DESIGN AND OPTIMIZATION OF NANO-SCALED

SILICON-GERMANIUM HETEROJUNCTION BIPOLAR TRANSISTORS

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

_____

Yun Shi

Certificate of Approval:

_____	_____
Foster Dai Associate Professor Electrical and Computer Engineering	Guofu Niu, Chair Professor Electrical and Computer Engineering
_____	_____
Victor P. Nelson Professor Electrical and Computer Engineering	Stephen L. McFarland Acting Dean Graduate School

DESIGN AND OPTIMIZATION OF NANO-SCALED SILICON-GERMANIUM HETEROJUNCTION BIPOLAR TRANSISTORS

Yun Shi

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

Auburn, Alabama
August 2005

DESIGN AND OPTIMIZATION OF NANO-SCALED

SILICON-GERMANIUM HETEROJUNCTION BIPOLAR TRANSISTORS

Yun Shi

_____

Signature of Author

_____

Date

Copy sent to:

Name Date

VITA

Yun Shi was born on February 19, 1977, in Beijing, P. R. China. She earned the Bachelor's degree in 1999 from Beijing University of Posts and Telecommunications, majoring in wireless communications. After graduation, she joined Great Dragon Telecom and worked on the development of internet management protocol (SNMP) using embedded operating system. In August 2001, she started her master program and joined the Alabama Microelectronics and Science Technology Center at Auburn University. Ever since then, she has been working on semiconductor device physics/modeling and TCAD. She earned the Master of Science degree in July 2003, and continued on the Ph. D program.

DISSERTATION ABSTRACT

DESIGN AND OPTIMIZATION OF NANO-SCALED SILICON-GERMANIUM HETEROJUNCTION BIPOLAR TRANSISTORS

Yun Shi

Doctor of Philosophy, August 2005
(M. S., Auburn University, Auburn, Al, USA. July 2003)
(B.E.E., Beijing University of Posts and Telecommunications, Beijing, P. R. China, July, 1999)

133 Typed Pages

Directed by Prof. Guofu Niu

In this work, we explore the design and optimization of nano-scaled SiGe HBTs. The cutoff frequency f_T and the maximum oscillation frequency f_{max} are optimized towards Terahertz. We first start with intrinsic device designs to obtain the initial one-dimensional doping profile, which features a 7 nm base width, a high base doping of $8\times10^{19}/\text{cm}^3$, and a 25 nm collector width. Using this profile, the impacts of Ge designs on device performance, including β, regional transit time, f_T, and $f_T\times BV_{CEO}$ product, are examined at the same film stability. After comprehensive comparisons, we conclude that graded Ge profile wins over box Ge profile in device performance metrics.

Then we focus on the 2D device scaling for the purpose of minimizing device parasitic effects and maintaining the high performance achieved by the highly scaled 1D design. The raised extrinsic base 2D structure is used, which is widely used in 200 GHz – 350 GHz SiGe HBT technologies. To better understand 2D parasitic effects,

the transit time analysis is extended to 2D, which provides a method to analyze the distributive capacitance. Two lateral scaling schemes, fixed total base width (w_{Base}) and fixed extrinsic base width ($w_{ex,B}$), are examined. When emitter window is scaled to 60 nm, f_{max} can be optimized to 1090 GHz for both scaling. However, fixed $w_{ex,B}$ scaling features less extrinsic base transit time $\tau_{B,ex}$, and hence the higher f_T. Therefore, fixed $w_{ex,B}$ lateral scaling is favored for those highly-scaled SiGe HBT designs.

ACKNOWLEDGMENTS

I am so grateful that there are so many people in my life who support and encourage me to finish this work. Without any of them, it is hard to imagine how I can stay focused and motivated through. First of all, I want to acknowledge my parents – for years, their unconditional love and continuous support has given me strength to grow and overcome difficulties. This dissertation is dedicated to them. Secondly, I want to thank my advisor Prof. Guofu Niu for giving me the opportunity to start my graduate studies in semiconductor physics/TCAD and providing me a strong base to undertake this work. I also thank my advisory committee, Dr. Foster Dai, Dr. Adit D. Singh, Dr. Victor Nelson, and Dr. John Williams for their interests and comments. Additional thanks to Tongan Wang, Rui Yan, Hua Yang, Weidong Tang for their long time encouragement. Thanks to Kejun Xia for his fruitful discussions and friendship. I am thankful for the financial assistance provided by the National Science Foundation under ECS-0119623 and ECS-0112923, the Semiconductor Research Corporation under SRC #2001-NJ-937, an IBM Faculty Partner Award, and the Alabama Microelectronics Science and Technology Center. Finally, I want to acknowledge the Lord Jesus Christ for making the race worth running and the prize worth living for.

Style manual or journal used IEEE Transactions on Electron Devices (together with the style known as "auphd"). Bibliography follows van Leunen's *A Handbook for Scholars*.

Computer software used The document preparation package TEX (specifically LATEX) together with the departmental style-file `auphd.sty`. The plots were generated using MICROSOFT VISIO ®, MATLAB®.

TABLE OF CONTENTS

List of Figures

LIST OF TABLES

CHAPTER 1

MODERN SIGE HBT OVERVIEW

The Silicon-Germanium heterojunction bipolar transistor (SiGe HBT) has gained worldwide attention, ever since the first high performance SiGe HBT with a cutoff frequency (f_T) of 75 GHz was grown using ultra-high vacuum/chemical vapor deposition (UHV/CVD) [1, 2] in 1990. Compared with Si BJT and Si CMOS, SiGe HBT features high speed, low noise, and high linearity, which make SiGe HBT technology more suitable and attractive to IC designs. After the first SiGe BiCMOS technology was invented in 1992 [3], single system-on-chip (SoC) integration became reality by growing the high-performance RF SiGe HBTs and the low-power, high-density Si CMOS on the same substrate. BiCMOS technology provides a broad market for SiGe HBT technology, from wired to wireless applications. In recent years, the explosion of telecommunications, such as 3G wireless and fiber communication, has been driving SiGe HBT technology growth from one generation to the next. Figure 1.1 shows the growth trend of SiGe HBTs. The cutoff frequency (f_T), which represents the device scaling and hence the technology generation, has been improved over 300 GHz [8], and a 600 GHz f_T SiGe HBT design [9] has been proposed.

Terahertz band applications have become interesting research areas in recent years. These applications cover medical and biological imaging [10], radio astronomy [11], upper atmosphere study [12], plasma diagnostics [13], and ultra-wide band communication systems that would benefit from the virtually infinite bandwidth with THz spectrum.

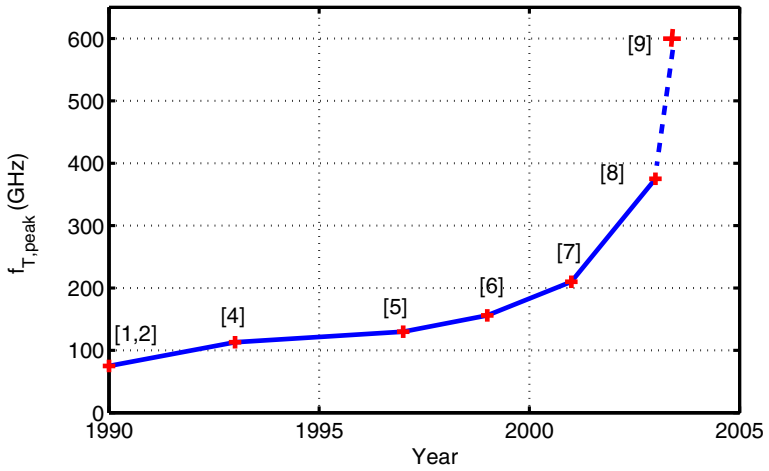


Figure 1.1: The growth of SiGe HBT technology.

Currently, GaAs Schottky barrier diodes and Gunn diodes have been used as key components in THz applications, such as mixers, power generators and oscillators [14–16]. However, they do not provide the ability to integrate the signal processing function with switching and amplification functions. Being a potential candidate, the operation frequency of SiGe HBT needs to be further improved through scaling.

In this work, we explore the SiGe HBT scaling issues for THz applications. First, the vertical profile scaling for an intrinsic device was examined. A nanometer-scale design was obtained, which can achieve over 1000 GHz f_T [17]. When the base width is scaled below 10 nm, the collector design becomes the bottleneck for speed. Higher doping (N_C) and smaller thickness (w_C) are proved to improve f_T at high injection, and hence achieve THz f_T. The Ge design issues for such highly scaled SiGe HBTs are also examined at the same film stability. The results show that a graded Ge profile surpasses the box Ge profile in all device performance metrics.

For THz SiGe HBTs, the emitter–collector transit time, which is $\tau_{EC} = 1/2\pi f_T$, is on the order of 0.1 ps. Therefore, the delay from extrinsic parasitic effects can degrade the ultra-high speed of the highly scaled intrinsic design. Besides f_T, the maximum oscillation frequency f_{max} is also relevant to circuit designs, and sensitive to extrinsic parasitic effects. A simplified expression is

$$f_{max} \approx \sqrt{\frac{f_T}{8\pi R_B C_{tc}}}, \tag{1.1}$$

where R_B is the base resistance, and C_{tc} is the CB junction depletion capacitance. In order to reduce parasitic effects, device structures have been improved through many innovations, such as shallow-trench-isolation (STI) [18], double-polysilicon technology (for self-aligned emitter structure) [19, 20], selectively ionized collector (SIC) [21], SiGe:C base [22], raised extrinsic base [23], and selectively-epitaxial grown (SEG) base [24]. These design innovations have been implemented in this work. Another method of reducing parasitic effects is lateral scaling. As we will see later, as f_T increases from 50 GHz to 300 GHz, the emitter window w_E decreases from 0.5 μm to 0.12 μm to achieve high f_{max}. The impact of lateral scaling on THz SiGe HBTs is examined, and 1000 GHz f_{max} is achieved when w_E is reduced to 60 nm.

1.1 SiGe HBT Operations with Device Scaling

The essential benefit of a SiGe HBT over Si BJT is that we can reduce the bandgap E_g by introducing Ge in the base. The result is an exponential increase of the effective

intrinsic carrier concentration, and hence the minority carrier density. By controlling the Ge profile, an accelerating electric field can be established in the base, which favors minority electron transport for an *npn* transistor, and thus helps reduce base transit time τ_B. Another advantage of using a SiGe base is that the low base sheet resistance $R_{B,\square}$, high β and low τ_B can be achieved simultaneously.

Much effort has been made for SiGe HBT scaling in order to improve device performance. A brief summary of the structure scaling trend and its impact on electrical performance are listed in Table 1.1. For the emitter design, phosphorus is preferred due

Table 1.1: SiGe HBT device scaling [25].

	Emitter	Base	Collector
Structure	Arsenic to Phosphorus Reduce emitter window	Narrow base Improve extrinsic design Higher Ge ramp Add Carbon	Increase doping Improve extrinsic design
Electrical	Lower R_E Reduced thermal budget for dopants drive-in	Lower R_B Lower τ_B	Higher avalanche Higher $J_{C,peak}$ Higher C_{CB}

to its low resistivity, which helps reduce emitter resistance R_E. For the base design, besides the aggressive base width w_B reduction, a higher Ge ramp is also used to help reduce base transit time τ_B. Additionally, base doping is increased, and the extrinsic base design is improved to reduce R_B. For the collector design, a higher doping and a narrow thickness design are used to reduce collector transit time τ_C. Inevitably, the Kirk effect is surpassed, but the breakdown voltage is reduced due to the higher CB space charge region (SCR) electric field. The impacts on device performance through growing technologies are compared in Table 1.2 for five generations of SiGe HBTs [26]. As

Table 1.2: Comparison of key performance parameters for different SiGe growth technologies [26].

Generation Index	1	2	3	4	5
w_E (μm)	0.5	0.25	0.18	0.13	0.12
f_T (GHZ)	47	47	120	210	375
f_{max} (GHz)	65	65	100	285	210
β	100	100	350	300	3500
BV_{CEO} (V)	3.4	3.4	1.8	1.7	1.4
BV_{CBO} (V)	10.5	10.5	6.5	5.5	5.0
$J_{C,peak}$ (mA/μm^2)	1.5	1.5	8	12	23

discussed, the emitter window w_E is reduced from 0.5 to 0.12 μm, f_T/f_{max} is improved to 375/210 GHz, and the Kirk effect is delayed, which is indicated by the increase of $J_{C,peak}$. However breakdown performance is sacrificed. BV_{CEO} is reduced from 3.4 V to 1.4 V. Next, we will briefly discuss how the performance of each device is impacted through vertical scaling.

1.1.1 f_T and Device Scaling

The cutoff frequency f_T, a figure-of-merit for device speed, can be generally expressed as

$$\frac{1}{2\pi f_T} = \tau_f + \tau_t, \tag{1.2}$$

$$\tau_f = \tau_E + \tau_B + \tau_{CSCR} \tag{1.3}$$

$$\tau_t = \frac{V_{th}}{J_C}(C_{te} + C_{tc}). \tag{1.4}$$

τ_f is the forward transit time due to electron diffusion for an *npn* SiGe HBT. τ_t is the depletion capacitance charging time associated with EB junction (C_{te}) and CB junction (C_{tc}). V_{th} is the thermal voltage and J_C is the collector current density. τ_f includes the emitter transit time τ_E, the base transit time τ_B, and the collector space charge region transit time τ_{CSCR}. For an *npn* bipolar transistor with constant base doping and Ge profile, τ_E and τ_B can be expressed as

$$\tau_E = \frac{1}{\beta}\frac{w_E^2}{2D_p}, \tag{1.5}$$

$$\tau_B = \frac{w_B^2}{2D_n}. \tag{1.6}$$

w_E and w_B are the quasi-neutral emitter and base width, β is the current gain, D_p is the minority hole diffusivity in the emitter and D_n is the minority electron diffusivity in the base. τ_E can be reduced significantly through increasing β. Depending on the emitter and base design, a sufficient β is required to make τ_E negligible. A main feature of scaled SiGe HBT is the significantly reduced base width w_B, which has decreased from 100 to 5 nm, accompanied by increased base doping N_B. Once the emitter and base are scaled down to the nano-scale, τ_{CSCR} becomes the bottleneck for speed. Generally, τ_{CSCR} can be expressed as [27]

$$\tau_{CSCR} = \frac{1}{X_{CSCR}}\int_0^{X_{CSCR}} \frac{X_{CSCR} - x}{v(x)}dx, \tag{1.7}$$

where X_{CSCR} is the space charge region width, and $v(x)$ is the local carrier velocity. For equilibrium transport, where the electron velocity saturates at V_{sat} in the collector SCR,

the right side of (1.7) equals "$X_{CSCR}/2V_{sat}$". Using first-order approximations [28]

$$X_{CSCR} \propto \frac{1}{\sqrt{N_C}} \tag{1.8}$$

Therefore, to fully benefit from base scaling, the collector doping N_C needs to be increased to reduce τ_{CSCR}.

1.1.2 $J_{C,peak}$ and Device Scaling

Observing (1.2) - (1.4), at the same J_C, f_T can be improved through the τ_f and junction capacitance reduction. Interestingly, however, when N_C is increased to reduce collector transit time τ_{CSCR}, the Kirk effect is delayed simultaneously. The reason is that the current density $J_{C,peak}$ at which base push out occurs [29] is proportional to N_C. As a result, not only does the peak f_T value increase, but also it happens at higher J_C. f_T rolls off when $J_C > J_{C,peak}$. $J_{C,peak}$ has increased from 3 mA/μm^2 to 23 mA/μm^2 as technology increases from 50 GHz to 375 GHz [25], as shown in Figure 1.2.

The concerns with the higher J_C are the potential device degradations associated with electromigration and self-heating, both of which are addressed by spreading current and power over an increasingly narrow emitter. Copper-interconnection and appropriate emitter layout can eliminate the electromigration. For device self-heating, it has been demonstrated in [25] that when the emitter width is scaled down proportionally to the increased $J_{C,peak}$, and the device perimeter is maintained constant, self-heating can be controlled to guarantee reliable device operation. Figure 1.3 shows one such example [25]. As $J_{C,peak}$ increases with device scaling, w_E is scaled proportional to $1/J_{C,peak}$, and

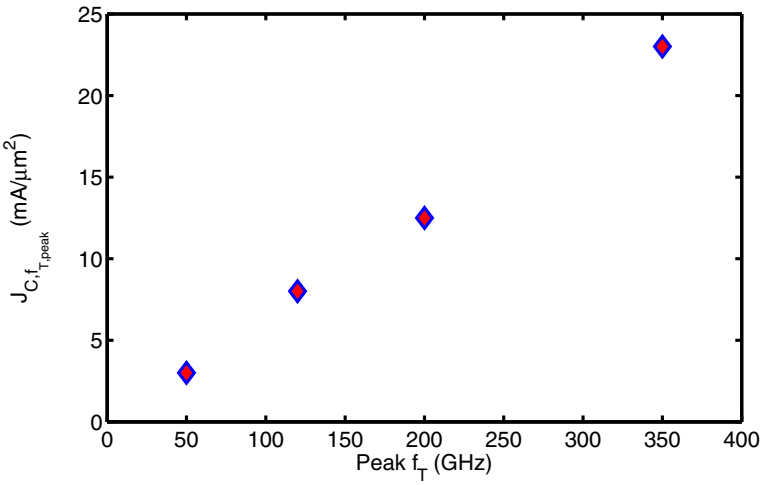


Figure 1.2: $J_{C,peak}$ versus $f_{T,peak}$ with SiGe HBT scaling technology [25].

emitter length l_E is chosen that f_T peaks at the same collector current $I_C = 2$ mA for each technology. As a result, the self-heating induced junction temperature increase can be maintained below 20 K. For the THz designs proposed in this work, $J_{C,peak}$ is as high as 40 mA/μm^2. A much more careful layout is required to maintain the device reliability with such high values of $J_{C,peak}$.

1.1.3 Avalanche and Device Scaling

The higher N_C also increases the CB SCR electric field. As we know, a higher field leads to a higher impact ionization rate, and thus a more severe avalanche effect. The

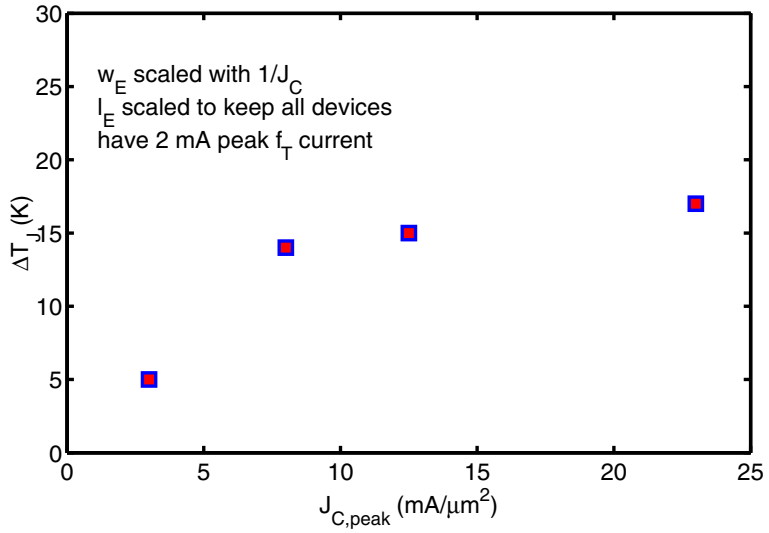


Figure 1.3: Increase of junction temperature with V_{CE}=1 V. L_E is scaled that all device have 2 mA peak f_T current [25].

avalanche effect can be measured using the M-1 factor as

$$M - 1 = \frac{I_{n,out}}{I_{n,in}} - 1, \tag{1.9}$$

$$= \frac{I_C(V_{CB})}{I_E - I_B(V_{BE})|_{V_{CB}=0}} - 1, \tag{1.10}$$

where $I_{n,out}$ is the electron current flowing out of the collector SCR, and $I_{n,in}$ is the electron current flowing into the collector SCR. This trend is shown in Figure 1.4 [25]. A higher $M - 1$ indicates the more severe avalanche effect.

The breakdown voltage BV_{CEO} equals the V_{CE} at which $M - 1 = 1/\beta$. With increasing β, the threshold $M - 1$ decreases. Therefore the BV_{CEO} decreases as well

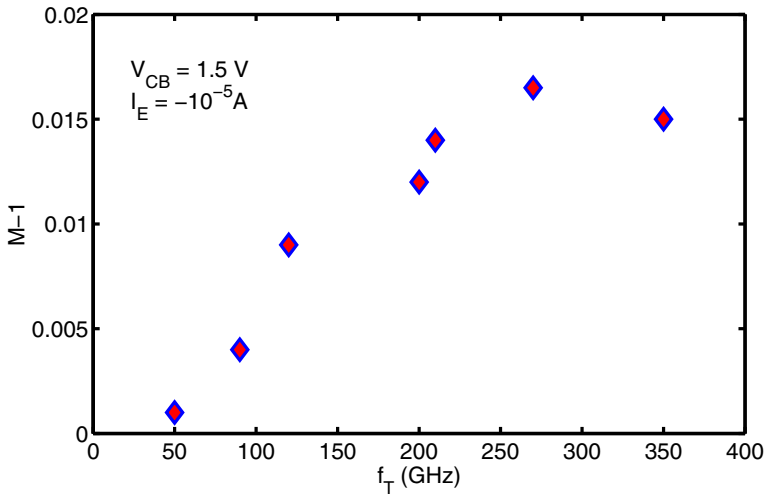


Figure 1.4: M-1 versus $f_{T,peak}$ at I_E=10^{-5}A and V_{CB} = 1.5 V [25].

with technology growth. As shown in Table 1.2, BV_{CEO} is reduced from 3.4 to 1.4 V when f_T increases from 47 to 375 GHz.

1.1.4 R_B and Scaling

Base resistance R_B impacts both RF noise performance and f_{max}. R_B reduction can be achieved through the lateral scaling as well as structural improvements, such as the double-poly process and the raised polysilicon base structure. Figure 1.5 shows the R_B improvement versus lithography node [30]. The data are normalized by emitter stripe length and emitter area, respectively. As shown, a 15–18% per generation decrease in length-normalized R_B is observed when w_E is reduced from 0.5 μm to 0.18 μm. The reason is the intrinsic base resistance $R_{B,in}$ reduction, which is inversely proportional to

w_E as

$$R_{B,in} = \frac{1}{12}\frac{w_E}{l_E}R_{B,\square}, \tag{1.11}$$

where l_E is the emitter length, and $R_{B,\square}$ is the base sheet resistance.

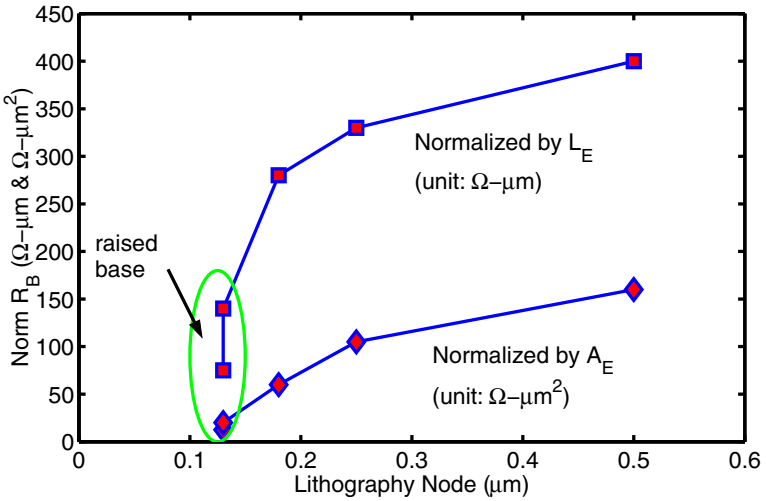


Figure 1.5: Base resistance (R_B) normalized by emitter stripe length and area versus lithography node [30].

For these early technologies, a "shared-base" design is used to form the extrinsic base. Figure 1.6 (a) shows the 2D cross section of a shared-base structure. The disadvantage is that the p^+ extrinsic base leads to a high extrinsic C_{tc}. The implantation process also damages the collector epi-layer and induces traps, which further increases C_{tc}. Starting from 0.13 μm node, a "raised-base" design is used to form the extrinsic base on top of the SiGe base, as shown in Figure 1.6 (b). Because the extrinsic C_{tc} is not related to the p^+ base, a higher doping can be used, which helps reduce R_B. As a result, a much larger 68% drop in R_B is observed for the length-normalized R_B.

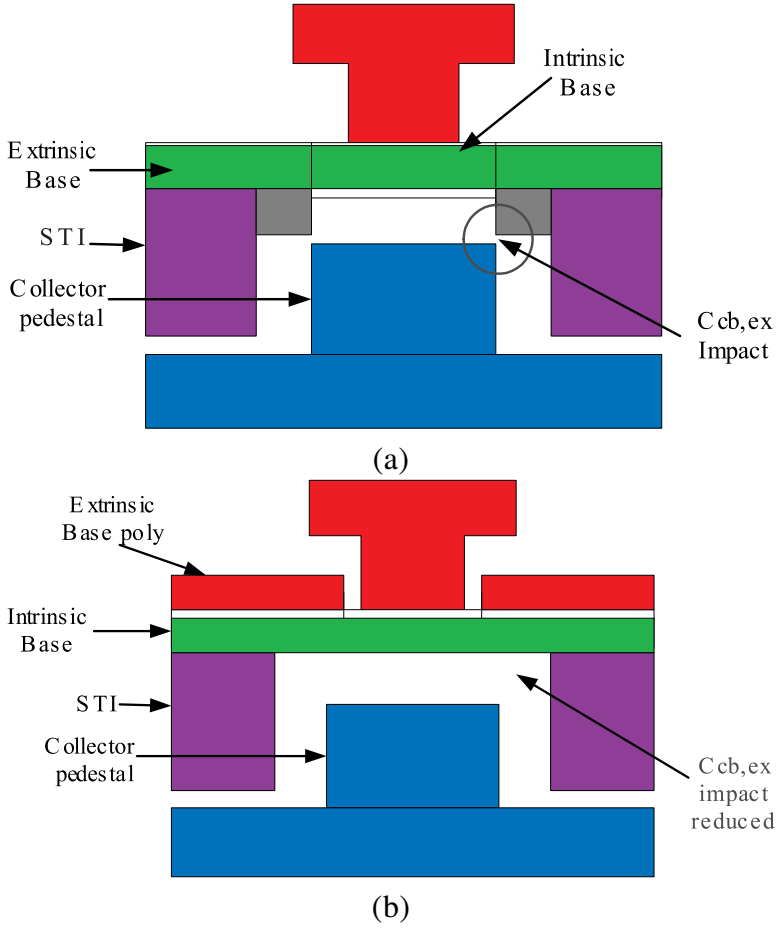


Figure 1.6: 2D schematic of device structure. (a) shared base and (b) raised base.

