INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI[®]
800-521-0600

**Design, Simulation and Fabrication of
a Self-Aligned SiGe Base Heterojunction Bipolar Transistors
For Low Power Operation**

by

Abdulabbas A. Hussain, B.Sc., M.Sc.

**A thesis submitted to the
Faculty of Graduate Studies and Research
In partial fulfillment of the requirements for the degree of**

**Master of Engineering
in Electronics**

**Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Electronics, Faculty of Engineering
Carleton University
Ottawa, Ontario
Canada**

June 1999

© A.A. Hussain 1999, Ottawa, Canada

**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-48450-5

Canada

**The undersigned hereby recommend to the Faculty of Graduate
Studies and Research acceptance of the thesis**

**“Design, Simulation and Fabrication of
a Self-Aligned SiGe Base Heterojunction Bipolar Transistors
For Low Power Operation”**

submitted by

Abdulabbas A. Hussain, B.Sc., M.Sc.

**in partial fulfillment of the requirements of the degree of
Master of Engineering**

Prof. N.G. Tarr

Thesis Supervisor

Prof. J.S. Wight

Chair, Department of Electronics

Carleton University

June 1999

**"IN THE NAME OF ALLAH,
THE BENIFICENT, THE MERCIFUL"**

Abstract

This thesis describes the design, fabrication and electrical characterization of a novel self-aligned SiGe base heterojunction bipolar transistor optimized for low-power operation. A theoretical analysis is presented revealing that the emitter-base junction capacitance dominates F_t in this operating regime. Guided by this initial analytic modelling, the device simulator BIPOLE3 was used to optimize doping profiles and other structural parameters. A graded Ge profile with a peak Ge content of 10% was assumed. Relatively light base and collector doping is required to optimize low power performance.

Experimental devices were fabricated by modifying an existing commercial process flow for silicon BJTs. New process modules were developed for field polysilicon formation, SiGe base deposition, and local interconnect formation. The SiGe base was formed by a UHV/CVD process. Electrical test results on the first experimental devices are very encouraging. DC characteristics are generally good, although F_t is approximately 30% lower than predicted by BIPOLE3.

Acknowledgements

I wish to convey sincere and deep appreciation to my thesis supervisor Dr. N.G. Tarr for his encouragement, support and for many helpful discussions and suggestions during the course of this work.

I would also like to thank the technical staff of the fabrication laboratory at Carleton University: Mr. Lyall Berndt for wafer diffusion/annealing processing; Mrs. Carol Adams for wafer pre-diffusion cleaning; Mr. Christopher Pawlowicz for direct write electron-beam lithography and etching. I would also like to express my gratitude to the professors, technical staff and students at the department of electronics for their help during the course of this work. Special thanks to Dr. D.J. Walkley for several consultations and help with difficulties I encountered.

Many members of the technical staff at Gennum Corporation certainly deserve recognition for their help and consultations: Mr. Andrew Cervin-Lawry for his supervision during the layout of the devices and fabrication; Mrs. Myriam Buchbinder for her help during fabrication and valuable comments; Dr. J. Kendall for AC measurements and useful discussions, and Mrs. W. Gustafson for wafer processing. Special thanks goes to Mr. D. Lynch for his valuable suggestions during the weekly meetings at Gennum.

Most importantly, I am deeply indebted to my forbearing wife and children, especially my daughter Berak and my son Belal during the preparation of this thesis. This work is dedicated to them.

Finally, I would like to thank, the department of Electronics at Carleton University, OSAP from the Ministry of Education and Training, Gennum Corporation of Burlington, Ontario and Micronet for their financial support. I extend my thanks to the Canadian Marconi Company for their bursaries.

Table of Contents

Abstract	iv
Acknowledgements	v
Table of Contents	vi
List of Figures	viii
List of Tables	x
List of Symbols	xii
CHAPTER 1. : INTRODUCTION	1
1.1 Basics of SiGe HBT Technology	1
1.2 Objective of the Present Work	3
CHAPTER 2. : SiGe HBT PHYSICS AND TECHNOLOGY	5
2.1 Physics of the SiGe HBT	5
2.2 Growth of SiGe on Si Substrate	9
2.3 The UHV/CVD Process	12
2.4 Surface Cleaning	13
2.5 SiGe Film Growth	14
2.6 Polycrystalline Silicon	14
2.7 Transit Times in Bipolar Transistors	15
2.7.1 One-Dimensional (Vertical) Analysis	15
2.7.2 Two-Dimensional (Lateral) Analysis	19
CHAPTER 3. : DEVICE ARCHITECTURE	22
3.1 Transistor Structures	22
3.2 Device Layout	25
CHAPTER 4. : DEVICE SIMULATIONS AND ANALYSIS	28
4.1 Introduction	28
4.2 Electrical Performance Specifications	30
4.3 Calculation of the Extrinsic Base of the NSA Devices	31
4.3.1 Model 1 (Overestimates Base Resistance)	32
4.3.2 Model 2 (Underestimates Base Resistance)	34
4.4 Delay Times Contributing to the Transition Frequency	34
4.4.1 Preliminary Calculations	35
4.4.2 Simulation Results	39

4.4.2.1	Effect of Doping Levels on the Delay Times	41
4.4.2.1.1	Vertical Analysis	41
4.4.2.1.2	Lateral Analysis	44
4.4.2.2	Effect of Transistor Size on the Delay Times	45
4.4.2.2.1	Vertical Analysis	45
4.4.2.2.2	Lateral Analysis	46
4.5	Effect of Doping Levels on the Electrical Characteristics	48
4.5.1	Emitter Doping Level	49
4.5.2	Epitaxial Collector Doping Level	50
4.6	Influence of Device Scaling on the Electrical Characteristics	58
4.7	Simulation Results of an Si-Base Device	60
4.8	General Transistor Characteristics	64
CHAPTER 5.:	DEVICE FABRICATION AND CHARACTERIZATION	69
5.1	Introduction	69
5.2	Test Wafer Layout	69
5.3	Short Loop Experiments	70
5.4	Field Poly Module	72
5.5	SiGe Base Deposition Module	79
5.6	Local Interconnect Module	83
5.7	Rapid Thermal Emitter Annealing	85
5.8	Device Characteristics and Discussion	88
5.8.1	DC Characteristics	88
5.8.2	AC Characteristics	91
CHAPTER 6. :	CONCLUSIONS AND RECOMMENDATIONS	96
6.1	Conclusions	96
6.2	Recommendations For Future Work	100
APPENDIX A. :		102
A.1	The pn One-Sided Step Junction	102
A.2	Depletion Layer Capacitance	103
APPENDIX B. :		104
B.1	Tables	104
REFERENCES		111

List of Figures

Figure	Description	Page
2.1	a:Energy band diagrams of a Si bipolar transistor (solid line) and SiGe HBT (dotted line). b:Doping profile of Ge across the base region	7
2.2	Representation of the strained and relaxed structures of SiGe on Si substrate	12
3.1	Cross-sectional view of the self-aligned HBT structure	25
3.2	Cross-sectional view of the non-self aligned HBT structure	25
3.3	Self-aligned HBT with two base contacts	27
3.4	Non-self aligned HBT with two base contacts	28
4.1	Extrinsic base cross-section	32
4.2	Model 1 for the NSA device	34
4.3	Model 2 for the NSA device	34
4.4	Typical simulated impurity profile for SA1X and NSA1X of SiGe HBT	67
4.5	Typical simulated Gummel plots of base and collector currents for SA1X and NSA1X of SiGe HBT	68
4.6	Typical simulated I_c - V_{ce} characteristics for SA1X and NSA1X of SiGe HBT	68
4.7	Simulated F_i and F_m as a function of I_c for SiGe HBT	69
5.1	Schematic cross-section of the field poly process sequence of the SiGe HBT	76
5.2	Schematic cross-section of the SiGe process sequence of the SiGe HBT	83
5.3	Schematic cross-section of the LOC_INT process sequence of the SiGe HBT	86

5.4	Experimental collector and base currents versus base-emitter bias for the SiGe HBT	91
5.5	Experimental collector current versus emitter-collector bias for different base currents for the SiGe HBT	92
5.6	Measured transition frequency versus collector currents for a SiGe HBT	94

List of Tables

Table	Description	Page
4.1	Mask dimensions of the vertical SiGe HBTs	41
4.2	Effect of doping levels on the delay times obtained from vertical analysis	45
4.3	Effect of doping levels on the delay times obtained from lateral analysis	45
4.4	Effect of a transistor size on the delay times obtained from vertical analysis	48
4.5	Effect of a transistor size on the delay times obtained from lateral analysis	48
4.6	Heavy emitter doping	52
4.7	Light emitter doping	53
4.8	Light emitter doping, light EPI doping	54
4.9	Heavy emitter doping, light EPI doping	57
4.10	Delay times for Si-base devices obtained from vertical analysis	64
4.11	Delay times for Si-base devices obtained from lateral analysis	64
5.1	Summary of the field poly results	79
5.2	Variation of sheet resistance with temperature	89
B.1 Part1	Heavy emitter doping, various geometries	107
B.1 Part2	Heavy emitter doping, various geometries	108
B.2 Part1	Light emitter doping, various geometries	109
B.2 Part2	Light emitter doping, various geometries	110
B.3 Part1	Light emitter doping, light EPI doping, various geometries	111
B.3 Part2	Light emitter doping, light EPI doping, various geometries	112

List of Symbols

A_{bc}	Base-collector junction area
A_{eb}	Emitter-base junction area
BV_{ceo}	Collector-emitter breakdown voltage
C_{diff}	Diffusion capacitance
C_{JC}	Base-collector junction capacitance
C_{JE}	Emitter-base junction capacitance
C_{JSC}	Substrate capacitance
D_n	Electron diffusion constant in the base
ΔE_g	Barrier height at a heterojunction
$\Delta E_{g,Ge(grade)}$	Band gap grading across the neutral base region
$\Delta E_{g,Ge(x=0)}$	Band gap reduction at the emitter-base depletion edge
$\Delta E_{g,Ge(x=W_b)}$	Band gap reduction at the base-collector depletion edge
EPI	Epitaxial collector layer
F_c	Grading coefficient
F_m	Maximum oscillation frequency
$(F_m)_p$	Peak value of the maximum oscillation frequency
F_t	Transition frequency
$(F_t)_p$	Peak value of the transition frequency
$F_{t,Si}$	Transition frequency for Si-base
$F_{t,SiGe}$	Transition frequency for SiGe-base
G_f	Grading factor
g_m	Transconductance
I_b	Base current
I_c	Collector current
$(I_c)_p$	Value of collector current giving highest F_t

ICs	Integrated circuits
J_c	Collector current density
k	Boltzmann constant
N_a	Ionized acceptor dopant concentration
N_c	Effective density of states in conduction band
N_d	Ionized donor dopant concentration
n_i	Intrinsic carrier concentration
N_v	Effective density of states in valence band
q	Electronic charge
R_B	Total base resistance
R_b	Intrinsic base resistance
R_{bex}	Extrinsic base resistance
r_c	Undepleted collector resistor
R_{cont}	Vertical resistance under the collector contact
r_e	Diffusion resistance
R_s	Sheet resistance
T	Absolute temperature
TEPI	Epitaxial layer thickness
τ_{AVL}	Delay time due to sidewall injection electron concentration into the base region outside the active base region
τ_{AVO}	Small signal delay time due to integral of minority carrier charge in the whole of active region
τ_b	Base-transit time
$\tau_{b,Si}$	Transit time for Si-base
$\tau_{b,SiGe}$	Transit time for SiGe-base
τ_c	Time constant for charging the base-collector capacitance through the collector resistance
τ_{cb}	Base-collector junction capacitance charging time
τ_e	Neutral emitter region charging time

τ_{eb}	Emitter-base junction charging time
τ_{ec}	Emitter to collector transit time
τ_{ESC}	Delay time due to minority carrier charge in the emitter-base space charge layer
t_{ox}	Oxide thickness
T_p	Heatpulse preset temperature
τ_{qbe}	Delay time due to free carrier charge in emitter-base space-charge layer
τ_{RCE}	Delay time due to C_{JC} and total series resistance
τ_{REE}	Delay time due to $(kT/q)(C_{JE}+C_{JC})$
τ_{scl}	Collector space charge layer transit time
τ_{SUB}	Delay time due to $(C_{JSC} R_{com}/\beta)$
τ_{tot}	Total delay time in a two-dimensional analysis
T_w	Wafer temperature
V_A	Early voltage
$V_{A,Si}$	Early voltage for Si-base
$V_{A,SiGe}$	Early voltage for SiGe-base
V_{bc}	Base-collector voltage
V_{bic}	Built-in voltage at the base-collector region
V_{bie}	Built-in voltage at the emitter-base region
V_T	Thermal voltage
V_{ce}	Collector-emitter voltage
V_{eb}	Emitter-base voltage
v_s	Saturated carrier velocity
W_e	Emitter depth
W_{scl}	Collector depletion width
X_{eb}	Width of the emitter-base junction
X_{J1}	Metallurgical junction depth at the emitter-base region
X_{J2}	Metallurgical junction depth at the base-collector region
ϵ_0	Permittivity of free space

ϵ_{si}	Permittivity of silicon
β	Current gain
γ	Ratio of the position-average density of states product ($N_c N_v$) between SiGe-base and Si-base transistors
γ_c	Base-collector junction coefficient constant
γ_e	Emitter-base junction coefficient constant
ρ	Wafer sheet resistance
η	Ratio of the position-average minority carrier mobility in the base of a SiGe HBT compared to a Si BJT

CHAPTER 1

INTRODUCTION

1.1 Basics of SiGe HBT Technology

Silicon bipolar transistor technology is about a half century old. For the past five decades, the speed of silicon bipolar junction transistors (BJTs) has gradually increased as a result of the decrease in their dimensions. Whether silicon devices have reached theoretical limits or not, the introduction of SiGe devices illustrates a discontinuity on the path to faster silicon [1]. Scientists and engineers at several institutions [1-3] have been studying new semiconductor materials for many years. One direction in these studies was to add small amounts of germanium (Ge) to silicon in order to create a new stable alloy ($\text{Si}_{1-x}\text{Ge}_x$, where x represents the mole fraction, often written SiGe for simplicity). This material is of considerable interest as it offers the potential to deliver improved performance compared to the conventional silicon devices.

The introduction of SiGe into conventional Si technology has attracted much attention because of the potential to obtain high device performance at low cost for many applications such as optical and wireless communication systems, local area networks and high speed digital circuits. High transition frequencies, F_t , high maximum oscillation frequencies F_m and high current gains can be achieved using SiGe heterojunction bipolar transistors (HBTs) rather than Si BJTs. Additional benefits such as higher Early voltages

and lower sheet resistance can also be achieved when SiGe is introduced in the base of the transistor. These high performance devices are ideal for meeting the requirements of the high-speed analog circuits required for RF and microwave applications [4,5].

Several groups have reported the transistor transition frequencies of SiGe-HBTs in excess of 100 GHz [6-8] with graded, narrow bases. Furthermore, the maximum oscillation frequency was increased to 160 GHz by minimizing both the total base resistance R_B and base-collector junction capacitance C_{JC} [9]. Minimizing R_B and C_{JC} is crucial to maximizing the F_m . Until recently, these performance levels were only achievable with the most advanced GaAs technology.

Transistors that are built in silicon germanium can run substantially faster than their pure silicon counterparts, and can be made at lower cost with greater opportunity for integration with standard Si processes than can high speed GaAs devices. Gallium arsenide processes exhibit excellent performance characteristics but suffer from higher pricing resulting from higher wafer costs, plus costs for epitaxy material and low device yields. SiGe opens the door to merging logic and RF on one die, and thus competes favorably with GaAs in performance across the range of frequencies where GaAs is now dominant. That merger stems from SiGe's ability to let bipolar and CMOS coexist on a single silicon wafer [10].

Adding Ge to the base of the transistor influences the band gap energy, band

discontinuities, scattering processes and electron-hole recombination processes. Ge reduces the band gap, lowering the barrier for electron injection into the base, which in turn allows heavier base doping to be used, lowering base resistance. Band gap grading introduces a drift field to aid the transport of electrons across the neutral base. The electron mobility in SiGe is about two to three times higher than that in pure Si, allowing more rapid electron transport across the base.

SiGe has the same crystallographic structure as Si [3] but its lattice constant is larger by about 4% for pure Ge. When SiGe is grown epitaxially on Si substrate in such a way that the film and substrate adopt a common lattice constant while maintaining perfect crystallinity across the growth interface, the composite film undergoes strain. One immediate concern is the effect of strained-layer epitaxy. For any given Ge content, there is a critical thickness of SiGe above which starts to relax toward its bulk lattice spacing and can generate defects which can devastate device performance. One possible solution to this problem is to prepare SiGe using ultra-high vacuum chemical vapor deposition UHV/CVD at relatively low temperature. This technique, which was developed by Meyerson and his coworkers [1] at IBM, was proven to be a successful technique to overcoming the above mentioned difficulties. This issue and related ones will be addressed in chapter two.

1.2 Objective of the Present Work

The main objective of this thesis is the design, simulation and fabrication of a vertical

SiGe base heterojunction bipolar transistor optimized for low power operation. While previous researchers have concentrated on maximizing the speed of SiGe HBTs irrespective of power consumption, this work attempts to utilize the SiGe base to enhance performance at low power levels. To realize the goal of high speed operation at low power, a novel self-aligned device structure is introduced.

The work described here is collaborative with Gennum Corporation of Burlington, Ontario. In particular, the novel self-aligned SiGe base device was developed by modifying an existing commercial process developed at Gennum. The author's contribution to the project included design of optimum doping and composition profiles for the new device using the BIPOLE3 simulator, modification of an existing L-Edit layout environment to include the new layers required for SiGe devices, design of a suite of test transistors, and development of several key process techniques related to SiGe base formation, and analysis of electrical characteristics of completed devices. Most device processing was carried out at Gennum, with some key steps completed at the National Research Council and Carleton University. To protect the commercial interests of Gennum, no details are given here on the baseline process on which the SiGe base is grafted.

The thesis is organized as follows. In chapter 2, a brief introduction to SiGe HBT technology and theoretical background are given. Chapter 3 presents device architectures. BIPOLE3 simulation results are presented in chapter 4. In chapter 5, details of the new developed process flow, fabrication techniques and electrical characterization results for completed devices are presented. The conclusions of this study are presented in chapter 6.

CHAPTER 2

SiGe HBT PHYSICS AND TECHNOLOGY

2.1 Physics of the SiGe HBT

The advantage of the SiGe heterojunction bipolar transistor (HBT) lies in the ability to modify the band gap of ordinary silicon semiconductor material for higher performance. Band gap engineering is a powerful tool used for creating faster transistors. Conventional silicon devices have an effectively fixed band gap of 1.12 eV which limits switching speed, compared to HBTs fabricated on III-V compound materials such as GaAs. With the addition and grading of the Ge (which has an energy gap of 0.66 eV [11]) concentration across the base region of a BJT, it is possible to modify the band gap and consequently enhance the performance of the silicon transistor.

