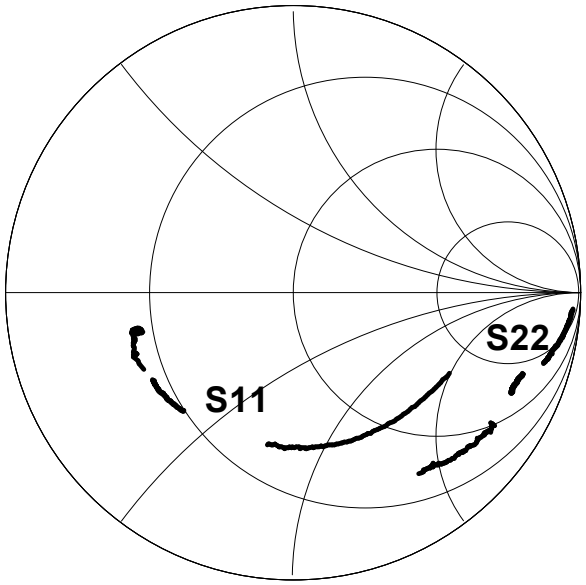
Given the practical limitations of realizing a unilateral transistor, one may question why the unilateral power gain is of any concern for transistor measurements. The utility of $U$ is found in its ability to predict the maximum frequency of oscillation of a transistor. For an HBT well described by a hybrid-p model, the unilateral power gain will show a well-behaved -20dB/decade roll-off independent of the reactances of the on-wafer embedding network and the transistor configuration (i.e. common-emitter versus common-base). This behavior motivates the use of $U$ to extrapolate the $f_{max}$ of a transistor. In contrast, the maximum stable

gain and maximum available gain of an HBT do not show prescribed roll-off characteristics. Clearly, the measurement of *U* in Figure 4.4 does not show a well-behaved -20dB/decade roll-off, and the negative resistance effects observed in measurements of $G_{22}$ and $G_{12}$ for the submicron device are not predicted by a standard hybrid- p circuit model.

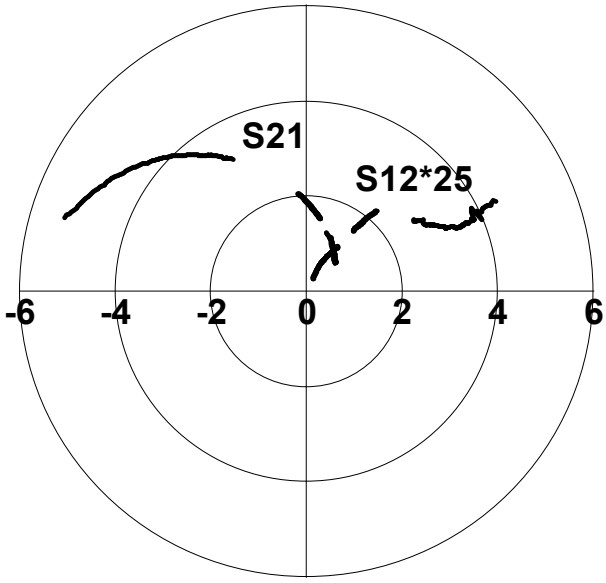
For completeness, the measured device S-parameters across all three of the measured frequency bands are presented in Figure 4.5. Note that while $S_{11}$, $S_{22}$, and $S_{21}$ show a relatively smooth variation across the three measurement bands, $S_{12}$ in the 140-220 GHz range appears to deviate significantly from the trajectory of the measurements in the lower frequency bands. The deviation of $S_{12}$ supports the suspicion that the measurement may be corrupted from excessive on-wafer probe-to-probe coupling as discussed further in Chapter 3.

The singularity in the transistor power gain is observed as the transistor bias current density is increased. For the device presented here, a singularity is measured at current densities as low as 0.55 mA/μm$^2$. At relatively low current densities, a significant reduction in the effective collector-base capacitance of the transistor is observed. For an intrinsic HBT described by a Tee-model with zero collector series resistance and zero extrinsic collector-base capacitance, the collector-base admittance $Y_{CB}$ is described in terms of the network Z-parameters as

$$Y_{CB} = \frac{1}{Z_{22} - Z_{21}} \qquad \text{Eqn. 4.13}$$




(a)




(b)

Figure 4.5: Measured S-parameters of transistor of Figure 4.4 in 6-45 GHz, 75-110 GHz and 140-220 GHz frequency bands.

The transferred-substrate technology has a zero series resistance Schottky collector contact and an extremely low extrinsic collector-base capacitance, justifying the assumptions of Eqn. 4.13. An effective collector-base capacitance can be defined as $C_{cb}$ = Im$(Y_{CB})/?$ . Figure 4.6 shows the effective $C_{cb}$ plotted versus frequency at increasing collector bias currents and a constant collector-base voltage. The data shows a clear trend of decreasing $C_{cb}$ with increasing bias current, an observation that is consistent with the capacitance cancellation model presented in the previous section. A total reduction of ~2 fF is observed over the range of applied bias currents.

Associated with the measured singularity in the unilateral power gain are observations of negative resistance effects in the common-emitter reverse conductance $G_{12}$ and in the output conductance $G_{22}$. To better illustrate this, consider $Y_{12}$ of an HBT described by a hybrid-p equivalent circuit model. To simplify the analysis, the device is assumed to have zero series emitter resistance, an assumption that reduces the terms in the expression for $Y_{12}$ but does not change the overall behavior of the parameter. Under this assumption, $Y_{12}$ expanded to second order in frequency is given by

$$Y_{12} = -\left(R_{cb} + w^2 C_{cb,i} C_{be} R_{bb}\right) - \mathrm{j}w\left(C_{cb,i} + C_{cb,x}\right) \qquad \text{Eqn. 4.14}$$

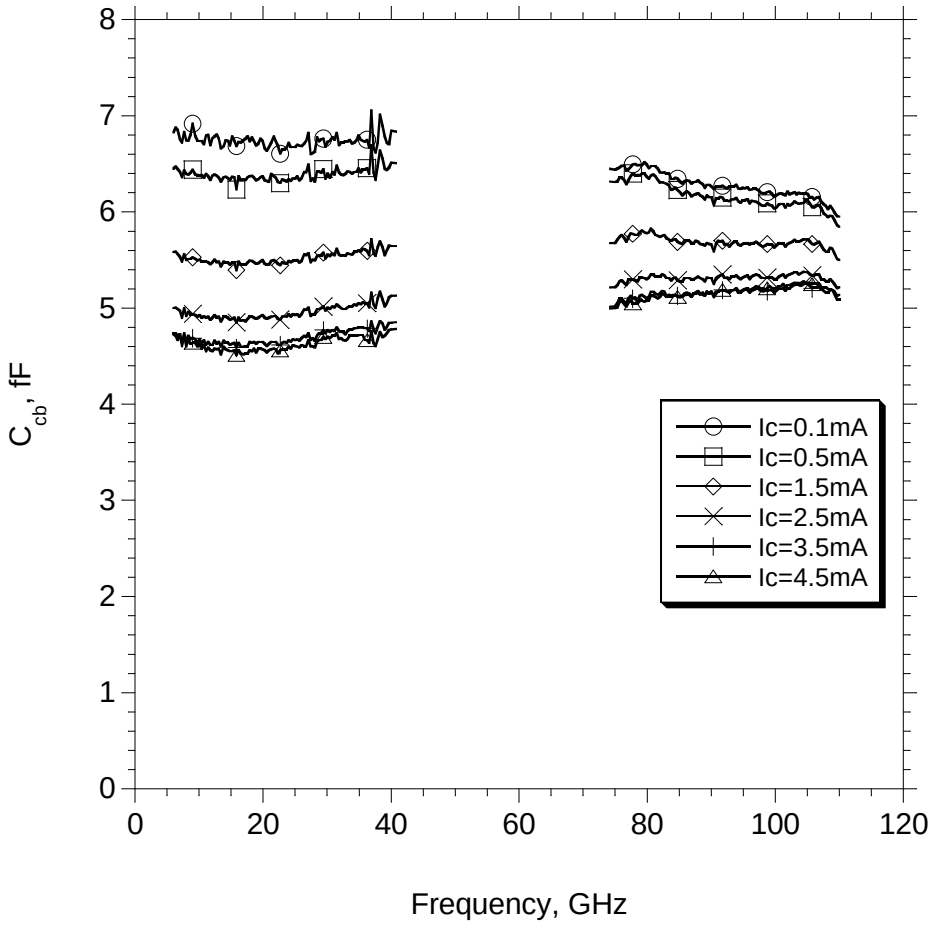


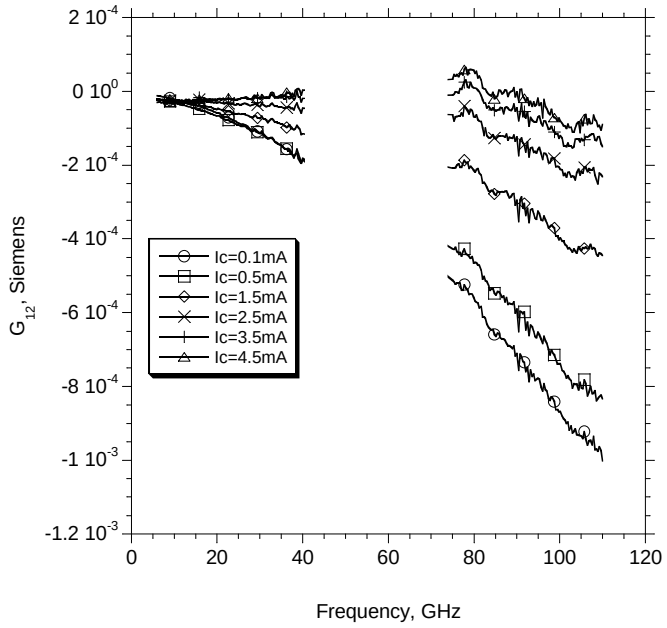

Figure 4.6: Effective base-collector capacitance $C_{cb}$ at varying $I_C$ and constant $V_{CB}$ = 0.35 V for transistor of Figure 4.4.

where $C_{be}$ is the base-emitter capacitance (junction and diffusion), $R_{bb}$ is the base resistance, $C_{cb,i}$ is the portion of collector-base capacitance internal to $R_{bb}$ in the circuit model, $C_{cb,x}$ is the remaining collector-base capacitance external to $R_{bb}$, and $R_{cb}$ is a finite collector-base resistance that arises in InGaAs collector HBTs from impact ionization in the collector region.

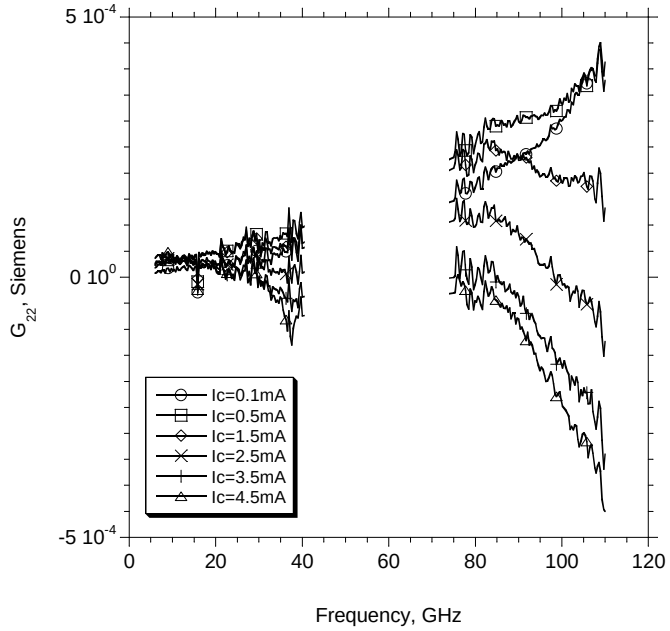
For an HBT described by a hybrid-p circuit model the real part of $Y_{12}$ ($G_{12}$), will show a parabolic variation with frequency and will always be negative. Figure 4.7, shows $G_{12}$ for the transistor of Figure 4.4 plotted versus frequency at varying collector currents and a constant collector-base voltage $V_{CB}$ = 0.35 V. The data shows that at low bias currents $G_{12}$ demonstrates the frequency dependence of Eqn. 4.14. However, as the current increases, the slope of $G_{12}$ changes, and eventually, positive $G_{12}$ is observed over portions of the frequency band. Similarly, $G_{22}$, also plotted in Figure 4.7, shows a trend towards negative output conductance with increasing bias current. In the following section, the dynamic capacitance cancellation model developed in the previous section is added to an HBT equivalent circuit model and similar trends in $G_{12}$ and $G_{22}$ are observed.

## *4.3 Equivalent Circuit Model*

The collector-base admittance that arises due to electron velocity modulation in the collector (Eqn. 4.9) may be added to the equivalent circuit model of a transferred-substrate HBT. However, before considering this addition, HBT measurements will be considered to determine whether the degree of capacitance cancellation experimentally observed in transferred-substrate HBTs is consistent with the developed model.




(a)




(b)

Figure 4.7: $G_{12}$ (a) and $G_{22}$ (b) plotted versus frequency at varying $I_C$ and constant $V_{CB}$ = 0.35 V for transistor of Figure 4.4.

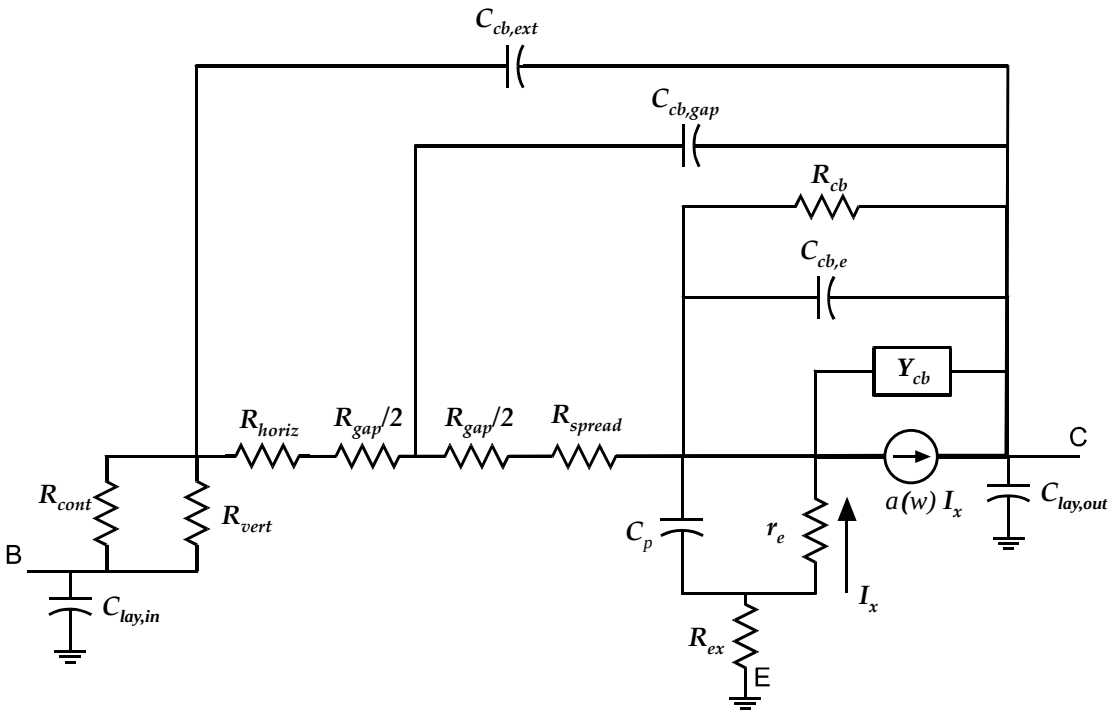
In the static capacitance cancellation model of Eqn. 4.11, the low frequency capacitance cancellation term is related to the collector current and the derivative of the collector transit time with respect to the collector-base voltage ($C_{cb,canc} = I_{C,0}\,\partial t_c / \partial V_{CB}\big|_{W_C}$). This derivative is taken under the condition of a constant collector depletion region thickness ($W_C$), and as pointed out in [5] this parameter is not generally available from measured transistor data. However, in the transferred-substrate technology the lightly-doped collector region is followed by a Schottky collector contact, and if fully depleted, we expect to observe little variation in the collector thickness. For the transistors reported in this work, the low frequency collector-base capacitance was measured at zero collector current and varying $V_{CB}$. It was found that the capacitance showed little variation at collector-base voltages greater than 0.35 V.

If the depleted collector thickness does not vary with applied $V_{CB}$, then an approximate measure of the term $C_{cb,canc} = I_{C,0}\,\partial t_c / \partial V_{CB}\big|_{W_C}$ can be determine from the change in the transistor current gain cutoff frequency in response to a small change in the collector-base voltage. For the transistor described in Figure 4.4, the HBT $f_t$ was measured at a constant collector current and the collector-base voltage was varied between 0.35V and 0.45V. At a collector current of 4.5 mA, an $f_t$ = 123 GHz and $f_t$ = 117 GHz were measured at $V_{CB}$ = 0.35V and $V_{CB}$ = 0.45V, respectively. If the decrease in $f_t$ is solely attributed to an increase in the collector transit time, then the change in collector transit time is

$\Delta t_c = \frac{1}{2p(117\,\mathrm{GHz})} - \frac{1}{2p(123\,\mathrm{GHz})} = 0.066\mathrm{psec}$, and the low frequency capacitance cancellation is approximated as $C_{cb,canc} = I_C \frac{\Delta t_c}{\Delta V_{cb}} = 3.0\,\mathrm{fF}$.

The estimated decrease in collector-base capacitance is slightly larger than the observed decrease of 2 fF measured at $V_{CB}$ = 0.35V (Figure 4.6). However, the measurement indicates that electron velocity modulation can account for the large relative decrease in the collector-base capacitance of submicron HBTs at increasing bias currents. Unfortunately, due to their poor breakdown and thermal characteristics, the validity of Eqn. 4.11 can only be tested over a limited bias range for InGaAs collector transferred-substrate HBTs. A more detailed experimental consideration of Eqn. 4.11 was performed in [5].

The transistor is modeled using the modified Tee-topology presented in Chapter 2 and shown again in Figure 4.8. The terms $C_{cb,e}$ and $C_{cb,gap}$ represent the portion of the collector-base capacitance directly under the emitter and the gap between the emitter and base contact, respectively. The values of these parameters are calculated using the estimated transistor geometry. The additional capacitance $C_{cb,ext}$ accounts for the remaining collector-base capacitance that is extracted from measurements at zero collector current and the applied base collector voltage such that $C_{cb,ext} = C_{cb0,meas} - C_{cb,e} - C_{cb,gap}$. The resistances $R_{spread}$, $R_{cont}$ and $R_{gap}$, and $R_{vert}$ were determined from the estimated transistor geometry, and sheet and contact resistances that were measured on the transistor epitaxy by the TLM method.




$r_e$ = 6.0O
$R_{ex}$ = 7.0 O
$R_{cb}$ = 25 kO
$R_{spread}$ = 1.1 O
$R_{gap}$ = 1.4 O
$R_{cont}$ = 7.5 O
$R_{vert}$ = 6.7 O
$R_{horiz}$ = 3.5 O

$C_{cb,e}$ = 2.1 fF
$C_{cb,gap}$ = 0.7 fF
$C_{cb,ext}$= 4.0 fF
$C_{lay,in}$ = 14.7 fF
$C_{lay,out}$ = 0.7 fF

$$C_p = C_{je} + \frac{5t_b}{6r_e}$$

$C_{je}$ = 52.5 fF

$$a(w) = a_o \left[\frac{\sin(wt_c)}{wt_c}\right] \exp\left[-jw\left(t_c + \frac{t_b}{6}\right)\right]$$

$\alpha_o$ = 0.925
$t_c$ = 0.49 psec
$t_b$ = 0.20 psec

Figure 4.8: Equivalent circuit model and parameter values used to simulate HBT of Figure 4.4. Admittance block $Y_{cb}$ implements capacitance cancellation model of Eqn. 4.9.

The collector-base admittance due to electron velocity modulation (Eqn. 4.9) is included in the model as the admittance block $Y_{cb}$ appearing in parallel with $C_{cb,e}$. For simulations, a low frequency collector-base capacitance cancellation ($I_C \partial t_c / \partial V_{CB}$) of 3 fF has been assumed. This value corresponds to the measured decrease in $C_{cb}$ for the device at the bias conditions of Figure 4.4. The remaining terms in the small-signal model were determined using a bias dependent extraction technique similar to that presented in [8]. A complete description of all of the parameter values is included Figure 4.8.

Figure 4.9 shows the simulated unilateral power gain, maximum available gain and short circuit current gain using the model of Figure 4.8. The measured data from the transistor of Figure 4.4 is also included on the plot. The simulated data shows a singularity in the unilateral power gain occurring at a slightly higher frequency than the measured singularity, and $U$ is observed to remain unbounded over the remainder of the simulated frequency range. In Figure 4.10, the modeled and simulated S-parameters are presented. Note that the model extraction used no numerical optimization to determine model parameters, and that the values for some of the parameters were based on estimated transistor geometries. In this context, the agreement between measurement and simulation is quite good.

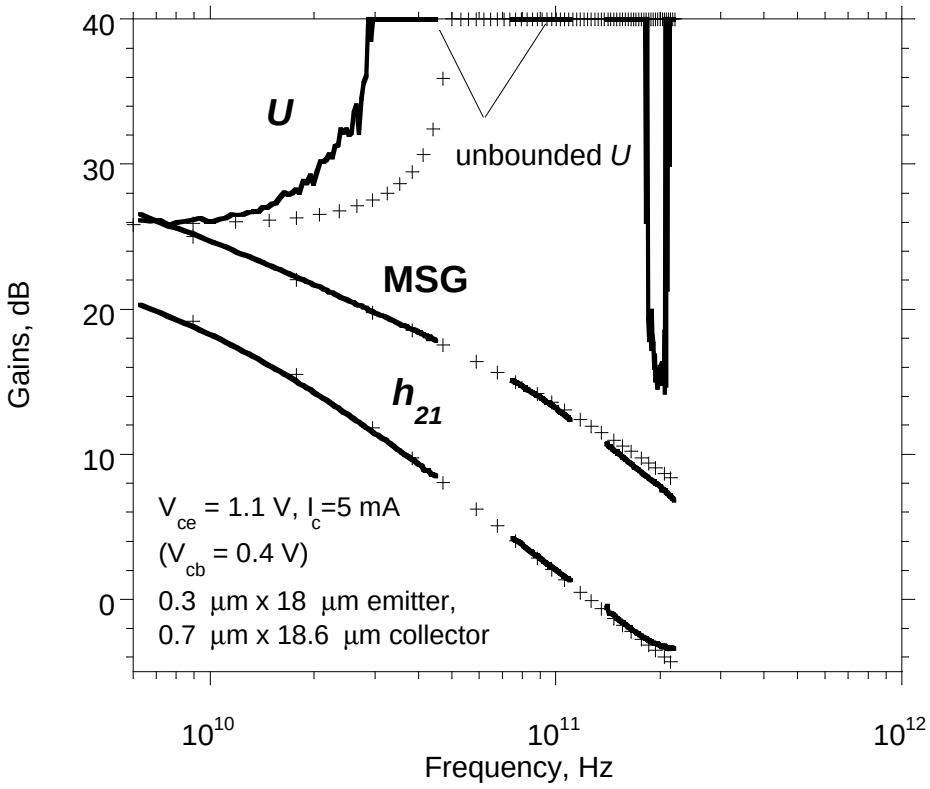
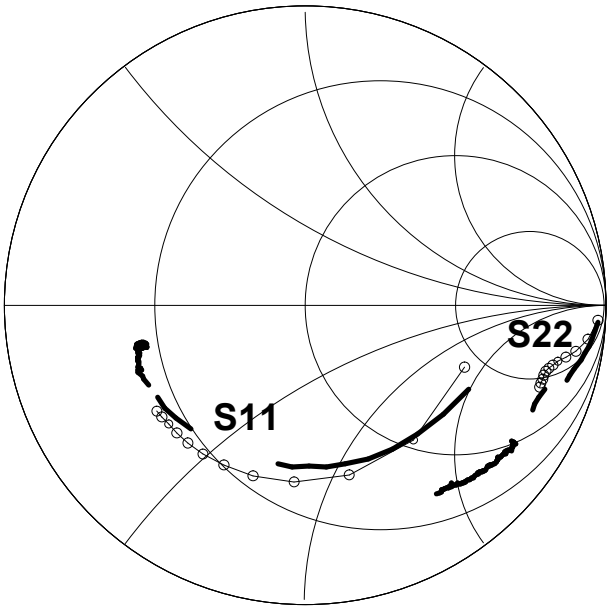
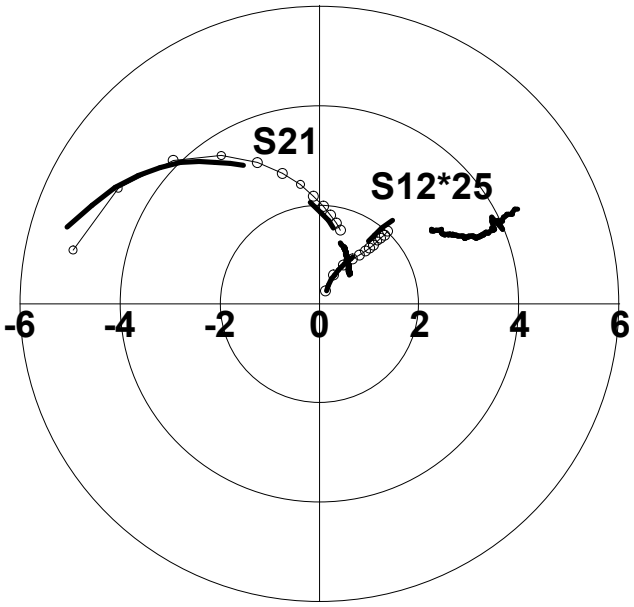



Figure 4.9: Measured (solid line) and simulated (crosses) unilateral power gain ($U$), short circuit current gain ($h_{21}$) and maximum stable gain (MSG) of transferred-substrate HBT.