1.1.5 f_{max} and Device Scaling

As discussed above, through vertical scaling, f_T can be improved dramatically. R_B and C_{tc} can be reduced by lateral scaling and device structure innovations as well. Additionally, the shallow-trench-isolation (STI) and selective-implanted-collector (SIC) structures can also help reduce extrinsic CB capacitance. Recall (1.1) that a reduction

of C_{tc} is desired to improve f_{max}. Overall, f_{max} has been increased from 65 GHz to 300 GHz for modern SiGe HBTs.

1.2 SiGe HBT Design Examples

From a design point of view, we devide the real device into intrinsic and extrinsic parts. Generally, the intrinsic device, which performs the main device functions, determines the device performance to first order. The extrinsic device, which is essential to link the current flow to electrodes, introduces parasitic effects and hence degrades overall performance. Therefore, the goal is to achieve high performance through intrinsic designs and minimize the parasitic effects through extrinsic designs.

1.2.1 200 GHz SiGe HBTs

Many research groups have contributed to the optimization of SiGe HBTs and have published the designs for 200 GHz devices. Table 1.3 shows the characteristics of four examples. As discussed previously, the thin base width w_B is for the purpose of reducing τ_B. Correspondingly, peak N_B is increased to the 10^{19}/cm^3 – 10^{20}/cm^3 to keep the low base sheet resistance $R_{B,\square}$.

The high collector doping N_C is required to achieve higher f_T, and as discussed, $J_{C,peak}$ increases as well. Figure 1.7 shows the impact of N_C on $f_{T,peak}$ and $J_{C,peak}$ for a 200 GHz SiGe HBT design with $w_B = 10$ nm. When N_C increases from 5×10^{16}/cm^3 to 10^{18}/cm^3, peak f_T is improved from 110 GHz to 240 GHz, and $J_{C,peak}$ increases from 1 mA/μm^2 to 4.5 mA/μm^2 [35].

Table 1.3: Characteristics of manufacturable 200 GHz SiGe HBT Technology.

Reference	[31]	[32]	[33]	[34]
Lithographic node (μm)	0.15	0.18	0.2	0.12
f_T/f_{max} (GHZ)	190/100	206/184	166/210	207/285
β	–	450	2000	400
BV_{CEO} (V)	1.7	1.8	1.7	1.7
$J_{C,peak}$ (mA/μm^2)	6.6	8	8	8.3
Peak N_B (/cm^3)	2×10^{20}	4×10^{19}	3.5×10^{20}	4×10^{19}
w_B (nm)	2.5	25	1.2	20
Ge_{peak}	10%	25%	22%	25%
$R_{B,\square}$ (k$\Omega/\square$)	2.5	3.0	2.0	2.5

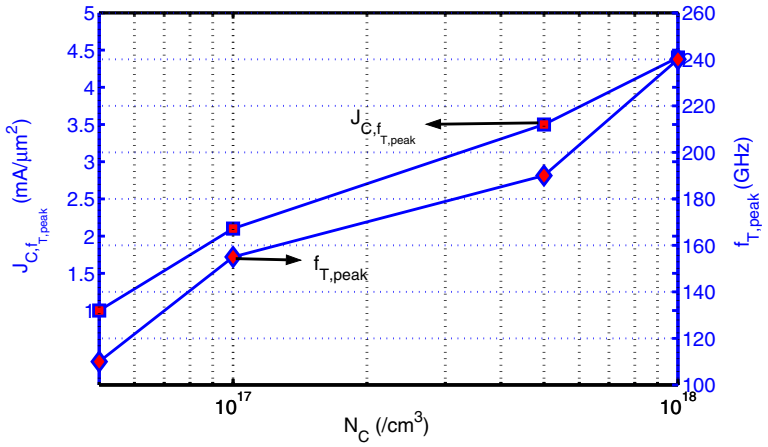


Figure 1.7: N_C impact on $J_{C,peak}$ and $f_{T,peak}$ for a 200 GHz SiGe HBT design with w_B = 10 nm [35].

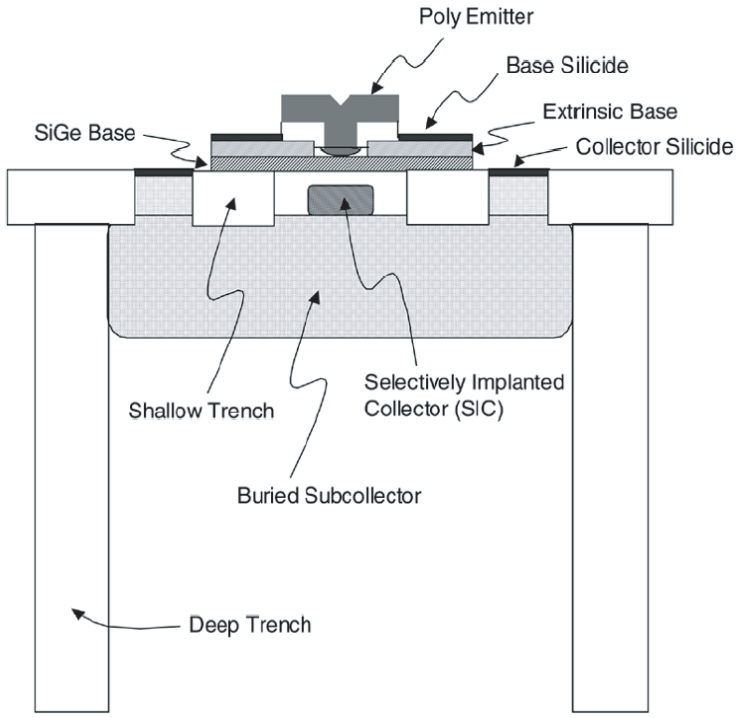


Figure 1.8: 2D schematic of a device structure used for 200 - 300 GHz SiGe HBTs [26].

1.2.2 Fabrication Process

The challenge of extrinsic device design is to minimize the parasitic effects and guarantee the high performance given by intrinsic designs. Figure 1.8 shows a 2D schematic of the device structure used for 200 – 300 GHz SiGe HBTs [26]. As shown, the double-polysilicon structure is used to make the emitter self-aligned to the extrinsic base. The SiGe base is grown non-selectively. A double-base contact structure is used, which reduces the base resistance by 1/4 compared to the single base contact structure. The extrinsic base is raised to the top of the intrinsic SiGe base. As a result, the CB

junction traps induced by p-type extrinsic base implantation are reduced. The base resistance is reduced as well, since the new structure allows a higher extrinsic base doping. The SIC is also used in order to reduce the extrinsic CB capacitance C_{tc}.

The structure shown in Figure 1.8 can be improved by using a selectively-grown epitaxial SiGe base. Figure 1.9 illustrates the process sequence. First, the polysilicon base is deposited and p^+ implantation is conducted. The pad oxide under the extrinsic base prevents the p^+ extrinsic base diffusion during the epitaxial base prebake and restricts the area so that the CB capacitance can be reduced. Another advantage is that the rapid-thermal-annealing (RTA) after implantation, which is a high temperature process, happens before base formation. This helps to maintain a heavily doped thin base. Figure 1.10 shows an SEM cross-section of a SiGe HBT with 220/230 GHz f_T/f_{max} [24], where the base was grown using selective SiGe epitaxy. In this work, we design the 2D structure based on those advanced features and fabrication capabilities.

1.3 Dissertation Contributions

The design and optimization of SiGe HBTs for high speed has been explored extensively in recent years. When the intrinsic vertical design is scaled down to the nanometer regime, there are new concerns and design issues. Especially for the extrinsic design, since now the intrinsic delay has now been reduced to 0.1 ps range, any comparable delay due to parasitic effects will sacrifice the ultra-high speed given by intrinsic design. Motivated by this challenge, we explore both the intrinsic 1D and extrinsic 2D design issues using MEDICI [37].

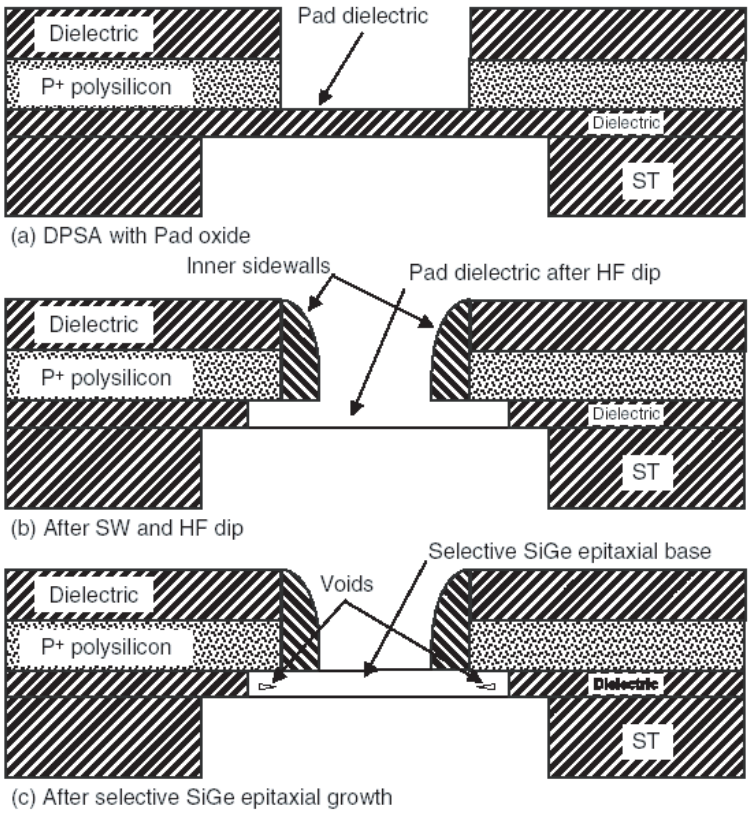


Figure 1.9: Selective SiGe epitaxial deposition [36].

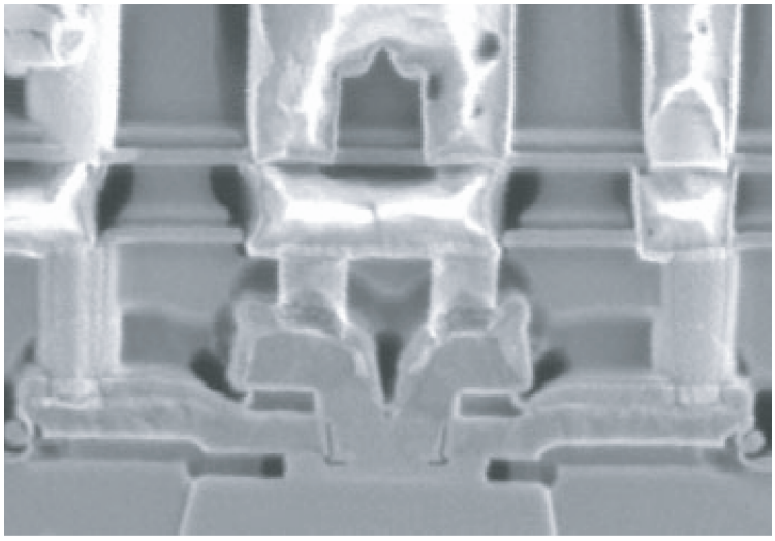


Figure 1.10: SEM picture of a 220/230 GHz f_T/f_{max} SiGe HBT [24].

In Chapter 2, we first run 1-D device simulations to achieve the initial designs efficiently. A nano-scaled SiGe HBT profile is obtained, which can achieve terahertz f_T. The spatial distribution of the total transit time is shown to be different from that in conventional devices. Extensive comparisons of the box and graded Ge profiles show that the graded profile leads to higher f_T, higher β and higher BV_{CEO} for the same total amount of Ge for nano-scale SiGe HBTs. A gradual n-collector to n^+ subcollector transition is shown to give higher f_T, BV_{CEO} and slower f_T roll-off.

The 2D design and parasitic effects are examined in Chapter 3. The delay contribution from each extrinsic device region is examined by using 2D regional transit time analysis. The device parasitic r_B and C_{tc} are extracted, and the dependence on lateral scaling is examined as well. 1.09 THz f_{max} is achieved when w_E is 60 nm. The overall

results would provide the guidelines for designing ultra fast SiGe HBTs with Terahertz f_T/f_{max}.

CHAPTER 2

1D PROFILE OPTIMIZATION

Device vertical scaling has been driven by the purpose of increasing cut-off frequency f_T. For the early technology, base transit time τ_B is the limitation of the speed. By reducing base width w_B, τ_B has been reduced significantly. Therefore, the collector space charge region (SCR) transit time, τ_{CSCR}, becomes the bottleneck, which makes the scaled SiGe HBT design more a collector design issue. A high collector doping N_C and a narrow collector epitexial layer w_C are desired to reduce τ_{CSCR}. When device dimension is reduced to nanometer scale, the non-equilibrium transport becomes significant due to both high electric field and rapid field change. Because of non-equilibrium transport, velocity overshoot occurs, and collector doping does not have to increase dramatically to achieve THz design. Hence, the device breakdown performance can be suitable for application requirements.

We first tune the physics models used in MEDICI, which can obtain reasonable device simulation results. Then, the design details for the vertical profile of THz f_T are addressed. The impact of Ge profile (box vs. graded) on THz device performance is compared at the same film stability. Since the collector design is more important for the THz design, we examine the impact of collector width w_C and n/n^+ transition on f_T and BV_{CEO}.

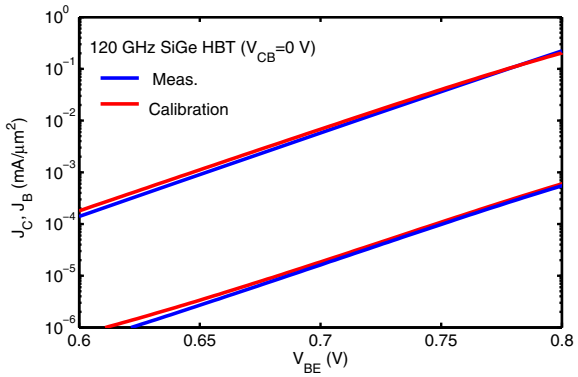
2.1 Physical Models

Before addressing design details and device characteristics analysis, we tune the physical models and their coefficients for device simulations. The validity is demonstrated by the simulations of a 120 GHz SiGe HBT.

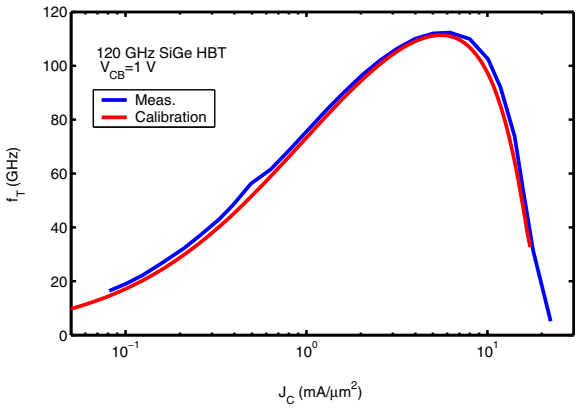
With device scaling, not only the high electric field, but also the rapid change of electric field over a short distance are present for scaled SiGe HBTs. As a result, the non-equilibrium carrier transport leads to high order phenomena, such as velocity overshoot. In order to model device function correctly, the energy balance (EB) equations are solved together with Poison's equation, and electron and hole continuity equations.

Table 2.1 lists the physical models and their Ge dependence. The effective density of states in SiGe is adjusted by fitting experimental I-V data of a 120 GHz HBT, as shown in Figure 2.1 (a). The affinity for SiGe is the same as for Si, meaning the Ge induced bandgap reduction is on the valence band completely. Slotboom BGN model is used to model the heavy doping induced bandgap narrowing [38, 39]. Due to the lack of experimental BGN data for SiGe, we assume here that the apparent BGN for SiGe is the same as for Si. Boltzmann statistics, as opposed to Fermi-dirac statistics, is used in order to ensure physically consistent modeling of minority carrier concentration, as discussed in [40]. The Philip's unified mobility model is used to be consistent with Slotboom's BGN model [41]. An electron relaxation time $\tau_n = 0.3$ ps is used. The simulated f_T can fit the measured f_T (2-D), as shown in Figure 2.1 (b).

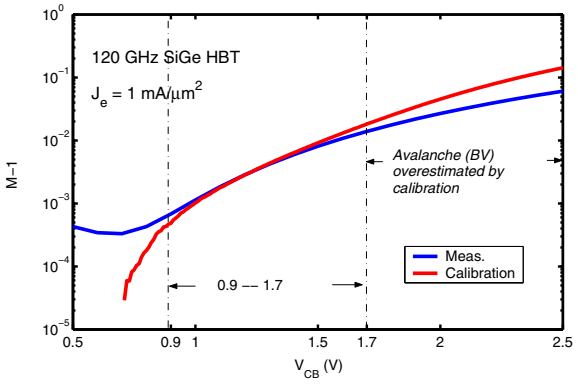
The calibration of avalanche multiplication factor M-1 over V_{CB} at fixed $J_E = 1$ mA/μm^2 is shown in Figure 2.1 (c). The model works well in the V_{CB} range of 0.9



(a)



(b)



(c)

Figure 2.1: Simulation of a 120 GHz SiGe HBT using tuned physical model coefficients. (a) Gummel, (b) $f_T - J_C$, and (c) $M - 1$ versus V_{CB}.

Table 2.1: Physical Models at T=300K

	Model	Ge Dependence (x)
Density of States	$N_{C,Si}$=2.89×10^{19}/cm^3	$N_{C,SiGe} = N_{C,Si}$
	$N_{V,Si}$=1.04×10^{19}/cm^3	$N_{V,SiGe} = 0.4 \times N_{V,Si}$
Bandgap	$E_{g,Si}$=1.12 eV	$E_{g,SiGe} = E_{g,Si}$−0.74x
Affinity	χ_{Si}=4.17 eV	$\chi_{SiGe} = \chi_{Si}$
BGN	Slotboom BGN	Same as Si
Statistics	Boltzmann Statistics	–
G/R	SRH and Auger	–
Mobility	Philips Unified Mobility model	Same as Si
Relaxation Time	τ_n=0.3 ps	Same as Si

V – 1.7 V, after which the avalanche effect is overestimated by simulations, and hence indicates that the simulated breakdown voltage is lower than measured. To the first order, we expect the simulated breakdown voltage to be reasonable for terahertz designs. Self-heating is more dependent on the 2D layout, and therefore is not included in the intrinsic device performance.

2.2 Intrinsic Design for THz f_T

Using the tuned physical models, we obtained a highly scaled 1D profile as shown in Figure 2.2 [42]. The metallurgical base width is 7 nm. the peak base doping is 8×10^{19}/cm^3. The base sheet resistance $R_\square$ is 8 kΩ/□. A retrograding collector profile is used for optimum high injection performance and breakdown voltage tradeoff. A 20 nm thick low-doped collector is used to reduce τ_C. Using this doping profile, both box

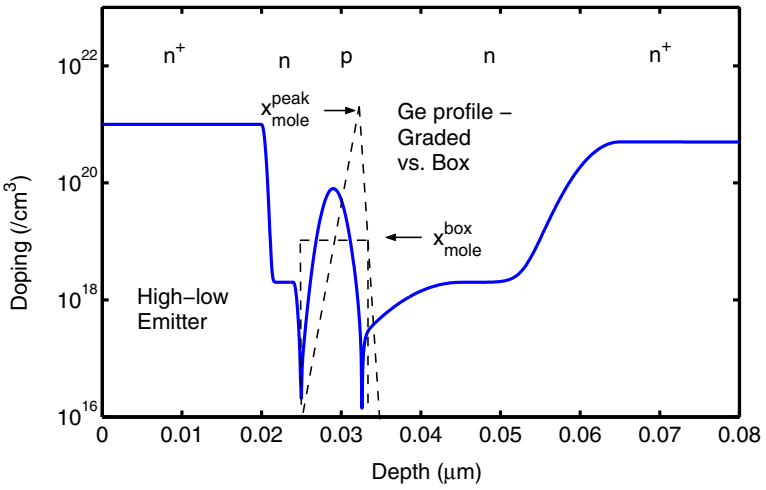


Figure 2.2: Nano-Scaled doping and Ge profile for Terahertz f_T. A low-doped emitter is used.

and graded Ge profile designs are examined at the same film stability. For the box Ge, x^{box}_{mole}=18%. For the graded Ge, x^{peak}_{mole}=27%.

Figure 2.3 shows the simulated $f_T - J_C$ curves. The Si-base design has 500 GHz peak f_T, the box and graded SiGe designs show 1.2 and 1.5 terahertz (THz) peak f_T, respectively. Next, we will address in detail the emitter, base and collector designs.

2.2.1 Emitter Design

In order to reduce emitter resistance R_E, a high emitter doping concentration of N_E = 10^{21}/cm^3 is used. A lightly doped emitter E_L, with $N_{El} = 3\times10^{18}$/cm^3, is sandwiched between the heavily doped emitter and base, as shown in Figure 2.2. The purpose is to reduce tunneling current and G/R current as well. To see the improvement, the base current density J_B is simulated for the EB profile with and without low-doped emitter,

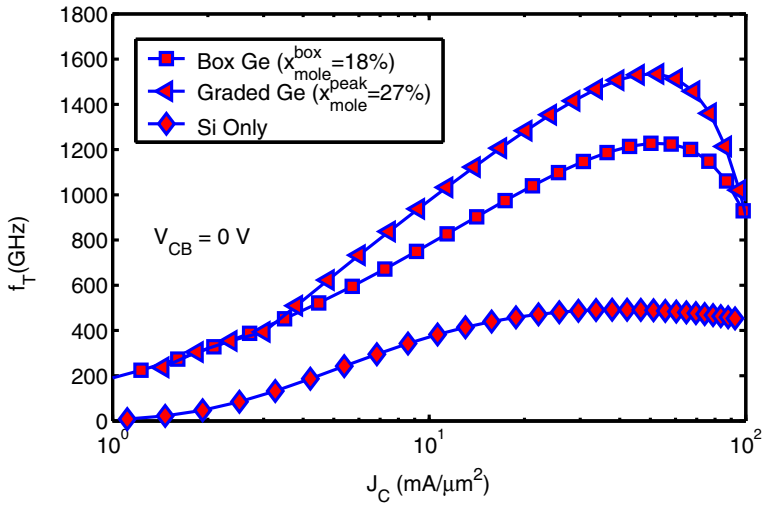


Figure 2.3: f_T - J_C curves for Si BJT, box Ge and Graded Ge designs. $V_{CB} = 0$ V. Total Ge are kept the same.

respectively. The box Ge profile is used in both. The results are compared in Figure 2.4. When the tunneling model is turned off, shown as the solid lines, the J_B for the high-low emitter design is more ideal at the low V_{BE} range of 0.3 - 0.5 V. The reason is that the doping concentration at EB junction is smaller when a low-doped emitter is used, which leads to a longer carrier lifetime and hence a less G/R current. When tunneling is included, shown as the dash lines, both J_B values are increased due to the EB junction tunneling currents. For the design without the high-low emitter, the tunneling is much more severe due to the heavier doping concentration at the junction.

Another advantage is that the low-doped emitter is depleted, and the EB SCR is located at this region. Therefore, the base width modulation due to the change of V_{BE} is less. Figure 2.5 (a) and (b) compare the β between two emitter structures for both box

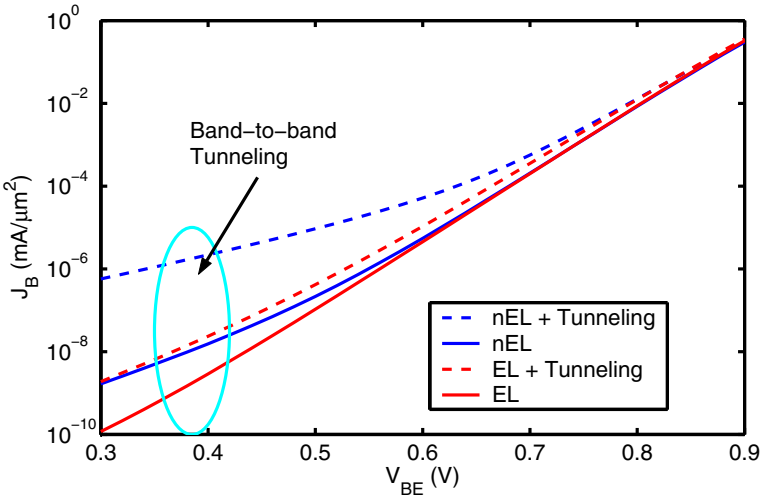
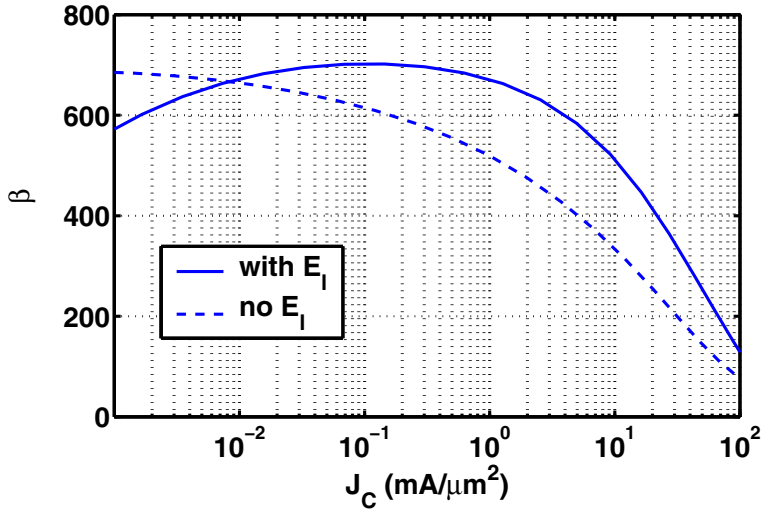


Figure 2.4: Comparisons of tunneling for two emitter designs. Box Ge profile is used.