The main difference between the SiGe HBT and Si BJT is illustrated in the energy band diagram of Fig.2.1a of a linearly graded Ge doping profile in Fig.2.1b [12]. Figure 2.1a shows schematic energy band diagrams of Si BJT (solid line) and SiGe HBT (dotted line) showing a band gap modification [12]. ΔE_g is the band gap difference between Si and SiGe; $\Delta E_{g,Ge}(x=0)$ and $\Delta E_{g,Ge}(x=W_b)$ represents the amount of band gap reduction at the emitter-base and base-collector depletion edges in the neutral base, respectively. $\Delta E_{g,Ge}(\text{grade}) = \Delta E_{g,Ge}(x=W_b) - \Delta E_{g,Ge}(x=0)$ represents the band gap grading across the

Figure 2.1 a: Energy band diagrams of a Si bipolar transistor (solid line) and SiGe HBT (dotted line). b: Doping profile of Ge across the base region [12]

neutral base region. This band gap grading introduced into the base region of a SiGe HBT induces a drift field which aids minority carrier transport through the base, resulting in higher frequency operation [12]. This means that the base band gap of SiGe HBTs can be engineered to enhance device performance, thereby making it suitable for a wide range of high speed analog and RF applications. Harame *et al* [12] found in SiGe HBTs, the Ge introduced into the base region reduces the base band gap by about 75 meV per 10% Ge compared to a Si BJT. It is this reduced band gap of SiGe strained layers compared to Si that is directly exploited in the SiGe heterojunction bipolar transistor. The ΔE_g is determined not only by the Ge concentration but also by the strain created between the

SiGe and Si substrate. For identically constructed devices, the ratio of the current gain β between a SiGe HBT and Si BJT is given by [12],

$$\beta_{\text{SiGe}} / \beta_{\text{Si}} = [\eta \gamma \Delta E_{g,\text{Ge}}(\text{grade}) / kT] \exp[\Delta E_{g,\text{Ge}}(x=0)/kT] / [1 - \exp(-\Delta E_{g,\text{Ge}}(\text{grade}) / kT)], \quad (2.1)$$

a ratio which is larger than unity for finite Ge content. η represents the ratio of the position-averaged minority carrier mobility in the base region of a SiGe HBT compared to a Si BJT, γ represents the ratio of position-averaged density of states product ($N_c N_v$) between the two transistors, k is Boltzmann constant, and T is the absolute temperature.

Equation 2.1 indicates that the smaller base band gap in a SiGe HBT exponentially increases the number of minority carriers injected into the base, causing an increase in the collector current for the same forward-bias.

As mentioned above, qualitatively, the grading of the Ge atoms across the quasi-neutral base induces a drift-field in the base which accelerates the electrons injected from the emitter to the collector, thereby decreasing the base transit time compared to a Si BJT. Harame *et al* [12] found for a constant base doping and using the theoretical equation derived by Kroemer [13] for the base transit time, the ratio of base transit time for SiGe ($\tau_{b,\text{SiGe}}$) and Si ($\tau_{b,\text{Si}}$) devices is

$$\tau_{b,\text{SiGe}} / \tau_{b,\text{Si}} \approx [2kT/(\Delta E_{g,\text{Ge}}(\text{grade}))][1 - (kT/\Delta E_{g,\text{Ge}}(\text{grade}))(1 - \exp(-\Delta E_{g,\text{Ge}}(\text{grade})/kT))], \quad (2.2)$$

which, for finite Ge grading, is less than unity, and thus enhances the transition frequency F_t . Since the base-transit time τ_b is usually a major component in the F_t equation,

$$F_{t,\text{SiGe}} > F_{t,\text{Si}} \quad (2.3)$$

An additional benefit of using a graded-Ge profile in a SiGe HBT is an enhancement in the Early voltage V_A - a measure of ease with which the majority carrier profile in the base is depleted by an applied base-collector bias. Prinz and Sturm [14] developed an analytical model to explain the higher values of the product βV_A in Si/SiGe/Si HBTs. The output conductance (hence Early voltage) is a measure of how much the neutral base profile can be depleted with reverse bias on the collector-base junction, and is manifested in the rise of the collector current density J_c with collector bias at fixed base-emitter voltage. The ratio of Early voltage in the SiGe HBT to that of the Si BJT is given by [13]

$$V_{A,\text{SiGe}}/V_{A,\text{Si}} \approx \exp[\Delta E_{g,\text{Ge}}(\text{grade})/kT][(1 - \exp(-\Delta E_{g,\text{Ge}}(\text{grade})/kT))/(\Delta E_{g,\text{Ge}}(\text{grade})/kT)], \quad (2.4)$$

Equation 2.4 indicates that for finite grading across the quasi-neutral base, the exponential dominates the functional dependence on Ge content, and the ratio in eq.2.4 is larger than unity, yielding an improvement for a SiGe HBT compared to a Si BJT. In

other words, the base region becomes harder to deplete than for the comparably doped silicon transistor. Large Early voltage means the transistor has high output resistance, which is useful for analog design applications such as RF communications.

From the above discussion, we can see that β , F_t and V_A are increased significantly when SiGe is introduced in the base region of the HBT, and this leads to improvement in the device performance.

High transition frequency F_t and high oscillation frequency F_m are simultaneously required for both digital and analog ICs. The relationship between F_t and F_m is given by [15],

$$F_m = [F_t / (8 \pi R_B C_{JC})]^{1/2}, \quad (2.5)$$

where R_B is the total base resistance and C_{JC} is the collector-base capacitance. SiGe HBTs have high F_m compared to Si BJTs since the base in the former transistors can be heavily doped while maintaining a high injection efficiency due to the band gap difference between silicon and strained SiGe. This means that the high current gain in the SiGe base of the HBT can be traded off for higher base doping without compromising device performance.

2.2 Growth of SiGe on Si Substrate

The nature of the crystal structure resulting from the growth of SiGe on the Si substrate

plays an important role on the device performance. There is a mismatch between the lattice constant of the Ge (5.658 Å) and Si (5.431 Å) [16], which is equal to about 4%. As a result of the mismatch between the lattice of the SiGe film and Si substrate, two distinct classes of growth are possible: pseudomorphic growth and dislocated growth. Figure 2.2 shows two dimensional representations of the two classes of growths including the Si substrate and the relaxed SiGe layer. Below is a qualitative description of the such growths, and the details can be found in several references [1,17,18].

In class one pseudomorphic (incommensurate) growth, a compressive strain will be developed as a result of the lattice mismatch between the epitaxial SiGe and the Si substrate. Consequently, the lattice constant of the SiGe will be tetragonally distorted, to match the in-plane lattice constant as shown in Fig.2.2c. Pseudomorphic growth of SiGe on Si can continue for a thickness determined primarily by the mismatch of the relaxed lattice constants of the SiGe and Si, and the substrate temperature. If the SiGe thickness exceeds a certain value or the substrate is exposed to sufficiently high temperatures for long periods of time, at which the pseudomorphically grown layer is no longer thermodynamically stable, the lattice constant relaxes to its original value. This means that the strain in the SiGe layer will be relaxed and misfit dislocations will form at the interface between the silicon and SiGe, as illustrated in Fig.2.2d. One disadvantage of the strain relaxation in heterojunction bipolar transistors is that it lowers the band gap difference between the Si and SiGe resulting in a degradation in injection efficiency and a reduction in current gain, as illustrated in eq.2.1. Also misfit dislocations that accompany

Fig.2.2 Representation of the strained and relaxed structures of SiGe on Si substrate

strain relaxation can lead to leakage and/or recombination at the interfaces where there are a high density of dislocations [19].

It appears from the above discussion that the SiGe strained layer is thermodynamically stable to the formation of unwanted film defects only when the Ge content is kept below a specific upper bound. Increasing the Ge content will decrease the critical thickness of the SiGe. Since Ge is required only in the base region of the transistors, and since many of the properties of bipolar transistors depend exponentially on changes in the band gap, only a small Ge concentration is required. In the present work a maximum concentration of 10% will be used during formation of the transistor base. This means that SiGe stability constraints are satisfied in the fabricated devices. Since the base thickness is less than 1000 Å, the devices are stable under the Mathews-Blakeslees criterion [20].

2.3 The UHV/CVD Process

SiGe alloy layers may be grown on Si epitaxially by several techniques [1,21-24] that compete in terms of cost, ease of thickness and dopant control, flexibility, scalability and contamination by impurities. The most common ones are molecular beam epitaxy (MBE), low pressure chemical vapor deposition (LPCVD), atmospheric pressure CVD and ultra-high vacuum chemical vapor deposition (UHV/CVD). The challenge of growing SiGe alloys is currently met by all of the above-mentioned methods, with UHV/CVD being the most widely used and the most amenable to high-volume production. This technique was first demonstrated by Meyerson [21]. More specifically, Meyerson and co-workers at

IBM [1] combined the ultra-high vacuum of MBE and a conventional CVD process by exploiting the role of hydrogen in regulating the chemistry at the Si and SiGe growth interface.

The UHV/CVD developed by Meyerson [1,21] offers several advantages over other growth techniques, such as low temperature (<550 °C) deposition, *in situ* p-type doping, highly precise film thickness, composition and dopant concentration profiles that can be produced of almost arbitrary abruptness. The result has been the ability to produce a high performance SiGe heterojunction bipolar transistor.

2.4 Surface Cleaning

The UHV/CVD technique is dependent on a simple yet vitally important surface cleaning and preparation procedure to avoid the growth of native oxide which would interfere with epitaxy. Full details on this cleaning technique have been given elsewhere [1,22]. The technique relies on hydrogen passivation of the bare silicon surface. This passivation is accomplished by immersing the substrates in dilute hydrofluoric acid and then removing them without any subsequent rinse. Photoemission spectroscopy studies [25] shows that HF etching renders the silicon surface passive, free of oxide, and stabilized by the presence of a layer of chemisorbed atomic hydrogen. The hydrogen adlayer formed by an HF dip reduces the reactivity of water vapor with the silicon surface by approximately thirteen orders of magnitude [1,25]. The hydrogen adlayer acts as a chemical barrier to recontamination. This means that in UHV/CVD there is no “native oxide” to be removed

from a silicon surface prior to epitaxial growth, provided the temperature/time history of the load sequence prevents hydrogen desorption until growth is initiated.

2.5 SiGe Film Growth

After the Si wafer surfaces were hydrogen passivated, they were placed in a quartz wafer boat and inserted into the load lock chamber, which is pumped to below 1×10^{-6} torr [1]. A hydrogen flow at a corresponding pressure in excess of 200 mtorr was used during wafer boat transfer into the main chamber. This prevents the cross-contamination resulting from residual gas in the load chamber. Hydrogen flows terminated when the main chamber was isolated from the load chamber. Growth is then initiated immediately. Gaseous sources employed are silane (SiH_4), germane (GeH_4) and diborane (B_2H_6), using helium as a diluent of dopants to produce films of desired chemical content. Deposition of SiGe alloys is performed at a temperature of about 525 °C and a total operating pressure of 1.0 mtorr. With appropriate control and rapid pumping of the gas mixture mentioned above, high quality of SiGe with a $\pm 2.0\%$ chemical uniformity and film thickness better than $\pm 1.0\%$ have been demonstrated [1,25].

2.6 Polycrystalline Silicon

Polycrystalline silicon (polysilicon) is an important material in integrated circuit technology, because of its high-temperature compatibility with single-crystal Si and its semiconducting properties. For almost three decades, polysilicon has been used in metal oxide semiconductor and bipolar technologies [26,27]. It may be used as the gate in MOS

devices, as the emitter, base contact or collector contact in bipolar junction transistors, and also in on-chip resistors and capacitors, and interconnect lines [26-29]. In bipolar transistors using polysilicon emitter contact technology, control of the interface between polysilicon and a silicon substrate is one of the key parameters to insure repeatability performance [30].

Polysilicon is so useful because it forms an adherent oxide, absorbs and re-emits dopants, has good step coverage if deposited by CVD, matches mechanical properties of Si single crystal, is compatible with HF, and many other properties, which are beyond the scope of the present work. These properties make polysilicon uniquely suited to be a primary local interconnect material, and its ability to form silicides extends that application to longer interconnect lines. In the present work, the polysilicon will be used to form the emitter of both the self-aligned and non-self aligned HBT, and the external base for the non-self aligned HBT. Another advantage of the polysilicon in the present work, is to protect the field isolation from HF attack during the passivation process.

2.7 Transit Times in Bipolar Transistors

2.7.1 One-Dimensional (Vertical) Analysis

The figure of merit commonly used to characterize the high-frequency behavior of a bipolar transistor is the cut-off or transition frequency F_T . This is the frequency at which the transistor incremental current gain drops to unity. The total transit time, τ_{ec} , from the emitter contact to the collector contact of the one-dimensional bipolar transistor is related

to the transition frequency by [15,16],

$$F_t = 1/(2\pi \tau_{ec}), \quad (2.6)$$

τ_{ec} is made up of six delay components corresponding to different regions of the transistor. The physical formula of each component is given as follows [15,16].

1. Neutral emitter region charging time τ_e

$$\tau_e = W_e W_b / (F_c D_n), \quad (2.7)$$

where W_e and W_b are the emitter depth and neutral base width respectively, F_c is a grading coefficient (typically between 1 and 5), and D_n is the electron diffusion constant in the base.

2. Emitter-base junction charging time τ_{eb}

This is the delay time due to emitter-base junction capacitance, C_{JE} , times the 'diffusion resistance' r_e ($=1/g_m$, where g_m is the transconductance, $g_m = I_c/V_T$, where I_c and V_T are collector current and thermal voltage respectively),

$$\tau_{eb} = r_e C_{JE} \quad (2.8)$$

3. Base-transit time τ_b

For a Si device having uniform doping, simple theory [31,32] predicts that for a given

base width, τ_b is given by

$$\tau_b = (W_b^2 / 2 D_n) + (W_b / v_s), \quad (2.9a)$$

According to eq.2.2, there is a drift field resulting from the grading of the Ge across the quasi-neutral base. This means that the first term of eq.(2.9a) should be multiplied by such a grading factor G_f , in this case eq.(2.9a) should be read as,

$$\tau_b = (W_b^2 / 2 D_n) G_f + (W_b / v_s), \quad (2.9b)$$

where $G_f = (2kT/\Delta E_g)[1 - (kT/\Delta E_g)(1 - \exp(-\Delta E_g/kT))]$. In eq.(2.9b), v_s is the saturated carrier velocity and the remaining parameters have their usual meaning.

4. Collector space charge layer transit time τ_{scl}

$$\tau_{scl} = W_{scl} / 2 v_s, \quad (2.10)$$

where W_{scl} is the collector depletion width.

5. Base-collector junction capacitance charging time τ_{cb}

$$\tau_{cb} = r_e C_{JC}, \quad (2.11)$$

where C_{JC} is the base-collector junction capacitance.

6. RC time constant for base-collector junction τ_c

$$\tau_c = r_c C_{JC}, \quad (2.12)$$

τ_c is the time constant for charging the base-collector capacitance C_{JC} through the collector resistance. Here r_c is the undepleted collector resistance, i.e the resistance of the collector region between the base-collector space-charge layer and the heavily doped substrate or buried layer.

There is an additional delay τ_{ϕ_e} due to free carrier charge in emitter-base space-charge layer [15]. There is no explicit analytical formula for this delay. According to Roulston [15], this delay is very small and can be obtained from BIPOLE3 simulation [33].

Thus, the total delay, τ_{ec} , is given by adding all of the above components:

$$\tau_{ec} = W_e W_b / (F_c D_n) + r_c C_{JE} + (W_b^2 / 2 D_n) G_f + (W_b / v_s) + W_{sc1} / 2 v_s + r_c C_{JC} + r_c C_{JC} \quad (2.13)$$

After manipulating the above equation, the final total delay becomes,

$$\tau_{ec} = W_e W_b / (F_c D_n) + (C_{JE} + C_{JC}) / g_m + (W_b^2 / 2 D_n) G_f + (W_b / v_s) + W_{sc1} / 2 v_s + r_c C_{JC} \quad (2.14)$$

The contribution of each delay component to the total delay will be discussed in chapter four according to the specification imposed by each design. The simulation results

obtained from the BIPOLE3 [33] program will give us an approximate value for each component of the delay. It should be noted that the delay due to the substrate capacitance, C_{JSC} , which is equal to C_{JSC} times r_c was not included in the total delay given by eq.2.14, and will be addressed in the two-dimensional analysis given below.

The second important parameter of the bipolar transistor is the maximum oscillation frequency, F_m . It is defined as the frequency at which the maximum available power gain of the transistor drops to unity. For bipolar transistors a convenient approximation for F_m is given by eq.2.5.

2.7.2 Two-Dimensional (Lateral) Analysis

In the vertical (one-dimensional) analysis all the delays given by eqs. 2.7-2.12 were computed as “charge density/current density” [15]. In addition, the transition frequency F_t in vertical analysis was determined directly from the sum of all six vertical delay times given by eq.2.7 to eq.2.12. F_t in a real transistor is always less than would be predicted using eqs.2.6 and 2.14 because these equations do not take into account the following factors: (1)sidewall injected charge and current, (2)sidewall capacitance for the emitter-base junction, (3)extrinsic base region capacitance, and (4)collector substrate resistance and capacitance (for IC devices).

It is very difficult to incorporate the above four factors in simple “hand calculations” of BJT performance. Lateral effects can, however, be taken into account using the BIPOLE3

program. For this purpose the BIPOLE3 manual [33] defines a new set of delay times in lateral analysis, and these are given as follows:

(1) τ_{AVO} : is a small signal delay time due to the integral of minority carrier charge in the whole of the active (intrinsic) region. This include the emitter, base, emitter-base and base-collector space charge layers. The delay time τ_{AVO} is related directly to the vertical simulation delay time. In the BIPOLE manual, the delay time τ_{AVO} is defined as the delay due to the intrinsic region diffusion capacitance C_{dif} . Roulston [15] has derived an equation relating C_{dif} to the delay times τ_e , τ_{qbe} , τ_b , τ_{scI} and the transconductance g_m as follows:

$$C_{dif} = g_m (\tau_e + \tau_{qbe} + \tau_b + \tau_{scI}), \quad (2.15)$$

(C_{dif} / g_m) is equivalent to the delay τ_{AVO} , thus

$$\tau_{AVO} = \tau_e + \tau_{qbe} + \tau_b + \tau_{scI} \quad (2.16)$$

(2) τ_{AVL} : is the delay time due to sidewall injected electron concentration into the base region outside the active (intrinsic) base region.

(3) τ_{ESC} : is the delay time due to minority carrier charge in the emitter-base space charge layer.

(4) τ_{RCE} : is the delay time due to C_{JC} and total collector series resistance. This delay is equivalent to total junction capacitance divided by small signal transconductance g_m .

(5) τ_{REE} : is the delay time due to $(kT/q)(C_{JE} + C_{JC})$, where q is the electronic charge and the remaining parameters have their usual meaning.

(6) τ_{SUB} : is the delay time due to $(C_{JSC} R_{cont}/\beta)$, where R_{cont} is the vertical resistance under the collector contact, and β is the current gain.

Adding all of the above delay times, the total delay time τ_{tot} in a two-dimensional analysis is thus given by,

$$\tau_{tot} = \tau_{AVO} + \tau_{AVL} + \tau_{ESC} + \tau_{RCE} + \tau_{REE} + \tau_{SUB} \quad (2.17)$$

The transition frequency F_t is then given by,

$$F_t = 1/(2 \pi \tau_{tot}) \quad (2.18)$$

In many practical devices, lateral effects must be taken into account to accurately predict F_t . In chapter four, BIPOLE3 program will be used to simulate the HBT and BJT devices using both vertical and lateral analysis.