(a)



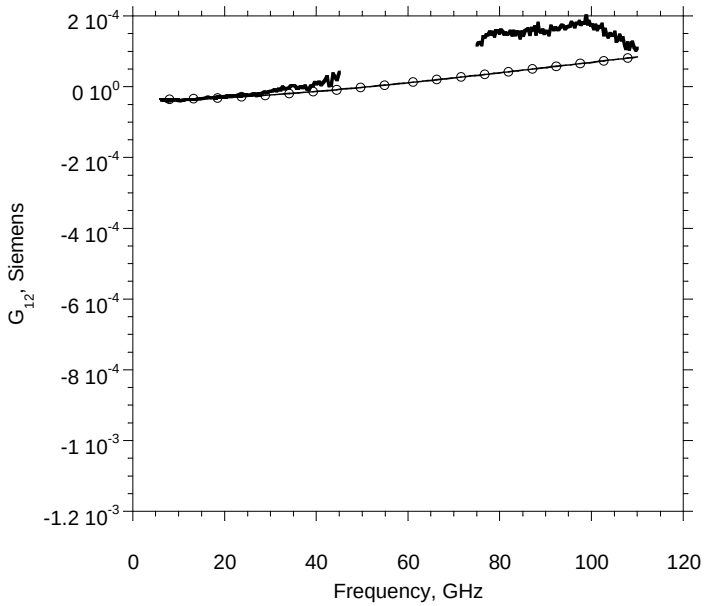

(b)

Figure 4.10: Transistor S-parameters measured in 6-45 GHz, 75-110 GHz and 140-220 GHz frequency bands (solid lines) and simulated S-parameters (6-110 GHz) using the transistor model of Figure 4.8.

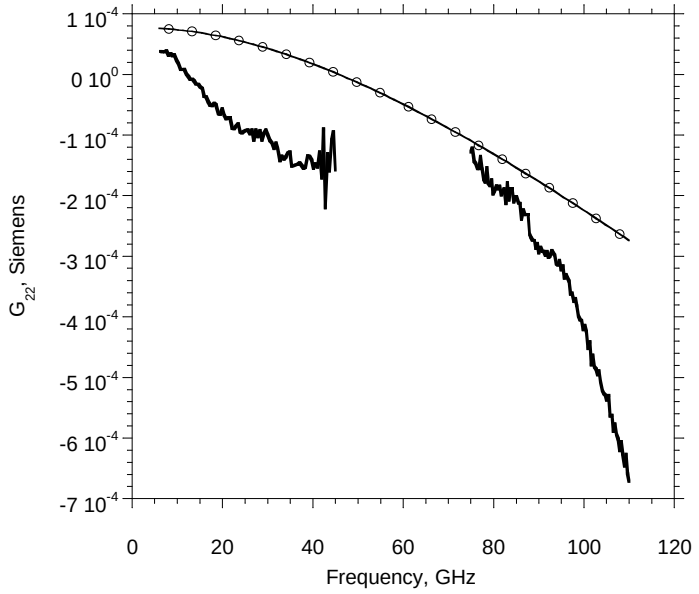
From simulations, the real parts of the transistor Y-parameters also show the same negative resistance trends as those observed in the measured device. Figure 4.11 shows $G_{12}$ and $G_{22}$ of the simulated and measured transistors. The agreement between measurement and simulation is reasonable given the constraints on the measurements and model parameters. It is again noted that the $Y_{cb}$ model is approximate given the uniform collector velocity model used in the derivation. Additionally, an accurate measurement of $G_{12}$ and $G_{22}$ is difficult since these terms are extremely small in the low parasitic transferred-substrate technology.

## 4.4 Conclusions

Power gain singularities have been observed in the measurements of submicron InGaAs-collector HBTs. Associated with these singularities are trends towards negative conductance in the common-emitter output conductance and positive conductance in the common-emitter reverse transmission characteristics. These trends cannot be predicted by standard HBT circuit models. The HBTs also exhibit a decrease in the effective collector-base capacitance with increasing current density. A dynamic model for collector-base capacitance cancellation due to electron velocity modulation in the collector has been developed. This model was incorporated with a small-signal equivalent HBT circuit model and singularities in the simulated unilateral power gain were observed.




(a)




(b)

Figure 4.11: Transistor (a) $G_{12}$ and (b) $G_{22}$ measured in 6-45 GHz, 75-110 GHz frequency bands (solid lines) and simulated (circles) using the transistor model of Figure 4.8.

The negative resistance trends observed in some transferred-substrate HBTs are atypical of III-V transistors. In a standard mesa-HBT, the reverse and output conductances of the transistor are dominated by the large extrinsic collector-base capacitance. The transferred-substrate process eliminates a majority of the extrinsic collector-base capacitance, and the intrinsic junction properties play a more important role in the device characteristics. Therefore, it is expected that second-order transport effects, like those described in this chapter, must be modeled to fully describe the transistor performance. However, the negative resistance effects described in this chapter are relatively small, and one does not see a large change in simulated S-parameters whether using a dynamic or static capacitance cancellation model.

The negative resistance trends do cause a large variation in the measurement of the transistor unilateral power gain with the consequence that the transistor $f_{max}$ cannot be extrapolated from measurements of $U$. However, the transistors exhibit a large available gain at the frequency limits of commercially available network analyzers, and transistor amplifiers presented in the next chapter further support the high frequency performance of the technology. Ultimately, it is through direct measurement and the use of devices in circuit applications that the maximum usable frequency of highly scaled HBTs will be determined.

## REFERENCES


1. R. Pullela, Q. Lee, B. Agarwal, D. Mensa, J. Guthrie, L. Samoska, M. Rodwell, " A > 400 GHz fmax Transferred Substrate Heterojunction Bipolar Transister IC Technology," *IEEE Electron Device Letters*, March 1998, vol. 19, no. 3, p.77.

2. Q.Lee, S.C. Martin, D. Mensa, R.P. Smith, J. Guthrie, S. Jaganathan, T. Mathew, S. Krishnan, S. Ceran, and M.J.W. Rodwell, "Submicron Transferred-Substrate Heterojunction Bipolar Transistors with Greater than 800 GHz fmax," C*onference Proceedings Indium Phosphide and Related Materials Conference*, Davos, Switzerland, May 16-20, 1999.

3. Q. Lee, S.C. Martin, D. Mensa, R.P. Smith, J. Guthrie, S. Jaganathan, Y. Betser, T. Mathew, S. Krishnan, L. Samoska, and M.J.W. Rodwell, "Submicron-Transferred-Substrate Heterojunction Bipolar Transistors with Greater than 1 THz fmax" *Conference Digest 57th Device Research Conference*, University of California, Santa Barbara, CA, June 28-30, 1999, Post Deadline.

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

5. Y. Betser and D. Ritter, "Reduction of the base collector capacitance in InP/GaInAs heterojunction bipolar transistors due to electron velocity modulation", *IEEE Transactions on Electron Devices*, vol. 46 , no. 4, April 1999

6. T. Ishibashi, "Nonequilibrium electron transport in HBTs," *IEEE Transactions on Electron Devices*, vol. 48, no. 11, Nov. 2001, pp. 2595-2605.

7. S.J. Mason, "Power gain feedback amplifier," *IRE Trans. Circuit Theory*, vol. CT-1, 1954, pp. 20-25.

8. S.J. Spiegel, D. Ritter, R.A. Hamm, A. Feygenson, and P.R. Smith, "Extraction of InP/InGaAs heterojunction, bipolar transistor small-signal equivalent circuit," *IEEE Transactions on Electron Devices*, vol. 42, no. 6, June 1995. pp. 1059-1064.

# Chapter 5 Ultra-high Frequency Amplifiers

The wide bandwidth and high available gain demonstrated by transferred-substrate HBTs makes them a promising candidate for G-band (140-220 GHz) electronics. Electronics in the 140-220 GHz band have applications in wideband communication systems, remote atmospheric sensing and planetary exploration. State-of-the-art amplifier results in this frequency range have been realized using deep submicron InP-based high electron mobility transistors (HEMTs). These results include: a 3-stage amplifier with 30 dB gain at 140 GHz [1], a 3-stage amplifier with 12-15 dB gain from 160-190 GHz [2] and a 6-stage amplifier with 20 ± 6 dB from 150-215 GHz [3].

In this chapter, small-signal amplifier designs in the transferred-substrate HBT technology are described. A single-stage amplifier was realized with 6.3 dB gain at 175 GHz [4]. Multi-stage amplifiers were subsequently designed with a three-stage amplifier exhibiting 12.0 dB gain at 170 GHz and a second design exhibiting 8.5 dB gain at 195 GHz [5]. In this section, amplifier results are presented and design considerations for ultra-high frequency tuned circuits are considered.

## *5.1 Device Characteristics*

Amplifier designs were fabricated in the transferred-substrate HBT technology described in Chapter 2. The epitaxial layer structure was identical to that described by Table 2.1. The transistors used in the amplifier designs had nominal emitter-base junction dimensions of 0.4 x 6 $\mu m^2$ and nominal collector-base junction dimensions

of 0.7 x 6.4 $\mu m^2$. Such devices have typical DC current gains ß of ~20 and common emitter breakdown voltages of ~1.5V at an emitter current density of 1 mA/ $\mu m^2$. In the amplifier designs, transistors were biased at an emitter current density of 1.5-2 mA/ $\mu m^2$ and at a collector-emitter voltage of 1.2-1.3 V. The RF gains of a transistor from a multi-stage amplifier process run are shown in Figure 5.1. The device bias conditions are $V_{CE}$ = 1.25 V and $I_C$ = 3.2 mA, and the transistor is observed to have a maximum stable gain of > 7dB at 200 GHz. The $f_t$ of the device is measured to be 180 GHz. The negative resistance trends described in the previous chapter were also observed in these transistors; however, these effects did not have a large influence on the amplifier design or performance.

## *5.2 Circuit Topologies*

A schematic for the single-stage amplifier is shown in Figure 5.2. The amplifier utilized a common-emitter topology with transmission line-based input and output impedance matching networks. The matching networks were synthesized using a series transmission line with an open circuit shunt stub. The input and output were designed to be simultaneously matched to a 50 O characteristic impedance at the design frequency.

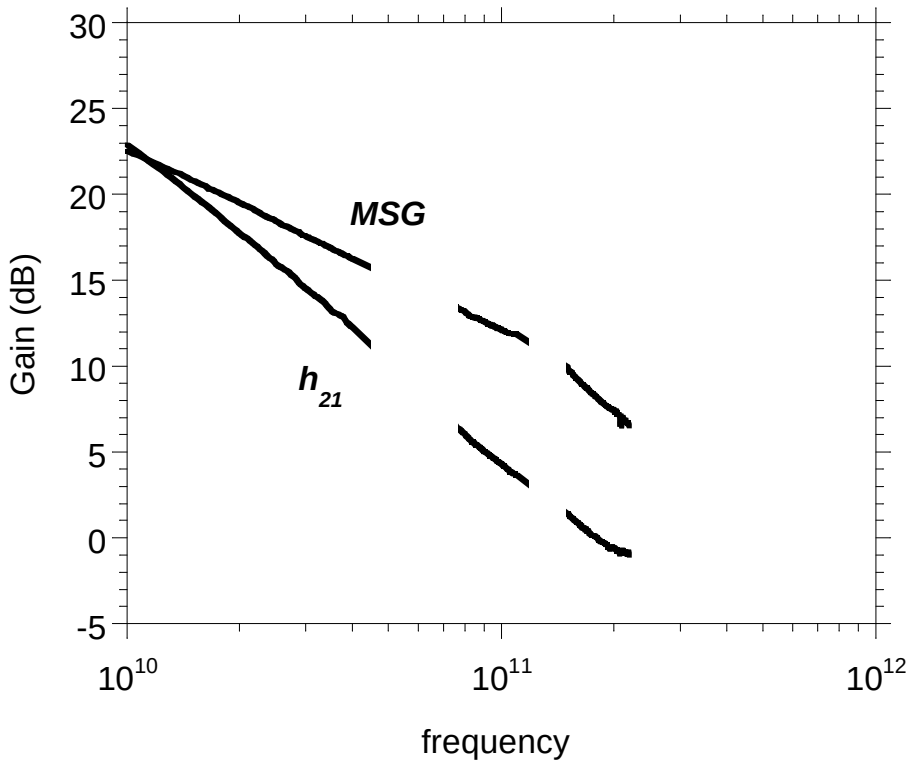



Figure 5.1: HBT short circuit current gain ($h_{21}$) and maximum stable power gain (*MSG*) measured in the 10-40 GHz, 75-110 GHz and 140-220 GHz bands. Device bias conditions $V_{CE}$ = 1.25 V, $I_C$ = 3.2 mA.

Low frequency stabilization was realized with a shunt resistor at the output of the transistor. A quarter-wavelength high impedance line to a shunt MIM capacitor bypassed the resistor at the intended design frequency. In the 140-220 GHz frequency range, resistive loss in the matching network was sufficient to stabilize the device. Bias Tees in the on-wafer probes were used to provide DC bias to the amplifier. First generation single-stage amplifiers were designed for an intended design frequency of 200 GHz. A chip photograph of a fabricated single-stage amplifier is shown in Figure 5.3.

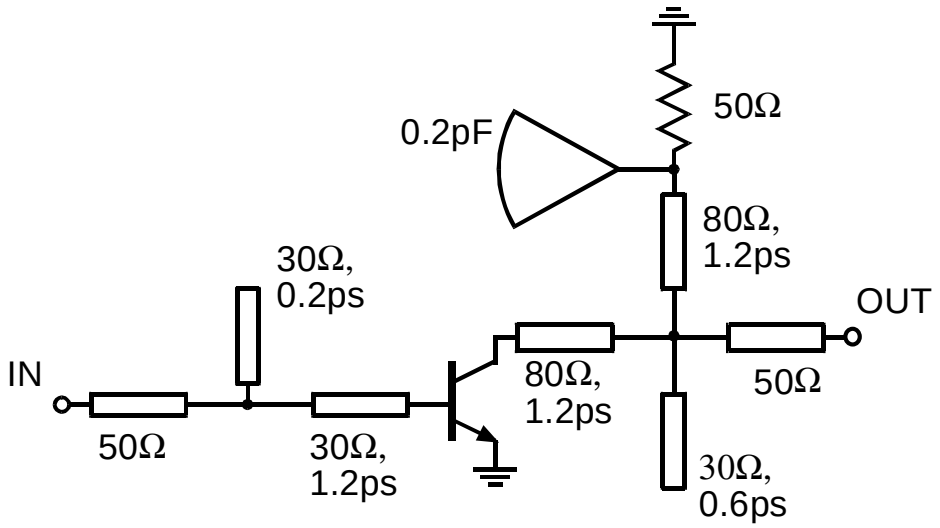



Figure 5.2: Circuit schematic of single-stage amplifier design.

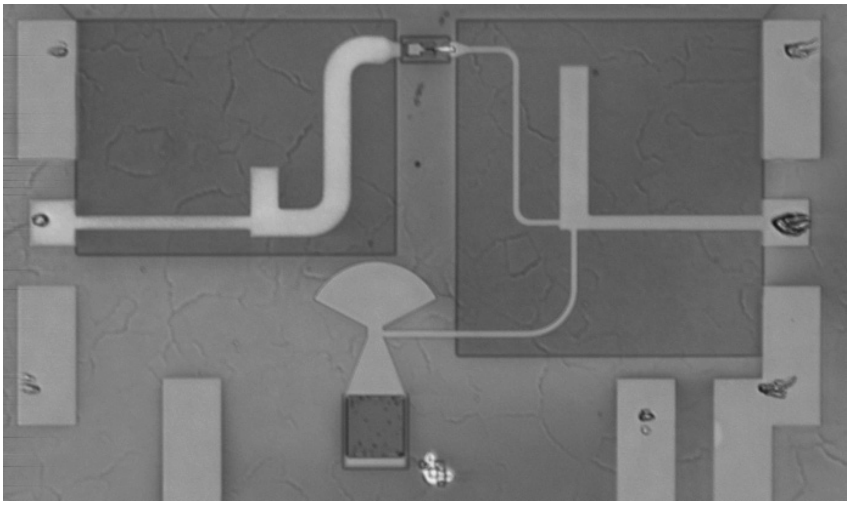


Figure 5.3: Chip photograph of single-stage amplifier. Cell dimensions are 0.69 mm x 0.35 mm.

After the successful realization of single-stage amplifiers, three-stage amplifiers were subsequently designed. A circuit schematic of a multi-stage amplifier design is shown in Figure 5.4. Inter-stage MIM capacitors, with nominal values of 75 fF, provide DC isolation between stages. To simplify the design, separate supply lines were used to provide DC bias to the base and collector of each device. The same stabilization scheme used for the single-stage amplifier was used for each of the transistors in the multi-stage design. Two multi-stage amplifiers were designed with intended design frequencies of 175 GHz and 200 GHz. A chip photograph of a fabricated multi-stage amplifier is shown in Figure 5.5.

## *5.3 Millimeter-Wave Design Considerations*

Circuits were designed using Agilent Technologies Advanced Design System simulation software [6]. The high design frequencies presented challenges in active and passive device modeling. Initial single-stage amplifiers were designed using a hybrid-p transistor model. This model was based on a low frequency extraction of previously fabricated submicron transferred-substrate HBTs [7]. It was found that HBT circuit models extended to the 140-220 GHz frequency band showed poor agreement with measurements.

Figure 5.6 shows measured and simulated results of a transferred-substrate HBT in the 6-45 GHz and 140-220 GHz frequency bands. The simulated results utilized a hybrid-p equivalent circuit model (Figure 2.5). The model parameters were determined using a bias dependent extraction technique using the measured data from the 6-45 GHz band. The extraction method is similar to that presented in [8].

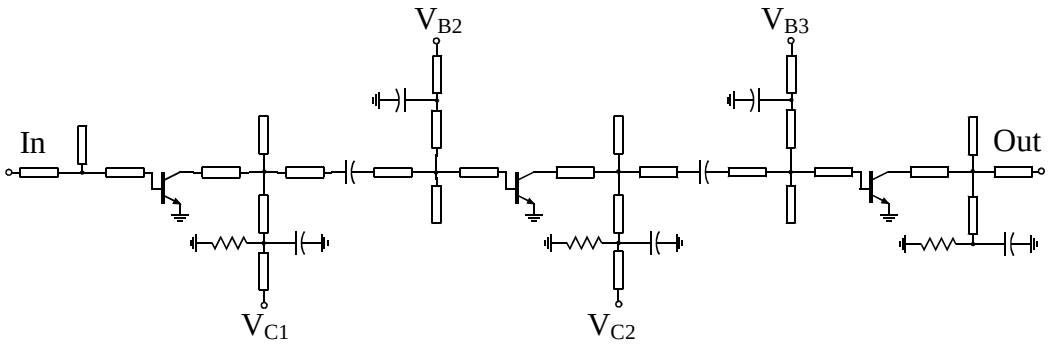



Figure 5.4: Circuit schematic of three-stage amplifier design.

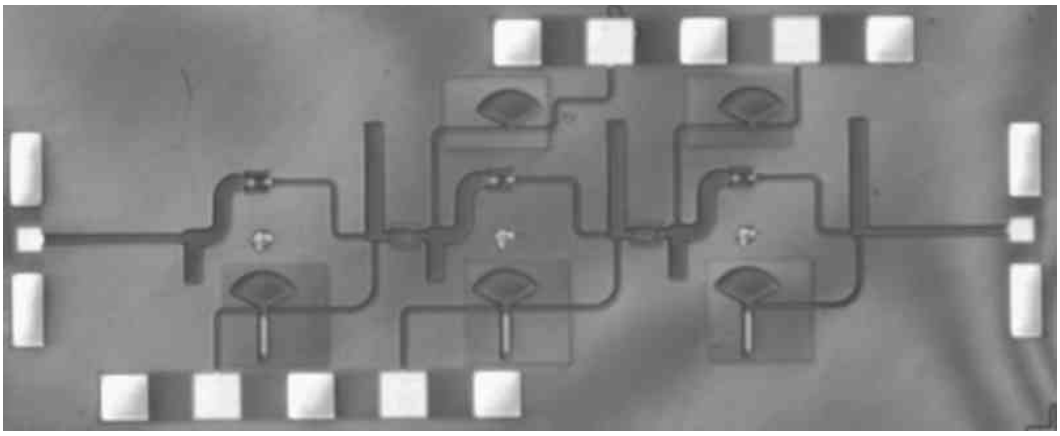


Figure 5.5: Chip photograph of three-stage amplifier. Cell dimensions are 1.66 mm x 0.35 mm.

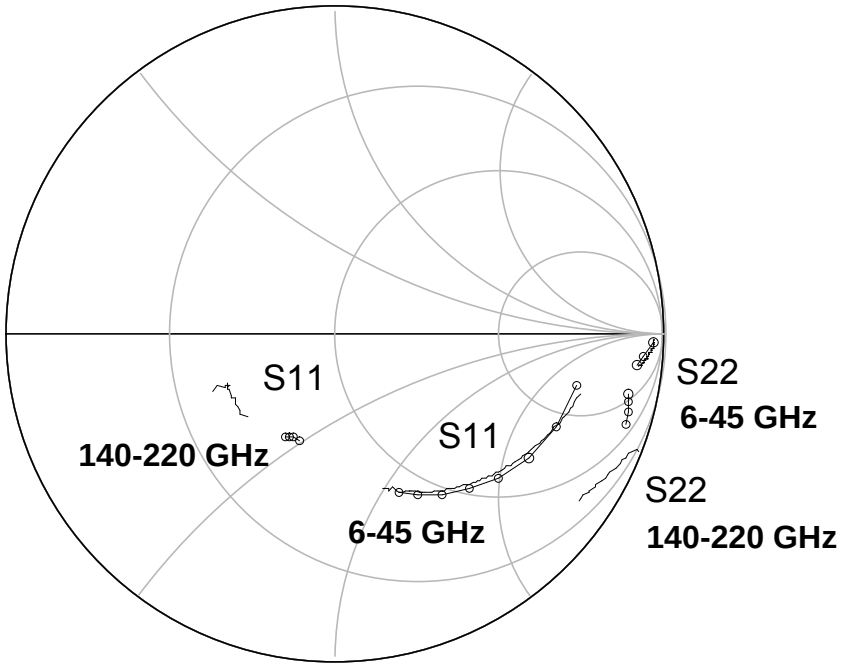



Figure 5.6: S-parameters of device measurements (solid lines) and simulations of hybrid- p model (circles) from 6-45 GHz and 140-220 GHz.

While good agreement was achieved between simulated and measured transistor parameters from 6-45 GHz, a large discrepancy was observed between measured and simulated results in the 140-220 GHz band. The capacitance cancellation model presented in Chapter 4 was not included in the simulated model. However, it was found that the addition of the model did not improve the agreement with the measured results.

The source of the discrepancy between measured transistor parameters and transistor model simulations in the 140-220 GHz band was not determined. Developing an accurate device model is necessary to better understand transistor operation at these and higher operating frequencies. However, measured device S-

parameters can be used in small-signal tuned circuit applications, and this was the approach taken for subsequent multi-stage amplifier designs.

The passive microstrip matching networks were designed using a combination of electromagnetic (EM) simulation and empirical computer aided design (CAD) models. A planar method-of-moments EM simulator was used to model the unique MIM capacitor structures and any microstrip discontinuities in the circuit layout (i.e. Tee junctions and bends). Standard microstrip CAD models were used to model the straight transmission line segments. This approach allowed CAD optimization routines to be used for the matching network design by modifying the lengths of the straight line segments. Optimization using EM simulators is computationally intensive, and full EM simulations of the matching networks were performed only as a final design verification.