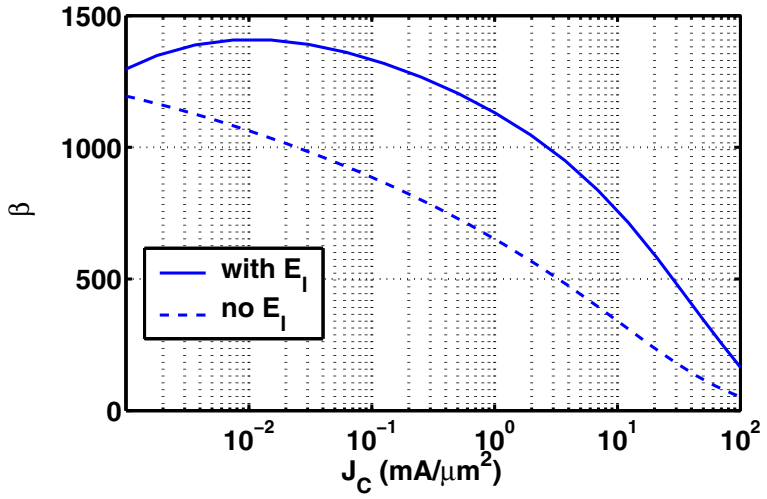
and graded Ge profiles. The low-emitter structure has a more linear β curve. The box Ge profile is better than the graded Ge profile, due to less Ge ramp effect. However, the process of forming a thin layer of low-doped emitter using *in-situ* growth is difficult and expensive. Usually, phosphorus is used, which features a low resistivity compared with arsenic.

2.2.2 Base Design

To mimic the profile after fabrication, a non-uniformly doped base profile is used, with the peak base doping as high as 8×10^{19}/cm^3. In order to reduce the base transit time τ_B dramatically, the base width is scaled as narrow as 7 nm. The fabrication innovations, such as "HCI-free LPCVD", have made it possible to grow a high quality thin Boron layer with doping level in the 10^{20}/cm^3 range [31]. The base sheet resistance,



(a)



(b)

Figure 2.5: Comparisons of β for two emitter structures, with and without low-doped emitter. (a) Box Ge profile and (b) Graded Ge profile.

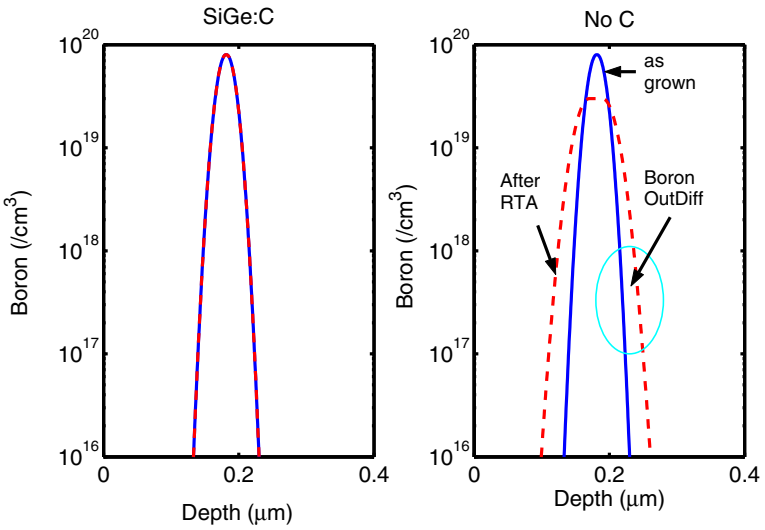


Figure 2.6: Comparisons of the Boron out-diffusion effects between SiGe and SiGe:C.

$R_{\square}$, is 8 kΩ/□, which is higher than the $R_{\square}$ of 2.5 – 3 kΩ/□ for 200 GHz technology. A main challenge of maintaining such a heavily doped thin base is the so called base out-diffusion during the following high temperature process. Benefitting from the fabrication technology improvement, Carbon has been added into the SiGe layer to eliminate the base out-diffusion and guarantee that the nano-scaled base design will be manufacturable [43, 44]. Figure 2.6 illustrates the benefit of using SiGe:C. As discussed in [44], usually a Carbon concentration of 10^{20}/cm^3 is required to achieve substantial suppression of the Boron out-diffusion, while the active Boron concentration is not impacted.

2.2.3 Collector Design

The collector design is very involved, which determines collector transit time τ_C, the onset of Kirk effect and hence J_C at peak f_T ($J_{C,peak}$), and the breakdown voltage. Collector doping N_C impact on transit time τ_C has been examined extensively [35] [45]. Using first order approximation, as addressed in [28],

$$\begin{aligned} \tau_C &= \tau_{CSCR} + \tau_{tc} \\ &\propto \frac{1}{N_C}, \end{aligned} \tag{2.1}$$

where τ_{CSCR} is the collector SCR transit time and τ_{tc} is the CB depletion capacitance charging time. Intuitively, N_C should be increased with scaling to reduce τ_C. In this work, a retrograding collector profile is used for optimum high injection performance and breakdown voltage tradeoff. A 20 nm thick moderately doped collector is used to reduce τ_C [46]. The narrow collector is fully depleted, and thus the n/n^+ transition has impacts on f_T and BV_{CEO}, as will be addressed later.

For the nano-scale base design, the base must be designed to avoid base punch-through. Besides, the breakdown voltage BV_{CEO} should be practical. To examine these issues, a simple 2D simulation is performed. Figure 2.7 shows the simulated output curves for a graded Ge profile design with $x_{mole}^{peak} = 5\%$, and $I_B = 50$ μA to 200 μA, in 50 μA step. The device shows no base punch-through, and a 1.65 V BV_{CEO} is observed. The measured BV_{CEO} for a 350 GHz f_T SiGe HBT is 1.4 V [47]. A higher BV_{CEO} in this design is due to the lower β. In this design β is 720, while in [47], β is 3500.

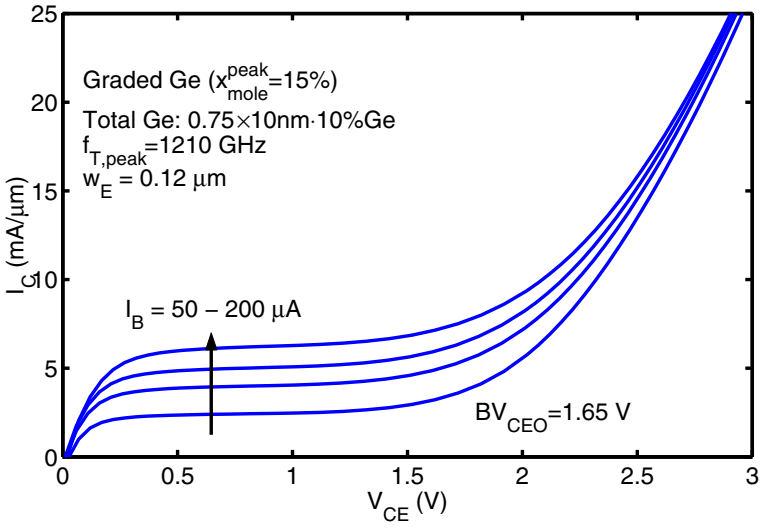


Figure 2.7: Output curves of the nano-scale SiGe HBT with a x_{mole}^{peak}=5% graded Ge profile.

2.3 1D Parasitic Analysis

Recall that f_T is inversely proportional to the overall transit time as

$$\frac{1}{2\pi f_T} = \sum \tau_n. \tag{2.2}$$

To probe device internal delay, a regional analysis of transit time is conducted to separate and extract each τ component. The transit delay can also be modeled using the distributive capacitance, which can be extracted using

$$C = \tau \times g_m. \tag{2.3}$$

g_m is the transconductance and can be extracted using the real part of y_{21} at low frequency. The capacitance associated with the junction depletion region is called the depletion capacitance. The capacitance associated with the excess minority carrier storage is called the diffusion capacitance.

2.3.1 1D Transit Time

Transit time analysis is widely used in understanding the spatial distribution of transit time $\Delta\tau_n$ and $\Delta\tau_p$, which are defined as,

$$\Delta\tau_n = q\Delta n/\Delta J_C, \tag{2.4}$$

$$\Delta\tau_p = q\Delta p/\Delta J_C. \tag{2.5}$$

Δn, Δp and ΔJ_C are electron, hole and conduction current density changes for a small *ac* excitation at the base, when V_{CE} is fixed [48]. For traditional designs with wider w_B, there is always a region in the base where $\Delta\tau_n$=$\Delta\tau_p$, as can be seen in Figure 2.8 for a 50 GHz f_T SiGe HBT [48]. However, for the nano-scale base design, such a region of $\Delta\tau_n$=$\Delta\tau_p$ does not exist, and $\Delta\tau_n \ll \Delta\tau_p$ in the whole base, as shown in Figure 2.9 for the Si design. Due to charge conservation, the total electron and hole charge changes must be the same, meaning $\int_0^l \Delta\tau_n dx = \int_0^l \Delta\tau_p dx$. Therefore, as shown in Figure 2.8, the $\Delta\tau_p$ integration area A in the base must be equal to the shaded area B in the collector. With scaling, the integration of $\Delta\tau_n$ in the p-type base is much smaller than in the n-type collector. As a result, $\Delta\tau_p$, which is related to area A, must be much larger than

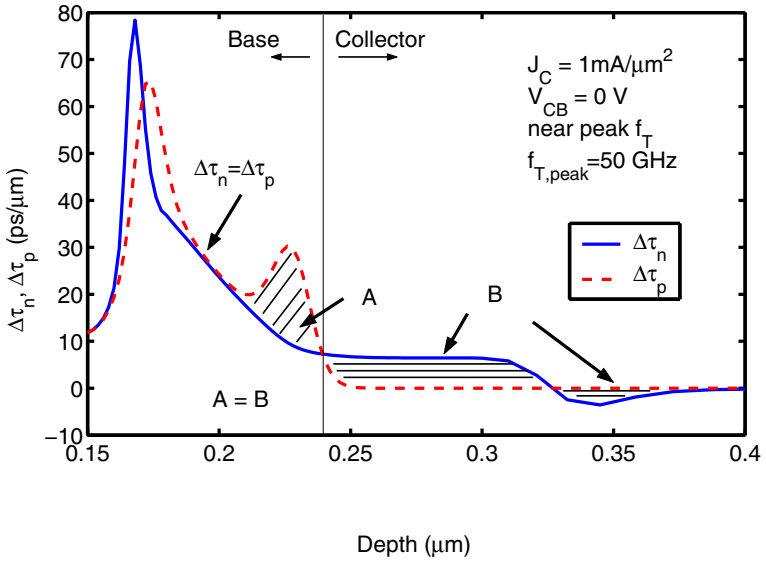


Figure 2.8: The $\Delta\tau_n$ and $\Delta\tau_p$ distributions for a 50 GHz SiGe HBT design [48].

$\Delta\tau_n$ in the base to balance the greater $\Delta\tau_n$ distribution in the collector, and hence the $\Delta\tau_n$=$\Delta\tau_p$ region disappears at nano-scale base width. Figure 2.9 also shows clearly that the emitter transit time severely limits f_T in the Si design. An EB junction high injection barrier due to heavy base doping induced BGN is observed.

Consequently, the boundaries used for traditional regional transit time analysis [49] cannot be defined for THz designs. In [50] the electrical junctions x^*_{EB} and x^*_{CB}, where $\Delta\tau_n$ and $\Delta\tau_p$ are crossed in junctions, are used as the boundaries to define each transit time components . The base transit time is

$$\tau^*_B = \int_{x^*_{EB}}^{x^*_{CB}} \Delta\tau_n|_{V_{CB}} dx. \tag{2.6}$$

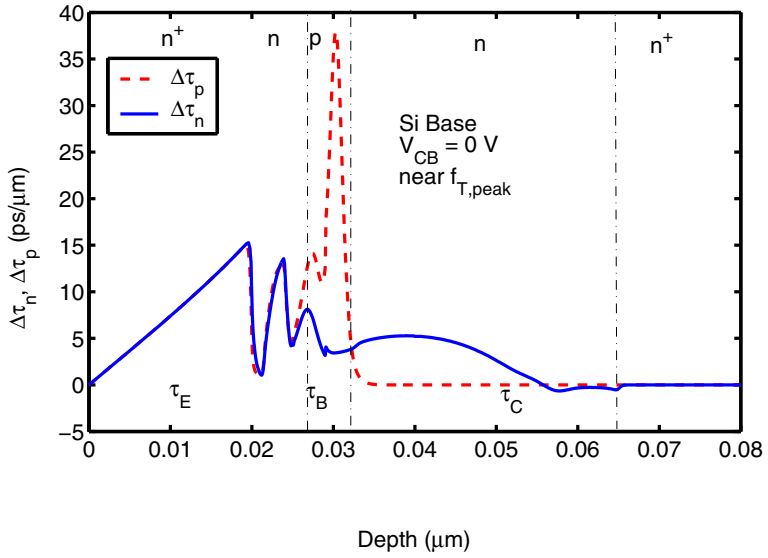


Figure 2.9: The $\Delta\tau_n$ and $\Delta\tau_p$ distributions for Si BJT design.

The τ_B^* consists of neutral base transit time τ_B and part of neutral EB SCR transit time. Since τ_B is very small for THz SiGe HBT designs, the neutral transit time may take an appreciable amount in τ_B^*. Therefore, in this work, we use the peak $\Delta\tau_n$ and $\Delta\tau_p$ positions as the depletion boundary to distinguish emitter, EB SCR, base and collector region, as shown in Figure 2.10.

x_{EE}, x_{EB} and x_{CB} are defined to distinguish each transit time components, where

1. x_{EE} is the peak $\Delta\tau_n$ position at the high–low emitter interface;

2. x_{EB} is the peak $\Delta\tau_p$ position at the emitter side of the base;

3. x_{CB} is the peak $\Delta\tau_p$ position at the collector side of the base.

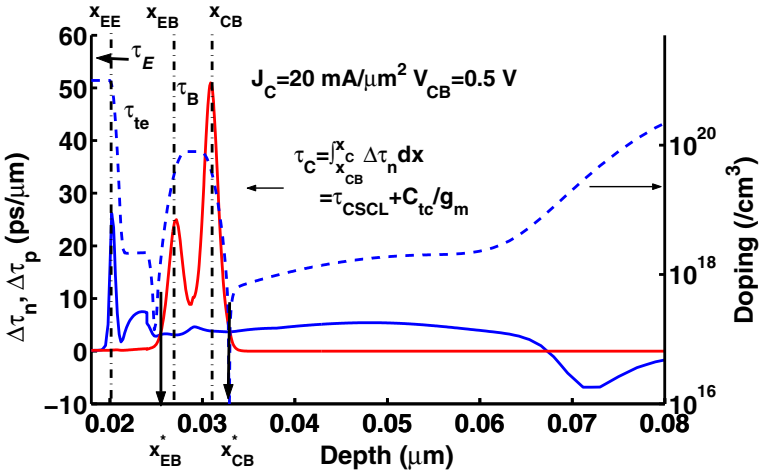


Figure 2.10: τ_E, τ_{EB}, τ_B, and τ_C definitions

Consequently, we divide emitter–collector transit time τ_{EC} into emitter transit time τ_E, EB SCR transit time τ_{te}, base transit time τ_B and collector transit time τ_C.

$$\tau_E = \int_0^{x_{EE}} \Delta\tau_n|_{V_{CE}} dx, \tag{2.7}$$

$$\tau_{te} = \int_{x_{EE}}^{x_{EB}} \Delta\tau_n|_{V_{CE}} dx, \tag{2.8}$$

$$\tau_B = \int_{x_{EB}}^{x_{CB}} \Delta\tau_n|_{V_{CE}} dx, \tag{2.9}$$

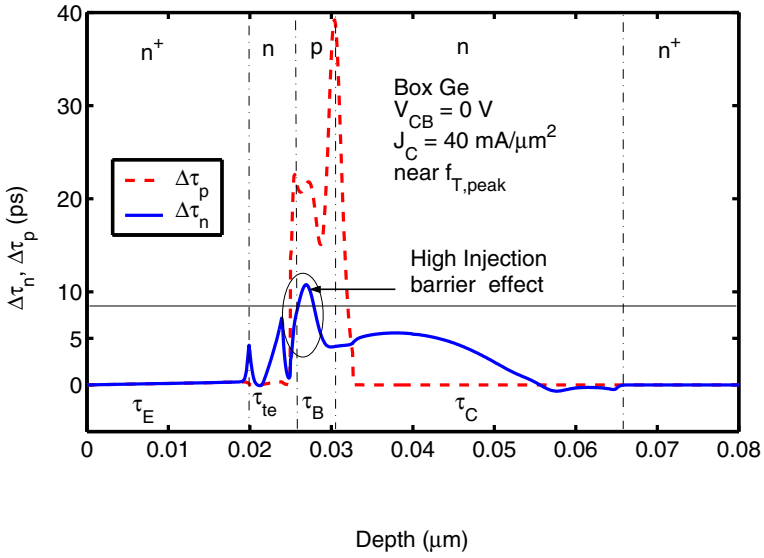
$$\tau_C = \int_{x_{CB}}^{x_C} \Delta\tau_n|_{V_{CE}} dx = \tau_{CSCR} + \tau_{tc}. \tag{2.10}$$

Figure 2.11 (a) and (b) show the $\Delta\tau_n$, $\Delta\tau_p$ distribution near peak f_T for the box and graded SiGe HBTs. Compared with Si, τ_E is reduced dramatically, due to high β, which is responsible for more than 500 GHz f_T increase over Si. However, because of

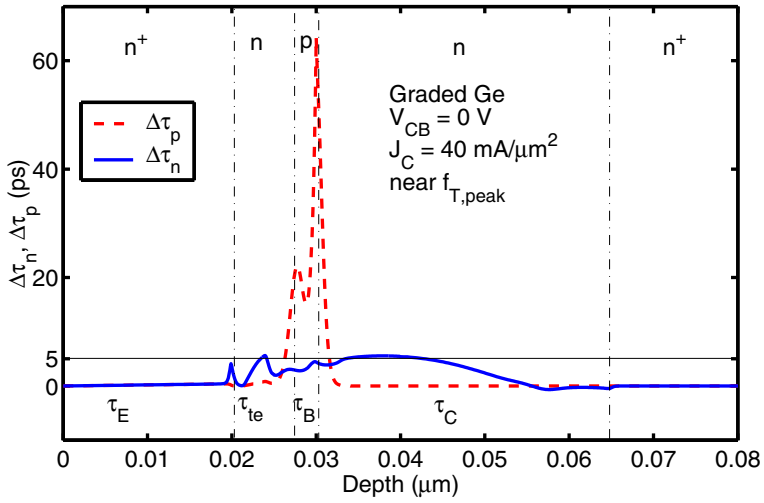
the higher barrier for holes in the base of the SiGe HBTs, $\Delta\tau_n > \Delta\tau_p$ in the low doped emitter, while $\Delta\tau_n = \Delta\tau_p$ for Si.

The J_C dependence of each τ component for Si BJT is shown in Figure 2.12 (a), in comparison with a $x_{mole}^{peak} = 27\%$ graded SiGe HBT. For Si BJT, τ_E and τ_C dominate. The advantage of a SiGe base is primarily more than 20× τ_E reduction. However, both EB and CB junction high injection barrier effects are more severe in the SiGe HBT, which leads to the τ_B increase at high J_C. No clear τ_B increase is observed for Si design. Figure 2.12 (b) compares J_C dependence of τ components between the graded and box profiles. τ_B is two times larger in the box profile. τ_C is dominant in both and hence the τ_{EC} curves follow the τ_C curves. When Kirk effect occurs at high J_C, τ_C increases more dramatically than τ_B, and is the main reason for f_T roll-off.

To examine the high injection barrier effect for box and graded Ge profiles, we simulate the electron, hole, Δn and $\Delta\tau_n$ distributions with increasing J_C. Figure 2.13 (a) and (b) show the electron and hole distributions, together with doping for box Ge profile. When J_C increases from 14 mA/μm^2 to 100 mA/μm^2, the electron storage in the base increase. The hole concentration at the EB junction side of the base is higher than the background p-type doping, which is due to the Ge induced band barrier. Figure 2.14 (a) shows the Δn distributions. Most Δn occurs at the EB side of the p-type base and the CB junction barrier effect gets more severe with increasing J_C. The $\Delta\tau_n$ curves are compared in Figure 2.14 (b), which shows clearly that for the box Ge profile, most τ_B increase is due to the EB junction barrier induced electron storage increase.

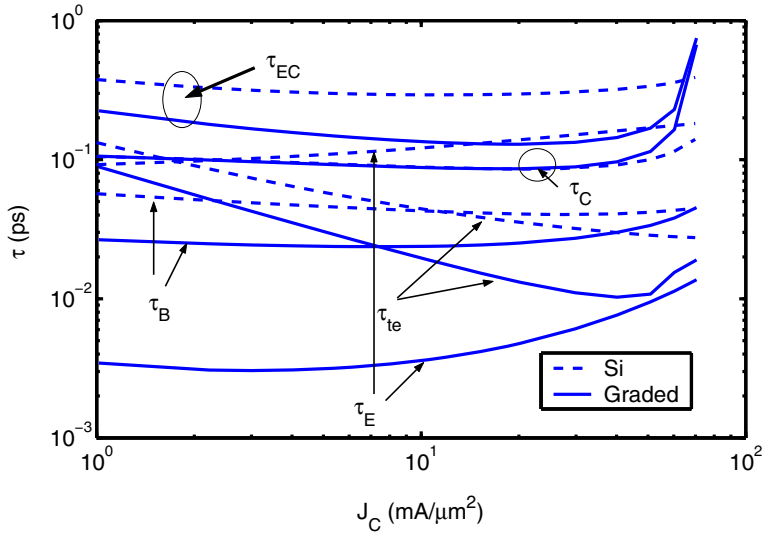


(a)

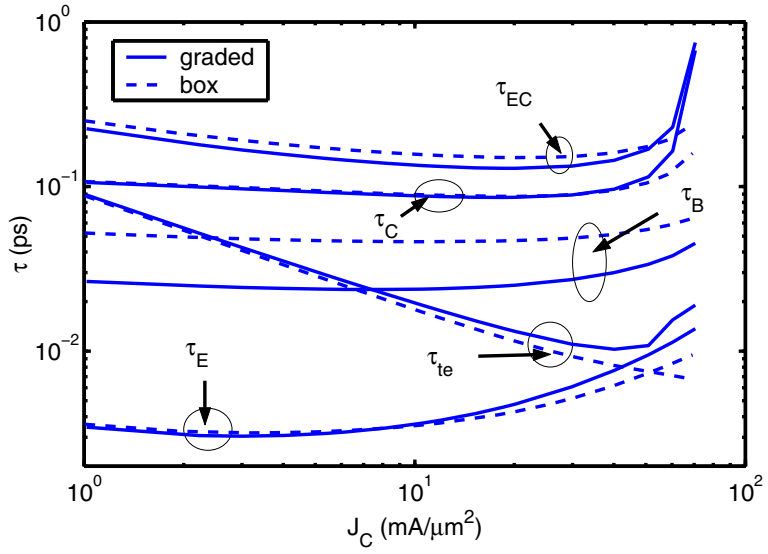


(b)

Figure 2.11: The $\Delta\tau_n$ and $\Delta\tau_p$ distributions for (a) box Ge design and (b) graded Ge.

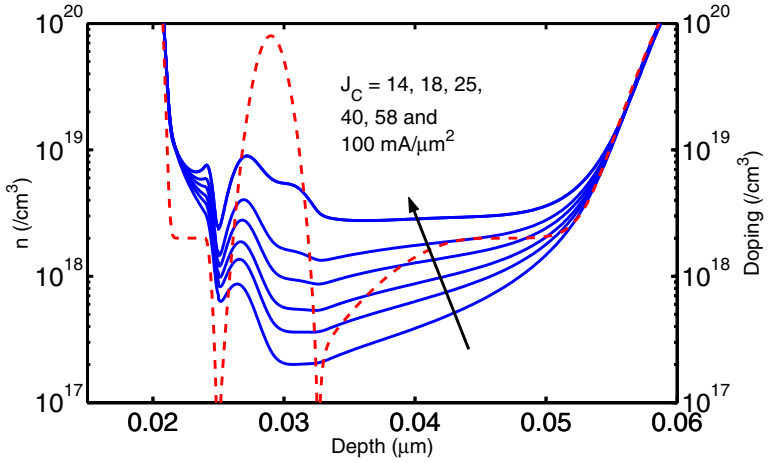


(a)

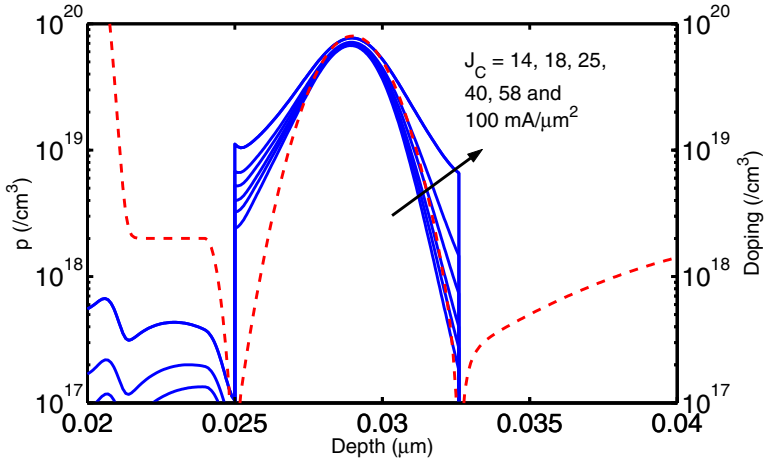


(b)

Figure 2.12: (a) Comparisons of τ components versus J_C characteristics between Si BJT and graded SiGe HBT designs. (b) Comparisons of τ components versus J_C characteristics between graded Ge and box Ge designs.