CHAPTER 3

DEVICE ARCHITECTURE

3.1 Transistor Structures

Two basic HBT structures were developed in the present work. Both structures follow a relatively conventional bipolar process flow prior to base formation. Devices are built over a heavily doped n^+ buried layer diffused into a lightly doped p-type substrate. A much more lightly doped n^- epitaxial (epi) collector layer is grown over the buried layer. The buried layer is contacted with a deep collector sinker diffusion. Transistors are isolated with a deep boron diffusion penetrating through the thickness of the epi collector to the p-type substrate. The active region in which the base and emitter are formed is defined by semi-recessed local oxidation of the silicon (LOCOS) isolation. A layer of p^+ polysilicon ("field poly") extends from the active region over the field oxide to allow contact to the base to be made in the field region.

The base region is formed by the UHV/CVD technique described in chapter two. The deposited base layer grows epitaxially in the active region, and adds to the thickness of the field poly in the field region. A polysilicon emitter is used. Silicidation with platinum is used to reduce the resistance of the emitter stripe and exposed extrinsic base regions both in the active region and over the field.

The two device structures considered differ in the alignment of the silicided base to the emitter. The first structure, which will be referred to as the “self-aligned” (SA) structure, is shown in cross-section Fig.3.1. In this device the silicided base region is self-aligned to the emitter polysilicon stripe. The low-resistance silicided extrinsic base is separated from the intrinsic base by just the width of the oxide spacer formed on the emitter sidewall, providing a very low value of extrinsic base resistance. Use of a deposited base allows base contact to be made over the field oxide, so that the extrinsic portion of the base-collector junction on either side of the emitter stripe has a width equal to the minimum lithographic dimension ($1.5\ \mu\text{m}$). This provides a considerable reduction in C_{JC} compared to a typical single-poly self-aligned BJT [34]. The main drawback of the self-aligned structure is the risk that the extrinsic SiGe base in the active region will be partly or completely consumed during the emitter poly etch. This risk is compounded by the need to ensure removal of the emitter poly “stringer” from the field poly sidewall, and the fact that it is difficult to devise a plasma etch chemistry that attacks polysilicon but not SiGe. If the emitter poly etch breaks through the SiGe base, contact to the base is lost. Similar risks are associated with the formation of the oxide spacer on the emitter sidewall.

The non-self aligned (NSA) structure shown in cross-section in Fig.3.2 avoids the problem of emitter poly etch endpointing by extending the emitter polysilicon over the edge of the device well. In this structure the active emitter area is defined by opening a window in a deposited oxide separating the emitter polysilicon from the base. This

Figure 3.1 Cross-sectional view of the self-aligned HBT structure

Figure 3.2 Cross-sectional view of the non-self aligned HBT structure

structure is much more tolerant to process variation, but suffers from increased base resistance since the silicided region does not extend to the edge of the intrinsic base.

3.2 Device Layout

The layout of each SA and NSA structure, which is presented below, consists of several mask levels for the process steps. The L-Edit program [35] was used to design the required mask layout. In the current HBT technology, the 1X npn has a single poly emitter with active area $1.5 \times 3.0 \mu\text{m}^2$. The HBT process is optimized around the emitter polysilicon. Figures 3.3 and 3.4 illustrate the mask levels for each structure. Part of the layers shown in Figs.3.3 and 3.4 will be described in detail in chapter five, whereas layers common to Gennum's existing bipolar process will not be described. The main difference between the SA and NSA structures is the difference in the base contacting scheme. Figure.3.3 illustrates the layout of the SA structure, where the base of the transistor consists of two components. The first component is due to the intrinsic base situated directly underneath the emitter poly, and the second component is due to the extrinsic base situated just outside the emitter poly, and bounded by the field poly.

Two base contacts were introduced to a 1X npn self-aligned and non self-aligned structures (Fig.3.3 and 3.4), and were placed laterally on each side of the emitter poly without compromising the device width. The main feature of the two-base contact devices, is that the areas of these devices are smaller than the one contact device. In this case the parasitic capacitances were reduced, therefore enhancing the device performance.

Figure 3.3 Self-aligned HBT with two base contacts

Figure 3.4 Non-self aligned HBT with two base contacts

CHAPTER 4

DEVICE SIMULATIONS AND ANALYSIS

4.1 Introduction

Simulation of any semiconductor device before fabrication is a very important step, because preliminary performance estimates can be obtained from simulations very quickly, and at much lower expense than the results obtained from the fabrication processes. The complete process flow is often very lengthy, usually taking several weeks and sometimes several months depending on the number of the masks used and the complexity of each process step. Two types of simulation software are commonly used in the development of new semiconductor technologies: process simulators and device simulators. Process simulators predict structural details such as doping profiles and the thickness of oxide layers on the basis of process parameters such as implant doses and furnace cycle times and temperatures. The most widely used simulators in this category are SUPREM3 and SUPREM4 [36]. Device simulators take structural information as inputs, and predict device electrical characteristics. Modern simulators can provide quite accurate predictions of device characteristics, particularly if there has been an opportunity for calibration. Even without a calibration, the simulator program still provides valuable information compared to the results obtained from the analytical equations.

In the present work, BIPOLE3 has been selected as a device simulator program to design

structures for the new SiGe HBT technology. BIPOLE3 [33] was developed by Prof. D.J. Roulston and co-workers at the University of Waterloo, Ontario, to provide a convenient and rapid means of predicting terminal electrical characteristics of bipolar transistors from input consisting of fabrication data such as impurity profiles, mask dimensions and recombination data. This would help gain insight into critical device characteristics such as the transition frequency, oscillation frequency, current gain, Early voltage and most of the parasitic resistances and capacitances of the transistor. It should be noted that BIPOLE3 does not have models precisely describing the novel HBT structures considered here, particularly in regard to the extrinsic base. However, in terms of the simplicity, cost and speed, BIPOLE3 is arguably the best simulator that adequately describes the present HBT.

The BIPOLE3 bipolar device simulation program is based on obtaining a set of differential equations describing current flow in the vertical (emitter-base-collector) direction and a separate differential equation describing current flow in the lateral direction [37,38]. In the vertical 'x' direction, the device is divided into five regions: (1) neutral emitter, (2) emitter-base space charge layer, (3) neutral base, (4) base-collector space charge layer, and (5) neutral collector. In the three quasi neutral regions, Poisson's equation is not solved, but only the transport and continuity equations. The solution for majority carrier current is then performed in the lateral 'y' direction in the neutral base region, using the results obtained from vertical analysis. The aim of the whole analysis is to predict the electrical characteristics in terms of the physical parameters. The details of

the theoretical analysis in the vertical and lateral directions are beyond the scope of this thesis and are given in the references [37,38]. All of the following results were based on a up-to-date version of the BIPOLE3 program (v.2.5.0, 23 March 1997 [33]).

4.2 Electrical Performance Specifications

One of the objectives of the present work is to determine doping and Ge concentration profiles for a SiGe base HBT meeting the following electrical specifications:

1. $F_t = F_m = 20$ GHz at $I_c = 200$ μ A and $V_{ce} = 1.0$ V.
2. A breakdown voltage $BV_{ceo} \geq 5$ V.
3. A current gain β of approximately 200.

These electrical specifications were to be met with an equipment set providing a minimum lithographic feature size of 1.5 μ m. The target emitter window size for a “1X” device was 1.5 μ m x 3.0 μ m.

For a 1X device a collector current of 200 μ A corresponds to collector current density J_c of just 4.4×10^3 A/cm². In contrast the high performance devices reported by researchers at IBM [12] reach peak F_t at $J_c \sim 10^5$ A/cm². Further, these devices typically have breakdown voltages of 2 to 3 V. The design problem faced here is therefore very different to that encountered in developing an HBT intended to provide maximum speed without concern for power dissipation or break down voltage BV_{ceo} .

4.3 Calculation of the Extrinsic Base of the NSA Devices

Before modeling the NSA devices, it is worthwhile to shed the light on the BIPOLE3 calculations of the external base resistance for the SA devices. BIPOLE3 does not have models precisely describing the novel HBT structures considered here, particularly in regard to the extrinsic base. Figure 4.1 shows the extrinsic base cross-section of an npn SA device where the P^+ region is separated from the emitter diffusion by a distance ESBX. The deficiency in the BIPOLE3 program is that the sheet resistance in the ESBX region is computed from the intrinsic region impurity profile, and this approximation does not reflect the true impurity profile in our SA devices. This means that the total measured base resistance may differ from the predicted base resistance. In the case of NSA devices, the BIPOLE3 program can not be used unless the extrinsic base resistance is modeled, because it is only suitable for SA devices. For this reason two approximate models were suggested and used throughout the simulations, and their results are presented in the following sections. Below is a brief discussion of each model.

Figure 4.1 Extrinsic base cross-section [33]

4.3.1 Model 1 (Overestimates Base Resistance)

Model 1 overestimates the extrinsic base resistance of the NSA devices. In this model there are three parameters. These parameters are used by BIPOLE3 program, and have the following meaning:

(1) $REXTSQ=4.0 \Omega/sq$, it represents the silicided region of the extrinsic base sheet resistance according to Fig.4.2. Figure 4.2 shows part of the layout of the NSA device illustrated in Fig.3.4. The magnitude of the $REXTSQ$ is given above, based on previous measurements of silicided devices.

(2) ESB: is the distance from the emitter window to the base contact metal.

(3) ESBX: is the distance from the emitter window to emitter poly edge.

In model 1 shown in Fig.4.2, the value of the $REXTSQ$ was fixed during simulation, whereas the values of ESB and ESBX were varied depending on the geometry of the device. The drawback of this model comes from the fact that BIPOLE3 computes the sheet resistance over distance ESBX from the intrinsic base region impurity profile directly under the emitter. These calculations are not accurate, because the impurity level under the emitter poly is less than that outside it, along the path towards the base contact, which is situated outside the field oxide. This is the reason that the simulation gives a higher value for the external base resistance than is actually anticipated.

Figure 4.2 Model 1 for the NSA device

Figure 4.3 Model 2 for the NSA device

4.3.2. Model 2 (Underestimates Base Resistance)

In model 2 shown in Fig.4.3, the parameters REXTSQ and ESB have different meaning as given below.

(1)REXTSQ: describes the sheet resistance of the field poly and SiGe poly stack. From short loop experiments, the value of the REXTSQ for polysilicon (thickness=3000 Å) was found to be 165 Ω/sq . It is expected that this value will be reduced if the thickness of the SiGe layer is added to the thickness of the field poly, because the sheet resistance is inversely proportional with the thickness.

(2)ESBX: is the distance from the emitter window to field poly edge.

The ESB has the same meaning as given in model 1. In model 2, it is assumed that the silicided region of the HBT, which is just outside the field poly, has negligible resistance, because it is insignificant compared with the remaining components of the total extrinsic base. In model 2, the magnitudes of the ESB and ESBX also depends on the geometry of the device. In the following analysis, both models were employed to calculate the extrinsic base resistance of the NSA devices.

4.4 Delay Times Contributing to the Transition Frequency

In order to gain insight into the mechanisms that govern F_t , vertical and lateral analyses have been conducted under various conditions to determine the magnitude of each delay component that contributes to F_t in the HBT.

4.4.1 Preliminary Calculations

The transition frequency given by eq.2.6 contains several components representing the delay times in various regions of the transistor. Before performing detailed BIPOLE3 simulations, it is desirable to evaluate each component in order to identify which delay time makes a significant contribution to the F_T .

Preliminary calculations have been carried out to select doping levels to serve as a starting point for BIPOLE3 simulations. The doping levels in the emitter, base and the collector were found to be equal to $3 \times 10^{20} / \text{cm}^3$, $1 \times 10^{18} / \text{cm}^3$ and $2.8 \times 10^{16} / \text{cm}^3$, respectively. The following one-dimensional hand calculations reveal why these doping levels were chosen. The self-aligned SA1X device was chosen for this analysis.

1. *Charging time in the quasi-neutral emitter region τ_e*

From eq.2.7, $\tau_e = W_e W_b / (F_c D_n)$,

Here the emitter stripe depth $W_e = 500 \text{ \AA}$, $W_b = 450 \text{ \AA}$, $D_n = 2.3 \text{ cm}^2/\text{s}$ and $F_c = 5$, so the delay time $\tau_e = 1.94 \text{ ps}$. To reduce this delay either W_e or W_b or both should be reduced. The emitter depth can be reduced using a well controlled rapid thermal annealing process instead of the conventional furnace anneal. Reducing the base width below $450 \text{ \AA}$ is limited by base punch through, and is not practical. It should be noted that there is a composition factor, resulting from germanium-induced band gap reduction at the base-emitter depletion edge [39]. The structure studied here has no Ge content at the base emitter junction, so this factor was not included in the calculation of τ_e . According to

Patton *et al* [39], the composition factor gives a significant reduction in τ_e . However, BIPOLE3 program takes into account such factor during the simulation.

2. Emitter-base junction charging time τ_{eb}

From eq.2.8, $\tau_{eb} = C_{JE} / g_m$,

where $g_m = I_c / V_i$, for $I_c = 200 \mu A$ and $V_i = 0.0259 V$, thus $g_m = 7.7 \times 10^{-3} S$. The emitter-base junction capacitance C_{JE} is given by [15],

$$C_{JE} = \epsilon_{Si} A_{eb} / X_{eb}, \quad (4.1)$$

where $\epsilon_{Si} = 11.9 \epsilon_o$ (ϵ_o is the permittivity of free space) is the silicon permittivity, A_{eb} ($\approx 4.5 \mu m^2$) is the emitter-base junction area and X_{eb} is the width of the junction. By using eq.A1.1 (Appendix A), $X_{eb} = 0.02 \mu m$ for $V_{eb} = 0.77 V$. Substituting the values of ϵ_{Si} , A_{eb} and X_{eb} into eq.4.1, $C_{JE} = 23.6 fF$, thus the delay time $\tau_{eb} = 3.0 ps$. This is a very significant delay which would slow down the device. In order to minimize it, the doping level in the base must be reduced to a minimum. This is one of the main reasons which led to the use of a light doping level in the base of the transistor.

3. Base-transit time τ_b

From eq.2.9b, $\tau_b = (W_b^2 / 2 D_n) G_f + (W_b / v_s)$,

G_f is the grading factor taken from eq.2.9b. For $W_b = 450 \text{ \AA}$, $D_n = 6.78 \text{ cm}^2/\text{s}$, and the reduction in the base band gap $\Delta E_g = 0.075 \text{ eV per } 10\% \text{ Ge}$ [12], thus $\tau_b = 1.13 ps$. This delay is also important and can only be reduced by decreasing the width of the base. In

the present work it is difficult to reduce the W_b below 450 Å, because of the occurrence of base punch through, as explained earlier. Also, since minimizing C_{JC} requires a lightly doped base, if W_b is too small the base resistance will be excessive and F_m will be degraded.

4. Collector space charge layer transit time τ_{sc1}

From eq.2.10, $\tau_{sc1} = W_{sc1} / 2 v_s$,

Using eq.A1.2 (Appendix A), $W_{sc1} = 0.3 \mu\text{m}$ and $v_s = 10^7 \text{ cm/s}$, thus $\tau_{sc1} = 1.5 \text{ ps}$. Although this delay is not dominant, it can not be neglected. This delay can only be reduced by increasing the doping level in the collector, which would increase the junction capacitance C_{JC} .

5. Base-collector junction capacitance charging time τ_{cb}

From eq.2.11, $\tau_{cb} = r_e C_{JC} = C_{JC} / g_m$,

The base-collector junction capacitance C_{JC} is given by [15],

$$C_{JC} = \epsilon_{Si} A_{bc} / W_{sc1}, \quad (4.2)$$

The base-collector junction area A_{bc} of the device SA1X $\approx 13.0 \mu\text{m}^2$, $W_{sc1} = 0.3 \mu\text{m}$,

$C_{JC} = 4.7 \text{ fF}$, thus $\tau_{cb} = 0.6 \text{ ps}$. This delay is relatively insignificant.

6. Collector charging time τ_c

From eq.2.12, $\tau_c = r_c C_{JC}$,

For $r_c \approx 50 \Omega$, $C_{JC} = 4.7 \text{ fF}$ as given above, thus $\tau_c = 0.23 \text{ ps}$. This delay is small and can be

neglected.

By substituting all of the above delay times into eq.2.6, the F_t was found to be equal to 19 GHz. Furthermore, the above analysis indicated that the dominant delay time is due to the emitter-base junction charging time τ_{eb} . Having τ_{eb} as the dominant delay is unusual, and reflects the very low current density at which the device is to be operated. To control τ_{eb} very light base doping must be used.

The unity power gain oscillation frequency F_m is related to F_t by eq. 2.5,

$$F_m = [F_t / (8 \pi R_B C_{JC})]^{1/2},$$

The above equation shows that F_m depends on the parasitic elements ($R_B C_{JC}$). To increase F_m , both R_B and C_{JC} must be decreased. Decreasing the C_{JC} is possible by increasing the W_{sc1} through decreasing in the collector doping. The base resistance R_B can be decreased by increasing the doping level in the base, and this degrades F_t because the junction capacitance C_{JE} will increase and ultimately introduce an unacceptable delay in the τ_{eb} . This case is much worse in the non-self aligned devices, because R_B in these devices is much larger than R_B in the self-aligned devices.

In all of the above analysis, the lateral effects that have been discussed in section 2.7.2 were not included in the hand calculations, because they are difficult to include analytically. It is possible, however, to include these effects using lateral analysis in the BIPOLE3 program.

4.4.2 Simulation Results

The purposes of the present simulations are two-fold. The first is to select the best doping profiles and junction depths in order to meet the required specifications. The second is to determine the effect of device geometry on performance to determine the benefits which might be expected with improved lithographic resolution. Several dimensions for the length and width of both emitter mask and field poly (polysilicon), in addition to the emitter poly width, have been investigated and are illustrated in Table 4.1. In Table 4.1, SA and NSA are labels used to define the self-aligned and non-self aligned transistors respectively.

In order to simulate the device characteristics using BIPOLE3, several input parameters must be provided. The most important parameters are summarized as follows: the doping in each region of the transistor; whether the doping is uniform or graded across each region (in the present simulation a uniform profile was considered in each region); the mask dimensions and device geometry; thickness of the epitaxial layer; width of the transistor base and the Ge concentration and its profile across the base; the depths of the emitter-base and base-collector junctions. In addition, appropriate values for the control parameters of BIPOLE3 program must be specified. The metallurgical junction depth X_{j2} , which is required in the simulation, was calculated using the following procedure:

Table 4.1. Mask dimensions of the SiGe HBTs

Device Name	E ¹ .mask length μm	E ¹ .mask width μm	F ² .poly width μm	F ² .poly length μm	E ¹ .poly width μm
SA05X	2.5	0.5	1.5	4.0	0.5
SA1X	3.0	7.0	3.0	4.5	1.5
SA1X125	3.0	6.75	2.25	4.5	1.25
SA1XFP25	3.0	6.5	2.5	4.5	1.5
SA125X125	1.25	6.75	2.25	4.5	1.25
SA125X25	2.5	6.75	2.25	4.5	1.25
SA15X15	1.5	7.0	3.0	4.5	1.5
NSA1X	3.0	1.5	3.0	4.5	4.0
NSA1X125	3.0	1.25	2.25	4.0	3.25
NSA1XFP25	3.0	1.5	2.5	4.5	3.5
NSA125X125	1.25	1.25	2.75	2.75	3.75
NSA125X25	1.5	1.25	2.75	3.0	3.75
NSA15X15	1.5	1.5	3.0	3.0	4.0

1: E means emitter

2: F means field

The width of the depletion region at the emitter-base junction X_{eb} and at the base-collector junction X_{bc} was calculated using a step junction-type model. The metallurgical junction depth at the base-collector region (X_{J2}) is the sum of the parameters X_{J1} , W_b , X_{eb} , and X_{bc} . In the present work, the X_{J1} (the metallurgical junction depth at emitter-base region) was considered to be equal to $0.05 \mu\text{m}$ and the width of the quasi-neutral base W_b was $0.045 \mu\text{m}$. The junction depths X_{eb} and X_{bc} depend on the dopant concentration in each region, and can be calculated analytically using eqs.A1.1 and A1.2 using a depletion approximation approach. In this case X_{J2} can easily be obtained for a given value of W_b . The depletion layer capacitances' C_{JE} and C_{JC} were calculated using eqs. A2.1 and A2.2 respectively.