To verify the accuracy of the passive element modeling approach, a test structure consisting of the amplifier input and output matching networks cascaded together without an active device was included on-wafer. Figure 5.7 shows the measured S-parameters of the test structure and simulation results of the same structure using the modeling approach described above. Very close agreement is seen between measured and simulated results.

The accuracy of the microstrip CAD models and electromagnetic simulations at such high frequencies can be attributed in part to the wiring environment provided by the transferred-substrate process. The thin (5μm) low dielectric constant ($e_r \sim 2.7$) BCB substrate ensures a single-mode propagation environment. The design of the

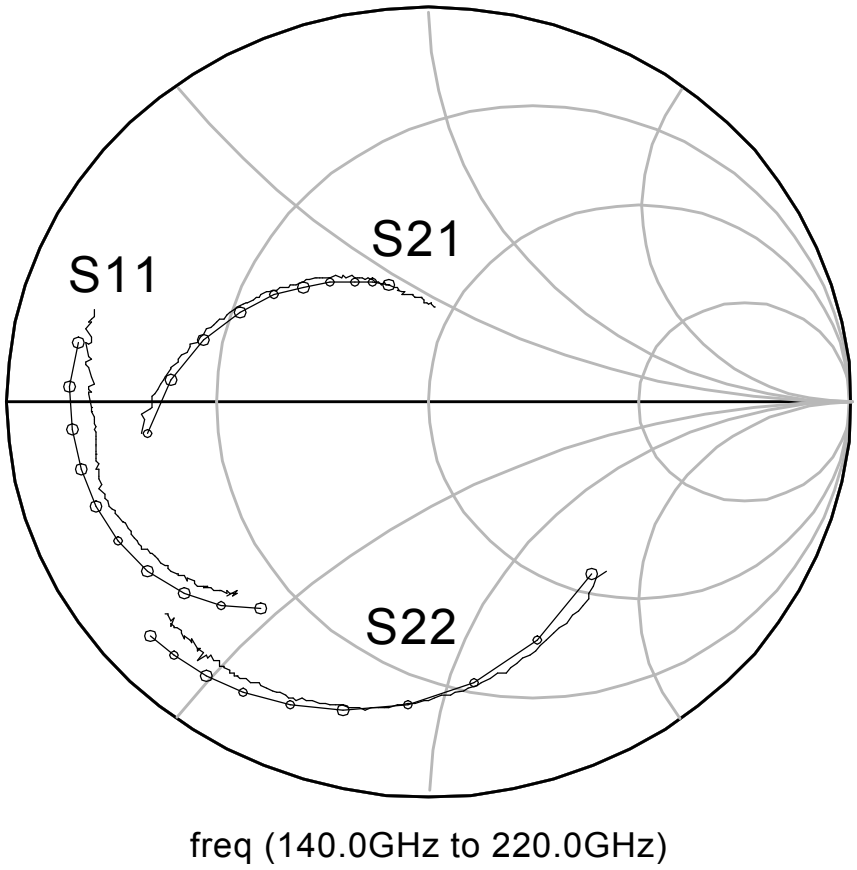



Figure 5.7: Measured (solid) and modeled (circle) S-parameters of matching network test structure.

wiring environment was motivated by the initial application of the transferred-substrate process to mixed-signal and digital integrated circuits. In these applications, low line-to-line crosstalk, low ground access inductance and good thermal heat-sinking are desirable. A thin substrate thickness is beneficial for the aforementioned requirements, but for high frequency tuned-circuit applications, it was found that these advantages were offset by excessive resistive loss in the transmission line matching networks.

As the substrate thickness is thinned, the width of the microstrip line necessary to realize a given characteristic impedance will be reduced and resistive losses will be increased. In the transferred-substrate microstrip environment, the width of a 50 O transmission line is 12.5 μm and resistive losses are high (~12 dB/cm at 200 GHz). For the single-stage amplifier designs presented here, circuit simulations showed a 2.0 dB higher gain if lossless transmission line models were used. The resistive losses could be reduced by increasing the BCB substrate thickness. However, this path was not pursued due to concerns of incorporating changes to the transferred-substrate process.

## *5.4 Amplifier Results*

The amplifiers were measured on-wafer from 140-220 GHz. The measurements were made using an HP8510C Vector Network Analyzer (VNA) with Oleson Microwave Labs Millimeter Wave VNA Extensions. The test set extensions are connected to GGB Industries coplanar wafer probes via a short length of WR-5 waveguide. The amplifier measurements were calibrated on-wafer using TRL calibration standards.

Figure 5.8 shows the measured gain, and input and output return loss of a single-stage amplifier design. The bias conditions for the transistor were $V_{CE}$ = 1.2 V and $I_C$ = 4.8 mA. The amplifier was found to have a peak gain of 6.3 dB at 175 GHz, with a gain of better than 3 dB from 140 to 190 GHz. Both the input and output return loss were better than 10 dB at 175 GHz. The gain-per-stage of the result represents one of the highest reported in any transistor technology for this frequency band.

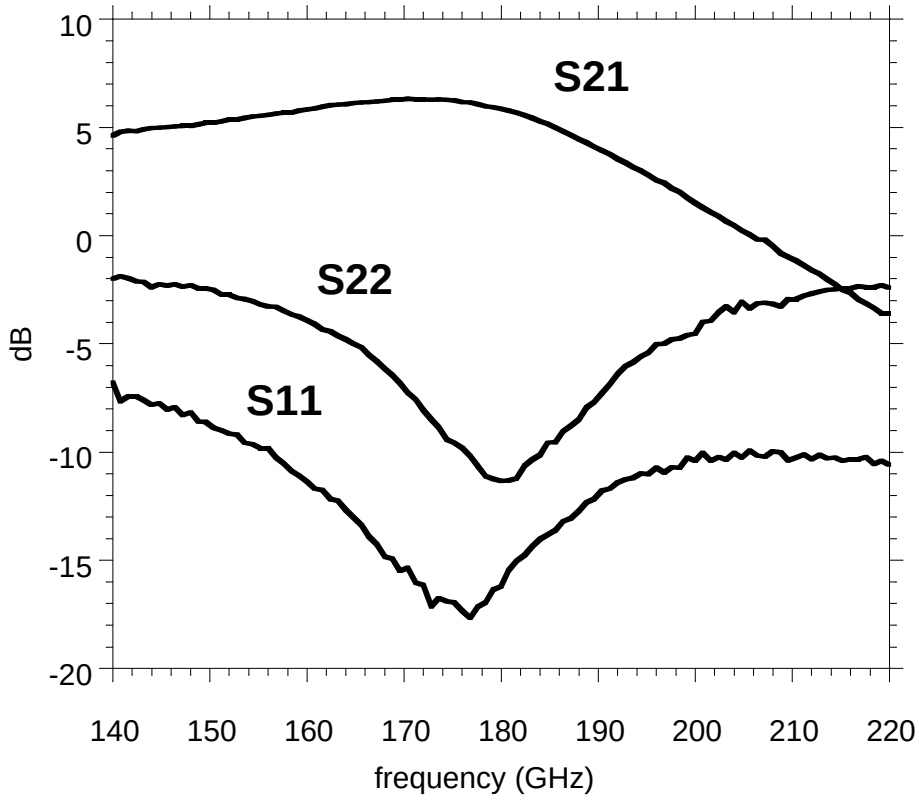
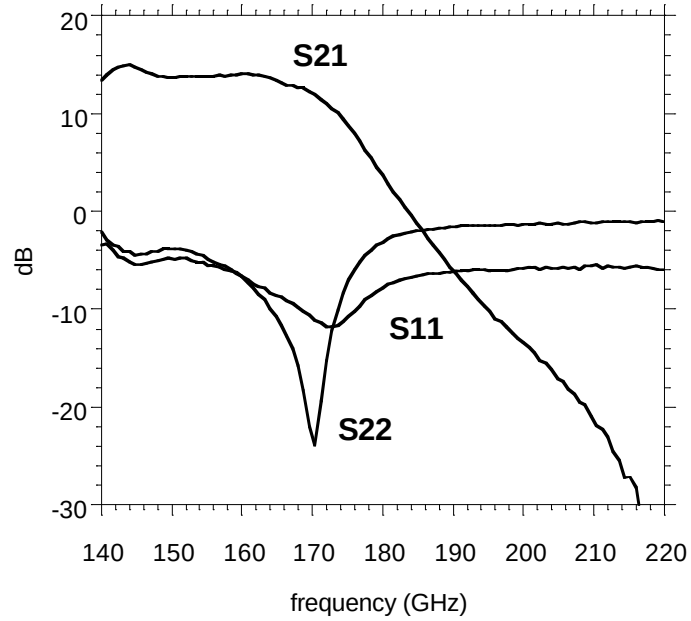
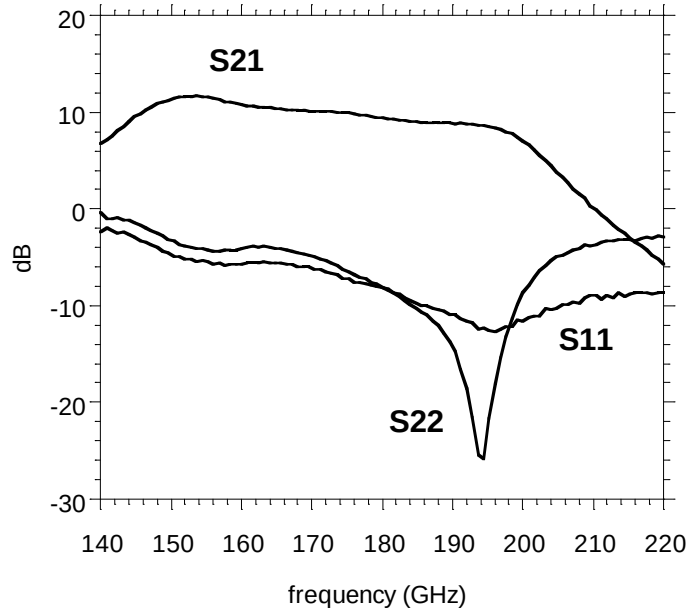



Figure 5.8: Measured S-parameters of single-stage amplifier.

Multi-stage amplifiers were fabricated in subsequent process runs. Figure 5.9(a) shows the measured gain, and input and output return loss of the 175 GHz amplifier design. Figure 5.9 (b) shows the same parameters measured for the 200 GHz amplifier design. For both amplifier designs, the transistors in the circuit were biased at $I_C$ = 3.2 mA, and $V_{CE}$ = 1.25 V.




(a)




(b)

Figure 5.9: Measured S-parameters of (a) 175 GHz and (b) 200 GHz multi-stage amplifier designs.

The 175 GHz amplifier design had a gain of 12.0 dB at the output match frequency of 170 GHz. A peak gain of 15.0 dB was measured at 144 GHz, and the gain was greater than 10 dB to 175 GHz. The 200 GHz amplifier design exhibited a gain of 8.5 dB at the output match frequency of 195 GHz. A peak gain of 11.7 dB was measured at 154 GHz and the gain was greater than 7.0 dB to 200 GHz.

The single stage amplifier showed a downward shift of ~25 GHz from the intended design frequency of 200 GHz. This shift was attributed to the hybrid-p transistor model used in the design cycle, which as discussed earlier was found to deviate significantly from measurements of devices in the 140-220 GHz band. The multi-stage amplifier designs were based on measured device S-parameters from the first generation single-stage amplifier process run. A downward shift of ~5 GHz from the intended design frequencies was observed for the multi-stage amplifier designs, and the peak gains of the designs were also less than those predicted from simulations. Transistor measurements from the multi-stage amplifier process run showed higher extrinsic emitter resistance and lower available power gain than the transistors used in the design cycle. Single-stage amplifier designs on this wafer demonstrated a peak gain of 3.5 dB at 175 GHz.

Figure 5.10 shows a circuit simulation of the 175 GHz multi-stage amplifier using measured transistor S-parameters from the multi-stage amplifier process run. The close agreement with measured amplifier results indicates that device variation is responsible for the downward shift from the design frequency and verifies the amplifier matching network design.

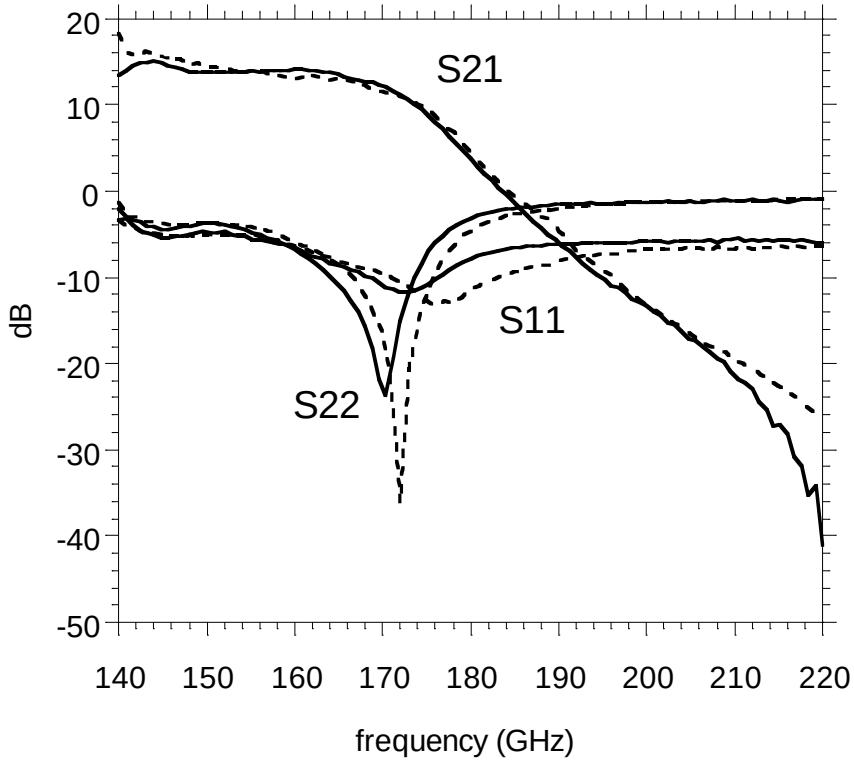



Figure 5.10: S-parameters of measured 175 GHz amplifier (solid lines) and circuit simulation of amplifier using measured transistor S-parameters (dashes).

### *5.5 Conclusions*

High-gain small-signal amplifiers have been realized in the 140-220 GHz frequency range. The transferred-substrate process provided a low parasitic HBT technology with high available gain in this frequency range, and the thin dielectric microstrip wiring environment enabled accurate modeling of the passive matching networks using standard microwave CAD tools. The drawbacks of the transferred-substrate technology were described in detail in Chapter 2. While the levels of integration required for millimeter-wave tuned-circuit applications are much less than those

required for mixed-signal circuits, the effort (i.e. failed process runs) required to yield the circuits described in this chapter was still very high.

The transferred-substrate devices do demonstrate the potential for a highly scaled HBT technology to compete with InP-based HEMTs for millimeter-wave circuit applications. Recent work on mesa-HBTs at UCSB has produced an InP double-heterojunction transistor with extrapolated $f_t$ and $f_{max}$ of 280 GHz, and >400 GHz, respectively [9]. The device had a measured maximum stable gain of ~5 dB at 175 GHz. The superior high frequency performance of the transistor is achieved through aggressive scaling of the base mesa width. Additionally, an extremely low base contact resistance is achieved through the use of highly carbon-doped InGaAs layers and palladium-based Ohmic contacts. In addition to its excellent RF performance, the device had a common-emitter breakdown voltage of 6 V, making it a promising candidate for millimeter-wave power applications. In the following chapter, a mesa-HBT technology is described that attempts to further reduce the parasitics of the transistor while increasing device yield and manufacturability.

## REFERENCES


1. C. Pobanz, M. Matloubian, M. Lui, H.-C. Sun, M. Case, C.M. Ngo, P. Janke, T. Gaier, and L. Samoska, "A high gain D-band InP HEMT amplifier," *IEEE Journal of Solid State Circuits*, vol. 34, no. 9, September 1999.

2. R. Lai, M. Barsky, R. Groundbacher, P.H. Liu, T.P. Chin, M. Nishimoto, R. Elmajarian, R. Rodriguez, L. Tran, A. Gutierrez, A. Oki, and D. Streit, "InP HEMT amplifier development for G-band (140-220 GHz) applications," *Proceedings 2000 International Electron Devices Meeting*, Dec., San Francisco, CA.

3. S. Weinreb, T. Gaier, R. Lai, M. Barsky, Y. C. Leong, and L. Samoska, "High-gain 120-215-GHz MMIC amplifier with integral waveguide transitions," *IEEE Microwave and Guided Wave Leters*, vol. 9, no. 7, July 1999.

4. M. Urteaga, D. Scott, T. Mathew, S. Krishnan, Y. Wei and M. Rodwell, "Single-stage G-band amplifier with 6.3 dB gain at 175 GHz," *2001 GaAs IC Symposium Digest*, October, Baltimore, MD.

5. M. Urteaga, D. Scott, S. Krishnan, Y. Wei, M. Dahlström, Z. Griffith, N. Parthasarathy, M. Rodwell, "Multi-stage G-band (140-220 GHz) InP HBT amplifiers," *2002 GaAs IC Symposium Digest*, October, Monterey, CA.

6. *Advanced Design System*, Palo Alto, CA: Agilent Technologies.

7. Q.Lee, S.C. Martin, D. Mensa, R.P. Smith, J. Guthrie, S. Jaganathan, T. Mathew, S. Krishnan, S. Ceran, and M.J.W. Rodwell, "Submicron Transferred-Substrate Heterojunction Bipolar Transistors with Greater than 800 GHz fmax," C*onference Proceedings Indium Phosphide and Related Materials Conference*, Davos, Switzerland, May 16-20, 1999

8. S.J. Spiegel, D. Ritter, R.A. Hamm, A. Feygenson, and P.R. Smith, "Extraction of InP/InGaAs heterojunction, bipolar transistor small-signal equivalent circuit," *IEEE Transactions on Electron Devices*, vol. 42, no. 6, June 1995. pp. 1059-1064.

9. M. Dahlström, M. Urteaga, S. Krishnan, N. Parthasarathy, M.J.W. Rodwell, “Ultra-Wideband DHBTs using a Graded Carbon-Doped InGaAs Base,” *Proceedings 2002 Indium Phosphide and Related Materials Conference, Post-Deadline,* May, Stockholm.

# Chapter 6 HBT Design for Digital Logic Speed

Submicron devices fabricated in the transferred-substrate process demonstrated the potential for a low parasitic HBT technology in millimeter-wave tuned circuit applications. The technology is well-suited for these types of applications given its high available gain at millimeter-wave frequencies and the low levels of integration required for these types of circuits. However, the application of wide bandwidth transistors extends beyond RF and millimeter-wave wireless transceiver circuits. Gigahertz frequency mixed-signal and digital integrated circuits have applications in fiber optic transceiver systems [1] and military communications systems [2], and transistor counts in these types of circuits may approach 10,000 devices.

In the mixed-signal and digital IC market, compound semiconductor devices compete with Si /SiGe bipolar transistors [3]. Given the investment into the infrastructure of silicon-based technologies, it is not surprising that Si/SiGe technologies are well suited for realizing high levels of integration. The process flows in silicon-based bipolar technologies are also well suited for submicron device scaling and extrinsic parasitic reduction. Figure 6.1 shows a cross-section of a Si/SiGe HBT taken from [3]. Specific process steps that facilitate the fabrication of submicron devices will now be considered in some detail to illustrate the differences with a typical III-V HBT process flow.

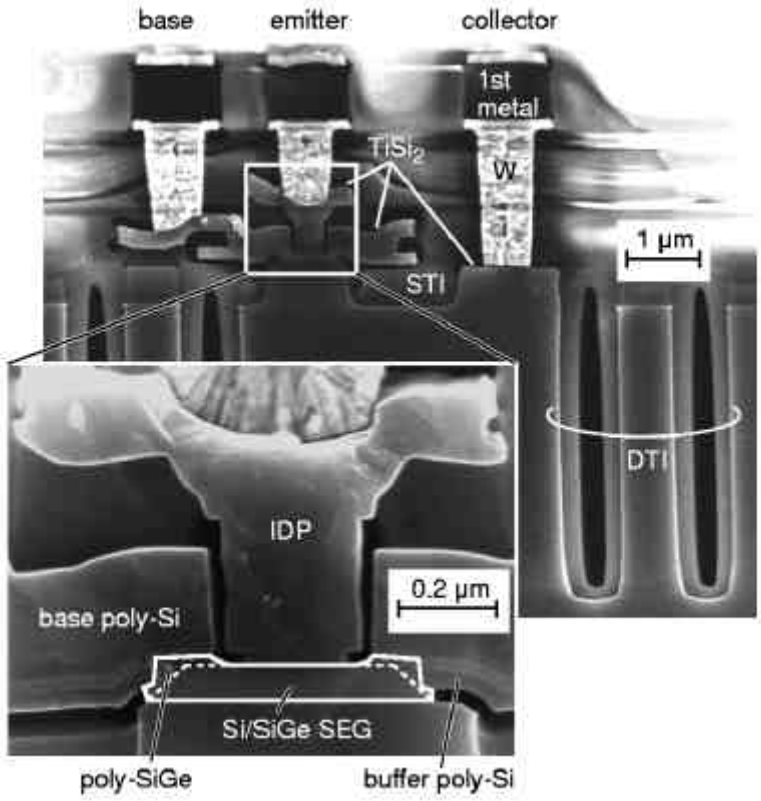



Figure 6.1: Cross-section of Si/SiGe bipolar transistor taken from [3]. Enlarged section shows regrown polysilicon emitter and base contact layers.

In a SiGe bipolar process, multiple selective area and/or non-selective area regrowths are used to define active junctions and polycrystalline contacting areas. By their nature, these additive processes allow tight control of lateral dimensions and the flexibility to add low $e_r$ dielectrics between contacting layers and the active junctions. The inset of Figure 6.1 demonstrates how a submicron (0.2 µm) emitter-base junction may be formed with a larger extrinsic emitter contacting area through selective area regrowth. In contrast, III-V process flows rely on subtractive etch processes to remove semiconductor layers in a top down fashion. Controlling

junction dimensions to submicron scales requires precise control of the lateral etch rate of the semiconductor, a challenge given that most III-V HBT processes rely on some wet chemical etching. Additionally, the top down process flow makes scaling of any underlying features progressively more difficult.

Si/SiGe bipolar processes also make use of ion implantation to control the doping profile in the collector region and to selectively implant the $n^+$ subcollector region. The collector implant allows the thickness of the collector depletion region to be carefully controlled. Ion implantations are seldom used in an InP HBT process flow, although isolation implants are sometimes used to produce high resistivity regions in GaAs HBT process flows.

The Si/SiGe process flow also makes use of trench isolation etches with dielectric refills for device isolation and parasitic reduction. The trench process consists of a dry semiconductor etch with a well-controlled sidewall profile, followed by a conformal coating of a dielectric ($SiO_2$ or $Si_3N_4$) to fill the trench. A CMP process is then typically used to planarize the dielectric back to some desired thickness. The deep trench isolation (labeled DTI in Figure 6.1) provides electrical isolation between devices and is used to reduce the parasitic substrate capacitance. The trench isolation processes allows the SiGe transistor to maintain a low vertical profile. This is important since process planarity is critical for the large scales of integration and multi-level interconnect formation. In an InP HBT process, a mesa-isolation etch is generally used to isolate transistors. Typical mesa-HBTs fabricated at UCSB have a vertical profile (from the bottom of the mesa to the top of the emitter

metal) of ~1.6 μm. However, at the levels of integration currently needed for InP HBTs, the lack of a more planar process flow is not likely to effect process yield. The shallow trench isolation (labeled STI in Figure 6.1) is used to reduce the parasitic collector-base capacitance lying under the base contacts. Similar approaches have been used for III-V HBT processes [4]. However, these approaches rely on an RIE etchback rather than a CMP etchback to planarize the dielectric.

The advanced process technologies used in SiGe HBTs gives them a clear advantage over III-V devices in achievable levels of integration and allows them to be competitive with III-V devices in digital circuit speed despite disadvantages in material properties. This discrepancy has led to initiatives to improve the underlying technology behind InP-based HBTs and develop process flows that more closely resemble those of SiGe devices [2]. At UCSB, work in this area is taking place on a number of fronts. Non-selective area regrowth of polycrystalline n-type and p-type material has been demonstrated in the InP system using molecular beam epitaxy [5,6]. These materials may be used to form low access resistance extrinsic contacts to the emitter and base layers in a SiGe-like process flow. MBE regrowth has also been used to form submicron emitter-base junctions, and RF devices have been realized utilizing this process [7]. At the time of this writing, further improvements in regrowth interface quality are necessary to minimize device leakage currents. Ion implantation of InP layers has also been investigated to improve HBT process flows. A collector pedestal implant process mirroring that of a SiGe HBT would dramatically reduce the extrinsic collector-base capacitance of the transistor.

Additionally, a silicon dopant implantation of the subcollector region, or an iron isolation implantation around the subcollector may be used to provide a planar device isolation technology.