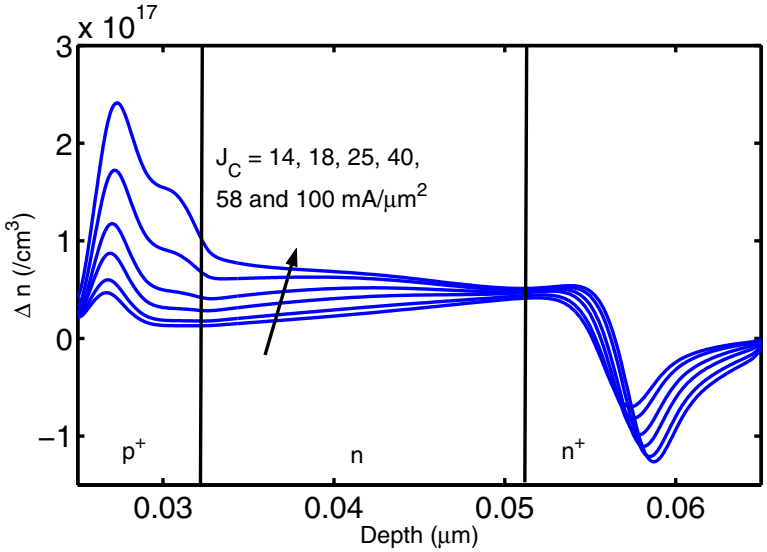


(a)

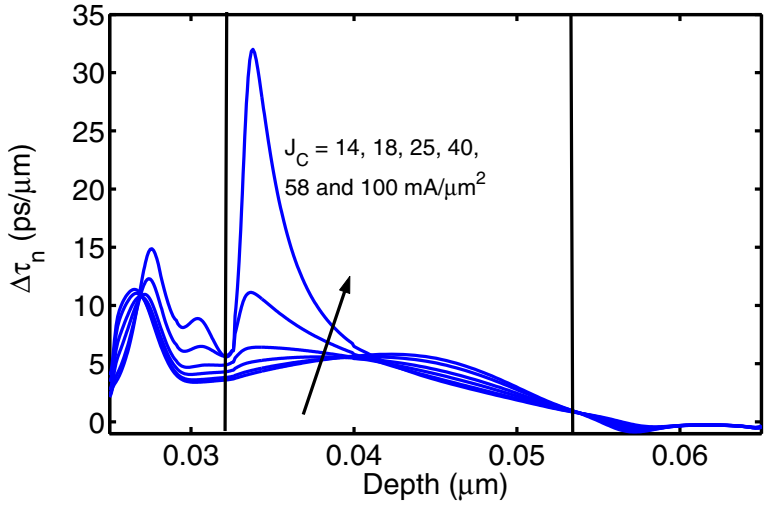


(b)

Figure 2.13: Barrier effect for Box Ge profile. (a) Electron concentration with increasing J_C and (b) Hole concentration with increasing J_C.

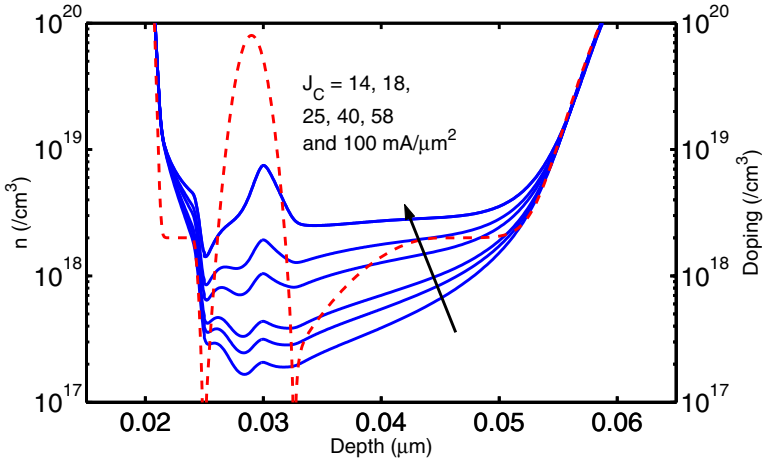


(a)

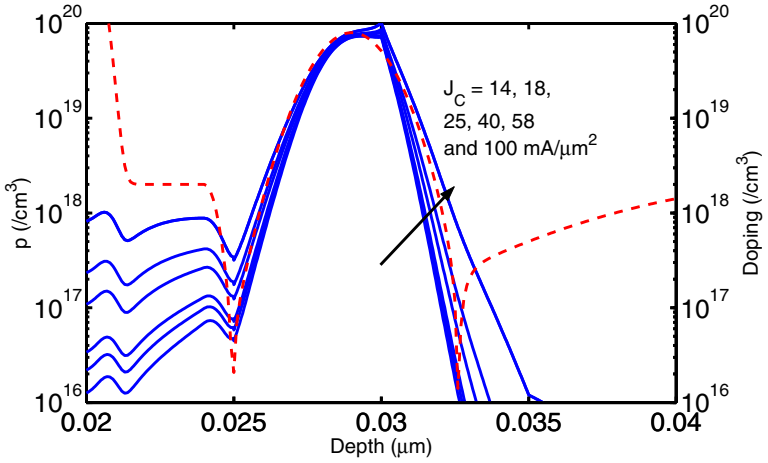


(b)

Figure 2.14: Barrier effect for Box Ge profile. (a) Δn distribution with increasing J_C and (b) $\Delta\tau_n$ distribution with increasing J_C.

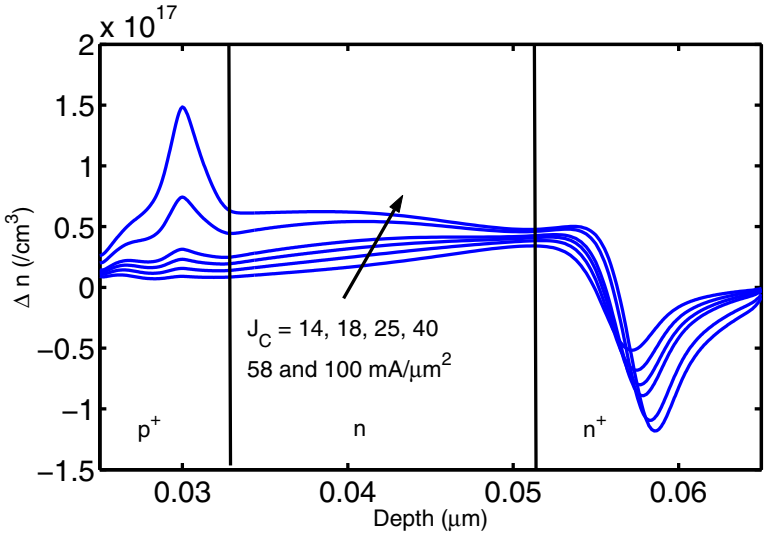


(a)

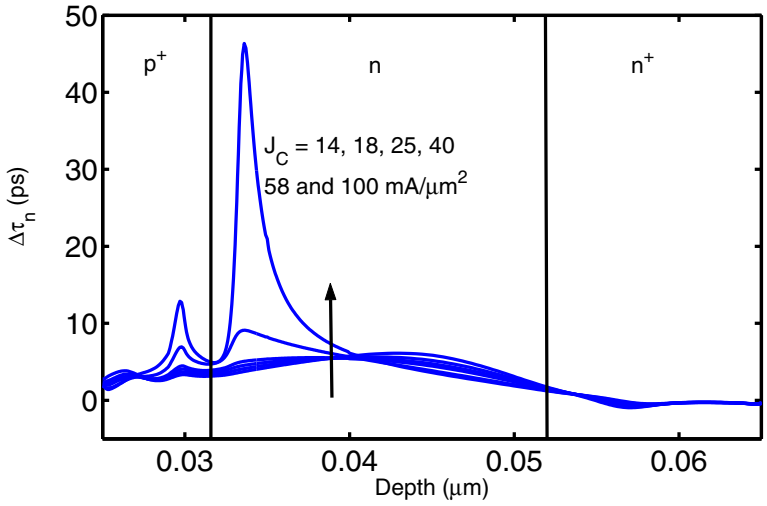


(b)

Figure 2.15: Barrier effect for Graded Ge profile. (a) Electron concentration with increasing J_C and (b) Hole concentration with increasing J_C.



(a)



(b)

Figure 2.16: Barrier effect for Graded Ge profile. (a) Δn distribution with increasing J_C and (b) $\Delta\tau_n$ distribution with increasing J_C.

Figure 2.15 (a) and (b) compare the electron and hole concentrations with increasing J_C for the graded Ge profile. Because the Ge at the EB and CB junction for the graded profile is much smaller than for the box profile, the barrier for holes is smaller. Therefore, the hole concentration near junctions are smaller for graded Ge. Figure 2.16 (a) and (b) compare the Δn and $\Delta\tau_n$ distributions for graded Ge. With increasing J_C, CB junction high injection barrier is much more severe, which leads to an increase of Δn and $\Delta\tau_n$ at the CB side of the base. We conclude that while the EB junction high injection barrier effect is responsible for the τ_B increase for the box profile, the CB junction high injection barrier effect is responsible for the τ_B increase for the graded profile.

2.3.2 Distributive Capacitance

Recall (2.3), an alternative way to analyze different delay times is to calculate the capacitance corresponding to each transit time as $C = g_m\tau$. This helps to associate distributive analysis with traditional lumped element equivalent circuit analysis. For 1D intrinsic design, we have EB depletion capacitance C_{te}, CB depletion capacitance C_{tc}, emitter diffusion capacitance C_{dE}, base diffusion capacitance C_{dB}, and collector SCR

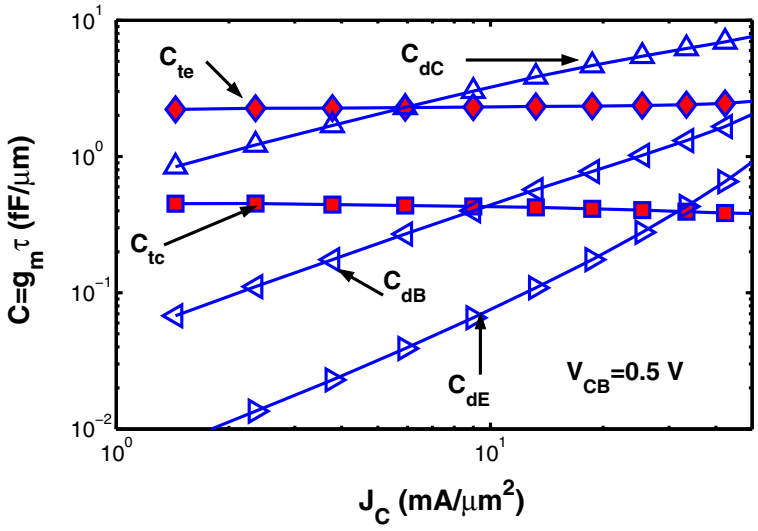


Figure 2.17: C_{te}, C_{tc}, C_{dE}, C_{dB} and C_{dC} versus J_C at $V_{CB} = 0.5$ V.

diffusion capacitance C_{dC}.

$$C_{te} = g_m \times \tau_{te}, \qquad \text{(depletion capacitance)} \qquad (2.11)$$

$$C_{tc} = g_m \times \tau_{tc}, \qquad \text{(depletion capacitance)} \qquad (2.12)$$

$$C_{dE} = g_m \times \tau_E, \qquad \text{(diffusion capacitance)} \qquad (2.13)$$

$$C_{dB} = g_m \times \tau_B, \qquad \text{(diffusion capacitance)} \qquad (2.14)$$

$$C_{dC} = g_m \times \tau_{CSCR}, \qquad \text{(diffusion capacitance)} \qquad (2.15)$$

At each J_C, the transconductance g_m of a SiGe HBT is extracted using the real part of y_{21} at low frequency f [51]. Figure 2.17 compares capacitance versus J_C curves for each C defined above for the THz design. The junction depletion capacitance is

equivalent to a parallel-plate capacitor of separation w_d,

$$C_t = \frac{\epsilon_{si}}{w_d}, \tag{2.16}$$

where w_d is the depletion width. C_t is independent of J_C. C_{tc} is the CB junction depletion capacitance. Notice that the C_{te} extracted in this work actually includes the EB junction neutral transit time τ_n [9]. However, since the junction depletion capacitance dominates, as shown, we did not extract τ_n. The diffusion capacitance is defined as

$$\begin{aligned} C_d &= \frac{dQ_{excess}}{dV_{BE}} = \frac{dI\tau}{dV_{BE}} \\ &= \frac{I}{V_{th}}\tau \end{aligned} \tag{2.17}$$

For emitter, the "I" is the base current, which can be expressed as "I_C/β". For base and collector SCR, "I" is the collector current. τ is the τ_E, τ_B and τ_{CSCR} respectively. Since as shown in Figure 2.12, τ_E, τ_B and τ_{CSCR} are constant from low to medium injection, the diffusion capacitance C_{dE}, C_{dB} and C_{dC} should be a linear function of J_C, which is correctly shown in Figure 2.17.

2.4 Design Issues

Next, we examine how Ge profile impacts device performance. Since the collector design primarily determines the τ_C, the collector width and the collector-subcollector

transition impact on transistor performance are also examined. Finally, the base optimization issue is addressed for the purpose of reducing base sheet resistance $R_{B,\square}$, which impacts both RF noise performance and f_{max}.

2.4.1 Impact of Ge

Discussions on the choice between the box and graded profiles have been addressed in many works, often using first order theories assuming uniform doping and drift-diffusion transport [52]- [54]. We revisit this issue here for nano-scale designs. To make a meaningful comparison, we compare the box and graded profiles at the same total Ge, so that the SiGe film stability is kept the same. We use the unit of "10nm·10%Ge" to specify the amount of total Ge.

At the SiGe film thickness of 7 nm, the maximum Ge that can be used to maintain the strained-layer is 2.5 "10nm·10%Ge" [48]. In this work, we examined the device performance up to the total Ge of 1.5 "10nm·10%Ge", where the film is under thermodynamic stability. Figure 2.18 compares f_T and β between the graded and box Ge profiles. For our doping profile design, the graded Ge profile shows a higher β than the box profile, which is the opposite of the prediction by first order analytical theories. An inspection of simulation details and additional simulations with different options show that this is related to the non-uniform base doping as well as the very thin base thickness. Figure 2.19 shows $f_{T,peak}$ versus β for the graded and box Ge profile. For the same β, the graded profile gives higher $f_{T,peak}$. Besides, as indicated in Figure 2.18, for the same performance, the graded profile uses less Ge and thus is more stable. For the same peak

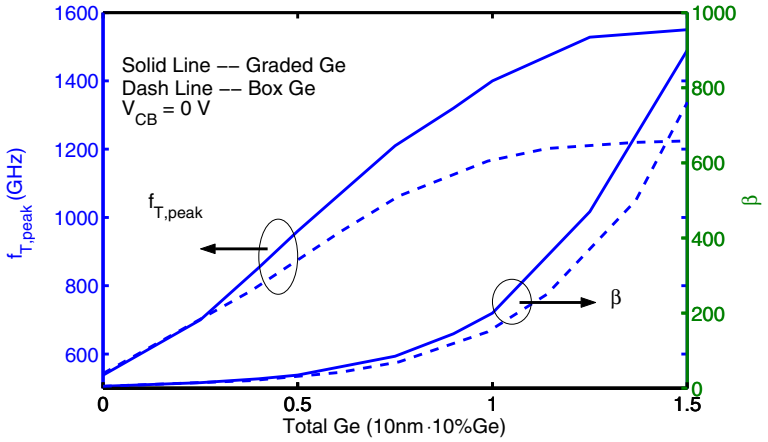


Figure 2.18: Comparisons of the $f_{T,peak}$ and β versus total Ge for box and graded Ge profiles

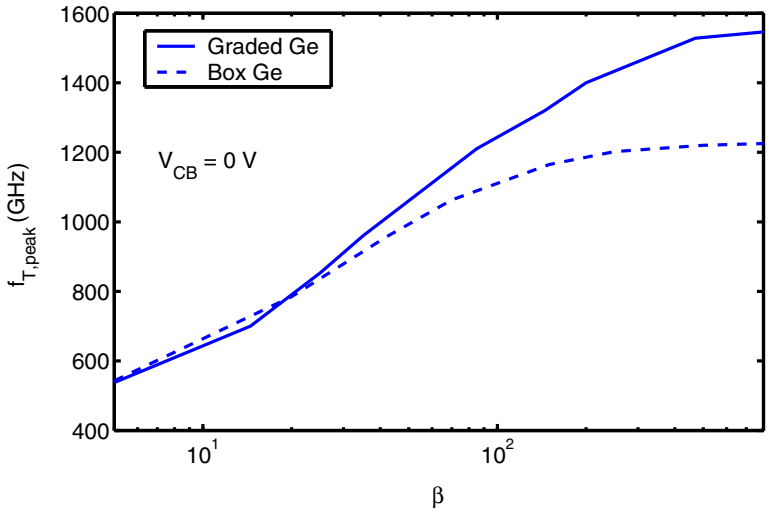


Figure 2.19: $f_{T,peak}$ versus β for box and graded Ge profiles

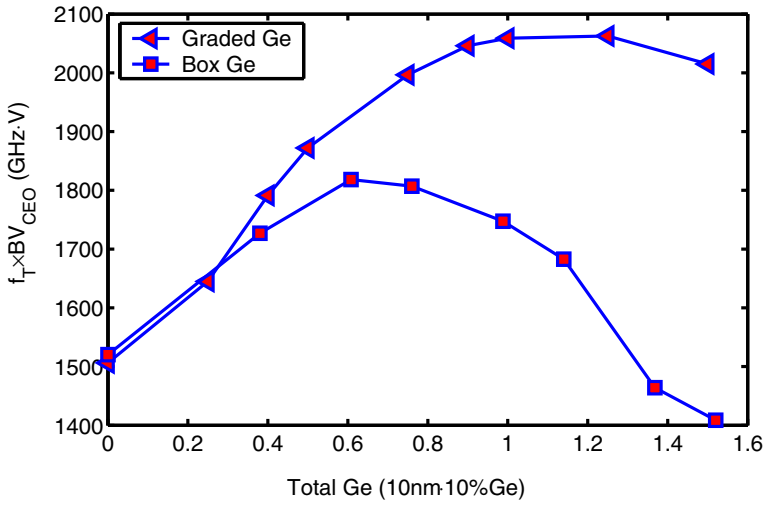
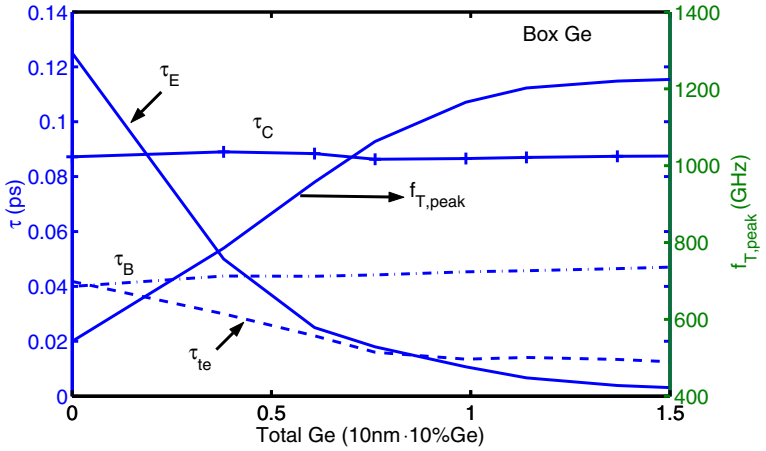


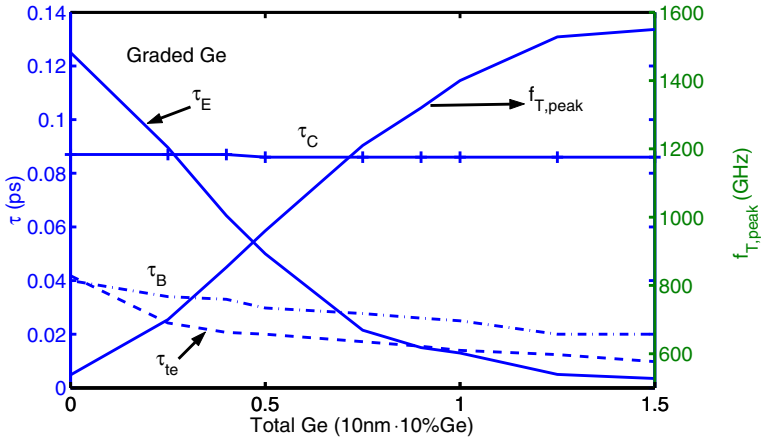
Figure 2.20: Comparison of $f_T \times BV_{CEO}$ versus total Ge between box and graded Ge profiles.

f_T, the graded profile has lower β, indicating a higher BV_{CEO}. Therefore, the graded Ge profile is preferred for higher device performance, as well as better stability. Figure 2.20 shows the $f_T \times BV_{CEO}$ product as a function of total Ge. The graded profile has a higher $f_T \times BV_{CEO}$ than the box profile, particularly at higher total Ge (and hence higher f_T). A 2000 GHz·V $f_T \times BV_{CEO}$ is achieved using the graded profile. The results are the opposite of the prediction by first order analytical theories. The reason is related to the non-uniform base doping as well as the very thin base thickness.

The Ge content dependence of τ components and peak f_T are shown in Figure 2.21 (a) for the box profile, and in Figure 2.21 (b) for the graded profile. τ_C is the same for both box and graded Ge profile, regardless of Ge content. τ_E is reduced dramatically with increasing Ge, which corresponds to a rapid $f_{T,peak}$ improvement. However τ_E becomes the smallest component when total Ge is above 10nm·10%Ge. Further increasing



(a)



(b)

Figure 2.21: τ_E, τ_{te}, τ_B, τ_C and f_T versus total Ge for (a) box Ge and (b) graded Ge.

Ge has no dramatic impact on reducing τ_{EC}. A major difference between the box and graded profiles is the base transit time τ_B. When total Ge is increased from 0 to 1.5 $\times$10nm·10%Ge, τ_B is increased by 0.007 ps in box profile design due to increasing EB barrier height, and τ_B is reduced by 0.015 ps in graded profile design due to increasing base accelerating field.

2.4.2 Impact of Collector Design

To examine the w_C impact, we fix emitter, base and, particularly, CB junction profiles, and reduce w_C from 200 to 150, 100, 50, and 30 nm, while the same n/n^+ transition is kept. Using the 1D transit time analysis, the $\Delta\tau_n$ and $\Delta\tau_p$ are extracted again to distinguish τ_{CSCR} and τ_{Ctc}. But Δv is at the emitter electrode, while V_{CB} is fixed. Because now the electron change in collector SCR is only due to injection, CB junction depletion capacitance C_{tc} charging time is not included in the collector $\Delta\tau_n$ integration. Therefore

$$\tau_{CSCR} = \tau_C|_{V_{CB}} \tag{2.18}$$

$$\tau_{tc} = \tau_C|_{V_{CE}} - \tau_C|_{V_{CB}}. \tag{2.19}$$

Applying the 1D transit time analysis mentioned above, τ_{CSCR} and τ_{tc} are calculated for each w_C at the same J_C of 20 mA/μm^2. Besides, τ_E, τ_B and τ_{te} are also extracted. Figure 2.22 compares the results. τ_{EC} is dominated by τ_{CSCR}, which decreases with decreasing w_C. Figure 2.23 compares $f_T - J_C$ curves. Mainly because of τ_{CSCR} reduction, a higher f_T is achieved at the same J_C for the design with smaller w_C.

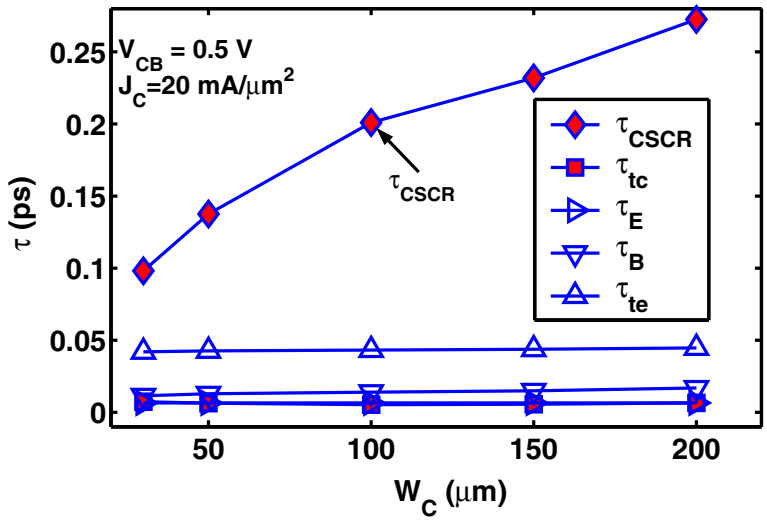


Figure 2.22: w_C impact on transit times components.