4.4.2.1 Effect of Doping Levels on the Delay Times

4.4.2.1.1 Vertical Analysis

By adjusting the vertical profile of a HBT structure, it is possible to extract several important parameters which describe the performance of the device. Several device geometries have been simulated, and the results of two typical HBT SA125X and SA1X devices are presented in Table 4.2. This table shows how the delay times change as the doping levels in the active regions of the transistor change. The changes in the delay times τ_e , τ_{qbe} and τ_{cb} are insignificant. This means that the main components of the delay times are due to the sum of the emitter-base junction charging time, τ_{eb} , the base transit time τ_b , and the collector space charge transit time, τ_{sc1} . The base transit time seems to be independent of the variation of the doping levels, whereas the delay times τ_{eb} and τ_{sc1} are

slowly changing with the doping levels. The delay time τ_{eb} can be reduced by decreasing the emitter-base capacitance C_{JE} , and this can be achieved by using a lightly doped base. This is the only profile-parameter that can be used to minimize the effect of C_{JE} on the performance of the devices. On the other hand, it is possible to reduce the effect of the C_{JE} by reducing the size of the device as in the case of the SA125X device in order to increase the collector current density at $I_c = 200 \mu A$. More detail on this point will be given in section 4.6.

The base transit time can be decreased by decreasing the thickness of the SiGe layer but at the expense of the break down voltage BV_{ceo} (due to the increase in the current gain β), and also at the risk of base punch through. As mentioned earlier, the minimum base width used in the present simulations was $450 \text{ \AA}$. The collector space charge layer time τ_{sc1} , can be reduced by decreasing the collector depletion width W_{sc1} . This can be achieved by increasing the doping level in the collector. The consequences of this procedure is that the break down voltage BV_{ceo} will decrease and the base-collector junction capacitance C_{JC} will increase. These results will be demonstrated in section 4.5. Table 4.2 indicates that as the size of the transistor changed, the delay time τ_{eb} significantly increased. According to eq.4.1, the junction area of the device SA125X is smaller than that of the device SA1X, thus the magnitude of the junction capacitance C_{JE} for the SA125X device is smaller than that of the SA1X device. The magnitudes of the C_{JE} for these devices will be given in section 4.6.

Table 4.2 Effect of doping levels on the delay times obtained from vertical analysis

Doping Levels /cm ³	Device Name	τ_e ps	τ_{cb} ps	τ_{gbe} ps	τ_b ps	τ_{sci} ps	τ_{cb} ps	F_t GHz
ED ¹ = 6.0×10^{20} BD ² = 1.0×10^{18} EPI ³ = 2.8×10^{16}	SA125X SA1X	0.160 0.160	3.190 4.110	0.020 0.023	2.000 2.000	1.796 1.700	0.081 0.087	21.96 19.70
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X SA1X	0.209 0.203	3.350 4.320	0.031 0.034	2.098 2.065	1.500 1.600	0.110 0.103	21.80 19.11
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 1.0×10^{16}	SA125X SA1X	0.198 0.193	3.599 4.070	0.031 0.033	2.040 2.016	2.300 2.300	0.020 0.021	19.43 18.43

Table 4.3 Effect of doping levels on the delay times obtained from lateral analysis

Doping Levels /cm ³	Device Name	τ_{AVO} ps	τ_{AVL} ps	τ_{RCE} ps	τ_{REE} ps	F_t GHz	C_{diff} fF	g_m x10 ⁻² S
ED = 6.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X SA1X	4.010 3.980	0.295 0.244	0.128 0.140	3.530 4.200	19.98 18.58	29.05 29.28	0.720 0.735
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X SA1X	3.990 4.020	0.304 0.253	0.141 0.152	3.600 4.360	19.80 18.11	28.32 29.62	0.709 0.728
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 1.0×10^{16}	SA125X SA1X	4.480 4.500	0.342 0.285	0.032 0.038	3.400 4.220	19.28 17.60	31.00 31.92	0.699 0.709

1: ED means emitter doping

2: BD means base doping

3: EPI means epitaxial collector doping

4.4.2.1.2 Lateral Analysis

The lateral simulations are important because they represent the true terminal characteristics of the device. It is also interesting to compare the results obtained from this analysis where the lateral effects were included to that obtained from the vertical analysis where the lateral effects were neglected.

The simulation results of the lateral analysis of the two devices SA125X and SA1X are presented in Table 4.3. The delay times τ_{ESC} (due to the minority carrier charge in the emitter-base space charge layer) and τ_{SUB} are $2-3 \times 10^{-15}$ seconds and 6×10^{-17} seconds respectively, and thus can be ignored. The delay time τ_{AVL} due to sidewall injection diffusion capacitance, and the delay τ_{RCE} due to C_{JC} and total collector series resistance are also small but are much larger than the delay times τ_{ESC} and τ_{SUB} . The dominant delay times are due to the τ_{AVO} and τ_{REE} . These delays have already been discussed in section 2.7.2. It should be noted that the size of the device SA125X is slightly smaller than the device SA1X, and therefore the total delay times of the former device are always shorter than the latter device, thus its transition frequency F_t is larger.

The results obtained from the vertical and lateral analysis, which are presented in Tables 4.2 and 4.3 respectively, are correlated with each other. The delay times τ_{REE} and τ_{RCE} obtained from the lateral analysis can be correlated with the delays τ_{cb} and τ_{cb} respectively. The magnitudes of the delays τ_{REE} and τ_{RCE} only are slightly larger than the delays τ_{cb} and τ_{cb} . This means that the lateral effects are insignificant at the low collector

currents investigated here. The magnitude of the delay time τ_{AVO} was found to be higher by 5-9 % than the magnitude of the sum of the four delay times τ_e , τ_{qbe} , τ_b and τ_{scl} . In other words, the magnitude of the F_t obtained from the vertical analysis is higher by 5-9% than F_t obtained from the lateral analysis. This relatively small difference between the results of vertical and lateral analysis can be explained in terms of sidewall capacitance which is not taken into account in the vertical analysis. Emitter de-biasing [15,40], which could also cause differences between vertical and lateral analysis, is not likely to be important at the low collector currents considered here. For $I_c = 200 \mu A$, $I_b = 1 \mu A$, so for a base resistance of 200Ω , V_{be} at the center of the emitter stripe is approximately $200 \mu V$ less than at the edge. This difference in V_{be} is negligible.

4.4.2.2 Effect of Transistor Size on the Delay Times

The influence of the device scaling on the delay times is presented in this section. Two devices having different geometries, SA125X25 and S15X15, were selected. The analysis of the remaining devices was found to be similar, and so will not be repeated here. The effect of device scaling on other transistor parameters will be discussed in section 4.6.

4.4.2.2.1 Vertical Analysis

The results of two devices SA125X25 and S15X15 are presented in Table 4.4. Table 4.4 indicates that the delay times, τ_e , τ_{qbe} , and τ_{cb} are insignificant. The remaining delays are similar to those results presented in Table 4.2 with the following differences. The magnitudes of all the delay times for all the devices of Tables 4.2 and 4.4, which had

different geometries, except the delay τ_{eb} due to the emitter-base junction capacitance, are almost independent of the transistor size. The τ_{eb} was found to be strongly dependent on the size of the transistor.

The preliminary vertical analysis indicated that the device SA15X15, which has the smallest size among the devices listed in Tables 4.2 and 4.4, has the smallest value of τ_{eb} . The reduction in the τ_{eb} can be attributed to the reduction in the emitter-base capacitance C_{JE} . The magnitudes of the junction capacitance C_{JE} for each device obtained from the vertical analysis will be presented in section 4.6.

4.4.2.2 Lateral Analysis

The simulation results of the lateral analysis for the devices SA125X25 and SA15X15 are presented in Table 4.5. The delay times τ_{ESC} and τ_{SUB} (not shown in Table 4.5) are in the order of $(2-3) \times 10^{-15}$ seconds and $(5-8) \times 10^{-17}$ seconds respectively, and thus can be neglected. The delay times τ_{AVL} and τ_{RCE} are also insignificant. Thus, the remaining delays τ_{AVO} and τ_{REE} are the main components of the total delay. All these observations are similar to those results presented in Table 4.3. The delay time τ_{REE} which is related to the τ_{eb} in the vertical analysis is strongly dependent on the size of the transistor. It decreased as the dimension of the device decreased. This means that there is a consistency between the results obtained from the vertical and lateral analysis with the lateral results being more accurate than that obtained from the vertical analysis. It should also be noted that the magnitude of the delay time τ_{AVO} obtained from the lateral analysis

Table 4.4 Effect of a transistor size on the delay times obtained from vertical analysis

Doping Levels /cm ³	Device Name	τ_e ps	τ_{cb} ps	τ_{qbc} ps	τ_b ps	τ_{sc1} ps	τ_{cb} ps	F_t GHz
ED = 6.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X25 SA15X15	0.160 0.160	2.847 2.198	0.018 0.016	2.018 2.064	1.736 1.829	0.077 0.066	23.21 25.13
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X25 SA15X15	0.212 0.216	2.998 2.265	0.029 0.025	2.100 2.100	1.556 1.694	0.102 0.081	22.75 24.94
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 1.0×10^{16}	SA125X25 SA15X15	0.203 0.211	3.038 2.214	0.029 0.025	2.070 2.100	2.230 2.200	0.020 0.019	20.97 23.51

Table 4.5 Effect of a transistor size on the delay times obtained from lateral analysis

Doping Levels /cm ³	Device Name	τ_{AVO} ps	τ_{AVL} ps	τ_{RCE} ps	τ_{REE} ps	F_t GHz	C_{diff} fF	g_m $\times 10^{-2}$ S
ED = 6.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X25 SA15X15	4.210 4.514	0.314 0.284	0.119 0.134	3.029 2.516	20.74 21.36	29.94 31.27	0.710 0.691
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 2.8×10^{16}	SA125X25 SA15X15	4.130 4.490	0.314 0.286	0.142 0.160	3.165 2.660	20.53 20.95	28.80 31.12	0.698 0.679
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 1.0×10^{16}	SA125X25 SA15X15	4.450 4.510	0.342 0.292	0.034 0.043	3.030 2.880	20.26 20.60	31.15 31.06	0.694 0.681

is in agreement with the prediction of eq.2.16 with minor differences equal to about 5-9%. This difference is expected because there are several factors which are not taken into account during the vertical analysis. The results presented in Table 4.5 also indicated that the magnitude of the delay τ_{AVO} is equal to (C_{dif}/g_m) , and this result is consistent with eq.2.15 of chapter two.

4.5 Effect of Doping Levels on the Electrical Characteristics

In section 4.4.1 the delay time in each region of the transistor was identified using hand calculations. Several solutions were suggested to minimize the delay times. In this section the BIPOLE3 program will be used to examine the results discussed in section 4.4.1. For this reason, extensive investigations have been carried out in order to identify the best doping levels which meet the required specifications that have been mentioned earlier in section 4.2.

In order to find out the effect of doping levels on the transistor performance in a certain region, it is required to change only the doping levels in that region. In the present investigation, the doping level in the base was kept constant, and equal to $1.0 \times 10^{18}/\text{cm}^3$. The selection of the doping level in the base was not arbitrary, it was based on both the hand calculations (section 4.4.1) and the preliminary simulation results. It was found that increasing the doping level in the base reduces the base resistance and increases the F_m significantly, but it degrades the F_t because C_{JE} increases. This point has already been discussed in the earlier sections.

The simulation results of the best doping levels, which meet the specifications, in the emitter, base and collector regions of the HBT are illustrated in Tables 4.6 to 4.9. The analysis given in these tables is for 1X SA1X and NSA1X devices. These devices can be fabricated using a simple process flow and the lithographic equipment available at Gennum. For those devices having geometries different from the devices SA1X and NSA1X, which are presented in Tables B.1-B.4 (Appendix B), it is also possible to fabricate some of them using very fine adjustment during the alignment and lithographic processes. The Tables B.1 to B.4 contain information about the transistor parameters for the self-aligned and non-self aligned structures of various geometries. These Tables indicate that there are several transistors which meet and exceed the required specifications. Among the best performance transistors which significantly exceeds the specifications was the SA05X. This device, however, can not be made with the current available equipment. The 0.5 μm poly-emitter stripe width of the device SA05X can only be defined using 0.5 μm lithography equipment or an electron beam direct write technique, which is currently available at Carleton University.

4.5.1 Emitter Doping Level

The results of simulations for SA1X and NSA1X devices are presented in Tables 4.6 and 4.7. The remaining devices are given in Table B.1 of Appendix B. In Tables 4.6, 4.7 and B.1, all the parameters have their usual meaning except the following: R_{bex} is the extrinsic base resistance, and $(F_t)_p$ and $(F_m)_p$ are the peak values of the transition and oscillation frequencies respectively, and $(I_c)_p$ is the value of collector current giving highest F_t .

The simulation results indicated that when the emitter doping levels were changed from $6 \times 10^{20}/\text{cm}^3$ to $3 \times 10^{20}/\text{cm}^3$, the F_i and F_m decrease accordingly. The main reason for these changes is that the parasitic capacitances at the emitter-base junction C_{JE} was slightly increased. The increases in the C_{JE} can be attributed to the decrease in the width of the depletion region X_{eb} , since the latter depends on the doping levels according to eq.A1.1 (Appendix A). However, the changes in F_i and F_m were insignificant ($\sim 1\text{-}2\%$) and can be neglected. Furthermore, there is a noticeable increase in the current gain β as a result of the reduction in the emitter doping level. This result is in agreement with the Roulston predictions [15], in that in simple theory β is proportional to the emitter doping level in a vertical npn transistor.

4.5.2 Epitaxial Collector Doping Level

The Tables 4.7-4.8 and B.2-B.3 (Appendix B) indicated that when the doping level in the collector (epitaxial layer, usually referred to as EPI) decreased from $2.8 \times 10^{16}/\text{cm}^3$ to $1.0 \times 10^{16}/\text{cm}^3$, the magnitudes of the following parameters are changed accordingly as described below.

(1) Table 4.2 indicated that τ_e and τ_{eb} decreased by about 5% and 6% respectively, whereas τ_{sc1} increased by about 30%. In spite of these changes, the decrease of F_i was insignificant.

Table 4.6 Heavy emitter doping

$I_c = 200 \mu A$	$XJ_1 = 0.05 \mu m$	$BV_{ceo} = 6.0 V$
$V_{bc} = 1.0 V$	$XJ_2 = 0.1384 \mu m$	$BV_{cbo} = 18.7 V$
$ED = 6.0 \times 10^{20} cm^{-3}$	$W_b = 0.045 \mu m$	Reach through = 4.19 V
$BD = 1.0 \times 10^{18} cm^{-3}$	$TEPI = 0.59 \mu m$	Ge = 10%
$EPI = 2.8 \times 10^{16} cm^{-3}$		

Parameters	SA1X SiGe	NSA1X SiGe	SA1X ⁺ Si
F_t (GHz)	18.9	18.9	14.1
F_m (GHz)	18.6	12.0, 14.0*	17.4
R_b (Ω)	190	190	620
R_{bex} (Ω)	260	910, 610*	150
C_{JE} (fF)	19.59	19.59	17.93
C_{JC} (fF)	4.52	4.52	2.9
C_{JSC} (fF)	12.1	12.1	10.8
$(F_t)_p$ (GHz)	26.5	26.5	17
$(F_m)_p$ (GHz)	23.7	17.4, 17.5*	20.3
$(I_c)_p$ (μA)	948	948	597
β	565	565	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} cm^{-3}$	$XJ_1 = 0.05 \mu m$	$BV_{ceo} = 9.6 V$
$BD = 4.0 \times 10^{18} cm^{-3}$	$XJ_2 = 0.1398 \mu m$	$BV_{cbo} = 21.0 V$
$EPI = 1.0 \times 10^{16} cm^{-3}$	$W_b = 0.07 \mu m$	Reach through = 4.19 V
	$TEPI = 0.59 \mu m$	

Table 4.7 Light emitter doping

$I_c = 200 \mu A$	$XJ_1 = 0.05 \mu m$	$BV_{\infty} = 5.7 V$
$V_{bc} = 1.0 V$	$XJ_2 = 0.1381 \mu m$	$BV_{cbo} = 19.0 V$
$ED = 3.0 \times 10^{20} cm^{-3}$	$W_b = 0.045 \mu m$	Reach through = 4.79 V
$BD = 1.0 \times 10^{18} cm^{-3}$	$TEPI = 0.59 \mu m$	Ge = 10%
$EPI = 2.8 \times 10^{16} cm^{-3}$		

Parameters	SA1X SiGe	NSA1X SiGe	SA1X+ Si
F_t (GHz)	18.5	18.4	14.1
F_m (GHz)	18.2	11.6, 13.6*	17.4
R_b (Ω)	190	190	620
R_{bex} (Ω)	260	910, 620*	150
C_{JE} (fF)	20	20	17.93
C_{JC} (fF)	4.56	4.56	2.9
C_{JSC} (fF)	12.1	12.1	10.8
$(F_t)_p$ (GHz)	25.2	25.3	17
$(F_m)_p$ (GHz)	23	14.5, 17.0*	20.3
$(I_c)_p$ (μA)	941	941	597
β	809	809	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} cm^{-3}$	$XJ_1 = 0.05 \mu m$	$BV_{\infty} = 9.6 V$
$BD = 4.0 \times 10^{18} cm^{-3}$	$XJ_2 = 0.1398 \mu m$	$BV_{cbo} = 21.0 V$
$EPI = 1.0 \times 10^{16} cm^{-3}$	$W_b = 0.07 \mu m$	Reach through = 4.19 V
	$TEPI = 0.59 \mu m$	

Table 4.8 Light emitter doping, light EPI doping

$I_c = 200 \mu A$	$XJ_1 = 0.05 \mu m$	$BV_{ceo} = 7.5 V$
$V_{bc} = 1.0 V$	$XJ_2 = 0.1359 \mu m$	$BV_{cbo} = 19.2 V$
$ED = 3.0 \times 10^{20} cm^{-3}$	$W_b = 0.045 \mu m$	Reach through = 1.92 V
$BD = 1.0 \times 10^{18} cm^{-3}$	$TEPI = 0.59 \mu m$	Ge = 10%
$EPI = 1.0 \times 10^{16} cm^{-3}$		

Parameters	SA1X SiGe	NSA1X SiGe	SA1X+ Si
F_t (GHz)	18	18	14.1
F_m (GHz)	23	14.5, 17.1*	17.4
R_b (Ω)	190	190	620
R_{bex} (Ω)	250	1300, 600*	150
C_{JE} (fF)	20	20	17.93
C_{JC} (fF)	2.99	2.99	2.9
C_{JSC} (fF)	12.3	12.3	10.8
$(F_t)_p$ (GHz)	21.9	21.9	17
$(F_m)_p$ (GHz)	25.9	16.1, 19.0*	20.3
$(I_c)_p$ (μA)	331	331	597
β	841	841	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} cm^{-3}$	$XJ_1 = 0.05 \mu m$	$BV_{ceo} = 9.6 V$
$BD = 4.0 \times 10^{18} cm^{-3}$	$XJ_2 = 0.1398 \mu m$	$BV_{cbo} = 21.0 V$
$EPI = 1.0 \times 10^{16} cm^{-3}$	$W_b = 0.07 \mu m$	Reach through = 4.19 V
	$TEPI = 0.59 \mu m$	

(2) There is an appreciable increase in the F_m of the NSA devices due to the change in its main components, C_{JC} and the total base resistance ($R_b + R_{bex}$). The decrease in the base-collector capacitance C_{JC} can be attributed to the increase in the depletion width at the collector junction X_{bc} . The increase in the total base resistance ($R_b + R_{bex}$) of the NSA devices is expected. It starts from the edge of the emitter poly and extends above the field oxide up to the base contact.