The aforementioned processes represent significant departures from a standard mesa-HBT process flow. In this chapter, more evolutionary improvements in a mesa-HBT process flow are described. These improvements include: the use of a thin emitter epitaxial structure to facilitate submicron junction formation, the use of dielectric sidewall spacers and refractory metals to define a self-aligned base Ohmic contact and the use of a deep trench etch and dielectric refill to reduce the extrinsic collector-base capacitance associated with the base contacting area. In addition to reducing the device parasitics, these process improvements were designed to improve the yield and manufacturability of submicron devices. It is further hoped that some of these approaches may later be combined with the more aggressive SiGe-like processes described earlier. Prior to considering the HBT process flow, the delay terms in an emitter-coupled logic gate are considered to determine the transistor parasitics that limit mixed-signal and digital logic speed.

## *6.1 Logic Gate Delay Analysis*

As will be shown in the analysis below, traditional transistor figures-of-merit $f_t$ and $f_{max}$ do not accurately predict the maximum clocking frequency of a digital logic gate in a given technology. A hand analysis of a logic gate propagation delay is complicated by the large signal switching transients that must be considered. A linearized model of the switching events is necessary to simplify the analysis.

Numerous analyses of emitter-coupled logic (ECL) gate delay have been performed for InP-based HBTs fabricated at UCSB [8, 9,10,11]. The results of the delay analysis presented in [11] are repeated here with particular attention paid to the transistor parasitics that dominate the delay expressions.

A static frequency divider configured in a divide-by-two configuration is used as a metric for evaluating the digital logic speed in a transistor technology. A circuit schematic for an (ECL) master slave flip-flop is shown in Figure 6.2. In a divide-by-two configuration, the outputs of the flip-flop are fed back to the inputs with the opposite phase so that output changes with every rising clock edge. The critical delay paths for the circuit have been marked with arrows. For the circuit to operate correctly in a divide-by-two configuration, it is assumed that the inputs to the differential pairs in the upper stages must pass the switching point (50%) by the time the differential pairs in the lower stages switch. Therefore, the delays associated with the lower level emitter followers, which are common to all clock inputs, may be neglected. Under these assumptions the delay over the paths outlined in Figure 6.2, represents one-half of the delay for the maximum operating frequency of the divider.

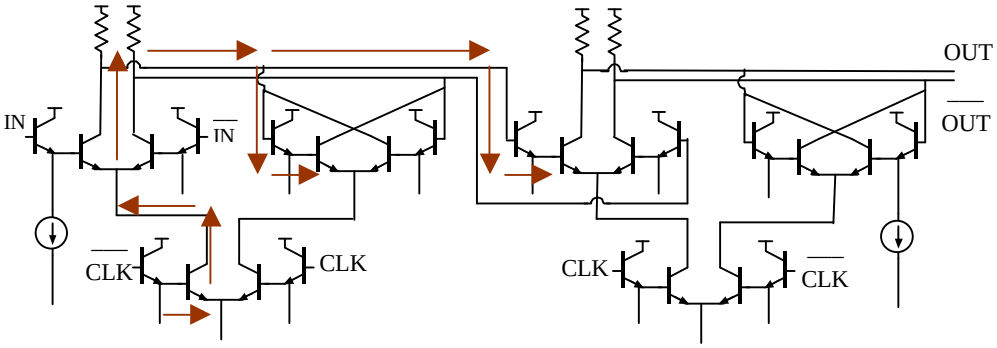



Figure 6.2: Circuit schematic of master-slave flip flop. Arrows represent critical delay paths for divide-by-two operation.

The analysis of the gate delay is performed using the charge control method, where delays associated with the charging time of each node in the signal path are summed. The charging time for each node is determined by the amount of charge ($Q$) necessary to switch the node from the initial to switching state (50%) and by the current charging the node ($I$) such that $t = \Delta Q/2I$ [12,13]. Using this linearized approach, node impedances are given by $R = \Delta V/\Delta I$ and $C = \Delta Q/\Delta V$. It should be noted that this analysis ignores terms that vary to second order in frequency, and is therefore only approximate.

To proceed with the analysis, a logic voltage swing must be specified and HBT model parameters must be determined. The transistor logic swing ? $V_L$ must be sufficiently large to provide noise margin for proper operation. The current ratio between two switching transistors in a differential pair is determined by the internal $V_{be}$ of the transistors, and the ratio between the on an off-state currents of the differential pair is proportional to $e^{\frac{qV_{be}}{KT}}$. For the analysis that follows, it is assumed

that $\Delta V_{be,\text{int}} = 6kT/q$ will provide sufficient noise margin for proper logic operation. In addition to the having a sufficient $\Delta V_{be,\text{int}}$, it is important that the latch also exhibit a finite voltage gain so that logic levels can be regenerated. In the presence of an extrinsic emitter resistance, the latch gain is approximated as $A_v \approx R_L / R_{ex}$, where $R_L$ is the ECL load resistance. Circuit simulations show that a latch gain of ~4 is sufficient for proper divider operation. Therefore, the selected logic swing is given by

$$\Delta V_L = 6kT/q + 4I_o R_{ex} = 6kT/q + 4J_o r_e \qquad \text{Eqn. 6.1}$$

where $I_o$ is the switched current, $J_o$ is the current density normalized to the emitter area, and $?_e$ is the normalized emitter resistivity.

The transistors are modeled by a hybrid-p transistor model like that shown in Figure 2.5. Linearized elements are used to describe the transistor input capacitance and transconductance. The emitter junction capacitance is determined by the average capacitance over a logic switching cycle such that

$$c_{je} \equiv \frac{\Delta Q}{\Delta V} = \frac{1}{\Delta V}\int_{V_{be,on}-\Delta V}^{V_{be,on}} c_{je}(V)dV \qquad \text{Eqn. 6.2}$$

where $V_{be,on}$ is the emitter-base turn on voltage, and $c_{je}$ is the emitter-base junction capacitance normalized with respect to junction area such that $C_{je} = c_{je} * A_e$.

The base-collector junction is assumed to be fully depleted across applied bias conditions, and the device is assumed to operate at a current density below the

Kirk threshold such that the base-collector capacitance may be assumed to remain constant during switching events. Additionally, it is assumed that the device has a fixed emitter to collector ratio such that the collector-base capacitance is proportional to the emitter area $C_{cb} = c_{cb} * A_e$. The transistor's large signal transconductance is given by $g_m = I_o / \Delta V$ and the large signal diffusion capacitance is therefore $C_{be,diff} = I_o (t_c + t_b) / \Delta V = I_o t_f / \Delta V$.

A current density of $J_o$ is assumed for the upper-level current switching HBTs, and a current density of $J_o/2$ is assumed for the lower-level clock switching current switching pairs and all emitter followers. The area of the upper-level differential pair transistors is denoted $A_e$, and the area of the lower level pairs is then $A_e/2$. It is found through SPICE simulations and hand analysis that circuit speed does not have a large dependence on the area of the emitter followers, and in this analysis the emitter follower are assumed to have an area of $2A_e$.

Using the charge control method and the transistor parameters described above, an approximate expression for the maximum operating frequency of the static frequency divider can be determined. The details of the derivation are described in [11] and are not repeated here since digital circuit design was not emphasized in this work. Instead the results of the analysis are considered, and the consequences on transistor design are analyzed.

Table 6.1 summarizes the delay coefficients ($a_{ij}$) determined from the charge control analysis. The total delay of the gate is determined by summing the products

| | $c_{je}$ | $c_{cb,x}$ | $c_{cb,i}$ | $t_f J_o / \Delta V_L$ | $t_{bus} J_o / \Delta V_L$ |
|---|---|---|---|---|---|
| $\Delta V_L / J_o$ | 1 | 6 | 6 | 1 | 1 |
| $kT/qJ_o$ | 0.5 | 1 | 1 | 0.5 | 0 |
| $?_e$ | -0.25 | 0.5 | 0.5 | 0.5 | 0 |
| $r_{bb}$ | 0.5 | 0 | 1 | 0.5 | 0 |

Table 6.1: Delay coefficients ($a_{ij}$) for master-slave flip-flop found by hand analysis. Maximum divide-by-two frequency is given by $T_{clock} = 1/2 f_{clock} = \sum a_{ij} r_i c_j$

of the delay coefficients, the charging resistances and the node capacitances such that $T_{clock} = 1/2 f_{clock} = \sum a_{ij} r_i c_j$ . The terms in Table 6.1 have been described in the text above with the exception of $t_{bus}$, which is the time of flight delay on the transmission line bus that connects the upper level collector nodes. In Table 6.1, the transistor transit time delay $t_f$ and the transmission line time of flight $t_{bus}$ are normalized with respect to the current density and logic swing to facilitate the description of the delay coefficients.

It is instructive to consider the terms in Table 6.1 in the context of a state-of-the-art InP HBT, the parameters of which will be described below. The device is assumed to have an emitter junction width of 0.5 μm and an emitter resistivity of 20 O- μm$^2$. The emitter depletion thickness is 200 Å . The base is 400 Å thick doped at 4 x $10^{19}$ cm$^{-3}$ and has 50 meV of compositional bandgap grading. A base contact resistivity of 30 O- μm$^2$ is assumed, and the base contacts are assumed to extend 0.5

μm on either side of the emitter metal. The extrinsic base-collector capacitance due to the base pad contacting area is assumed to be 50% of the total collector-base capacitance. This value is consistent with submicron mesa devices fabricated at UCSB. The collector thickness is 1500 Å, and the device is assumed to operate at the corresponding Kirk current density, $J_o$ = 4 mA/ μm$^2$. The transmission line delay $t_{bus}$ is assumed to be 0.6 psec, a value based on layouts of dividers fabricated at UCSB.

A logic voltage swing of ? $V_L$ = 500 mV is selected to satisfy Eqn. 6.1. By defining the voltage logic swing and the operating current density, the emitter length is determined by the selection of the load resistance. For high-frequency digital circuits, it is desirable to terminate long transmission lines with matched impedances to prevent reflections and ringing on the lines. For the divide-by-two circuits designed at UCSB, doubly terminated transmission lines are used to connect the collector nodes on the upper logic levels. The lines are terminated in their characteristic impedance, and the highest transmission line impedance realizable in an on-wafer environment is typically ~100 O. The effective load resistance for the latch will be the parallel combination of the two terminations, and therefore, $R_L$ = 50 O. The emitter length is then given by $L_e = \Delta V_L / W_e J_0 R_L = 4.75\mu\text{m}$.

For the transistor described above, the $f_t$ and $f_{max}$ predicted using the methods described in Chapter 2 are 288 GHz, and 367 GHz respectively. The relative contribution of the delay elements to the delay of the divide-by-two circuit can be determined using the coefficients in Table 6.1. The percentage contribution from

| | $c_{je}$ | $c_{cb,x}$ | $c_{cb,i}$ | $t_f J_o / \Delta V_L$ | $t_{bus} J_o / \Delta V_L$ | **Total** |
|---|---|---|---|---|---|---|
| $\Delta V_L / J_o$ | 15.6% | 21.2% | 29.7% | | | 66.5% |
| $\Delta V_L / J_o$ | | | | 10.0% | 14.2% | 24.2% |
| $kT/qJ_o$ | 0.4% | 0.2% | 0.3% | 0.7% | | 1.1% |
| $?_e$ | -0.7% | 0.3% | 0.4% | 2.7% | | 0.9% |
| $r_{bb}$ | 3.2% | | 2.0% | 5.1% | | 7.2% |
| **Total** | 18.5% | 21.6% | 32.4% | 13.1% | 14.2% | 100% |

Table 6.2: Percentage contribution of delay terms to the overall delay of static frequency divider using HBT parameters described in text. Predicted maximum frequency operation of divider is 158 GHz.

each of the terms is outlined in Table 6.2. The maximum operating frequency of the circuit is predicted to be 119 GHz for the HBT parameters described above. It should be noted that SPICE simulations typically predict a maximum operating frequency that is considerably less (on the order of 15-20%) than that predicted by hand analysis. However, the hand analysis does provide a good approximation of the relative contributions of the delay terms.

Examining the terms in Table 6.2 illustrates why the transistor figures-of-merit $f_t$ and $f_{maz}$ do a poor job of predicting the maximum operating frequency of the static frequency divider. The percentage contribution of the HBT transit time delays ($t_f = t_b + t_c$) to the divider delay are a relatively modest 13.1%. By contrast, the transit time delays contribute 76% to the total forward delay ($t_{ec}$) of the transistor that

determines the HBTs $f_t$ ($t_{ec}$ = 1/2p $f_t$). The discrepancy is explained by the observation that when operating in a digital circuit, the emitter-base diffusion capacitance that is determined by $t_f$ is reduced by the ratio of ? $V_L$ to *kT/q*, (10:1 for the example here).

The large signal voltage swings that reduce the contributions of the HBT transit times to the divider's delay, subsequently increase the contributions of the junction capacitances ($C_{cb}$ and $C_{je}$) through the delay terms $(\Delta V_L / I_o)C_{cb}$ and $(\Delta V_L / I_o)C_{je}$. These delay terms are minimized by operating the HBTs at high current densities. As discussed in Chapter 2, the maximum current density is determined by the onset of base pushout, and is increased by thinning the collector. If the switching transistors are operated at the Kirk limit, the delay term $(\Delta V_L / I_o)C_{cb}$ is given by

$$(\Delta V_L / I_o)C_{cb} = \frac{A_c}{A_e}\frac{\Delta V_L}{(V_{cb} + V_{bi})}\frac{T_c}{4v_{eff}} \qquad \text{Eqn. 6.3}$$

where Eqn. 2.4 has been used to express the maximum current density. Eqn. 6.3 shows that the delay term is minimized by thinning the collector thickness while increasing the operating current density, and that the delay term may be scaled independently of any lateral scaling of junction dimensions. However, if the collector is thinned without either concurrent lateral scaling of the collector-base capacitance, or improvements in Ohmic contact resistances, the magnitudes and relative contributions of the $R_{ex}C_{cb}$ and $R_{bb}C_{cb}$ delay terms will increase. Therefore, the ratios

of the emitter and base resistances to the load resistance ( $R_{ex}I_o/\Delta V_L$ and $R_{bb}I_o/\Delta V_L$ ) become good metrics for monitoring the relative contributions of these terms. For the HBT parameters described above, the $R_{ex}J_o/\Delta V_L$ and $R_{bb}J_o/\Delta V_L$ ratios are 0.17 and 0.41, respectively.

From Table 6.2, the contribution of delay terms containing the emitter resistance to the overall delay appears to be low (0.9%). However, the form of the coefficients has masked the contributions of $R_{ex}$ to the minimum logic swing ? $V_L$. In Table 6.3, Eqn. 6.1 has been used to replace the delay coefficient $(\Delta V_L/J_o)$ with $(6kT/qJ_o + 4r_e)$, and the percentage contributions of the delay elements have been recalculated. Note that the for the transit time and transmission line bus delays, the coefficient $(\Delta V_L/J_o)$ has not been replaced since these terms have been normalized with respect to the coefficient. The terms in Table 6.3 show a much larger contribution of terms containing the emitter resistance to the overall delay of the circuit (~46%).

The analysis presented above has indicated that the operating current density and collector-base capacitance are critical parameters in determining the bandwidth of a logic gate in state-of-the-art InP HBTs. The ratio of the emitter resistance and the base resistance to the load resistance are found to provide good metrics for monitoring the relative contributions of resistances that charge the junction capacitances. The emitter resistance in particular is found to play a key role in

| | $c_{je}$ | $c_{cb,x}$ | $c_{cb,i}$ | $t_f J_o / \Delta V_L$ | $t_{bus} J_o / \Delta V_L$ | **Total** |
|---|---|---|---|---|---|---|
| $\Delta V_L / J_o$ | | | | 10.1% | 14.3% | 24.4% |
| $kT/qJ_o$ | 5.4% | 6.9% | 9.7% | 0.3% | | 22.3% |
| $?_e$ | 9.9% | 14.7% | 20.6% | 0.8% | | 46.0% |
| $r_{bb}$ | 3.2% | | 2.0% | 2.0% | | 7.3% |
| **Total** | 18.5% | 21.6% | 24.5% | 13.2% | 14.3% | 100% |

Table 6.3: Percentage contribution of delay terms to the overall delay of static frequency divider using HBT parameters described in text, with coefficient $(\Delta V_L / J_o)$ replaced by $(6kT/qJ_o + 4r_e)$ for relevant terms.

determining logic gate speed when its contribution to the minimum logic swing is considered.

Given these observations, it is not surprising that Si/SiGe HBTs have been competitive with InP HBTs in digital logic speed. Si/SiGe HBTs have the following attributes that make them well suited for digital logic circuits: low extrinsic emitter resistance due to polysilicon emitter regrowth, low extrinsic collector-base capacitance due to collector pedestal implant and shallow trench isolation, and high operating current densities (~ 10 mA/μm$^2$) through the use of thin collector depletion regions. In turn, these attributes suggest areas where significant improvements can be made the design of InP HBTs.

## *6.2 Scaled Mesa-HBT Technology*

In this section, modifications to a mesa-HBT process flow are described that are designed to reduce device parasitics while improving the yield and manufacturability of the transistor. The process flow modifications include: the use of a thin emitter epitaxial structure to facilitate submicron junction formation, the use of dielectric sidewall spacers and refractory metals to define a self-aligned base Ohmic contact, and the use of a deep trench etch and dielectric refill to reduce the extrinsic collector-base capacitance associated with the base contacting area. Prior to considering the process improvements the HBT epitaxial structure is described.

### 6.2.1 Epitaxial Layer Design

A typical layer structure for the scaled mesa-HBTs fabricated in this work is shown in Table 6.4. The material used for the scaled mesa-HBT designs was grown by molecular beam epitaxy by commercial epitaxy vendor IQE. The layer structure is described from the top down. The emitter layer structure was designed to be compatible with submicron emitter-base junction formation. An anisotropic dry etch is desirable to eliminate lateral undercut during emitter mesa formation. However, as discussed in Chapter 2, carbon is the preferred dopant for the p-type InGaAs base, and carbon doping is incompatible with the methane/hydrogen/argon RIE that was used in the transferred-substrate process.

| Layer | Composition | Dopant | Thickness |
|---|---|---|---|
| Emitter Cap | InAs | Si: 2x $10^{19}$ $cm^{-3}$ | 150 Å |
| Emitter Cap | $In_{0.53}Ga_{0.47}$ As | Si: 2x $10^{19}$ $cm^{-3}$ | 100 Å |
| $N^+$ Emitter | InP | Si: 1x $10^{19}$ $cm^{-3}$ | 100 |
| $N^-$ Emitter | InP | Si: 8x $10^{17}$ $cm^{-3}$ | 250 |
| Base-Emitter Grade | $In_{0.45}Ga_{0.55}$ As / $In_{0.52}Al_{0.48}As$ | Si: 8x $10^{17}$ $cm^{-3}$ | 203 Å |
| Base-Emitter Grade | $In_{0.45}Ga_{0.55}$ As / $In_{0.52}Al_{0.48}As$ | C: 2x $10^{18}$ $cm^{-3}$ | 67 Å |
| Base | $In_{0.45}Ga_{0.55}As$ ? $In_{0.53}Ga_{0.47}$ As | C: 4x $10^{19}$ $cm^{-3}$ | 400 Å |
| Collector Setback | $In_{0.53}Ga_{0.47}$ As | Si: 2x $10^{16}$ $cm^{-3}$ | 200 Å |
| Base-Collector Grade | $In_{0.53}Ga_{0.47}$ As / $In_{0.52}Al_{0.48}As$ | Si: 2x $10^{16}$ $cm^{-3}$ | 240 Å |
| Collector Pulse Doping | InP | Si: 3x $10^{18}$ $cm^{-3}$ | 30 Å |
| Collector | InP | Si: 2x $10^{16}$ $cm^{-3}$ | 1030 Å |
| Sub-collector | $In_{0.53}Ga_{0.47}$ As | Si: 2x $10^{19}$ $cm^{-3}$ | 50 Å |
| Sub-collector | InP | Si: 2x $10^{19}$ $cm^{-3}$ | 3000 Å |
| Substrate | InP | Semi-Insulating | |

Table 6.4:Scaled-mesa HBT epitaxial layer structure.

Chlorine-based RIE chemistries are often used to etch InP-based compounds, and such etches will not result in hydrogen passivation. However, the reaction products from a $Cl_2$ RIE are non-volatile at room temperature and require a heated RIE plate [14]. The lack of such a tool at UCSB during most of this work precluded the use of a dry etch for the emitter mesa formation. However, a heated $Cl_2$-based

inductively coupled plasma (ICP) etch was used for the trench etch described later in this chapter.

To minimize lateral undercut during a wet emitter mesa etch, the emitter layers were kept very thin. The total emitter thickness was reduced to ~800 Å compared to a ~2500 Å thickness for the transferred-substrate HBTs. The emitter cap layer is typically formed with a heavily doped InGaAs layer to facilitate the formation of low resistivity emitter Ohmic contacts. In an attempt to further reduce the emitter contact resistance, a thin InAs emitter cap was used for some HBT designs. InAs has a narrower bandgap than InGaAs, and contact resistivities $< 1 \times 10^{-7}$ O-cm$^2$ have been obtained to thin strained InAs layers on InGaAs [15]. In the following chapter, device results will be presented that compare the extrinsic emitter resistance measured for a device with an InGaAs emitter cap versus a device with a composite InAs/InGaAs cap. The remaining emitter stack consists of a 100 Å thick InP layer doped at $2 \times 10^{19}$ cm$^{-3}$, followed by a 250 A InP layer doped at $8 \times 10^{17}$ cm$^{-3}$. A 280 Å chirped superlattice (CSL) InGaAs/InAlAs grade is used to remove the conduction band discontinuity from the wideband emitter to the base.

The InGaAs base is 400 Å thick nominally doped at $5 \times 10^{19}$ cm$^{-3}$ and contains ~50 meV of compositional grading to reduce base transit time. The selection of a 400 Å base thickness is motivated by the logic gate delay analysis presented in the previous section. As described in the analysis, the transistor transit time delays are found to contribute a relatively small fraction towards the logic gate delay. For a range of base thicknesses from 200-500 Å, the predicted divider speed is found show

less than 5% variation, as increases/decreases in the base transit time are offset by decreases/increases in charging times associated with the base resistance. For the same range of base thicknesses, a 35% variation in transistor $f_t$ is observed, further supporting the observation that transistor figures-of-merit may show poor correlation with digital logic speed. A thicker base is desirable to increase process tolerances. Potential issues with thin base transistors include: spiking of the base Ohmic contacts through to the collector, and the difficulty of stopping in the base layer when performing a non-selective etch of the base-emitter grade.

A great deal of effort at UCSB has gone into the design of the base-collector grade for double-heterojunction transistors. The grade used in this work was designed by M. Dahlström and design considerations are described in detail in [16,17]. The grade consists of a 200 Å InGaAs setback layer followed by a 240 Å InGaAs/InAlAs CSL, with a lattice period of 15 Å. A delta-doped InP layer at the end of the grade offsets the quasi-electric field introduced by the grade and delays the onset of the Kirk effect. It has been found that the doping of the delta-doped layer is critical for determining the depletion thickness of the collector region and must be controlled precisely. The total collector thickness for the mesa-HBTs reported in this work is 1500 Å.

A thin 50 Å InGaAs subcollector etch stop is followed by a thicker InP subcollector. Thermal considerations motivate the use of a minimal amount of InGaAs in the subcollector. The thickness of the InGaAs etch stop has been found to play a critical role in determining the thermal resistance of the transistor [18].

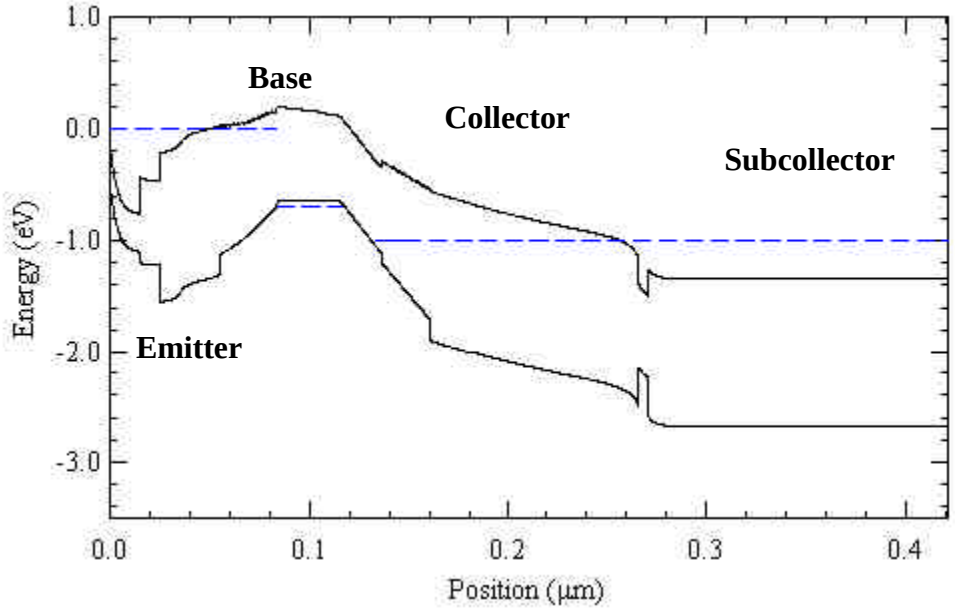



Figure 6.3: Band diagram of double-heterojunction mesa-HBT device with device epitaxy described by Table 6.4. Bias conditions: $V_{be}$ = 0.7 V, $V_{ce}$ = 1.0 V.