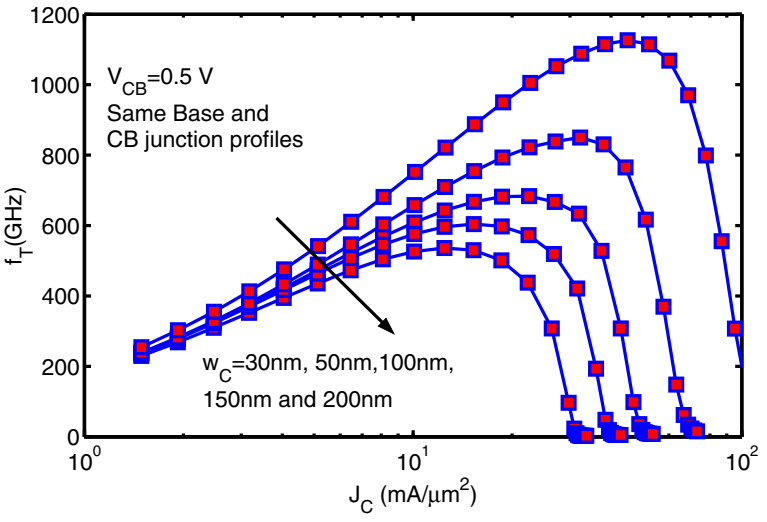
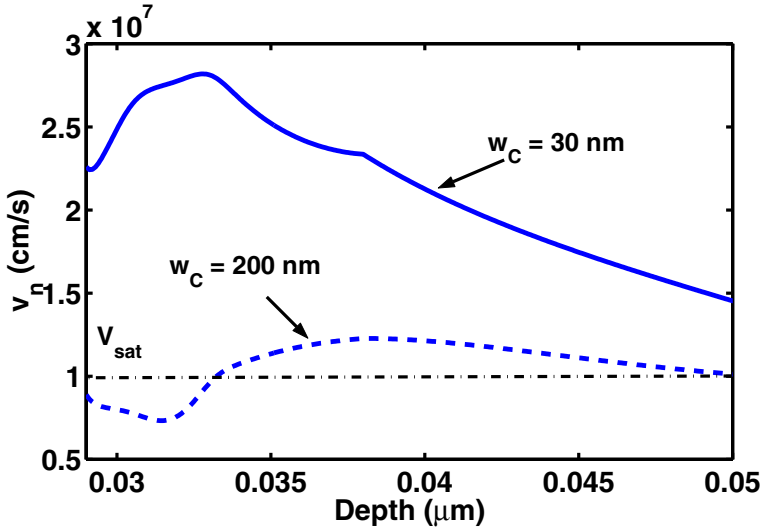


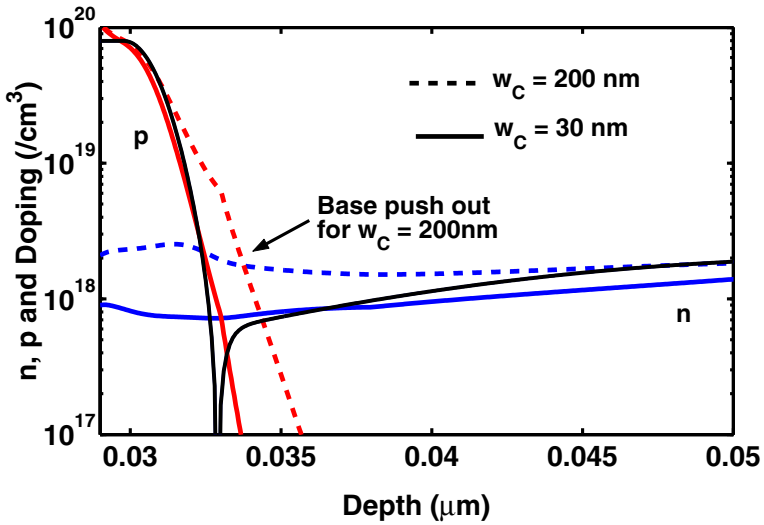
Figure 2.23: f_T-J_C curves for w_C = 200 nm, 150 nm, 100 nm, 50 nm, and 30 nm. V_{CB} = 0.5 V.

Interestingly, the onset of Kirk effect is delayed with decreasing w_C, as indicated by a higher $J_{C,peak}$ for smaller w_C. This is a result of higher electric field in the CB SCR for smaller w_C, which leads to a higher electron velocity. Figure 2.24 (a) compares the electron velocity v_n profiles in the collector for w_C = 200 and 30 nm designs. J_C = 40 mA/μm^2, V_{CB} = 0.5 V. Because of the non-equilibrium transport, velocity overshoot occurs. The smaller the w_C, the higher the v_n. Consequently, electron concentration n is higher for larger w_C. p is also higher for larger w_C, and hence base push-out (Kirk effect) is easier to happen. A w_C smaller than 30 nm is required to delay the onset of Kirk effect, and hence to achieve THz intrinsic design.

For the design with narrow collector width, the n/n^+ transition gradient directly impacts the effective low-doped collector thickness, and thus should have an impact on f_T and BV_{CEO}. Figure 2.25 compares the electric field distribution in the collector for five n/n^+ transitions with increasing gradient, for a x_{mole}^{peak}=27% graded Ge profile. Since a rapid charge distribution change is associated with an abrupt transition, a higher electric field is presented at the n/n^+ junction. BV_{CEO} is thus degraded from 1.55 V for coll1 to 1.36 V for coll5. Figure 2.26 compares τ_C-J_C for each design. Increasing n/n^+ gradient effectively increases w_C, and hence decreases f_T. Meanwhile, a more rapid f_T roll-off is observed for an abrupt n/n^+ transition. Therefore, a gradual n/n^+ transition is desired for high f_T, high BV_{CEO} and weaker f_T roll-off.



(a)



(b)

Figure 2.24: Comparisons of (a) electron velocity profile, and (b) n/p distributions inside CB SCR between $w_C = 200$ nm and 30 nm designs. $J_C = 40$ mA/μm^2, $V_{CB} = 0.5$ V. A base push-out is observed for $w_C = 200$ nm.

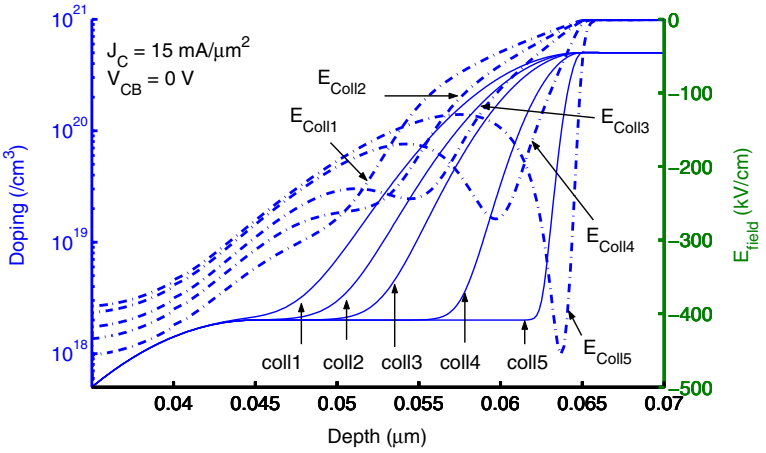


Figure 2.25: Electric field distribution for five sub-collector designs.

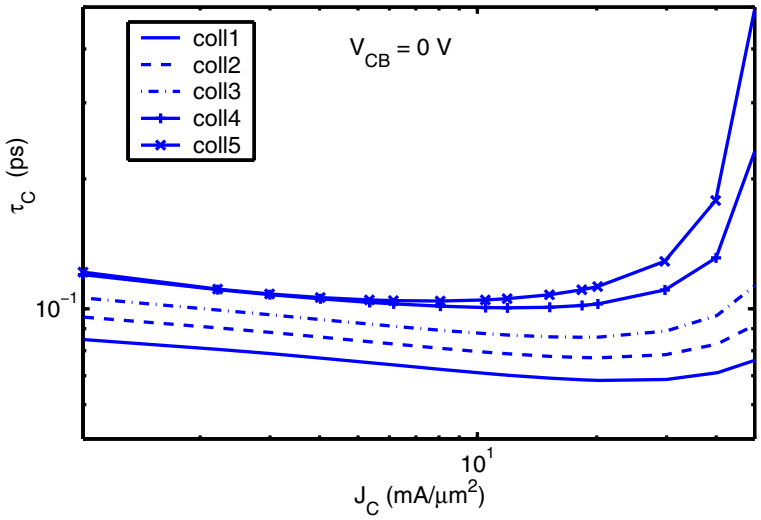


Figure 2.26: τ_C - J_C for five sub-collector designs.

2.4.3 Base Profile Optimization

Compared with 200 GHz SiGe HBT designs, we notice that the base sheet resistance $R_{B\square}$ for THz design is much higher, primarily because of the narrow base width. $R_{B\square}|_{200G}$ ≈2.5 kΩ/□, and $R_{B\square}|_{THz}$ ≈8 kΩ/□.

The base sheet resistance for an *npn* transistor is extracted as

$$R_{B\square} = \frac{1}{q\int_0^{w_B} N_B(x)\mu_p(x)dx}, \tag{2.20}$$

where w_B is the base width and μ_p is the hole mobility. Since $R_{B\square}$ is inversely proportional to base doping N_B and base width w_B, we can increase either N_B or w_B to optimize $R_{B\square}$. Because it is more challenging to form a very thin layer of heavily doped base, in this work, we increase w_B instead, which compromises τ_B and hence f_T. Figure 2.27 compares the optimized base profile with original design. The peak base doping is the same, and the base width is increased. $R_{B\square}$ is reduced to 3.5 kΩ/□, which is close to the value for 200 GHz designs. However, peak f_T is reduced also to 1.1 THz.

The reason to have a low $R_{B\square}$ is to reduce intrinsic base resistance and hence the total base resistance. For a double base contact structure, the intrinsic base resistance is

$$R_{B,in} = \frac{1}{12}\left(\frac{w_E}{l_E}\right) R_{B\square}. \tag{2.21}$$

The lateral scaling of w_E can also reduce $R_{B,in}$. With reducing $R_{B,in}$, the extrinsic base resistance $R_{B,ex}$ can become appreciable, which is related to the base current flowing through the extrinsic base region. In order to reduce $R_{B,ex}$, device structure innovation

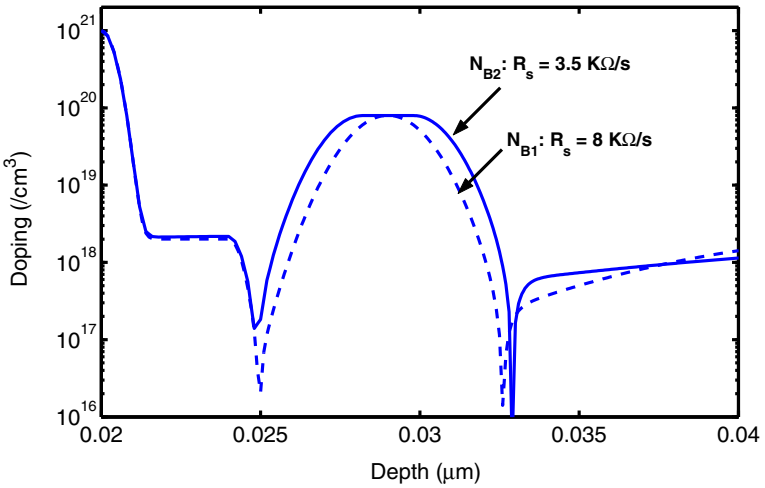


Figure 2.27: Optimized base profile for lower $R_{B\square}$

such as the raised extrinsic base and the lateral scaling of the extrinsic base are implemented.

2.5 Summary

In this chapter, the intrinsic profile design for THz SiGe has been discussed. The high-low emitter structure is shown to provide the advantage of reducing tunneling and G/R current, as well as reducing the V_{BE} modulations on base width. As a result, more linear J_B and J_C can be obtained, which gives better β. When emitter and base transit time is reduced dramatically with scaling, the collector SCR transit time becomes the limit to speed. In order to achieve THz f_T intrinsic design, collector width has to be reduced below 30 nm. The sub-collector transition also impacts τ_C and BV_{CEO}. A

gradual n-n^+ transition is favored, which results in smaller τ_C and higher BV_{CEO}. Comprehensive comparisons between graded and box Ge profile designs are examined. We conclude that the graded Ge profile wins in all aspects of the device performance matrix for THz designs. Next, we will extend the design to 2D to examine the impact of parasitic effects on device performance.

CHAPTER 3

2D DEVICE DESIGNS

The 2D device structure must be optimized and improved with intrinsic vertical scaling. For the THz intrinsic design, the total transit time τ_{EC} is on the order of 0.1 ps. The delay from extrinsic parasitic effects can be appreciable, and hence degrades device performance significantly. Many research groups have been devoted to device optimization and have proposed innovative 2D device structures [55, 57]. In this work, the design with optimized base profile is used, which has lower base sheet resistance $R_{B\square}$, and hence lower $R_{B,in}$.

Then the 2D structure parasitic effects are studied using MEDICI. The 1D transit time analysis is extended to 2D. The distributive capacitance is revisited for the 2D structure. The intrinsic and extrinsic base resistance, $R_{B,in}$ and $R_{B,ex}$, and CB junction capacitance C_{tc} are extracted. The raised extrinsic base structure has been chosen, which can reduce $R_{B,ex}$. $R_{B,in}$ and C_{tc} can be reduced through lateral scaling. Two lateral scaling schemes are examined. When w_E is scaled to 60 nm, f_{max} can be improved to 1090 GHz.

3.1 2D Device Simulations

Figure 3.1 shows the MEDICI 2D input device structure [59]. As discussed, with device optimization, the intrinsic base resistance $R_{B,in}$ is reduced, so that the extrinsic base resistance $R_{B,ex}$ can be appreciable and dominate the total base resistance. As

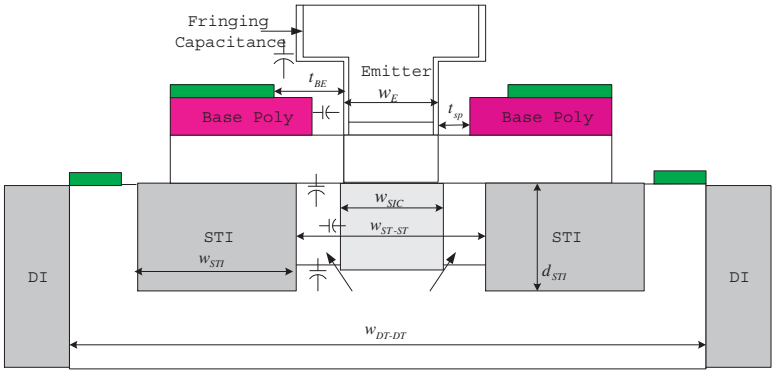


Figure 3.1: Schematic of device cross section using a raised base structure. w_E=120 nm, t_{EB}=140 nm, t_{sp}=40 nm, w_{STI}=0.4 μm, d_{STI}=0.25 μm, w_{ST-ST}=0.38 μm, and w_{DT-DT}=1.7 μm.

shown in [58], the extrinsic base-poly resistance $R_{B,poly}$, which is underneath the base contact and emitter spacer, can be reduced by reducing the spacing distance t_{BE}. As a result, R_B can be reduced and hence f_{max} can be improved. In this work, a t_{EB} of 140 nm is used. The STI design impacts the extrinsic CB capacitance. A smaller STI area is desired to reduce the capacitance. The STI thickness is 0.25 μm and the STI width is 0.4 μm. A smaller spacing between two STIs can help reduce the extrinsic device induced parasitic delays. However, the associated stress hurts the mobility and hence degrades the high speed. Currently, the spacing between the STI edges is 0.38 μm. The emitter window w_E is 0.12 μm. The y-parameters are simulated as a function of frequency, from which h_{21} and Mason's unilateral gain U are calculated. f_T and f_{max} are then obtained from the extrapolation of $|h_{21}|$ and U.

In some reports, drift-diffusion (DD) simulations were found to be accurate for SiGe HBTs with peak f_T above 200 GHz [60]. Our experience, however, shows that

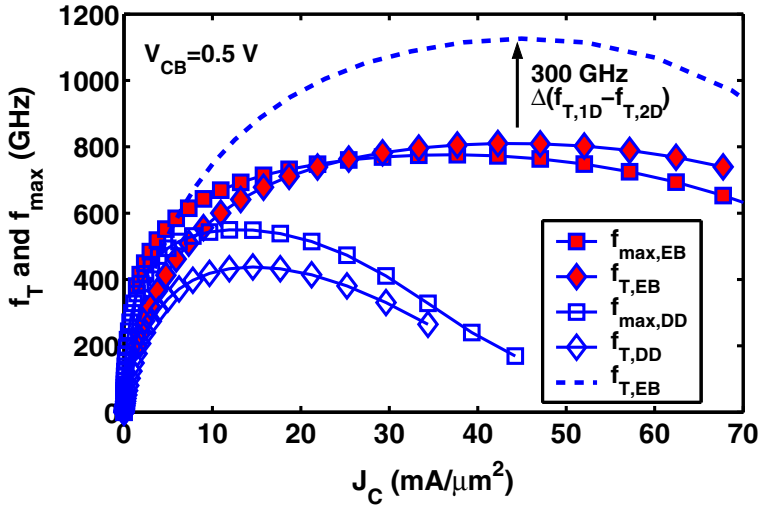
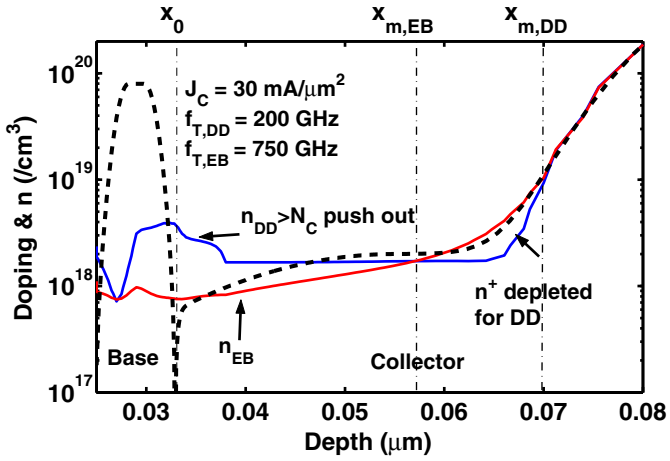
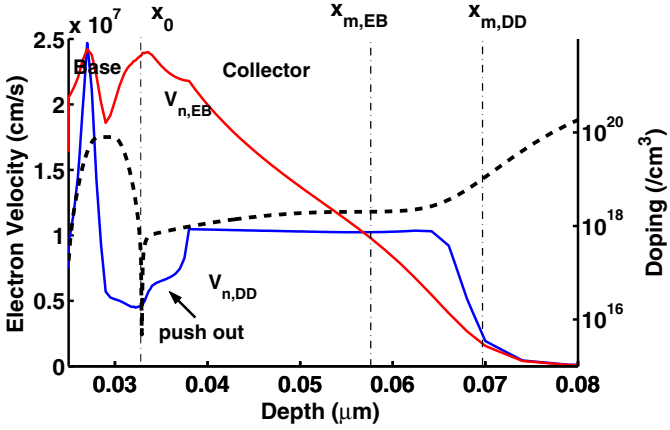


Figure 3.2: Comparisons of f_T/f_{max} - J_C curves obtained between 2D DD and EB simulations. Also shown is the 1D EB f_T result.

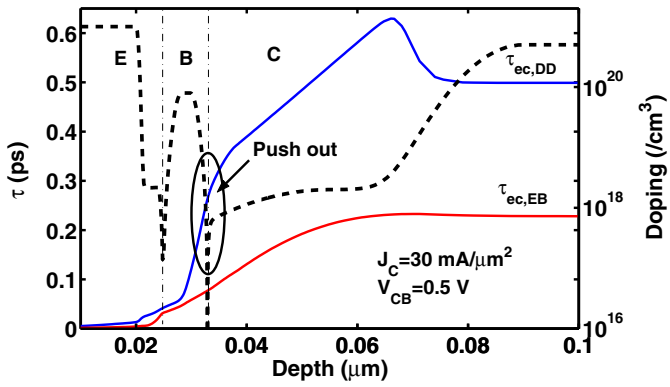
non-equilibrium transport is significant. Figure 3.2 compares the 2D $f_T/f_{max}-J_C$ curves from energy balance (EB) and drift-diffusion (DD) simulations. Peak f_T is 800 GHz and peak f_{max} is 745 GHz for EB simulations. The 1D f_T from EB simulation is also shown. A 200 GHz peak f_T reduction is observed due to extrinsic parasitic effects. Observe that the f_T and f_{max} roll off at lower J_C in DD simulation, because for DD simulations, $n_{DD} > N_C$ in CB SCR as shown in Figure 3.3 (a), and base push out occurs. The SCR boundary of DD simulation, $x_{m,DD}$, moves towards n^+ buried layer and n^+ is depleted. $x_{m,EB} < x_{m,DD}$. Figure 3.3 (b) compares the electron velocity between DD and EB simulations. Non-equilibrium transport leads to velocity overshoot, therefore $V_{n,EB}$ is greater than $V_{n,sat}$ in SCR for EB simulations. While for DD simulations, $V_{n,DD}$ is lower than $V_{n,sat}$ near CB junction because of the greater n_{DD} accumulated in the push



(a)



(b)



(c)

Figure 3.3: Comparisons of simulation details between DD and EB on (a) electron distribution, (b) electron velocity, and (c) τ distribution. $J_C = 30$ mA/μm^2, $V_{CB} = 0.5$ V.

out region. Generally, collector SCR transit time τ_{CSCR} is related to V_n as [27]

$$\tau_{CSCR} = \int_{x_0}^{x_m} \frac{1}{V_n(x)} \left(1 - \frac{x}{x_m - x_0}\right) dx, \tag{3.1}$$

where x_0 and x_m indicate the collector SCR boundaries. At equilibrium, $V_n(x) = V_{n,sat}$ and $\tau_{CSCR} = (x_m - x_0)/2V_{n,sat}$. (3.1) indicates that the velocity close to x_0 has more weight. As $V_{n,EB} \gg V_{n,DD}$ at the start, τ_{CSCR} is lower for EB simulations. Figure 3.3 (c) compares τ_{EC} profiles between DD and EB. For DD simulations, a dramatic increase of τ is observed at the push out region, and a large τ decrease at the n^+ is observed. The reason is that $x_{m,DD}$ is pushed to the buried layer, the n/n^+ transition layer is depleted, and a negative $\Delta\tau_n$ appears at the end of CB SCR [48]. This non-equilibrium transport delays the onset of Kirk effect, and is significant for THz HBTs, as the collector doping does not need to dramatically increase.

Considering impact ionization at the CB junction, the non-local effect due to the non-equilibrium transport equivalently delays the maximum impact ionization rate [61–63]. Therefore, a practical device breakdown performance can be maintained in scaled devices with over 200 GHz peak f_T [64]. This is particularly important for applications where transient current density can significantly go beyond the peak f_T current density, such as in large signal power amplifiers [65].

To include the non-local effect, an effective electric field E_{eff}, instead of local electric field, is used in the avalanche analysis. The energy of electrons is derived from

approximations to the EB equations as [61]

$$\Delta w_n = \frac{3}{5} q \int_0^x E(u) \exp\left(\frac{u-x}{\lambda_n}\right) du, \tag{3.2}$$

where Δw_n is the average electron energy, $E(u)$ is the local electric field, and λ_n is the energy relaxation length for electrons. E_{eff} is then calculated as

$$E_{eff} = \frac{5}{3} \frac{\Delta w_n}{q \lambda_n}. \tag{3.3}$$

The M-1 curves extracted using E_{eff} between THz and a 200 GHz design are compared in Figure 3.4. $M-1$ at higher V_{CB} is about the same between THz and 200 GHz HBTs. Consequently, the CB breakdown voltage of THz HBTs is expected to be similar as the 200 GHz HBTs.

3.2 2D Parasitic Analysis

To better understand how parasitic effects limit f_T/f_{max}, we extract the base resistance, and CB capacitance from MEDICI simulation results. Also, we extend the 1D transit time analysis into 2D. Similarly, the 2D distributive capacitance is extracted.

3.2.1 R_B and C_{CB}

Given the increasing difficulty of base resistance extraction using impedance circle methods in scaled HBTs, we extract the extrinsic and intrinsic base resistance, $R_{B,ex}$ and

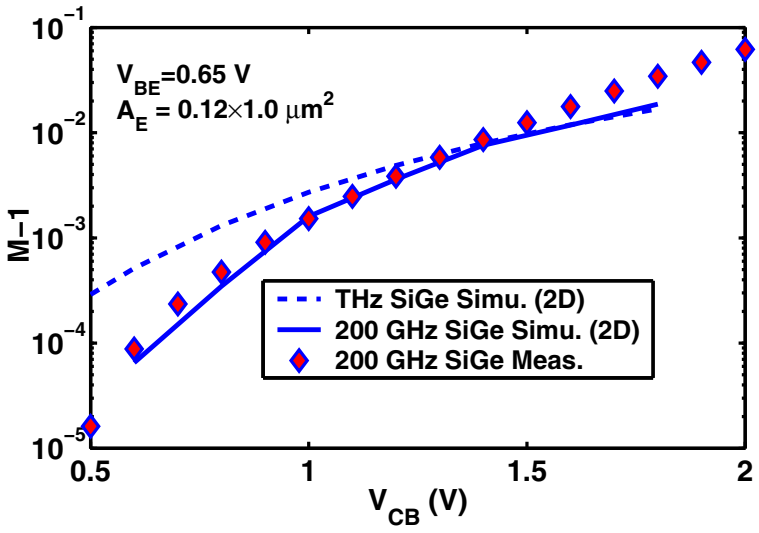


Figure 3.4: Simulated $M-1$ versus V_{CB} curves for 200 GHz and THz SiGe HBTs. Also shown is the measured data for 200 GHz design.

$R_{B,in}$, as follows

$$R_{B,ex} = \Delta\phi_{fp}/I_B, \tag{3.4}$$

$$R_{B,in} = \frac{1}{12} \times \frac{R_\square}{l_E/w_E} \tag{3.5}$$

where $\Delta\phi_{fp}$ is the quasi-hole Fermi potential drop across the extrinsic base, I_B is the base current, $R_\square$ is the intrinsic base sheet resistance, and l_E=1 μm. The extracted $R_{B,in}$ is 33.6 Ω-μm, and $R_{B,ex}$ is 70 Ω-μm. The total r_b is comparable to the 200 GHz HBT reported in [55], which has a r_b of 110 Ω-μm. An ideal ohmic base contact is assumed in this work, and the silicide resistance is not included in R_B. Therefore, our analysis represents the best case after we make the perfect contact.

In order to extract C_{CB}, "ωIm(Z_{22}-Z_{21})" is plotted versus "ω^2", the y-intercept equals "$-1/C_{CB}$" [66]. Using this method, a C_{CB} of 1.14 fF/μm is extracted, slightly smaller than the 1.7 fF/μm reported in [55], likely due to the smaller spacing between the STI and the smaller area of the selectively implanted collector (SIC).

3.2.2 2D Transit Time Analysis

Extending the 1D regional transit time analysis into 2D, the total τ_{EC} can be expressed as

$$\tau_{EC} = \frac{\partial Q}{\partial I_C} = \frac{q}{\Delta I_C} \int_{x1}^{x2} \int_{y1}^{y2} \Delta n dx dy, \tag{3.6}$$

where the Δn and ΔI_C are for a small ΔV_{BE} increase at a given V_{CE}. x_1, x_2, y_1 and y_2 are the boundaries for the 2D regions. Vertically, we use Y^*_{EE}, Y^*_{EB} and Y^*_{CB} as boundaries, which correspond to x_{EE}, x_{EB} and x_{CB} in Figure 2.10, respectively. Laterally, the emitter window boundaries $X_{E,st}$ and $X_{E,en}$ are used to separate the intrinsic and the extrinsic device. Then τ_{EC} is divided into seven components as shown in Figure 3.5,

1. τ_{Ef}: the fringing capacitance charging time induced by EB spacers;
2. τ_E: the emitter transit time ;
3. τ_{te}: the EB junction capacitance charging time;
4. $\tau_{B,in}$: the intrinsic base transit time;
5. $\tau_{B,ex}$: the extrinsic base transit time;

6. $\tau_{C,in}$: the intrinsic collector transit time;

7. $\tau_{C,ex}$: the extrinsic collector transit time.