(3) The breakdown voltage BV_{ceo} is significantly increased from 5.7 V to 7.5 V, whereas, the reach through voltage is decreased from 4.79 V to 1.92 V as a result of the decreasing in the doping level of the EPI layer. Wolf [41] indicated that when the epitaxial layer thickness is too small, under increasing reverse-bias voltage the depletion region of the C-B junction will entirely penetrate the epitaxial layer before the predicted BV_{cbo} voltage has been reached. If the voltage is increased beyond this point, the depletion region must enter the heavily doped buried layer, resulting in BV_{cbo} values smaller than those predicted merely by the doping in the EPI layer.

(4) The percentage of the increase and decrease in the F_i , F_m , β and other transistor parameters fluctuating from one transistor to another depends mainly on the size of the transistor. The most sensitive areas which affect the performance of the device are the length and width of the field poly and the length and width of the emitter stripe mask. Table B.2 (Appendix B) indicates that when there dimensions become smaller, the performance of the device improves, because the parasitic capacitances and resistances

are reduced. Furthermore, the small geometry devices may run at higher current density for a given collector current.

Table 4.9 shows another doping profile, where the doping levels in the collector and the base are similar to that presented in Table 4.8 except the doping level in the emitter was increased from $3.0 \times 10^{20}/\text{cm}^3$ to $6.0 \times 10^{20}/\text{cm}^3$. The simulation results indicated that the performance of the devices shown in Table 4.9 is slightly improved, but is insignificant. It should be noted that the values of the current gain β depicted in all tables are high. According to Roulston [42], the author of the BIPOLE3 program, the values of β should be smaller after fabrication than the numerically calculated by BIPOLE3, partly because the control parameters of the SiGe in the BIPOLE3 program were not fully established. Also to be noted that, the current gain is strongly dependent on the minority carrier lifetime in the emitter and surface recombination velocity [15], which is difficult to predict, because it is processing dependent.

There are two extra parameters left, which are not included in all tables, the Early voltage (V_A) and the pinch sheet base resistance. Early voltages calculated from BIPOLE3 simulations were found to be between 87-140 V. The fluctuations in the values of the V_A are due to the actual variation in the doping levels used during the simulations. For example, the Early voltage of all the devices shown in Tables 4.7 and B.2 was found to be 87 V, whereas for the devices in Tables 4.8 and B.3, $V_A = 140$ V. Early voltage for Si-base devices was found to be equal to 28 V. The pinch sheet resistance for all the

Table 4.9 Heavy emitter doping, light EPI doping

$I_c = 200 \mu A$	$XJ_1 = 0.05 \mu m$	$BV_{ceo} = 7.7 V$
$V_{bc} = 1.0 V$	$XJ_2 = 0.1362 \mu m$	$BV_{cbo} = 19.2 V$
$ED = 6.0 \times 10^{20} cm^{-3}$	$W_b = 0.045 \mu m$	Reach through = 1.92 V
$BD = 1.0 \times 10^{18} cm^{-3}$	$TEPI = 0.59 \mu m$	Ge = 10%
$EPI = 1.0 \times 10^{16} cm^{-3}$		

Parameters	SA1X SiGe	NSA1X SiGe	SA1X+ Si
F_t (GHz)	18.7	18.7	14.1
F_m (GHz)	23.4	14.8, 17.7*	17.4
R_b (Ω)	190	190	620
R_{bex} (Ω)	250	890, 600*	150
C_{JE} (fF)	19.57	19.57	17.93
C_{JC} (fF)	2.79	2.79	2.9
C_{JSC} (fF)	12.3	12.3	10.8
$(F_t)_p$ (GHz)	23.6	24	17
$(F_m)_p$ (GHz)	26.9	16.8, 19.8*	20.3
$(I_c)_p$ (μA)	466	466	597
β	619	619	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} cm^{-3}$	$XJ_1 = 0.05 \mu m$	$BV_{ceo} = 9.6 V$
$BD = 4.0 \times 10^{18} cm^{-3}$	$XJ_2 = 0.1398 \mu m$	$BV_{cbo} = 21.0 V$
$EPI = 1.0 \times 10^{16} cm^{-3}$	$W_b = 0.07 \mu m$	Reach through = 4.19 V
	$TEPI = 0.59 \mu m$	

present SiGe-base devices were found to be between 11.8 k Ω /sq to 14.1 k Ω /sq, whereas for the Si-base device, the pinch sheet resistance was found to be 19.9 k Ω /sq.

The general conclusions for this section are given below:

1. The performance of most of the devices that are presented in Table 4.6 is slightly better than all of the devices presented in Tables 4.7 to 4.9, with the results of the devices shown in Table 4.8 as a second best performance. This can be explained in terms of the appropriate selection of the doping levels across the transistor regions. Evidence for that can be seen from the performance of the devices shown in Table 4.6 as compared to the same devices shown in Tables 4.6 and 4.7. The emitter implantation time for the devices shown in Table 4.6 is twice than that shown in Table 4.7, and may slow down the production. Modern implantation equipment using high current density may resolve this type of delay.
2. Generally, all the simulation results presented in Tables 4.6-4.9 and part1 of Table B.1 of Appendix B indicate that without exception, the performances of the SA devices are better than the corresponding NSA devices simulated under same conditions. The transition frequency F_t of most of the SA devices is slightly lower than that of the NSA devices, because C_{JC} in the SA devices is slightly larger than that of the NSA devices. This is due to the fact that the geometries of the SA devices are slightly different from the corresponding geometries of the NSA devices. Furthermore, the oscillation frequencies F_m of the NSA devices are smaller than that of all the SA devices, because the total base resistance in the NSA devices is much larger than that of the SA devices, though the C_{JC}

in the NSA devices is smaller than that of the corresponding SA devices. Generally, the magnitude of the total base resistance ($R_b + R_{bex}$) times C_{JC} for SA devices is smaller than for the corresponding NSA devices.

3. The difference between the vertical and lateral analysis was found 5-9%, which is insignificant. This means that the crowding effect [15,40] is negligible. This point was discussed earlier.

4.6 Influence of Device Scaling on the Electrical Characteristics

Earlier it was noted that the main challenge faced in the design of the HBTs reported here was the need to deliver high performance at relatively low current density. In order to investigate the effect of mask dimensions and consequently the size of a transistor on its performance, several SA and NSA HBT devices with different dimensions have been designed and subjected to simulation. These devices are listed in Tables B.1 to B.4 of Appendix B. Additional detail on the dimensions of these devices is given in Table 4.1. In these devices the emitter mask length (EML) was reduced from the normal length (3 μm) to 2.5 μm , 1.5 μm , and 1.25 μm . The emitter width poly (EWP) was also reduced from 1.5 μm to 1.25 μm . As a result of these changes the remaining device geometry for each device was reduced accordingly.

The devices listed in Tables B.1 to B.4 were simulated under various doping levels.

Among the listed devices is the self-aligned device SA05X , which has the best performance compared with the remaining devices. This device has $F_t = 26.2$ GHz and

$F_m = 37.2$ GHz as shown in part1 of Table B.1. The superiority of this device can be attributed to the significant decrease in the intrinsic base-resistance R_b and the parasitic capacitances, especially at the emitter-base junction C_{JE} .

The feature of the devices presented in the Tables B.1 to B.4, is that the capacitance of all the SA and NSA devices at the emitter-base junction C_{JE} is much smaller than the C_{JE} of all the devices presented in Tables 4.6–4.9 simulated under similar doping levels. The reduction in the size of these devices reduces the parasitic resistances and capacitances, and leads to improvement in their performance. It should be noted that Table B.3 indicates that F_t and F_m for $I_c = 200$ μ A for some devices are below 10 GHz. However, the peak values of the F_t and F_m , which are located at $I_c = 97$ to 139 μ A for such devices were found to be above 20 GHz. This would suggest that this type of devices are only useful for very low power applications, specifically for $I_c < 200$ μ A.

The simulation results presented in Tables B.1 and B.4 indicate that as the emitter doping level increases the performance of the device is enhanced. These results are consistent with the results presented in section 3.7 and illustrated in Tables 4.6–4.9. Generally, the influence of doping levels on the performance of all the devices presented in Tables B.1 to B.4 are similar to those devices presented in Tables 4.6 to 4.9.

The general conclusion to this section is given as follows: The length and width of the field poly, length and width of the emitter stripe were found to play important role on the

performance of the device. As these parameters become smaller, performance of the device is significantly improved. This is due to the reduction in the R_b , C_{JE} and C_{JC} combined with the requirement for operation at $I_c = 200 \mu A$.

4.7 Simulation Results of an Si-Base Device

It was found useful to compare the simulation results of the SiGe-base devices with the corresponding Si-base, to justify the contention that the SiGe-base devices have superior performance relative to the Si-base devices. One of the npn 1X self-aligned devices was chosen for this task. The result of the simulations are presented in the third column of the Tables 4.6-4.9 and last row of the Tables B.1-B.4. Without exception, the performance of all the SA and NSA SiGe-base devices are much better than the Si-base devices.

Evidence for that can be seen, for example, in the total base-resistance of the SiGe-base devices being much smaller than the corresponding resistance in the Si-base devices. The current gain of the SiGe-base devices is much higher than for the Si-base devices, and this result is consistent with eq.2.1. The Early voltage was found to be equal to 28 V for the Si-base, compared to 87 to 140 V for SiGe-base devices.

In order to explain the difference between the performance of the Si-base and the SiGe-base devices, the contribution of each delay in the F_t of the Si-base devices should be determined. Vertical and lateral simulations were carried out, and the results are illustrated in Tables 4.10 and 4.11. For the purpose of the comparison, the simulation results of the SiGe-base of the SA1X device are repeated in Tables 4.10 and 4.11. During

the simulation, it was found that in order to avoid the base punch through, the width of the quasi-neutral base should be made equal to 700 Å, and the doping level in the base should be equal to $4 \times 10^{18} / \text{cm}^3$. This doping is four times higher than for SiGe-base devices. It should be noted that the Si devices have a diffused base and thus the doping is non-uniform across the base and has a Gaussian profile, whereas in the SiGe devices the doping is graded uniformly across the base. Because of these differences, comparison between the simulation results of the Si-base and SiGe-base devices will be qualitative. The simulation results presented in Tables 4.10 and 4.11 can be summarized as follows:

- (1) There is full agreement for F_t obtained from vertical and lateral analysis. This means that the lateral effects are insignificant.
- (2) The delay time τ_e in the neutral emitter is much larger in the Si-base than that of SiGe-base. According to eq.2.7, the τ_e is proportional with the base width W_b and the width of the Si-base is much larger than the width of the SiGe-base devices as mentioned earlier.
- (3) The emitter-base junction charging time τ_{eb} is the dominant delay. It is about 8% larger in Si-base than that of SiGe-base.
- (4) The delay due to a free carrier charge in emitter-base space charge τ_{qbe} is much larger in Si-base than in SiGe-base devices. According to Roulston [15], the delay τ_{qbe} is proportional with $1/I_c$. Since β is proportional to I_c and β in Si-base devices is much smaller than β in SiGe-base devices (Tables B1.-B.4), thus the delay τ_{qbe} should be higher in Si-base compared to SiGe-base devices.

(5) The delay time in the base τ_b of Si-base devices is less than half of the delay in SiGe-base devices. This would suggest that the τ_b is not an important delay for Si-base devices. This result is inconsistent with the essence of eq.2.2, where the ratio of $\tau_{b, SiGe} / \tau_{b, Si}$ is less than unity. The inconsistency can be explained in terms of the doping profile in the base of both the Si and SiGe devices. As mentioned above, the doping profile in the SiGe devices is uniform across the emitter, base and collector, and does not obey a typical Gaussian profile. The base in the Si device has a non-uniform profile and follows the Gaussian profile. This would suggest that when a reverse bias is applied across the collector, the depletion region at the base-collector junction will enter the quasi-neutral base region. In consequence, the thickness of the base will reduce and thus allow the carriers to reach the collector much faster and the τ_b will become shorter. The drawback of this feature is that the Si devices will suffer from low punch through, and this significantly reduces the Early voltage. This conclusion is consistent with the present results, where the Early voltage was found to be much larger in SiGe-base devices than that in the Si-base devices.

(6) The transit time due to collector space charge layer τ_{sc1} is also significant, and is about 8% larger in Si-base than in SiGe-base devices.

(7) The collector charging time τ_{cb} is insignificant (≈ 0.01 ps), thus it can be neglected.

(8) Table 4.11 shows that the delay τ_{AVO} is about 29% higher in Si-base than in SiGe-base devices, because the delays τ_e and τ_{qbc} which are part of the delay τ_{AVO} are much larger in Si-base than in SiGe-base devices (Table 4.10).

Table 4.10 Delay times for Si-base devices obtained from vertical analysis

Doping Levels /cm ³	Device Name	τ_e ps	τ_{cb} ps	τ_{qbc} ps	τ_b ps	τ_{rd} ps	τ_{db} ps	F_t GHz
ED = 3.0×10^{20} BD = 4.0×10^{18} EPI = 1.0×10^{16}	SA1X: Si	1.821	4.46	1.54	0.91	2.5	0.01	14.16
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 1.0×10^{16}	SA1X: SiGe	0.193	4.07	0.03	2.02	2.3	0.02	18.43

Table 4.11 Delay times for Si-base devices obtained from lateral analysis

Doping Levels /cm ³	Device Name	τ_{AVO} ps	τ_{AVL} ps	τ_{RCE} ps	τ_{REE} ps	F_t GHz	C_{diff} fF	g_m x10 ⁻² S
ED = 3.0×10^{20} BD = 4.0×10^{18} EPI = 1.0×10^{16}	SA1X: Si	6.35	0.24	0.03	4.61	14.16	44.89	0.693
ED = 3.0×10^{20} BD = 1.0×10^{18} EPI = 1.0×10^{16}	SA1X: SiGe	4.5	0.29	0.04	4.22	17.6	31.92	0.709

4.8 General Transistor Characteristics

Figure 4.4 shows a typical simulation of the uniform vertical profile for the SA1X and NSA1X of the SiGe-HBTs. This structure has an emitter doping level of $3 \times 10^{20} / \text{cm}^3$ at the poly/mono interface, a base doping level of $1 \times 10^{18} / \text{cm}^3$ and an epi-layer that has a doping level of $2.8 \times 10^{16} / \text{cm}^3$. The Ge content (not shown in Fig.4.4) is ramped from 0 to 10% across the base. The spacers at the emitter-base and base-collector regions are also not shown in Fig.4.4, because BIPOLE3 program can't provide simulations for the spacers.

A typical Gummel simulation plot for the SA1X and NSA1X of the SiGe-HBTs at $V_{bc} = 1.0 \text{ V}$ is shown in Fig.4.5. The emitter area of the device shown in Fig.4.5 is $1.5 \times 3.0 \mu\text{m}^2$.

Figure 4.6 shows a typical plot of the I_c versus the collector-emitter bias (V_{ce}) for SA1X and NSA1X SiGe HBTs. The simulations were conducted at different values of I_b from $0.2 \mu\text{A}$ to $0.5 \mu\text{A}$. The simulation results indicated that the emitter-to-collector breakdown voltage BV_{ceo} at $V_{bc} = 1.0 \text{ V}$ is 5.7 V . The base-to-collector breakdown voltage is 19.0 V . The Early voltage for the device shown in Fig.4.6 is 87 V .

High frequency performance variations with device geometry and doping for SiGe-HBTs have been discussed in sections 4.4 to 4.6. Typical transition frequency F_t and maximum frequency of oscillation F_m versus collector current I_c , at $V_{bc} = 1.0 \text{ V}$, are shown in Fig.4.7.

Figure 4.7(a) demonstrates the high performance of the SA device, where F_t and F_m have peak values equal to 26.5 GHz and 23.7 GHz respectively at $I_c = 948 \mu\text{A}$. In the NSA device shown in Fig.4.7(b), the F_t and F_m have peak values equal to 25.3 GHz and 14.5 GHz respectively at $I_c = 941 \mu\text{A}$. The significant decreases in F_m is due to the increase in the extrinsic-base resistance of the devices.

Figure 4.4 Typical simulated impurity profile for SA1X and NS1X of SiGe HBT

Figure 4.5 Typical simulated Gummel plots of base and collector currents for SA1X and NS1X of SiGe HBT

Figure 4.6 Typical simulated I_c - V_{ce} characteristics for SA1X and NS1X of SiGe HBT

Figure 4.7 Simulated F_t and F_m as a function of I_c for SiGe HBT

CHAPTER 5

DEVICE FABRICATION AND CHARACTERIZATION

5.1 Introduction

This chapter describes the processing experiments carried out to realize the SiGe HBTs designed in chapter 4. The strategy followed in developing the new SiGe HBT technology was to keep the process architecture as close to existing Gennum BJT technologies as possible, making modifications only when necessary. The main modifications required were the development of three new process modules: field poly formation, SiGe base deposition and local interconnect formation. In addition, some minor modifications to layer thickness and annealing conditions in the baseline BJT technology were found desirable to optimize device performance. The majority of the processing experiments were carried out in Gennum's fabrication line in Burlington, Ontario. Some additional experiments were completed in Carleton University's Microelectronics fabrication facility. In order to protect information proprietary to Gennum, no details on the baseline BJT process flow on which the SiGe base was grafted will be given here.

5.2 Test Wafer Layout

Before processing could begin, it was necessary to generate a complete test wafer layout, including passive devices, transistors designed for DC measurements and transistors in special pad frames intended for high frequency AC s-parameter measurements. For the

passive components, it is required to design a test structure in order to be able to measure the sheet resistance, the capacitance, and the linewidth for most of the device layers. The main objective is to ensure that the values of the device parameters are within the required specifications. Four point van der Pauw structures [43] were used to measure the sheet resistance, whereas four point Kelvin structures were used to measure the linewidth of each layer.