Subcollector thicknesses of 1500 Å or 3000 Å were used in this work, with overall device planarity motivating the use of thinner layers. A band diagram of the layer structure in forward active mode of operation is shown in Figure 6.3

### 6.2.2 Scaled mesa-HBT Process Flow

A top-down view of the mesa-HBT footprint is shown in Figure 6.4, and a cross-section of the device before device passivation is shown in Figure 6.5. The most obvious deviations from a standard mesa-HBT process flow are the use of dielectric sidewall spacers to form the self-aligned emitter-base junction and the use of a trench isolation to isolate the parasitic base-pad capacitance. The process improvements

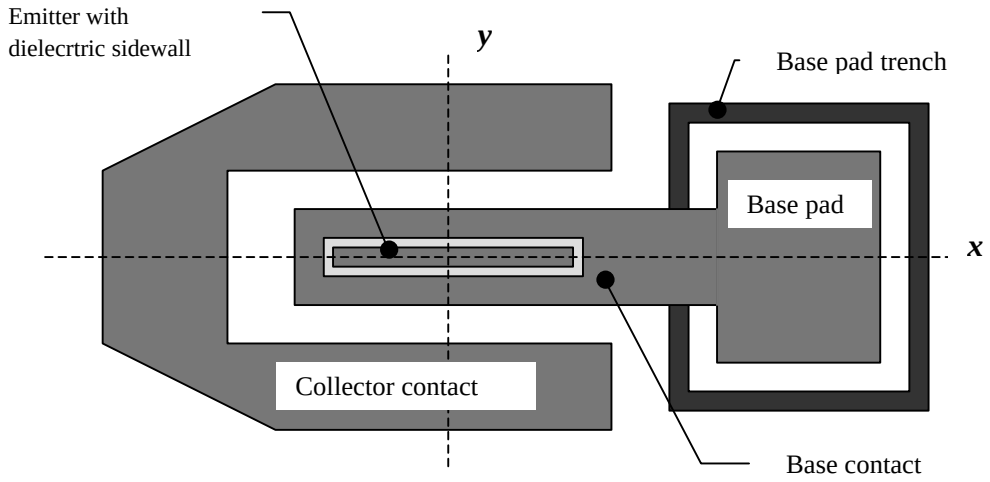



Figure 6.4: Top-down view of scaled mesa-HBT footprint

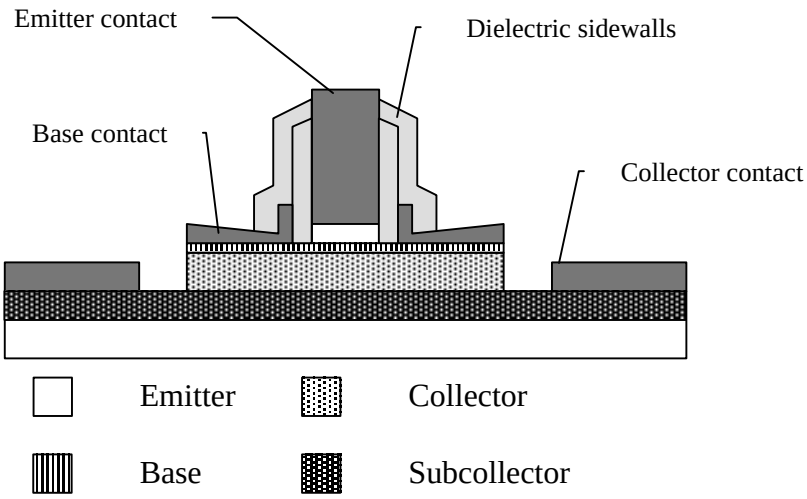



Figure 6.5: Cross-section of scaled mesa-HBT. Cross-section is taken in y-plane designated in Figure 6.4.

implemented in the scaled mesa-HBT process will be described in detail in the following sections.

### Emitter-base Junction Formation

The HBT process flow used to form the self-aligned emitter-base junction is shown in Figure 6.6. The process begins with the deposition of the emitter metal. The planarization/etchback process that is used to form the self-aligned base-contact requires a tall emitter post for process latitude. The metal liftoff process commonly used in III-V systems is ill suited for the formation of submicron features with large height to width ratios. Small features take on a trapezoidal shape due to metal deposition around the photoresist opening. As metal is deposited, the opening will close and this can result in submicron features being shorter than intended and in the formation of metal strands on the top of the feature. Some work was done to examine an etched emitter contact using a thick sputtered tungsten film. However, the etch profile of submicron features could not be well-controlled using the etch tools available at UCSB. Therefore, a liftoff process was used to form the emitter contacts for devices fabricated in this work. A Ti/Pt/Au/Pt emitter stack is used with thicknesses of 200Å/400Å/10,000Å/200Å. The top platinum layer is added to protect the gold during the base metal etchback process.

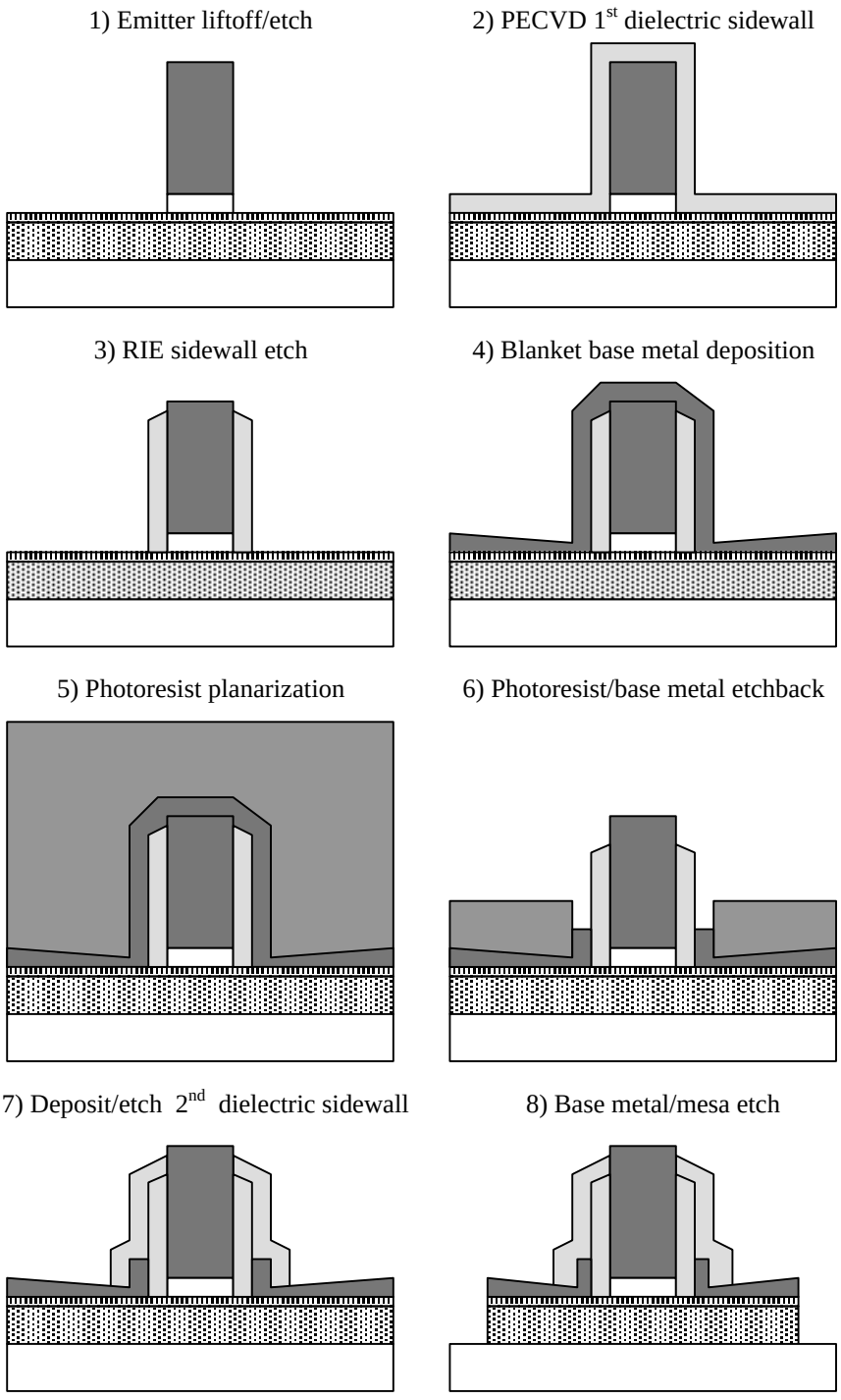



Figure 6.6: Process flow for emitter-base junction formation in scaled mesa-HBT technology.

A self-aligned emitter mesa-etch is performed using the emitter contact as the etch mask. The emitter cap is etched using a citric acid/hydrogen peroxide mixture, and the InP emitter is etched using a hydrochloric/phosphoric acid mixture. The thin emitter layers are found to be effective in controlling the lateral undercut of the emitter semiconductor, with ~0.05 μm of undercut being observed. The InP etch stops selectively in the base-emitter grade when a sufficient InGaAs composition has been reached.

A dielectric sidewall spacer is used to separate the emitter from the base contact. The sidewall spacer is formed using a conformal plasma-enhanced vapor deposition (PECVD) of a dielectric film followed by a reactive ion etch of the film (Figure 6.6 steps 2 and 3). The reactive ion etch is anisotropic with a slow lateral etch rate, and therefore, when the dielectric is etched clear in the field of the wafer a sidewall film will remain. Since the sidewall dielectric will passivate the semiconductor surface between the emitter and base contact, the interface properties will play an important role in determining leakage currents in the transistor.

A number of published works have reported on the surface passivation properties of PECVD deposited $Si_xN_y$ and $SiO_2$ films on InP-based materials. Unfortunately, a clear consensus regarding these properties has not been established. Degradation in HBT current gain when passivated with PECVD deposited films has been reported in [19, 20, 21, 22, 23]. In [19] and [20], severe degradation in HBT current gain was reported when InP/InGaAs HBTs were passivated with PECVD

deposited $Si_xN_y$ films. However, in [21] and [22] the current gain degradation was found to be less severe and more pronounced at low current densities. The deposition of SiO or $SiO_2$ films on the InGaAs base was found to cause a large increase in base leakage current in [19], [21], [22] and [23]. In [23], $SiO_2$ was deposited on a depleted InP ledge through which the base Ohmic contacts were annealed. This process is similar to the ledge process commonly used in GaAs HBTs [24] and was found to produce a low leakage device.

The increase in base leakage current when passivating with $Si_xN_y$ and $SiO_2$ films has been attributed to the pinning of the surface Fermi level along exposed surfaces [22]. If the surface Fermi level is pinned close to the conduction band, the accumulation of electrons near the surface will form a parasitic conduction path from the emitter to the base contact. Work in [25] suggests that the deposition of $Si_xN_y$ films tends to pull the surface Fermi level of InGaAs towards the conduction band creating a surface leakage path. This effect was found to be much less appreciable for InAlAs surfaces, with the surface Fermi-level being pinned closer to midgap.

The sidewall spacer process flow was motivated by published work regarding passivation with dielectric films. Rather than placing the dielectric on the InGaAs base, the sidewall is placed on the InAlAs-containing base-emitter grade. The grade is expected to be fully depleted and act as a passivation ledge. A $Si_xN_y$ film is also

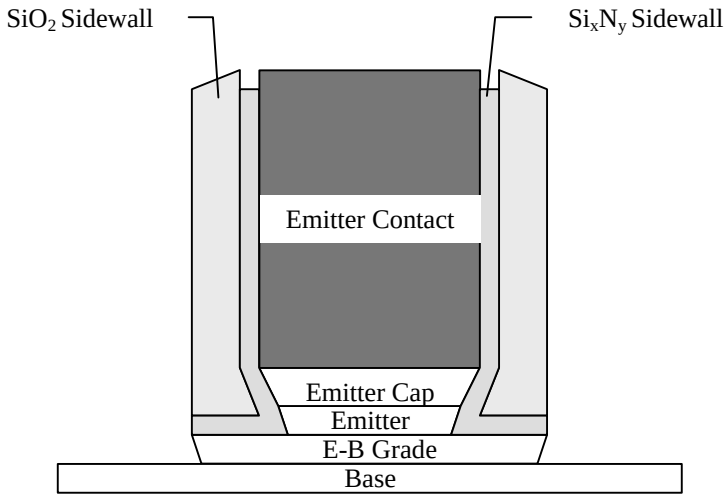



Figure 6.7: Detailed cross-section of emitter-base junction after step 3 in process flow of Figure 6.6.

placed in contact with the ledge rather than a $SiO_2$ film. The dry etch rate of $Si_xN_y$ films is significantly greater than $SiO_2$ films, and for the base metal etchback process it was important to keep the integrity of the dielectric sidewall. For this reason a double sidewall process was developed, where a $Si_xN_y/SiO_2$ (300Å/1000Å) layer was deposited. DC results that will be presented in the next chapter show that the passivation provided a relatively low leakage emitter-base junction. A detailed cross-section of the emitter-base junction with dielectric sidewalls is shown in Figure 6.7.

After deposition of the first dielectric sidewall, a citric-acid based etch is used to etch through the emitter-base grade into the base semiconductor. The etch rate is slow (~10 Å/sec) allowing the non-selective etch to be stopped in the thin base region. The deposition of the base Ohmic contacts is similar to the scheme reported in [4] for InGaP/GaAs HBTs. In [4], a WSi base Ohmic contact was blanket

deposited on the wafer and removed from the top and sidewalls of the emitter contact using a planarization/etchback process. It was found in [4] that a thin titanium layer inserted as the first layer in the base Ohmic contact dramatically reduced the base contact resistance. In this work, a Pt/W (30Å/ 1200Å) base Ohmic contact is used. The thin platinum layer is deposited using electron-beam evaporation, and the tungsten layer is then deposited in a magnetron sputtering system.

As discussed in Section 2.2, the insertion of thin palladium and platinum layers as contacting layers to the InGaAs base is found to produce excellent Ohmic contacts. Experiments performed at UCSB showed a Pt/Ti/Pt/Au p-type contact to produce contact resistivities comparable to the Pd/Ti/Pd/Au scheme described in Section 2.2 (~10 $\Omega$-$\mu m^2$). A palladium layer was not used in this work because of the difficulty in etching palladium with the fluorine-based RIE used to etch tungsten. In [26], the etching of palladium with a $SF_6$ RIE was found to produce a severe redeposition of a fluorinated layer. Attempts to etch thin palladium films in the fluorine-based RIE system at UCSB were unsuccessful. Similar problems were not observed for thin platinum films.

The sputter deposition of the tungsten film produces a conformal metal deposition that covers the top of the emitter contact and the dielectric sidewalls. To prevent base-emitter short circuits, contact between the base and emitter metal must be broken. This is accomplished using a planarization and etchback process, where a thick planarizing polymer is spun on to the wafer and etched back to expose the

emitter post. A fluorine-based RIE is then used to etch the tungsten from the top and sides of the emitter contact. Two critical parameters for the polymer are the degree of planarization it provides and the RIE selectivity versus tungsten. Ideally, the polymer would provide 100% planarization with 100% etch selectivity versus tungsten.

Engineered polymer films such as BCB or spin-on-glass are designed to provide good planarization properties over relatively large features. For example, BCB is specified to provide ~90% planarization over features < 100 μm. However, these dielectrics are generally etched in the fluorine-based plasmas used to etch tungsten films. The non-selectivity reduces the process margin during the etchback, and complicates the removal of the film after the etchback is completed. In this work, a thick (~4 μm) photoresist layer is used as a planarizing film. The photoresist was found to provide good local planarization over the submicron emitter features, although a precise value for the degree of planarization was not measured.

The photoresist is etched back to the top of the emitter post using an $O_2$ RIE. An $SF_6$/Ar etch is then used to etch the Pt/W base contact. The photoresist was found to have ~1:2 etch selectivity with respect to tungsten in the RIE process. In order to ensure a complete break between the base and emitter metal the photoresist is etched back to a final thickness of ~3000 Å. After completion of the etchback process, the photoresist is removed in acetone.

Even after the etchback process, strands of base metal may remain on the dielectric sidewalls of the emitter. This provides another potential source of base-emitter short circuits when contact is made to the emitter post. The emitter is

contacted using the same polyimide planarization and etchback process described in Chapter 2.2 for transferred-substrate HBTs. To prevent the Metal 1 contact to the emitter post from contacting any remaining base metal on the sidewalls, a second $SiO_2$ sidewall is deposited to cover the metal. The polyimide etchback is performed in an $O_2$ plasma that will not etch the $SiO_2$ sidewall.

After the etchback process, the base metal and base mesa are defined in a single photolithography step. The base metal is removed from the field using an $SF_6$/Ar RIE, with a laser interferometer used for etch endpoint detection. The same photoresist mask is then used for a wet etch of the base mesa. The base, collector setback and base-collector grade are etched in a hydrogen peroxide/phosphoric acid/ DI water mixture that stops selectively on the InP collector. The remaining InP collector is then etched using a hydrochloric/phosphoric acid etch that stops selectively on the InGaAs subcollector etch stop.

Figure 6.8 shows cross-sections of the transistor after emitter-base junction formation taken using the FEI Focused Ion Beam (FIB) system in the UCSB Materials department. The images show the effectiveness of the thin emitter semiconductor in controlling the lateral undercut and illustrate the dielectric sidewalls separating the emitter and base contacts. Issues related to device performance using this process will be considered in the following chapter.

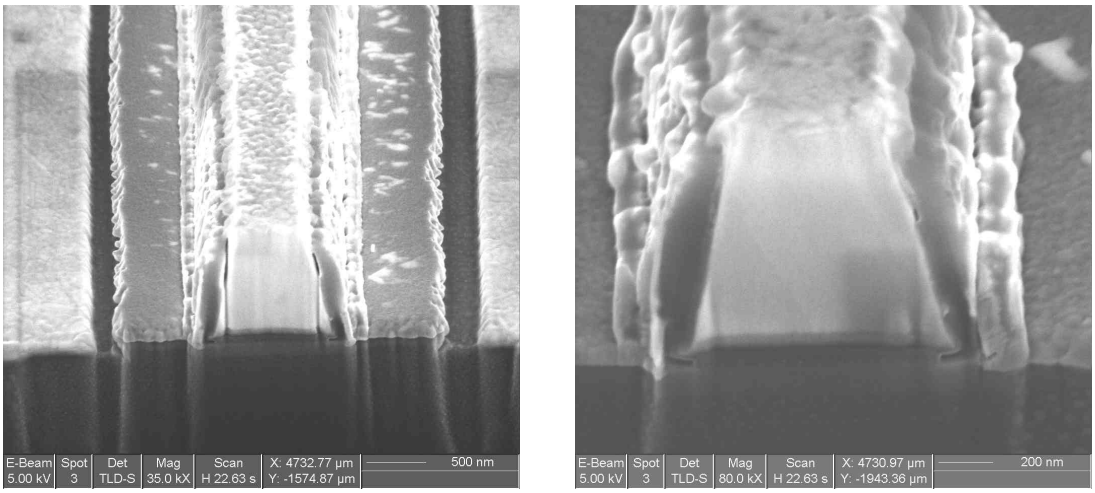



Figure 6.8: Cross-sections of scaled mesa-HBT emitter-base junction taken using Focused Ion Beam (FIB) system at UCSB.

### Base Pad Trench Isolation

In the digital logic analysis performed at the beginning of this chapter, the collector-base capacitance was found to be a critical factor in determining gate delay. The collector-base capacitance has contributions from the intrinsic capacitance lying under the base-emitter junction and the extrinsic capacitance lying under the base-emitter gap and base Ohmic contacts. In addition to the base contacts that run parallel to the device, an additional area of metal is required to make contact to the base metal. This area is denoted as the base pad in Figure 6.4, and the capacitance of the area may represent a significant fraction of the total collector-base capacitance for a submicron device.

The required size of the base pad depends on the contacting scheme and the required process tolerances. The pad may be contacted using a base post in which

case the size of the pad depends on the minimum size of the base post and the minimum separation between the post and the emitter. In this process, the base post is deposited to be the same height as the emitter post, so that it may be contacted using an etchback process after device passivation. Alternatively, a via through the passivating layer may be used to contact the base pad. In this case, the minimum pad size depends on the minimum via size and the necessary overlap with the underlying metal.

For the transferred-substrate devices described earlier in this work, a via through the polyimide passivation was used to contact the base pad. The minimum size for the via was 2 x 2 $\mu m^2$ and an overlap tolerance of 1 μm was used. In this case, the minimum size for the base pad is 16 $\mu m^2$. This value can be compared to the area lying under the base-mesa associated with the length of the emitter for the digital logic transistor described in Section 6.1. This device had an emitter length of 4.75 μm and a base mesa width of 1.5 μm, for a total area of 7.1 $\mu m^2$. For this transistor, the base pad would represent ~70% of the total collector-base capacitance.

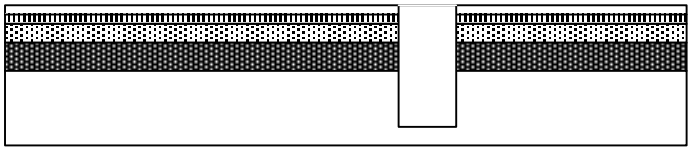
A base post process would typically have less stringent minimum feature size and alignment tolerances than a base via process. In the following chapter, mesa-HBT results will be described where the base post was defined using electron beam lithography. This process allowed the base post area to be scaled to ~1 $\mu m^2$ and to be placed within 0.5 μm of the emitter contact. The excess base pad area associated with the post was then ~2.25 $\mu m^2$, or ~24% of the total collector-base capacitance for the transistor described above.

As an alternative to scaling the size of the base pad, the contribution of the base pad to the base-collector capacitance can be eliminated if the capacitance is isolated from the active device. A number of approaches have been suggested to isolate the capacitance in III-V HBTs. One approach is to use a narrow base metal strip to connect the base mesa with the base pad [27, 28, 29]. The semiconductor beneath the strip is then undercut during the wet chemical base mesa and isolation etches so that it hangs as a suspended airbridge structure. By removing the subcollector region between the active device and the base pad region, the base pad capacitance does not contribute to the collector-base capacitance of the transistor. While effective in eliminating the base pad capacitance, this approach seems ill suited for large levels of integration.

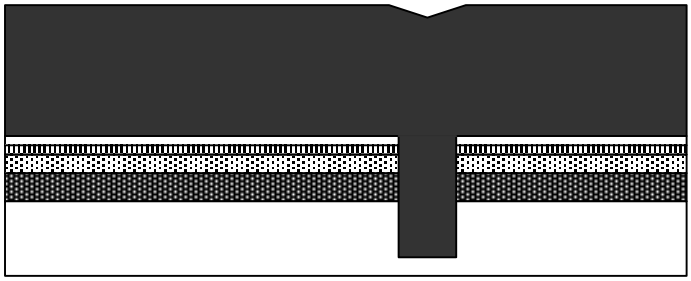
Another approach to eliminating the pad capacitance is to use a dielectric planarization process [4]. In this process, the base mesa etch is performed prior to depositing the base metal. A planarizing dielectric is then deposited and etched back to be level with the base semiconductor. The base metal can then be deposited to contact the base semiconductor with the extrinsic base pad area lying on top of the dielectric. In this approach, the semiconductor beneath the base pad has been replaced by a lower dielectric constant dielectric. The challenge in this process lies in controlling the etchback to the base semiconductor.