τ_{Ef} is extracted as the integration along the emitter perimeter. τ_E equals the integration above Y^*_{EE} for the rest of emitter region. τ_{te}, $\tau_{B,in}$ and $\tau_{C,in}$ are calculated using

$$\tau_{te} = \frac{q}{\Delta I_C} \int_{X_{E,st}}^{X_{E,en}} \int_{Y^*_{EE}}^{Y^*_{EB}} \Delta n dy dx, \tag{3.7}$$

$$\tau_{B,in} = \frac{q}{\Delta I_C} \int_{X_{E,st}}^{X_{E,en}} \int_{Y^*_{EB}}^{Y^*_{CB}} \Delta n dy dx, \tag{3.8}$$

$$\tau_{C,in} = \frac{q}{\Delta I_C} \int_{X_{E,st}}^{X_{E,en}} \int_{Y^*_{CB}}^{Y^*_{C}} \Delta n dy dx. \tag{3.9}$$

Since SIC width is assumed to be the same as w_E in our case, $X_{E,st}$, $X_{E,en}$, Y^*_{CB} and Y_C* define the SIC region. Y^*_C is the n/n^+ interface. $\tau_{B,ex}$ is the integral for the extrinsic base, which is induced by the lateral electron injection through the emitter side-wall. $\tau_{C,ex}$ is the integral for the extrinsic collector, which is outside of the SIC region. Applying (9) and (10) for 2D τ analysis, we can further distinguish SCR transit time and depletion capacitance charging time for both $\tau_{C,in}$ and $\tau_{C,ex}$, which are denoted as $\tau_{CSCR,in}$, $\tau_{tc,in}$, $\tau_{CSCR,ex}$ and $\tau_{tc,ex}$, respectively. An example of Δn distribution near peak f_T is shown in Figure 3.6, where most of the Δn occurs in the intrinsic base and collector regions.

Since MEDICI only outputs the Δn associated with each mesh node in the "tiff" format [37], we have to post-process the raw data in order to extract Δn and device geometry informations. Then, based on the 2D regional boundaries defined above, we run our own Matlab 2D integration script to calculate each 2D τ. Figure 3.7 shows

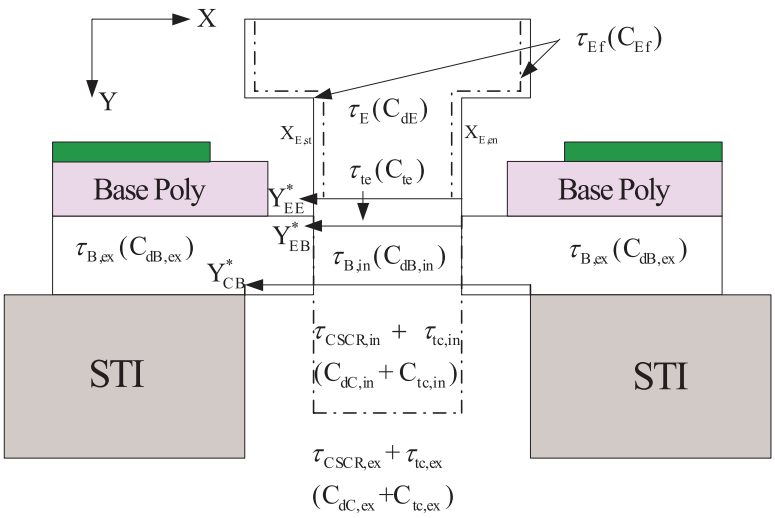


Figure 3.5: Definition of 2D τ components and related capacitance. Y^*_{EE}, Y^*_{EB} and Y^*_{CB} correspond to X_{EE}, X_{EB} and X_{CB} in Figure 3.

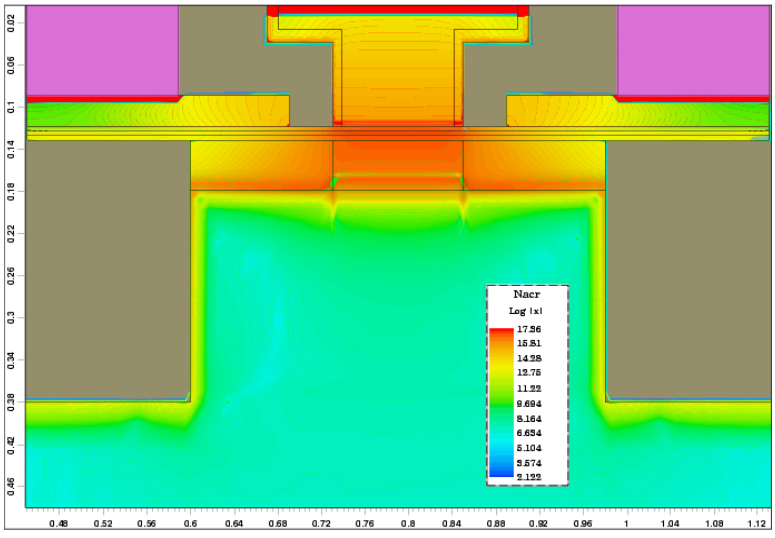


Figure 3.6: Simulated 2D Δn distribution near peak f_T, V_{CB}=0.5 V

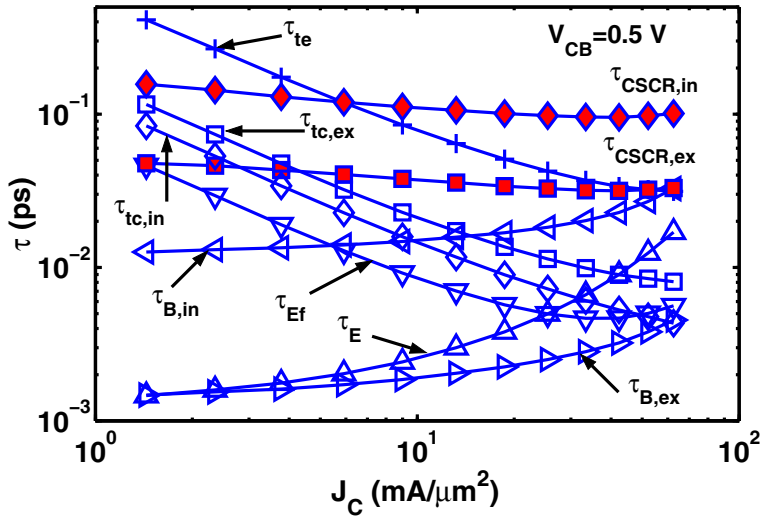


Figure 3.7: τ_E, τ_{Ef}, τ_{EB}, $\tau_{B,in}$, $\tau_{B,ex}$, $\tau_{CSCR,in}$, $\tau_{CSCR,ex}$, $\tau_{tc,in}$ and $\tau_{tc,ex}$ versus J_C at V_{CB}=0.5 V.

the 2D τ components defined above versus J_C at V_{CB}=0.5 V. $\tau_{CSCR,in}$ and $\tau_{CSCR,ex}$ are not sensitive to biasing. Both τ_{te} and $\tau_{CSCR,in}$ are appreciable from low to high J_C. τ_{te} limits τ_{EC} at lower J_C, while $\tau_{CSCR,in}$ does at higher J_C. τ_E, $\tau_{B,in}$ and $\tau_{B,ex}$ increase at high J_C when high injection occurs. τ_{Ef}, τ_{te}, $\tau_{tc,in}$, and $\tau_{tc,ex}$ show the depletion capacitance charging behavior. $\tau_{tc,ex}$ is larger than $\tau_{tc,in}$ The contribution from each region is compared in Figure 3.8 in percentage. The $f_T - J_C$ characteristic is primarily determined by τ_{te}, $\tau_{CSCR,in}$ and $\tau_{CSCR,ex}$. Extrinsic collector design dominates the total extrinsic delay, and leads to 200 GHz peak f_T reduction compared with 1D simulation.

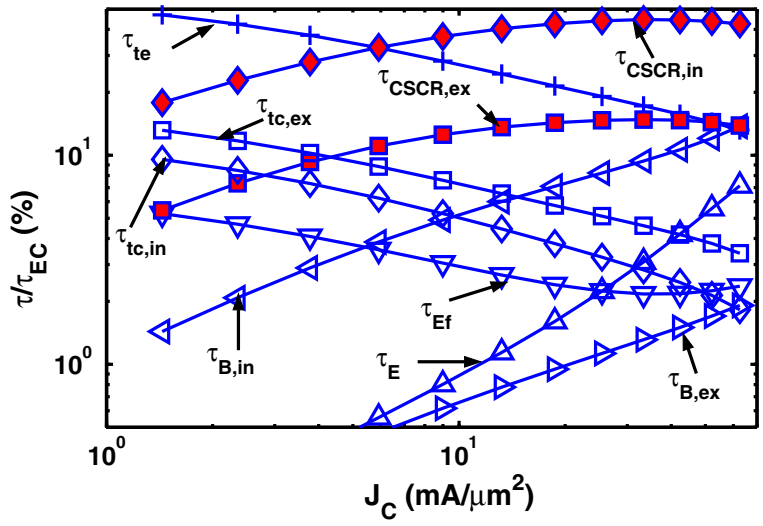


Figure 3.8: Comparison of τ contribution from each region. V_{CB}=0.5 V.

3.2.3 2D Distributive Capacitance

Similar to 1D distributive capacitance, 2D distributive capacitance can be extracted. Table 3.1 lists the definition of each capacitance component. Figure 3.9 shows capacitance versus J_C curves. In our analysis, C_{te} includes EB junction depletion capacitance and neutral capacitance as well. However, the neutral capacitance is very small and the depletion capacitance dominates. As shown, C_{te}, C_{Ef}, $C_{tc,in}$ and $C_{tc,ex}$ are appreciable for those highly scaled SiGe HBT designs. C_{tc}, the sum of $C_{tc,in}$ and $C_{tc,ex}$, is 1.10 fF/μm, which agrees with 1.14 fF/μm extracted using [Z] parameters. C_{te}>C_{tc}, and $C_{tc,ex}$>$C_{tc,in}$. The diffusion capacitance, which is associated with excess carrier injection, increases with J_C exponentially. $C_{C,in}$ and $C_{C,ex}$ are dominant. Overall, collector design is shown to be more important for high speed SiGe HBT designs.

Table 3.1: 2D Distributive Capacitance

C_{dE}	$g_m\tau_E$	Emitter diffusion capacitance
C_{dB}	$g_m\tau_B$	Base diffusion capacitance
$C_{dC,in}$	$g_m\tau_{CSCR,in}$	Intrinsic Collector SCR Capacitance
$C_{dC,ex}$	$g_m\tau_{CSCR,ex}$	Extrinsic Collector SCR Capacitance
C_{te}	$g_m\tau_{te}$	EB junction capacitance
C_{ef}	$g_m\tau_{Ef}$	EB fringing capacitance
$C_{tc,in}$	$g_m\tau_{tc,in}$	Intrinsic CB capacitance
$C_{tc,ex}$	$g_m\tau_{tc,ex}$	Extrinsic CB capacitance

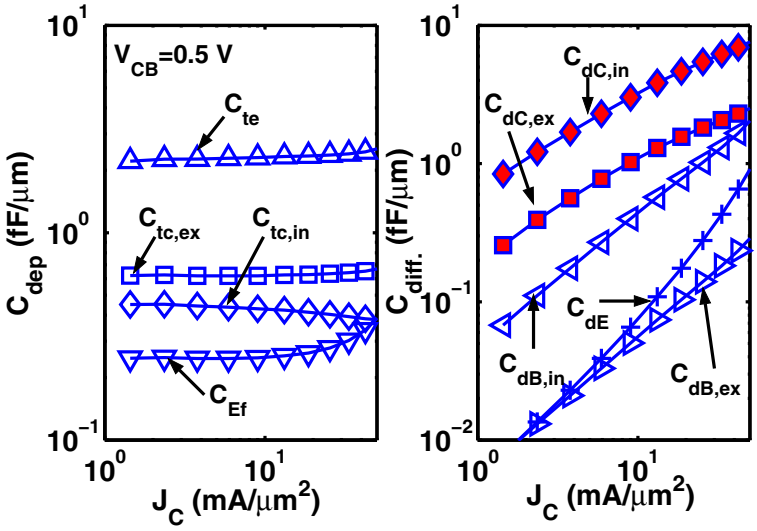


Figure 3.9: Capacitance versus J_C curves. V_{CB}=0.5 V.

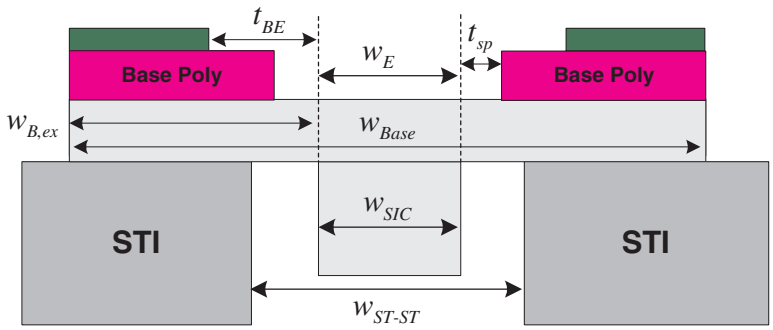
Table 3.2: Device performance for fixed w_{Base} lateral scaling.

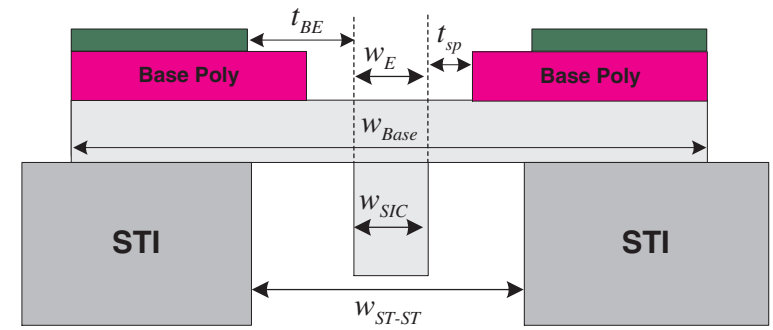
w_E(nm)	120	100	80	60
$R_{B,in}$(Ω-μm)	33.6	28	22.4	16.8
$R_{B,ex}$(Ω-μm)	70	65.6	64	62
C_{CB} (fF/μm)	1.14	1.12	1.09	1.06
f_T (GHz)	787	771	750	718
f_{max} (GHz)	745	871	966	1090

3.3 Lateral Scaling

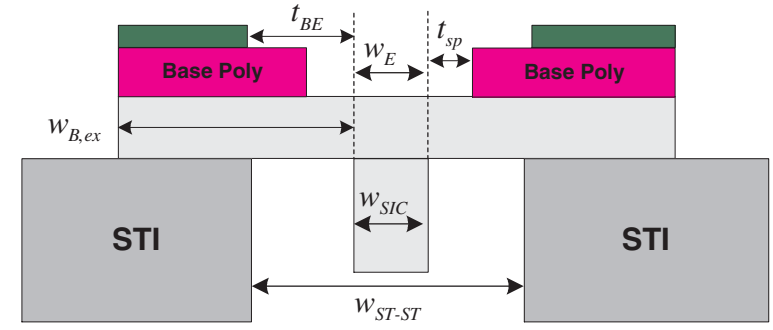
Lateral scaling is necessary to reduce $R_{B,in}$ and to reduce current crowding. To examine the impact of lateral scaling, we shrink the emitter window w_E from 120 nm to 100 nm, 80 nm and 60 nm, respectively, as illustrated in Figure 3.10. We also assume that the SIC width is the same as w_E. The spacer thickness t_{sp}, and the spacing between base contact and emitter t_{EB} are both fixed during scaling. The active collector area is defined by the STI spacing w_{ST-ST}. We will first consider scaling w_E for fixed total base width, w_{Base}. Besides, w_{STI}, d_{STI}, w_{ST-ST} and w_{DT-DT} are all kept the same. The w_{ST-ST} of 0.38 μm is already small. Further shrinking will lead to the device performance degradation due to stress.

For the fixed w_{Base} scaling, extrinsic base width $w_{B,ex}$ increases with scaling. Therefore, at the same V_{BE}, base current I_B is more for the scaled devices, and thus the $R_{B,ex}$ is smaller as shown in Table 3.2. C_{CB} is extracted from [Z] parameters, and the values decrease because of the smaller SIC area. Overall, through lateral scaling, a 355 GHz f_{max} improvement is obtained. However, f_T decreases from 787 GHz to 718 GHz.





(a) fixed w_{Base}



(b) fixed $w_{B,ex}$

Figure 3.10: Illustration of lateral scaling. (a) w_{Base} is fixed. (b) $w_{B,ex}$ is fixed.

Table 3.3: Device performance for fixed $w_{B,ex}$ lateral scaling.

w_E (nm)	120	100	80	60
$R_{B,in}$(Ω-μm)	33.6	28	22.4	16.8
$R_{B,ex}$(Ω-μm)	70	70	70	70
C_{CB} (fF/μm)	1.14	1.10	1.08	1.05
f_T(GHz)	787	785	770	760
f_{max}(GHz)	745	871	968	1090

Next, we keep $w_{B,ex}$ the same during scaling, as shown in Figure 3.10 (b). The advantage is a smaller overlap between base and collector through STI, and less peak f_T reduction. The disadvantage is that $R_{B,ex}$ will remain the same (as opposed to decreasing) with scaling. Table 3.3 compares the device performance for different w_E.

3.3.1 2D versus 1D Intrinsic

To compare the impacts of 1D intrinsic design and 2D design on transit time through lateral scaling, we define the 1D intrinsic transit time and extrinsic delays as

$$\tau_{E,in} = \tau_E + \tau_{te}, \text{(Intrinsic emitter transit time)} \tag{3.10}$$

$$\tau_{E,ex} = \tau_{Ef}, \text{(Extrinsic emitter transit time)} \tag{3.11}$$

$$\tau_{B,in} = \tau_{B,in}, \text{(Intrinsic base transit time)} \tag{3.12}$$

$$\tau_{B,ex} = \tau_{B,ex}, \text{(Extrinsic base transit time)} \tag{3.13}$$

$$\tau_{C,in} = \tau_{CSCR,in} + \tau_{tc,in}, \text{(Intrinsic collector transit time)} \tag{3.14}$$

$$\tau_{C,ex} = \tau_{CSCR,ex} + \tau_{tc,ex}, \text{(Extrinsic collector transit time)} \tag{3.15}$$

Figure 3.11 compares each intrinsic and extrinsic τ for the two scaling methods at w_E = 120, 100, 80, and 60 nm, respectively. Both $\tau_{C,ex}$ and $\tau_{B,ex}$ increase with reducing w_E for fixed w_{Base} scaling. However, for fixed $w_{B,ex}$ scaling, $\tau_{B,ex}$ is the same. Therefore, the increase of τ_{EC} with lateral scaling is less for fixed $w_{B,ex}$ scaling compared with fixed w_{Base} scaling. As a result, f_T at w_E = 60 nm for fixed $w_{B,ex}$ scaling is higher than fixed w_{Base} scaling. The intrinsic and extrinsic collector depletion capacitance and SCR capacitance are also extracted using 2D transit time analysis. The results are compared in Figure 3.12 (a) and (b), respectively. Since w_{ST-ST} is kept the same for both scalings, similar collector capacitances are observed. Because of the smaller SIC area, both $C_{tc,in}$ and $C_{CSCR,in}$ decrease with reducing w_E, However, $C_{tc,ex}$ increases, which makes the total collector depletion capacitance C_{CB} not scale with lateral scaling. $C_{CSCR,ex}$ is

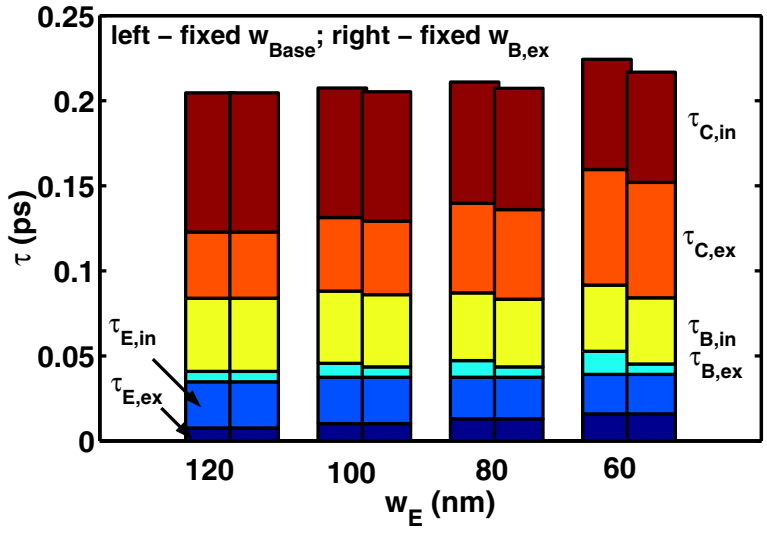
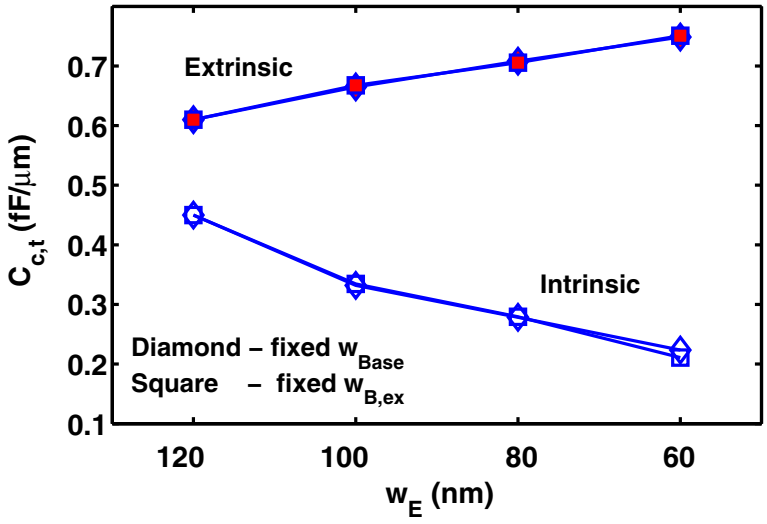


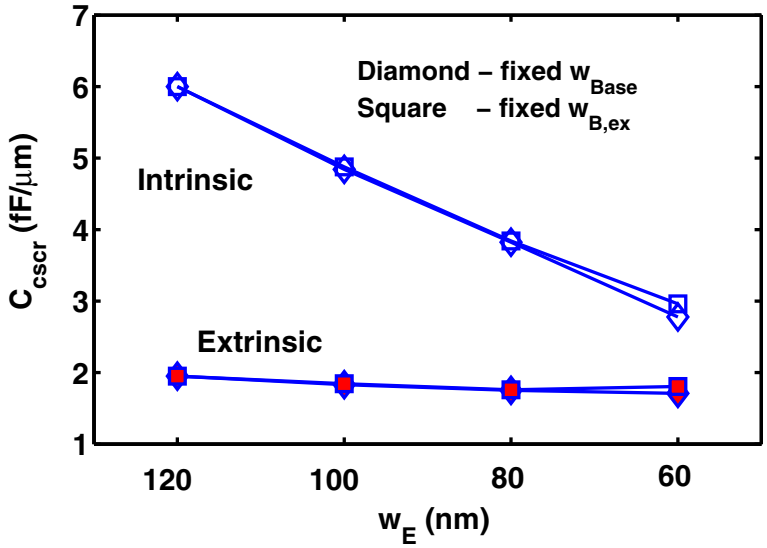
Figure 3.11: Comparisons of intrinsic and extrinsic τ_E, τ_B and τ_C at peak f_T between two lateral scaling.

not sensitive to lateral scaling. In the extrinsic collector, most Δn happens along SIC boundaries, which is not impacted by lateral scaling. Overall, the low $R_{B,ex}$ with fixed w_{Base} scaling is traded off by its higher $\tau_{B,ex}$, and the two scaling schemes lead to the same f_{max} of 1090 GHz at $w_E = 60$ nm. A fixed $w_{B,ex}$ scaling has led to a better overall performance, because of higher f_T.

For both scaling schemes, the extrinsic transit time becomes an increasingly larger portion of the total transit time as w_E decreases from 120 to 60 nm, as shown in Figure 3.13. If we do a linear extrapolation, τ_{in} and τ_{ex} curves will meet near $w_E = 50$ nm. After this point, extrinsic design dominates the total τ_{EC}.



(a)



(b)

Figure 3.12: Comparisons of (a) $C_{tc,in}$ and $C_{tc,ex}$, and (b) $C_{CSCR,in}$ and $C_{CSCR,ex}$ for fixed w_{Base} and fixed $w_{B,ex}$ scalings.

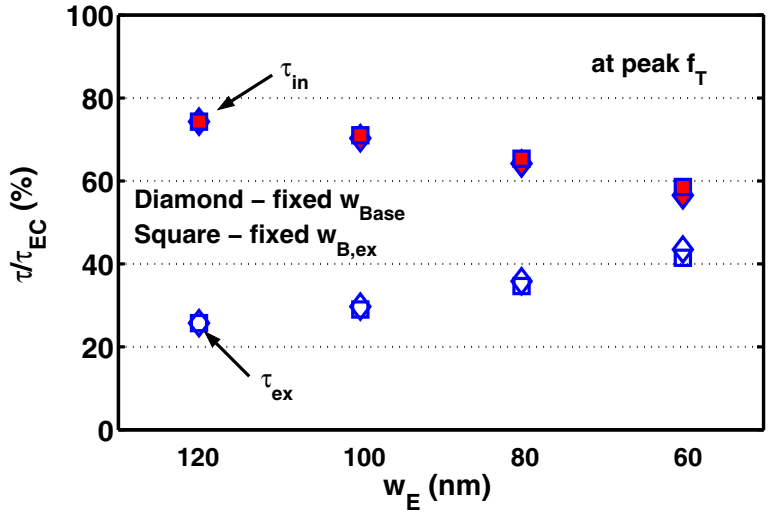


Figure 3.13: Comparisons of intrinsic and extrinsic τ contributions at peak f_T.