The layout was accomplished using the L-Edit program. It was first necessary to prepare a new L-Edit technology file describing the SiGe HBT process, including design rules. While laying out actual transistors, dimensions were minimized in accordance with these design rules to minimize parasitic capacitance.

5.3 Short Loop Experiments

The purpose of the short loop (preliminary) experiments is to optimize the conditions of the depositions, etching, diffusions, oxidations and lithography. For example, in the process integration, the thickness and the concentration of the epitaxial collector are important parameters which determine the breakdown voltages of the bipolar transistor. These parameters should be selected in such a way to meet as closely as possible the device specifications.

In order to integrate a new process module in an existing technology, it is necessary to perform short loop and mechanical experiments in order to be able to identify and

optimise the process under investigation until it meets the specification. The short loop and the mechanical experiments can be regarded as preliminary experiments. The meaning of the short loop experiments can be elucidated by the following example. If a specific thickness of an oxide layer is required to be grown on a certain layer, then the annealing temperature and deposition time should be optimised until it meets the specification. A mechanical experiment is one intended to establish device structures (such as layer thickness, etch parameters, etc.) without providing electrical test results. Thus, from the short loop and mechanical experiments, we can obtain optimised conditions before we start full-wafer processing.

In the present work, we have performed several short loops and mechanical experiments for the SiGe HBT technology. Three new modules have been developed, and the feature of each process for each new layer are discussed below. The remaining process modules of the new HBT technology are common to other Gennum technologies. The new modules are as follows:

1. Field poly module
2. SiGe deposition module
3. Local interconnect module

Each module consists of two or more process steps. Process details for each module are given below.

5.4 Field Poly Module

The UHV/CVD method for SiGe deposition requires an unusual precleaning procedure in which the substrate is dipped in dilute hydrofluoric acid and then removed without any rinse step. This procedure passivates dangling bonds on the Si surface with hydrogen. Since bare silicon is strongly hydrophobic, a wafer slowly pulled from hydrofluoric acid solution will emerge in a completely dry condition. This cleaning procedure can also be applied to wafers with patterned oxides on their surfaces (as would typically be the case in a full process) provided less than 15% of the surface area is oxide. To achieve less than 15% oxide surface coverage in a commercial bipolar process, it is necessary to coat most of the wafer surface with a sacrificial polysilicon layer that will be referred to as “field poly”.

Apart from being essential to the UHV/CVD precleaning process, the field poly offers advantages to the HBT process flow. For example, the structure of the deposited SiGe on the field poly is different from the structure of the SiGe deposited onto an oxide layer, because the former substrate has a polycrystalline structure whereas the latter substrate has an amorphous structure. Since the nature of the substrate has an influence on the growth of the deposited layer, it is expected that the sheet resistance of the SiGe deposited on the field poly, to a certain extent, is slightly less than that of the SiGe deposited onto the field oxide. Furthermore, the thermal expansion coefficient of the field poly should be closer to the thermal expansion coefficient of the SiGe than the thermal expansion of the field oxide. This means the thermal stress arising from the difference in

the thermal expansion of the film (field poly) and that of the substrate (field oxide) will be much less in the case of the field poly surface than in the field oxide surface. Thermal stress results from the difference in the thermal expansion coefficients of the film and the substrate when the film/substrate is cooled from a high temperature to room temperature. Consequently, the sheet resistance of the SiGe layer on the field poly will be less than that on the field oxide. This feature is especially important for the non-self aligned HBT, because in this architecture the resistance of the field poly is part of the total extrinsic-base resistance of the transistor base. This point was discussed in chapter 4. Most important, in the present technology, the width of the field poly defines the single crystal SiGe region of the self-aligned transistors.

The process module for field poly formation consists of the following eight steps:

1. Pre-field poly clean and field poly deposition
2. Field poly cap oxide
3. Field poly implant
4. Field poly anneal
5. Cap oxide etch
6. Field poly mask
7. Dry poly field etch
8. Field poly mask resist strip

Cross sectional views across a poly finger after each of the above steps are given in Fig.5.1(a-e). A brief description of the above process flow follows. Wafers are first RCA

cleaned and then loaded into a low pressure chemical vapour deposition (LPCVD) tube. Fig.5.1a shows the field poly layer deposited onto the field oxide. The deposition conditions of the polysilicon were obtained from the short loop experiments. After deposition, the wafers were inspected in a high intensity light to ensure that they appear uniformly free from any defects such as streaks or haze. Following this step, the cap oxide (500 Å) was thermally grown on the field poly by means of wet oxidation using predetermined conditions (Fig.5.1b). Next, ion implantation was used to implant boron (B^+) into the field poly through the cap oxide, using an energy of 30 keV and a dose of 3×10^{15} ions/cm² (Fig.5.1c). After ion implantation, the wafers were pre-diffusion cleaned and furnace annealed at 950 °C for 40 minutes (Fig.5.1d). Following the drive-in, the cap oxide was removed using a plasma etch. The field poly inside the active region was etched using a plasma dry etch, based on the results obtained from the short loop experiments (Fig.5.1e). Finally, a poly mask resist strip (using an oxygen plasma ashing process) was performed to conclude the field poly process.

The cap oxide serves two functions. First, it blocks the implanted B-impurity from evaporation from the field poly layer during the field poly anneal process, and thus reduces its sheet resistance. Secondly, it reduces the defects on the surface of the field poly resulting from the boron implantation damage, and allows a higher implant energy to be used. With increasing implant energy, the dopant penetrates deeper and becomes less sensitive to several sources of process variation. It should be noted that the growth of the capping oxide onto the field poly consumes part of it. According to Wolf [44], the

(a)After field poly deposition

(b)After field poly cap oxide

(c)After field poly implant

Figure 5.1(a-c) Schematic cross-section of the field poly process sequence of the SiGe

HBT

(d) After field poly anneal & cap oxide etchback

(e) After field poly mask/etch

Figure 5.1(d-e) Schematic cross-section of the field poly process sequence of the SiGe

HBT

oxide consumes silicon to a depth of about $0.44 t_{ox}$, where t_{ox} is the grown oxide thickness. For $t_{ox} = 500 \text{ \AA}$, the silicon consumption was found to be equal to about $220 \text{ \AA}$. The consumption of silicon resulting from growth of the capping oxide was taken into account during the field poly deposition.

A large set of samples of field poly layers has been prepared under various conditions. A summary of the results for these samples together with their conditions is given in Table 5.1. This table shows effects of various parameters on the sheet resistance of the field poly. Including are thickness of the field poly, implantation types and their doses and energies, annealing time and temperature and finally capping of the field poly. To assess quantitatively the effect of these parameters on the characteristics of the field poly, a comprehensive microstructural analysis is required, which is beyond the scope of this thesis. The main feature of the results presented in Table 5.1 can be explained as follows.

Table 5.1 indicates that the sheet resistance R_s of the field poly is inversely proportional to the film thickness. Also to be noted is that, the optical microscope examinations indicated that when the annealing temperature is raised to 1000°C , most of the wafers become hazy. This problem is more pronounced for uncapped samples annealed at 1000°C . However, this type of problem completely disappeared when the annealing temperature was reduced to 950°C . It is expected that the high temperature annealing process causes vertical grain growth of the field poly, which enhances the roughness of the interface, giving haze [45].

Table 5.1 Summary of the field poly results

Sample No.	Thickness Å	Implant	Energy keV	Dose cm ⁻²	Annealing Time min	Temp °C	Oxide cap	R _s Ω/sq.
1	1500	B ⁺	25	3E15	40	950	----	366
2	3000	B ⁺	25	3E15	40	950	----	177
3	1500	BF ₂	50	5E15	40	950	----	402
4	3000	BF ₂	50	5E15	40	950	----	182
5	2900	BF ₂	50	5E15	40	950	----	188
6	3300	B ⁺	25	3E15	40	950	----	158
7	3300	B ⁺	25	3E15	40	1000	----	341*
8	3020	B ⁺	30	3E15	40	950	✓	165
9	3020	B ⁺	30	3E15	40	1000	✓	128*
10	3020	BF ₂	100	5E15	40	1000	✓	113*
11	3100	BF ₂	100	5E15	20	950	✓	150
12	3100	B ⁺	30	3E15	40	950	✓	163
13	3100	B ⁺	30	3E15	20	950	✓	169

*: Means the wafer was hazy

Table 5.1 also indicates that when a BF_2 implant is used, the sheet resistance of the field poly becomes less than when B^+ is implanted. Unfortunately, the presence of the fluorine (from the BF_2 implant) may enhance boron penetration through the oxide layer into the underlying substrate. The spreading resistance analysis indicated that when BF_2 was used the boron penetrates to the underlying substrate through the oxide layer. Elemental boron is therefore considered to be superior to BF_2 as an implant species to dope the field poly.

According to the above analysis, the optimized conditions for preparing the field poly are to cap it by an oxide, then implant it with B^+ with a dose of 3×10^{15} ions/cm² using an energy of 30 keV, and finally furnace anneal it at 950 °C for 40 minutes. It should be noted that, during the emitter processing, it was found that thickness of the field poly should be less than 2000 Å in order to reduce the emitter stringers.

5.5 SiGe Base Deposition Module

The SiGe base module consists of two processes given below.

1. SiGe Deposition
2. Sacrificial low temperature oxide deposition (SAC_LTO)

The SiGe layer was prepared using a UHV/CVD technique in a “Sirius” hot wall reactor from Leybold AG. This system is very similar to that developed by Meyerson [20].

Before SiGe deposition, the thin oxide, which covered the opening of the field poly in the n-epi region, was etched away (Fig.5.2a) until the surface became hydrophobic. Next, the

wafers were loaded immediately in the UHV/CVD chamber for deposition of the SiGe layer (Fig.5.2b).

On the basis of simulations with a version of SUPREM3 calibrated for use with SiGe base regions, the group responsible for growth of the SiGe epitaxial base recommended the insertion of additional buffer and spacer layers on the emitter and collector sides of the base. The SiGe scheme starts by depositing a 200 Å silicon buffer layer onto the epitaxial collector layer at the base-collector region. This layer, which is an extension of the existing field poly layer for the non self-aligned structure, is useful because it decreases the sheet resistance of the field poly, and it covers any surface roughness that may arise from the previous process steps. Following the buffer layer, a 100 Å spacer layer of undoped SiGe (10%Ge) is deposited onto the previous silicon layer. The main purpose of this layer is to accommodate the out-diffusion of boron from reaching the collector during annealing of the field poly or emitter poly. The surface of this layer is very clean, free from any oxide or contamination, and can be regarded as an excellent nucleation surface for the next doped and graded SiGe layer. Then, the main layer of 700 Å of graded SiGe (0%-10%) doped with boron of concentration of 1×10^{18} atoms/cm³ is deposited. Finally, a 400 Å deposition of undoped silicon spacer above the previous layer at the base-emitter junction was performed to finish the structure. The feature of the spacer is to prevent arsenic out-diffusion from the emitter from reaching the base. Additional benefit is to ensure that the SiGe is not oxidized during the emitter poly annealing. It has been found that the presence of a small amount of Ge during silicon

oxidation process enhanced the oxidation rates [46]. In addition the Ge was found to be segregated to the growth interface during oxidation. Pile-up of Ge at the interface is likely to lead to a leakage under electrical stress. In order to control these effects, the presence of the undoped spacers whether Si or SiGe at the base-emitter or base-collector junctions serves as a sacrificial layer for preventing oxide growth. The spacer also helps reduce the emitter-base junction capacitance, because the emitter poly, which has the highest doping concentration, is not deposited directly on the doped and graded SiGe layer. Furthermore, the spacers also allow for dopant diffusion during the emitter anneal. It should be noted that the Ge profile was ramped from 0% at the emitter-base junction to 10% at the base-collector junction. The Ge and boron profiles were chosen to optimise the required F_t and other transistor parameters.

The final step in the SiGe process is to deposit a sacrificial low temperature oxide (SAC_LTO) layer on the wafers to protect the SiGe base from photoresist during the local interconnect patterning (Fig.5.2c). The thickness of the SAC_LTO was between 1300 Å to 1700 Å. The SAC_LTO will be etched away in the later stages.

In the self-aligned architecture, the SiGe layer under the emitter poly has a single crystal structure and determines the intrinsic base-resistance, whereas the SiGe in the silicided region determines the extrinsic base-resistance. In the non self-aligned architecture, the extrinsic resistance has three components. The first is the resistance of the SiGe from the edge of the emitter poly to the edge of the field poly inside the base active region (the

(a) After cap oxide etch

(b) After SiGe deposition

(c) After SAC_LTO deposition

Figure 5.2(a-c) Schematic cross-section of the SiGe process sequence of the SiGe HBT

SiGe in this region has a single crystal structure). The second is the resistance of the SiGe/field poly stack from the edge of the field poly inside the active region to the silicided region outside the field poly region (the SiGe in this region has a polycrystalline structure). The third component is due to the resistance of the silicided region connecting to the contact, and the SiGe layer in this region has a polycrystalline structure.

5.6 Local Interconnect Module

As mentioned above, SiGe grown over field poly has a polycrystalline structure, whereas the growth of SiGe in the epitaxial Si-base has a single crystal structure. We refer to the field region where the SiGe has a polycrystalline structure as a local interconnect. Both structures are connected with each other as if there is one continuous conducting layer. The local interconnect is made up of SiGe overlying field poly. The local interconnect is silicided later in the process flow for both self-aligned and non-self aligned architectures. For both self-aligned and non-self aligned architectures, the base contacts were placed on local interconnect regions outside the base active region. The local interconnect process module consists of five steps:

1. Local interconnect mask
2. Sacrificial LTO etch
3. SiGe and field poly etch
4. Local interconnect resist strip
5. Sacrificial LTO etchback

(a) After local-interconnect mask/etch & SiGe/field poly dry etch

(b) After photoresist strip & SAC_LTO etchback

Figure 5.3(a-b) Schematic cross-section of the LOC_INT process sequence of the SiGe HBT

The process flow starts by masking the wafer using the local interconnect mask (Fig.5.3a). Next, part of the sacrificial LTO on the field poly is removed using a BOE (buffer oxide etch) wet etch, followed by removal of the underlying SiGe/field poly stack using a plasma dry etch (Fig.5.3a). Fig.5.3a shows a lateral etch taking place resulting from the isotropic nature of the wet etch. The adverse effects of the lateral etch at this stage is not important because physically it does not affect other layers on the wafer. The next step is to strip the photoresist using the oxygen plasma etch. Finally, the sacrificial LTO is removed using a wet etch method (Fig.5.3b) and this finishes the local interconnect process flow module.

5.7 Rapid Thermal Emitter Annealing

The technique of rapid thermal annealing (RTA) is becoming widely applied in the semiconductor industry and serves, among other things, to minimize the diffusion between a film and the substrate or between the layers of a device. This technique can activate the dopant in a few seconds with minimal redistribution, compared to minutes in a conventional furnace annealing, which substantially broadens the profile. RTA is therefore desirable as a means of activating the emitter implant in SiGe HBTs.

To calibrate the existing rapid thermal annealing AG Associates Heatpulse 210 System [47], the following procedure was adopted. Boron was implanted into n-type <100> wafers using an energy of 35 keV and dose of 5×10^{15} ions/cm². These implant conditions were chosen to produce a layer in which sheet resistance is a strong function of anneal

temperature for temperatures close to 1025 °C. Immediately prior to introduction into the Heatpulse chamber, wafers were given an RCA type clean ending with a dip in 10:1 H₂O:HF to remove any interfacial oxide. Several tests were conducted on a test wafer to establish the annealing conditions. Argon gas was used during the annealing cycles. Finally, the real wafers were introduced into the RTA chamber and annealed individually between 945 °C and 1070 °C for 10 seconds.

The sheet resistance of each wafer was measured using the four-point van der Pauw method. The measured sheet resistances were found to be between 28 and 81 Ω/sq, and these values correspond to a temperature range between 1070 °C and 945 °C at the wafer centre. The results of the calibration are given in Table 5.2.

In the present work, the emitter poly anneal was the only significant thermal step following SiGe base deposition. It was carried out using furnace annealing at 850 °C for 20 minutes, followed by RTA in argon atmosphere 970 °C-1000 °C for 20 seconds. The wafers were RCA cleaned and dipped in HF acid solution before RTA as described above. For some additional wafers the emitter was exclusively furnace annealed at 900 °C for 20 minutes.

Table 5.2 Variation of sheet resistance with temperature

T_p (°C)	T_w (°C)	ρ (Ω/sq)
1025	1070	28.2
980	1050	35.6
950	1025	43.6
940	1006	47.1
930	990	53.0
920	985	54.9
900	980	59.3
880	957	71.2
870	945	81.2

T_p : Heatpulse preset temperature

T_w : Wafer temperature

ρ : Wafer sheet resistance

5.8 Device Characteristics and Discussion

This section describes the DC and AC measurement results provided by Gennum for 1X self-aligned and non-self aligned SiGe HBTs. The doping levels in the emitter, base and epi-layer are $3 \times 10^{20} / \text{cm}^3$, $1 \times 10^{18} / \text{cm}^3$ and $1 \times 10^{16} / \text{cm}^3$ respectively. The Ge concentration is ramped from 0 to 10% across the base. More details on the SiGe deposition conditions were given in section 5.5.

5.8.1 DC Characteristics

Gummel plots of collector and base currents are shown in Fig.5.4 for self-aligned (Fig.5.4(a)) and for non-self aligned (Fig.5.4(b)) SiGe HBTs. Figure 5.4 indicates good constancy of current gain down to a very low current with the non-self aligned HBT being the best. The I-V behaviour demonstrates the quality of the epitaxial material used and highlighting the low current capabilities of the SiGe HBTs process. This means that the metastable strained SiGe layers can withstand some high-temperature processing without deleterious electrical effects.

Variation of collector current with collector-emitter voltage for different base currents for SiGe HBT (SA and NSA) is shown in Fig.5.5. This figure shows the emitter-collector breakdown occurs near $BV_{ceo} = 7 \text{ V}$, which considerably exceeds design targets.

From the experiment, it was observed that the current gain β for the NSA1X device is equal to 385 at a collector current $I_c = 200 \mu\text{A}$. The high experimental value of β was about half of the predicted value by the BIPOLE3 program. The experimental value of β as

Figure 5.4 Experimental collector and base currents versus base-emitter bias for the SiGe HBT

Figure 5.5 Experimental collector current versus collector-emitter bias for different base currents for the SiGe HBT

compared to BIPOLE3 strengthens confidence in the accuracy of the simulator. However, the difference in the numerical value of β may be attributed to the following factors: (1) β is strongly process dependent and (2) the BIPOLE3 program was not calibrated. Taken into account that β is known to be strongly process dependent, and the BIPOLE3 program was not calibrated, the experimental value of β as compared to BIPOLE3 is encouraging.

5.8.2 AC Characteristics

The frequency response of the SiGe HBT was evaluated by measuring s-parameters using a high frequency network analyzer in conjunction with a Cascade RF wafer probe. After correcting for the effect of the pad capacitance, the transition frequency F_t for the SA1X and NSA1X SiGe HBTs measured at a collector current $I_c = 200 \mu\text{A}$ were 12.2 GHz and 13.1 GHz respectively. The doping levels for the fabricated devices were as follows: Emitter doping= $3 \times 10^{20}/\text{cm}^3$, base doping= $1 \times 10^{18}/\text{cm}^3$, and epitaxial collector doping = $1 \times 10^{16}/\text{cm}^3$. The transition frequency dependence on the collector current is shown in Fig.5.6 for both SA1X and NSA1X devices. The difference between the simulated and experimental values in the F_t is given below.