In this work, a trench isolation etch was used to isolate the base pad capacitance from the active device. The process flow for the trench isolation is outlined in Figure 6.9. The steps in the trench process flow begin before the

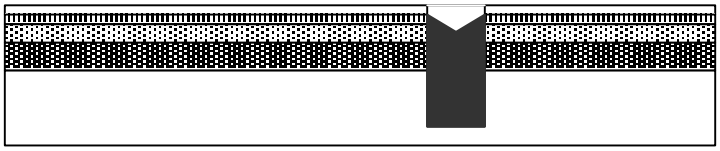
1) ICP Trench Etch



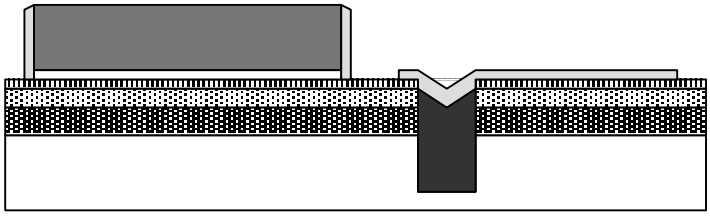
2) Dielectric Refill



3) Planarization Etchback



4) Emitter/sidewall definition w/ trench mask



5) Base contact definition

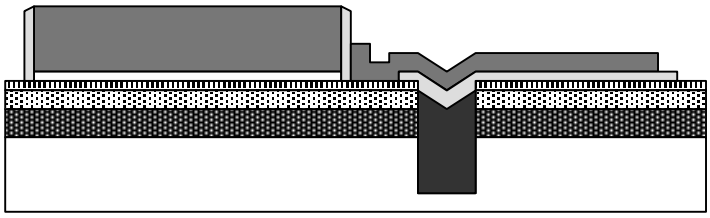


Figure 6.9: Process flow for base pad trench isolation in scaled mesa-HBT technology. Cross section is taken in x-plane of Figure 6.4.

deposition of the emitter contacts. As diagrammed in Figure 6.4 the trench etch surrounds the base pad area. A $Si_xN_y$ etch mask is used to define the trench, and the mask is defined using a photoresist mask and a $CF_4$ based RIE. For this work, a 0.8 μm wide trench was defined. The trench etch was performed using a Unaxis VLR inductively coupled plasma (ICP) system. The ICP tool has a heated substrate chuck that was set to 200°C for the etch. The high temperature is necessary to produce volatile $InCl_x$ compounds during the etch [14]. The etch is performed using a $Cl_2/N_2$ etch chemistry, and the etch conditions were found to produce a high etch rate (~1 μm/min) with a straight sidewall profile. To isolate the base pad capacitance, the trench etch must go through the subcollector and into the semi-insulating substrate. For the devices fabricated in this work, the trench etch depth was ~ 1.3 μm.

The trench was filled using a spin-on-glass (SOG) dielectric. Spin-on-glasses are silicon-based polymers that are commonly used for gap fill and planarization in silicon processes. The properties of the polymers can be engineered for specific applications. The spin-on-glass used in this work (Honeywell 512B) was a siloxane-based polymer designed to fill submicron gaps with high aspect ratios. Spin-on-glasses are designed to be compatible with silicon processes, and some process modifications were necessary for their application in a III-V HBT process.

The recommended cure temperature of the SOG was 425°C for 1 hour, a temperature that was beyond the capabilities of the ovens available in the UCSB cleanroom. Instead, a 6 hour 350°C cure was performed, and using this cure, the SOG film was found to be stable during the remaining temperature cycles of the

HBT process. In some initial process runs, the SOG was found to lose adhesion in the trench after the after the base metal deposition. The trench would lift causing a break in the base metal where it crossed the trench. Depositing an initial $SiO_2$ adhesion layer on the wafer before depositing the SOG solves this problem. The $SiO_2$ layer is removed from the field of the wafer during the planarization etchback.

The planarization etchback is performed using a $CF_4/O_2$ RIE. A laser interferometer is used to determine when the SOG and $SiO_2$ adhesion layer have been removed from the field of the wafer. An ~25% overetch was performed to ensure that the entire wafer surface was cleared. As a consequence of the overetch and the initial planarization profile, the SOG is recessed slightly below the top of the wafer surface. Since the emitter and base layers are thin, there exists a potential problem if the SOG is etched back past the base-collector interface. In this case, base metal that runs over the trench may contact the exposed collector semiconductor along the sidewall of the trench, creating a base-collector short circuit. This problem was avoided by adding a mask step during the emitter sidewall definition. Patterning the sidewall dielectric so that it remained covering the trench ensured that collector semiconductor would not be exposed within the trench. The base metal could then be safely deposited and patterned using the process described in the previous section.

An SEM image of an HBT after the base mesa etch with the base pad isolation trench is shown in Figure 6.10. Although the trench appears level with the

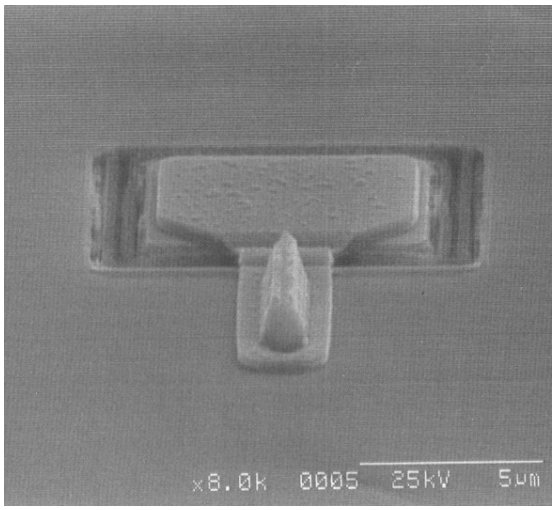



Figure 6.10: Scaled mesa-HBT after base mesa etch with base pad isolation trench.

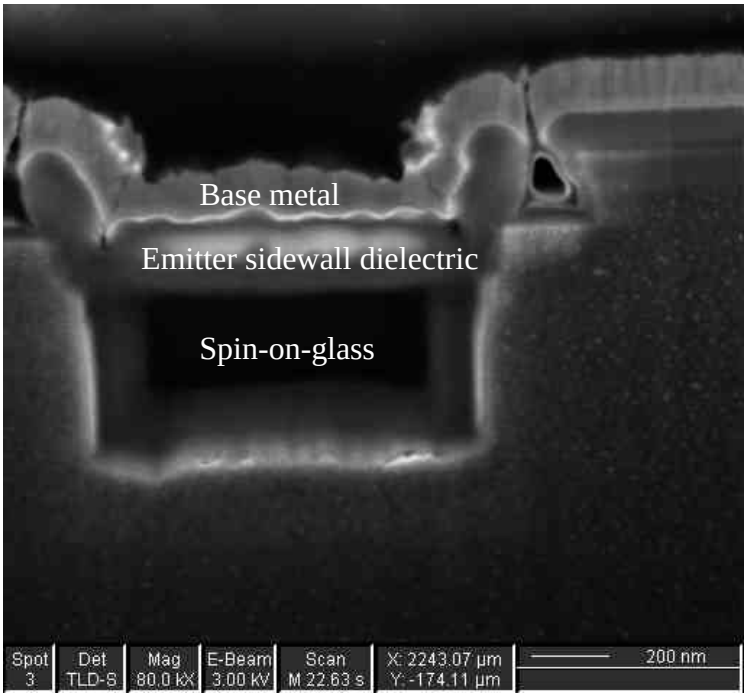



Figure 6.11: FIB cross-section of isolation trench showing base metal step coverage.

semiconductor surface, it has some topography associated with it due to the deposition of the $SiO_2$ adhesion layer and the emitter dielectric sidewall. The profile of the trench is better illustrated in the FIB cross-section of Figure 6.11. The cross-section shows potential problems with base metal step coverage over the trench. The development of the trench process was somewhat limited by the tools available in a university cleanroom environment, and it is expected that improvements in trench profile and etchback control could be obtained with improved dry etch tools.

### Mesa-HBT Process: Collector Contact to First Level Interconnect

After the base mesa etch, the remaining steps in the HBT process closely follow a standard III-V HBT process flow. A two-sided Ti/Pt/Au collector contact is deposited around the base mesa. The transistor is then electrically isolated by etching through the InP subcollector into the semi-insulating substrate. The device is passivated and planarized using the polyimide process described for transferred-substrate devices in Chapter 2. For device process runs, a single layer of metal interconnect is deposited.

## *6.3 Conclusions*

In this chapter, HBT design considerations for digital logic speed were presented. Due to the large signal switching characteristics of a digital logic gate, the charging of junction capacitances is found to have a larger relative contribution to the gate delay than to the transistor's traditional figures-of-merit. Conversely, the HBT transit times are found to have a smaller relative contribution. High current density

operation and low extrinsic emitter resistance were also found to be critical to increasing clock rates. Based on the requirements for a digital logic transistor, the process flow for a scaled-mesa HBT technology was presented. The technology used dielectric sidewall spacers to form a self-aligned base-emitter junction and an isolation trench etch to decrease the extrinsic collector-base capacitance. In addition to parasitic reduction, the process flow was also designed to increase the yield and manufacturability of a III-V HBT. In the following chapter, DC and RF results of HBTs fabricated using the scaled-mesa HBT process flow are presented.

## REFERENCES


1. J. Sitch, R. Surridge, “HBT ICs for OC192 Equipment,” *International Journal of High Speed Electronics and Systems*, vol. 9, no. 2, 1998, pp. 549-566.

2. J. C. Zolper, “Challenges and opportunities for InP mixed signal circuit technology,” *Conference Proceedings 2003 Indium Phosphide and Related Materials Conference*, May, Santa Barbara, CA, pp. 8-11.

3. K. Washio, “SiGe HBT and BiCMOS technologies for optical transmission and wireless communication systems,” *IEEE Transactions on Electron Devices*, vol. 50, no. 3, March 2003.

4. T. Oka, K. Hirata, K. Ouchi, H. Uchiyama, K. Mochizuki, and T. Nakamura, “Small-scaled InGaP/GaAs heterojunction HBT’s with WSi/Ti base electrode and buried ,” *IEEE Transactions on Electron Device Letters*, vol. 45, no. 11, Nov. 1998, pp. 2276-2282.

5. D. Scott, M. Urteaga, N. Parthasarathy, J.H. English, M. Rodwell, “Moleculay beam deposition of low-resistance polycrystalline InAs,” *Proceeding Lester Eastman Conference*, University of Delaware, Newark, Delaware, Aug. 6-8 2002.

6. Y. Dong, D.W. Scott, Y. Wei, A.C. Gossard, M. Rodwell, “Low resistance p-type polycrystalline GaAs grown by molecular beam epitaxy, *To be Published Journal of Crystal Growth.*

7. D. Scott, Y. Wei, Y. Dong, N. Parthasarathy, M. Rodwell, *Submitted IEEE Electron Device Letters*, November, 2002.

8. R. Pullela, *Digital Integrated Circuits in the Transferred-substrate HBT Technology*, Ph.D. Dissertation, University of California Santa Barbara, June 1998.

9. Q. Lee, *Ultra-high Bandwidth Heterojunction Bipolar Transistors and Millimeter-wave Digital Integrated Circuits*, Ph.D. Dissertation, University of California Santa Barbara, June 1999.

10. T. Mathew, *High Speed Digital ICs in Transferred-substrate HBT Technology*, PH.D. Dissertation, University of California Santa Barbara, August 2001.

11. M.J.W. Rodwell, M. Urteaga, Y. Betser, D. Scott, M. Dahlstrom, S. Lee, S. Krishnan, T. Mathew. S. Jaganathan. Y. Wei, D. Mensa, J. Guthrie, R. Pullela, Q. Lee, B. Agarwal, U. Bhattacharya, S.

Long “Scaling of InGaAs/InAIAs HBTs for High Speed Mixed-Signal and mm-Wave ICs”, *International Journal of High Speed Electronics and Systems*, Vol. 11, No. 1, pp. 159-215.

12. D.A. Hodges and H.G. Jackson, *Analysis and Design of Digital Integrated Circuits, 2nd Edition*, McGraw-Hill, 1983.

13. P.K. Tien, “Propagation delay in high speed silicon bipolar and GaAs digital circuits,” *Internation Journal of High Speed Electronics and Systems*, 1 (1), 1990, pp. 101-124.

14. J.E. Schramm, *Reactive Ion Etching of Indium-based Compounds Using Methane/Hydrogen/Argon*, Ph.D. Dissertation, University of California Santa Barbara, June 1995.

15. D. Mensa, *Improved Current-Gain Cutoff Frequency and High Gain-Bandwidth Amplifiers in Transferred-Substrate HBT Technology*, Ph.D. Dissertation, University of California Santa Barbara, Sept. 1999.

16. M. Dahlström, *Ultra High Speed InP Heterojunction Bipolar Transistors*, Ph.D. Dissertation, Royal Institute of Technology, Stockholm, Sweeden, 2003

17. M. Dahlström, X.-M. Fang , D. Lubyshev, M. Urteaga, S. Krishnan, N. Parthasarathy, Y. M. Kim, Y. Wu, J.M. Fastenau, W.K. Liu, M.J.W Rodwell, “Wideband DHBTs using a graded carbon-doped InGaAs base,” *IEEE Electron Device Letters*, Vol. 24, No. 7, July 2003, pp. 433- 435.

18. I. Harrison, M. Dahlström, S. Krishnan, Z. Griffith, Y.M. Kim, M. Rodwell, “Thermal limitations of InP HBTs in 80 and 160 GBit integrated circuits,” *Technical Digest Indium Phoshpide and Related Material Conference*, Santa Barbara, CA, May 12-16, 2002, pp. 160-163.

19. D. Caffin, L. Bricard, J.L. Courant, L.S. How Kee Chun, B. Lescaut, A.M. Duchenois, M. Meghelli, J.L. Benchimol, P. Launay, “Passivation of InP-based HBTs for high bit rate circuit applications,” *Conference Digest1997 Indium Phosphide and Related Materials Conference*, May, Cape Cod, MA, pp. 637-640.

20. A. Ouacha, M. Willander, B. Hammarlund, R.A. Logan, “Effect of surface passivation with SiN on the electrical properties of InP/InGaAs/ heterojunction bipolar transistors,” *Japanese Journal of Applied Physics*, vol. 74, no. 9, Nov. 1993, pp. 5601-5605.

21. H. Fukano, Y. Takanashi, M. Fujimoto, “Surface currents in InP/InGaAs heterojunction produced by passivation film formation,” *Japanese Journal of Applied Physics*, vol. 32, Part 2, no. 12B, Dec. 1993, pp. 1788-1791.

22. T. Kikawa, S. Takatani, H. Masuda, T. Tanoue, “Passivation of InP-based heterostructure bipolar transistors in relation to surface fermi level,” *Japanese Journal of Applied Physics*, vol. 38, Part 1, no.2B, Feb. 1999, pp. 1195-1199.

23. S. Driad, W.R. Mckinnon, S. Laframboise, S.P. McAlister, “Improved InGaAs/InP double-heterojunction bipolar transistors using a thin-emitter structure design,” *Microwave and Optical Technology Letters*, vol. 21, no. 4, May 1999, pp 235-238.

24. R. J. Malik, L.M. Lunardi, R.W. Ryan, S.C. Shunk, M.D. Feuer, “Submicron scaling of AlGaAs/GaAs slef-aligned thin emitter heterojunction bipolar transistors (SATE-HBT) with current gain independent of emitter area,” *Electronics Letters*, 25, 1989, pp. 1175-1177.

25. M. Arps, H.-G. Bach, W. Passenberg, A. Umbach, W. Schlaak, “Influence of $SiN_x$ passivation on the surface potential of GaInAs and AlInAs in HEMT layer structures,” *Conference Digest 1996 Indium Phosphide and Related Materials Conference*, May, Schwabisch-Gmund , Germany, pp. 308-311.

26. F. Fracassi, R. d’Agostino, A. Caccuci, “Dry etching of palladium thin films in fluorine containing plasmas: X-ray photoelectron spectroscopy investigation,” *Journal of Vacuum Science Technology A*, 13(1), Jan/Feb 1995, pp. 63-66.

27. M. Ida, K. Kurishima, N. Watanabe, “Over 300 GHz $f_T$ and $f_{max}$ InP/InGaAs double heterojunction bipolar transistors with thin pseudomorphic base,” *IEEE Electron Device Letters*, vol. 24, no. 12, December 2002, pp. 694-696.

28. M. Dvorak, C. Bolognesi, O. Pitts, S. Watkins, “300 GHz InP/GaAsSb/InP double HBTs with high current capability and $BV_{CEO}$ = 6V, *IEEE Electron Device Letters*, vol. 22, no. 8 August 2001, pp. 361-363.

29. W. Hafez, J-W. Lai, M. Feng, “Vertical scaling of 0.25μm emitter InP/InGaAs single heterojunction bipolar transistors with $f_T$ of 452 GHz,” *IEEE Electron Device Letters*, vol. 24, no. 7, July 2003, pp.436-438.

# Chapter 7 Scaled Mesa-HBT Results

In this chapter, measurements of devices fabricated in the scaled-mesa HBT technology are presented. Results are presented for two types of devices. One set of devices utilized the base pad trench isolation for extrinsic collector-base capacitance reduction. For these devices, a majority of the processing was performed using the UCSB cleanroom facilities with the exception of the base metal planarization/etchback step, which was performed in the Rockwell Scientific Company cleanroom. For the second set of devices, a majority of the fabrication was performed at Rockwell Scientific, with the exception of the base metal deposition, which was performed at UCSB. For these devices, a submicron base post written by electron beam lithography was used to reduce the extrinsic collector-base capacitance. In the following sections, relevant DC and RF parameters are presented and particular attention is paid to the device parasitics that were determined to be important for digital logic speed in Chapter 6.

## *7.1 DC Device Results*

Figure 7.1 shows the DC-IV common-emitter characteristics of a submicron HBT fabricated at UCSB. The device layer structure is the same as that described in Table 6.4, and the device had emitter junction dimensions of 0.6 x 6 $\mu m^2$. The device is

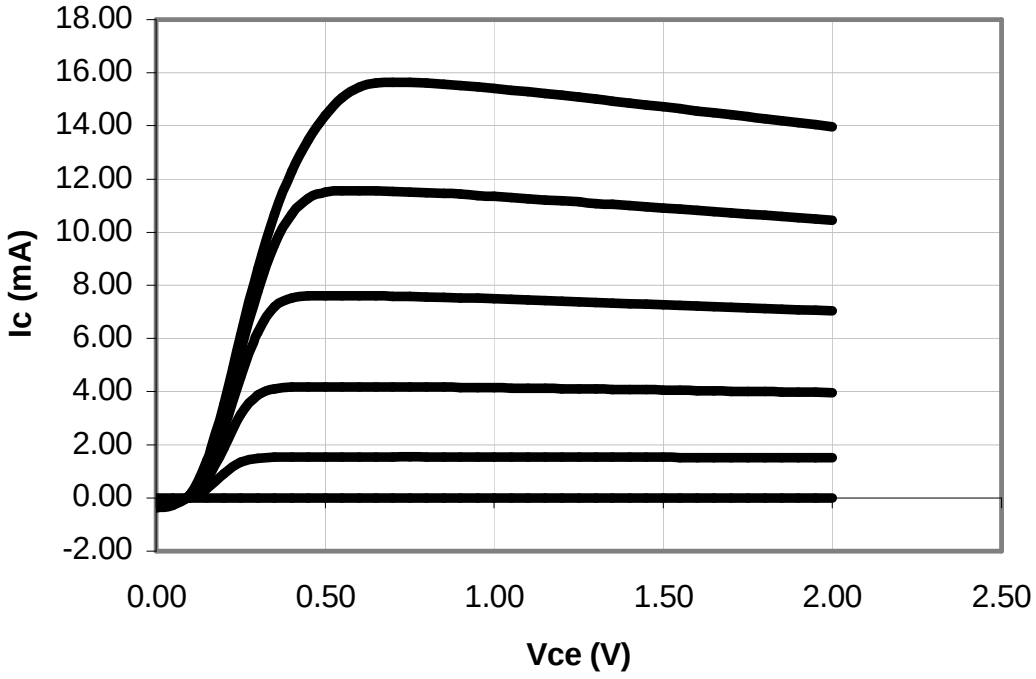



Figure 7.1: High current density common-emitter IV characteristics for scaled mesa-HBT. Emitter junction dimensions 0.6 x 6 $\mu m^2$.

found to have a DC current gain of ~50 with an open circuit common emitter breakdown voltage ($BV_{CEO}$) of > 6V. Importantly for digital logic applications, the transistor is capable of high current density operation with a relatively low collector offset voltage.

The common-emitter Gummel characteristics for the transistor are shown in Figure 7.2. The transistor is observed to have collector ideality factor ($n_c$) of ~1. The base ideality factor ($n_b$) is found to have a value of ~2 in the current range of 2 nA to 100 nA, and an ideality factor of ~1.5 in the current range of 500 nA to 50 µA. The collector base crossover current is measured to be ~ 500 nA.

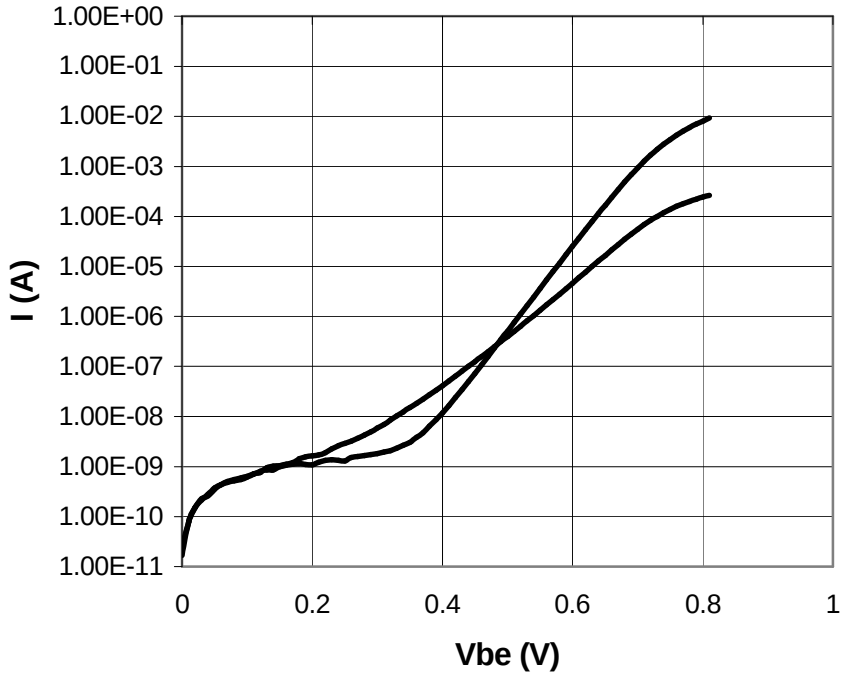



Figure 7.2: Common-emitter Gummel characteristics of scaled-mesa HBT fabricated at UCSB. Emitter junction dimensions 0.6 x 6 $\mu m^2$.

For a graded base-emitter junction, the base ideality factor is expected to be ~1 since the junction interface does not have an abrupt material transition that may give rise to recombination centers. For transferred-substrate devices with an InAlAs emitter and a graded base-emitter junction, base ideality factors of ~1 were generally observed. The $n_b$ = 1.5 region observed in the Gummel characteristics of Figure 7.2 is attributed to the base-emitter grade design from the InP emitter to the InGaAs base. Large area (~30 x 30 $\mu m^2$) HBTs fabricated using the same epitaxial design showed a base ideality factor of ~1.4, indicating the ideality factor is due to bulk and not surface recombination. Since the focus of this work was on process improvement and not epitaxial design, factors determining the base-emitter grade characteristics

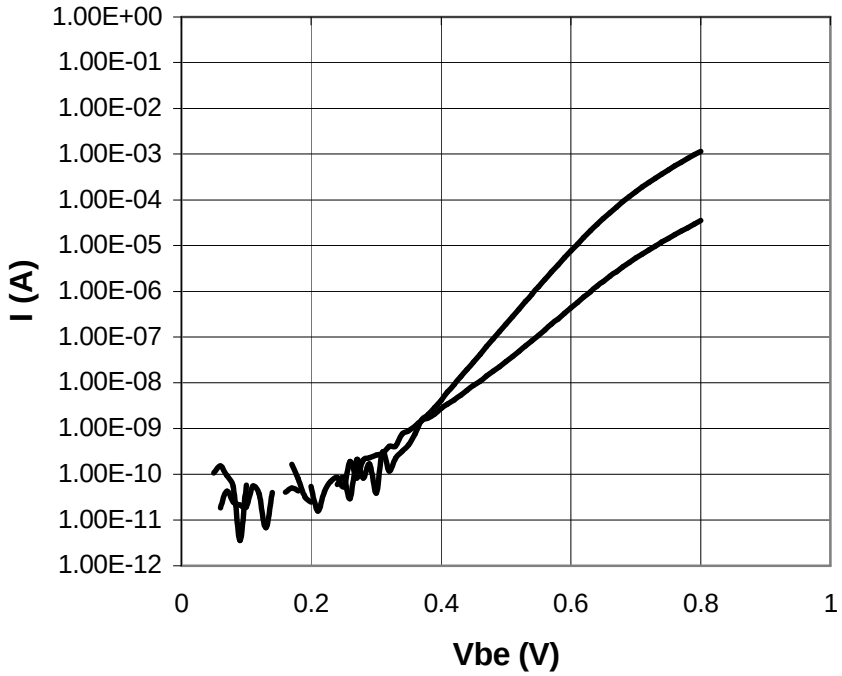



Figure 7.3: Common-emitter Gummel characteristics of scaled-mesa HBT fabricated at Rockwell Scientific. Emitter junction dimensions 0.7 x 3 $\mu m^2$.

were not considered in detail. However, devices fabricated at UCSB by M. Dahlström using a modified base-emitter grade have since shown an ideality factor of ~1.1 [1].