3.3.2 2D versus 2D Intrinsic

Another way of examining the parasitic effects impact is to define the internal 2D intrinsic device as shown in Figure 3.14, where the width is 20 nm wider than w_E in order to include the spreading effects. The two lateral scaling schemes are re-visited for the 2D intrinsic device. The peak values of f_T and f_{max} for 2D intrinsic device are extracted and compared with the data for a full 2D device. The results are shown in Figure 3.15.

Recall Figure 2.3, the 1D intrinsic f_T is 1100 GHz. Due to the spreading effects, the 2D intrinsic peak f_T is reduced to 980 GHz. The parasitic RC delays lead to more than 100 GHz peak f_T reduction. The solid arrows show the f_T reduction for the fixed w_{Base} lateral scaling and the dash arrows show the f_T reduction for the fixed $w_{B,ex}$ lateral scaling. As shown, the first case leads to more severe f_T reduction, which is due to

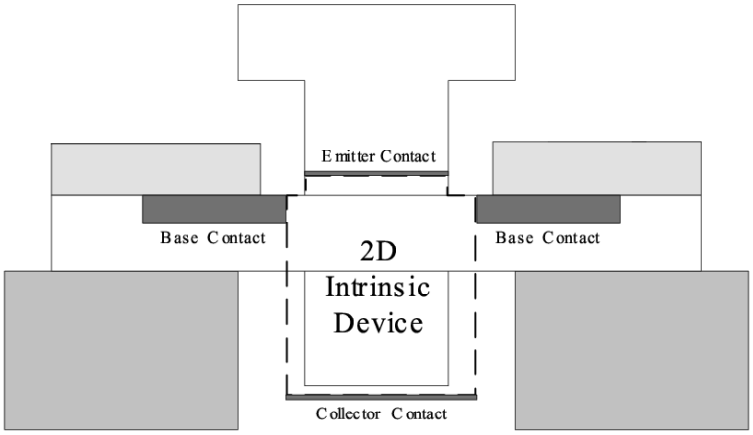


Figure 3.14: Illustration of the internal 2D intrinsic device.

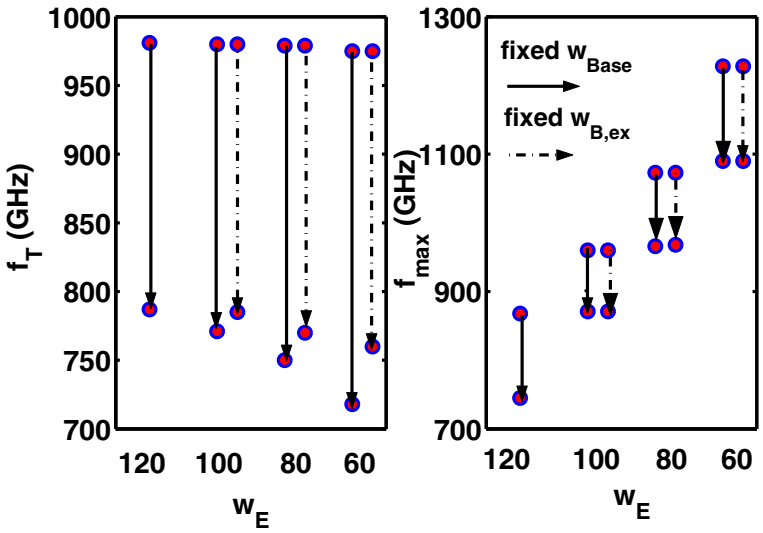


Figure 3.15: Comparisons of 2D intrinsic and extrinsic f_T/f_{max} between two scalings.

the higher $\tau_{B,ex}$ as discussed above. However, the similar f_{max} decrease is observed between two scaling methods. When w_E is reduced to 60 nm, while the 2D intrinsic f_{max} increases from 880 GHz to 1220 GHz, the full 2D f_{max} increases from 745 GHz to 1090 GHz.

3.4 Summary

We have examined the impact of extrinsic parasitic effects on the f_T and f_{max} of highly scaled SiGe HBTs to explore the structural requirements necessary to further improve f_{max} towards terahertz. A 2D regional transit time analysis shows that the extrinsic collector parasitics are the most dominant extrinsic transit time contributor, which accounts for 20% of τ_{EC} at peak f_T. The fixed w_{Base} and the fixed $w_{B,ex}$ lateral scaling schemes are examined to quantify the extrinsic parasitics geometry reduction needed to achieve f_{max} improvement. The non-equilibrium transport is shown to significantly reduce the electron concentration in the CB SCR, which alleviates Kirk effect and helps improving breakdown voltage.

CHAPTER 4

CONCLUSION

We have explored in this work the optimization of SiGe HBT towards THz f_T/f_{max}. The 1D intrinsic design is discussed first. The transit time analysis is used to probe the internal τ, and find the bottleneck for speed. Both base width w_B and collector width w_C have to be reduced to nanometer scale in order to reduce τ_B and τ_C. The base doping N_B as high as $8\times10^{1}9/\text{cm}^3$ is used to reduce base sheet resistance $R_{B,\square}$. The Ge design impacts on device performance are compared between box and graded Ge profile. The same film stability is used. Graded Ge profile can give higher f_T and higher BV_{CEO}.

Next, 2D design of SiGe HBT is addressed. The 1D transit time analysis is extended to 2D, which can quantify transit delay from each intrinsic and extrinsic part. The base resistance and CB capacitance are both extracted from MEDICI simulation results. Using those methods, the lateral scaling is examined. The necessary structure is revealed, which can achieve THz f_{max}.

In order to improve this work, we would like to propose to explore the 2D device structure and processing limitation in the future.

BIBLIOGRAPHY

[1] G. L. Patton, J. H. Comfort, B. X. Meyerson, E. F. Crabbe, G. J. Scilla, E. Fresart, J. M. C. Stork, J. Y.-C. Sun, D. L. Harame, J. Burghartz, "63-75 GHz f_T SiGe-base Heterojunction-Bipolar Technology," *Tech. Dig. IEEE Symp. VLSI Tech.*, pp. 49-50, 1990.

[2] G. L. Patton, J. H. Comfort, B. S. Meyerson, E. F. Crabbe, G. J. Scilla, E. Fresart, J. M. C. Stork, J. Y.-C. Sun, D. L. Harame, J. Burghartz, "75 GHz f_T SiGe Base Heterojunction Bipolar Transistors," *IEEE Electron Device Lett.*, vol. 11, pp. 171-173, 1990.

[3] D. L. Harame, E. F. Crabbe, J. D. Cressler, J. H. Comfort, J. Y.-C. Sun, S. R. Stiffler, E. Kobeda, J. N. Burghartz, M. M. Gilbert, J. C. Malinowski, A. J. Dally, S. Ratanaphanyarat, M. J. Saccamango, W. Rausch, J. Cotte, C. Chu, J. M. C. Stork, "A High-Performance Epitaxial SiGe-Base ECL BiCMOS Technology," *Tech. Dig. IEEE Int. Electron Device Meeting*, pp. 19-22, 1992.

[4] E. Crabbe, B. Meyerson, D. Harame, J. Stork, A. Megdanis, J. Cotte, J. Chu, M. Gilbert, C. Stanis, J. Comfort, G. Patton, and S. Subbanna, "113-GHz f_T Graded-Base SiGe HBT's," *IEEE Trans. on Electron Devices*, vol. 40, no. 11, pp. 2100 - 2101, 1993.

[5] K. Oda, E. Ohue, M. Tanabe, H. Shimamoto, T. Onai, and K. Washio, "130 GHz-f_T SiGe HBT Technology," *Tech. Dig. IEEE Int. Electron Device Meeting*, pp. 791 - 794, 1997.

[6] A. Gruhle, H. Kibbel, A. Schurr, D. Behammer, and U. Konig, "SiGe Heterojunction Bipolar Transistors with 156 GHz Transit Frequency," *Proc. of State-of-The-Art Program on Compound Semiconductors*, pp. 198 - 200, 1999.

[7] S.-J. Jeng, B. Jagannathan, J.-S. Rieh, J. Johnson, K. T. Schonenberg, D. Greenberg, A. Stricker, H. Chen, M. Khater, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, "A 210 GHz f_T SiGe Heterojunction Bipolar Transistor with a non-self-aligned (NSA) structure," *IEEE Electron Device Lett.*, vol. 22, no. 11, pp. 542 - 544, 2001.

[8] J. -S. Rieh, B. Jagannathan, H. Chen, K. T. Schonenberg, D. Angell, A. Chinthakindi, J. Florkey, F. Golan, D. Greenberg, S. -J. Jeng, M. Khater, F.

Pagette, C. Schnabel, P. Smith, A. Stricker, K. Vaed, R. Volant, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, "SiGe HBTs with Cut-off Frequency of 350 GHz," *Tech. Dig. IEEE Int. Electron Device Meeting*, pp.771-774, 2002.

[9] J. Slotboom, Keynote Speech, *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003.

[10] B. B. Hu and M. C. Nuss, "Imaging with Terahertz Waves," *Optics Lett.*, vol. 20, no. 16, pp. 1716 - 1718, 1995.

[11] S. Withington, "Thz Technology for Astronomy," *IEE Colloquium on Terahertz Technology*, pp. 1 - 3, 1995.

[12] R. G. Nitsche, R. U. Titz, and E. M. Biebl, "A Low-Noise 2.5 THz Heterodyne Receiver with Tunable Reflector Antenna for Atmospheric OH-spectroscopy," *IEEE MTT-s Digest*, pp. 1727 - 1730, 1998.

[13] B. H. Deng, C. W. Domier, A. J. H. Donne, K. C. Lee, N. C. Luhmann Jr., E. Mazzucato, T. Munsat, H. Park, and M. van de Pol, "THz Techniques in Plasma Diagnotics," *IEEE MTT-s Digest*, pp. 1587 - 1590, 2002.

[14] T. W. Crowe, R. J. Mattauch, H. P. Roser, W. L. Bishop, W. C. B. Peatman, and X. Liu, "GaAs Schottky Diodes for THz Mixing Applications," *Proceedings of IEEE*, vol. 80, no. 11, pp. 1827 - 1841, 1992.

[15] S. T. Allen, U. Bhattacharya, and M. J. W. Rodwell, "4 THz Sidewall-etched Varactors for sub-mm-wave Sampling Circuits," *Technical Digest of Gallium Arsenide Integrated Circuit Symposium*, pp. 285 - 287, 1993.

[16] H. Eisele, "Conventional and Novel Approaches to RF Power Generation with Two-terminal Devices at Terathertz Frequencies," *Proceedings of IEEE International Conference on Terahertz Electronics*, pp. 13 - 18, 2002.

[17] Y. Shi and G. Niu, "Vertical Profile Design and Transit Time Analysis of Nano-Scale SiGe HBTs for Terahertz f_T," *Proc. of the 2004 IEEE Bipolar/BiCMOS Circuits Technology Meeting*, pp. 213 - 216, 2004.

[18] M. Nandakumar, A. Chatterjee, S. Sridhar, K. Joyner, M. Rodder, and I-C. Chen, "Shallow Trench Isolation for Advanced ULSI CMOS Technologies," *Tech. Dig. IEEE Int. Electron Device Meeting*, pp. 133 - 136, 1998.

[19] D. D. Tang, P. M. Solomon, T. H. Ning, R. D. Isaac, and R. E. Burger, "1.25 μm Deep-Groove-Isolated Self-Aligned Bipolar Circuits," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 925 - 931, 1982.

[20] G. P. Li, T. H. Ning, C. T. Chuang, M. B. Ketchen, D. D. Tang, and J. Mauer, "An Advanced High-performance Trench-Isolated Self-Aligned Bipolar Technology," *IEEE Trans. on Electron Devices*, vol. ED-34, pp. 2246 - 2254, 1987.

[21] H. N. Yu, "Transistor with Limited-area Base–Collector Junction," U. S. Patent Re. 27 045 re-issued, Feb. 2, 1971.

[22] L. D. Lanzerotti, J. Sturm, E. Stach, R. Hull, T. Buyuklimanli, and C. Magee, "Suppression of Borom Outdiffusion in SiGe HBT's by Carbon Incorporation," *Tech. Dig. IEEE Int. Electron Device Meeting*, pp. 249 - 252, 1996.

[23] G. Freeman, B. Jagannathan, S.-J Jeng, J.-S Rieh, A. Stricker, D. Ahlgren, and S. Subbanna, "Transistor Design and Application Considerations for > 200-GHz SiGe HBTs," *IEEE Trans. on Electron Devices*, vol. 50, pp. 645 - 655, 2003.

[24] P. Chevalier, C. Fellous, L. Rubaldo, D. Dutartre, M. Laurens, T. Jagueneau, F. Leverd, S. Bord, C. Richard, D. Lenoble, J. Bonnouvrier, M. Marty, A. Perrotin, D. Gloria, F. Saguin, B. Barbalat, R. Beerkens, N. Zerounian, F. Aniel, and A. Chantre, "230 GHz Self-Aligned SiGeC HBT for 90 nm BiCMOS Technology," *Proc. of the 2004 IEEE Bipolar/BiCMOS Circuits Technology Meeting*, pp. 225 - 228, 2004.

[25] G. Freeman, J.-S. Rieh, B. Jagannathan, Z. Yang, F. Guarin, A. Joseph, "Device Scaling and Application Trends for Over 200 GHz SiGe HBTs," *Silicon Monolithic Integrated Circuits in RF Systems*, pp. 6 - 9, 2003

[26] J. -S Rieh, B. Jagannathan, D. Greenberg, M. Meghelli, A. Rylyakov, F. Guarin, Z. Yang, D. Ahlgren, G. Freeman, P. Cottrell, D. Harame, "SiGe Heterojunction Bipolar Transistors and Circuits toward Terahertz Communication Applications," *IEEE Trans. on Microwave Theory and Techniques*, vol. 52, pp. 2390 - 2408, 2004.

[27] S. E. Laux and W. Lee, "Collector Signal Delay in the Presence of Velocity Overshoot," *IEEE Electron Device Lett.*, vol. 11, No. 4, pp. 174 - 176, 1990.

[28] P. H. C. Magnee, "SiGe:C HBT Device Technology: What is the Big Deal?" *2003 Bipolar/BiCMOS Circuits and Technology Meeting*, Short Course, pp. 1 - 81, 2003.

[29] C. T. Kirk, "A Theory of Transistor Cutoff Frequency (f_T) Falloff at High Current Densities," *IEEE Trans. on Electron Devices*, pp. 164 - 174, 1962.

[30] D. Greenberg, S. Sweeney, B. Jagannathan, G. Freeman, and D. Ahlgren, "Noise Performance Scaling in High-Speed Silicon RF Technologies", *Digest of Silicon Monolithic Integrated Circuits in RF Systems*, pp. 22 - 25, 2003.

[31] Y. Kiyota, T. Hashimoto, T. Udo, A. Kodama, H. Shimamoto, R. Hayami and K. Washio, "190-GHz f_T 170-GHz f_{max} SiGe HBT with Heavily Doped Base by HCI-Free Selective Epitaxy," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 139 - 142, 2002.

[32] T. F. Meister, H. Schafer, K. Aufinger, R. Stengl, S. Boguth, R. Schreiter, M. Rest, H. Knapp, M. Wurzer, A. Mitchell, T. Bottner, and J. Bock, "SiGe Bipolar Technology with 3.9 ps Gate Delay," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 103 - 106, 2003.

[33] T. Tominari, S. Wada, K. Tokunaga, K. Koyu, M. Kubo, T. Udo, M. Seto, K. Ohhata, H. Hosoe, Y. Kiyota, K. Washio, and T. Hashimoto, "Study on Extremely Thin Base SiGe:C HBTs Featuring sub-5-ps ECL Gate Delay," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 107 - 110, 2003.

[34] B. Jagannathan, M. Khater, F. Pagette, J.-S. Rieh, D. Angell, H. Chen, J. Florkey, F. Golan, D. R. Greenberg, R. Groves, S. J. Jeng, J. Johnson, E. Mengistu, K. T. Schonenberg, C. M. Schnabel, P. Smith, A. Stricker, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, "Self-Aligned SiGe NPN Transistors with 285 GHz f_{max} and 207 GHz f_T in a Manufacturable Technology," *IEEE Electron Device Lett.*, vol. 23, No. 5 May, 2002.

[35] J. Gebner, F. Schwierz, D. Nuernbergk, and H. Mau, "The Influence of the Vertical Design of SiGe HBTs on the Cutoff Frequency," *IEEE Trans. on Electron Devices*, Accepted Aug. 14, 2002.

[36] D. L. Harame, "Device Structure and BiCMOS Integration," in *The Silicon Heterostructure Handbook*, pp. 213 - 229, in press.

[37] MEDICI user's manual, Synopsys.

[38] J. W. Slotboom, "The pn-product in Silicon," *Solid State Electronics*, vol. 20, pp. 279 - 283, 1977.

[39] D. B. M. Klaassen, J. W. Slotboom, and H. C. De Graaff, "Unified Apparent Bandgap Narrowing in n- and p-type Silicon," *Solid State Electronics*, vol. 35, pp. 125 - 129, 1992.

[40] Y. Shi, G. Niu, J. D. Cressler, and D. L. Harame, "On the Consistent Modeling of Band-gap Narrowing for Accurate Device-level Simulation of Scaled SiGe HBTs," *IEEE Trans. on Electron Devices*, vol. 50, pp. 1370 - 1377, 2003.

[41] D. B. M. Klaassen, "A Unified Mobility Model for Device Simulation – I. Model Equations and Concentration Dependence," *Solid State Electronics*, vol. 35, No. 7, pp. 953 - 959, 1992.

[42] Y. Shi and G. Niu, "Vertical Profile Design and Transit Time Analysis of Nano-Scale SiGe HBTs for Terahertz f_T," *Proc. of the 1993 IEEE Bipolar/BiCMOS Circuits Technology Meeting*, pp. 213 - 216, 2003.

[43] H. J. Osten, D. Knoll, B. Heinemann, B. Tillack, "Carbon Doping of SiGe Heterobipolar Transistors," *Digest of Silicon Monolithic Integrated Circuits in RF Systems*, pp. 19 - 23, 1998.

[44] H. J. Osten, G. Lippert, P. Gaworzewshi, and R. Sorge, "Impact of Low Carbon Concentration on the Electrical Properties of Highly Boron Doped SiGe Layers," *Appl. Phys. Lett*, vol. 71, pp. 1522 - 1524, Sept. 1997.

[45] P. Agarwal, H. G. A. Huizing, P. G. C. Magnee, "The Influence of Parasitic Resistances on the f_T-Optimization of High-Speed SiGe-HBTs," *IEEE ESSCIRC/ESSDERC*, pp. 291 - 294, 2003.

[46] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*, Cambridge University Press, 1998.

[47] J.-S. Rieh, B. Jagannathan, H. Chen, K. T. Schonenberg, D. Angell, A. Chinthakindi, J. Florkey, F. Golan, D. Greenberg, S.-J. Jeng, M. Khater, F. Pagette, C. Schnabel, P. Smith, A. Stricker, K. Vaed, R. Volant, D. Ahlgren, G. Freeman, K. Stein, S. Subbanna, "SiGe HBTs with Cut-off Frequency of 350 GHz," *Tech. Dig. IEEE Int. Electron Device Meeting*, pp. 771-774, 2002.

[48] J. D. Cressler and G. Niu, *Silicon-Germanium Heterojunction Bipolar Transistors*, Artech House.

[49] P. A. H. Hart, "Bipolar Transistor and Integrated Circuits," *Handbook on Semiconductors*, vol. 4, North-Holland Publishing Company, Amsterdam, 1981.

[50] J. J. H. Van Den Biesen, "A Simple Regional Analysis of Transit Times in Bipolar Transistors," *Solid-State Electronics*, vol. 29, pp. 529 - 534, 1986.

[51] K. Xia, G. Niu, D. Sheridan, and W. E. Ansley, "Ratio Based Direct Extraction of Small-Signal Parameters for SiGe HBTs," *2004 Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 144 - 147, 2004.

[52] J. N. Burghartz, T. O. Sedgwick, D. A. Grutzmacher, D. Nguyen-Ngoc, and K. A. Jenkins, "APVCB-grown Self-aligned SiGe-Base HBT's," *Proc. of the 1993 IEEE Bipolar/BiCMOS Circuits Technology Meeting*, pp. 55 - 62, 1993.

[53] L. C. M. van den Oever, L. K. Nanver, and J. W. Slotboom, "Design fo 200 GHz SiGe HBT's", *Proc. of the 2001 IEEE Bipolar/BiCMOS Circuits Technology Meeting*, pp. 78 - 81, 2001.

[54] L. C. M. van den Oever et al., *Proceedings of the SAFT/IEEE workshop*, pp. 119 - 122, 2000.

[55] G. Freeman, B. Jagannathan, S.-J. Jeng, J.-S. Rieh, A. D. Stricker, D. C. Ahlgren, and S. Subbanna, "Transistor Design and Application Considerations for > 200 - GHz SiGe HBTs," *IEEE Trans. on Electron Devices*, pp. 645 - 655, 2003.

[56] P. Chevalier, C. Fellous, L. Rubaldo, D. Dutartre, M. Laurens, T. Jagueneau, F. Leverd, S. Bord, C. Richard, D. Lenoble, J. Bonnouvrier, M. Marty, A. Perrotin, D. Gloria, F. Saguin, B. Barbalat, R. Beerkens, N. Zerounian, F. Aniel, and A. Chantre, "230 GHz Self-aligned SiGeC HBT for 90 nm BiCMOS Technology," *Bipolar/BiCMOS Circuits and Technology, Proceedings of 2004 Meeting*, pp. 225 - 228, 2004.

[57] B. Heinemann, R. Barth, D. Bolze, J. Drews, P. Formanek, Th. Grabolla, U. Haak, W. H*ö*ppner, D. Knoll, K. K*ö*pke, B. Kuck, R. Kurps, S. Marschmeyer, H. H. Richter, H. R*ü*cher, P. Schley, D. Schmidt, W. Winkler, D. Wolansky, H.-E. Wulf, and Y. Yamamoto, "A Low-parasitic Collector Construction for High-Speed SiGe:C HBTs," *IEEE International Electron Devices Meeting, Technical Digest*, pp. 251 - 254, 2004.

[58] M. Khater, J.-S. Rieh, T. Adam, A Chinthakindi, J. Johnson, R. Krishnasamy, M. Meghelli, F. Pagette, D. Sanderson, C. Schnabel, K. T. Schonenberg, P. Smith, K. Stein, A. Stricker, S.-J. Jeng, D. Ahlgren, and G. Freeman, "SiGe HBT Technology with f_{max}/f_T=350/300 and Gate Delay Below 3.3 ps," *IEEE International Electron Devices Meeting, Technical Digest*, pp. 247 - 250, 2004.

[59] Y. Shi and G. Niu, "2D Device Parasitic Analysis for 800/1000 GHz f_T/f_{max} SiGe HBTs", *Bipolar/BiCMOS Circuits and Technology, Proceedings of 2005 Meeting*, to be published, 2005.

[60] G. Freeman, A. Stricker, J.-S. Rieh, and D. Greenberg, "SiGe HBT Performance Limits," in *The Silicon Heterostructure Handbook*, pp. 569 - 580, in press.

[61] E. F. Crabbe, J. M. C. Stork, G. Baccarani, M. V. Fischetti, and S. E. Laux, "The Impact of Non-equilibrium Transport on Breakdown and Transit Time in Bipolar Transistors," *IEEE International Electron Devices Meeting, Technical Digest*, pp. 463 - 466, 1990.

[62] J. W. Slotboom, G. Streutker, M. J. V. Dort, P. H. Woerlee, A. Pruijmboom, D. J. Gravesteijn, "Non-local Impact Ionization in Silicon Devices," *IEEE International Electron Devices Meeting, Technical Digest*, pp. 127 - 130, 1991.

[63] P. Agarval, M. J. Goossens, V. Zieren, E. Aksen, and J. W. Slotboom, "Impact Ionization in Thin Silicon Diodes," *IEEE Electron Device Lett.*, pp. 807 - 809, 2004.

[64] G. Freeman, J.-S. Riech, B. Jagannathan, Z. Yang, F. Guarin, A. Joseph, "Device Scaling and Application Trends for Over 200 GHz SiGe HBTs," *Silicon Monolithic Integrated Circuits in RF Systems*, pp. 6 - 9, 2003

[65] J. Pan, G. Niu, A. Joseph, and D. L. Harame, "Impact of Profile Design and Scaling on Large Signal Performance of SiGe HBTs," *Bipolar/BiCMOS Circuits and Technology, Proceedings of 2004 Meeting*, pp. 209 - 213, 2004.

[66] M. Sotoodeh, L. Sozzi, A. Vinay, A. H. Khalid, Z. Hu, A. A. Rezazadeh, and R. Menozzi, "Stepping Toward Standard Methods of Small-Signal Parameter Extraction for HBT's," *IEEE Trans. on Electron Devices*, pp. 1139 - 1151, 2000.