The simulation results of the SA1X and NS1X devices, which had the same structure as the fabricated devices except for the base width W_b (Table 4.8), indicated that the F_t is equal to 18 GHz at $I_c = 200 \mu\text{A}$ and $V_{ce} = 1 \text{ V}$. The disagreement between the simulated and experimental results may be due to differences in the base width. Specifically, the simulation results were based on the neutral base width being equal to 450 Å. There is as

Figure 5.6 Measured transition frequency versus collector current for the SiGe HBT

yet no experimental measurement (either secondary ion mass spectroscopy (SIMS) or spreading resistance profile (SRP)) providing information on the base width of the experimental devices. In the presence of boron out-diffusion from the base into the surrounding spacer layers, the thickness of the neutral base might be expected to be as much as 700 Å. When the simulations are repeated for a base width equal to 700 Å, the transition frequency F_t was found to be equal to 15.2 GHz at $I_c = 200 \mu\text{A}$ and $V_{ce} = 1 \text{ V}$ for both SA1X and NS1X devices.

The analysis of the total delays revealed that the delay across both the neutral emitter τ_e and neutral base τ_b is about twice in the case of $W_b = 700 \text{ Å}$ than $W_b = 450 \text{ Å}$. Taken into account that BIPOLE3 was not calibrated, and was not an ideal simulator (as was discussed in chapter 4 for our devices, especially for non-self aligned devices), the simulation and experimental results are in good agreement. On the other hand, if we assumed that out-diffusion of boron occurred, which will be trapped by the spacers, the thickness of the base will increase. This means that if the simulations are repeated again for a base width larger than 700 Å, the simulation results will become very close to the experimental results.

In the case of the self-aligned devices, in addition to the problem of the base width discussed above, there are additional factors which have significant impacts on the F_t and other transistor parameters. These include the consequences of overetching during patterning of the emitter poly and emitter poly oxide sidewall spacer, and the presence of

an emitter polysilicon “stringer” on the field poly sidewall. The goal of the plasma etch step which defines the emitter stripe is complete removal of the emitter polysilicon from regions not protected by photoresist without damage to the underlying SiGe base layer. It is difficult to devise a plasma etch chemistry which attacks polysilicon yet does not etch SiGe alloy with Ge content of a few percent. (Recall that the Ge content is graded across the base, rising from 0% at the emitter-base junction to 10% at the base-collector junction). The problem is compounded by inevitable cross-wafer and wafer-to-wafer variations of 10% or more in etch rate in commercial plasma etchers. A similar problem arises during formation of the oxide spacer on the emitter sidewall that is required for the silicidation process. The latter problem is improved somewhat by the relatively selectivity achievable between oxide and silicon in plasma etching.

Reducing the thickness of the SiGe base in the region between the emitter and the field poly may lead to a situation in which the base is completely consumed during silicidation. This will create a Schottky diode in the extrinsic base-collector junction. Although this is not entirely undesirable (the diode acts as a Schottky clamp, preventing saturation of the base), it may lead to considerable variation in base-collector junction characteristics from device to device.

The emitter stringer results from the anisotropic nature of the etch used to pattern the emitter. Etches of this kind leave deposits on the sides of vertical structures. Since the emitter polysilicon is patterned before the emitter anneal, the stringer is likely to be

undoped. Further, the stringer is likely to be completely converted to PtSi during silicidation, slightly reducing base resistance. Unfortunately, at the place where the emitter stripe intersects the field poly sidewall, the presence of the emitter poly stringer may interfere with the formation of the emitter oxide sidewall spacer. A short-circuit between the silicided emitter and the silicided extrinsic base may form at this location. Since the cross-sectional area of the short is likely to be very small, it may appear in the base-emitter diode characteristics as an increase in leakage current at low bias. It is possible that the non-ideal current apparent at low V_{be} in Fig.5.4a arises in this way.

The best approach to removing the emitter stringer is to reduce the height of the field poly, and to slope the edge of the field poly sidewall by using an isotropic etch to define this feature. In the self-aligned device the field poly need only be thick enough to allow for silicidation.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

This thesis has described the design, fabrication and electrical characterization of a novel self-aligned SiGe alloy base heterojunction bipolar transistor. The device structure uses an emitter polysilicon stripe to define the width of the intrinsic device, provides self-alignment of the silicided extrinsic base to the emitter, and allows contact to the base to be made over the field oxide, minimizing parasitic capacitance. Although self-aligned single-polysilicon homojunction BJTs have been reported previously, this is the first demonstration of a comparable HBT. The new structure can be fabricated with a considerably simpler process flow than previously reported SiGe base HBTs. A more conservative non-self aligned structure which simulation predicts should offer somewhat inferior electrical performance was also fabricated.

The work is also novel in the extension of SiGe HBT design to far lower current densities and device power dissipation levels than considered in previous work. Previous researchers typically concentrated on optimizing the speed of the SiGe HBT irrespective of power dissipation, typically optimizing operation for current densities approaching 10^6 A/cm². The present study required optimization of speed for operation at 200 μ A collector current in a device with a 1.5 μ m by 3.0 μ m intrinsic area, corresponding to a

collector current density of just $4.4 \times 10^3 \text{ A/cm}^2$.

The development of the SiGe base HBT described here was a large project involving contributions from several researchers. The main contributions made by the author are:

1. Design of a composition and doping profile capable of providing high-speed performance at low current density. This part of the work was carried out using the BIPOLE3 device simulator.
2. Development of the key process modules required to implement the SiGe base HBT by modifying an existing homojunction BJT process architecture at Gennum Corporation.
3. Layout of a suite of high-frequency and low-frequency test transistors for use in process development. All experimental results reported here were obtained from this set of devices.
4. Analysis of electrical test data on prototype devices provided by Gennum to assess the success of the profile design, to test the accuracy of the BIPOLE3 simulator in describing the new SiGe HBT structure, and to make recommendations for structural modifications to improve performance in future process runs.

The most important points emerging from the present investigation can be summarized as follows:

1. The preliminary calculations and simulation results identified the dominant components of the delay times that influenced the performance of the device. These include, the emitter-base junction charging time τ_{eb} , base-transit time τ_b and collector space charge layer transit time τ_{scf} . Both one-dimensional (vertical) and two-dimensional (lateral) analysis were conducted using the device simulator BIPOLE3 program. The results of the transition frequency obtained from the vertical analysis are 5-9% higher than the transition frequency obtained from the lateral analysis. This would suggest that the lateral effects, which were ignored in the vertical analysis are not important in the current devices.
2. The simplified regional analysis of transit times in SiGe HBTs presented in this thesis identified the main structural parameters that significantly influence transistor electrical performance. This approach greatly facilitates the physical interpretation of numerical simulation on the F_t and F_m and provide insight into how F_t can be modelled. The analysis indicated that in order to achieve high performance at low collector current ($I_c=200 \mu A$), the SiGe-base must be lightly doped, whereas the emitter should be heavily doped. The optimized doping levels in the emitter, base and collector are $3 \times 10^{20}/cm^3$, $1 \times 10^{18}/cm^3$ and $2.8 \times 10^{16}/cm^3$, respectively. It was found that by simulation the above doping levels combined with a base-width W_b equal to $450 \text{ \AA}$ give $F_t \geq 18.5 \text{ GHz}$. Decreasing W_b below

450 Å gives unacceptably low breakdown voltage BV_{∞} . Increasing W_b was found to significantly degrade the speed of the devices.

3. The simulation results revealed very significant performance advantages can be obtained by reducing device geometries. For example, at $I_c = 200 \mu A$, F_t and F_m for the device SA05X (shown in Table B.1 Part 1 of Appendix B) are 26.2 GHz and 37.2 GHz respectively. The analysis indicated that the parasitic capacitances C_{JE} and C_{JC} and the total base resistance are much smaller for small-geometry devices than the typical 1X devices.

4. The simulation results conclusively indicate that the SiGe-base devices have much better performance than the corresponding Si-base devices. Specifically, at a given collector current, the transition frequency F_t , maximum oscillation frequency F_m , current gain β , Early voltage V_A , and the breakdown voltages BV_{∞} and BV_{cbo} , are much higher for SiGe-base devices than the corresponding Si-base devices. This indicates that SiGe HBTs can offer significant performance advantages even for low power operation.

5. Three process modules have been developed and integrated with the Gennum's existing BJT technologies to allow fabrication of the new SiGe base HBT. These modules include field poly formation, SiGe-base deposition and local interconnect formation.

6. The first examples of both the self-aligned and non-self aligned devices have been successfully fabricated at Gennum Corporation. Electrical test results for this set of devices are very encouraging. The self-aligned devices show non-idealities in the Gummel plot at low current density, probably as a result of peripheral leakage resulting from the emitter poly stringer on the field poly sidewall. DC characteristics for the non-self aligned devices were near ideal. At $I_c = 200 \mu A$ typical self-aligned devices gave $F_t = 12.2$ GHz while a non-self aligned device gave $F_t = 13.2$ GHz.

7. Difference in performance of simulated and experimental devices are believed to result from differences in the width of the SiGe-base. The simulation results were based on the base-width equal to $450 \text{ \AA}$ in order to achieve the target specifications. However, in order to ensure that the present devices should work at this stage of the current phase, the base-width was arbitrarily increased to $700 \text{ \AA}$. The penalty for this action is that both F_t and F_m of the SA and NSA devices were seriously degraded. The F_t for SA1X and NS1X devices was found to be equal to 12.2 GHz and 13.2 GHz respectively, at $I_c = 200 \mu A$. In addition to the base width W_b , there are several main reasons which lead F_t for SA1X to be smaller than F_t of the NSA device.

6.2 Recommendations For Future Work

When this work was initiated, no experimental data was available on boron and germanium diffusion in an SiGe HBT structure of the kind considered here. Now that a first set of test devices has been produced, doping profiles should be obtained using

SIMS and/ or SRP techniques before and after emitter annealing. This would provide empirical data on the effect of the emitter anneal on device structure for use in future designs. In addition, availability of actual doping profiles would permit precise comparison of the predictions of the BIPOLE3 simulator with the performance of an experimental device. This calibration of BIPOLE3 would be very useful for subsequent redesign of doping and composition profiles.

The available information on the doping and composition profiles actually produced in the first set of test devices suggests that the base width was closer to 700 Å than the design value of 450 Å. An attempt should be made to produce a second set of test devices with neutral base width as close as possible to 450 Å.

To make the self-aligned HBT structure manufacturable in an industrial environment, the issue of polysilicon emitter stringers on the field poly sidewall must be addressed. This problem can probably be solved by reducing the field poly height, rounding the shape of the field poly sidewall, and perhaps modifying the emitter polysilicon plasma etch.

APPENDIX A

A1. The pn One-Sided Step Junction

The space charge (depletion) widths in the emitter-base X_{eb} and the base-collector X_{bc} regions at thermodynamic equilibrium can approximately be calculated using a depletion layer approximation. Only the final formulae are given below, and the details are given elsewhere [16].

$$X_{eb} = [(2 \epsilon_{si} V_{bic}/q) (N_d/N_a (N_a + N_d))]^{1/2}, \quad (A1.1)$$

$$X_{bc} = [(2 \epsilon_{si} V_{bic}/q) (N_a/N_d (N_a + N_d))]^{1/2}, \quad (A1.2)$$

where the N_a and N_d are ionized acceptor and donor dopant concentrations, respectively; q is the electronic charge; ϵ_{si} ($=\epsilon_0\epsilon_r$) is the permittivity of silicon; ϵ_0 and ϵ_r are the permittivity of the free space and the relative permittivity of the silicon ($=11.9$) [15] respectively; V_{bic} and V_{bic} are the built-in-voltage at the emitter-base and base-collector regions respectively, and in general is given by [15],

$$V_{bi} = (kT/q) \ln (N_a N_d / n_i^2), \quad (A1.3)$$

where k is the Boltzmann's constant and T is the absolute temperature; for $kT= 0.0259$ eV at 300 K, n_i (intrinsic carrier concentration) for silicon was taken to be equal to $1.6 \times 10^{10} / \text{cm}^3$ [15]. As mentioned above eqs. A1.1 and A1.2 were derived under the assumption of thermodynamic equilibrium, so when an external bias voltage is applied

across the pn junction, eqs. A1.1 and A1.2 should be modified in such away that this bias should be subtracted from the built-in voltage V_{bi} .

A2. Depletion Layer Capacitance

For a p-n junction, the depletion or the space-charge layer capacitances are given by [15],

$$C_{JE} = C_{je0} [1 - (V_{be}/V_{bie})]^{\gamma_e}, \quad (A2.1)$$

$$C_{JC} = C_{jc0} [1 - (V_{bc}/V_{bic})]^{\gamma_c}, \quad (A2.2)$$

where C_{je0} and C_{jc0} are the depletion layer capacitances of the two junctions at zero bias; γ_e and γ_c are coefficient constants which are taken to be equal to 1/2 and 1/3 respectively [15]. These values based on the assumption that the emitter has a uniform doping profile whereas the emitter-base junction was more closely approximated by a linearly graded junction profile. The present simulation results are in agreement with Roulston results [15]. The remaining parameters in eqs. A2.1 and A2.2 have their usual meaning. The values of the parameters of the right side of both eqs. A2.1 and A2.2 were obtained from the simulation analysis.

APPENDIX B

B.1 Tables

Table B.1 Part1. Heavy emitter doping, various geometries

$I_c = 200 \mu A$	$XJ_1 = 0.05 \mu m$	$BV_{cbo} = 6.0 V$
$V_{bc} = 1.0 V$	$XJ_2 = 0.1384 \mu m$	$BV_{cbo} = 18.7 V$
$ED = 6.0 \times 10^{20} cm^{-3}$	$W_b = 0.045 \mu m$	Reach through = 4.19 V
$BD = 1.0 \times 10^{18} cm^{-3}$	$TEPI = 0.59 \mu m$	
$EPI = 2.8 \times 10^{16} cm^{-3}$	$Ge = 10\%$	

Device Name	F_i GHz	F_m GHz	R_o Ω	R_{ext} Ω	C_{JE} fF	C_{JC} fF	C_{JSC} fF	$(F_i)_p$ GHz	$(F_m)_p$ GHz	$(I_c)_p$ μA	β
SA05X EPW=0.5 μm FPW=1.5 μm	26.2	37.2	100	290	5.66	1.96	9.95	27.7	38.9	277	570
SA125X EPW=1.25 μm FPW=2.25 μm	20.5	23.1	170	250	16.4	3.38	11.6	26.8	28.3	793	570
NSA125X EMW=1.25 μm FPW =2.25 μm	20.7	16.7 20.7*	180	770 440*	16.4	3.01	11.3	26.9	16.7 25.1*	794	562
SAF25X EPW=1.5 μm FPW=2.5 μm	19.2	20.6	190	260	19.6	3.76	11.3	26.6	26.3	948	565
NSAF25X EMW=1.5 μm FPW =2.5 μm	19.2	14.7 17.2*	190	760 460*	19.6	3.76	11.4	26.6	17.9 21.7*	948	565
SA-1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.9	2.9	10.8	17.0	20.3	597	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} cm^{-3}$	$XJ_1 = 0.05 \mu m$	$BV_{cbo} = 9.6 V$
$BD = 4.0 \times 10^{18} cm^{-3}$	$XJ_2 = 0.1398 \mu m$	$BV_{cbo} = 21.0 V$
$EPI = 1.0 \times 10^{16} cm^{-3}$	$W_b = 0.07 \mu m$	Reach through = 4.19 V
	$TEPI = 0.59 \mu m$	

Table B.1 Part2. Heavy emitter doping, various geometries

The BV_{cbo} and the reach through of the following devices are slightly different from the devices listed in Table B.1 part1, and their magnitudes are given below:

$BV_{cbo} = 19.0 \text{ V}$

Reach through = 5.17 V

Device Name	F_i GHz	F_m GHz	R_o Ω	R_{be} Ω	C_{π} fF	C_{jc} fF	C_{jsc} fF	$(F_i)_p$ GHz	$(F_m)_p$ GHz	$(I_c)_p$ μA	β
SA125X125 EPW=1.25 μm EML=1.25 μm FPW=2.25 μm	22.7	17.7	230	630	7.04	3.5	11.6	25.0	18.9	328	557
NSA125X125 EMW=1.25 μm EML= 1.25 μm FPW =2.75 μm	23.3	14.0 16.2*	240	1600 1200*	7.04	2.6	10.7	25.4	14.9 17.3*	328	557
SA125X25 EPW=1.25 μm EML= 2.5 μm FPW=2.25 μm	21.3	21.8	190	330	13.8	3.4	11.6	26.2	25.7	659	565
NSA125X25 EMW=1.25 μm EML= 1.5 μm FPW =2.75 μm	22.7	13.9 16.1*	230	1500 1000*	8.41	2.9	10.7	25.5	15.1 17.6*	394	565
SA15X15 EPW=1.5 μm EML= 1.5 μm FPW=3.0 μm	21.7	15.4	240	560	10.0	4.6	12.1	25.4	17.2	471	563
NSA15X15 EMW=1.5 μm EML= 1.5 μm FPW =3.0 μm	22.5	13.1 15.2*	250	1500 1000*	10.0	4.6	11.0	26.0	14.5 16.9*	471	563
SA-1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.9	2.9	10.8	17.0	20.3	597	127

Table B.2 Part1. Light emitter doping, various geometries

$I_c = 200 \mu A$
 $V_{be} = 1.0 V$
 $ED = 3.0 \times 10^{20} \text{ cm}^{-3}$
 $BD = 1.0 \times 10^{18} \text{ cm}^{-3}$
 $EPI = 2.8 \times 10^{16} \text{ cm}^{-3}$

$XJ_1 = 0.05 \mu m$
 $XJ_2 = 0.1381 \mu m$
 $Wb = 0.045 \mu m$
 $TEPI = 0.59 \mu m$
 $Ge = 10\%$

$BV_{\infty} = 5.7 V$
 $BV_{cbo} = 19.0 V$
 $Reach \text{ through} = 4.79 V$

Device Name	F_i GHz	F_m GHz	R_b Ω	R_{ext} Ω	C_{JE} fF	C_{JC} fF	C_{JSC} fF	$(F_i)_p$ GHz	$(F_m)_p$ GHz	$(I_c)_p$ μA	β
SA05X EPW=0.5 μm FPW=1.5 μm	25.2	35.6	99	290	5.78	2.0	9.95	26.2	37.0	275	775
SA125X EPW=1.25 μm FPW=2.25 μm	20.1	22.6	170	250	16.8	3.39	11.6	25.4	27.5	788	792
NSA125X EMW=1.25 μm FPW=2.25 μm	20.3	16.1 20.1*	170	770 440*	16.8	3.06	11.3	25.5	19.5 24.5*	788	792
SAF25X EPW=1.5 μm FPW=2.5 μm	18.7	20.2	190	260	20.0	3.82	11.3	25.4	25.7	941	809
NSAF25X EMW=1.5 μm FPW=2.5 μm	18.7	13.8 16.7*	190	760 460*	20.0	3.8	11.4	25.4	17.4 21.1*	941	809
SA-1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.9	2.9	10.8	17.0	20.3	597	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} \text{ cm}^{-3}$
 $BD = 4.0 \times 10^{18} \text{ cm}^{-3}$
 $EPI = 1.0 \times 10^{16} \text{ cm}^{-3}$