The Gummel characteristics of Figure 7.2 also show a low current $n_b \sim 2$ region and a relatively high base collector crossover current. It is believed that these observations may be related to the base-emitter passivation scheme. As discussed in the previous chapter, conflicting published literature exists regarding the passivation of InP HBTs with PECVD deposited dielectrics. It is therefore expected that the passivation may depend critically on the surface preparation and deposition conditions that are used. Figure 7.3 shows the Gummel characteristics for a

transistor fabricated at Rockwell Scientific using the same dielectric sidewall passivation scheme and the same base-emitter epitaxy design as the device in Figure 7.2. For this transistor, the base collector crossover current has been reduced to ~ 2 nA, and the base ideality factor is found to be ~ 1.5 over the entire low current region. The results indicate the effectiveness of the sidewall dielectric passivation process and show that the process parameters influence this effectiveness. A study into how the process parameters (surface preparation, PECVD conditions) affect the base-emitter junction characteristics was not considered in this work.

## *7.2 Base Emitter Diode Yield*

The scaled-mesa HBT process was intended to increase the yield and manufacturability in a III-V HBT process flow. Reducing the process to practice was found to be a challenge given the resource constraints in a university cleanroom environment. The planarization/etchback process used to form the self-aligned emitter-base junction was found to produce good yield on a local level. Figure 7.4 shows the IV characteristics of 360 parallel base-emitter diodes formed using the sidewall process. In some cells on the same sample, almost all devices showed base-emitter short circuits. Non-uniformity in the planarization and etchback steps was the likely cause of the yield variations.

Two competing mechanisms resulted in non-uniformity in the base-emitter etchback process. The devices described in this work were fabricated on pieces (~ ¼ 's) of 3" and 4" inch wafers. The asymmetric shape of the substrates resulted in edge beading when the photoresist planarization was spun on the wafers, with the

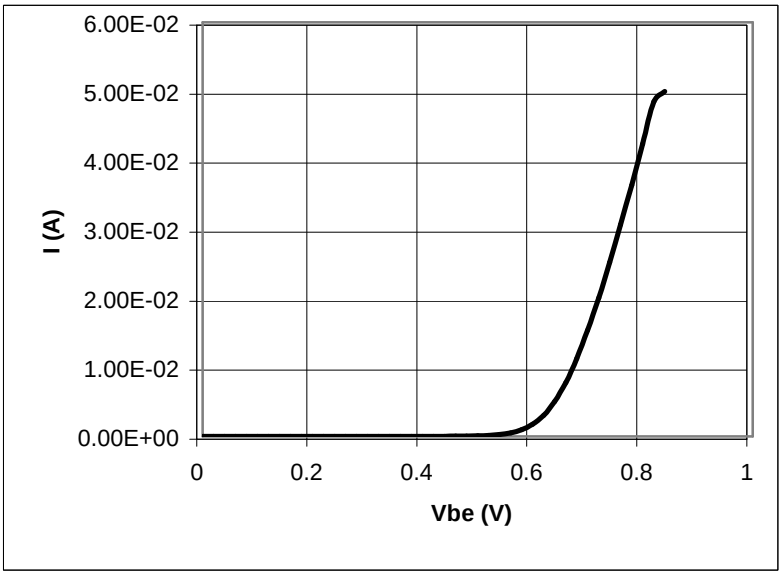



Figure 7.4: IV characteristics of 360 parallel base-emitter diodes fabricated using self-aligned sidewall spacer process. Emitter junction dimensions are 0.7 x 3 $\mu m^2$.

thickness of the photoresist increasing near the edges. The uniformity of the resist spin-up would be expected to improve if full wafers were used in the process.

In an RIE process, etch rates tend to increase near the edges and corners of a sample where the electric fields are larger. This effect was observed in the photoresist etchback process, where the photoresist was observed to clear in the corners of the wafer while a few thousand angstroms of resist remained in the center of the wafer. Some improvements in etch uniformity could likely be obtained by switching to rounded full wafers or by optimizing the RIE conditions.

Base-emitter diode yield was generally found to decrease towards center of the sample, indicating that the non-uniformity of the RIE dominated the process. The photoresist would typically be etched back to a thickness of ~3000Å in the center of

the wafer and a thickness of ~1000Å near the edges. The base-emitter short circuits are believed to be caused by strands of base metal remaining in contact with the top of emitter metal. This assertion is supported by the observation that many devices exhibited a leaky resistive path in parallel with the base-emitter diode that could be rendered open by increasing the base-emitter voltage. These results indicate the leakage path could not support moderate currents. An observation that is consistent with the presence of thin metal strands.

The etchback process latitude could be improved with a taller emitter contact, as the base metal in contact with the emitter would be exposed to the RIE for a longer time for the same final etchback thickness. Unfortunately, the liftoff process used to define the emitter contact limited the height of the metal for submicron features. A second possible approach to improving the yield of the step would be to incorporate a second planarization and etchback step into the process. This again would increase the exposure of the base metal on top of the emitter metal to the RIE.

While the self-aligned base-emitter junction process showed problems with global yield across a sample, the good local yield obtained indicates that the process is viable for high volume applications if issues with etchback uniformity and process latitude are addressed.

### *7.3 Contact Resistances*

As described in Chapters 2 and 6, the Ohmic contact resistances play a critical role in determining the high frequency performance of submicron HBTs. In this section,

evaluations of resistive terms for the extrinsic emitter, base and collector resistances are presented.

### 7.3.1 Emitter Resistance

The emitter resistance of fabricated HBTs was evaluated using the emitter flyback technique [2]. In this measurement, a base current is forced into the device with the collector held open. When the base-emitter diode is turned-on, the internal base-emitter voltage will remain approximately constant. The emitter resistance can then be extracted from the gradient of the forced current ($I_E$) plotted versus the collector-emitter voltage ($V_{CE}$).

Using the flyback method, the emitter resistance of fabricated HBTs was determined. As discussed in Chapter 6, the use of a thin InAs emitter cap layer was examined for reducing emitter contact resistance. Table 7.1 compares the emitter resistance extracted from two sets of submicron mesa-HBTs. The transistors had identical epitaxial designs except for the emitter cap layer. One set of devices had a 250Å InGaAs emitter cap doped at $2x10^{19}$ $cm^{-3}$, and these devices were found to have an average emitter resistivity of ~30 $\Omega$-$\mu m^2$. The other set of devices had a InAs/InGaAs emitter cap (150Å/100Å) doped at $2x10^{19}$ $cm^{-3}$, and these devices were found to have an average emitter resistivity of ~23 $\Omega$-$\mu m^2$. The improvements in emitter resistance are attributed to the low contact resistance to the InAs layer. For both emitter cap stacks, the emitter resistivity scales with emitter width an indication that the lateral undercut of the contact has been well-controlled.

| | Emitter Cap: InAs/InGaAs (150Å/100Å) | | Emitter Cap: InGaAs (250Å) | |
|---|---|---|---|---|
| Emitter Dimensions | $R_{ex}$ (Ω) | $?_e$ (Ω-μm$^2$) | $R_{ex}$ (Ω) | $?_e$ (Ω-μm$^2$) |
| 0.5 x 6 μm$^2$ | 10.5 | 22 | 7.4 | 32 |
| 0.6 x 6 μm$^2$ | 8.4 | 24 | 6.7 | 30 |
| 0.7 x 6 μm$^2$ | 7.2 | 23 | 5.4 | 30 |

Table 7.1: Emitter resistance extracted by emitter flyback method for HBTs fabricated with different emitter cap layers.

The results presented in Table 7.1 were for HBTs fabricated without the base pad trench isolation process. Devices fabricated using the trench process and the InAs/InGaAs emitter cap showed emitter resistivities of ~ 40 Ω-μm$^2$. The emitter contact is formed after the trench isolation etchback, and it is believed that damage or surface contamination from the etch resulted in the increased resistivity. The etch process conditions could likely be modified to improve the contact resistance. However, such improvements were not considered in this work.

### 7.3.2 Base Resistance

Scaled-mesa HBTs fabricated at UCSB suffered from a high base resistance, and subsequently a low transistor $f_{max}$. RF device results will be considered in the next section. Here, terms that contribute to the transistor base resistance will be considered and potential sources for the high base resistance observed in the devices will be investigated.

In the scaled-mesa HBT process flow, a tungsten metal stack is used to contact the base semiconductor. A base contact using only tungsten was evaluated and found to produce an extremely high contact resistance (> 1000 O-μm$^2$). As discussed in Chapter 6, a thin platinum layer was inserted beneath the tungsten to reduce the contact resistivity. The transmission line method (TLM) was used to determine the contact resistance of the Pt/W (30Å/1500Å) contact, and values in the range of 20-50 O-μm$^2$ were obtained. While not as low as those seen for evaporated Pd/Ti/Pd/Au contacts, this range of contact resistivities is suitable for fabricating high performance devices. For example, the digital logic transistor described in Section 6.1 assumed a contact resistivity of 30 O-μm$^2$.

In addition to the contact resistance, the metal sheet resistance of the base contact also affects transistor performance. In the scaled mesa-HBTs designs, the base-mesa width is aggressively scaled to reduce the extrinsic collector-base capacitance. The base metal resistance contributes an effective base resistance of $R_{b,metal} = r_{sheet} L_E / 6W_{b,metal}$ , where $?_{sheet}$ is the base metal sheet resistance, $L_E$ is the emitter length, and $W_{b,metal}$ is the width of the base metal on either side of the emitter contact. Bulk tungsten has a resistivity that is approximately four times worse than gold, and thin sputtered tungsten films typically have poorer resistivities than bulk tungsten. Using the magnetron sputter system at UCSB, bulk tungsten resistivities of ~25 μO-cm were obtained for thin tungsten films. This translates to a sheet resistance of ~2 O/square for a 1500 Å thick tungsten film. For the digital logic

transistor described in Section 6.1, this value of base metal resistance would contribute only 3 O towards the total base resistance of 24 O.

Based on the contact and metal sheet resistances, the Pt/W base metal stack appears well suited for realizing a high performance transistor. The base resistance in a fabricated transistor is better predicted by the measurement of pinched-TLM structures. In these structures (Figure 7.5), the emitter contact is left in the gap between TLM pads. The slope of the measured resistance versus TLM gap spacing gives the base semiconductor sheet resistance under the emitter. The intercept of the plot contains the resistive terms that remain constant for each gap spacing. These terms include the base contact resistance, the base semiconductor resistance under the dielectric sidewall and any contribution from the base metal resistance between the measurement probe and the gap (this term is expected to be small). Measurements of pinched-TLM structures fabricated at UCSB showed a much higher intercept than expected.

The contribution of the base contact resistance and base semiconductor resistance to the total base resistance can be normalized with respect to the emitter length. Assuming a base sheet resistance of 500 O, a contact resistivity of 30 O-μm$^2$ and a sidewall width of 1000Å, this normalized resistance is 55 O-μm for a single-sided base contact. If the base metal resistance is negligible, the intercept of the pinched TLM plot can be used to predict the same normalized resistance. Measurements of pinched TLM structures showed a normalized resistance of ~375 O-μm, approximately 7 times that predicted from the base parameters.

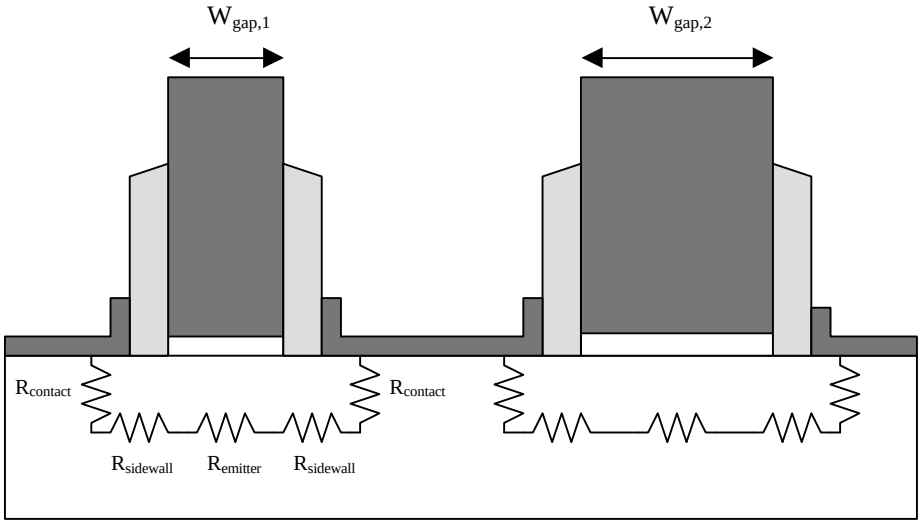



Figure 7.5: Schematic diagram of pinched-TLM structure with contributing resistive terms.

There was some concern that the deposition and/or removal of the dielectric sidewall may have contributed to the anomalously high base resistance. The PECVD of $Si_xN_y$ and $SiO_2$ films uses a silane ($SiH_4$) carrier gas, and the films are known to contain a high hydrogen concentration. As discussed in previous chapters, hydrogen is known to passivate carbon dopants in InGaAs. There was also concern that damage from the sidewall etchback may have resulted damage to the base, even though the sidewalls were deposited on the base-emitter grade. A TLM sample was prepared to determine whether the dielectric sidewall process caused any deleterious effects to the base resistance. The sample was exposed to the same deposition and etch conditions as in the dielectric sidewall process flow. Standard Pd/Ti/Pd/Au contacts were then deposited, and the TLM structures were formed. Negligible

differences in sheet and contact resistances were observed between the sample that underwent the sidewall process and a control sample that did not.

The suspected source of the high base resistance was determined by looking at cross-sections of transistors and TLM structures. The thickness of the as deposited tungsten base contact showed a dramatic thinning as it approached the tall emitter contact structures. The thinning of the base metal is worsened by the formation of the second dielectric sidewall (step 8 in Figure 6.6). As described in Section 6.2, the second sidewall is deposited to cover any base metal that may remain on the dielectric sidewalls. The sidewall etch uses a fluorine-based etch chemistry that will also etch the tungsten base metal. The sidewall etch is controlled using a laser interferometer and a slight overetch (~25%) is performed to ensure the dielectric is cleared from the top of the emitter contact. The laser interferometer monitors the etch in the field of the wafer, and the etch rate may be higher near the tall emitter features due to stronger electric fields. These factors would contribute to a reduction in the base metal thickness near emitter features. Figure 7.6 shows a FIB cross-section of the base-emitter junction of a fabricated transistor. The base metal shows a severe thinning near the emitter contact. This thinning was also observed in the TLM structures.

The thinning of the base metal increases the gap resistance between the emitter and the base contact and also increases the base metal sheet resistances. These factors are the likely cause of the anomalously high base resistance extracted from TLM measurements and of the low $f_{max}$ observed for mesa-HBTs fabricated at

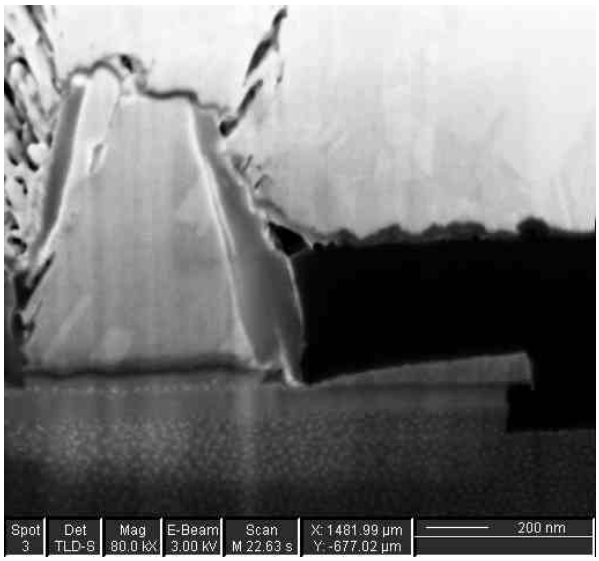



Figure 7.6: FIB image of scaled mesa-HBT showing thinning of base metal near emitter contact.

UCSB. Scaled mesa-HBTs fabricated at Rockwell Scientific had a second dielectric sidewall deposited after device passivation and planarization. In this process flow, the base metal is protected during the second sidewall etch. These devices showed a much higher $f_{max}$ than transistors fabricated at UCSB. RF device results from both types of devices will be presented in the following section.

### 7.3.3 Collector Resistance

As described in Chapter 6, a thin (50 Å) InGaAs layer was used in the subcollector for thermal management considerations. TLM measurements were performed to ensure that the collector contact resistance was sufficiently low despite the use of the thin InGaAs layer. A contact resistivity of 12 O-µm$^2$ was extracted from

measurements. A value that is similar to that has been obtained with thicker (>100 Å) InGaAs layers.

## *7.4 RF Results*

### 7.4.1 UCSB Fabricated Transistors

Mesa-HBTs fabricated at UCSB used the base pad trench isolation for extrinsic collector-base capacitance reduction. Device results are considered for a transistor with a nominal emitter junction width of 0.5 μm. For this device, the base-mesa width extended 0.5 μm on either side of the emitter contact. As described in the previous section, the transistors fabricated at UCSB suffered from two process problems: a slightly high extrinsic emitter resistance due the planarization etchback and an extremely high base resistance due to thinning of the base metal around the emitter contact

Figure 7.7 shows a plot of the unilateral power gain ($U$) and short circuit current gain ($h_{21}$) for a transistor with nominal emitter junction dimensions of 0.5 x 3 $\mu m^2$. The device is biased at a collector current $I_C$ = 9 mA, and a collector base voltage $V_{CB}$ = 0.4 V. Referenced to the emitter junction dimensions the device is operating at a current density $J_E$ = 6 mA/$\mu m^2$. The extrapolated $f_t$ and $f_{max}$ of the transistor are 247 GHz and 166 GHz, respectively. The base-collector capacitance of the transistor is extracted from the imaginary part of $Y_{12}$ at low frequencies and is found to be 6 fF. Operating at this bias condition, the transistor is found to have a low $C_{cb}/I_C$ ratio of 1 psec/V.

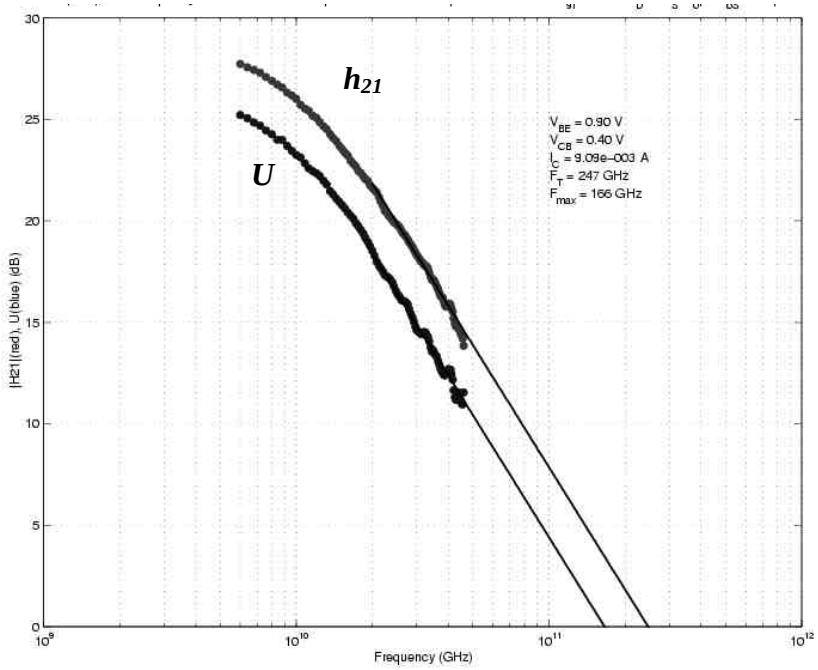



Figure 7.7: Unilateral power gain ($U$) and short circuit current gain ($h_{21}$) of submicron mesa-HBT fabricated at UCSB. Emitter junction dimensions: 0.5 x 3 $\mu m^2$.

The maximum operating current density, determined by the Kirk threshold, can be increased by applying a greater collector-base voltage (see Eqn. 2.4). In a standard ECL gate, the emitter followers may operate at a collector-base voltage of 0V. It is therefore important that digital logic transistors be able to sustain high current density operation at low collector-base voltages. Figure 7.8 shows the $f_t$ and $f_{max}$ of the transistor plotted versus current density at collector-base voltages of 0V, 0.2V and 0.4V.

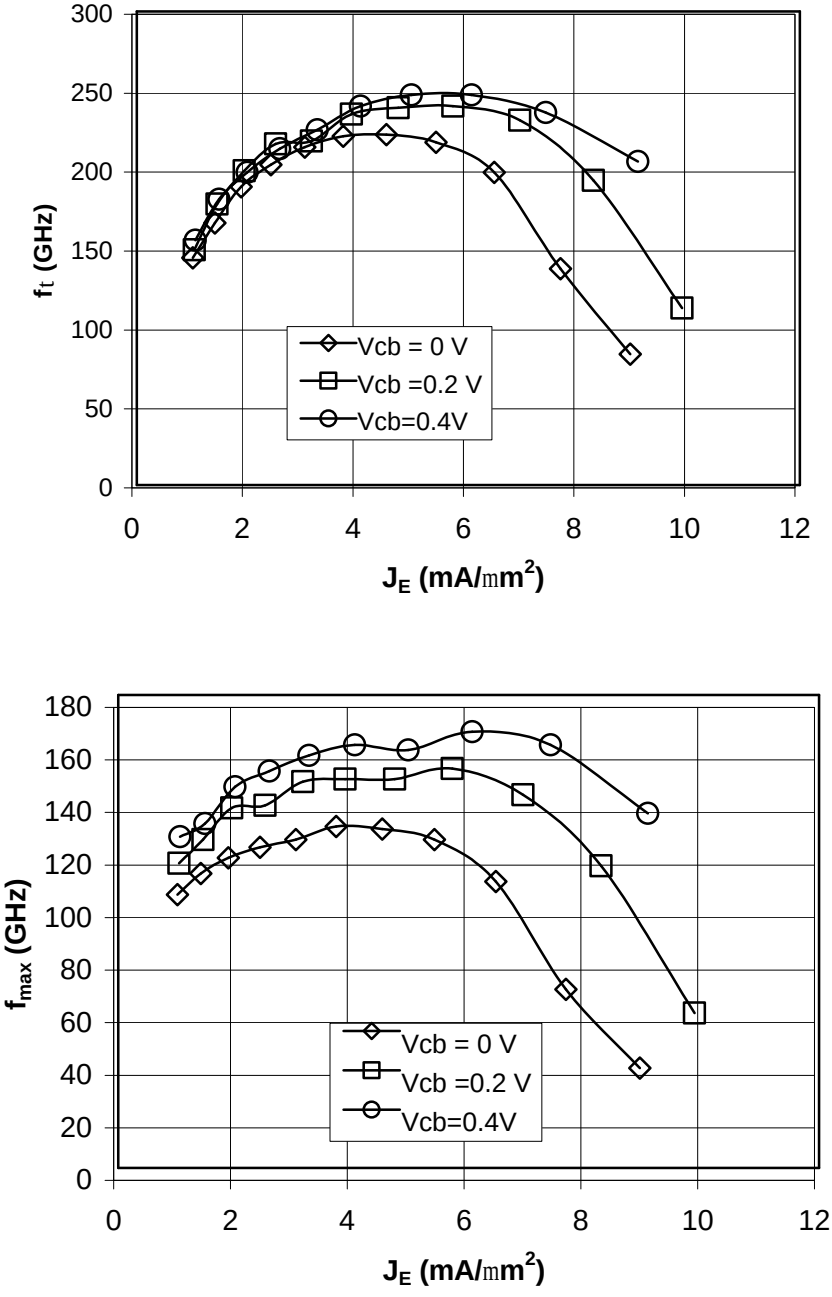



Figure 7.8: $f_t$ and $f_{max}$ plotted versus current density ($J_E$) at varying collector-base voltages ($V_{CB}$) for HBT described in Figure 7.7

The high current density operation and high $f_t$ exhibited by the transistor indicate a high effective electron velocity in the collector. As shown in Eqn. 2.1, a plot of 1/(2p $f_t$) versus 1/$I_C$ will have a y-intercept given by $t_c + t_b + (R_{ex} + R_c)C_{cb}$. Using the extracted $C_{cb}$ (from Im($Y_{12}$)), $R_{ex}$ (from the flyback method) and $R_c$ (from TLM measurements and the transistor geometry), the sum of $t_c + t_b$ is determined to be 0.33 psec for the transistor of Figure 7.7. If Eqn. 2.2 is used to estimate $t_b$, then $t_c$ and subsequently the effective collector velocity ($v_{eff} = T_c/2t_c$) can be determined. An effective collector velocity of 4.2 x $10^7$ cm/sec is extracted from the transistor data. That value is consistent with the high current operation that the transistor can sustain before the onset of the Kirk effect.