APPENDICES

APPENDIX A

MEDICI INPUT FILE

A.1 2D MESH

```
assign  name=sufix c.val="RaisedBase2_we=120nm"
$ done = 1 -> solve INI for tif file
$ done = 0 -> Do not
assign  name=done n.val=1
assign  name=Depthoff n.val=0.095
call    file=Layout
assign  name=gemax n.val=0.28
mesh    out.fil=@sufix".msh"

$  MESH
x.mesh  width=0.2  n.space=5 h2=0.02
x.mesh  width=0.25 h1=0.02 h2=0.02 h3=0.06
x.mesh  width=0.10 n.space=6 h2=0.015
x.mesh  width=@ox2x2-@conb1x2 h1=0.015 h2=0.005
x.mesh  width=@PolySix1-@ox2x2 h1=0.008 n.space=8
x.mesh  width=0.02  h1=0.005
x.mesh  width=0.04  h1=0.005 h2=0.002 h3=0.015
```

```
x.mesh  width=0.12   h1=0.002 h2=0.002 h3=0.02

x.mesh  width=0.04  h1=0.002 h2=0.005 h3=0.015

x.mesh  width=0.02  h1=0.005

x.mesh  width=@PolySix1-@ox2x2 h2=0.008 n.space=8

x.mesh  width=@ox2x2-@conb1x2 h1=0.005 h2=0.015

x.mesh  width=0.10  h1=0.015 n.space=6

x.mesh  width=0.25  h1=0.02 h2=0.02 h3=0.06

x.mesh  width=0.2  n.space=5 h1=0.02

y.mesh  depth=0.04 n.space=4 h2=0.005

y.mesh  depth=0.05 h1=0.01 h2=0.01 h3=0.02

y.mesh  depth=0.025 h1=0.008 h2=0.001 h3=0.004

y.mesh  depth=0.005 h1=4e-4

y.mesh  depth=0.006 h1=5e-4

y.mesh  depth=0.003 h1=2e-4

y.mesh  depth=@geX4-0.129 h1=5e-4

y.mesh  depth=0.032 h1=5e-4 h2=2e-3

y.mesh  depth=0.215 h1=4e-3 h2=0.01 h3=0.03

y.mesh  depth=0.98 h1=0.02 h2=0.2


$ REGION

region  num=AllINI oxide

region  num=Edge PolySi Polygon
```

```
+x.poly=(@PolySix1,@PolySix2,@PolySix3,@PolySix4,
+ @PolySix5,@PolySix6,@PolySix7,@PolySix8)
+y.poly = (@PolySiy1, @PolySiy2, @PolySiy3, @PolySiy4,
+ @PolySiy5, @PolySiy6, @PolySiy7, @PolySiy8)
region  num=Inner PolySi Polygon
+   x.poly = (@PolySix1+0.01, @PolySix2-0.01, @PolySix3-0.01, @PolySix4-0.01,
+  @PolySix5-0.01, @PolySix6+0.01, @PolySix7+0.01, @PolySix8+0.01)
+   y.poly = (@PolySiy1, @PolySiy2, @PolySiy3-0.01, @PolySiy4-0.01, @PolySiy5,
+  @PolySiy6, @PolySiy7-0.01, @PolySiy8-0.01)

region  num=4a sige y.min=@geX1 y.max=@geX2 x.mol=0 x.end=@gemax
+ y.linear x.min=@sigex1 x.max=@sigex2
region  num=4b sige y.min=@geX2 y.max=@geX3 x.mol=@gemax x.end=@gemax
+ y.linear x.min=@sigex1 x.max=@sigex2
region  num=4c sige y.min=@geX3 y.max=@geX4 x.mol=@gemax x.end=0
+ y.linear x.min=@sigex1 x.max=@sigex2
region  num=Epi1 silicon x.min=@ox2x2 x.max=@PolySix6
+ y.min=@geX4 y.max=0.18
region  num=Epi2 silicon x.min=@PolySix5 x.max=@ox6x1
+ y.min=@geX4 y.max=0.18
region  num=SIC  silicon x.min=@PolySix6 x.max=@PolySix5
+ y.min=@geX4 y.max=0.18
```

```
region  num=3c silicon Polygon
+   x.poly = (@si2x1, @si2x2, @si2x3, @si2x4, @si2x5,
+  @si2x6, @si2x7, @si2x8, @si2x9, @si2x10, @si2x11, @si2x12)
+   y.poly = (@si2y1, @si2y2, @si2y3, @si2y4, 0.18,
+  0.18,   @si2y7, @si2y8, @si2y9, @si2y10, @si2y11, @si2y12)
region  num=1a PolySi x.min=@conc1x1 x.max=@conc1x2
+  y.min=@conc1y1 y.max=0.25
region  num=1b PolySi x.min=@conc2x1 x.max=@conc2x2
+  y.min=@conc2y1 y.max=0.25
region  num=1c PolySi x.min=@conb1x1 x.max=@conb1x2+0.04
+ y.min=@conb1y1 y.max=@conb1y2
region  num=1d PolySi x.min=@conb2x1-0.04 x.max=@conb2x2
+ y.min=@conb2y1 y.max=@conb2y2
region  num=1f PolySi x.min=@Poly3x1 x.max=@Poly3x2
+  y.min=@Poly3y1 y.max=@Poly3y2
region  num=1g PolySi x.min=@Poly4x1 x.max=@Poly4x2
+ y.min=@Poly4y1 y.max=@Poly4y2


$ ELECTRODE
elec    name=emitter x.min=@PolySix1 x.max=@PolySix2 top
elec    name=base region=1c
elec    name=base region=1d
```

```
elec    name=coll region=1b

elec    name=coll region=1a


$ PROFILE

$****** PolySi Emitter ******

profile n-type n.peak=1e21 Polygon

+   x.poly = (@PolySix1, @PolySix2, @PolySix3, @PolySix4,

+  @PolySix5, @PolySix6, @PolySix7, @PolySix8)

+   y.poly = (@PolySiy1, @PolySiy2, @PolySiy3, @PolySiy4,

+  @depthoff, @depthoff, @PolySiy7, @PolySiy8)

+   n.char=1e-10

profile 1d.ascii in.fil=asHI_Collopt4.dat y.column=1 n.column=2

+ y.offset=@depthoff

+ y.min=@depthoff  x.min=@PolySix6 x.max=@PolySix5 x.char=1e-4

profile 1d.ascii in.fil=asLOW_Collopt4.dat y.column=1 n.column=2

+ y.offset=@depthoff y.min=@depthoff

+ x.min=@PolySix6 x.max=@PolySix5 x.char=1e-4

$****** PolySi Extrinsic Base

profile p-type n.peak=5e20 x.min=@Poly3x1 x.max=@Poly3x2

+ y.min=@Poly3y1 y.max=@Poly3y2 y.char=5e-3 x.char=1e-4

profile p-type n.peak=5e20 x.min=@Poly4x1 x.max=@Poly4x2

+ y.min=@Poly4y1 y.max=@Poly4y2 y.char=5e-3 x.char=1e-4
```

```
$****** Intrinsic Base
profile 1d.ascii in.fil=boron_Collopt4.dat y.column=1 p.column=2
+ y.offset=@depthoff y.min=@sigey1
+   x.min=@sigex1 x.max=@sigex2 x.char=1e-5
$****** Collector
profile n-type n.peak=1e16 x.min=@ox2x2 x.max=@PolySix6
+ y.min=@sigey1 y.max=0.195 x.char=1e-3 y.char=1e-10
profile n-type n.peak=1e16 x.min=@PolySix5 x.max=@ox6x1
+ y.min=@sigey1 y.max=0.195 x.char=1e-3 y.char=1e-10
profile 1d.ascii in.fil=phos_Collopt4.dat y.column=1 n.column=2
+ y.offset=@depthoff
+ x.min=@PolySix6 x.max=@PolySix5 x.char=1e-4
profile 1d.ascii in.fil=asBuri_Collopt4.dat y.column=1 n.column=2
+ y.offset=@depthoff
+   y.min=@sigey1 y.max=0.195 x.min=@PolySix6 x.max=@PolySix5 x.char=1e-4
profile n-type n.peak=5e20 Polygon
+   x.poly = (0, @conc1x2, @ox2x1, @ox2x2, @ox2x2, @PolySix6,
+ @PolySix6, @PolySix5, @PolySix5, @ox6x1, @ox6x1, @ox6x2,
+ @conc2x1, @conc2x2, @conc2x2, 0)
+   y.poly = (@ox2y1, @ox2y1, @ox2y2, @ox2y2, 0.175, 0.175, 0.195,
+ 0.195, 0.175, 0.175, @ox6y2, @ox6y2, @ox6y1, @ox6y1, 1.5, 1.5) n.char=1e-6
```

A.2 DC Solution

INI.inp - solve for the initial solutions

```
assign      name=sufix c.val="RaisedBase_1p5THz"
mesh        in.fil="../mesh/"@sufix".msh"
models      phumob bgn consrh auger et.model tmpmob ef.tmp
call        file=newslotboom
material    silicon ele.tauw=3e-13
symb        newton carrier=2 ele.temp
solve   v(emitter)=0 v(coll)=0 v(base)=0
solve   v(emitter)=0 v(base)=0.1 v(coll)=0.1
save    out.fil=@sufix"INIuncoup.sov"
symb    newton carr=2 ele.temp coup.ele
load    in.fil=@sufix"INIuncoup.sov"
solve   v(emitter)=0 v(base)=0.1 v(coll)=0.1
+ out.fil=@sufix"VBE=0p1V_VCB=0V.sov"
loop    steps=8
    assign  name=vb n.val=0.15 delta=0.05
    solve   v(emitter)=0 v(base)=@vb v(coll)=@vb
l.end
save    out.fil=@sufix"VBE=0p5V_VCB=0V.sov"
solve   v(emitter)=0 v(base)=0.5 v(coll)=0.5 elec=coll vstep=0.1 nstep=5
solve   v(emitter)=0 v(base)=0.5 v(coll)=1.0
```

```
+ out.fil=@sufix"VBE=0p5V_VCB=0p5V.sov"
```

DC4YParam.inp - solve and save each DC solutions.

```
assign    name=sufix c.val="RaisedBase_1p5THz"
COMMENT   Specify a rectangular mesh
mesh      in.fil="../mesh/"@sufix".msh"
symb      newton carriers=2 ele.temp coup.ele
models    bgn phumob consrh auger tmpmob et.model ef.tmp
call    file=newslotboom
material    silicon ele.tauw=3e-13
load    in.fil=@sufix"VBE=0p5V_VCB=0p5V.sov"
assign    name=nstep n.val=50
assign    name=vstep n.val=0.008
assign    name=VCB n.val=0.5
Loop steps=@nstep
    assign name=vb n.val=0.55  delta=@vstep
        assign name=index n.val=1 delta=1
        assign name=zfile c.val=@sufix"_DCsov"@index"_VCB=0p5V"
    log    out.fil=@zfile".log"
        solve  v(emitter)=0 v(coll)=@vb+@VCB v(base)=@vb
        save   out.fil=@zfile
        plot.1d x.ax=V(base) y.ax=I(coll) y.log out.fil=@zfile".iv"
L.end
```

A.3 *ac* Solution

```
assign    name=sufix c.val="RaisedBase_1p5THz"
COMMENT   Specify a rectangular mesh
mesh      in.fil="../mesh/"@sufix".msh"
symb      newton carriers=2 ele.temp coup.ele
models    bgn phumob consrh auger tmpmob et.model ef.tmp
call      file=newslotboom
material  silicon ele.tauw=3e-13
assign    name=nstep n.val=20
assign    name=idxINI n.val=31
Loop steps=@nstep
          assign name=index n.val=@idxINI    delta=1
          assign name=zfile c.val=@sufix@index
      log   out.fil=@zfile"acVCB=0p5V.log"
      Load  in.fil=@sufix"_DCsov"@index"_VCB=0p5V"
          solve  nfstep=30 mult.freq
+                AC.ANAL  TERM=(Base,coll) FREQ=10e9 fstep=1.2 Hi.freq
L.end
```

APPENDIX B

MATLAB SCRIPTS

B.1 Transit Time Analysis

B.1.1 $\Delta\tau_n$ and $\Delta\tau_p$ Extraction

```
clear all;

close all;

sufix = '../Graded/Graded';

q=1.6e-19;

for i=8:2:16

    nacfn = sprintf('Nac%d.dat', i);

    pacfn = sprintf('Pac%d.dat', i);

    Jcfn = sprintf('Jnac%d.dat', i);

    Icfn = sprintf('Ic%d.dat', i);


    Mnac = load(chf(nacfn));

    Mpac = load(chf(pacfn));

    Mjc = load(chf(Jcfn));

    MIc = load(chf(Icfn));


    deln = Mnac(:,4);
```

```
    delp = Mpac(:,4);
    Jc = Mjc(:,4);
    JCC(i) = MIc(1,2)./0.2.*1000;

    dtaun = q.*deln./Jc.*1e-4.*1e12; % ps/um
    dtaup = q.*delp./Jc.*1e-4.*1e12; % ps/um
    dist = Mjc(:,3);
    figure(1);
    plot(dist,-deln,'b');
    hold on;

    figure(2);
    plot(dist, -dtaun, 'b');
    hold on;
    %plot(dist,-delp,'r');
    %legend('\Delta\tau_p','\Delta\tau_n');
    xlabel('Depth (\mum)');
    ylabel('\Delta\tau_n, \Delta\tau_p (ps)');

    check = Int_tau_depth(dist, -dtaun);
end
```

B.1.2 1D τ_n and τ_p Extraction

```
clear all;

close all;

sufix = 'Box';

q=1.6e-19;

for i=1:17

    nacfn = sprintf('Nac%d.dat', i);

    pacfn = sprintf('Pac%d.dat', i);

    Jcfn = sprintf('Jnac%d.dat', i);

    Icfn = sprintf('Ic%d.dat', i);


    Mnac = load(chf(nacfn));

    Mpac = load(chf(pacfn));

    Mjc = load(chf(Jcfn));

    MIc = load(chf(Icfn));


    deln = Mnac(:,4);

    delp = Mpac(:,4);

    Jc = Mjc(:,4);

    JCC(i) = MIc(1,2)./0.2.*1000;


    dtaun = q.*deln./Jc.*1e-4.*1e12; % ps/um
```

```
    dtaup = q.*delp./Jc.*1e-4.*1e12; % ps/um
    dist = Mjc(:,3);

    jx1 = 0.02;
    jx2 = 0.025;
    jx3 = 0.0325;
    Tau_eh=0;
    Tau_el=0;
    Tau_b=0;
    Tau_c=0;

len=length(dist);
for k=1:len-1
        Taun = -(dtaun(k));
        if (dist(k)<jx1)
            dTeh = Taun*(dist(k+1)-dist(k)); % Emitter Transit time
            Tau_eh = Tau_eh + dTeh;
            continue;
        end

        if (dist(k)>jx1 & dist(k)<jx2) % Emitter transit time
            dTel = Taun*(dist(k+1)-dist(k)); % Emitter Transit time
```

```
            Tau_el = Tau_el + dTel;
            continue;
        else if (dist(k)>jx2 & dist(k)<=jx3) % Base Transit time
            dTb = Taun*(dist(k+1)-dist(k));
            Tau_b = Tau_b + dTb;
        else
            dTc = Taun*(dist(k+1)-dist(k)); % Collector charging time
            Tau_c = Tau_c + dTc;
        end
    end
end
TauEH(i) = Tau_eh;
TauEL(i) = Tau_el;
TauB(i) = Tau_b;
TauC(i) = Tau_c;
Total(i) = Tau_eh+Tau_el+Tau_b+Tau_c;
end


figure(1);
subplot(1,2,1);
semilogy(JCC, TauEH);
hold on;
```

```
semilogy(JCC, TauEL,'--');

semilogy(JCC, TauB,'.-');

semilogy(JCC, TauC, '-+');

figure(2);

loglog(JCC, Total);
```

B.1.3 2D τ Extraction

```
clear all;

close all;

%%%%%%%%%%%%%%%%%%%%

% Region ID

% (2)  - Poly Edge

% (3)  - Poly Inner

% (4)  - SiGe 4a

% (5)  - SiGe 4b

% (6)  - SiGe 4c

% (7)  - Si Epi1

% (8)  - Si Epi2

% (9)  - Si SIC

% (10) - Si 3C

% (13) - Poly 1f

% (14) - Poly 1g
```

```
% (16) - Poly Edge
%%%%%%%%%%%%%%%%%%%%

Region = [4,5,6, 13,14];
q = -1.6e-19;

dir = './WE=100nm/E_';
filename = sprintf('%sdQ_Base_Jc=30.dat', dir);
delete(filename);
fid = fopen(filename, 'a+');
fprintf(fid,'%QEBjx, QBaseEx, QBaseIn, QCollEx, QCollIn, QTotal \n');

%%%%%%%%%%%%%%%%%%%%%%%%
%%% Interation - 26:2:48 to get all QE vs. biasing
%%%%%%%%%%%%%%%%%%%%%%%%
for Index4Biasing = 1:1

Coorfn = sprintf('%sVBcoordinate_VCB=0p5V_Jc=30.data', dir);
Trifn  = sprintf('%sVBtriangle_VCB=0p5V_Jc=30.data', dir);
Nodefn = sprintf('%sVBNodeIndex_VCB=0p5V_Jc=30.data',dir);
Datafn = sprintf('%sVBData_VCB=0p5V_Jc=30.data', dir);
```

```
Coordinate = load(chf(Coorfn, 2));

Tri = load(Trifn);

NodeIndex = load(Nodefn);

Data = load(Datafn);

InteQ = 0;


%%% Triangle "2, TriIndex, RegionID, c1, c2, c3"

%%% The unit of each region is made up of triangles

RegionID = Tri(:,3);

c1 = Tri(:,4);

c2 = Tri(:,5);

c3 = Tri(:,6);


%%% Coordinate "1, CoIndex, X, Y"

CoIndex = Coordinate(:,2);

X = Coordinate(:,3);

Y = Coordinate(:,4);


len = length(RegionID);


for m=1:length(Region) % for each region defined in the input file
```

```
    clear dInteQ;
    dQEBjx = 0;
    dQBaseEx = 0;
    dQBaseIn = 0;
    dQBasePoly = 0;
    dQCollEx = 0;
    dQCollIn = 0;

for i=1:len
    if(RegionID(i)~=Region(m))
        continue;
    else
        %%%% get info. of c1 c2 c3
        target = [c1(i),c2(i),c3(i)];
        lenCoord = length(CoIndex);
        for j=1:3
            for k=1:lenCoord
                if(CoIndex(k)~=target(j))
                    continue;
                else
                    x_ax(j) = X(k);
                    y_ax(j) = Y(k);
```

```
            break;
        end
    end
end
d1 = sqrt((x_ax(1)-x_ax(2))^2+(y_ax(1)-y_ax(2))^2);
d2 = sqrt((x_ax(1)-x_ax(3))^2+(y_ax(1)-y_ax(3))^2);
d3 = sqrt((x_ax(3)-x_ax(2))^2+(y_ax(3)-y_ax(2))^2);
temp = (d1+d2+d3)/2;
ds = sqrt(temp*(temp-d1)*(temp-d2)*(temp-d3)); % unit- um^2
ds = ds*1e-8; % unit- cm^2

%%%% get data
lenData = length(NodeIndex);
for j=1:3
    for k=1:lenData
        if(NodeIndex(k)~=target(j))
            continue;
        else
            data_nac(j) = Data(k,15);
            % Get ac_electron n (unit: /cm^3)
            break;
        end
```

```
    end
end
avg_nac = (data_nac(1)+data_nac(2)+data_nac(3))/3;
% unit of n - /cm^3
dInteQ(i) = (q*avg_nac)*ds; % unit: Coul/cm
dInteQ(i) = dInteQ(i)/1e4; % unit: Coul

if( Region(m)==13 | Region(m)==14 )
    dQBasePoly = dInteQ(i);
    continue;
else
    %%% Extract delta N for each region definition -
        %for extrinsic base, EB jx, and intrinsic Base
    if(max(x_ax)<=0.73 | min(x_ax)>=0.85)
    % left/right-side of SiGe Base
        if ( max(y_ax) <= 0.1260 )
            dQBaseEx(i) = dInteQ(i);
        else
            dQCollEx(i) = dInteQ(i);
        end
        continue;
    end
```

```
            if (min(x_ax)>0.73 & max(x_ax)<0.85)
            % Middle of SiGe Base
                if (max(y_ax)<=0.1222)
                    dQEBjx(i) = dInteQ(i);
                else if (min(y_ax)>0.1222 & max(y_ax)<=0.126)
                        dQBaseIn(i) = dInteQ(i);
                    else
                        dQCollIn(i) = dInteQ(i);
                    end
                end
                continue;
            end
        end
    end
end

Result(m) = sum(dInteQ);
QEBjx(m) = sum(dQEBjx);
QBaseEx(m) = sum(dQBaseEx);
QBaseIn(m) = sum(dQBaseIn);
QBasePoly(m) = sum(dQBasePoly);
```

```
QCollEx(m) = sum(dQCollEx);
QCollIn(m) = sum(dQCollIn);
end

DataSave = [sum(QEBjx), sum(QBaseEx)+sum(QBasePoly),
    sum(QBaseIn), sum(QCollEx), sum(QCollIn), sum(Result)];
StrSave = num2str(DataSave);
fprintf(fid,'%s\n', StrSave);
end
fclose(fid);
disp('FINISH!');
```

B.2 Distributive Capacitance

```
clear all;
close all;

dJc = load('dJc.txt');
dIc = dJc.*0.12./1e8;
dQE = load('dQ_Emitter_Biasing.dat');
dQB = load('dQ_Base_Biasing.dat');
dQC = load('dQ_Coll_Biasing.dat');
dQCVe = load('dQ_Coll_BiasingVe.dat');
```

```
dQCVe = -1.*dQCVe;

gm = load('../../PostProcessing/GmJc.dat');

gm = gm';

Taud = load('../../PostProcessing/TaudJc.dat');

RB = load('../../PostProcessing/RBJc.dat');

Taud = Taud'./gm;

TauEh = dQE(:,1)./dIc.*1e12;

TauEl = ( dQE(:,2) + dQB(:,1) )./dIc.*1e12;

TauEf = dQE(:,3)./dIc.*1e12;

TauBin = dQB(:,3)./dIc.*1e12;

TauBex = dQB(:,2)./dIc.*1e12;

TauCin = (dQCVe(:,1)+dQB(:,5))./dIc.*1e12;

TauCex = (dQCVe(:,2)+dQB(:,4))./dIc.*1e12;

TauCin_dep = (dQC(:,1) - dQCVe(:,1))./dIc.*1e12;

TauCex_dep = (dQC(:,2) - dQCVe(:,2))./dIc.*1e12;

TauEC = TauEh + TauEl + TauEf + TauBin

    + TauBex + TauCin + TauCex + TauCin_dep + TauCex_dep;


j = 1;

for i=26:2:48

    fn = sprintf('../Calibration/RaisedBase2_we=120nm_DCsov%d_VCB=0p5V.iv', i);

    Bias = load(chf(fn, 5));
```

```
    Jc(j)=Bias(1,2)/0.12*1000;
    j = j+1;
end
Tau_x = (TauEl+TauBin+TauCin*0.8);
CdE=gm.*TauEh;
Cte = gm.* TauEl;
CEf = gm.*TauEf;
CdBin = gm.*TauBin;
CdBex = gm.* TauBex;
CdCin = gm.* TauCin;
CdCex =gm.* TauCex;
```

B.3 $M-1$ Postprocessing

```
clear all;
close all;
EMedici = load(chf('VBE=0p65V_VCB=1p8V_vac.ef'));
Xjxn = 0.165; % um
Xjxp = 0.126; % um
VCB4Save=1.8;
% Extract Edata in the CB jx
index = 1;
dist = EMedici(:,1);
```

```
for k=1:length(dist)
    if(dist(k)>= Xjxp & dist(k)<=Xjxn)
        xax(index) = EMedici(k,1);
        yax(index) = EMedici(k,2);
        index = index+1;
    else
        continue;
    end
end

%%%%%%%%%%%% actual data to work with
Edata = [xax; yax];
Edata = Edata';
len = length(xax);
xp = xax(1);
xn = xax(len);
%%%%%%%%%%%%%%%%%%%%%%%%
Eeffdata(1) = 1e4;
alphaN(1) = 0;
alphaP(1) = 0;
for i=2:len
    Eeffdata(i) = Eeff(xp, xax(i), Edata); % V/cm
```

```
    alphaNP = IIalpha(Eeffdata(i));
    alphaN(i) = alphaNP(1); % 1/cm
    alphaP(i) = alphaNP(2); % 1/cm
end
% Calculate inner-integral
expTerm(1) = 0;
func = alphaN - alphaP; % 1/cm
func = func./1e4; % 1/um
func = [xax; func];
func = func';
for i=2:len
    tempRlt = Integral(xp, xax(i), func);
    expTerm(i) = exp(tempRlt);
end

% Calculate outer-integral
clear func;
func = alphaN .* expTerm; % 1/cm
func = func./1e4; % 1/um
%Rlt = -Integral(Xjxp, Xjxn, func);
for i=1:length(func)-1
    avg = (func(i)+func(i+1))/2;
```

```
    delx = xax(i+1)-xax(i);
    area(i) = avg * delx;
end
Rlt = sum(area);

Mn = 1/(1-Rlt);
MM1 = Mn - 1;
disp(MM1);

if(1)
MM14Save=MM1.*1e6;
outputData = num2str([VCB4Save, MM14Save]);
fname = 'RaisedBase_we=120nmVBE=0p65V.MM1';
fid = fopen(fname, 'a+');
fwrite(fid, outputData);
fprintf(fid, '\n');
fclose(fid);
end

function result = IIalpha(E)
E = abs(E);
% for electrons
```

```
an = 7.03e5; %1/cm

bn = 1.231e6; %V/cm


% for holes

if ( E<=4e5)

    ap = 1.582e6;

    bp = 2.036e6;

else if (E>4e5 & E<= 6e5)

        ap = 6.71e5;

        bp = 1.693e6;

    end

end


alphaN = an * exp(-bn/E);

alphaP = ap * exp(-bp/E);

result = [alphaN alphaP];


% Calbulate the effective Electric Filed at position x

% Considering the non-local effect

function result = Eeff (xp, x, Edata)


Lambdan = 65; % nm
```

```
Lambdan = 65e-3; % um


depthE = Edata(:,1);
lenE = length(depthE);
xaxE(1) = Edata(1,1);
yaxE(1) = Edata(1,2)/1e4;
indexE = 2;
for kE=2:lenE
    if(depthE(kE)<=x)
        xaxE(indexE) = Edata(kE,1); % um
        yaxE(indexE) = Edata(kE,2)/1e4; % V/um
        indexE = indexE+1;
    else
        break;
    end
end
nE = length(xaxE)-1;
for i=1:nE
    V2E = yaxE(i+1)*exp((xaxE(i+1)-x)/Lambdan);
    V1E = yaxE(i)  *exp((xaxE(i)  -x)/Lambdan);
    avgE = (V1E+V2E)/2;
    delxE = xaxE(i+1)-xaxE(i);
```

```
    areaE(i) = avgE * delxE;
end
result = sum(areaE)/Lambdan; % V/um
result = result/1e-4; % V/cm

function result = Integral (st, en, data)
index = 2;
dist = data(:,1);
xaxInt(1) = data(1,1);
yaxInt(1) = data(1,2);
for k=2:length(dist)
    if(dist(k)<=en)
        xaxInt(index) = data(k,1);
        yaxInt(index) = data(k,2);
        index = index+1;
    else
        break;
    end
end
n = length(xaxInt)-1;
for i=1:n
    %clera avg delx;
```

```
    avgInt = (yaxInt(i)+yaxInt(i+1))/2;
    delxInt = xaxInt(i+1)-xaxInt(i);
    areaInt(i) = avgInt * delxInt;
end
result = sum(areaInt);
```