$XJ_1 = 0.05 \mu m$
 $XJ_2 = 0.1398 \mu m$
 $W_b = 0.07 \mu m$
 $TEPI = 0.59 \mu m$

$BV_{\infty} = 9.6 V$
 $BV_{cbo} = 21.0 V$
 $Reach \text{ through} = 4.19 V$

Table B.2 Part2. Light emitter doping, various geometries

Device Name	F_i GHz	F_m GHz	R_s Ω	R_{ext} Ω	C_{TE} fF	C_{JC} fF	C_{JSC} fF	$(F_i)_p$ GHz	$(F_m)_p$ GHz	$(I_o)_p$ μA	β
SA125X125 EPW=1.25 μm EML=1.25 μm FPW=2.25 μm	22.5	17.0	230	630	7.18	3.4	11.6	24.6	18.1	328	765
NSA125X125 EMW=1.25 μm EML= 1.25 μm FPW =2.75 μm	23.0	13.5 15.7*	240	1600 1200*	7.18	2.6	10.7	25.0	14.3 16.7*	328	765
SA125X25 EPW=1.25 μm EML= 2.5 μm FPW=2.25 μm	20.9	20.9	190	333	14.0	3.4	11.6	25.3	24.5	657	786
NSA125X25 EMW=1.25 μm EML= 1.5 μm FPW =2.75 μm	22.5	13.4 15.6*	230	1500 1000*	8.57	2.8	10.7	25.1	14.5 16.9*	394	776
SA15X15 EPW=1.5 μm EML= 1.5 μm FPW=3.0 μm	21.2	14.7	240	560	10.2	4.6	12.1	24.4	16.3	470	776
NSA15X15 EMW=1.5 μm EML= 1.5 μm FPW =3.0 μm	22.0	12.5 14.5*	250	1500 1000*	10.2	3.1	11.0	24.9	13.8 16.0*	470	776
SA-1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.9	2.9	10.8	17.0	20.3	597	127

Table B.3 Part1. Light emitter doping, light EPI doping, various geometries

$I_c = 200 \mu A$
 $V_{bc} = 1.0 V$
 $ED = 3.0 \times 10^{20} cm^{-3}$
 $BD = 1.0 \times 10^{18} cm^{-3}$
 $EPI = 1.0 \times 10^{16} cm^{-3}$

$XJ_1 = 0.05 \mu m$
 $XJ_2 = 0.1359 \mu m$
 $Wb = 0.045 \mu m$
 $TEPI = 0.59 \mu m$
 $Ge = 1\%$

$BV_{\infty} = 7.5 V$
 $BV_{cbo} = 19.2 V$
 $Reach\ through = 1.92 V$

Device Name	F_i GHz	F_a GHz	R_o Ω	R_{ext} Ω	C_{JE} fF	C_{JC} fF	C_{JSC} fF	$(F)_p$ GHz	$(F_a)_p$ GHz	$(I_c)_p$ μA	β
SA05X EPW=0.5 μm FPW=1.5 μm	4.9	19.8	97	290	5.78	1.3	10.2	22.7	41.3	97	799
SA125X EPW=1.25 μm FPW=2.25 μm	19.6	28.7	170	250	16.8	2.25	11.8	22.1	30.8	277	825
NSA125X EMW=1.25 μm FPW=2.25 μm	19.7	20.1 25.2*	170	1300 430*	16.8	1.99	11.5	22.2	21.4 27.0*	277	825
SAF25X EPW=1.5 μm FPW=2.5 μm	18.2	25.5	190	250	20.0	2.49	11.5	22.0	28.5	331	841
NSAF25X EMW=1.5 μm FPW=2.5 μm	18.2	17.2 20.9*	190	1200 450*	20.0	2.49	11.7	22.0	19.1 23.3*	331	841
SA-1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.9	2.9	10.8	17.0	20.3	597	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} cm^{-3}$
 $BD = 4.0 \times 10^{18} cm^{-3}$
 $EPI = 1.0 \times 10^{16} cm^{-3}$

$XJ_1 = 0.05 \mu m$
 $XJ_2 = 0.1398 \mu m$
 $W_b = 0.07 \mu m$
 $TEPI = 0.59 \mu m$

$BV_{\infty} = 9.6 V$
 $BV_{cbo} = 21.0 V$
 $Reach\ through = 4.19 V$

Table B.3 Part2. Light emitter doping, light EPI doping, various geometries

$$I_c = 200 \mu A$$

$$V_{be} = 1.0 V$$

$$ED = 3.0 \times 10^{20} \text{ cm}^{-3}$$

$$BD = 1.0 \times 10^{18} \text{ cm}^{-3}$$

$$EPI = 1.0 \times 10^{16} \text{ cm}^{-3}$$

$$XJ_1 = 0.05 \mu m$$

$$XJ_2 = 0.1359 \mu m$$

$$Wb = 0.045 \mu m$$

$$TEPI = 0.59 \mu m$$

$$Ge = 10\%$$

$$BV_{ceo} = 7.5 V$$

$$BV_{cbo} = 19.2 V$$

$$\text{Reach through} = 1.92 V$$

Device Name	F_i GHz	F_a GHz	R_o Ω	R_{in} Ω	C_{JE} fF	C_{JC} fF	C_{JSC} fF	$(F)_p$ GHz	$(F_a)_p$ GHz	$(I_c)_p$ μA	β
SA125X125 EPW=1.25 μm EML=1.25 μm FPW=2.25 μm	5.8	10.8	230	610	7.18	2.22	11.8	21.1	20.8	116	775
NSA125X125 EMW=1.25 μm EML= 1.25 μm FPW =2.75 μm	6.0	8.8 9.6*	240	1600 1100*	7.18	1.65	10.9	21.6	16.2 18.9*	116	775
SA125X25 EPW=1.25 μm EML= 2.5 μm FPW=2.25 μm	20.8	26.7	180	320	14.0	2.22	11.8	22.0	27.5	232	813
NSA125X25 EMW=1.25 μm EML= 1.5 μm FPW =2.75 μm	7.9	10.0 11.7*	230	1400 1000*	8.55	1.8	10.9	21.7	16.4 19.1*	139	786
SA15X15 EPW=1.5 μm EML= 1.5 μm FPW=3.0 μm	14.6	15.3	230	540	10.19	2.95	12.3	21.1	18.8	166	792
NSA15X15 EMW=1.5 μm EML= 1.5 μm FPW =3.0 μm	16.2	13.3 15.5*	240	1400 1000*	10.2	1.99	11.2	21.6	15.6 18.1*	165	791
SA-1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.93	2.9	10.8	17.0	20.3	597	127

Table B.4 Heavy emitter doping, light EPI doping, various geometries

$I_c = 200 \mu A$
 $V_{bc} = 1.0 V$
 $ED = 6.0 \times 10^{20} \text{ cm}^{-3}$
 $BD = 1.0 \times 10^{18} \text{ cm}^{-3}$
 $EPI = 1.0 \times 10^{16} \text{ cm}^{-3}$

$XJ_1 = 0.05 \mu m$
 $XJ_2 = 0.1362 \mu m$
 $W_b = 0.045 \mu m$
 $TEPI = 0.59 \mu m$
 $Ge = 10\%$

$BV_{\infty} = 7.7 V$
 $BV_{cbo} = 19.2 V$
 $\text{Reach through} = 1.92 V$

Device Name	F_i GHz	F_m GHz	R_o Ω	R_{in} Ω	C_{π} fF	C_{jc} fF	C_{isc} fF	$(F_i)_p$ GHz	$(F_m)_p$ GHz	$(I_c)_p$ μA	β
SA05X EPW=0.5 μm FPW=1.5 μm	24.0	43.3	97	290	5.66	1.24	10.2	24.5	43.2	137	578
SA125X EPW=1.25 μm FPW=2.25 μm	20.5	29.1	170	250	16.4	2.09	11.8	23.8	32.0	391	608
NSA125X EMW=1.25 μm FPW=2.25 μm	20.6	20.5 25.7*	170	760 430*	16.4	1.86	11.5	24.3	22.4 28.2*	391	608
SAF25X EPW=1.5 μm FPW=2.5 μm	18.9	25.8	190	250	19.6	2.33	11.5	23.7	29.6	466	619
NSAF25X EPW=1.5 μm FPW=2.5 μm	18.9	17.6 21.3*	190	740 450*	19.57	2.33	11.7	24.1	20.0 24.4*	466	619
SA1X* EPW=1.5 μm FPW=3.0 μm	14.1	17.4	620	150	17.9	2.9	10.8	17.0	20.3	597	127

* Underestimates base resistance

+ Si-base doping profiles:

$ED = 3.0 \times 10^{20} \text{ cm}^{-3}$
 $BD = 4.0 \times 10^{18} \text{ cm}^{-3}$
 $EPI = 1.0 \times 10^{16} \text{ cm}^{-3}$

$XJ_1 = 0.05 \mu m$
 $XJ_2 = 0.1398 \mu m$
 $W_b = 0.07 \mu m$
 $TEPI = 0.59 \mu m$

$BV_{\infty} = 9.6 V$
 $BV_{cbo} = 21.0 V$
 $\text{Reach through} = 4.19 V$

REFERENCES

- [1] B.S. Meyerson, "UHV/CVD Growth of Si and SiGe Alloys: Chemistry, Physics, and Applications", *Proceedings of the IEEE*, **80** (1992) 1592.
- [2] A. Pruijmboom, C.E. Timmering, J.M.L. Van Rooij-Mulder, D.J. Gravesteijn, W.B. de Boer, W.J. Kersten, J.W. Slotboom, C.J. Vriezema, and R. de Kruif, "Heterojunction bipolar transistor with $\text{Si}_{1-x}\text{Ge}_x$ base", *Microelectronic Engineering*, **19** (1992) 427.
- [3] T.E. Whall and E.H.C. Parker, "Silicon germanium heterostructures- advanced materials and devices for silicon technology", *J. Mat. Sci.: Materials in Electronics*, **6** (1995) 249.
- [4] R. Gotzfried, F. Beisswanger and S. Gerach, "Design of RF Integrated Circuits Using SiGe Bipolar Technology", *IEEE BCTM* 3.1, p.51, 1997.
- [5] L.E. Larson, "Silicon Bipolar Transistor Design and Modeling for Microwave Integrated Circuit Applications", *IEEE BCTM* 9.1, p.142, 1996.
- [6] E.F. Crabbe, B.S. Meyerson, J.M.C. Stork and D.L. Hareme, "Vertical Profile Optimization of Very High Frequency Epitaxial Si- and Si-Ge-base Bipolar Transistors", *IEDM Tec.Dig.*, p.83, 1993.
- [7] J.D. Warnock, "Silicon Bipolar Device Structures for Digital Applications: Technology, Trends, and Future Directions", *IEEE Trans. Electron Devices*, pp.377, 1995.
- [8] K. Oda, E. Ohue, M. Tanabe, H. Shimamoto, T. Onai and K. Washio, "130-GHz F, SiGe HBT Technology", *IEDM Tec.Dig.*, p.371, 1997.
- [9] A. Schuppen, U. Erben, A. Gruhle, H. Kibbel, H. Schumacher, and U. Konig, "Enhanced SiGe Heterojunction Bipolar Transistors with 160 GHz- F_m ", *IEDM Tech. Dig.*, p.743, 1995.
- [10] D.C. Ahlgren, G. Freeman, S. Subbanna, R. Groves, D. Greenberg, J. Malinowski, D. Nguyen-Ngoc, S.J. Jeng, K. Stein, K. Schonenberg, D. Kiesling, B. Martin, S. Wu, D.L. Hareme, and B. Meyerson, "A SiGe HBT BICMOS Technology for Mixed Signal RF Applications", *IEEE BCTM* 12.3, p.195, 1997.
- [11] A.G. Milnes and D.L. Feucht, "Heterojunctions and Metal-Semiconductor Junctions", *Academic Press, New York*, chp.1, 1972.

- [12] D.L. Harame, J.H. Comfort, J.D. Cressler, E.F. Crabbe, J.Y.-C. Sun, B.S. Meyerson and T. Tice, "Si/SiGe Epitaxial-Base Transistors- Part I: Materials, Physics, and Circuits", *Trans. Electron Devices*, **42** (1995) 455.
- [13] H. Kroemer, "Two integral relaxations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region", *Solid-State Electron*, **28** (1985) 1101.
- [14] E.J. Prinz and J.C. Sturm, "Analytical Modeling of Current Gain-Early Voltage Products in Si/Si_{1-x}/Ge_x Heterojunction Bipolar Transistors", *IEDM Tech. Dig.*, p.853, 1991.
- [15] D.J. Roulston, "Bipolar Semiconductor Devices", McGraw-Hill Publishing Company, New York, chp.9, 1990.
- [16] S.M. Sze, "Physics of semiconductor devices", John Wiley & Sons, New York, 1981.
- [17] C.G. Van-de-Walle and R.M. Martin, "Theoretical calculations of heterojunction discontinuities in the Si/Ge system", *Phys. Rev. B.*, **34** (1986) 5621.
- [18] D.C. Houghton, J.-M. Baribeau and N.L. Rowell, "Misfit dislocation injection, interfacial stability and photonic properties of Si-Ge strained layers", *J. of Materials Science: Materials in Electronics*, **6** (1995) 280.
- [19] C.A. King et al, "Electrical and material quality of Si_{1-x}Ge_x / Si P-N heterojunctions produced by limited reaction processing", *IEEE Electron Device Lett.*, **10** (1989) 159.
- [20] J. Matthews and A. Blakeslee, "Defects In Epitaxial Multilayers", *J. Crystal Growth*, **27** (1974) 118.
- [21] B.S. Meyerson, "Low temperature silicon epitaxy by UHV/CVD", *Appl. Phys. Lett.*, **48** (1986) 797.
- [22] S.S. Iyer, G.L. Patton, S.L. Delage, S. Tiwari and J.M.C. Stork, "Silicon-Germanium Base Heterojunction Bipolar Transistors by Molecular Beam Epitaxy", *Technical Digest of the 1987 International electron Device Meeting*, p.874, 1987.
- [23] J.N. Burghartz, et al, "An Ultra-Low Thermal Budget SiGe Base Bipolar Technology", *Technical Digest of 1993 Symposium on VLSI Technology*, p.59, 1993.

- [24] M. Hong, et al, "High-Performance SiGe Epitaxial Base Bipolar Transistors Produced by a Reduced-Pressure CVD Reactor", *IEEE Electron Device Lett.*, 14 (1993) 450.
- [25] B.S. Meyerson, F.J. Himpsel and J. Uram, "Bistable conditions for low-temperature Silicon epitaxy", *Appl. Phys. Lett.*, 57 (1990) 1034.
- [26] T. Kamins, "Polycrystalline Silicon for Integrated Circuit Applications", Kluwer Academic Publishers, Boston, p.155, 1988.
- [27] P. Asburn, "Design and Realization of Bipolar Transistors", John Wiley & Sons, Toronto, chp.4, 1988.
- [28] P.-E. Hellberg, S.-L. Zhang, and C.S. Peterson, "Work Function of Boron Doped Polycrystalline $\text{Si}_x\text{Ge}_{1-x}$ Films", *IEEE Electron Device Letters*, 18 (1997) 456.
- [29] S.J. Jeng, D.C. Ahlgren, G.D. Berg, B. Ebersman, G. Freeman, D.R. Greenberg, J. Malinowski, D. Nguyen-Ngoc, K.T. Schonenberg, K.J. Stein, D. Colavito, M. Longstreet, P. Ronsheim, S. Subbanna, D.L. Harame, "Impact of Extrinsic Base Process on NPN HBT performance and Polysilicon Resistor in Integrated SiGe HBTs", *IEEE BCTM 12.1*, 187, 1997.
- [30] G.L. Patton, J.C. Bravman and J.D. Plummer, "Physics, Technology, and Modeling of Polysilicon Emitter Contacts for VLSI Bipolar Transistors", *IEEE Trans. Electron Devices*, 33 (1986) 1754.
- [31] D.J. Roulston, "Low current base-collector boundary conditions in GHz frequency transistor", *Solid State Electronics*, 18 (1975) 845.
- [32] J.J.H. Van Den Biessen, "A simple regional analysis of transit times in bipolar transistors", *Solid State Electronics*, 29 (1986) 529.
- [33] BIPOLE3 User Manual: is a product of BIPSIM Inc., developed by Prof. D.J. Roulston at University of Waterloo, Ontario, Canada.
- [34] C.H.H. Emons, G.A.M. Hurkx, H.E.J. Pijpers, M.S. Peter, R. Koster, and J.W. Slotboom, "A new high-performance CMOS-compatible reduced-area bipolar transistor", *IEEE BCTM 2.2*, p.44, 1997.
- [35] L-Edit is developed by Tanner Research, Inc., Tanner EDA Division, CA, USA.
- [36] SUPREM-3 and SUPREM-4 are Technology Modeling Associates, version of the process simulator, originally developed at Stanford University, U.S.A.

- [37] D.J. Roulston, S.G. Chamberlain, J. Sehgal, "Simplified Computer Aided Analysis of Double Diffused Transistors Including Two-Dimensional High Level Effects", *IEEE Trans. Electron Devices*, **19** (1972) 809.
- [38] D.J. Roulston, A.A. Eltoukhy, "Bipolar Device Simulation Using a Combination of 1D Regional and 2D Regional Numerical analysis", *NASECODE III Proceedings*, Boole Press, Dublin, p.236, 1983.
- [39] G.L. Patton, D.L. Harame, J.M.C. Stork, B.S. Meyerson, G.J. Scilla and E. Ganin, "Graded-SiGe-Base, Poly-Emitter Heterojunction Bipolar Transistors", *IEEE Electron Device Letters*, **10** (1989) 534.
- [40] N.H. Fletcher, "Aspects of the Design of Power Transistors", *Proc.IRE*, **43** (1955) 551.
- [41] S. Wolf and R.N. Tauber, "Silicon Processing For The VLSI Era", Lattice Press, Sunset Beach, California, Vol II, 1990.
- [42] D.J. Roulston, private communication, April, 1998.
- [43] L.J. van der Pauw, "A method of Measuring Specific Resistivity and Hall effect of Discs of Arbitrary Shape", *Phil. Res. Rep.*, **13** (1958) 1.
- [44] S. Wolf, "Silicon Processing For The VLSI Era", Lattice Press, Sunset Beach, California, Vol I, 1986.
- [45] K. Shinada, S. Mori and Y. Mikata, "Reduction in Polysilicon Oxide Leakage Current by Annealing Prior to Oxidation", *J. Electrochem. Soc.*, **132** (1988) 2185.
- [46] G.L. Patton, S.S. Iyer, S.L. Delage, E. Ganin and R.C. McIntosh, "Oxidation of strained SiGe layers grown by MBE", *Mat. Res. Soc. Symp. Proc.*, **102** (1988) 295.
- [47] Rapid Thermal Processing System Operations Manual (Heatpulse 210T-02), AG Associates, Palo Alto, CA, USA.