### 7.4.2 Rockwell Scientific Fabricated Transistors

Mesa-HBTs fabricated at Rockwell Scientific used electron-beam lithography to define the emitter contact, the base post and the base mesa. The use of e-beam lithography allowed the base post to be aggressively scaled and placed close to the emitter contact reducing the extrinsic collector-base capacitance. Device results are considered for transistors with nominal emitter junction widths of 0.7 μm. For these devices, the base-mesa width extended 0.5 μm on either side of the emitter contact. The devices used the same base contact scheme that was used for transistors fabricated at UCSB. One significant difference in the process was that the second sidewall deposition took place after device passivation and planarization. The passivation layer then protected the base metal during the sidewall RIE. These

devices showed a much higher $f_{max}$ than transistors fabricated at UCSB, indicating an improvement in base resistance.

It should be noted that these transistors did not use the on-wafer TRL calibration described in Chapter 3. Instead an off-wafer TRL calibration was performed using a calibration substrate, and the on-wafer embedding network was deembedded from measurements using the measurements of an open circuit and short circuit network.

Figure 7.9 shows a plot of the unilateral power gain ($U$) and short circuit current gain ($h_{21}$) for a transistor with nominal emitter junction dimensions of 0.7 x 3 $\mu m^2$. The device is biased at a collector current $I_C$ = 12.6 mA and a collector base voltage $V_{CB}$ = 0.4 V. Referenced to the emitter junction dimensions the device is operating at a current density $J_E$ = 6 mA/ $\mu m^2$. The extrapolated $f_t$ and $f_{max}$ of the transistor are 292 GHz and 314 GHz, respectively. The base-collector capacitance of the transistor is extracted from the imaginary part of $Y_{12}$ at low frequencies and is found to be 6.4 fF. Operating at this bias condition, the transistor is found to have a $C_{cb}/I_C$ ratio of 0.51psec/V. Figure 7.10 shows the $f_t$ and $f_{max}$ of the transistor plotted versus current density at collector-base voltages of 0V, 0.2V and 0.4V. The transistor is observed to sustain high current density (>4 mA/ $\mu m^2$) operation at $V_{CB}$ = 0V.

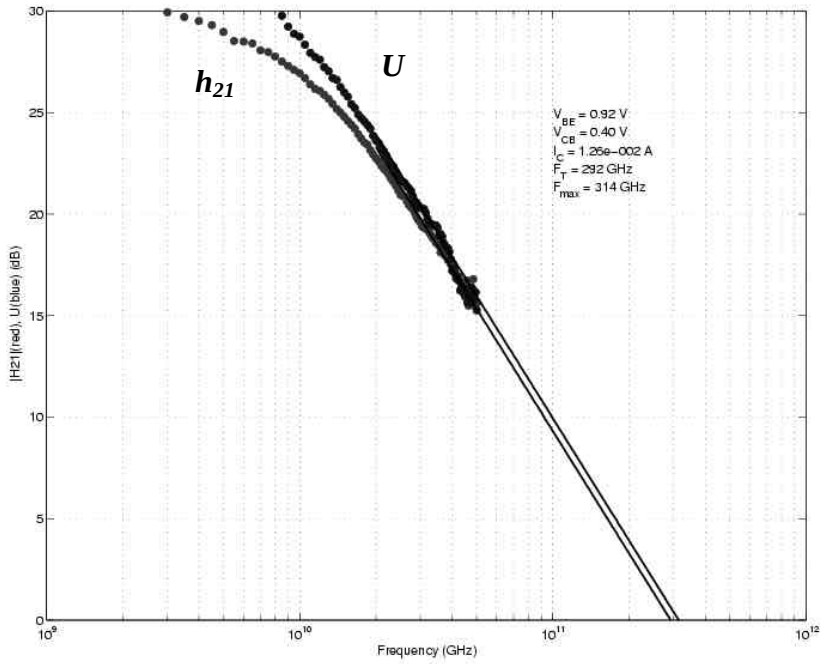



Figure 7.9: Unilateral power gain ($U$) and short circuit current gain ($h_{21}$) of submicron mesa-HBT fabricated at Rockwell Scientific. Emitter junction dimensions: 0.7 x 3 $\mu m^2$.

The collector-base capacitance can also be monitored for evidence of base pushout. Figure 7.11 shows the collector-base capacitance plotted versus current density at varying collector-base voltages. The higher value of $C_{cb}$ extracted at low collector-base voltages indicates the collector is not fully depleted at 0V $V_{CB}$.

The high $f_{max}$ exhibited by the Rockwell Scientific transistor indicates an improvement in the base resistance compared to the transistors fabricated at UCSB. However, the transistors' RF performances still seem to show a dependence on the sheet resistance of the base metal. Figure 7.12 shows a plot of the unilateral power

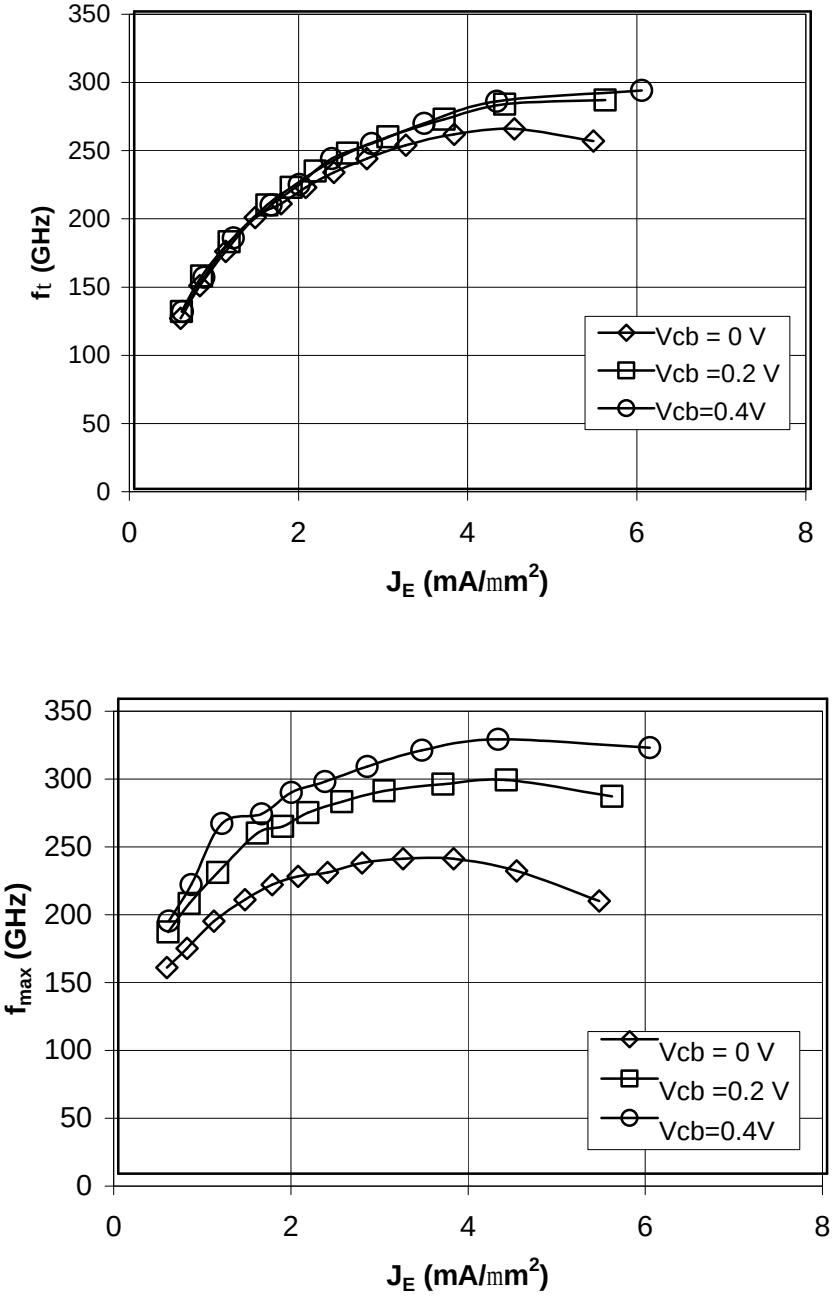



Figure 7.10: $f_t$ and $f_{max}$ plotted versus current density ($J_E$) at varying collector-base voltages ($V_{CB}$) for HBT described in Figure 7.9

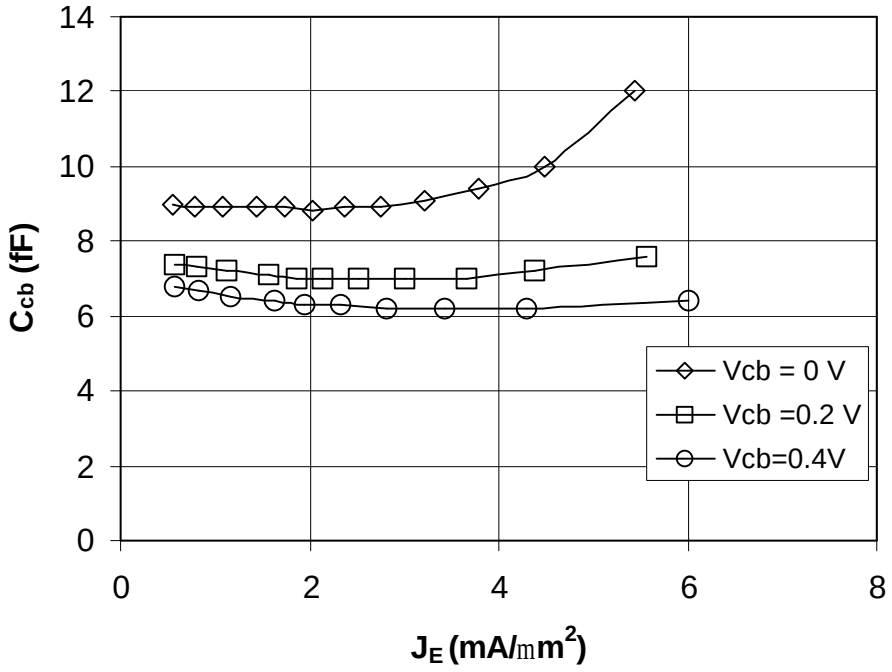



Figure 7.11: Collector-base capacitance ($C_{cb}$) extracted from Im($Y_{12}$)/? plotted versus current density ($J_E$) at varying collector-base voltages ($V_{CB}$) for HBT described in Figure 7.9

gain ($U$) and short circuit current gain ($h_{21}$) for a transistor with nominal emitter junction dimensions of 0.7 x 5 μm$^2$. The device is biased at a collector current $I_C$ = 20.5 mA ($J_E$ ~6 mA/ μm$^2$) and a collector base voltage $V_{CB}$ = 0.4 V. The extrapolated $f_t$ and $f_{max}$ of the transistor are 298 GHz and 239 GHz, respectively. The transistor shows a significantly lower $f_{max}$ than that observed for the 3 μm long emitter stripe device, an indication that the resistance contributed by the base metal along the length of the device is degrading the RF performance. The $f_t$ of the transistor is slightly higher than the shorter device because the extrinsic collector-base capacitance due to the base post is proportionally smaller. The base-collector

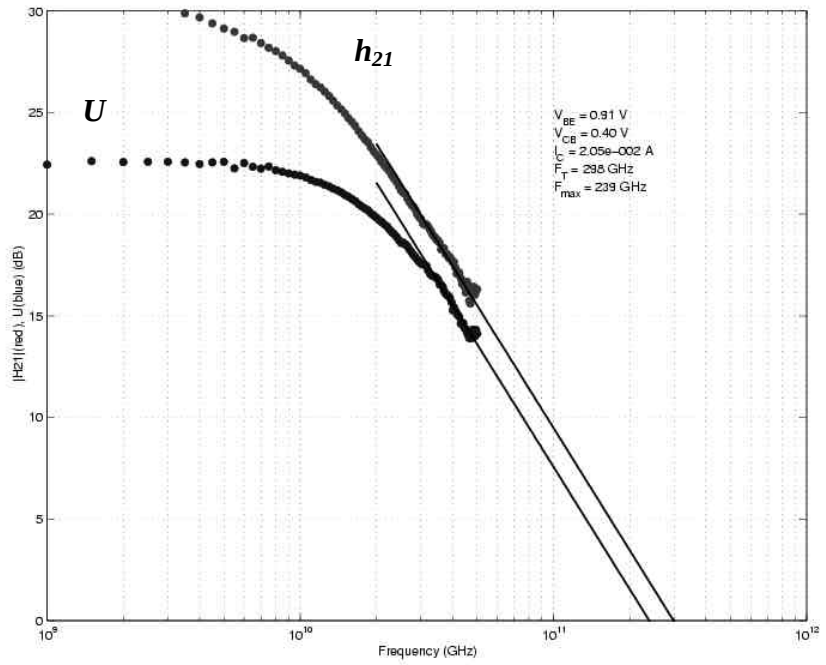



Figure 7.12: Unilateral power gain ($U$) and short circuit current gain ($h_{21}$) of submicron mesa-HBT fabricated at Rockwell Scientific. Emitter junction dimensions: 0.4 x 5 $\mu m^2$.

capacitance of the transistor is extracted from the imaginary part of $Y_{12}$ at low frequencies and is found to be 8.4 fF. Operating at this bias condition, the transistor is found to have a $C_{cb}/I_C$ ratio of 0.41psec/V.

## 7.5 Conclusions

Scaled mesa-HBTs have been fabricated with RF performance that compares favorably with state-of-the-art devices reported in the InP system. The HBT fabrication process has been modified to improve the manufacturability and scalability of submicron devices. The devices exhibit high $f_t$, high current density operation and low collector base capacitance. Devices fabricated at UCSB, suffered

from a high base resistance due to a thinning of the base metal that limited the transistor $f_{max}$. Devices fabricated at Rockwell Scientific used a modified base-emitter process flow that avoided these problems. A submicron HBT was realized with a simultaneous $f_t$ and $f_{max}$ of close to 300 GHz operating at an emitter current density of 6 mA/μm$^2$.

## REFERENCES

1. *Personal Communication* with Mattias Dahlström, October 2003.

2. W. Liu, *Handbook of III-V Heterojunction Bipolar Transistors*, John Wiley and Sons Inc., 1998, pp. 1089-1090.

# Chapter 8 Conclusions

## *8.1 Accomplishments*

This work has described the design, fabrication and characterization of submicron InP-based HBTs. Two unique device topologies were considered: a transferred-substrate HBT and a scaled mesa-HBT. Although there are significant differences in the process flows of the two devices, their designs are intimately tied. At the start of the work described in this thesis, the transferred-substrate process had been in development at UCSB for approximately 6 years and the process flow was well established. In this work, deep submicron scaling of the transistors was explored using electron-beam lithography. Highly-scaled transferred-substrate devices were found to exhibit negative resistance trends that are not predicted by standard HBT models. A theory to explain these trends due to electron velocity modulation in the collector was developed and correlated with measured results. A consequence of the negative resistance trends is that the transistor's $f_{max}$ cannot be extrapolated from low frequency measurements. The high available power gain of the transistors was confirmed through the design of small-signal amplifiers in the 140-220 GHz band. Single-stage and multi-stage amplifiers were design and fabricated, and these results represented the first HBT amplifiers reported in this frequency range.

The transferred-substrate process demonstrated the potential for a low parasitic III-V HBT technology in millimeter-wave tuned-circuit applications. However, severe problems limited the application of transferred-substrate devices

beyond small-scale demonstration circuits. The most significant problem was the inability to realize high levels of yield and integration in the technology. Developing a more manufacturable process was limited by the resources available in a university research environment. However, given the complexity of the process, the resources of the entire III-V community may not have been enough to develop a technology suitable for manufacture.

In addition to the technical challenges, the advantages of the transferred-substrate process were limited in mixed-signal and digital circuit applications. Despite clear advantages in RF figures-of-merit, transferred-substrate devices had demonstrated maximum logic speeds that were comparable InP mesa-HBTs and Si/SiGe bipolars. Transferred-substrate device suffered from a low maximum operating current densities because of their poor thermal characteristics, and the devices had large parasitic capacitances from the required thermal vias.

Based on these inherent problems, work on the transferred-substrate process at UCSB was stopped and efforts were redirected towards developing an InP HBT technology that was both scalable and manufacturable. In parallel with the work described here, research at UCSB has been pursued to develop process flows that more closely resemble those of a Si/SiGe bipolar transistor. Such a process flow would represent a revolutionary change in InP HBT fabrication methods. In this work, more evolutionary improvements in a mesa-HBT process were developed.

A scaled mesa-HBT process was developed that relies on well-established dielectric sidewall spacer and dry etch processes to form a self-aligned base-emitter

junction. A trench etch and dielectric refill has been incorporated to substantially reduce the extrinsic collector-base capacitance of the transistor. Additionally, the device epitaxy has been tailored to support a submicron process, with device parameters that are optimized for high digital logic speeds and not necessarily traditional transistor figures-of-merit ($f_t$, $f_{max}$). HBTs using this process flow have been fabricated at UCSB and Rockwell Scientific, and good RF results have been achieved.

## *8.2 Future Work*

The parameters of the scaled mesa-HBT were designed for high-speed digital logic applications based on the delay analysis of Chapter 6. The demonstration of high-speed digital circuits using the technology is necessary to validate the design approach. High-speed divider circuits using the process have been designed and are currently being fabricated at Rockwell Scientific Company and Global Communication Semiconductors (GCS).

From a technology perspective, further improvements could be made in the scaled mesa-HBT process flow, while stopping short of the revolutionary Si/SiGe-like process flow that has also been proposed. The planarization/etchback process used in the base-emitter junction formation was difficult and labor intensive. An ideal process flow would enable the base metal to be selectively deposited on the base semiconductor and not on the emitter contact or sidewalls. In silicon technologies, self-aligned salicides (silicon/metal alloys) may be formed selectively on semiconductor. An equivalent technology in a III-V material system would be

highly desirable. An alternative approach to forming the self-aligned base metal in a III-V system may be to use a selective chemical vapor deposition of a metal, or to attempt electroplating the base metal using the base semiconductor as the seed layer.

In this work, a trench isolation etch was used for base pad capacitance reduction, and this process could be extended to provide device isolation. Damage isolation implants of InP do not produce sufficiently high resistivity for device isolation. However, if a trench etch was used to electrically isolate the devices, an ion implantation could be added to create a moderately high resistance substrate that is sufficient for wiring capacitance reduction.

From an applications standpoint, the scaled-mesa HBTs in this work have focused on digital logic applications. However, in scaling device parasitics the mesa-HBTs have shown progressively higher levels of transistor power gain. This makes the devices a promising candidate for the millimeter-wave and sub-millimeter wave RF applications that the transferred-substrate technology was so well suited for. Additionally, double heterojunction transistors simultaneously offer large breakdown voltages, which should enable moderate levels of power generation.

As was discussed in the beginning of this work, the InP material system offers significant materials advantages over Si/SiGe devices. For digital applications, Si/SiGe bipolar devices offer comparable circuit bandwidths because they are aggressively scaled through advanced device processing. If scaled to the same degree, there is no question that InP devices would offer superior performance. However, the advanced processes used in a Si/SiGe foundry are enabled by the

enormous capital investment in the silicon CMOS infrastructure. An open question is whether a market exists to support the development of a similar infrastructure in the III-V semiconductor community.

# Appendix A

MATLAB source code for complex impedance correction of TRL calibration

```
digits(9);
% Read Excel File with measured Transistor S-parameters
raw_sparam=xlsread('TSHBT_r_Uncorr_Sparams\R3C2_E3x18_C7_W3_I50_V11_G.xls');
l=length(raw_sparam);
% Read Excel File with Complex Characteristic Impedance Correction Data
raw_zo=xlsread('Zo_r_140to220_mom.xls');
N=2048;
M=2048*6;
Zo_cal=57.0;

% Convert raw data to complex numbers
for n=1:101
   S11_meas(n)=raw_sparam(n,2)+i*raw_sparam(n,3);
   S12_meas(n)=raw_sparam(n,6)+i*raw_sparam(n,7);
   S21_meas(n)=raw_sparam(n,4)+i*raw_sparam(n,5);
   S22_meas(n)=raw_sparam(n,8)+i*raw_sparam(n,9);
   zo(n)=raw_zo(n,2)+i*raw_zo(n,3);

% Calculate measured y-parameters
   delta_meas(n)=(1+S11_meas(n))*(1+S22_meas(n))-S12_meas(n)*S21_meas(n);
   Y11_meas(n)=((1-
S11_meas(n))*(1+S22_meas(n))+S12_meas(n)*S21_meas(n))/delta_meas(n)/Zo_cal;
   Y12_meas(n)=-2*S12_meas(n)/delta_meas(n)/Zo_cal;
   Y21_meas(n)=-2*S21_meas(n)/delta_meas(n)/Zo_cal;
   Y22_meas(n)=((1+S11_meas(n))*(1-
S22_meas(n))+S12_meas(n)*S21_meas(n))/delta_meas(n)/Zo_cal;

% Calculate corrected Y-parameters
   Y11_corr(n)=Y11_meas(n)*Zo_cal/zo(n);
   Y12_corr(n)=Y12_meas(n)*Zo_cal/zo(n);
   Y21_corr(n)=Y21_meas(n)*Zo_cal/zo(n);
   Y22_corr(n)=Y22_meas(n)*Zo_cal/zo(n);
```

```
% Calculate corrected S-parameters referenced to 50 Ohms
   delta_corr(n)=(1+Y11_corr(n)*50)*(1+Y22_corr(n)*50)-Y12_corr(n)*50*Y21_corr(n)*50;
   S11_corr(n)=((1-Y11_corr(n)*50)*(1+Y22_corr(n)*50)+Y12_corr(n)*50*Y21_corr(n)*50)
/delta_corr(n);
   S12_corr(n)=-2*Y12_corr(n)*50/delta_corr(n);
   S21_corr(n)=-2*Y21_corr(n)*50/delta_corr(n);
   S22_corr(n)=((1+Y11_corr(n)*50)*(1-
Y22_corr(n)*50)+Y12_corr(n)*50*Y21_corr(n)*50)/delta_corr(n);

% Calculate Unilateral Power Gain
  U_numerator(n)=abs(Y21_corr(n)-Y12_corr(n))*abs(Y21_corr(n)-Y12_corr(n));
  U_denominator(n)=4*(real(Y11_corr(n))*real(Y22_corr(n))-real(Y21_corr(n))*real(Y12_corr(n)));
 if U_denominator(n) < 0
      mason(n)=0;
 else
      mason(n)=10*log10(U_numerator(n)/U_denominator(n));
 end

% Calculate MSG/MAG
   b(n)=1+abs(S11_corr(n))^2-abs(S22_corr(n))^2-abs(S11_corr(n)*S22_corr(n)-
S12_corr(n)*S21_corr(n))^2;
   k(n)=(1-abs(S11_corr(n))^2-abs(S22_corr(n))^2+abs(S11_corr(n)*S22_corr(n)-
S12_corr(n)*S21_corr(n))^2)/(2*abs(S12_corr(n)*S21_corr(n)));
   if (b(n) > 0) & (k(n) > 1)
      MSG_MAG(n)= 10*log10(abs(S21_corr(n))/abs(S12_corr(n))*(k(n)-(k(n)^2-1)^.5));
   else
      MSG_MAG(n)= 10*log10(abs(S21_corr(n))/abs(S12_corr(n)));
   end

% Calculate h21
   h21(n)= 10*log(Y21_corr(n)/Y11_corr(n));

% Calculate Ccb
   Ccb(n)= -imag(Y12_corr(n))/(2*3.14159*raw_sparam(n));

end
```

```
% Write corrected S-parameters to file in Touchstone (s2p) format.
fp = fopen('R3C2_E3x18_C7_W3_I50_V11_G.s2p','w');
 fprintf(fp,'%s\n', '# hz S ri R 50');
for n=1:101
   x=[raw_sparam(n); real(S11_corr(n)); imag(S11_corr(n)); real(S21_corr(n)); imag(S21_corr(n));
real(S12_corr(n)); imag(S12_corr(n)); real(S22_corr(n)); imag(S22_corr(n))];
   fprintf(fp,'%1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t
%+1.9E\n',x);
end

% Write U, MSG/MAG.
fclose(fp);
fp = fopen(R3C2_E3x18_C7_W3_I50_V11_G.txt','w');
 fprintf(fp,'%s\t %s\t %s\t %s\t %s\t %s\t %s\t %s\n', 'freq', 'U', 'MSG/MAG', 'H21', 'Ccb', 'ReY12',
'ReY22', 'ImY22');
for n=1:101
   x=[raw_sparam(n); mason(n); MSG_MAG(n); h21(n); Ccb(n); real(Y12_corr(n));
real(Y22_corr(n)); imag(Y22_corr(n))];
   fprintf(fp,'%1.9E\t %1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\t %+1.9E\n',x);
end
fclose(fp);
```